Chapter 2

DESIGNING ETHICAL PRACTICE IN BIOSURVEILLANCE

The Project Argus Doctrine

JEFF COLLMANN[1,*] and ADAM ROBINSON[2]

CHAPTER OVERVIEW

Biosurveillance entails the collection and analysis of information needed to provide early warning of outbreaks of infectious disease, both naturally occurring and intentionally introduced. Data derived from repositories containing various types of sensitive information may be required for this purpose, including individually identifiable, copyrighted, and proprietary information. The Project Argus Biosurveillance Doctrine was developed to ensure that ethical and legal principles guide the collection and handling of such information. Project Argus does not, however, use individually identifiable information or any material derived from individually identifiable information for any phase of the project. Further, Project Argus is not used for purposes of law enforcement, counterterrorism, or public health surveillance. This chapter details why and how the doctrine was developed and summarizes its guiding principles and key elements.

Keywords: Biosurveillance; Sensitive information; Information protection; Privacy

[1*] *O'Neill Institute for National and Global Health Law, Disease Prevention and Health Outcomes, School of Nursing and Health Studies, Georgetown University Medical Center, Box 571107, 3700 Reservoir Rd, NW, Washington, DC 20057–1107, USA, collmanj @georgetown.edu*

[2] *The MITRE Corporation, 7515 Colshire Drive, McLean, VA 22102, USA*

D. Zeng et al. (eds.), *Infectious Disease Informatics and Biosurveillance*, Integrated Series in Information Systems 27, DOI 10.1007/978-1-4419-6892-0_2,

1. INTRODUCTION

Biosurveillance entails the collection and analysis of information needed to provide early warning of outbreaks of infectious disease, both naturally occurring and intentionally introduced. Data derived from repositories containing various types of sensitive information may be required for this purpose, including individually identifiable, copyrighted, and proprietary information. Project Argus searches open media in all countries of the globe except the United States to find direct and indirect indications that local communities have identified and begun to respond to an emerging infectious disease such as SARS or pandemic influenza [1]. The ultimate goal is to provide early warning of such events so that countermeasures can be taken to limit the spread and mitigate the consequences of the disease. When originally planning Project Argus, we developed a "Biosurveillance Doctrine" to ensure that ethical principles would guide the collection and handling of such information. Specifically, our efforts included five steps:

- Development and analysis of scenarios for managing sensitive project information
- Examination of relevant laws, regulations, good practice, and case studies in the acquisition, analysis, and archiving of this information
- Development of administrative, physical, and technical policies and procedures for safely managing project information
- Development of technical design requirements for the Project Argus biosurveillance system and
- Development of a doctrine management process

Through these efforts, the system incorporated requirements for the ethical handling of sensitive information from the start, rather than retrofitting it later. Because the initial phase of the project focused only on the acquisition, archiving, analysis, and presentation of biosurveillance information, we limited our initial efforts to ensure ethical practice of/for these activities as reported in this chapter. Project Argus did not, when implemented, use individually identifiable information or any material derived from individually identifiable information for any phase of the project. Further, Project Argus is not used for purposes of law enforcement, counterterrorism, or public health surveillance. This chapter details why and how the doctrine was developed and summarizes its key components.

2. BACKGROUND

Project Argus developed the technical and doctrinal requirements for an integrated, multisource information system designed to perform global bio-

surveillance for epidemics; biological accidents; and bioattacks on humans, animals, and plants [1].

The ethical issues surrounding the development and maintenance of such a system have been a key consideration from the outset of Project Argus. No single public law or set of regulations governs the handling and protection of the broad range of sources, types, security classifications, and potential uses of the information to be collected and analyzed. Therefore, the Project Argus doctrine team was formed to develop the necessary guidance. The resulting biosurveillance doctrine sets forth explicit principles, management structures, policies, procedures, and technical design requirements intended to ensure the ethical handling and use of sensitive information by project participants. In this respect, a strong moral, organizational, and technical divide exists between Project Argus and initiatives that have drawn the censure of Congress, the media, and the American public for their failure to ensure such protections. After describing the methods used to develop the doctrine, we provide a high-level view of its guiding principles and key elements.

It should be noted that the version of the doctrine presented here applies only to the acquisition, archiving, analysis, and presentation of biosurveillance information in Project Argus. The doctrine team analyzed a broad set of sensitive information, including individually identifiable, copyrighted and public information. We include our analysis of and approach for handling this broad set of sensitive information for the sake of completeness and as a guide to others. Project Argus does not use any individually identifiable information in any phase of the project.

3. OVERVIEW: INFORMATION PROTECTION

Project Argus collects, archives, and interprets various types of information, including confidential or sensitive information that requires special handling. The leaders and sponsors of Project Argus required development of the Biosurveillance Doctrine to familiarize all project members, contractors, and partners with relevant laws, regulations, ethical principles, and good industrial practices governing use of sensitive information and ensure their compliance with their precepts. In addition to examining relevant cases such as the controversy about the Terrorism Information Awareness (TIA) program (see below), the Biosurveillance Doctrine team investigated issues associated with using specific types of sensitive information, including individually identifiable, proprietary, and copyrighted information. Certain guidance, such as the Principles of Fair Use of copyrighted materials, bears directly on the type of information that Project Argus acquires. Other guidance, such as the Security and Privacy Standards of the Health Insurance Portability and

Accountability Act (HIPAA) of 1996 and the European Privacy Directive, directly or indirectly affects how Project Argus shares information with potential partners. For example, no Project Argus investigators qualify as "covered entities" under HIPAA. Project Argus doctrine must, nonetheless, refer to HIPAA because it may potentially collaborate with health-care providers who should share patient information only with HIPAA-compliant partners. HIPAA and the European Privacy Directive also embody versions of good privacy and security practice. By aligning its practice with principles expressed in these regulatory regimes, Project Argus demonstrates good faith in protecting sensitive information obtained from its partners or through its own initiatives.

