NISTIR 8156

Improving Biometric and Forensic Technology: The Future of Research Datasets

Symposium Report

Melissa Taylor
Shannan Williams
George Kiebuzinski

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8156

National Institute of Standards and Technology
U.S. Department of Commerce

AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

NISTIR 8156

Improving Biometric and Forensic Technology: The Future of Research Datasets

Symposium Report

Melissa Taylor
Shannan Williams
Forensic Science Research Program
Special Programs Office
National Institute of Standards and Technology

George Kiebuzinski
Noblis, Inc.

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8156

March 2017

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Kent Rochford, Acting NIST Director and Under Secretary of Commerce for Standards and Technology

Table of Contents

1 Executive Summary 6
1.1 Background 6
1.2 Summary of Workshop Presentations and Discussions 6
1.3 Findings 7
2 Introduction 10
2.1 Sponsorship 10
2.2 Acknowledgements 11
2.3 Disclaimer 11
3 Workshop Overview 12
4 The Role of Datasets in Improving Biometric and Forensic Technology 12
4.1 Initiation of NIST/NIJ Collaborative Efforts 12
4.2 Challenges in the Collection and Use of Biometric and Forensic Datasets 13
4.3 Discussion 15
4.3.1 Research Challenges 15
4.3.2 Data Collection Issues 16
5 Privacy and Legal Issues in the Collection, Distribution, and Use of Biometric Datasets 17
5.1 Summary of Presentations 17
5.2 Discussion 19
5.2.1 Specific Recommendations 19
5.2.2 Research Challenges 20
5.2.3 Implementation Issues 20
6 NIST Biometric and Forensic Database Catalog 21
7 Modality Panel: Fingerprint/Palmprint 23
7.1 Summary of Presentations 23
7.2 Discussion 26
7.2.1 Recommendations 27
7.2.2 Research Challenges 27
7.2.3 Data Collection/Implementation Issues 27
8 Modality Panel: Iris 28
8.1 Summary of Presentations 28
8.2 Discussion 30
8.2.1 Recommendations 30
8.2.2 Research Challenges 30
8.2.3 Data Collection and Implementation Issues 31
9 Modality Panel: Voice 32
9.1 Summary of Presentations 32
9.2 Discussion 33
9.2.1 Recommendations 33
9.2.2 Research Challenges 33
9.2.3 Data Collection and Implementation Issues 33
10 Modality Panel: Handwriting 34
10.1 Summary of Presentations 34
10.2 Discussion 35
10.2.1 Recommendations 35
10.2.2 Research Challenges 35
10.2.3 Data Collection Issues 36
11 Face and Person at a Distance 37
11.1 Summary of Presentations 37
11.2 Discussion 41
11.2.1 Recommendations 41

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

11.2.2 Research Challenges 41
11.2.3 Data Collection and Implementation Issues 41
12 Statistical Significance in Biometric and Forensic Datasets 43
12.1 Summary of Presentations 43
12.2 Discussion 44
12.2.1 Recommendations 44
12.2.2 Research Challenges 44
13 Findings 45
14 References 48
14.1 Presentations at the workshop (with hyperlinks): 48
14.2 Related documents: 48
14.3 Speaker Biographies 48
14.4 Participant List 57

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Figures

Figure 0-1: Event logo. 6
Figure 4-1 : NIJ/NIST collaborative effort to address biometric and forensic dataset availability. 13
Figure 4-2: Examples of why to be a cautious user and developer of biometric datasets. 13
Figure 4-3: Basic questions about the use of biometric datasets. 14
Figure 4-4: NIST software testing using sequestered data. 15
Figure 5-1: Lab vs. operational data conditions. 17
Figure 5-2: Highlights from the Privacy Act of 1974. 18
Figure 6-1: Goals of the biometric and forensic dataset catalog. 21
Figure 6-2: Breakdown of catalogued datasets by modality. 21
Figure 6-3: Biometric and Forensic Research Catalog Homepage 22
Figure 7-1: Public domain fingerprint and palmprint databases. 23
Figure 7-2: Timeline of biometrics evaluations at NIST. 24
Figure 7-3: Current landscape of origins of biometric data. 26
Figure 8-1: Eye disorder iris data examples. 28
Figure 8-2: Microscopic view of pupil dilation. 29
Figure 8-3: Manual iris matching. 29
Figure 8-4: Iris image variations in color. 30
Figure 9-1: Speaker recognition error rates. 32
Figure 10-1: Biometric individuality. 34
Figure 10-2: Handwriting image data processing. 35
Figure 11-1: Heterogeneous face recognition. 38
Figure 11-2: Dataset collection overview. 39
Figure 11-3: Non-cooperative watchlist composition data. 40
Figure 12-1: Summary of presentation on purposes of databases. 44

Tables

Table 7-1: Difference between Public Datasets and Sequestered Datasets 25

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

1 Executive Summary

On January 26th and 27th, 2015, the National Institute of Standards and Technology (NIST), in collaboration with the National Institute of Justice (NIJ), held a workshop entitled **Improving Biometric and Forensic Technology: The Future of Research Datasets** at NIST's campus in Gaithersburg, Maryland. This document summarizes the discussions and conclusions of the workshop, held in response to a request by NIJ, including feedback obtained from workshop participants during facilitated panel discussions. The event webcast and presentation materials are available at: http://www.nist.gov/forensics/biometric-forensic-technology-webcast.cfm.

Figure 0-1: Event logo.

1.1 Background

A biometric refers to the measurement of physiological characteristics, including – but not limited to – fingerprints, iris patterns, and facial features. Biometric data is essential for the development and enhancement of automated systems that establish or verify a person's identity. Since the 1960s, NIST has promoted and participated in projects developing standards, databases, and evaluations of biometric technologies. An example is Special Database 27 (SD27), which has been widely used in the development of automated latent fingerprint matching systems.[1] Over the years, NIST's involvement has expanded from fingerprints to include such modalities as face, iris, voice, DNA, and multiple combinations of modalities.

Although NIST and other organizations in government, academia, and private industry have created a number of large and varied databases, research has been hindered by the high cost of data collection, privacy protection requirements, and the absence of standardized public datasets with a common set of metrics.

In response to the widely acknowledged lack of publicly available data for research, the National Institute of Justice sponsored and asked NIST to conduct the following activities:

1. **Create a comprehensive catalog of publicly available biometric and forensic datasets for the following modalities: fingerprints and palmprints, iris, face, person at a distance, voice, and handwriting.**
2. **Host a symposium to discuss the adequacy of existing publicly available datasets and the future needs in these areas.**

While biometric data is useful in a variety of applications (i.e. security and access control) this workshop focused on the role of biometric data in investigations involving forensic science. The modalities covered in this symposium and report include fingerprints, palmprints, face, iris, voice, and handwriting. The material presented at the symposium is summarized and the key findings are described in this report.

1.2 Summary of Workshop Presentations and Discussions

The workshop was attended by a wide range of researchers from academia and private industry. They described their views on the state of the art of data collection, the collection processes they use or have developed, and the need for designing a more effective collection process. The presentations examined

[11] *No longer available for distribution by NIST.*

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

the problems encountered in conducting research while at the same time safeguarding privacy. The presentations were organized into a number of sessions and panels:

- A plenary session devoted to the role of datasets in improving biometric and forensic technology
- A panel on privacy and legal issues in the collection, distribution, and use of biometric and forensic datasets
- A description of NIST's new catalog of publicly available biometric databases
- Data collection issues and challenges for specific modalities:
- A panel on statistical significance in biometric and forensic data

1.3 Findings

The workshop revealed an agreement on the need for a number of key actions to be potentially taken by research stakeholders, academia, law enforcement, or federal agencies such as NIST and NIJ:

1. Establish an Institutional Review Board (IRB) Toolkit Working Group

The legal and ethical issues related to protecting the privacy of individuals make it difficult to obtain approval to collect biometric and forensic data for use in research. Frequently, processes for protecting privacy are unclear and complex, making approval even more of a challenge since privacy provisions are often interpreted differently by government agencies and private organizations.

Guidance on and streamlining of the IRB process is needed to understand and consistently implement the policies defined by the U.S. Department of Health and Human Services' (HHS) Office for Human Research Protections (OHRP) and the Federal Policy for the Protection of Human Subjects (the "Common Rule").

