

Chapter 1

INTELLIGENCE AND SECURITY INFORMATICS (ISI): CHALLENGES AND OPPORTUNITIES

Chapter Overview

The tragic events of September 11th and the following anthrax contamination of letters caused drastic effects on many aspects of society. Academics in the fields of natural sciences, computational science, information science, social sciences, engineering, medicine, and many others have been called upon to help enhance the government's ability to fight terrorism and other crimes. Six critical mission areas have been identified where information technology can contribute, as suggested in the "National Strategy for Homeland Security" report, including: *intelligence and warning, border and transportation security, domestic counter-terrorism, protecting critical infrastructure, defending against catastrophic terrorism, and emergency preparedness and responses*. Facing the critical missions of national security and various data and technical challenges, we believe there is a pressing need to develop the science of "Intelligence and Security Informatics" (ISI). This chapter reviews ISI research challenges, compares ISI with the development of Biomedical Informatics, and suggests federal funding initiatives and research opportunities of relevance to ISI.

1.1 Introduction

The tragic events of September 11th and the following anthrax contamination of letters caused drastic effects on many aspects of society. Terrorism became the most significant threat to national security because of its potential to bring massive damage to our infrastructure, economy, and people.

In response to this challenge, federal authorities are actively implementing comprehensive strategies and measures in order to achieve the three objectives identified in the “National Strategy for Homeland Security” report (Office of Homeland Security, 2002): (1) preventing future terrorist attacks, (2) reducing the nation’s vulnerability, and (3) minimizing the damage and recovering from attacks that occur. State and local law enforcement agencies, likewise, are becoming more vigilant about the criminal activities that harm public safety and threaten national security.

Academics in the fields of natural sciences, computational science, information science, social sciences, engineering, medicine, and many others have been called upon to help enhance the government’s ability to fight terrorism and other crimes. Science and technology have been identified in the “National Strategy for Homeland Security” report as the keys to win the new counter-terrorism war (Office of Homeland Security, 2002). It is widely believed that information technology will play an indispensable role in making our nation safer (National Research Council, 2002) by supporting intelligence and knowledge discovery through collecting, processing, analyzing, and utilizing terrorism- and crime-related data (Chen et al., 2003a; Chen et al., 2004b). Based on the crime and intelligence knowledge discovered, the federal, state, and local authorities can make timely decisions to select effective strategies and tactics as well as allocate the appropriate amount of resources to detect, prevent, and respond to future attacks.

1.2 Information Technology and National Security

Six critical mission areas have been identified where information technology can contribute to the accomplishment of the three strategic national security objectives identified in the “National Strategy for Homeland Security” report (Office of Homeland Security, 2002):

- *Intelligence and Warning.* Although terrorism depends on surprise to damage its targets (Office of Homeland Security, 2002), terrorist activities are not random and impossible to track. Terrorists must plan and prepare before the execution of an attack by selecting a target, recruiting and training executors, acquiring financial support, and

traveling to the country where the target is located (Sageman, 2004). To avoid being preempted by authorities they may hide their true identities and disguise attack-related activities. Similarly, criminals may use falsified identities during police contacts (Wang et al., 2004a). Although it is difficult, detecting potential terrorist attacks or crimes is possible and feasible with the help of information technology. By analyzing the communication and activity patterns among terrorists and their contacts (i.e., terrorist networks), detecting deceptive identities, or employing other surveillance and monitoring techniques, intelligence and warning systems may issue timely, critical alerts and warnings to prevent attacks or crimes from occurring.

- *Border and Transportation Security.* Terrorists enter a targeted country through an air, land, or sea port of entry. Criminals in narcotics rings travel across borders to purchase, carry, distribute, and sell drugs. Information, such as travelers' identities, images, fingerprints, vehicles used, and other characteristics, is collected from customs, borders, and immigration authorities on a daily basis. Counter-terrorism and crime-fighting capabilities can be greatly improved by the creation of a "smart border," where information from multiple sources is shared and analyzed to help locate wanted terrorists or criminals. Technologies such as information sharing and integration, collaboration and communication, biometrics, and image and speech recognition will be greatly needed in such smart borders.
- *Domestic Counter-terrorism.* As terrorists, both international and domestic, may be involved in local crimes, state and local law enforcement agencies are also contributing to the missions by investigating and prosecuting crimes. Terrorism, like gangs and narcotics trafficking, is regarded as a type of organized crime in which multiple offenders cooperate to carry out offenses. Information technologies that help find cooperative relationships between criminals and their interactive patterns would also be helpful for analyzing terrorism. Monitoring activities of domestic terrorist and extremist groups using advanced information technologies will also be helpful to public safety personnel and policy makers.
- *Protecting Critical Infrastructure and Key Assets.* Roads, bridges, water supplies, and many other physical service systems are critical infrastructure and key assets of a nation. They may become the target of terrorist attacks because of their vulnerabilities (Office of Homeland Security, 2002). Moreover, virtual (cyber) infrastructure such as the Internet may also be vulnerable to intrusions and inside threats (Lee and