Good information security practice requires establishing administrative, physical, and technical controls to protect the confidentiality, integrity and availability of all project data. We imagined that research data from Project Argus might reside in various locations, including the ISIS Center at Georgetown University and MITRE (partners in the development of the Argus information system). Relevant information security policies from all such hosts and other project participants appear as appendices to the Project Argus Biosurveillance Doctrine as required. The ISIS Center houses several R&D projects that manage confidential information, including individually identified health information. The ISIS Center has established a risk-based information security program with controls to protect information of several types including public, commercially sensitive, research, and individually identifiable information. MITRE has rigorous controls reflecting its identity as a major Federally Funded Research and Development Center serving sensitive sectors of the U.S. government. Project Argus benefits from the general organizational controls and tailors specific controls to meet its own needs when appropriate. Memoranda of Understanding among participating organizations document their mutual obligations in protecting shared project information of any kind whenever necessary.

The Biosurveillance Doctrine Team's analysis of these general issues yielded some implications for Project Argus. Summaries of these implications follow to help the reader better understand the specific policies and procedures proposed for Project Argus.

3.1 Fair Information Practice Principles

Fair Information Practices represent an international consensus on appropriate handling of personal information. Various versions of these principles appear in the European Privacy Directive, the U.S. Privacy Act, and guidelines issued by the Organization for Economic Cooperation and

Development and the Canadian Standards Association. Key provisions include the following.

- Notice: At or before the time of collection, individuals shall be informed of the personal information to be collected, the purpose of the collection, and to whom the information may be disclosed.
- Consent: To the maximum extent possible, individuals shall consent to the collection of their personal information at or before the time of collection.
- Accuracy: Personal information shall be sufficiently accurate, complete, and current to serve the intended purpose.
- Security: Personal information shall be protected by safeguards appropriate to the sensitivity of the information.
- Access: Individuals shall have the opportunity to review the personal information held about them and records of its disclosure.
- Redress: Individuals shall have the opportunity to request correction of their personal information and to challenge compliance with stated practices.
- Limitation: Collection, use, disclosure, and retention of personal information shall be limited to that which is necessary for the intended purpose.

3.2 Proprietary Information

Copyrighted Information. Project Argus seeks, acquires, archives, and analyzes copyrighted materials, primarily through its web search technology (known as Apollo). The ISIS Center, the home base for Project Argus, qualifies as a non-profit, educational, research-oriented institution. Furthermore, Project Argus is a pilot study of limited scope. If its methods do not prove useful, the project will be discontinued. For materials not easily acquired retroactively due to their ephemeral nature, such as dynamic web pages, Project Argus downloads a copy of the web page and creates an archive of Hypertext Markup Language (html) files for future reference. To stay within the bounds of fair use, as defined in copyright law (17 U.S.C. §107), Project Argus acknowledges its use of copyrighted materials first by purchasing materials of interest when necessary, either directly from the publisher or through an aggregator. Project Argus excludes all website sections not relevant to its research requirements and labels all archived articles as "For Research Purposes Only." The archives will only exist for the life of the project. Project Argus does not distribute, republish, or disseminate the archived articles for any commercial or non-commercial purpose under any conditions. For these reasons, the limited use of copyrighted materials in Project Argus constitutes fair use.

Confidential Business Information. Project Argus may handle confidential business information in many forms. For example, Project Argus could potentially contract with commercial companies to provide aggregated data of various types. In such cases, Project Argus drafts contracts that reflect its own information protection and use policies as well as comply with federal law and policy. In all instances, Project Argus only uses the data for the defined purposes of the project and does not share the data with parties external to Project Argus.

3.3 Individually Identifiable Information

- **Protected Health Information (PHI).** The ISIS Center, the home base for Project Argus, does not qualify as a covered entity under HIPAA because it does not provide or pay for medical treatment of individuals. Under certain circumstances, Project Argus may receive PHI from covered entities such as Georgetown University Hospital or the Washington Hospital Center as part of conducting research in biosurveillance. The HIPAA Privacy Rule would require submission of a Human Subjects Review application of some type to the Georgetown University Medical Center's Institutional Review Board (IRB). Government sponsors might also require review of Project Argus' use and disclosure of PHI by a relevant IRB. Depending on the actual circumstances, the application may seek an expedited or full review. This has not yet occurred in the project but may occur in later phases.
- **Telephone call detail.** Although never used in Project Argus, we investigated methods for preparing aggregate telephone call data between regions of interest based on individually identifiable telephone call information. Each call on a telephone network generates a call detail record (CDR) that stores the telephone number of the phone that made the call (the originating number), the dialed number, the telephone number that received the call (the terminating number), the time at which the call was placed, and the duration of the call [2]. Telephone companies routinely use these statistical analyses to monitor network reliability and detect international fraud. Companies may also perform analysis of aggregated CDR on a contractual basis to third parties.

