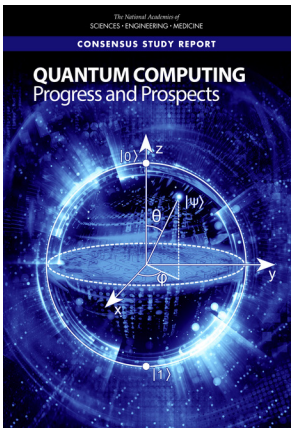
AUTHENTICITY, INTEGRITY, AND SECURITY IN A DIGITAL WORLD

FEBRUARY 19-20, 2019 | NATIONAL ACADEMY OF SCIENCES BUILDING | WASHINGTON, DC 20418

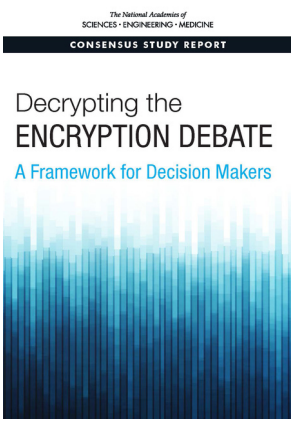
List of selected reports from the National Academies Press related to the meeting topic.



QUANTUM COMPUTING: PROGRESS AND PROSPECTS (CSTB 2018)

Quantum mechanics, the subfield of physics that describes the behavior of very small (quantum) particles, provides the basis for a new paradigm of computing. First proposed in the 1980s as a way to improve computational modeling of quantum systems, the field of quantum computing has recently garnered significant attention due to progress in building small-scale devices. However, significant technical advances will be required before a large-scale, practical quantum computer can be achieved. *Quantum Computing: Progress and Prospects* provides an introduction to the field, including the unique characteristics and constraints of the technology, and assesses the feasibility and implications of creating a functional quantum computer capable of addressing real-world problems. This report considers hardware and software requirements, quantum algorithms, drivers of advances in quantum computing and quantum devices, benchmarks associated with relevant use cases, the time and resources required, and how to assess the probability of success.

DECRYPTING THE ENCRYPTION DEBATE: A FRAMEWORK FOR DECISION MAKERS (CSTB 2018)



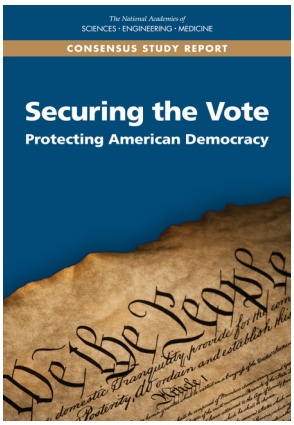
Encryption protects information stored on smartphones, laptops, and other devices - in some cases by default. Encrypted communications are provided by widely used computing devices and services - such as smartphones, laptops, and messaging applications - that are used by hundreds of millions of users. Individuals, organizations, and governments rely on encryption to counter threats from a wide range of actors, including unsophisticated and sophisticated criminals, foreign intelligence agencies, and repressive governments. Encryption on its own does not solve the challenge of providing effective security for data and systems, but it is an important tool. At the same time, encryption is relied on by criminals to avoid investigation and prosecution, including criminals who may unknowingly benefit from default settings as well as those who deliberately use encryption. Thus, encryption complicates law enforcement and intelligence investigations. When communications are encrypted "end-to-end," intercepted messages cannot be understood. When a smartphone is locked and encrypted, the contents cannot be read if the phone is seized by investigators. Decrypting the Encryption Debate reviews how encryption is used, including its applications to cybersecurity; its role in protecting privacy and civil liberties; the needs of law enforcement and the intelligence community for information; technical and policy options for accessing plaintext; and the international landscape. This book describes the context in which decisions about providing authorized government agencies access to the plaintext version of encrypted information would be made and identifies and characterizes possible mechanisms and alternative means of obtaining information.

DATA MATTERS: ETHICS, DATA, AND INTERNATIONAL RESEARCH COLLABORATION IN A CHANGING WORLD-PROCEEDINGS OF A WORKSHOP (GUIRR 2018)

In an increasingly interconnected world, perhaps it should come as no surprise that international collaboration in science and technology research is growing at a remarkable rate. As science and technology capabilities grow around the world, U.S.-based organizations are finding that international collaborations and partnerships provide unique opportunities to enhance research and training. International research agreements can serve many purposes, but data are always involved in these collaborations. The kinds of data in play within international research agreements varies widely and may range from financial and consumer data, to Earth and space data, to population behavior and health data, to specific project-generated data—this is just a narrow set of examples of research data but illustrates the breadth of possibilities. The uses of these data are various and require accounting for the effects of data access, use, and sharing on many different parties. Cultural, legal, policy, and technical concerns are also important determinants of what can be done in the realms of maintaining privacy, confidentiality, and security, and ethics is a lens through which the issues of data, data sharing, and research agreements can be viewed as well. A workshop held on March 14-16, 2018, in Washington, DC explored the changing opportunities and risks of data management and use across disciplinary domains. The third workshop in a series, participants gathered to examine advisory principles for consideration when developing international research agreements, in the pursuit of highlighting promising practices for sustaining and enabling international research collaborations at the highest ethical level possible. The intent of the workshop was to explore, through an ethical lens, the changing opportunities and risks associated with data management and use across disciplinary domains—all within the context of international research agreements. This publication summarizes the presentations and discussions from the workshop.

