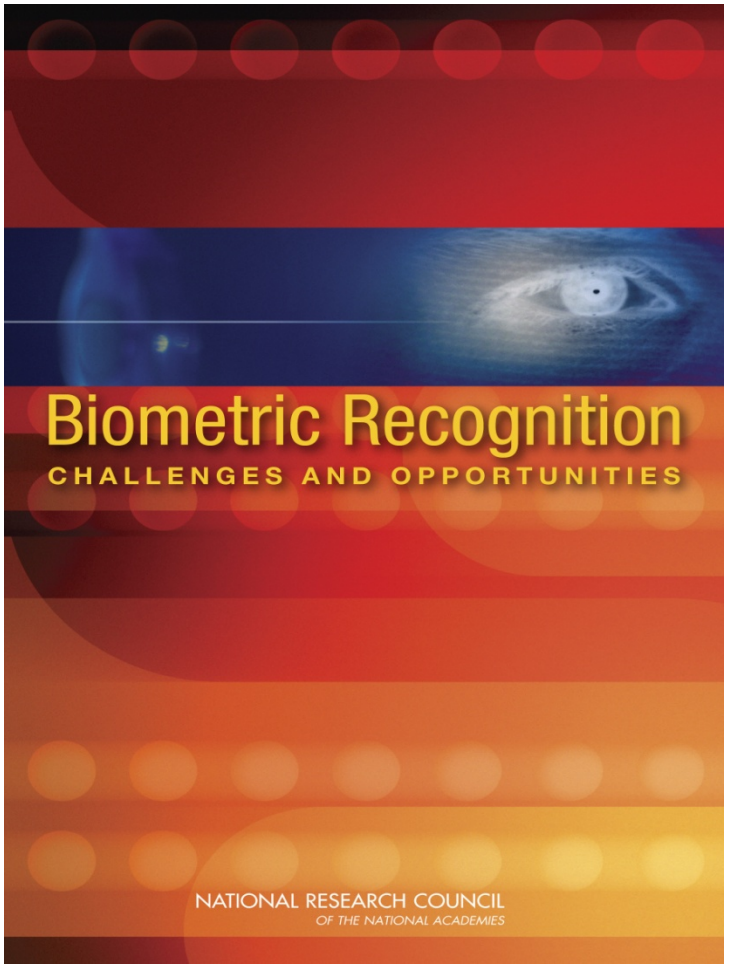
Biometric Recognition Challenges and Opportunities



17 December 2010 @ NIST

Joseph Pato, Committee Chair, Whither Biometrics

Computer Science and Telecommunications Board

National Research Council

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Study Background & Committee

Charge*

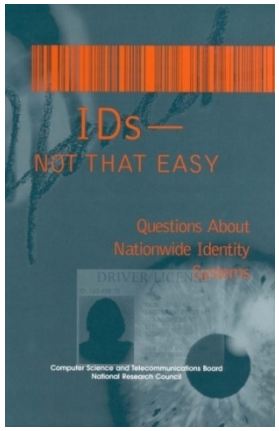
- Provide a comprehensive assessment of biometrics that examines current capabilities, future possibilities, and the role of government in their development.
- Explore the technical and policy challenges associated with the development, evaluation, and use of biometric technologies and systems that incorporate them.
- Examine associated research challenges and identify a multi- and inter-disciplinary research agenda to begin to meet them.
- Examine multiple stakeholders and points of view on multiple technologies, applications, and implementation issues

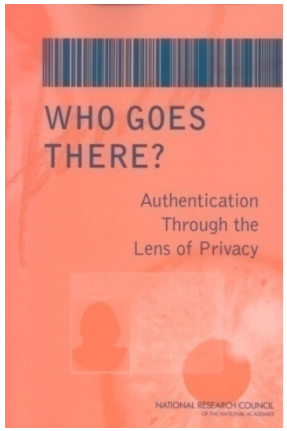
* Project sponsors: DHS, DARPA, CIA, NSF

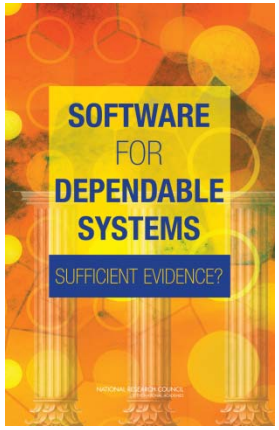
Background

- Grounded in previous CSTB work
  - *IDs—Not That Easy: Questions about Nationwide Identity Systems*
  - *Who Goes There? Authentication Through the Lens of Privacy*
  - *Software for Dependable Systems: Sufficient Evidence?*

