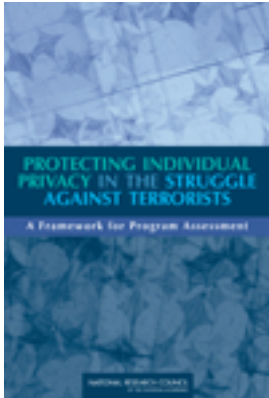


Free Executive Summary



Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment

Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, National Research Council

ISBN: 978-0-309-12488-1, 376 pages, 6 x 9, paperback (2008)

This free executive summary is provided by the National Academies as part of our mission to educate the world on issues of science, engineering, and health. If you are interested in reading the full book, please visit us online at <http://www.nap.edu/catalog/12452.html>. You may browse and search the full, authoritative version for free; you may also purchase a print or electronic version of the book. If you have questions or just want more information about the books published by the National Academies Press, please contact our customer service department toll-free at 888-624-8373.

All U.S. agencies with counterterrorism programs that collect or "mine" personal data -- such as phone records or Web sites visited -- should be required to evaluate the programs' effectiveness, lawfulness, and impacts on privacy. A framework is offered that agencies can use to evaluate such information-based programs, both classified and unclassified. The report urges Congress to re-examine existing privacy law to assess how privacy can be protected in current and future programs and recommends that any individuals harmed by violations of privacy be given a meaningful form of redress. Two specific technologies are examined: data mining and behavioral surveillance. Regarding data mining, the report concludes that although these methods have been useful in the private sector for spotting consumer fraud, they are less helpful for counterterrorism because so little is known about what patterns indicate terrorist activity. Regarding behavioral surveillance in a counterterrorist context, the report concludes that although research and development on certain aspects of this topic are warranted, there is no scientific consensus on whether these techniques are ready for operational use at all in counterterrorism.

This executive summary plus thousands more available at www.nap.edu.

Copyright © National Academy of Sciences. All rights reserved. Unless otherwise indicated, all materials in this PDF file are copyrighted by the National Academy of Sciences. Distribution or copying is strictly prohibited without permission of the National Academies Press <http://www.nap.edu/permissions/>. Permission is granted for this material to be posted on a secure password-protected Web site. The content may not be posted on a public Web site.

Executive Summary

In a democratic society it is vitally important that citizens and their representatives be able to make an informed judgment on how to appropriately balance privacy with security. This report seeks to contribute to that informed judgment.

September 11, 2001, provided vivid proof to Americans of the damage that a determined, fanatical terrorist group can inflict on our society. Based on the available information about groups like Al Qaeda, most importantly their own statements, it seems clear that they will continue to try to attack us. Further attacks by such groups, and indeed by domestic terrorists like Timothy McVeigh, could be as serious as, or even more serious than, September 11 and Oklahoma City. Because future terrorist attacks on the United States could cause major casualties as well as severe economic and social disruption, the danger they pose is real, and it is serious. Thus, high priority should be given to developing programs to detect intended attacks before they occur so that there is a chance of preventing them.

At the same time, the nation must ensure that its institutions, information systems, and laws together constitute a trustworthy and accountable system that protects U.S. citizens' rights to privacy.

In this report, the Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals examines the role of data mining and behavioral surveillance technologies in

2 PROTECTING INDIVIDUAL PRIVACY IN THE STRUGGLE AGAINST TERRORISTS

counterterrorism programs,¹ and it provides a framework for making decisions about deploying and evaluating those and other information-based programs on the basis of their effectiveness and associated risks to personal privacy.

The most serious threat today comes from terrorist groups that are international in scope. These groups make use of the Internet to recruit, train, and plan operations, and they use public channels to communicate. Therefore, intercepting and analyzing these information streams might provide important clues regarding the nature of the terrorist threat. Important clues might also be found in commercial and government databases that record a wide range of information about individuals, organizations, and their transactions, movements, and behavior. But success in such efforts will be extremely difficult to achieve because:

- The information sought by analysts must be filtered out of the huge quantity of data available (the needle in the haystack problem); and
- Terrorist groups will make calculated efforts to conceal their identity and mask their behaviors, and will use various strategies such as encryption, code words, and multiple identities to obfuscate the data they are generating and exchanging.

Modern data collection and analysis techniques have had remarkable success in solving information-related problems in the commercial sector; for example, they have been successfully applied to detect consumer fraud. But such highly automated tools and techniques cannot be easily applied to the much more difficult problem of detecting and preempting a terrorist attack, and success in doing so may not be possible at all. Success, if it is indeed achievable, will require a determined research and development effort focused on this particular problem.

Detecting indications of ongoing terrorist activity in vast amounts of communications, transactions, and behavioral records will require technology-based counterterrorism tools. But even in well-managed programs such tools are likely to return significant rates of false positives, especially if the tools are highly automated. Because the data being analyzed are primarily about ordinary, law-abiding citizens and businesses, false positives can result in invasion of their privacy. Such intrusions raise valid concerns about the misuse and abuse of data, about the accuracy

¹In this report, the term “program” refers to the system of technical, human, and organizational resources and activities required to execute a specific function. Humans—not computers—are always fully responsible for the actions of a program.

of data and the manner in which the data are aggregated, and about the possibility that the government could, through its collection and analysis of data, inappropriately influence individuals' conduct. Intruding on privacy also risks ignoring constitutional concerns about general search, as reflected in the Fourth Amendment. The committee strongly believes that such intrusion must be minimized through good management and good design, even if it cannot be totally eliminated.