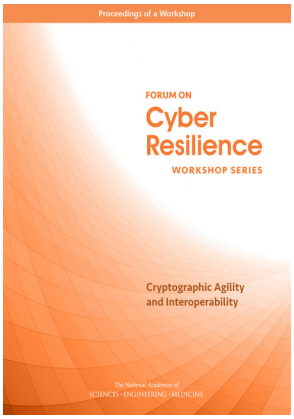
OPEN SCIENCE BY DESIGN: REALIZING A VISION FOR THE 21ST CENTURY RESEARCH (BRDI 2018)

Openness and sharing of information are fundamental to the progress of science and to the effective functioning of the research enterprise. The advent of scientific journals in the 17th century helped power the Scientific Revolution by allowing researchers to communicate across time and space, using the technologies of that era to generate reliable knowledge more quickly and efficiently. Harnessing today's stunning, ongoing advances in information technologies, the global research enterprise and its stakeholders are moving toward a new open science ecosystem. Open science aims to ensure the free availability and usability of scholarly publications, the data that result from scholarly research, and the methodologies, including code or algorithms, that were used to generate those data. *Open Science by Design* is aimed at overcoming barriers and moving toward open science as the default approach across the research enterprise. This report explores specific examples of open science and discusses a range of challenges, focusing on stakeholder perspectives. It is meant to provide guidance to the research enterprise and its stakeholders as they build strategies for achieving open science and take the next steps.



SECURING THE VOTE (CSTL/CSTB 2018)

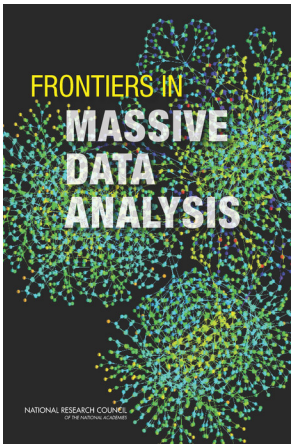
During the 2016 presidential election, America's election infrastructure was targeted by actors sponsored by the Russian government. *Securing the Vote: Protecting American Democracy* examines the challenges arising out of the 2016 federal election, assesses current technology and standards for voting, and recommends steps that the federal government, state and local governments, election administrators, and vendors of voting technology should take to improve the security of election infrastructure. In doing so, the report provides a vision of voting that is more secure, accessible, reliable, and verifiable.



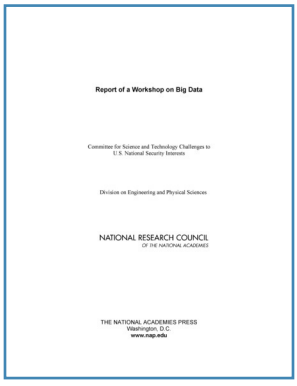
CRYPTOGRAPHIC AGILITY AND INTEROPERABILITY: PROCEEDINGS OF A WORKSHOP (CSTB 2017)

In May 2016, the National Academies of Sciences, Engineering, and Medicine hosted a workshop on Cryptographic Agility and Interoperability. Speakers at the workshop discussed the history and practice of cryptography, its current challenges, and its future possibilities. This publication summarizes the presentations and discussions from the workshop.

FRONTIERS IN MASSIVE DATA ANALYSIS (BMSA 2013)

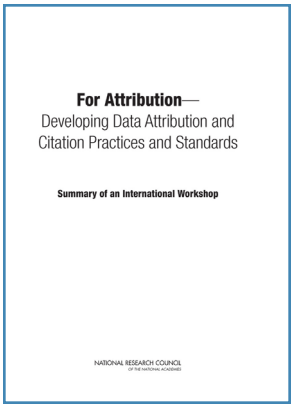


Data mining of massive data sets is transforming the way we think about crisis response, marketing, entertainment, cybersecurity and national intelligence. Collections of documents, images, videos, and networks are being thought of not merely as bit strings to be stored, indexed, and retrieved, but as potential sources of discovery and knowledge, requiring sophisticated analysis techniques that go far beyond classical indexing and keyword counting, aiming to find relational and semantic interpretations of the phenomena underlying the data. *Frontiers in Massive Data Analysis* examines the frontier of analyzing massive amounts of data, whether in a static database or streaming through a system. Data at that scale—terabytes and petabytes—is increasingly common in science (e.g., particle physics, remote sensing, genomics), Internet commerce, business analytics, national security, communications, and elsewhere. The tools that work to infer knowledge from data at smaller scales do not necessarily work, or work well, at such massive scale. New tools, skills, and approaches are necessary, and this report identifies many of them, plus promising research directions to explore. *Frontiers in Massive Data Analysis* discusses pitfalls in trying to infer knowledge from massive data, and it characterizes seven major classes of computation that are common in the analysis of massive data. Overall, this report illustrates the cross-disciplinary knowledge—from computer science, statistics, machine learning, and application disciplines—that must be brought to bear to make useful inferences from massive data.