  … 20+ years of CSTB work in cybersecurity and information systems…







Whither Biometrics Committee

- JOSEPH N. PATO, Hewlett-Packard Labs, *Chair*
- BOB BLAKLEY, Gartner
- JEANETTE BLOMBERG, IBM Almaden Research Center
- JOSEPH P. CAMPBELL, Massachusetts Institute of Technology, Lincoln Laboratory
- GEORGE T. DUNCAN, Carnegie Mellon University
- GEORGE R. FISHER, Prudential-Wachovia (retired)
- STEVEN P. GOLDBERG, Georgetown University Law Center*
- PETER T. HIGGINS, Higgins & Associates, International
- PETER B. IMREY, Cleveland Clinic and Case Western Reserve University
- ANIL K. JAIN, Michigan State University
- GORDON LEVIN, Disney
- LAWRENCE D. NADEL, Noblis
- JAMES L. WAYMAN, San Jose State University

- *LYNETTE I. MILLETT, CSTB Study Director*

Committee Approach

- Be neutral with respect to application and technology wherever possible
- Take a broad systems view
- Consider issues at scale
- Acknowledge privacy issues, but focus on engineering and cultural considerations
- Learn from information security community
- Informed by previous NRC work
- Informed by related forensic efforts
  - But focused on automated biometric recognition, not forensic applications

Principal Conclusions

Recognize Inherently Probabilistic Nature of Systems

- Biometric systems contend with uncertainty at nearly every stage of system operation
- Some fraction of interactions will produce incorrect or indeterminate results
- Error and exception processing critical to system success
- Qualitatively different from traditional discrete information system bugs or glitches

Rigorous Systems Approach Necessary for Success

- Biometric components are embedded in larger socio-technical systems
  - Environment, operations, adjudication mechanisms, policy choices, requirements, data management, user interfaces, maintenance, . . .
- Effectiveness depends as much on social context as it does on technical & engineering factors
- Biometric systems should be designed and evaluated relative to their specific intended purposes and contexts rather than generically or in isolation.

Scientific Basis

Needs Strengthening

- The distinctiveness of biometric characteristics is not well understood at global population scales.
  - Develop a science of human individual distinctiveness
- The effects of human behavior and interaction on large-scale biometric system performance (and vice versa) is not well explored.
- Numerous opportunities for research – see Chapter 5

A Well-Designed Biometric System

- Takes into account that recognition is based on similarity and **probabilistic** not absolute matching; presumptions and burdens of proof are correspondingly **conservative**
- Anticipates a **lifecycle** corresponding to changes in presentation distributions, stability of traits, and technology
- Assesses the **reliability** of information associated with a recognition **independently** of the **confidence** in correct recognition
- Provides **exception handling** as robust as the primary biometrics process and handles errors **gracefully**, without violating dignity, privacy or due process rights
- Publicly states explicit **security**, **privacy** and policy goals
- Recognizes that biometric traits are inherently **not secret** and will minimize risks to privacy and of misrecognition arising from this fact