The difficulty of detecting the activity of terrorist groups through their communications, transactions, and behaviors is hugely complicated by the ubiquity and enormity of electronic databases maintained by both government agencies and private-sector corporations. Retained data and communication streams concern financial transactions, medical records, travel, communications, legal proceedings, consumer preferences, Web searches, and, increasingly, behavioral and biological information. This is the essence of the information age—it provides us with convenience, choice, efficiency, knowledge, and entertainment; it supports education, health care, safety, and scientific discovery. Everyone leaves personal digital tracks in these systems whenever he or she makes a purchase, takes a trip, uses a bank account, makes a phone call, walks past a security camera, obtains a prescription, sends or receives a package, files income tax forms, applies for a loan, e-mails a friend, sends a fax, rents a video, or engages in just about any other activity. The proliferation of security cameras and means of tagging and tracking people and objects increases the scope and nature of available data. Law-abiding citizens leave extensive digital tracks, and so do criminals and terrorists.

Gathering and analyzing electronic, behavioral, biological, and other information can play major roles in the prevention, detection, and mitigation of terrorist attacks, just as they do against other criminal threats. In fact the U.S. government has increased its investment in counterterrorism programs based on communications surveillance, data mining, and information fusion. Counterterrorism agencies are particularly interested in merging several different databases (information fusion) and then probing the combined data to understand transactions and interactions of specific persons or organizations of interest (data mining). They would also like to identify individuals (through data mining and behavioral surveillance) whose transactions and behavior might indicate possible terrorist links.

Such techniques often work well in commercial settings, for example for fraud detection, where they are applied to highly structured databases and are honed through constant use and learning. But the problems confronting counterterrorism analysts are vastly more difficult. Automated identification of terrorists through data mining (or any other known

methodology) is neither feasible as an objective nor desirable as a goal of technology development efforts.

One reason is that collecting and examining information to inhibit terrorists inevitably conflicts with efforts to protect individual privacy. And when privacy is breached, the damage is real. The degree to which privacy is compromised is fundamentally related to the sciences of database technology and statistics as well as to policy and process. For example, there is no way to make personal information in databases fully anonymous. Technical, operational, legal, policy, and oversight processes to minimize privacy intrusion and the damage it causes must be established and uniformly applied. Even under the pressure of threats as serious as terrorism, the privacy rights and civil liberties that are the cherished core values of our nation must not be destroyed.

The quality of the data used in the difficult task of preempting terrorism is also a substantial issue. Data of high quality are correct, current, complete, and relevant, and so they can be used effectively, economically, and rapidly to inform and evaluate decisions. Data derived by linking high-quality data with data of lesser quality will tend to be low-quality data. Because data of questionable quality are likely to be the norm in counterterrorism, analysts must be cognizant of their effects, especially in fused or linked databases, and officials must carefully consider the consequent likelihood of false positives and privacy intrusions.

The preliminary nature of the scientific evidence, the risk of false positives, and operational vulnerability to countermeasures argue for behavioral observation and physiological monitoring being used at most as a preliminary screening method for identifying individuals who merit additional follow-up investigation. Although laboratory research and development of techniques for automated, remote detection and assessment of anomalous behavior, for example deceptive behavior, may be justified, there is not a consensus within the relevant scientific community nor on the committee regarding whether any behavioral surveillance or physiological monitoring techniques are ready for use at all in the counterterrorist context given the present state of the science.

The committee has developed and provides in Chapter 2 a specific framework for evaluation and operation of information-based counterterrorism programs to guide deployment decisions and facilitate continual improvement of the programs.

National security authorities of course should always adhere to the law, but the committee recognizes that laws will have to be reviewed and revised from time to time to ensure that they are appropriate, up to date, and responsive to real needs and contemporary technologies.

With these several concerns and issues in mind, the committee makes the following recommendations.

Recommendation 1. U.S. government agencies should be required to follow a systematic process (such as the one described in the framework proposed in Chapter 2) to evaluate the effectiveness, lawfulness, and consistency with U.S. values of every information-based program, whether classified or unclassified, for detecting and countering terrorists before it can be deployed, and periodically thereafter. Under most circumstances, this evaluation should be required as a condition for deployment of information-based counterterrorism programs, but periodic evaluation and continual improvement should *always* be required when such programs are in use. The committee believes that the framework presented in Chapter 2 defines an appropriate process for this purpose.