One may compile aggregated data of call volumes from one specified region to another. It is not highly granular information that an analyst could use to deduce the identities of individual callers in a designated area. When telecommunications carriers are required to provide call-identifying

information, it is by court order and is limited to specific individuals and forms of communication. Because the purpose of Project Argus is to detect indications of societal disruption, only aggregated regional call data is of use. Project Argus has never used telephone data of any kind in its work; but, evaluated these measures for the sake of completeness and scholarly relevance.

- **Individual financial information.** None of the data streams initially proposed for Project Argus analysis was financial in nature. There could come a time, however, when aggregated financial data, such as the number of automated teller machine transactions in a given period for a given region, could prove useful for detecting societal disruption. Privacy regulations such as the ones in the Gramm–Leach–Bliley Act could serve as a model in the future but are not needed at this time.
- **Intelligence on U.S. Persons.** The mission of Project Argus includes detecting social disruption, not tracking individuals. Thus, Project Argus has and will not develop, pilot, or evaluate means for identifying individuals for law enforcement, crime prevention, or public health surveillance purposes on U.S. or foreign persons. Project Argus implemented policies and procedures designed to minimize the incidental, or unintentional, collection and to dispose of information about U.S. persons.
- **European Privacy Directive.** As the foregoing discussions of specific types of individually identifiable information imply, the United States implements a sectoral approach to privacy. By contrast, the European Privacy Directive handles all individually identifiable information with a single, comprehensive approach and restricts the flow of information to countries that do not provide substantially equivalent protections. The United States and the European Union (EU) have adopted "Safe Harbor" provisions with which U.S. entities must comply in order to transact business involving personal information between the United States and EU member states. The Biosurveillance Doctrine Team examined the European Privacy Directive and the Safe Harbor provisions because they represent a major instance of the Fair Information Practice Principles, described below, with which Project Argus aspires to comply.
- **Aggregation of Information from Disparate Databases.** Project Argus recognizes the theoretical possibility that it might generate individually identifiable data in the course of combining otherwise de-identified data from disparate databases. In general, Project Argus does not seek to create individually identifiable data from any sources. Project Argus investigators, furthermore, discard any such data that appears incidentally as a function of intended or unintended project procedures.

4. METHODS

The Project Argus doctrine team drew general inspiration from the work of William Odom, who recommends organizing the intelligence community to reflect the phases of the intelligence cycle: topic selection and data collection, analysis, use, and evaluation [3]. The doctrine described in this chapter focused only on the acquisition, archiving, analysis, and presentation of biosurveillance information. Eventually, the doctrine will provide end-to-end protection of information through all phases of biosurveillance (see Figure 2-1).

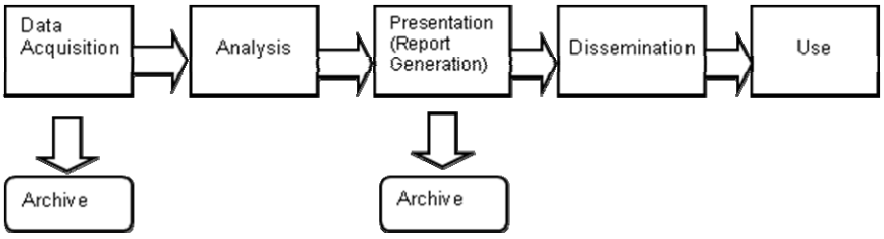

Figure 2-1. End-to-end protection of biosurveillance information.

The doctrine team conducted five types of activities to carry out its charge. The first was an analysis of some typical scenarios that Project Argus team members may confront when managing the sensitive but unclassified information originally imagined for analysis in the project. One scenario was developed for each of five information types: telecommunications information, information from open-source media, remote-sensing information, changes in website content, and air transportation information. The results of the scenario analyses informed our second activity, an examination of laws, regulations, good practice, and case studies in the acquisition, analysis, and archiving of this information, such as the Privacy and Security Rules of the Health Insurance Portability and Accountability Act (HIPAA) of 1996 [4–5]. We developed the doctrine in three steps based on these efforts: we authored administrative, physical, and technical policies and procedures for acquiring, analyzing, archiving, and protecting project information; we created technical design requirements for the system; and we developed a doctrine management process. Through these measures, the system incorporated requirements for the ethical handling of sensitive information from the start, rather than retrofitting them later.

4.1 Information Scenarios

The five scenarios address information that falls under one or more of five broad types of information. Each scenario traces handling of the information

through collection, archiving, analysis, and presentation (report generation). Three of the five scenarios are presented below; collectively they illustrate all the issues encountered in the scenario analyses.

- **Scenario 1: Telecommunications Information**
  - ***Scenario type*:** Managing individually identifiable information
  - ***Source of information*:** Individually identifiable call detail records (CDRs)
  - ***Biosurveillance information*:** Ongoing deidentified summaries of calls between selected regions of the world
  - ***Analytic objective*:** Identify significant deviations from baseline rates of calls between these regions of the world

Step 1. Acquire information

Project Argus eventually decided not to use telephone call data for any purpose. The doctrine team developed an approach to deidentifying telephone data before this decision was made. Each call on a telephone network generates a CDR that stores the number of the telephone used to make the call (the originating number), the dialed number, the telephone number that received the call (the terminating number), the time at which the call was placed, and the duration of the call. CDRs constitute International Telecommunications Company (ITC) proprietary business information because the information is generated in the course of ITC's business and is collected and stored by ITC pursuant to its arrangements with its customers. ITC regularly and legally monitors CDRs for a variety of routine business purposes, such as ensuring network reliability and detecting international fraud, and compiles CDR reports. ITC cannot provide the CDRs or CDR reports to third parties such as Project Argus without the permission of the customers involved. To make it possible to establish the baseline as well as the ongoing rate of telephone traffic between regions of the world of interest to Project Argus, it was established that ITC aggregate the CDRs in a manner that removes all individually identifiable data (see below) and retain the CDRs themselves. It was also established that no one from Project Argus ever participate in acquiring the data, see the CDRs or handle any individually identifiable information.