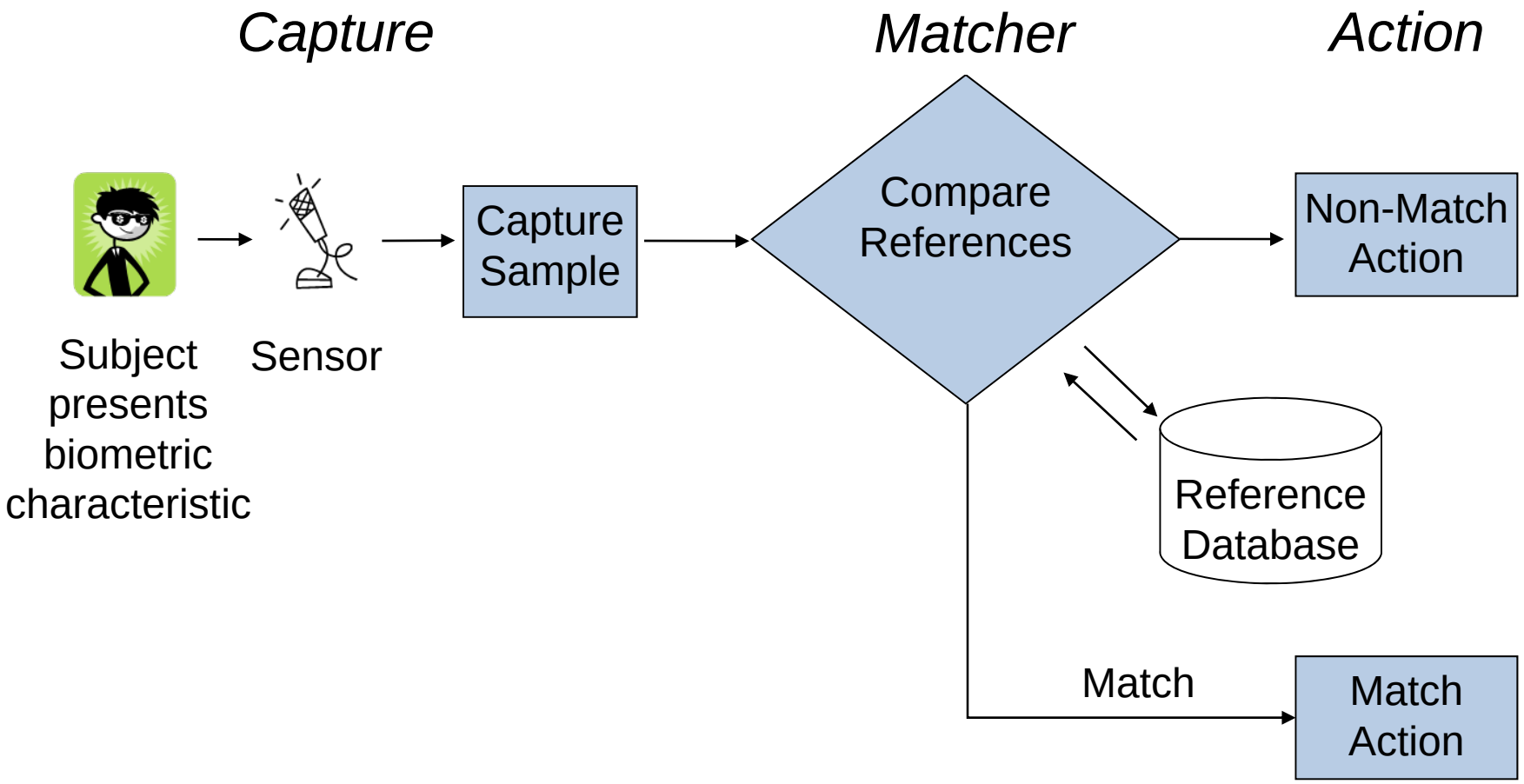
Fundamental Concepts

Definition

- Biometrics is the automated recognition of individuals based on their behavioral and biological characteristics.*

*ISO/IEC JTC1/SC37 Standing Document 2:
Harmonized Biometric Vocabulary

Operational Overview



Fundamental Dogma

- An individual is more similar to him- or herself over time than to anyone else.

Grounded Fundamental Dogma

- An individual is more likely similar to him- or herself over time than to anyone else likely to be encountered.

… to some degree of confidence

Stability and Distinctiveness

- Some traits appear stable over time, while others can change significantly even over short periods of time.

- Underlying distinctiveness and stability of biometric traits are not well understood at large scales.

Numerous Sources of Uncertainty

- Physical change in person over time
- Interface & environment change
- Motivation & social factors
- Processed features
  - Noisy data, sensors, algorithms
- Data integrity

Biometric systems perform decision-making under uncertainty

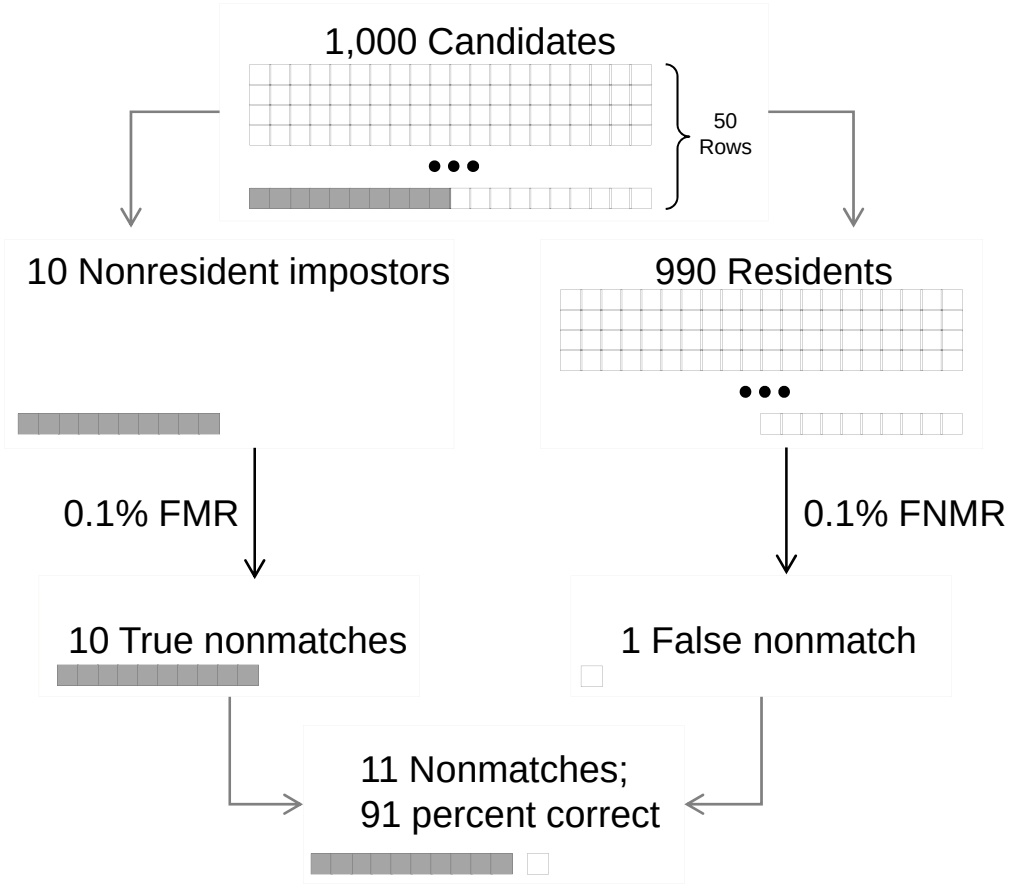
Error Rates Can be Misleading

- Error rates capture component and system performance in a narrow sense
  - FAR, FRR
  - FMR, FNMR
- Confidence in system behavior depends not just on error rates, but also relies on knowing prior probability of a false claim.
  - FMR and FNMR do not predict the impostor base rate