BIG DATA: A WORKSHOP REPORT (DEPS 2012)

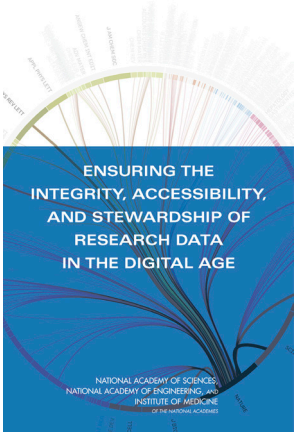
In 2012, the Defense Intelligence Agency (DIA) approached the National Research Council's TIGER standing committee and asked it to develop a list of workshop topics to explore the impact of emerging science and technology. From the list of topics given to DIA, three were chosen to be developed by the Committee for Science and Technology Challenges to U.S. National Security Interests. The first in a series of three workshops was held on April 23-24, 2012. This report summarizes that first workshop which explored the phenomenon known as big data. The objective for the first workshop is given in the statement of task, which explains that that workshop will review emerging capabilities in large computational data to include speed, data fusion, use, and commodification of data used in decision making. The workshop will also review the subsequent increase in vulnerabilities over the capabilities gained and the significance to national security. The committee devised an agenda that helped the committee, sponsors, and workshop attendees probe issues of national security related to so-called big data, as well as gain understanding of potential related vulnerabilities. The workshop was used to gather data that is described in this report, which presents views expressed by individual workshop participants. *Big Data: A Workshop Report* is the first in a series of three workshops, held in early 2012 to further the ongoing engagement among the National Research Council's (NRC's) Technology Insight-Gauge, Evaluate, and Review (TIGER) Standing Committee, the scientific and technical intelligence (S&TI) community, and the consumers of S&TI products.



FOR ATTRIBUTION: DEVELOPING DATA ATTRIBUTION AND CITATION PRACTICES AND STANDARDS–SUMMARY OF AN INTERNATIONAL WORKSHOP (BRDI 2012)

The growth of electronic publishing of literature has created new challenges, such as the need for mechanisms for citing online references in ways that can assure discoverability and retrieval for many years into the future. The growth in online datasets presents related, yet more complex challenges. It depends upon the ability to reliably identify, locate, access, interpret, and verify the version, integrity, and provenance of digital datasets. Data citation standards and good practices can form the basis for increased incentives, recognition, and rewards for scientific data activities that in many cases are currently lacking in many fields of research. The rapidly-expanding universe of online digital data holds the promise of allowing peer-examination and review of conclusions or analysis based on experimental or observational data, the integration of data into new forms of scholarly publishing, and the ability for subsequent users to make new and unforeseen uses and analyses of the same data-either in isolation, or in combination with, other datasets. The problem of citing online data is complicated by the lack of established practices for referring to portions or subsets of data. There are a number of initiatives in different organizations, countries, and disciplines already underway. An important set of technical and policy approaches have already been launched by the U.S. National Information Standards Organization (NISO) and other standards bodies regarding persistent identifiers and online linking. The purpose of the symposium was to examine a number of key issues related to data identification, attribution, citation, and linking to help coordinate activities in this area internationally, and to promote common practices and standards in the scientific community.

ENSURING THE INTERGRITY, ACCESSIBILITY, AND STEWARDSHIP OF RESEARCH DATA IN THE DIGITAL AGE (COSEPUP 2009)



As digital technologies are expanding the power and reach of research, they are also raising complex issues. These include complications in ensuring the validity of research data; standards that do not keep pace with the high rate of innovation; restrictions on data sharing that reduce the ability of researchers to verify results and build on previous research; and huge increases in the amount of data being generated, creating severe challenges in preserving that data for long-term use. *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age* examines the consequences of the changes affecting research data with respect to three issues - integrity, accessibility, and stewardship-and finds a need for a new approach to the design and the management of research projects. The report recommends that all researchers receive appropriate training in the management of research data, and calls on researchers to make all research data, methods, and other information underlying results publicly accessible in a timely manner. The book also sees the stewardship of research data as a critical long-term task for the research enterprise and its stakeholders. Individual researchers, research institutions, research sponsors, professional societies, and journals involved in scientific, engineering, and medical research will find this book an essential guide to the principles affecting research data in the digital age.

About the Government-University-Industry Research Roundtable (GUIRR)

GUIRR's mission is to convene senior-most representatives from government, universities, and industry to define and explore critical issues related to the national and global science and technology agenda that are of shared interest; to frame the next critical question stemming from current debate and analysis; and to incubate activities of on-going value to the stakeholders. The forum is designed to facilitate candid dialogue among participants, to foster self-implementing activities, and, where appropriate, to carry awareness of consequences to the wider public.

Government | University | Industry
RESEARCH ROUNDTABLE

For more information about GUIRR, visit our web site at www.nas.edu/guirr
500 Fifth Street, N.W. Washington, D.C. 20001 • guirr@nas.edu