Step 2. Aggregate data

ITC could aggregate the CDR data by preparing graphs that illustrate the volume of calls between regions of the world specified by Project Argus analysts. This step deidentifies the data and, thereby, makes the aggregate results available for such purposes as biosurveillance. The preparation of these graphs does not constitute a routine business practice for ITC; however, ITC agreed to provide the graphs as part of its participation in Project Argus.

Because regions with small call volume would not yield meaningful results, Project Argus agreed to specify a minimum call volume required per geographic area for ITC to produce a data point. Although ITC retained the ability to identify the individuals represented in the graphs through the CDRs, through this mechanism Project Argus could establish a technical design requirement to prevent project analysts from recovering individually identifiable call data from the graphs. ITC and Project Argus agreed to produce a Memorandum of Understanding (MOU) to specify the terms and conditions for the production, transfer, and use of the graphs had the project been implemented.

Step 3. Produce graphical reports and deliver to Project Argus

The protocol specified that ITC produce monthly graphical reports of call volumes for regions specified by Project Argus and transmit the reports to analysts in the ISIS Center. Had call volumes between regions of interest equaled or exceeded an Argus-defined threshold, ITC would have shifted to daily reporting.

Step 4. Archive information

According to the protocol, Project Argus analysts at the Imaging Science and Information Systems (ISIS) Center would receive, index, and digitally store the above reports, producing a comprehensive archive of all information received from ITC. From an information security perspective, the ITC reports contain sensitive but unclassified information. The ISIS Center has established policies and procedures for protecting the confidentiality, integrity, and availability of sensitive information (see http://www.isis.georgetown.edu). During periods when call volumes fall below the alert threshold, requirements for data integrity and timeliness remain consistent with the everyday research and development (R&D) environment of the ISIS Center. When call volumes exceed the alert threshold and ITC reports arrive daily, Project Argus should consider escalating these requirements. The project reevaluates its information security requirements, considers the need for new requirements, and recommends special controls if needed. Archives will exist only for the life of the project.

Step 5. Analyze aggregated ITC information

The ISIS Center proposed to analyze the aggregated call volume information received from ITC and compare them with other datasets so as to identify anomalies that may represent indications and warnings of an emerging bioevent in a region of interest. Project Argus produces and posts several types of reports on a restricted website "For Official Use Only."

- **Scenario 2: Open-Source Media Information**
  - ***Scenario type*:** Managing foreign copyrighted information
  - ***Source of information*:** Foreign news media websites
  - ***Biosurveillance information*:** Whole news stories and abstracted information about disease and social disruption in regions of interest outside the United States
  - ***Analytic objective*:** Identify direct indicators of disease and indirect indicators of social disruption secondary to an emerging bioevent

Step 1. Acquire electronic data

Using geographic selection criteria based on Project Argus research requirements, analysts identify online newspapers and other websites of interest to the project. They also identify the information on these websites that is not relevant to the project. For example, on a community website, only sections that might publish articles about school closings would be of interest to the project; sections reporting on local sports scores and entertainment would not be as relevant. Argus engineers write a script specifying the interval of retrieval and the content to be excluded for isisMiTAP, an integrated suite of human-language technologies that processes semistructured textual data. On the established retrieval schedule, isisMiTAP retrieves and stores in its archive all content ("articles") from the identified websites that has not been specifically excluded. Project Argus treats all these reports as if they were copyrighted. To comply with the principles of fair use, the project excludes all website sections irrelevant to its research requirements and labels all archived articles as "for research purposes only." Thus archived articles are not distributed, republished, or disseminated for any commercial or noncommercial purpose under any conditions. Moreover, as with aggregated CDR data, Project Argus archives the articles only for the duration of the project.

Step 2. Translate and catalogue articles of interest

isisMiTAP processes the articles retrieved and presents the information they contain to users through various interfaces. isisMiTAP processing includes machine translation of foreign-language content, information extraction in the form of identifying named entities and keywords (e.g., diseases, locations, people), categorization (binning and posting to a news server), archiving (storing raw and derivative files on disk), and indexing (to support full-text searches). Human linguists translate selected articles to allow for more complete understanding. Analysts prepare summaries of selected articles upon demand.

Step 3. Archive information

Project Argus created and maintains a temporary electronic archive of selected articles at the ISIS Center, including all original-language texts and English translations, article summaries, and web links. All retrieved and archived articles are treated as "confidential – copyrighted" information subject to appropriate administrative, physical, and technical information security controls. The requirements for data integrity and timeliness will remain consistent with the everyday R&D environment of the ISIS Center until a bioevent is suspected, at which point the need to escalate the requirements will be considered. Project Argus reevaluates its information security requirements, considers the need for new requirements, and recommends special controls when needed. For example, a research and development prototype must not typically operate at all times and can tolerate some downtime. Were Project Argus to evolve into a mission-critical operational unit, it would require a business continuity plan that includes tactics to recover from interruptions in its IT system that currently does not exist.