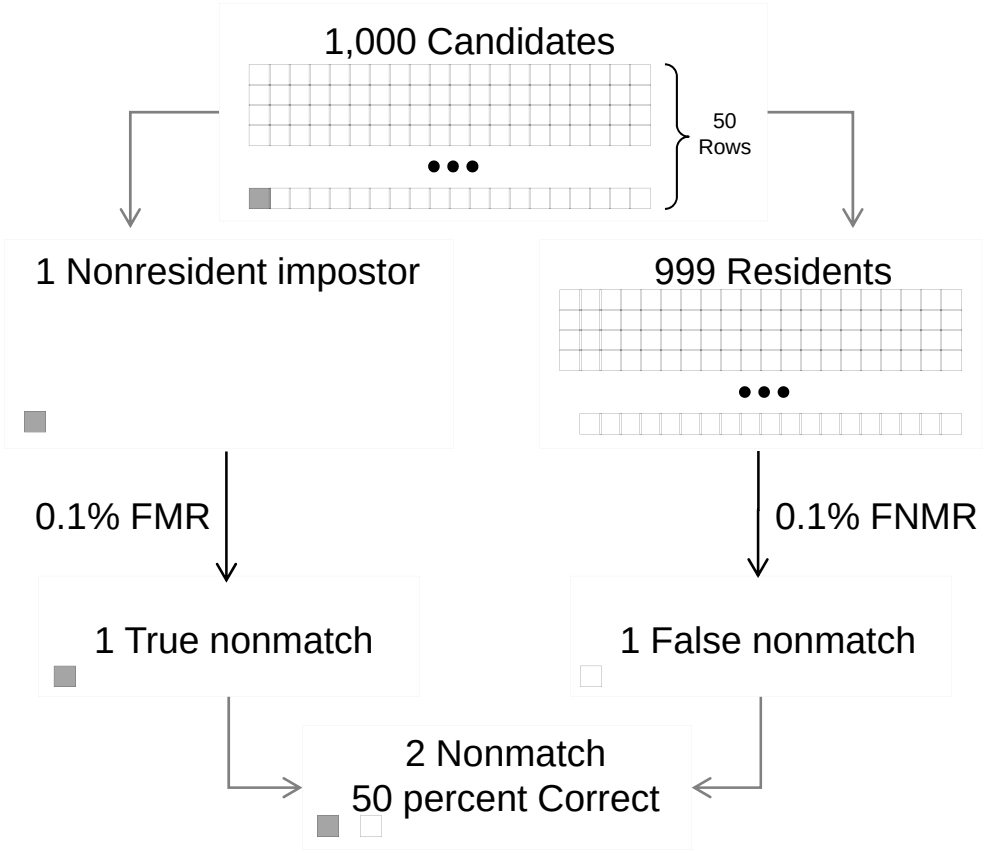
Confidence in Recognition Decisions

- College dorm access control scenario
  - 0.1% FMR
  - 0.1% FNMR
- Imposters attempt to gain access by posing as a legitimate resident
- Confidence in rejections depends on imposter base rate and declines with a reduction in the number of imposters