2. Enhance and Maintain the Biometric and Forensic Research Database Catalog

NIST has completed the initial Biometric and Forensic Research Database Catalog. Taxonomy for the catalog should be published and a plan to maintain and update it should be developed.

3. Identify and Develop Datasets for Rapid Public Release

Researchers need more and larger publicly available datasets for algorithm development. The most frequently mentioned public dataset was the NIST Special Database 27 latent fingerprint dataset, which was created about 1996 and publicly released in 1999. SD27 has been incredibly useful, but is relatively small and has been overused, thereby diminishing its usefulness. That dataset needs to be augmented with additional latent prints and exemplars collected using up-to-date technology. Currently, SD27 is no longer distributed by NIST.

Collection of new datasets is a time consuming and expensive process and requires a two-pronged approach:

1. Identify and develop datasets for rapid public release. This effort focuses on identifying the types of data that could have expedited approvals, dataset collection processes that can be expedited, or datasets that are readily available. This effort would include:
    a. Coordination with other organizations in making data public, including those organizations that have legal approval to share data with NIST.
    b. Encouraging universities working with local and state law enforcement agencies to make datasets public.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

2. Determine which areas of research have the greatest need for new collections and identify those datasets that may take longest to develop.

A separate effort should focus on data about unusual physical conditions that are not encountered often enough to develop specific algorithms. Such conditions, which include “keyhole” irises and scars or dysplasia in friction ridge skin, might be handled using special processing.

4. Coordinate Data Gap Analysis

Because biometric and forensic datasets are so costly to assemble, significant gaps in needed data should be identified before a major new dataset is created. The analysis should identify the key areas of research, development, and system acquisition that are hindered by lack of data and highlight how data collections could be shared among users to minimize costs. All of the datasets will need to be categorized using the taxonomy in the Data Catalog.

Data collection coordinated by NIST or NIJ should prioritize and balance broad-based needs against the costs of collection. For example, the lack of public (or large sequestered) latent palm datasets is impeding the advancement of an entire modality, whereas the lack of datasets for a new acquisition device might be only of niche interest to a few vendors or researchers.

It is recommended that a person or small team coordinate, request, and collate information from a variety of sources to identify needs and priorities for data among these modalities:

- Face
- Person at a distance
- Friction ridge (finger and palm)
- Iris
- Voice
- Handwriting

5. Develop Best Practices in the Collection, Dissemination, and Use of Biometric and Forensic Datasets

Recommendations about data collection practices, pitfalls, and distribution issues have been compiled in a document, Best Practices in the Collection, Dissemination, and Use of Biometric and Forensic Datasets.

The types of problems — and recommendations — discussed generally fall into these categories:

- Collection: Problems include ad hoc or sloppy data collection; biased procedures; inadequate quality control; removing or filtering data. These can be addressed through planning, quality assurance procedures, and documentation.
- Documentation: Datasets and any revisions should be accompanied by complete documentation so that consumers can easily comprehend their source, potential use, and limitations. Standardized ways of characterizing datasets (e.g., taxonomy and metadata) may be of benefit.
- Use: Care must be taken to ensure that datasets are not misused or misrepresented, and that conclusions are not overstated.
- Complying with standards: All data collections must meet regulatory requirements along with applicable standards.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Resources: Include adequate funding and other resources so that the dataset can be updated as often as needed to avoid obsolescence.

6. Develop Sequestered Data Testing Plan

The sparse amount of data for performance testing can be mitigated by using sequestered data handled by trusted and authorized testing laboratories. This practice allows algorithm developers to submit their technology to the testing laboratory for evaluation without making the data public. The data must be continually updated to reflect new modalities, sources of information, new collection technologies, and applications.

To provide algorithm developers with easy access to test results and to ensure timely evaluations of revised algorithms, these tests should be available on an ongoing basis, similar to the distribution of NIST's Proprietary Fingerprint Template Evaluations. Additional funding is needed to support expanding repositories of sequestered data and evaluate new technologies.

7. Develop a Path Forward

Acting on these recommendations and suggestions will require a plan that provides a set of prioritized steps for implementation and that identifies potential early successes.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

2 Introduction

The lack of specialized datasets and the inaccessibility of large datasets drastically limit progress in forensic science and biometric technology. Data is an essential requirement for the development of algorithms to test and analyze biometric matching and evaluate accuracy for proof of concept. Testing is usually performed with small datasets to provide rapid turnaround as the algorithm undergoes numerous improvement cycles. Frequently, the datasets may emphasize a specific characteristic to test a particular feature of an algorithm. As the algorithm improves, the developer and the user conduct tests to see how it will perform in a real-world environment. Large sets of operational data are used in performance testing to provide statistically meaningful results. Restrictions on using operational data are a limiting factor in terms of exploring and improving systems.

The lack of sufficient research data is a major limitation to research, the development of new technology, and the process for making informed policy decisions in forensic science and biometrics. To address the problem, NIST, with funding from NIJ, convened a workshop entitled: **Improving Biometric and Forensic Technology: The Future of Research Datasets.** In addition to focusing attention on the issue, the workshop identified critical areas of need and obtained advice from researchers conducting data collection to understand how they address privacy issues. The Biometric and Forensic Dataset Catalog was also developed to provide a compilation of publicly available databases and their associated metadata enabling researchers to search for data relevant to their research uses. The symposium provided a forum for the discussion of the adequacy of existing publicly available datasets, as input for developing a multi-agency roadmap for collection and dissemination of datasets in the future. The modalities addressed included fingerprints and palmprints, iris, face, person at a distance, voice, and handwriting.

2.1 Sponsorship

The NIJ is the research, development, and evaluation agency of the U.S. Department of Justice and is dedicated to researching crime control and justice issues. NIJ provides objective, independent, evidence-based knowledge and tools to meet the challenges of crime and justice. The Office of Investigative and Forensic Sciences is the Federal Government's lead agency for forensic science research and development. It also administers programs that support crime laboratories and law enforcement agencies in their efforts to expedite processing of high-volume cases, reduce backlog, and provide training in new technologies. Forensic science program areas include Research and Development in Basic and Applied Forensic Sciences, Paul Coverdell Forensic Science Improvement Grants, DNA Capacity Enhancement and Backlog Reduction, Solving Cold Cases with DNA, Post-Conviction DNA Testing of DNA to Exonerate the Innocent, National Missing and Unidentified Persons System, and Using DNA to Identify the Missing.

A nonregulatory agency of the Department of Commerce, NIST promotes U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve quality of life. It accomplishes these actions for the forensic science community through the Forensic Science Research Program, located within the Special Programs Office (SPO) at NIST. The Forensic Science Research Program directs research efforts to develop performance standards, measurement tools, operating procedures, guidelines, and reports that will advance the field of forensic science.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

2.2 Acknowledgements

The authors of this document gratefully acknowledge the speakers and participants of this workshop whose discussions formed the basis of this report. The following individuals are also acknowledged for their contributions to this document's development and review:

Austin Hicklin, Noblis
Brad Wing, (former) ITL, NIST
Michael Garris, ITL, NIST
Rich Cavanaugh, Director, SPO, NIST
Mark Stolorow, Director of Organization of Scientific Area Committee Affairs, NIST
Susan Ballou, Manager, Forensic Science Program, NIST
Reva Schwartz, Research Manager, Forensic Science Program, NIST

2.3 Disclaimer

Opinions or points of view expressed in this report are those of the workshop participants and organizers do not necessarily reflect the official position or policies of the U.S. Department of Commerce. Mention of commercial products or services in this report does not imply approval or endorsement by the National Institute of Standards and Technology, nor does it imply that such products or services are necessarily the best available for the purpose.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

3 Workshop Overview

The workshop on **Improving Biometric and Forensic Technology: The Future of Research Datasets** was held on 26 and 27 January, 2015, at the Gaithersburg, Maryland facilities of the National Institute of Standards and Technology (NIST). The workshop included discussions on the following topics:

- The Role of Datasets in Improving Biometric and Forensic Technology.
- Privacy and Legal Issues in the Collection, Distribution, and Use.
- Discussion of New Catalog of Publicly Available Datasets.
- Statistical Significance Issues.