Periodically after a program has been operationally deployed, and in particular before a program enters a new phase in its life cycle, policy makers should apply a framework such as the one proposed in Chapter 2 to the program before allowing it to continue operations or to proceed to the next phase. Consistency with relevant laws and regulations, and impact on individual privacy and civil liberties—as well as validity, effectiveness, and technical performance—should be rigorously assessed. Such review is especially necessary given that the committee found little evidence of any effective evaluation performed for current programs intended to detect terrorist activity by automated analysis of databases. (If such evidence does exist, it should be presented in the appropriate oversight forums as part of such review.) Periodic review may result in significant modification of a program or even its cancellation.

Any information-based counterterrorism program of the U.S. government should be subjected to robust, independent oversight. All three branches of government have important roles to play to ensure that such programs adhere to relevant laws. All such programs should provide meaningful redress to any individuals inappropriately harmed by their operation.

To protect the privacy of innocent people, the research and development of any information-based counterterrorism program should be conducted with synthetic population data. If and when a program meets the criteria for deployment in the committee's illustrative framework described in Chapter 2, it should be deployed only in a carefully phased manner, e.g., being field tested and evaluated at a modest number of sites before being scaled up for general use. At all stages of a phased deployment, data about individuals should be rigorously subjected to the full safeguards of the framework.

Recommendation 2. The U.S. government should periodically review the nation's laws, policies, and procedures that protect individuals' private

information for relevance and effectiveness in light of changing technologies and circumstances. In particular, Congress should reexamine existing law to consider how privacy should be protected in the context of information-based programs (e.g., data mining) for counterterrorism. Such reviews should consider establishment of restrictions on how personal information can be used. Currently, legal restrictions are focused primarily on how records are collected and assessed, rather than on their use.

PROTECTING INDIVIDUAL PRIVACY IN THE STRUGGLE AGAINST TERRORISTS

A Framework for Program Assessment

Committee on Technical and Privacy Dimensions of Information
for Terrorism Prevention and Other National Goals

Committee on Law and Justice and Committee on National Statistics
Division on Behavioral and Social Sciences and Education

Computer Science and Telecommunications Board
Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Support for this project was provided by the Bureau of Transportation Statistics, with assistance from the National Science Foundation under sponsor award number SES-0112521; the Department of Homeland Security, with assistance from the National Science Foundation under sponsor award number SES-0411897; the National Center for Education Statistics, with assistance from the National Science Foundation under sponsor award number SBR-0453930; and the National Science Foundation under sponsor award numbers SRS-0632055 and IIS-0441216. Additional funding was provided by the Presidents' Circle Communications Initiative of the National Academies.

Library of Congress Cataloging-in-Publication Data

Protecting individual privacy in the struggle against terrorists : a framework for program assessment.

p. cm.

Includes bibliographical references.

ISBN 978-0-309-12488-1 (pbk.) — ISBN 978-0-309-12489-8 (pdf) 1. Terrorism—United States—Prevention. 2. Surveillance detection—United States. 3. Privacy, Right of—United States. 4. Technological innovations—Law and legislation—United States.

HV6432.P76 2008

363.325'163--dc22

2008033554

This report is available from

Committee on Law and Justice *or*
Computer Science and Telecommunications Board
National Research Council
500 Fifth Street, N.W.
Washington, DC 20001

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>.

Copyright 2008 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

**COMMITTEE ON TECHNICAL AND PRIVACY
DIMENSIONS OF INFORMATION FOR TERRORISM
PREVENTION AND OTHER NATIONAL GOALS**

WILLIAM J. PERRY, Stanford University, *Co-chair*
CHARLES M. VEST, National Academy of Engineering, *Co-chair*
W. EARL BOEBERT, Sandia National Laboratories
MICHAEL L. BRODIE, Verizon Communications
DUNCAN A. BROWN, Johns Hopkins University
FRED H. CATE, Indiana University
RUTH A. DAVID, Analytic Services, Inc.
RUTH M. DAVIS, Pymatuning Group, Inc.
WILLIAM H. DuMOUCHEL, Lincoln Technologies, Inc.
CYNTHIA DWORK, Microsoft Research
STEPHEN E. FIENBERG, Carnegie Mellon University
ROBERT J. HERMANN, Global Technology Partners, LLC
R. GIL KERLIKOWSKA, Seattle Police Department
ORIN S. KERR, George Washington University Law School
ROBERT W. LEVENSON, University of California, Berkeley
TOM M. MITCHELL, Carnegie Mellon University
TARA O'TOOLE, University of Pittsburgh Medical Center
DARYL PREGIBON, Google, Inc.
LOUISE RICHARDSON, Harvard University
BEN A. SHNEIDERMAN, University of Maryland
DANIEL J. WEITZNER, Massachusetts Institute of Technology

Staff

BETTY M. CHEMERS, Committee on Law and Justice
CAROL PETRIE, Committee on Law and Justice
JULIE ANNE SCHUCK, Committee on Law and Justice
MICHAEL L. COHEN, Committee on National Statistics
HERBERT S. LIN, Computer Science and Telecommunications Board
JANICE M. SABUDA, Computer Science and Telecommunications
Board (through April 2008)

COMMITTEE ON LAW AND JUSTICE (DBASSE)