Stolfo, 1998). Criminals and terrorists are increasingly using cyberspace to conduct illegal activities, share ideology, solicit funding, and recruit. In addition to physical devices such as sensors and detectors, advanced information technologies are needed to model the normal behaviors of the usage of these systems and then use the models to distinguish abnormal behaviors from normal behaviors. Protective or reactive measures can be selected based on the results to secure these assets from attacks.

- *Defending Against Catastrophic Terrorism.* Terrorist attacks can cause devastating damage to a society through the use of chemical, biological, or radiological weapons. Biological attacks, for example, may cause contamination, infectious disease outbreaks, and significant loss of life. Information systems that can efficiently and effectively collect, access, analyze, and report data about catastrophe-leading events can help prevent, detect, respond to, and manage these attacks (Damianos et al., 2002).
- *Emergency Preparedness and Responses.* In case of a national emergency, prompt and effective responses are critical to reduce the damage resulting from an attack. In addition to the systems that are prepared to defend against catastrophes, information technologies that help design and experiment with optimized response plans (Lu et al., 2003), identify experts, train response professionals, and manage consequences are beneficial in the long run. Moreover, information systems that facilitate social and psychological support to the victims of terrorist attacks can also help society recover from disasters.

Although it is important for the critical missions of national security, the development of information technology for counter-terrorism and crime-fighting applications faces many problems and challenges.

1.3 Problems and Challenges

Currently, intelligence and security agencies are gathering large amounts of data from various sources. Processing and analyzing such data, however, has become increasingly difficult. By treating terrorism as a form of organized crime, these challenges can be categorized into three types:

- *Characteristics of criminals and crimes.* Some crimes may be geographically diffused and temporally dispersed. In organized crimes such as transnational narcotics trafficking, criminals often live in different countries, states, and cities. Drug distribution and sales occur in different places at different times. Similar situations exist in other

organized crimes (e.g., terrorism, armed robbery, and gang-related crime). As a result, an investigation must cover multiple offenders who commit criminal activities in different places at different times. This can be fairly difficult given the limited resources that intelligence and security agencies have. Moreover, as computer and Internet technologies advance, criminals are utilizing cyberspace to commit various types of cybercrimes under the disguise of ordinary online transactions and communications.

- *Characteristics of crime and intelligence related data.* A significant source of challenge is information stovepipe and overload resulting from diverse data sources, multiple data formats, and large data volumes. Unlike other domains such as marketing, finance, and medicine in which data can be collected from particular sources (e.g., sales records from companies, patient medical history from hospitals), the intelligence and security domain does not have a well-defined data source. Both authoritative information (e.g., crime incident reports, telephone records, financial statements, immigration and customs records) and open source information (e.g., news stories, journal articles, books, web pages) need to be gathered for investigative purposes. Data collected from these different sources often are in different formats ranging from structured database records to unstructured text, image, audio, and video files. Important information such as criminal associations may be available but contained in unstructured, multilingual texts and remains difficult to access and retrieve. Moreover, as data volumes continue to grow, extracting valuable and credible intelligence and knowledge becomes a difficult problem.
- *Characteristics of crime and intelligence analysis techniques.* Current research on the technologies for counter-terrorism and crime-fighting applications lacks a consistent framework addressing the major challenges. Some information technologies including data integration, data analysis, text mining, image and video processing, and evidence combination have been identified as being particularly helpful (National Research Council, 2002). However, the question of how to employ them in the intelligence and security domain and use them to effectively address the critical mission areas of national security remains unanswered.

Facing the critical missions of national security and various data and technical challenges, we believe there is a pressing need to develop the science of “Intelligence and Security Informatics” (ISI) (Chen et al., 2003a; Chen et al., 2004b), with its main objective being the “development of

advanced information technologies, systems, algorithms, and databases for national security-related applications, through an integrated technological, organizational, and policy-based approach” (Chen et al., 2003a).