The workshop also included panels on several biometric and forensic modalities, including fingerprints, iris, voice, handwriting, and face/person-at-a-distance which discussed:

- State of the Industries
- Existing Data Collection Projects
- NIST Research and Cataloguing Efforts
- Future Needs

The detailed workshop speaker list and agenda are available at: http://nist.gov/forensics/biometric-research-datasets-event.cfm. Hyperlinks to additional documents and activities that relate to the workshop are provided in the references section of this report.

4 The Role of Datasets in Improving Biometric and Forensic Technology

The meeting began with a discussion on the importance of improving biometric and forensic technology. Susan Ballou, Forensic Science Manager at NIST's Special Programs Office, reviewed her experience in the forensic laboratory and the critical importance of data. A plenary session followed, with panelists Mark Greene, Program Manager at NIJ; Austin Hicklin, Fellow at Noblis; and Brad Wing of the NIST Information Technology Laboratory (ITL).

4.1 Initiation of NIST/NIJ Collaborative Efforts

Mark Greene (NIJ) provided the background and context for NIJ's efforts to improve biometric datasets. The agency seeks to address the limitations of existing biometric data collections in developing a new latent fingerprint database. The new database would help build the next generation of latent fingerprint algorithms, Greene said. The same limitations apply to all biometric datasets, he noted. Datasets are also highly restricted by the regulations imposed by the privacy protection provisions. Citing the SD27 Latent Fingerprint Set, originally created to evaluate the International Automated Fingerprint Identification System (IAFIS) performance in the late 1990s, Greene pointed out that it is now the most over-trained database in biometrics. As of March 2017, SD27 is no longer distributed by NIST.

Greene also described the goals to address these challenges in NIJ's collaboration with NIST (Figure 4-1).

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

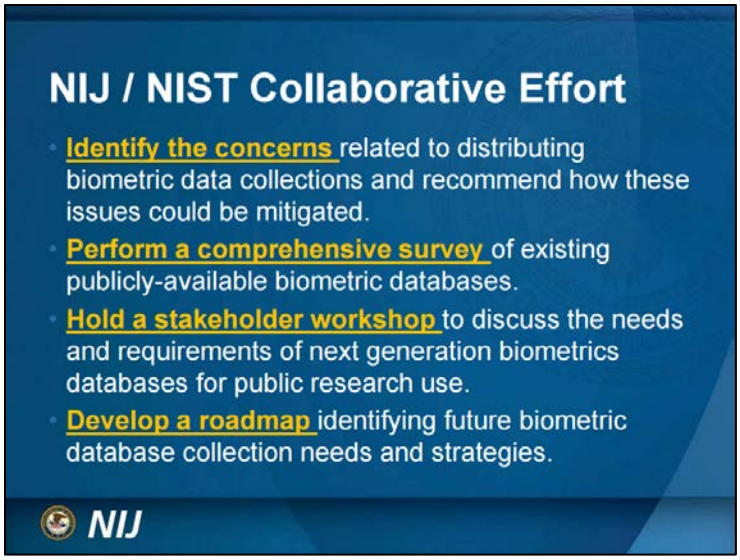


Figure 4-1 : NIJ/NIST collaborative effort to address biometric and forensic dataset availability.

4.2 Challenges in the Collection and Use of Biometric and Forensic Datasets

Austin Hicklin and Brad Wing presented an overview of the challenges in compiling biometric datasets and how they are used. Hicklin pointed out that datasets must be appropriate for a given purpose and provide a good match for the proposed application—both in terms of the intended population and the devices used to capture the data. Datasets to be used for developing proofs of concept or for emerging technology need not meet the rigorous requirements for determining algorithm applicability in an operational application.

Hicklin encouraged researchers to adopt a healthy skepticism when using datasets (Figure 4-2). If datasets are improperly collected or misused, things can go very wrong. These problems are aggravated if the dataset characteristics are not known or are not described.

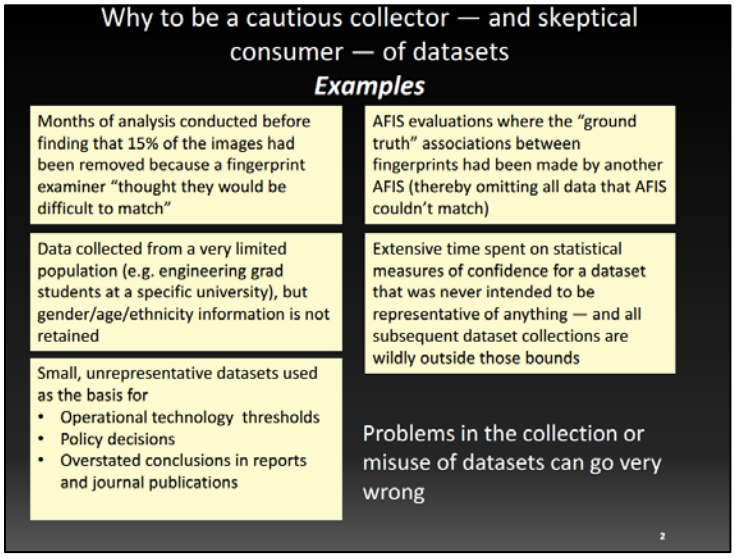


Figure 4-2: Examples of why to be a cautious user and developer of biometric datasets.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

To address the misuse of data and to help researchers and operational personnel with their data collection, Hicklin proposed the development of a best practice guide for biometric practitioners. The guide would list major dataset collections, caveats, and note potential pitfalls in their use.

For some modalities, Software Development Kit (SDK) testing programs have overcome the limitations imposed by the unavailability of datasets in operational environments, Hicklin noted. In these programs, researchers and developers submit their algorithms for testing on a large sequestered dataset maintained by the government. By using sequestered data, sensitive data are not released; the submitted algorithms are tested against the sensitive data by the data host agency, and the subsequent results are released to the algorithm submitter.

Wing cited several key questions (Figure 4-3) that should be posed in developing a comprehensive data collection and availability framework.

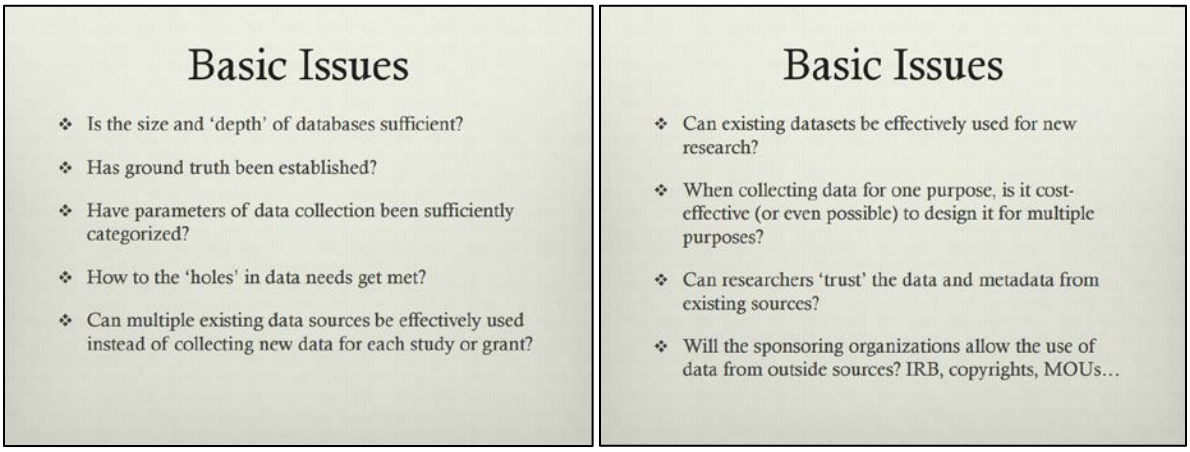


Figure 4-3: Basic questions about the use of biometric datasets.