JAMES Q. WILSON, University of California, Los Angeles (Emeritus),
Chair

PHILIP J. COOK, Terry Sanford Institute of Public Policy, Duke
University, *Vice Chair*

DAVID H. BAYLEY, University of Albany, State University of New York

RICHARD J. BONNIE, University of Virginia Law School

MARTHA CRENSHAW, Wesleyan University

ROBERT D. CRUTCHFIELD, University of Washington

JOHN J. DIIULIO, JR., University of Pennsylvania

STEVEN N. DURLAUF, University of Wisconsin, Madison

JOHN A. FERREJOHN, Stanford University

ARTHUR S. GOLDBERGER, University of Wisconsin, Madison

BRUCE HOFFMAN, RAND Corporation

ROBERT L. JOHNSON, New Jersey Medical School

JOHN H. LAUB, University of Maryland

TRACEY L. MEARES, University of Chicago

TERRIE E. MOFFITT, University of London

MARK H. MOORE, Harvard University

RUTH PETERSON, Ohio State University

RICHARD ROSENFELD, University of Missouri–St. Louis

ROBERT J. SAMPSON, Department of Sociology, Harvard University

JEREMY TRAVIS, Jay College of Criminal Justice, New York

CHRISTY VISHER, The Urban Institute

CAROL PETRIE, Director

BETTY CHEMERS, Senior Program Officer

LINDA DePUGH, Program Associate

COMMITTEE ON NATIONAL STATISTICS (DBASSE)

WILLIAM F. EDDY, Department of Statistics, Carnegie Mellon
University, *Chair*

KATHARINE ABRAHAM, University of Maryland

ROBERT BELL, AT&T Research Laboratories

WILLIAM DuMOUCHEL, Lincoln Technologies, Inc.

JOHN HALTIWANGER, University of Maryland

V. JOSEPH HOTZ, University of California, Los Angeles

KAREN KAFADAR, University of Colorado, Denver, and Health
Sciences Center

DOUGLAS MASSEY, Princeton University

VIJAY NAIR, University of Michigan, Ann Arbor

JOSEPH NEWHOUSE, Harvard University

SAMUEL H. PRESTON, University of Pennsylvania

KENNETH PREWITT, Columbia University

LOUISE RYAN, Harvard University

NORA CATE SCHAEFFER, University of Wisconsin, Madison

ALAN ZASLAVSKY, Harvard University Medical School

CONSTANCE F. CITRO, Director

COMPUTER SCIENCE AND TELECOMMUNICATIONS BOARD (DEPS)

JOSEPH F. TRAUB, Columbia University, *Chair*
PRITHVIRAJ BANERJEE, Hewlett Packard Company
FREDERICK R. CHANG, University of Texas, Austin
WILLIAM DALLY, Stanford University
MARK E. DEAN, IBM Almaden Research Center
DEBORAH ESTRIN, University of California, Los Angeles
KEVIN KAHN, Intel Corporation
JAMES KAJIYA, Microsoft Corporation
RANDY H. KATZ, University of California, Berkeley
JOHN E. KELLY III, IBM
SARA KIESLER, Carnegie Mellon University
PETER LEE, Carnegie Mellon University
TERESA H. MENG, Stanford University
WILLIAM H. PRESS, University of Texas, Austin
PRABHAKAR RAGHAVAN, Yahoo! Research
ALFRED Z. SPECTOR, Google, Inc.
ROBERT F. SPROULL, Sun Microsystems, Inc.
PETER SZOLOVITS, Massachusetts Institute of Technology
ANDREW J. VITERBI, Viterbi Group, LLC
PETER WEINBERGER, Google, Inc.

JON EISENBERG, Director
KRISTEN R. BATCH, Associate Program Officer
RENEE HAWKINS, Financial and Administrative Manager
HERBERT S. LIN, Chief Scientist
LYNETTE I. MILLETT, Senior Program Officer
MORGAN R. MOTTO, Program Associate
ERIC WHITAKER, Senior Program Assistant

For more information on CSTB, see its Web site at <http://www.cstb.org>, write to CSTB, National Research Council, 500 Fifth Street, N.W., Washington, DC 20001, call (202) 334-2605, or e-mail the CSTB at cstb@nas.edu.

Preface

In late 2005, the National Research Council (NRC) convened the Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals. Supported by the U.S. Department of Homeland Security and the National Science Foundation, the committee was charged with addressing information needs of the government that arise in its deployment of various forms of technology for broad access to and analysis of data as it faces the challenges of terrorism prevention and threats to public health and safety. Specifically of interest was the nexus between terrorism prevention, technology, privacy, and other policy issues and the implications and issues involved in deploying data mining, information fusion, and behavioral surveillance technologies. The study sought to develop a conceptual framework that policy makers and the public can use to consider the utility, appropriateness, and empirical validity of data generated and analyzed by various forms of technology currently in use or planned in the near future. The committee notes that the development of this framework did not include the development of systems for preventing terrorism. By design and in response to the charge for the study, this report focuses on data mining and behavioral surveillance as the primary techniques of interest.