Step 4. Analyze archived articles

Project Argus analysts identify, categorize, and evaluate the significance of salient events in the archived articles. Project Argus produces and posts several types of reports on a restricted website "For Official Use Only."

- **Scenario 3: Remote-Sensing Information**
  - ***Scenario type*:** Managing U.S. federal government information
  - ***Source of information*:** Website of the U.S. National Aeronautics and Space Administration (NASA)
  - ***Biosurveillance information*:** Aggregated remote-sensing information on weather conditions in selected countries of interest
  - ***Analytic objective:*** Identify the existence of weather conditions favorable to the development and spread of selected infectious diseases, such as Rift Valley Fever

Step 1. Acquire information

The U.S. National Oceanic and Atmospheric Administration (NOAA) collects weather data using remote satellite sensing and distributes the data to NASA, which analyzes it and posts it at http://www.nasa.gov. The U.S. government makes these reports available to the public for unlimited use, at no cost. When the doctrine team conducted this analysis, Argus analysts intended to download the aggregated weather reports on a monthly basis from the NASA website to the Argus electronic archive. In practice, these specific reports did not prove useful. Analysts do periodically consult the NASA MODIS Rapid

Response System website to retrieve remote-sensing images of recent fires, volcano eruption, and storms around the globe, data that shares the same properties of aggregated weather data for the purposes of fair use.

Step 2. Archive information

Project Argus downloads various public data from the NASA website to Argus's electronic archive. The ISIS Center's policies and procedures for protecting the confidentiality, integrity, and availability of the content of its archives that include sensitive information apply to such data. As with articles from foreign media, the requirements for data integrity and timeliness of remote-sensing information remain consistent with the everyday R&D environment of the ISIS Center until a bioevent is suspected, at which point Project Argus reevaluates its information security requirements for this information.

Step 3. Analyze remote-sensing information

Argus analysts examine the NASA information to establish the presence or absence of conditions favorable to the development and spread of infectious diseases such as Rift Valley Fever. Project Argus produces and posts several types of reports on a restricted website "For Official Use Only."

4.2 Laws, Regulations, and Good Practice in Managing Sensitive Information

Guided by the results of the scenario analyses, the doctrine team examined laws, regulations, and best practice pertaining to the management of sensitive information federal laws and regulations, including executive orders, international directives, agency policies and procedures, Congressional testimony, government and nongovernment reports, best practices and ethical principles. Certain guidance, such as the principles of fair use of copyrighted materials, bears directly on the types of information Project Argus acquires. Other guidance, such as the Security and Privacy Rules of HIPAA and the European Privacy Directive, may affect directly or indirectly how the project shares information with potential partners. For example, no Project Argus investigators qualify as "covered entities" under HIPAA. The Project Argus doctrine must nonetheless refer to HIPAA because at some point the project may collaborate with healthcare providers who should share patient information only with HIPAA-compliant partners. HIPAA and the European Privacy Directive also embody good privacy and security practice. By aligning its practice with principles expressed in these regulatory regimes, Project Argus demonstrates good faith in protecting sensitive information obtained from its partners or through its own initiatives.

4.3 Case Study: The Terrorism Information Awareness Program

Just prior to the launch of Project Argus, the U.S. Congress discontinued funding for the TIA program, a large counterterrorism effort organized by the Defense Advanced Research Projects Agency (DARPA). At first glance, the TIA program, with its focus on counterterrorism and law enforcement, has little in common with Project Argus. Yet both initiatives had to address common issues related to privacy and security, as well as functionality. Whether beginning with individually identifiable information, such as CDR data, or discovering identities in the course of analysis, as with public health or medical surveillance information, organizations conducting both terrorist investigations and biosurveillance must obey relevant privacy laws; establish pertinent policies and procedures; train their workforces; and implement risk-based administrative, physical, and technical privacy and security safeguards. In preparing the Project Argus Doctrine, the team examined TIA and other programs for lessons regarding these core controls [6–23].

In addition to implementing one of the key lessons learned from these case studies – the need to address such issues in policies and procedures from the outset of a project – the Project Argus doctrine team conducted a detailed analysis of such programs to identify other potential pitfalls and lessons learned. We recognize that the American public basically accepts as legitimate the aims of both scientific research and counterterrorism. However, individual programs must carefully assess and clearly explain the tradeoffs that exist between individual and societal welfare in specific instances, particularly in times of threat and conflict.

Thus biosurveillance investigators must not take for granted the good will of their subjects, their institutions, or their funding agencies. Rather, they must take personal responsibility for ensuring the implementation of appropriate privacy controls. We identified specific means to that end in our research, including:

- Incorporate privacy and security controls into technical design requirements for computerized biosurveillance information systems.
- Take full advantage of the privacy functions of the Institutional Review Board (IRB). As suggested by the report on TIA of the Department of Defense Inspector General, the IRB is fully equipped to advise and monitor researchers on privacy policies, procedures, and practices. In most academic medical research institutions, HIPAA has strengthened the IRB's awareness of and competence to deal with privacy issues.