Citing the CASIA-IrisV1 dataset, which replaced pupil images with a "circular region of constant intensity to mask out specular reflections from the NIR illuminators," Wing posed the question: Under what circumstances should altered data be suitable for research? He posed a similar question about outliers and degraded data. Some databases become so familiar to researchers that their systems become tuned to ensure 'success.' NIST developed SDK evaluations of fingerprint software using sequestered datasets to help combat this problem, as described by Wing in Figure 4-4.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

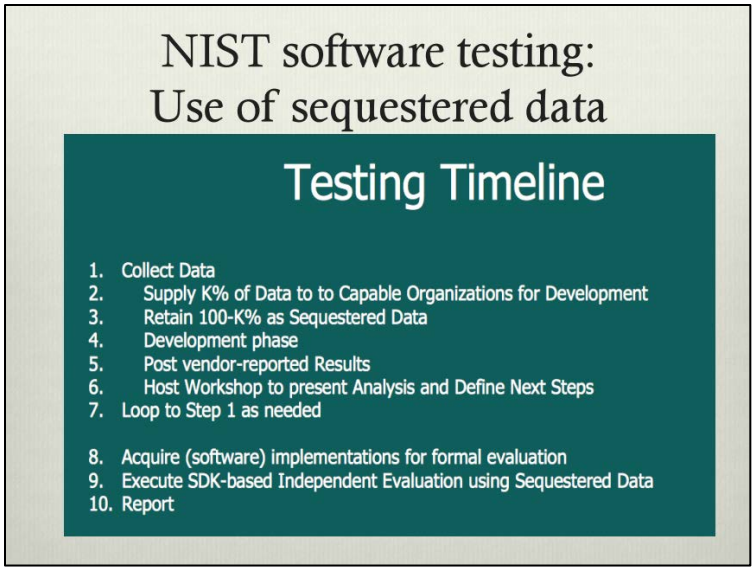


Figure 4-4: NIST software testing using sequestered data.

Wing also described several issues that, if ignored, can limit the quality and representativeness of the data collection and the privacy of an individual. Cross-cultural issues, for example, may arise since capture procedures for some biometrics may touch sensitive religious issues such as removing face veils. Ensuring that data is kept private is particularly important in DNA testing, which may reveal such socially sensitive issues as paternity and familial relationships. Cost can also be an issue, depending on the amount of data needed, ergonomic considerations, and device/camera representativeness.

4.3 Discussion

Session speaker and audience discussion highlighted the following needs:

- For all biometric modalities, the research community needs access to more public datasets.
- Concerns about how biometric data collections are shared and distributed need to be addressed and recommendations as to how these issues could be mitigated need to be developed.
- The practice of using sequestered data for performance testing (called SDK testing) should be expanded.
- NIST's new Biometric and Forensic Database catalog can serve as a one-stop source of information for the availability and characteristics of datasets. However, the catalog needs to be analyzed to identify data gaps and should be maintained and updated to include new databases and related information.

4.3.1 Research Challenges

- Biometric practitioners need a best practice guide that compiles major dataset collections, caveats, and notes potential dataset usage pitfalls. The guide would ensure that the datasets:
    - Represent real-world operational data
    - Include a number of subjects/sources with multiple representations
    - Include reliable metadata—sex, date of birth, date of capture, capture technology, resolution, finger position, nationality or race, pressure, moisture, rotation, etc.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Represent the target population and target technology— diversity in age, sex, capture technology, race, etc.
- Have sufficient support so that the collection process can be ongoing and provide a basis for longitudinal studies and emerging technologies, e.g., new data capture devices such as contactless fingerprint scanners

4.3.2 Data Collection Issues

- The workshop identified several data collection issues and recommendations:
  - Ground truth (attribution of source) should be established for the data collection effort.
  - Policies should be developed to address the alteration of ground truth data (tattoo change, surgical changes, unknown compression/decompression cycles, etc.)
  - New dataset collections need to be identified and compiled. Should a list be maintained and should there be a published list of data collection plans by public agencies?
  - Appropriate capture devices should be well described and represented in the dataset.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

5 Privacy and Legal Issues in the Collection, Distribution, and Use of Biometric Datasets

The protection of human subjects is a major concern in the collection and use of biometric data. Privacy provisions are regulated by the federal government and defined by the Common Rule, the standard of ethics by which government-funded research involving human subjects is held and with which nearly all academic institutions have agreed to comply. These regulations govern the operation of Institutional Review Boards (IRBs) which provide oversight of activities involving human research. The Rule, and the regulations for the IRBs, were developed for the protection of research subjects. Biometric data was not in its original scope; it was only later extended to subjects contributing biometric data. The application of these rules to biometrics has proven to be a major source of inconsistency and confusion.

To address the multiple privacy issues and concerns, the workshop convened a privacy panel moderated by Henry Wixon, Office of Chief Council at NIST. Panelists included Mike Garris, a senior scientist at NIST ITL, Jeremy Dawson, Research Assistant Professor at the Lane Department of Computer Science and Electrical Engineering at West Virginia University, Dorothy Glancy, Professor of Law at Santa Clara University, and Greg Michaud, Director of Michigan State Police Forensic Science Division.

5.1 Summary of Presentations

Wixon observed that technology almost always races ahead of the laws that regulate it. This can lead to a bending of regulations and serious violations of the Common Rule. Dawson supported Wixon's assertion that many privacy rules are derived from medical research and are not a good fit for biometric research.

Dawson described the database collection efforts conducted by WVU in multiple modalities. Collections were initiated in response to specific research demands and typically required the creation of regular and specialized or outlier (pathological) datasets for the Federal Bureau of Investigation (FBI), Department of Homeland Security (DHS), and Office of Naval Research (ONR). Figure 5-1 describes the differences Dawson observed between lab-based data and operational data.

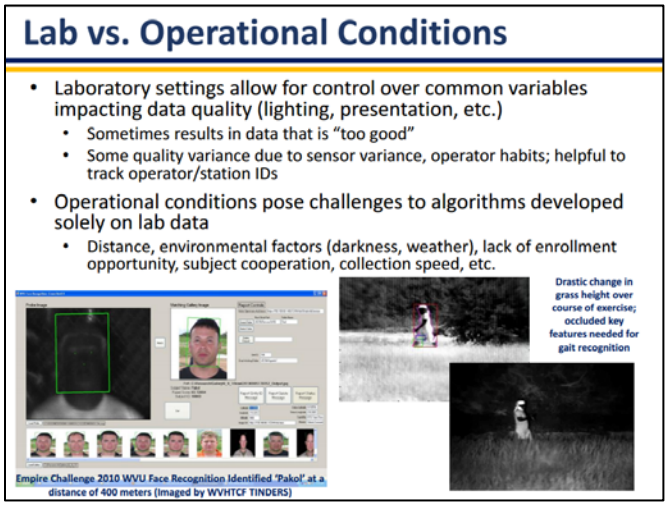


Figure 5-1: Lab vs. operational data conditions.

Garris provided an introduction to the privacy issues based on his experience and practice at NIST. He summarized the lessons learned over the last twenty years that biometric data has been collected. This period saw major institutional changes in privacy policy practice related to the Common Rule. To satisfy the privacy requirements, the research community must ensure that the data collection process does not

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

physically or psychologically harm human subjects, protects basic rights of privacy, and safeguards data samples in support of open research, Garris said.

The IRB process is not always a smooth, uneventful process, he noted. For example, outside organizations requesting data may use the originating entity's IRB documentation; however, the level of scrutiny, requirements, and process varies by research organization. Researchers should apply the highest standards for drafting and maintaining documents, including linking current versions of required elements with updated documentation process (e.g., increase in # subjects). Garris cautioned not to cut corners on the privacy review process because eventually it is likely to inhibit the use of the data.

Garris' experience with IRBs also revealed that user-agreements are often required with research datasets. These should be brief and easy to understand and should include a process for adaptation to new uses and/or updating with similar data. Given the differences in the implementation of the IRBs, it is necessary to develop and publish standard templates and best practices for IRB documentation.

Garris observed that not all biometrics need the same level of privacy protection. Facial images require a stringent level of protection because facial characteristics are so easily recognized. Additional protection required for this modality may include tighter access control, lock-down of the data both physically and logically, disabling of online/external search and export services, and development of modality-specific policies, procedures, and training.