The committee interpreted its charge as helping government policy makers to evaluate and make decisions about information-based programs to fight terrorism or serve other important national goals, and it thus sought to provide a guide for government officials, policy makers, and technology developers as they continue to explore new surveillance

tools in the service of important national security goals. Chapter 1 scopes the issues involved and introduces key concepts that are explored in much greater depth in the appendixes. Chapter 2 outlines a framework for a systematic assessment of information-based programs being considered or already in use for counterterrorist purposes (and other important national needs, such as law enforcement and public health) in terms of each program's effectiveness and its consistency with U.S. laws and values. Chapter 3 provides the committee's conclusions and recommendations. The appendixes elaborate extensively on the scientific and technical foundations that underpin the committee's work and the legal and organizational context in which information-based programs necessarily operate. The committee regards the appendixes as essential elements of the report.

Note that although the committee heard from representatives from many government agencies, this report does not evaluate or critique any specific U.S. government program. Rather, it is intended to provide policy makers with a systematic framework for thinking about existing and future operational information-based programs, especially in a counterterrorist context.

Nowhere is the need for this study and the framework it proposes more apparent than in the history of the Total Information Awareness (TIA) program. Indeed, the TIA program and the issues it raised loomed large in the background when this committee was appointed, and although the TIA program was terminated in September 2003, it is safe to say that the issues raised by this program have not been resolved in any fundamental sense. Moreover, many other data mining activities supported by the U.S. government continue to raise the same issues: the potential utility of large-scale databases containing personal information for counterterrorist and law enforcement purposes and the potential privacy impact of law enforcement and national security authorities using such databases. A brief history of the TIA program is contained in Appendix J.

The committee consisted of 21 people with a broad range of expertise, including national security and counterterrorism, intelligence and counterintelligence, privacy law and information protection, organizations and organizational structure, law enforcement, statistics, information technology, cognitive psychology, terrorism, database architecture, public health, artificial intelligence, databases, cryptography, machine learning and statistics, and information retrieval.

From 2005 to 2007, the committee held six meetings, most of which were intended to enable it to explore a wide range of points of view. For example, briefings and other inputs were obtained from government officials at all levels, authorities on international law and practice relat-

ing to policy, social scientists and philosophers concerned with collection of personal data, experts on privacy-enhancing technologies, business representatives concerned with the gathering and uses of personal data, and researchers who use personal data in their work. Several papers were commissioned and received, as well as a number of contributed white papers.

Preparation of the report was undertaken on an unclassified basis. Although a number of classified programs of the U.S. government make use of data mining, the fundamental principles of data mining themselves are not classified, and these principles apply to both classified and unclassified applications. Thus, at the level of analysis presented in this report, the fact that some of the U.S. government's counterterrorist programs are classified does not materially affect the analysis provided here. In addition, the U.S. government operates a variety of classified programs intended to collect data that may be used for counterterrorist purposes. However, as collection programs, they are out of the scope of this report, and all that need be noted is that they produce data relevant to the counterterrorist mission and that data mining and information fusion technologies must process.

This study could not have been undertaken without the support of the government project officers, Larry Willis, U.S. Department of Homeland Security, and Larry Brandt and Brian D. Humes, National Science Foundation, who recognize the complex issues involved in developing and using new technologies to respond to terrorism and other national efforts, such as law enforcement and public health, and the need to think through how this might best be done.

Given the scope and breath of the study, the committee benefited greatly from the willingness of many individuals to share their perspectives and expertise. We are very grateful to the following individuals for their helpful briefings on technologies for data mining and detection of deception: Paul Ekman, University of California, San Francisco; Mark Frank, University of Buffalo; John Hollywood, RAND Corporation; David Jensen, University of Massachusetts; Jeff Jonas, IBM; David Scott, Rice University; John Woodward, RAND Corporation; and Thomas Zeffiro, Georgetown University. Useful insights on the use of these technologies in the private sector were provided by Scott Loftnesness, Glenbrook Partners, and Dan Schutzer, Financial Services Technical Consortium. William Winkler, Census Bureau, helped the committee understand the technologies' potential impact on federal statistical agencies.

Background briefings on relevant privacy law and policy were provided by Henry Greely, Stanford University; Barry Steinhardt, American Civil Liberties Union; Kim Taipale, Center for Advanced Studies in Science and Technology Policy; and Lee Tien, Electronic Frontier Founda-

tion. We also benefited from the expert testimony of Whitfield Diffie, Sun Microsystems; John Pike, Global Security; and Jody Westby, Global Cyber Risk, on the role of information technologies in counterterrorism. In addition to counterterrorism, the impact and implications of data mining for law enforcement and public health were important foci of the committee's work. In the public health area, the following persons contributed to the committee's understanding: James Lawler, Homeland Security Council, White House; Farzad Mostashari, New York City Public Health Department; Patricia Quinlisk, State of Iowa; and Barry Rhodes and Lynn Steele, Centers for Disease Control and Prevention. Useful insights on the role of law enforcement in counterterrorism were provided in presentations made by Roy Apseloff, National Media Exploitation Center; Michael Fedarcyk, Federal Bureau of Investigation (retired); and Philip Reiting, Microsoft. We found extremely helpful the international perspectives of Joe Connell, New Scotland Yard (retired), and Ravi Ron, former head of Israel's Ben Gurion Airport.