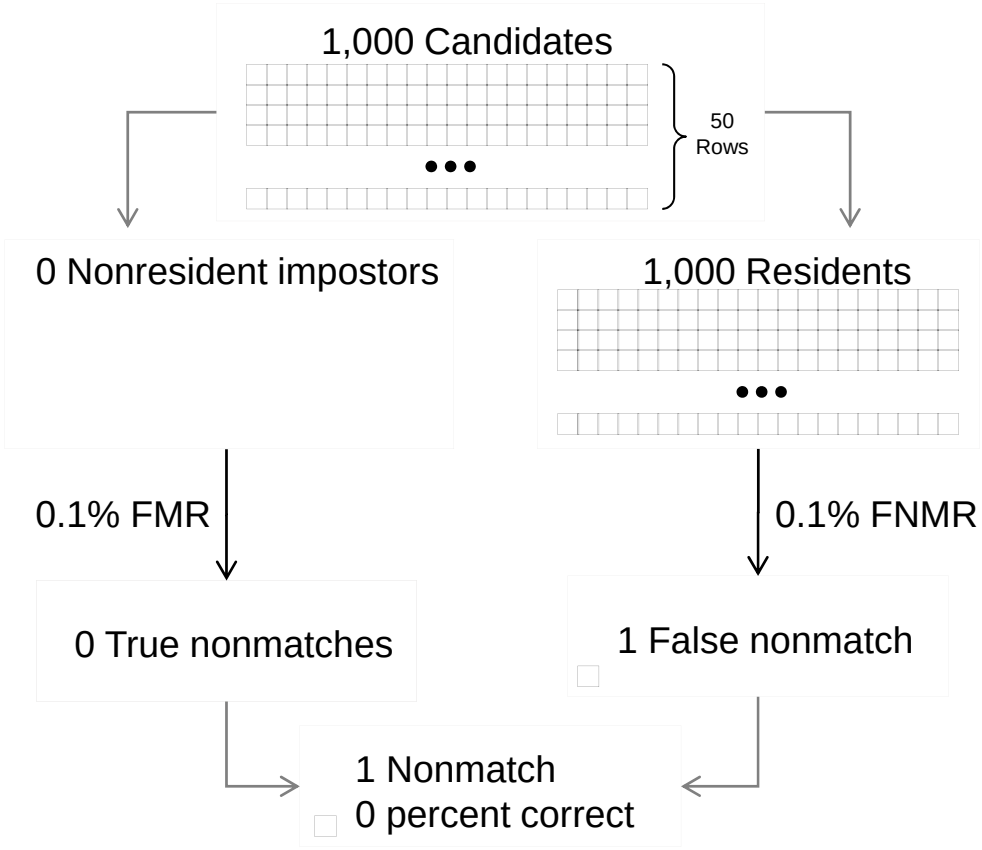
91% Confidence With 10 Imposters per 1000



50% Confidence With 1 Imposter per 1000



0% Confidence With 0 Imposters per 1000



Similar Issues for Watchlists

- What confidence to have in a match when users are presenting to be checked against a watchlist
  - Most will not match
  - Need prior probability of expected matches to know what confidence to have in match
- Increasing size of watchlist cannot be expected to improve all aspects of system performance
- As watch list size increases we should be less confident that a match is correct
  - with implications for the push towards increased interoperability

Engineering Biometric Systems in Context

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

System Life-Cycle Considerations

- Issues for all systems
  - Training
  - Commissioning
  - Component fault replacement
  - Decommissioning
- Issues for large scale, long-lived systems
  - Technology refresh – hardware and software
  - Data quality, currency, & integrity
  - Changes in target population
  - Evolving threat models
  - Policy flexibility

Immediate Operational Contexts

- Operational requirements and choices affect system design decisions and system effectiveness
  - User Context
  - Application Context
  - Technology Context
  - Performance Context

User Context

Parameter	<Less Challenging		More Challenging>
Data Subject			
Awareness	Very		Not very
Motivation	Cooperative	Indifferent	Uncooperative
Training	Well trained		Not very
Habituation	Very		Not very
Party Benefiting	Both	User / Consumer	Owner / Agency

= Thumb access to PDA
= Access to Elbonian Fitness Center
= Elbonian border watch list

Application Context

Parameter	<Less Challenging		More Challenging>
Supervision	Direct	Remote	None
Claim type	Positive		Negative
Recognition type	Verification	One to few	Identification

= Thumb access to PDA
= Access to Elbonian Fitness Center
= Elbonian border watch list

Technology Context

Parameter	<Less Challenging		More Challenging>
Environment	Controlled		Variable
Engagement	Active	Passive with cooperation	Passive
Sample Capture	Overt		Covert
Dataset	Proprietary		Open Standards

= Thumb access to PDA
= Access to Elbonian Fitness Center
= Elbonian border watch list

Performance Context

Parameter	<Less Challenging		More Challenging>
Throughput	Low	Medium	High
Sensitivity to error rate	Low	Medium	High

= Thumb access to PDA
= Access to Elbonian Fitness Center
= Elbonian border watch list

These Contexts Matter

- Stating that a system is a “biometric system” or uses “biometrics” says little
  - … about what the system is for or
  - … how difficult it is to successfully implement or
  - … its likelihood of successful deployment

Research needed: Taxonomy of systems & design implications

Biometric Systems and Trustworthiness

Typical Security Goals in Biometric System

- Determine that an observed trait belongs to a living human who is present and acting intentionally
- Accurate comparison of the observed trait to reference data maintained in the system (within desired confidence levels)
- Backend data security, integrity, etc.