Biometrics research is evolving into human identity research. This is likely to involve the collection and use of multiple biometric modalities (e.g., fingerprints combined with face photo and iris images and other measures.) and soft biometrics based on biographic information. Using multiple types of data poses special challenges for formulating protection policies for human subjects. It may not be possible to have a single policy for all biometric data collections.

Glancy, an author of the Biometric Research Databases and the Privacy Act of 1974, reviewed the history of privacy protection. She noted that the Privacy Act of 1974 only applies to Federal executive branch agencies, including the Executive Office of the President. Figure 5-2 highlights key provisions of the Act.

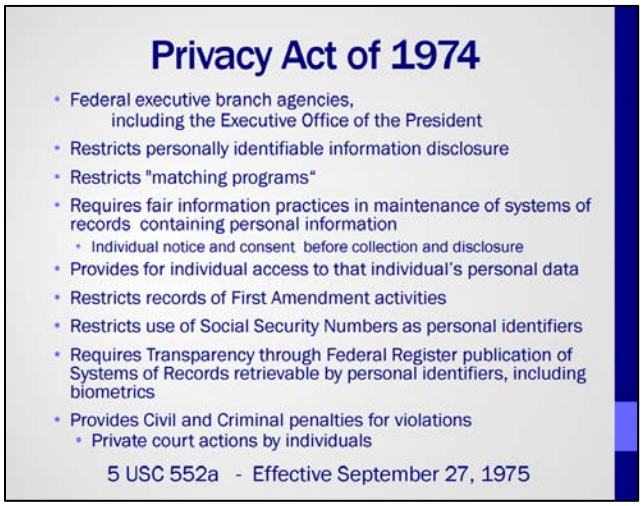


Figure 5-2: Highlights from the Privacy Act of 1974.

The Act provides for exceptions for research involving analysis of existing data and other materials, if they are publicly available, or in cases where the collected data cannot identify test subjects. The exception process is addressed by the IRBs, which have been created by multiple research organizations.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Universal Principles spelled out by the Act include:

- People should be informed if their personal information is being collected.
- Personal information should only be used for the purpose for which it was collected.
- Personal information should only be collected for a clearly identified purpose.

The Privacy Act of 1974 requires that:

- Each record is individually identified.
- Personal information is provided.

Michaud noted that the Michigan State Police rely on a Michigan law that allows the collection of operational biometric data for research purposes. He works in partnership with scientist Anil Jain of Michigan State University to develop datasets for research purposes. His research focuses on latent fingerprints, facial recognition, DNA, biometric fusion, and firearms identification. He also takes advantage of a Michigan law that allows the use of DMV facial data for criminal investigations.

5.2 Discussion

Policies defined by the U.S. Department of Health and Human Services (HHS) / Office for Human Research Protections (OHRP) Federal Policy for the Protection of Human Subjects are known as the "Common Rule"). It was developed for the medical community and is difficult to apply to biometrics and forensic research, Wixon and Dawson pointed out. HHS is currently reviewing the Rule and considering possible updates. Wixon, who is engaged in the update effort, suggested that NIST coordinate with HHS and inform the agency of the issues faced by the biometric and forensic communities. Greene suggested that the wider research community develop a consolidated position on this issue.

In response to a moderator's question, the panel discussed whether an IRB performed by one group should be automatically acceptable to another user. Wixon opined that sharing was only allowable if the same facts underlying the IRB apply. Dawson said that in his experience IRBs are not transferable.

There is no stock answer about whether deceased persons have privacy protection. While the 1974 Privacy Act does not protect the biometric records of deceased persons, several state laws do. Thus, use of deceased data must be decided on a case-by-case basis, depending on the source of the data and the state where it might be used.

In the discussion on privacy, Wing presented two consent forms: one from Cornell University, the other from Notre Dame University. The two forms have different levels of detail, underlining the observation made by several speakers that the IRB processes are not uniform.

5.2.1 Specific Recommendations

The following recommendations were made by session speakers and/or identified in the subsequent discussions following the presentations:

- NIST and NIJ should develop a consolidated position on proposed changes related to the Common Rule update process.
- NIST and NIJ should coordinate with HHS on revising the Common Rule. Representatives of the NIJ and NIST General Counsel should participate in the process.
- A standardized approach and common templates should be developed to ensure privacy protection, allowing data collection to proceed in a timely manner. The approach should address the following:
    - IRB templates

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Consent forms
- Data User Agreements for research datasets which include:
    - Consistent policies for in-house research.
    - Consistent policies for public use (due to sharing of data).

5.2.2 Research Challenges

The workshop participants identified several challenges likely to affect the creation of new datasets for research use:

- There is no universally accepted standard for data collection and use. (Legal issue may prohibit the use of data collected by a private company, even if the company had the consent of participants.) Developing such a standard should be investigated.
- Current policy does not take into account that different biometrics may require different levels of privacy protection. For example, because facial images are inherently recognizable, they require a high level of privacy protection. Data on other modalities may not need the same level of protection.
- The research community may need to develop a policy document that addresses the use of publicly available databases that includes guidance on biometric samples from deceased individuals. This may be especially important in light of recent state laws governing the privacy of such data.

5.2.3 Implementation Issues

- Consider developing policies that account for the necessary degree of privacy protection as a function of how the data is to be used (e.g., laboratory use, performance testing, etc.).
- Decide where standardized consent forms and templates should be posted.
- Determine language on use of data for purposes for which they were not originally intended (data sharing). Note the right to use and publish the data may be different between the user and the owner of the data. Both owners and users of the data need to address this issue.
- Develop an approach to define restrictions on the use of the data.
- Investigate rules for combining data obtained under different IRBs and determinations.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

6 NIST Biometric and Forensic Database Catalog

Shannan Williams, project manager for the Forensic Science Research Program at NIST, summarized the Biometric and Forensics Research Database Catalog Project. This project was sponsored by NIJ and is being implemented by NIST. It is a comprehensive catalog of publicly available biometric and forensic datasets for the following modalities: finger/palmprints, iris, face, person at a distance, voice, and handwriting. Williams described the key goals (Figure 6-1) of the effort and the sources of the data, which include web searches, journal articles, and universities with well-known programs in biometrics.

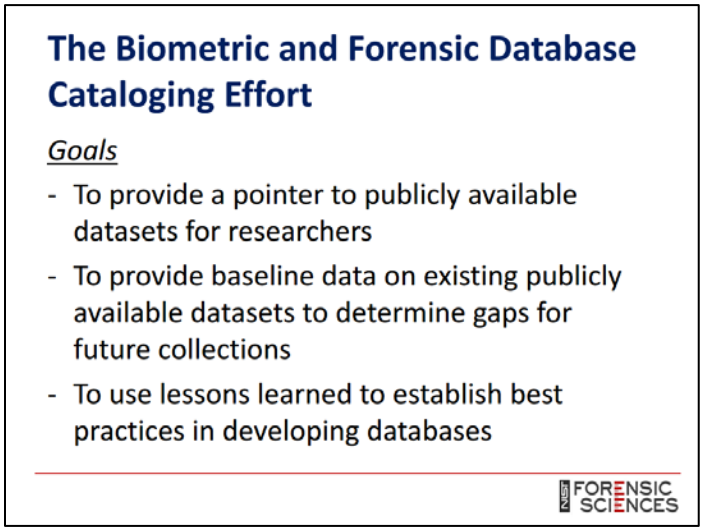


Figure 6-1: Goals of the biometric and forensic dataset catalog.

At the time of the presentation a total of 362 datasets have been identified and 165 have been posted on https://tsapps.nist.gov/BDbC/. Figure 6-2 describes the breakdown by modality of the datasets contained in the catalog. The Data Catalog website was created by in-house developers at the NIST Office of System Information Management (OISM). It can be used internally for data entry and externally to conduct searches. Users can conduct searches via text or taxonomy.