This study also benefited considerably from briefings by government officials involved on a daily basis with the issues at the heart of the study. We particularly want to thank Randy Ferryman and Admiral Scott Redd from the National Counter Terrorism Center and Clint C. Brooks (retired) from the National Security Agency, who shared their vision of how the nation should conduct its counterterrorism activities while maintaining its democratic ideals. Numerous staff members from the Department of Homeland Security (DHS) also shed important light on government activities relating to terrorism prevention, including Mel Bernstein, Timothy Keefer, Hyon Kim, Sandy Landsberg, John V. Lawler, Tiffany Lightbourn, Grace Mastalli, Allison Smith, and Lisa J. Walby. Toby Levin was particularly helpful in sharing timely and relevant information on the work of the DHS Privacy Office, and the committee appreciated the interest of the DHS Data Privacy and Integrity Advisory Committee in its work and their willingness to keep members abreast of their activities and role in protecting privacy.

The committee also thanks Michael D. Larsen of Iowa State University and Peter Swire of Ohio State University, who responded to its request for white papers, and Amy Corning and Eleanor Singer, University of Michigan, who prepared an informative paper on public opinion.

This study involved NRC staff from three different NRC units. We would like to thank them for their valuable assistance to this project as well as for their collegiality, which contributed to a far richer experience for all involved. Betty Chemers of the NRC's Committee on Law and Justice served as study director and organized and facilitated the meetings, Michael Cohen of the Committee on National Statistics provided technical expertise on statistical and data mining issues, and Herbert

Lin of the Computer Science and Telecommunications Board undertook the difficult job of turning the committee's writing contributions into a coherent whole and working with the co-chairs to mediate and resolve intellectual disagreements within the committee. Carol Petrie provided guidance and support throughout the study process. We would also like to thank Julie Schuck and Ted Schmitt for their research assistance and Jennifer Bishop, Barbara Boyd, Linda DePugh, and Janice Sabuda for their administrative support. Finally, we greatly appreciate the efforts undertaken by Eugenia Grohman, Susan Maurizi, Kirsten Sampson Snyder, and Yvonne Wise to complete the review and editing processes and bring this report to fruition.

Charles M. Vest and William J. Perry, *Co-chairs*
Committee on Technical and Privacy
Dimensions of Information for Terrorism
Prevention and Other National Goals

Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Steve M. Bellovin, Columbia University,
R. Stephen Berry, University of Chicago,
David L. Carter, Michigan State University,
Richard F. Celeste, Colorado College,
Hermann Habermann, Bureau of the U.S. Census (retired),
David Jensen, University of Massachusetts, Amherst,
Alan F. Karr, National Institute of Statistical Sciences,
Diane Lambert, Google, Inc.,
Butler Lampson, Microsoft Corporation,
Michael D. Larsen, Iowa State University,
Lance Liebman, Columbia Law School,
Patricia Quinlisk, State of Iowa,
Jerome Reiter, Duke University,

Andrew P. Sage, George Mason University,
Paul Schwartz, University of California, Berkeley,
Eugene Spafford, Purdue University,
Robert D. Sparks, California Medical Association Foundation,
William O. Studeman, Northrop Grumman Mission Systems, and
Peter Weinberger, Google, Inc.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by William H. Press, University of Texas at Austin, and James G. March, Stanford University. Appointed by the National Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

EXECUTIVE SUMMARY	1
1 SCOPING THE ISSUE: TERRORISM, PRIVACY, AND TECHNOLOGY	7
1.1 The Nature of the Terrorist Threat to the United States, 7	
1.2 Counterterrorism and Privacy as an American Value, 8	
1.3 The Role of Information, 11	
1.4 Organizational Models for Terrorism and the Intelligence Process, 15	
1.5 Activities of the Intelligence Community and of Law Enforcement Agencies, 17	
1.6 Technologies of Interest in This Report, 19	
1.6.1 Data Mining, 20	
1.6.2 Behavioral Surveillance, 24	
1.7 The Social and Organizational Context, 26	
1.8 Key Concepts, 27	
1.8.1 The Meaning of Privacy, 27	
1.8.2 Effectiveness, 29	
1.8.3 Law and Consistency with Values, 30	
1.8.4 False Positives, False Negatives, and Data Quality, 35	
1.8.5 Oversight and Prevention of Abuse, 41	
1.9 The Need for a Rational Assessment Process, 42	