1.4 Intelligence and Security Informatics vs. Biomedical Informatics: Emergence of a Discipline

Comparing ISI with Biomedical Informatics, an established academic discipline addressing information management issues in biological and medical applications (Shortliffe and Blois, 2000; Chen et al., 2005), we found tremendous analogies between these two disciplines. Table 1-1 summarizes the similarities and differences between ISI and Biomedical Informatics.

Table 1-1. Analogies between ISI and Biomedical Informatics.

		Biomedical Informatics	ISI
Challenges	Domain-specific	<ul style="list-style-type: none"> Complexity and uncertainty associated with organisms and diseases Critical decisions regarding patient well-being and biomedical discoveries 	<ul style="list-style-type: none"> Geographically diffused and temporally dispersed organized crimes Cyber-crimes on the Internet Critical decisions related to public safety and homeland security
	Data	<ul style="list-style-type: none"> Information stovepipe and overload HL7 XML standard PHIN MS messaging Patient records, diseases data, medical images 	<ul style="list-style-type: none"> Information stovepipe and overload Justice XML standard Criminal incident records Multilingual intelligence open sources
	Technology	<ul style="list-style-type: none"> Ontologies and linguistic parsing Information integration Data and text mining Medical decision-support systems and techniques 	<ul style="list-style-type: none"> Information integration Criminal network analysis Data, text, and web mining Identity management and deception detection
Methodology		KDD	KDD
Contributions	Scientific	<ul style="list-style-type: none"> Computer and information science, sociology, policy, legal Clinical medicine and biology 	<ul style="list-style-type: none"> Computer and information science, sociology, policy, legal Criminology, terrorism research
	Practical	<ul style="list-style-type: none"> Public health Patient well-being Biomedical treatment and discovery 	<ul style="list-style-type: none"> Crime investigation and counter-terrorism National and homeland security

In terms of data characteristics, they both face the information stovepipe and information overload problem. In terms of technology development, they both are searching for new approaches, methods, and innovative use of existing techniques. In terms of scientific contributions, they both may add new insights and knowledge to various academic disciplines.

Most importantly, as a consistent research framework based on knowledge management and data mining has begun to emerge in biomedical informatics (Chen et al., 2005), ISI also needs a framework to guide its research. Facing the unique challenges (and associated opportunities) of information overload and the pressing need for advanced criminal and intelligence analyses and investigations, we believe that the Knowledge Discovery from Databases (KDD) methodology (Fayyad and Uthurusamy, 2002), which has achieved significant success in other information-intensive, knowledge-critical domains including business, engineering, biology, and medicine, could be critical in addressing the challenges and problems facing ISI. More details about such a research framework will be discussed in the following chapter.

1.5 Federal Initiatives and Funding Opportunities in ISI

Similar to biocomputing and biomedical informatics, a new, emerging, and critical discipline such as ISI not only can spark the imagination and excitement of society, but it also draws the attention of federal funding agencies. As a testament to the importance of ISI-related research, the United States National Science Foundation (NSF) publicly stated its strong commitment to national security research (as stipulated in its founding mandate), in addition to its traditional leadership role in basic science and engineering research and education. Many federal research and funding agencies in the United States have established new research programs that aim to address different facets of national security research. Without intending to be comprehensive, we summarize some significant past and ongoing federal funding programs of relevance to ISI, especially for academic researchers in universities and research institutes. There is much information about national security-related funding opportunities for commercial companies and vendors which will not be covered in this chapter.

- **National Science Foundation (NSF):** The NSF has issued several Information Technology Research (ITR) program announcements with a national security focus. The computing (CISE) and behavioral (SBE) divisions of NSF are encouraging multi-disciplinary research projects of relevance to ISI. The NSF/CIA KDD (Knowledge Discovery and

Dissemination) program is a good example of joint NSF and intelligence community funding initiatives. Most of the NSF-funded projects stress scientific innovation.