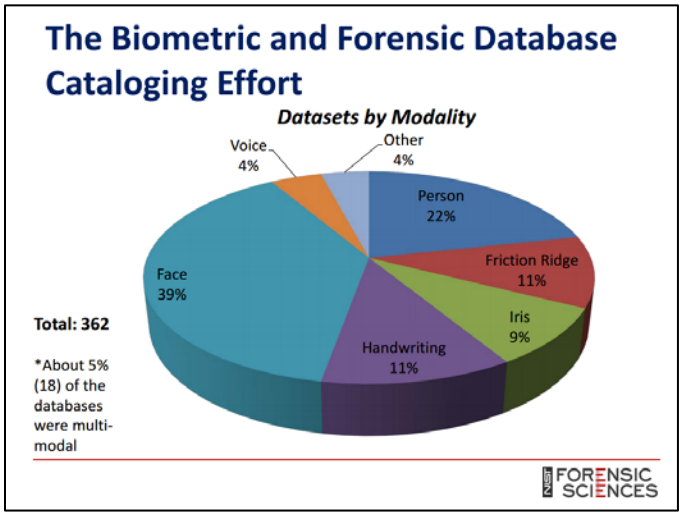


Figure 6-2: Breakdown of catalogued datasets by modality.

Williams also described the website's interface and provided screenshots as shown in Figure 6-3.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Figure 6-3: Biometric and Forensic Research Catalog Homepage

Additional work is required to complete the Catalog. Williams requested feedback on the existing catalog- how well it is categorized, and whether it is useful. The information is expected to be included in a Best Practice guide for developing future databases.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

7 Modality Panel: Fingerprint/Palmprint

The fingerprint/palmprint panel included Anil Jain, University Distinguished Professor, Department of Computer Science and Engineering, Michigan State University; Jeremy Dawson, Research Assistant Professor, West Virginia University; Craig Watson, NIST-ITL; and Elham Tabassi, NIST-ITL.

7.1 Summary of Presentations

Jain observed that friction ridge data comes from a diversity of sources and has many applications. Sources include federal agencies, state forensic laboratories (e.g., Michigan State Police), national registries (e.g., Unique Identification Authority of India), biometric customers, vendors, and academic laboratories. Despite the large number of applications and the relative maturity of friction ridge biometrics, only a few datasets are in the public domain, as shown in Figure 7-1.

Public Domain Fingerprint Databases

Type	Database	Data size	Year released	Comments
Latent & exemplar fingerprints	NIST SD27	258 latents & mates	2000	Latent matching research
	WVU	449 latents & mates	2008-2010	
Rolled	NIST SD4	2,000 X 2 fingerprints	1992	Classification and matching research
	NIST SD14	27,000 X 2 fingerprints	2002	
Digital video	NIST SD24	100 fingers	1998	50 males and 50 females
High resolution	NIST SD30	360 fingerprints	2001	Both 500 and 1000 ppi
Optical FTIR, capacitive & sweep	FVC	17,760 fingerprints of 2,360 fingers	2000, 2002, 2004, 2006	Including real and synthetic fingerprints
	Cross device	77,760 fingerprints of 720 fingers	2012	9 sensors
Spoof	LivDet	44,000 prints; spoof made of 5 materials	2009,2011, 2013	Fake and real fingerprints

Public Domain Palmprint Databases

Type	Database	Data size	Year released	Comments
Low resolution (~100 ppi)	PolyU palmprint	7,752 palmprints	2003	386 different palms
	PolyU Multispectral palmprint	6,000 palmprints	2008	250 subjects
High resolution (500 ppi)	THUPALMLAB	1,280 palmprints	2011	80 subjects
Latent palmprint	LPIDB v1.0	380 latent palmprints and mates of 100 different palms	2014	28 males and 23 females

Figure 7-1: Public domain fingerprint and palmprint databases.

Jain noted that many research questions require large-scale sets of friction ridge data. New applications and new technology place a continuous demand for new data. Even established applications, such as latent fingerprint matching, which was implemented in the 1980s, need further research. Public datasets are needed to:

- Train algorithms to learn (i) friction ridge models, and (ii) fingerprint similarity.
- Perform scalability studies of an Automated Fingerprint Identification System (AFIS).
- Analyze individuality of friction ridge patterns.
- Determine persistence of friction ridge accuracy using longitudinal studies.
- Determine minimum age at which fingerprints can be collected.
- Support performance testing using new capture devices and legacy data.
- Improve recognition of spoofing or avoidance attempts.
- Develop new algorithm concepts.

Watson described the friction ridge datasets created by NIST (NIST Special Database 4, 9, 14, 24, 27, 27A, 29, and 30) in response to requests by other federal agencies. He described their characteristics, sources, formats, and distribution limitations. The differences in the datasets addressed pattern variation, resolution, and compression algorithms. He noted that most of the datasets are based on scanned paper cards; there are no true livescan data. The reliance on inked card data is not representative of the modern

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

operational environment and may therefore produce erroneous performance data. All data are from deceased subjects and are available on CD-ROM. NIST currently no longer distributes SD27 to the public.

NIST often uses a phased approach for performance testing using sequestered data, Watson said. The phased approach enables vendors to develop and test the interface, learn about their algorithms' performance on a set of operational data, and train their algorithms, before running the final performance test. Watson also described a timeline of biometrics evaluations that NIST has used since 2002, as shown in Figure 7-2.

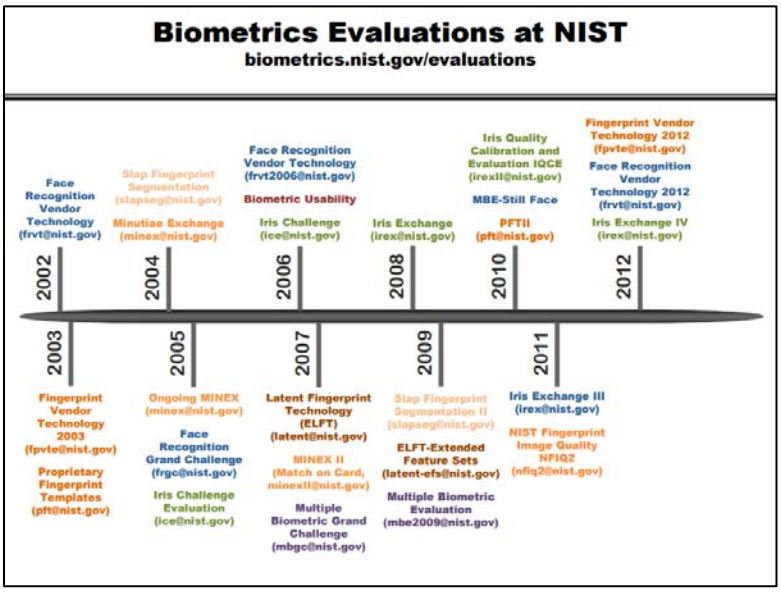


Figure 7-2: Timeline of biometrics evaluations at NIST.

Table 7-1 contrasts the differences between public and sequestered data and their uses, as described by Watson.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Table 7-1: Difference between Public Datasets and Sequestered Datasets

Public Datasets	Sequestered Datasets
Can distribute to all users	Cannot distribute to anyone
Smaller sample sizes (hundreds/thousands)	Larger sample sizes (millions)
Small storage capacity required	Increased storage capacity required
Mostly derived from older data (deceased subjects)	Operational data, often derived from on-going operations
When used for development – processed on participant hardware When used for validation — processed on participant and NIST hardware	Used for large-scale evaluation — run on NIST owned hardware only 8
Results are not always published	Results published in public reports

In his briefing, "WVU Biometric Data Collection Projects," Dawson described several fingerprint data collection efforts. The data is used for interoperability and human factors testing across data types. The WVU data collections were conducted in a laboratory environment focused on special conditions and/or characteristics, and were intended to be tested on emerging technologies.

The laboratory settings used for the data capture allow for control over common variables impacting data quality (lighting, presentation, etc.). Training can also greatly influence data quality and needs to be standardized. Dawson recognized that operational conditions pose challenges to algorithms developed solely on laboratory data. Many factors need to be considered, including distance from/to the image captured, the environment (darkness, weather), lack of enrollment opportunity, subject cooperation, and collection speed. Data collection in the university environment poses a challenge because of the lack of a representative demographic distribution, Dawson observed.