2	A FRAMEWORK FOR EVALUATING INFORMATION-BASED PROGRAMS TO FIGHT TERRORISM OR SERVE OTHER IMPORTANT NATIONAL GOALS	44
2.1	The Need for a Framework for Evaluating Information-Based Programs, 44	
2.2	Evaluating Effectiveness, 47	
2.3	Evaluating Consistency with U.S. Law and Values, 52	
2.3.1	Data, 53	
2.3.2	Programs, 54	
2.3.3	Administration and Oversight, 56	
2.4	A Note for Policy Makers: Applying the Framework in the Future, 57	
2.5	Summary of Framework Criteria, 59	
2.5.1	For Evaluating Effectiveness, 59	
2.5.2	For Evaluating Consistency with Laws and Values, 61	
2.5.3	For Developing New Laws and Policies, 63	
3	CONCLUSIONS AND RECOMMENDATIONS	67
3.1	Basic Premises, 67	
3.2	Conclusions Regarding Privacy, 71	
3.2.1	Protecting Privacy, 71	
3.2.2	Distinctions Between Capability and Intent, 75	
3.3	Conclusions Regarding the Assessment of Counterterrorism Programs, 75	
3.4	Conclusions Regarding Data Mining, 76	
3.4.1	Policy and Law Regarding Data Mining, 76	
3.4.2	The Promise and Limitations of Data Mining, 77	
3.5	Conclusions Regarding Deception Detection and Behavioral Surveillance, 82	
3.6	Conclusions Regarding Statistical Agencies, 84	
3.7	Recommendations, 86	
3.7.1	Systematic Evaluation of Every Information-Based Counterterrorism Program, 86	
3.7.2	Periodic Review of U.S. Law, Policy, and Procedures for Protection of Privacy, 95	

APPENDIXES

A	Acronyms	105
B	Terrorism and Terrorists	111
B.1	The Nature of Terrorism, 111	

B.2	Some Tactics of Terrorism, 113	
B.3	A Historical Perspective on Terrorism, 114	
B.4	Explaining Terrorism, 114	
B.5	Al Qaeda and the Terrorist Threat to the United States, 115	
B.6	Terrorists and Their Supporting Technologies, 118	
B.7	Looking to the Future, 119	
C	Information and Information Technology	120
C.1	The Information Life Cycle, 120	
C.1.1	Information Collection, 120	
C.1.2	Information Correction and Cleaning, 121	
C.1.3	Information Storage, 122	
C.1.4	Information Analysis and Use, 122	
C.1.5	Information Sharing, 122	
C.1.6	Information Monitoring, 123	
C.1.7	Information Retention, 124	
C.1.8	Issues Related to Data Linkage, 126	
C.1.9	Connecting the Information Life Cycle to the Framework, 126	
C.2	The Underlying Communications and Information Technology, 128	
C.2.1	Communications Technology, 128	
C.2.2	Information Technology, 129	
C.2.3	Managing Information Technology Systems and Programs, 131	
D	The Life Cycle of Technology, Systems, and Programs	133
E	Hypothetical and Illustrative Applications of the Framework to Various Scenarios	137
E.1	Airport Security, 137	
E.1.1	The Threat, 137	
E.1.2	A Possible Technological Approach to Addressing the Threat, 138	
E.1.3	Possible Privacy Impacts, 139	
E.1.4	Applying the Framework, 140	
E.2	Syndromic Surveillance, 141	
E.2.1	The Threat, 141	
E.2.2	A Possible Technological Approach to Addressing the Threat, 141	
E.2.3	Possible Privacy Impacts, 142	
E.2.4	Applying the Framework, 144	

F	Privacy-Related Law and Regulation: The State of the Law and Outstanding Issues	150
F.1	The Fourth Amendment, 150	
F.1.1	Basic Concepts, 150	
F.1.2	Machine-Aided Searches, 151	
F.1.3	Searches and Surveillance for National Security and Intelligence Purposes That Involve U.S. Persons Connected to a Foreign Power or That Are Conducted Wholly Outside the United States, 152	
F.1.4	The Miller-Smith Exclusion of Third-Party Records, 153	
F.2	The Electronic Communications Privacy Act, 154	
F.3	The Foreign Intelligence Surveillance Act, 155	
F.4	The Privacy Act, 156	
F.5	Executive Order 12333 (U.S. Intelligence Activities), 159	
F.6	The Adequacy of Today's Electronic Surveillance Law, 160	
F.7	Further Reflections from the Technology and Privacy Advisory Committee Report, 164	
G	The Jurisprudence of Privacy Law and the Need for Independent Oversight	166
G.1	Substantive Privacy Rules, 167	
G.1.1	Privacy Challenges Posed by Advanced Surveillance and Data Mining, 167	
G.1.2	Evolution of Regulation of New Technologies, 172	
G.1.3	New Surveillance Techniques That Raise Privacy Questions Unaddressed by Constitutional or Statutory Privacy Rules, 175	
G.1.4	New Approaches to Privacy Protection: Collection Limitation Versus Use Limitation, 175	
G.2	Procedural Privacy Rules and the Need for Oversight, 176	
G.2.1	Oversight Mechanisms of the U.S. Government, 177	
G.2.2	A Framework for Independent Oversight, 179	
G.2.3	Applying Independent Oversight for Government Agencies to Protect Privacy, 182	
G.2.4	Collateral Benefits of Oversight, 184	
H	Data Mining and Information Fusion	185
H.1	The Need for Automated Techniques for Data Analysis, 185	
H.2	Preparing the Data to Be Mined, 189	