Trustworthiness

- Model threats to understand potential points of attack and estimate probability of attack
- Assume biometric traits are public information
- Manage the trustworthiness of the entire recognition process

Multiple Potential Targets

- Resource being protected by biometric system
  - will have particular vulnerabilities and likely threats separate from mode of protection
- Biometric system itself
  - will have particular vulnerabilities and likely threats separate from what it's protecting
- Both drive analysis and decisions about how best to provide overall security

Social, Cultural, and Legal Implications

Interaction Between Systems & Individuals

- Motivating participation by individuals
  - Clear benefits for participants
  - Limits on system uses
- Facilitating individual participation
  - Planning for diversity
  - Ease of use (beyond sensor interface)
  - Graceful exception handling and accommodations for failures

Societal Impact

- Community acceptance is influenced by sense of proportionality related to perceived or actual side effects
  - Universality and the potential for disenfranchisement
  - Potential for record linkage and the loss of anonymity
  - Covert surveillance and the potential for abuse of power
  - Constraints on individuality and identity

Legal Issues

- Remediation
  - Identity fraud by falsifying, altering, or concealing biometric traits
  - Inappropriate denial of due process rights resulting from improper recognition
  - Responsibility of system operators to minimize misuse of biometric samples
- Reliability
  - Effects of depiction of technology in popular culture
  - Role of expert testimony and Frye & Daubert standards
- Privacy
  - Legal issues overlap (but do not encompass) cultural issues; both matter

Information Sharing

- Information sharing is attractive
  - Administrative efficiencies & business purposes
  - Research uses
- Information sharing requires caution
  - Biometric data are personally identifiable information
  - Biometric data can serve to correlate disparate databases (to a degree of confidence)
  - Access to data may allow discovery of *doppelgangers* or discovery of enrolled users

Data Policies

- Guidance from 2002 *IDs – Not That Easy* remains relevant
  - What is the *purpose?*
  - What is the *scope of the population?*
  - What is the *scope of the data?*
  - Who would be the *users?*
  - What *types of use?*
  - Is participation *mandatory*?
  - What legal structures *protect integrity, privacy and due process?*

Research Opportunities & Public Policy Considerations

Technical and Engineering Research Opportunities

- Distinctiveness and stability of underlying phenomena
  - both absolutely and under common conditions of capture
- Modality-related
  - Sensors, segmentation, invariant representation, robustness
- Human factors and affordance
- Testing and evaluation
  - Test data
  - Usability Testing
- Information & system security
- Scale – numbers, geography, time, and so on

Social Science Research Opportunities

- Individuals
  - Performance & effectiveness
  - Behaviors & affordance
- Society
  - Social impacts, direct and indirect
  - Community acceptance
- Develop data that predict how well a system will perform
  - Experimental studies
  - Field studies, ethnography

Public Policy Considerations

- Feasibility of large-scale deployment
  - Does existing technology adequately satisfy the problem?
  - Do deployment plans integrate adequate risk management?
  - Does national origin of technology create undue risk?
  - Is there an adequate biometrics workforce?
- Are social impact assessments helpful?
- Do deployments represent serious potential for identity theft?
- Can authoritarian regimes exploit human recognition; how could such a risk be mitigated?
- What considerations limit or facilitate research use of biometric data?

Lessons from Other Large-Scale Systems

Lessons from Other Large-Scale Systems

- Many factors contributing to the success of large-scale systems are not unique to biometric systems

- Opportunity to cross-fertilize biometrics community with expertise from other domains such as:
  - information security
  - medical diagnostics
  - manufacturing
  - systems engineering

Biometrics Share Characteristics with Other Large-Scale Systems

- Key success factors for large-scale systems:
  - Good project management
  - Alignment of capabilities with underlying need and operational environment
  - Thorough threat and risk analysis

- Common contributors to failures:
  - Inappropriate technology choices
  - Lack of sensitivity to user perceptions and needs
  - Poor understanding of population issues
  - Lack of a viable business case

Software, Systems, and Demonstrating Dependability

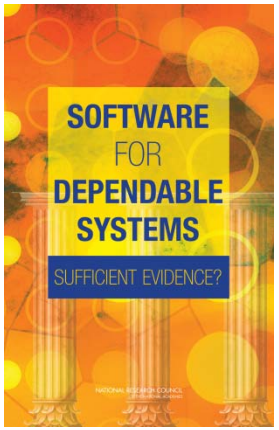
- Biometric systems are software systems (among other things), and thus
  - general lessons in developing robust software systems apply.

- How to demonstrate reliability of mission-critical systems?

- How to evaluate whether a system is dependable (meeting its stated goals)?

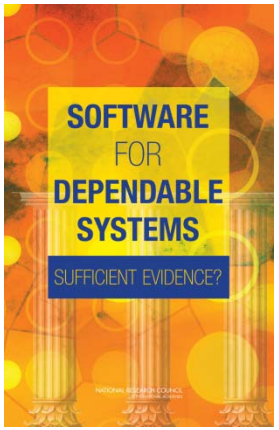
What We Know About Software Systems

- Extent of failures to date
  - software has already resulted in critical system failures
  - death, injury and major economic loss
- Roots of failure
  - bugs in code account only for 3% of failures blamed on software
  - most failures blamed on interactions with operators, environment
  - often poor understanding of requirements
- Development strategies
  - building dependable software is difficult and costly
  - quality is highly variable
  - certification regimes and standards have mixed record