In her presentation, "Usefulness of Current Biometric Datasets," Tabassi provided a high-level view of the data collection environment and described her vision of how a data collection improvement program would address problems and challenges. Understanding how the data was obtained is a critical element in determining its usefulness and applicability. Tabassi said. She described three types of datasets, differentiated by their origin, and characteristics as described by Figure 7-3.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

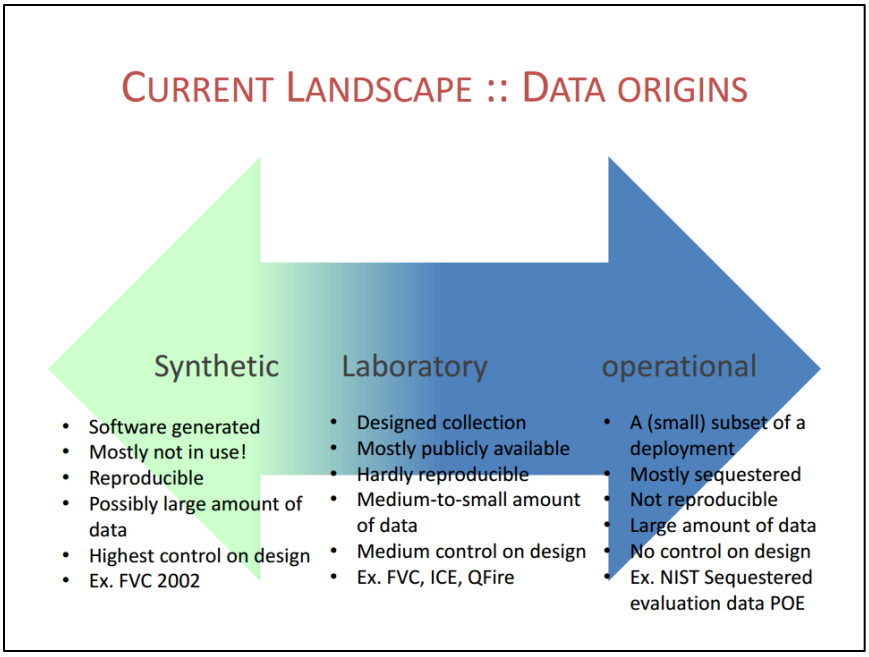


Figure 7-3: Current landscape of origins of biometric data.

Tabassi described a wish list for an all-purpose data collection in which data should:

- Be representative of real-world operational data.
- Include a large number of subjects/sources with multiple representations.
- Include reliable metadata –sex, date of birth, date of capture, capture technology, resolution, finger position, nationality or race, pressure, moisture, rotation, etc.
- Be diverse (representative of the target population and target technology) — age, sex, capture technology, race, etc.
- Be ongoing and extendable to provide a basis for:
  - Longitudinal Studies
  - Emerging Technologies, e.g., contactless fingerprints.

To achieve her "wish list" and ensure accurate characterization of operational real-world data, Tabassi proposed that researchers follow these suggestions:

- Design the collection so that it targets a feature in fingerprint analysis known as "learnt clusters," perhaps via uniform data collection protocol.
- Obtain a better understanding of required sample size and the associated uncertainty in measuring the error rates.
- Improve uniformity of reporting.
- Improve data integrity in laboratory collection.
- Develop a guidance document on consolidation —enrolled (gallery) and search (probe) compositions.

7.2 Discussion

The following summarizes the recommendations, challenges, and various data collection issues identified during the plenary session:

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

7.2.1 Recommendations

- Develop capabilities for algorithm performance testing with large friction ridge datasets, including palm print data.
- Develop a framework to maximize the sharing of data; this will reduce the burdens of both cost and time in data collection.
- Facilitate the ongoing development of datasets created in a laboratory environment that is focused on special conditions and/or characteristics and are intended to be tested in emerging technologies.
- Establish a dataset maintenance program.
- Develop a policy on how long data should be retained, taking into account the obsolescence of data capture devices.
- Develop policies on releasing data to outside users (physical distribution).

7.2.2 Research Challenges

The workshop identified several research challenges that are likely to affect and possibly hamper the creation of new datasets:

- Ground truth errors do occur, especially if consolidated data is used. Identifying errors and their origin is made more difficult by the need for anonymity to protect privacy. A better approach for resolving ground truth errors is needed.
- When new and legacy data are compared analysis requirements are needed to determine the actual impact on performance.

7.2.3 Data Collection/Implementation Issues

Operational conditions pose challenges to algorithms developed solely on lab data. Many factors need to be considered, including contact versus contactless data capture, environmental factors, lack of enrollment opportunity, subject cooperation, and collection speed. Implementation of data collection requires that the following concerns are addressed:

- Datasets need to be continually created to support emerging technologies (e.g., new contact fingerprint devices, 3D data capture, and contactless devices for image compatibility analysis).
- Datasets should be collected in a laboratory setting so that common variables affecting data quality, such as lighting and presentation, can be controlled.
- Guidelines are needed to address cost of distribution.
- Develop guidelines to ensure greater uniformity in training, which will enhance the quality of data collection.
- Develop a framework that will allow latent practitioners to create training data.
- Develop dataset of pathological cases (outlier) datasets for algorithms development and analysis.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

8 Modality Panel: Iris

The Iris Panel included Stephanie Schuckers, Clarkson University; George W. Quinn, NIST-ITL; and Arun Ross, Michigan State University. Patrick Grother, NIST-ITL, moderated a panel discussion on future needs.

8.1 Summary of Presentations

Schuckers gave an overview of the state of the art of iris data collection, based on her own work and that of co-author Brian Walczak, both participants in the Center for Identification Technology Research (CITeR). Quinn and Ross added their perspective on the state of the art in iris identification technology. All of the panelists identified issues for further study, and identified new directions for iris research. Schuckers reviewed the science of iris matching, including advances in modeling iris physiology and ongoing studies to improve accuracy of the matching. The briefing also addressed security issues related to spoofing of the iris image. Schuckers proposed methods to detect liveness and to detect or mitigate image alteration due to contact lenses. She emphasized the wide variety of capture techniques used in iris identification, each posing a different set of requirements for data collection.

Other issues that require study include how disease, dilation (enrollment and probe differences), aging (permanence of iris patterns), and eye disorders affect identification and data quality. Figure 8-1 displays examples of eye disorder images. Researchers have not yet determined the accuracy with which the iris of an individual can be identified nor the impact of scaling on the false-positive identification rate. The use of mobile iris capture devices also poses a data collection challenge.

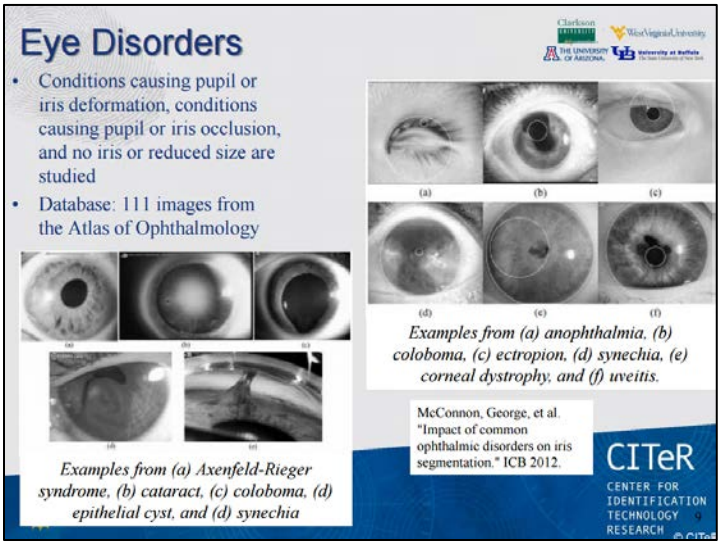


Figure 8-1: Eye disorder iris data examples.

Schuckers supported the need for greater database sharing, noting that an effective way of doing so is by using sequestered data for algorithm testing. Sequestered data provides additional benefits by making the results comparable and reproducible.

Quinn stated that Iris recognition is highly accurate when the quality of the images is good. However, scientists need to better understand when and how data quality impairs performance. The IREX III supplemental report found that among the 1,013 searches that failed to return the correct mate (for any of the top 3 matching algorithms), each had problems with the images. Performance of an operational system will be determined by the poorest quality samples and how frequently they occur, Quinn noted.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

He agreed with other speakers that dilation, aging, and eye disorders are areas for further research, and provided additional detail about the dilation of the pupil as shown in Figure 8-2. He pointed out that 45 of 1,013 unsuccessful searches probably failed due to abnormal pupil shapes. Other issues in iris identification include differences in illumination (multi-spectral data capture), specular highlights, and performance variations as a function of eye color.