- Department of Homeland Security (DHS): DHS probably has the largest number of research initiatives of relevance to ISI due to its agency mission. Four university-based homeland security research centers have been established. Many new ad hoc initiatives such as terrorism informatics research, bioagent surveillance, smart border, biometrics, deception detection, and critical infrastructure protection are also under active development. DHS-funded projects tend to be more problem-specific. Due to a lack of staff support, personnel turnover, and inexperience in funded research, the DHS proposal review process is significantly less structured than that of NSF.
- Department of Defense (DOD) and Intelligence Community: After the disastrous ending of the TIA (Total Information Awareness) program spearheaded by former Admiral Poindexter, the DOD and the intelligence community have not publicized many new research activities of relevance to ISI. One exception may be the ARDA (Advanced Research & Development Activity) program that aims to develop advanced information technologies for the intelligence community. The abovementioned joint NSF/CIA KDD program that draws top-notch researchers to perform unclassified national security and intelligence research is an excellent model for advanced and potentially high-impact ISI research.
- Center for Disease Control and Prevention (CDC) and National Institutes of Health (NIH): Not surprisingly, CDC and NIH are supporting national security research of relevance to infectious diseases and bioagent surveillance. The National Library of Medicine has issued solicitations in crisis management of relevance to public health. Similar to DHS, CDC appropriates significant funding to state and local jurisdictions for their public health, disease surveillance, and emergency response needs. Although such projects are often short-term and implementation-oriented in nature, university researchers can benefit from collaborating with state and local public health agencies, and vice versa.
- Department of Justice (DOJ): DOJ and its research arm, National Institute of Justice (NIJ), traditionally fund research and development projects of relevance to local and state public safety and law enforcement agencies. Due to the importance of domestic counter-terrorism, focused counter-terrorism research programs have been developed at NIJ. Most of the NIJ projects are of smaller scale and single-PI in nature due to the

agency's limited funding ability. Most projects need to demonstrate significant public-safety relevance and value.

- Office of Naval Research (ONR), Air Force Research Labs, private foundations, and others: There are many other funding opportunities with other traditional armed forces-specific federal funding agencies. Most of them have either adjusted or are adjusting their existing research programs to be of relevance to the new homeland security mission. Several private foundations have also begun to support selected homeland security-related research, in particular, in citizen responses, disaster relief, and education.

For our readers we include a listing of sample federal program announcements of relevance to ISI in the section below. The list is not intended to be comprehensive, but to illustrate the abundant research and funding opportunities in ISI. Readers are advised to perform periodic searches on the web for ongoing program announcements.

- National Science Foundation (NSF), Information Technology Research (ITR) Program: The NSF ITR Program has been successful in opening up opportunities at the frontiers of IT research and education. In its fourth year, FY 2003, the program stimulated research on the fundamental challenges facing the continued expansion and utilization of IT across the sciences and engineering, the creation of novel uses of IT, the interaction of IT with society at large, and the use of IT to enhance security and reduce society's vulnerabilities to catastrophic events, whether natural or man-made. In FY 2004, the focus is "ITR for National Priorities." Particular emphasis is placed on the distributed systems, grids and infrastructures that support the attainment of these national priorities.
- Department of Homeland Security (DHS): Through the Homeland Security Centers of Excellence Program, DHS is investing in university-based partnerships to develop centers of multi-disciplinary research where important fields of inquiry can be analyzed and best practices developed, debated, and shared. DHS selected the University of Southern California to house the first HS-Center, known as the Homeland Security Center for Risk and Economic Analysis of Terrorism Events (CREATE). Texas A&M University and its partners were selected as the HS-Center for Foreign Animal and Zoonotic Disease Defense. The University of Minnesota and its partners were selected as the HS-Center for Food Protection and Defense, which will address agro-security issues related to post-harvest food protection. The University of Maryland's team and its partners were selected recently as the fourth HS-

center to study the social, economic, and psychological dimensions of terrorism research.

The DHS Innovative Architectures for Unified Incident Command and Decision Support Program seeks proposals for an innovative information management and sharing architecture that answers the growing needs of the emergency responder community. This solicitation seeks to confront the technical challenges associated with the development of such innovative, modular, scaleable, and secure information management architecture.

The DHS Security Research and Development Program seeks the development and deployment of technologies to protect the nation's cyber infrastructure, including the Internet and other critical infrastructures that depend on computer systems for their mission.