- H.3 Subject-Based Data Mining as an Extension of Standard Investigative Techniques, 192
- H.4 Pattern-Based Data Mining Techniques as Illustrations of More Sophisticated Approaches, 193
- H.5 The Evaluation of Data Mining Techniques, 198
 - H.5.1 The Essential Difficulties of Evaluation, 199
 - H.5.2 Evaluation Considerations, 200
- H.6 Expert Judgment and Its Role in Data Mining, 205
- H.7 Issues Concerning the Data Available for Use with Data Mining and the Implications for Counterterrorism and Privacy, 207
- H.8 Data Mining Components in an Information-Based Counterterrorist System, 208
- H.9 Information Fusion, 209
- H.10 An Operational Note, 211
- H.11 Assessment of Data Mining for Counterterrorism, 213

- I Illustrative Government Data Mining Programs and Activity 218
 - I.1 Total/Terrorism Information Awareness (TIA), 219
 - I.2 Computer-Assisted Passenger Prescreening System II (CAPPS II) and Secure Flight, 219
 - I.3 Multistate Anti-Terrorism Information Exchange (MATRIX), 222
 - I.4 Able Danger, 224
 - I.5 Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE), 226
 - I.6 Automated Targeting System (ATS), 228
 - I.7 The Electronic Surveillance Program, 229
 - I.8 Novel Intelligence from Massive Data (NIMD) Program, 230
 - I.9 Enterprise Data Warehouse (EDW), 231
 - I.10 Law Enforcement Analytic Data System (NETLEADS), 232
 - I.11 ICE Pattern Analysis and Information Collection System (ICEPIC), 232
 - I.12 Intelligence and Information Fusion (I2F), 233
 - I.13 Fraud Detection and National Security Data System (FDNS-DS), 233
 - I.14 National Immigration Information Sharing Office (NIISO), 234
 - I.15 Financial Crimes Enforcement Network (FinCEN) and BSA Direct, 234
 - I.16 Department of Justice Programs Involving Pattern-Based Data Mining, 235

J	The Total/Terrorist Information Awareness Program	239
J.1	A Brief History, 239	
J.2	A Technical Perspective on TIA's Approach to Protecting Privacy, 243	
J.3	Assessment, 247	
K	Behavioral-Surveillance Techniques and Technologies	250
K.1	The Rationale for Behavioral Surveillance, 250	
K.2	Major Behavioral-Detection Methods, 251	
K.2.1	Facial Expression, 252	
K.2.2	Vocalization, 254	
K.2.3	Other Muscle Activity, 255	
K.2.4	Autonomic Nervous System, 255	
K.2.5	Central Nervous System, 257	
K.3	Assessing Behavioral-Surveillance Techniques, 258	
K.4	Behavioral and Data Mining Methods: Similarities and Differences, 259	
L	The Science and Technology of Privacy Protection	263
L.1	The Cybersecurity Dimension of Privacy, 263	
L.2	Privacy-Preserving Data Analysis, 266	
L.2.1	Basic Concepts, 266	
L.2.2	Some Simple Ideas That Do Not Work in Practice, 268	
L.2.3	Private Computation, 269	
L.2.4	The Need for Rigor, 270	
L.2.5	The Effect of Data Errors on Privacy, 273	
L.3	Enhancing Privacy Through Information-System Design, 275	
L.3.1	Data and Privacy, 275	
L.3.2	Information Systems and Privacy, 276	
L.4	Statistical Agency Data and Approaches, 277	
L.4.1	Confidentiality Protection and Public Data Release, 278	
L.4.2	Record Linkage and Public Use Files, 279	
M	Public Opinion Data on U.S. Attitudes Toward Government Counterterrorism Efforts	281
M.1	Introduction, 281	
M.2	Data and Methodology, 284	
M.3	Organization of This Appendix, 287	
M.4	General Privacy Attitudes, 288	
M.5	Government Surveillance, 291	

M.5.1	Trends in Attitudes Toward Surveillance Measures, 291	
M.5.2	Communications Monitoring, 294	
M.5.3	Monitoring of Financial Transactions, 300	
M.5.4	Video Surveillance, 301	
M.5.5	Travel Security, 302	
M.5.6	Biometric Identification Technologies, 303	
M.5.7	Government Use of Databases and Data Mining, 304	
M.5.8	Public Health Uses of Medical Information, 306	
M.6	The Balance Between Civil Liberties and Terrorism Investigation, 310	
M.6.1	Civil Liberties Versus Terrorism Prevention, 311	
M.6.2	Privacy Costs of Terrorism Investigation, 315	
M.6.3	Personal Willingness to Sacrifice Freedoms, 316	
M.6.4	Concerns About Uses of Expanded Powers, 317	
M.7	Conclusions, 319	
M.8	Annex, 322	
M.8.1	Details of Cited Surveys, 322	
M.8.2	Research of Organization/Sponsor Name Abbreviations, 322	
M.8.3	List of Surveys, 324	
M.8.4	References, 334	
N	Committee and Staff Biographical Information	335
O	Meeting Participants and Other Contributors	