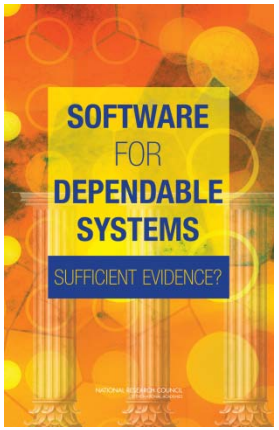
What We Don't Know

- Incomplete and unreliable data about
  - extent and frequency of software failures
  - efficacy of development approaches
  - benefits of certification schemes
- Consequences
  - mandating particular process does not guarantee dependability
  - avoid being too prescriptive about particular tools or techniques
  - put in place mechanisms for collecting industry-wide evidence
  - make evidence focus of dependable system development



Three E's for Dependable Software Systems

- Be explicit
  - properties established
  - assumptions about domain and usage
  - level of dependability
- Develop and present evidence
  - **dependability case** that properties hold
  - scientifically justifiable claims
  - open to audit by a third-party
- Exploit expertise
  - approach is technology-independent
  - demand for evidence stretches today's best practices
  - deviate from best practice only with good reason



Distinctive Nature of Biometric Systems and Problem Space

- *Similar* to… medical diagnostics, digital authentication systems, manufacturing production lines, mission-critical tightly-coupled software systems, information kiosks…

  … but *not* strictly analogous to any of the above

- *And* poised to address major public policy challenges…

  …while interacting with large portions of the population…

  …*about* complex personal notions of identity, privacy, and privilege.

Acknowledge and Embrace

- Acknowledge the importance of what is being asked of this technology

  … especially when used at scale for broad security or public policy goals

- Embrace the challenges inherent

  … in multidisciplinary research – deep science, hard engineering, social aspects, and

  … in deploying reliable, broadly-trusted systems.

Questions?

For more information…

Joe Pato, Committee Chair, joe.pato@hp.com
Lynette Millett, Study Director, lmillett@nas.edu

www.cstb.org

Reports available at: www.nap.edu

Extra Slides

CSTB on Cybersecurity and Trustworthiness

- **Critical Code: Software Producibility for Defense** (2010) assesses the growing importance of software for national security and examines how the U.S. Department of Defense can most effectively meet its future software needs.
- **Proceedings of a Workshop on Deterring Cyberattacks: Informing Strategies and Developing Options** (2010) examines governmental, economical, technical, legal, and psychological challenges involved in deterring cyber attacks.
- **Biometric Recognition: Challenges and Opportunities** (2010) presents a broad and comprehensive assessment of biometric recognition systems -- articulating design and operational considerations as well as outlining a research agenda to bolster the scientific and engineering underpinnings of these systems.
- **Toward Better Usability, Security, and Privacy of Information Technology** (2010) identifies research opportunities and ways to embed usability considerations in design and development related to security and privacy, and vice versa.
- **Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities** (2009) concludes that although cyberattack capabilities are an important asset for the United States, the current policy and legal framework for their use is ill-formed, undeveloped, and highly uncertain and that U.S. policy should be informed by an open and public national debate on technological, policy, legal, and ethical issues they pose.
- **Toward a Safer and More Secure Cyberspace** (2007) explores the nature of online threats, considers some of the reasons why past research for improving cybersecurity has had less impact than anticipated, and offers a strategy for future research aimed at countering cyber attacks.
- **Software for Dependable Systems: Sufficient Evidence?** (2007) discusses how the growing use and complexity of software necessitates a different approach to dependability and recommends an evidence-based approach to achieving greater dependability and confidence.
- **Who Goes There? Authentication Through the Lens of Privacy** (2003) describes and examines issues, concepts, and techniques for authentication from the perspective of how they implicate privacy—and how adverse impacts on privacy might be contained.
- **Critical Information Infrastructure Protection and the Law: An Overview of Key Issues** (2003) discusses antitrust, FOIA, and liability as factors in protecting critical information infrastructure, given technical and economic conditions.
- **IDs -- Not That Easy: Questions About Nationwide Identity Systems** (2002) outlines challenging policy, process, and technological issues presented by nationwide identity systems.
- **Cybersecurity Today and Tomorrow: Pay Now or Pay Later** (2002) Recaps highlights from past CSTB security reports with a focus on issue identification and practical guidance.
- **Trust in Cyberspace** (1999) provides an assessment of the state of the art procedures for building trustworthy networked information systems; proposes directions for research in computer and network security, software technology, and system architecture; and assesses current technical and market trends in order to better inform public policy as to where progress is likely and where incentives could help.
- **Realizing the Potential of C4I: Fundamental Challenges** (1999) addresses the intersecting arenas of security, interoperability, and DOD culture and processes as they relate to challenges in command, control, communications, computers, and intelligence.
- **Cryptography's Role in Securing the Information Society** (1996) describes the growing importance of encryption, relating a government interests to interests in the spread and control of encryption, and recommends policy changes.
- **Computers at Risk: Safe Computing in the Information Age** (1991), an enduring primer for information security, explains key concepts and terms, outlines the technology and procedures that give rise to and can alleviate security problems, relates security to complementary concerns such as privacy and safety, and describes the private and public sector institutional contexts.