- Devote great care to preparing the privacy and security portions of the IRB review forms, particularly the informed consent form. The IRB forms can function for an individual research project much like the privacy impact assessment prepared by federal agencies, helping to identify and propose mitigation plans for privacy risks associated with a project. The informed consent form provides an ideal vehicle for explaining to subjects a project's privacy protections.
- When affiliated with a medical center, cultivate an effective relationship with the center's HIPAA privacy and security officers. Like the privacy ombudsmen in federal agencies, these individuals facilitate communication on these matters among researchers, subjects, the institutions involved in the work and external agencies, such as the Office of Civil Rights and the Department of Health and Human Services.
- Consider using an external project advisory board when conducting research or using "data mining" methods that could raise privacy concerns. If properly composed and chartered, such a group can provide useful expertise in policy, privacy, and legal matters beyond a researcher's own institution and enhance the credibility of a project's good-faith efforts in the event of controversy.
- Formally develop and document in writing privacy and security policies and procedures for the research project or its parent unit. As HIPAA and the report of the Department of Defense Inspector General emphasize, these written policies and procedures should explain protections identified in the IRB forms, including administrative, physical, and technical controls for privacy and security.
- Work with relevant information security officers in the home institutions of project members to establish sound controls protecting the confidentiality, integrity, and availability of research data, including individually identifiable information.
- Train project team members in the ethical principles and institutional policies and procedures governing information privacy and security in the project.

5. RESULTS AND ANALYSIS

Based on the analyses described above, we developed three key elements of the Project Argus Biosurveillance Doctrine: policies and procedures, technical requirements, and doctrine management.

5.1 Policies and Procedures

The Argus policies and procedures express the information protection and use guiding principles in succinct form.

General Policy Statements. Project Argus will not acquire, archive, analyze, or distribute information in a form that is prohibited by applicable laws, regulations, or good ethical practice. The project will establish a risk-based IA program to protect the confidentiality, integrity, and availability of all project-related information, including public and unclassified but sensitive information; the project will not handle classified information of any type. Participating institutions will apply their own IA policies in acquiring, archiving, analyzing, and distributing any Project Argus information at their locations. Memoranda of Agreement (MOA) will establish the conditions for sharing information among collaborating organizations. Sharing is defined as providing collaborators access to an original owner's information by any means, including remote electronic access to an owner's archive or transfer of an owner's information to a collaborator's archive.

Specific Policies and Procedures. The scenarios discussed earlier established conditions governing the acquisition, analysis, archiving, and distribution of individually copyrighted, identifiable, proprietary, confidential, publicly available, and business information. For the doctrine, these conditions were translated into specific policies and procedures for each of the five types of information, such as the following ones for individually identifiable information.

- Original owners of individually identifiable information bear responsibility for complying with applicable laws, regulations, and good ethical practice in making such information in their possession available to Project Argus investigators.
- Original owners of individually identifiable information bear responsibility for obtaining permission from subjects as necessary or required for the use of such information in Project Argus.
- Only original owners of individually identifiable information may view, change, analyze, or otherwise use such information in Project Argus unless otherwise agreed upon and justified in writing.
- Original owners of individually identifiable information will limit access to the purposes of Project Argus, its data, data parameters, and data destinations to those members of their own organizations with a need to know such things.
- Project Argus will strive to acquire, archive, analyze, and distribute only aggregated or deidentified information from original owners of individually identifiable information.

- Written MOAs between original owners of individually identifiable information and Georgetown University on behalf of Project Argus will incorporate any relevant rules for the acquisition, archiving, analysis, and distribution of such information to be shared in the course of the project.
- All Project Argus participants, including original owners of individually identifiable information, will abide by the terms and conditions of relevant MOAs regarding the acquisition, archiving, analysis, and distribution of such information obtained from an original owner.
- If Project Argus investigators should inadvertently acquire individually identifiable information, they will discard it and seek no further information about the individual.
- Project Argus participants will receive training in the appropriate handling of any such inadvertently acquired information before incorporating information into the project archive.
- No Project Argus participant will share individually identifiable information with any person, within or external to the project, who is unauthorized to view, receive, or otherwise use it. Nor will they share it with any such organizations or entities.
- Project Argus will classify individually identifiable information as "confidential – individually identifiable" and treat it and any associated information according to the protections for confidential information on the ISIS Center network.

5.2 Technical Requirements

To enable and enforce compliance with the policies and procedures detailed above, the doctrine team developed technical requirements to guide engineers in designing the Project Argus biosurveillance information system to incorporate technical controls that enable and enforce compliance with the policies and procedures detailed above. Consistent with the IA philosophy of defense in depth – which requires multiple, overlapping privacy and security protections – the system's specific application controls will function within an administrative, physical, and technical infrastructure that complements and sustains them. Some of the technical requirements apply to all types of information; while others apply only to specific types, as follows.

- All types of information:
  - Requirements for data integrity and timeliness will remain consistent with the everyday R&D environment of the ISIS Center, unless special circumstances arise, such as a suspected bioevent.

  - Technical requirements will reflect good industry practice for protecting unclassified sensitive information, such as individually identifiable, proprietary, copyrighted, and research information.
  - Project Argus will regularly reevaluate its IA requirements, consider the need for new requirements, and recommend special controls if needed.
- "Confidential – individually identifiable" information:
  - In collaboration with Project Argus investigators, original owners of such information will develop, implement, and monitor the performance of methods for aggregating or deidentifying individually identifiable information when such aggregated information is required for research.
  - Project Argus participants will not be able to reidentify the subjects of the aggregated individually identifiable information to which they have access. Methods used for deidentifying information will minimize the chances of such reidentification.
  - Project Argus participants who are not authorized original owners of such information will not have access to the individually identifiable information from which aggregated information is prepared.
- "Confidential – proprietary" information: Project Argus will develop, implement, and monitor the performance of technical measures designed to enforce any special requirements listed in an MOA between Georgetown University and an original owner of such information.
- "Confidential – copyright" information: Project Argus will develop, implement, and monitor the performance of technical measures designed to enforce the time limit on storage of copyrighted information in the project archives.
- Publicly available information: Project Argus will develop, implement, and monitor the performance of technical measures designed to enforce any special requirements for protecting publicly available information that becomes research data, as deemed necessary by a relevant risk assessment.