- National Institutes of Health (NIH), National Library of Medicine (NLM), Informatics for Disaster Management Program: The National Library of Medicine, the National Institute of Mental Health, and the National Institute of Biomedical Imaging and Bioengineering support informatics research that addresses biomedical information management problems relevant to management of disasters. Disasters can be caused by nature or by man, through accident or by malice. Terrorism, particularly bioterrorism, is now an important focus of federal activity, but terrorism is only one of a number of threats to public safety classified as disasters. Disaster management is heavily dependent on efficient flow of information. How best to utilize information technology in a disaster situation poses a number of problems for which relevant informatics research is necessary.
- Center for Disease Control and Prevention (CDC), National Center for Infectious Diseases (NCID), Bioterrorism Extramural Research Grant Program: The main functions of the CDC NCID Office include: program coordination, surveillance integration and informatics activities, and analytical activities of relevance to infectious diseases. The NCID Office of Extramural Research funded \$8.4 million dollars for nine biodefense research grants in FY03. The NCID areas of interest for bioterrorism and infectious disease research include: surveillance and epidemiology, interruption of transmission, and environmental detection.
- Department of Defense (DOD), Advanced Research and Development Activity (ARDA) Program: The Advanced Research and Development Activity (ARDA) program is an Intelligence Community (IC) center for conducting advanced research and development related to information technology (IT) (information stored, transmitted, or manipulated by

electronic means). ARDA sponsors high risk, high payoff research designed to produce new technology to address some of the most important and challenging IT problems faced by the intelligence community. The research is currently organized into five technology thrusts: Information Exploitation, Quantum Information Science, Global Infosystems Access, Novel Intelligence from Massive Data, and Advanced Information Assurance. Recent solicitations include: Advanced Question and Answering for Intelligence; Video Analysis and Content Extraction (VACE) Phase II; Proactive and Predictive Information Assurance for Next Geneva Systems; Advanced Countermeasures for Insider Threat; Information Assurance for the U.S. Intelligence Community; and Advanced Question and Answering Phase I.

- Department of Justice (DOJ), National Institute of Justice (NIJ): The National Institute of Justice (NIJ) is the research, development, and evaluation agency of the U.S. Department of Justice and a component of the Office of Justice Programs. NIJ provides objective, independent, evidence-based knowledge and tools to enhance the administration of justice and public safety.

The NIJ's Situational Aspects of Crime Program seeks research that examines situational characteristics and the events that lead up to criminal acts in order to identify target points for prevention and intervention. Research under this solicitation should focus on the characteristics of criminal events or the interactions between the characteristics of situations and individuals.

Their Terrorism and Transnational Crime Program seeks research and evaluation that will inform national policy and practice related to terrorism, transnational criminal behavior, and any connections between them. Proposed research should aim to improve criminal justice and first responder strategies for prevention of, preparation for, response to, and mitigation of terrorist incidents at the federal, state, and local levels.

1.6 Future Directions

The emergence of a new discipline such as ISI would require careful cultivation and development by many top-notch researchers and practitioners from many different disciplines, including (but not limited to): computer science, information science, information systems, electrical engineering, social science, law, public policy, criminal justice, terrorism research, psychology, behavioral and economic sciences, management science,

bioinformatics, public health, etc. There is an abundance of opportunities for developing new and innovative funded ISI-related projects.

Regardless of which funding programs you may be considering for your research, there are some common characteristics among successfully funded and (eventually, after execution) high-impact projects:

- Unique and critical scientific or engineering innovations: You need to clearly distinguish your research from others.
- Important problems and significant partners: You need to address important national security problems and demonstrate your commitment to address these problems with the help and support of your local, state, and federal agency partners.
- From small to large: Most funded projects began humbly with proof-of-concept level funding.
- A multi-disciplinary team: After initial success, a multi-disciplinary team of computer scientists, system developers, social scientists, policy and legal experts, domain (intelligence and security) experts, and such will be needed to implement a full-scale, multi-phased, complex national security-related project.
- Aim high: The following tangible (but somewhat lofty) project goals are always good for your project team to aim at: (1) publishing your project findings in *Science*, *Nature*, or *Proceedings of the Academy of Science* (for its scientific contributions); and (2) being featured in a *New York Times* or *USA Today* front page article (for its societal impact).

1.7 Questions for Discussion

1. What are some ways to identify potentially relevant federal funding programs? How can the topics to focus on be determined? How do you interact with relevant program managers and directors for feedback? How do you get invited to a PI-only meeting when you are not a PI yet (but would like to become a PI)? How frequently do you need to travel to DC for project meetings?
2. What are some ways to identify critical local, state, and federal partners for your research project? How do you convince them of your value? How do you get their letters of support, data, and commitment of time and personnel? How can you make the system work for all involved?

3. What are some ways to publish your work in this area? Which conferences and journals would be best? How can you publicize your work with the media and the press?



<http://www.springer.com/978-0-387-24379-5>

Intelligence and Security Informatics for International
Security

Information Sharing and Data Mining

Chen, H.

2006, XVIII, 182 p., Hardcover

ISBN: 978-0-387-24379-5