Medical Diagnostics

- Individual components in general usage are rarely as sensitive and specific as under testing
- Confirming a test by repetition is less valuable than confirming it by a different test
- Limitations in individual components can vitiate the effectiveness of others.
- Effectiveness of a system is highly population-specific

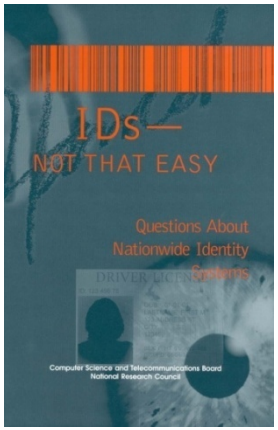
Manufacturing

- System objectives must be clear and competing priorities resolved
- The full range of the operational environment should be anticipated
- Ongoing operational testing and blind challenges of operational systems
- Biometric systems may be considered a production line
  - Individuals presenting for recognition are input
  - High quality decisions are output

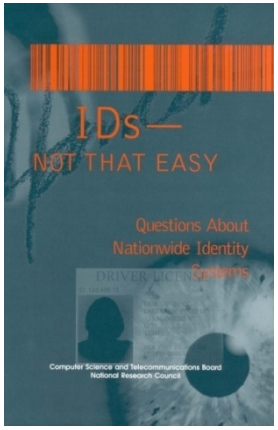
Dependable Software Systems
Sufficient Evidence?

Large-Scale Identity Systems
It's Not That Easy

- What would be **the purpose** of the system?
  - Assumptions differ about ends
- **Who** would be issued an ID?
- **What data** would be collected?
- Who would **use** the system?
- What **types of use** would be allowed?
- Is participation **voluntary or mandatory**?
- What **legal structures** would be needed?



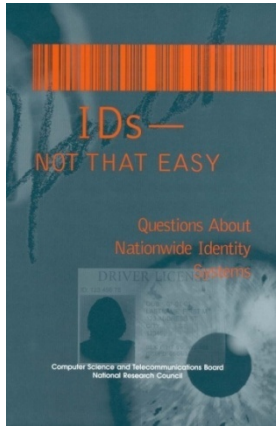
What does Identity Provide?



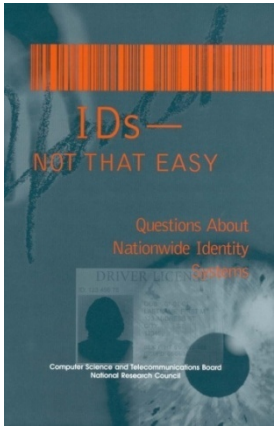
- What is identity?
- Individuals often have multiple identities
- What identity information is relevant to the purpose of the system?
- When are group identities (e.g., “older than 21”) appropriate? How is identity initially established?
- What is the meaning of the ID?
- Where does the identity information reside?
- What kinds of access to and analysis of that information is allowed or disallowed?

Who Uses the System? For What?

- Under what circumstances may someone request an ID? Access data within the system?
- Would the private sector be allowed to use the system?
- Who will manage, oversee, and maintain the system?
- Will public-private partnerships be required?
- What are the implications of broad vs. narrow sets of users?
- What questions can be asked of the system?
- Will data mining be allowed? By whom?
- Will real-time correlation of system transactions be needed?
- How will information be linked to other systems?
- What are the privacy implications of allowing broad queries and/or linkages?



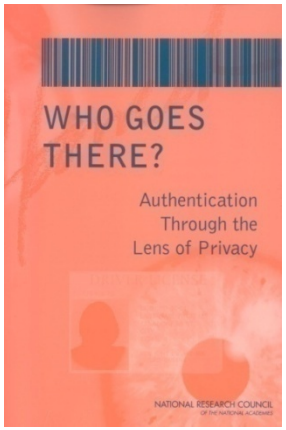
Data and Backend Systems (Not just “IDs”)



- Would consolidation of other databases be necessary?
- Centralization creates a single target for adversaries
- What are the risks from denial of service attacks
- How are potential privacy invasions mitigated if correlative capabilities allowed?
- What are the tracking, surveillance, and prediction requirements?
- Is high availability necessary? What are the implications?
- What backups and redundancy would be necessary?
- Differing levels of access and query capabilities for different users?
- Procedurally, how would maintenance and administration work?

Major Findings from Authentication study

- Context, scope, implementation matter greatly
- Local contexts/uses usually more sensitive to privacy considerations
- Secondary uses are particularly problematic



- Toolkit for thinking through design is provided
- Checklist for evaluating/designing authentication systems is presented

When Designing a Privacy-Sensitive Authentication System

- Authenticate only for necessary, well-defined purposes
- Minimize the scope of data collected
- Minimize the retention interval of data collected
- Articulate what entities will have access to the collected data
- Articulate what kinds of access to and use of the data will be allowed
- Minimize the intrusiveness of the process
- Overtly involve the individual to be authenticated in the process
- Minimize the intimacy of the data collected
- Ensure that the use of the system is audited and that the audit record is protected against modification and destruction
- Provide for individuals to check on and correct information held and used for authentication