5.3 Doctrine Management Process

The Project Argus doctrine team realized that promulgating policies, procedures and technical requirements alone would not necessarily assure compliance. We sought to institutionalize the information protection through a series of organizational measures including routine doctrine team meetings, special oversight procedures, specifications for groups working with but not members of Project Argus and, critically, training for Project Argus staff.

- **Doctrine Team Meetings.** The doctrine team stayed active throughout the initial development of Project Argus and developed a concept of operations for sharing Argus results with its customers. The team has reported to the oversight boards described below.
- **Oversight Boards.** Project Argus established internal and external oversight boards to review its reports and provide overall project guidance including compliance with the Biosurveillance Doctrine. The internal board consists of the doctrine team members and all Project Argus task team leaders. The external board consists of representatives of key government and academic stakeholders.
- **Doctrine Specifications for User Communities.** Sponsors, consumers, and unwitting participants all will have varying needs throughout the life of Project Argus. The doctrine team was instrumental in establishing the organizational conditions for safe sharing of biosurveillance data among a range of government agencies that has proven invaluable in numerous actual biothreat scenarios.
- **Doctrine Training and Awareness Program.** The doctrine team developed a code of ethics that embodies the key elements of the doctrine and served as the foundation for a comprehensive training and awareness program for all Project Argus participants that remains active.

6. CONCLUSION

The American public basically accepts as legitimate the aims of scientific research and counterterrorism, including biosurveillance for pandemics. Individu programs must carefully assess and clearly explain the tradeoffs that exist between individual and societal welfare involved in specific instances, particularly in times of heightened concern about possible attacks by an elusive and possibly indigenous foe. To this end, Project Argus has developed the Biosurveillance Doctrine described in this chapter, designed to ensure the appropriate acquisition, analysis, protection, and use of sensitive information, particularly individually identifiable, copyrighted, and proprietary information. The doctrine will be reviewed and revised as necessary throughout the life of the project. Through these efforts, Project Argus aims to keep faith with its sponsoring agencies, Congress, and the American people.

ACKNOWLEDGMENTS

The authors wish to thank Alexander W. Joel, Esq., Glen C. Graber, Ph.D. and Ted Cooper, M.D. for their helpful comments on earlier drafts of this

chapter. Georgeann Higgins and Sandra Sinay made important contributions. The Intelligence Technology Innovation Center (ITIC) funded all work associated with this chapter. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the ITIC.

QUESTIONS FOR DISCUSSION

1. What are the advantages and disadvantages of planning information protection and use policies and procedures before designing an information technology project?
2. What are the similarities and differences in protecting the various types of sensitive information including proprietary, copyrighted and individually identifiable information?
3. Do you think that government or commercial organizations are capable of protecting or appropriately using sensitive or personal information?
4. Are codes of good information practice sufficient to protect sensitive information in a company or government organization? If not, what else do you think is necessary?
5. What difficulties can you imagine might arise in monitoring compliance with an organization's information protection and use policies and procedures after an information management system is deployed?
6. If you were in charge of developing an IT system to handle sensitive information, how would you incorporate information protection issues into your design process?

REFERENCES

1. Wilson, J., Polyak, M., Blake, J., and Collmann, J. A Heuristic Indication and Warning Staging Model for Detection and Assessment of Biological Events. *Journal of American Medical Informatics Association* 15(2) (2008 Mar/Apr), 158–171.
2. Fisher, K., Hogstedt, K., Rogers, A., and Smith, F. (2002). *Hancock 2.0.1 Manual*, AT&T Labs Shannon Laboratory http://www.research.
3. Odom, W.E. (2003). *Fixing Intelligence for a More Secure America.* New Haven, CT: Yale University Press.
4. Department of Health and Human Services; Office of the Secretary. (2002). Final Rule. 45 CFR Part 160, 162, and 164, standards for privacy of individually identifiable health information. *Federal Register* 67, no. 157 (14 August), 53181–53273.

5. Department of Health and Human Services; Office of the Secretary. (2003) Final rule. 45 CFR Part 160, 162, and 164, security standards. *Federal Register* 68, no. 34 (20 February), 8333–8381.
6. Crews, Jr., C. The Pentagon's total information awareness project: Americans under the microscope? *National Review Online*, (2002 25 November).
7. Cooper, T., and Collmann, J. (2005). Managing Information Security and Privacy in Health Care Data Mining. In, Advances in Medical Informatics: Knowledge Management and Data Mining in Biomedicine, New York, NY: Springer Science; pp. 95–137.
8. Defense Advanced Research Projects Agency; Information Awareness Office. (2003). Report to Congress regarding the terrorist information awareness program: in response to consolidated appropriations resolution, Pub. L. No. 108–7, Division M, § 111(b); 20 May.
9. Department of Defense; Office of the Inspector General, Information Technology Management. (2003). Terrorist information awareness program. Report D-2004–033; 12 December.
10. Department of Health and Human Services; Office for Human Research Protections. (2004). Guidance on research involving coded private information or biological specimens. 10 August. Available at http://www.hhs.gov/ohrp/humansubjects/guidance/cdebiol.pdf.
11. Department of Health and Human Services; National Institutes of Health Office for Protection from Research Risks. (2001). Protection of human subjects. 45 C.F.R. § 46.
12. Department of Health and Human Services. (2003). Protecting personal health information in research: understanding the HIPAA privacy rule. NIH publication no. 03-5388.
13. Federal Trade Commission. (1999). In brief: the financial privacy requirements of the Gramm-Leach-Bliley Act. Available at http://www.ftc.gov/bcp/conline/pubs/buspubs/glbshort.htm.
14. Fisher, K., Hogstedt, K., Rogers, A., and Smith, F. (2002). *Hancock 2.0.1 Manual*. Florham Park, NJ: AT&T Labs Shannon Laboratory.
15. Mack, G., Bebee, B., Shafi, I., Wenzel, G., Medairy, B., and Yuan, E. (2002) Total Information Awareness Program (TIA) System Description Document (SDD) v1.1. Defense Advanced Research Projects Agency Information Awareness Office. White Paper; 19 July.
16. National Institutes of Health. (2004). Clinical research and the HIPAA privacy rule. NIH publication no. 04–5495.
17. Safire, W. You are a suspect. *The New York Times*, (2002 14 November). Available from: http://www.commondreams.org/views02/1114-08.htm.
18. Simons, B., and Spafford, E. Letter to Honorable John Warner, Chairman, Senate Committee on Armed Forces; (2003 23 January).
19. Stanley, J., and Steinhardt, B. (2003) Bigger monster, weaker chains: the growth of an American surveillance society. American Civil Liberties Union, Technology and Liberty Program. White Paper.
20. Sweeney, L. ed. (2003). Navigating computer science research through waves of privacy concerns. Carnegie Mellon University, School of Computer Science, Pittsburgh, Technical report, CMU CS 03-165, CMU-ISRI-03-102.
21. Taipale, K. Data mining and domestic security: connecting the dots to make sense of data. *The Columbia Science and Technology Law Review* 5 (2003), 5–83.
22. Taylor, S. Big brother and another overblown privacy scare. *Atlantic Online*, (2002 10 December).
23. *The Washington Post*. Total information awareness. Editorial, (2002 16 November).

SUGGESTED READING

1. Cooper, T., and Collmann, J. (2005) Managing Information Security and Privacy in Health Care Data Mining. In, *Advances in Medical Informatics: Knowledge Management and Data Mining in Biomedicine*, New York, NY: Springer Science; pp. 95–137.
2. Collmann, J., and Cooper, T. Breaching the Security of the Kaiser Permanente Internet Patient Portal: The Organizational Foundations of Information Security. *Journal of the American Medical Informatics Association* 14 (2007 Mar), 239–43.
3. Cooper, T., Collmann, J., and Neidermeier, H. Organizational repertoires and rites in health information security. *Cambridge Quarterly of Healthcare Ethics*, Volume 17 <http://journals.cambridge.org/action/displayJournal?jid=CQH&volumeId=17&bVolume=y#loc17>, Issue 04 <http://journals.cambridge.org/action/displayIssue?jid=CQH&volumeId=17&seriesId=0&issueId=04>, October 2008, pp. 441-452.
4. Department of Health and Human Services; Office of the Secretary. (2003). Final rule. 45 CFR Part 160, 162, and 164, security standards. *Federal Register* 68, no. 34 (20 February), 8333–8381.
5. Department of Health and Human Services. (2003). Protecting personal health information in research: understanding the HIPAA privacy rule. NIH publication no. 03-5388.
6. Odom, W.E. (2003) *Fixing Intelligence for a More Secure America.* New Haven, CT: Yale University Press.
7. Sweeney, L. ed. (2003). Navigating computer science research through waves of privacy concerns. Carnegie Mellon University, School of Computer Science, Pittsburgh, Technical report, CMU CS 03–165, CMU-ISRI-03-102.
8. Taipale, K. Data mining and domestic security: connecting the dots to make sense of data. *The Columbia Science and Technology Law Review* 5 (2003) 5–83.

ONLINE RESOURCES

1. US Fair Trade Commission, Fair Information Practice Principles, http://www.ftc.gov/reports/privacy3/fairinfo.shtm
2. US Department of Commerce, Safe Harbor, http://www.export.gov/safeharbor/sh_overview.html
3. Health Information Management System Society, *HIMSS Privacy and Security Toolkit* http://www.himss.org/ASP/privacySecurityTree.asp?faid=78&tid=4#PSToolkit#PSToolkit
4. US Department of Health and Human Service, HIPAA Privacy Support, http://www.hhs.gov/ocr/hipaa/
5. US Department of Health and Human Service, HIPAA Privacy and Security Rules, http://aspe.hhs.gov/ADMNSIMP/
6. Caralli, R.A., Stevens, J.F., Wallen, C.M., White, D.W., Wilson, W.R., and Young, L.R. *Introducing the CERT Resiliency Engineering Framework: Improving the Security and Sustainability Processes*, Software Engineering Institute, http://www.sei.cmu.edu/publications/documents/07.reports/07tr009.html

http://www.springer.com/978-1-4419-6891-3

Infectious Disease Informatics and Biosurveillance

Zeng, D.; Chen, H.; Castillo-Chavez, C.; Lober, W.B.; Thurmond, M. (Eds.)

2011, LI, 488 p., Hardcover
ISBN: 978-1-4419-6891-3