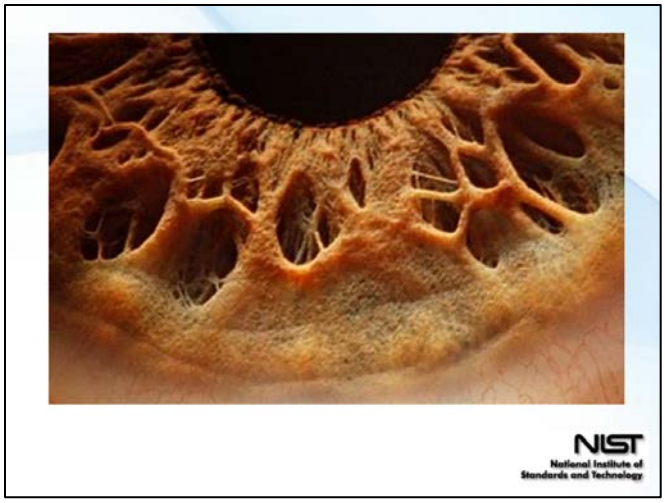


Figure 8-2: Microscopic view of pupil dilation.

Quinn identified several emerging areas of research, including:

- Forensic applications
- Surgical alterations
- Neo-natal iris matching

These concepts were further developed by Ross, who spoke about iris forensics. His presentation, "Iris as a Forensic Modality: The Path Forward," explored the multiple issues and research in using iris as forensic evidence. Little research has been conducted in this emerging field, he noted. Ross said he would like to develop a framework for using iris images as identification in a court of law.

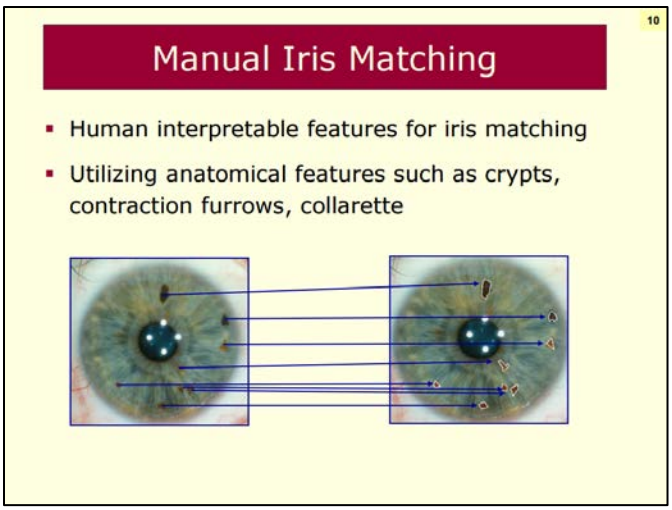


Figure 8-3: Manual iris matching.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Iris forensics also includes ocular forensics, the study of the regions surrounding the iris. Such studies include the eyesclera, eyelids, eyelashes, eyebrow, and skin texture. Figure 8-3 describes manual matching of iris features. Iris forensics also includes variations in color. Ross noted that the concept of iris color is not well defined and that color response to different illumination sources needs to be studied, as shown in Figure 8-4.

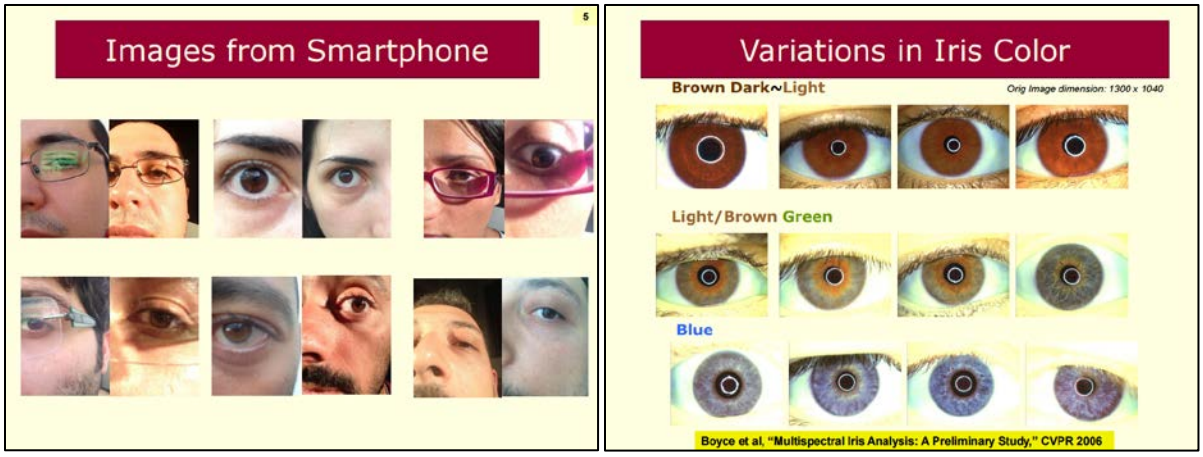


Figure 8-4: Iris image variations in color.

8.2 Discussion

During the panel discussion the speakers agreed that there was a need for large databases to test aging impact. There was also agreement that cross-sensor performance is an issue—although greater adherence to standards has somewhat mitigated the impact.

Dataset gaps exist in analysis of ethnic variation and of disease impact on iris characteristics. There are no public datasets for the study of aging and iris characteristics, particularly of the old and very young. Data on aging needs to be collected in a systematic manner so that sensor variability does not invalidate aging studies. Dawson commented that collection of age-related data may need to be limited to operational data. He observed that genetic research on the iris will have an impact on the IRB approval process, which, does not anticipate future use of the data. Researchers need to study how data obtained with a mixed type of illumination (visible and IR) impacts performance; particularly in forensic research.

The following summarizes the recommendations, challenges, and various data collection issues identified during the plenary session:

8.2.1 Recommendations

- Sequestered data should be used for SDK testing.
- A standardized dataset will spur the development and testing of algorithms that can further iris image research, including the study of common protocols, common data, and reproducible research.
- The data-sharing process among researchers should be improved.

8.2.2 Research Challenges

The workshop identified several research challenges that are expected to impact the creation of new datasets:

- Aging impact on iris recognition.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Iris matching in a court of law — (Can manual iris matching be sufficient for acceptance in a court of law?).
- Visual Iris Matching involving:
  - Texture analysis
  - Determining whether sex and ethnicity can be deduced from the iris image
  - Determining if biological age can be deduced from the iris image
  - Determining whether disease can be deduced from the iris image
  - Use of macro-feature searching of an iris database.
  - Use of macro-features to extract gender attributes.
- Image Forensics (automated process- new research areas):
  - Human interpretable features for iris matching using macro-features
  - Deducing illumination source from image
  - Resolving identity using “in-the-wild” images.
  - Determining imaging source of iris data.
  - Determining genetics of human iris.
  - Post-mortem (recognition boundaries as a result of deterioration)
  - Deducing dataset source from the images.

8.2.3 Data Collection and Implementation Issues

- Larger datasets are needed by researchers, particularly to address aging effects.
- The need for several new types of datasets were identified, including those that contain:
  - Iris capture at standoff distance (identified issues include shutter flutter and subject movement)
  - Off-angle capture
  - Periocular area datasets
  - Highly diverse capture devices for cross-sensor comparisons across multiple lighting spectra
  - Various means of illumination
  - Disease, dilation (enrollment and probe differences), abnormal pupils, surgical alterations and eye disorders
  - Neo-natal iris matching
  - Purkinje images[2]
- Datasets are needed to accommodate new forms of biometrics related to the iris, such as:
  - Eye movement biometrics (liveness and movement pattern recognition)
  - Conjunctival vasculature (sclera characteristics)
  - Periocular and ocular measurement
- Multiple-event data are required to enhance data quality measurement and improvement algorithms.

[2] *Purkinje images are reflections of objects from the structure of the eye. They are also known as Purkinje reflexes and as Purkinje-Sanson images.*

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

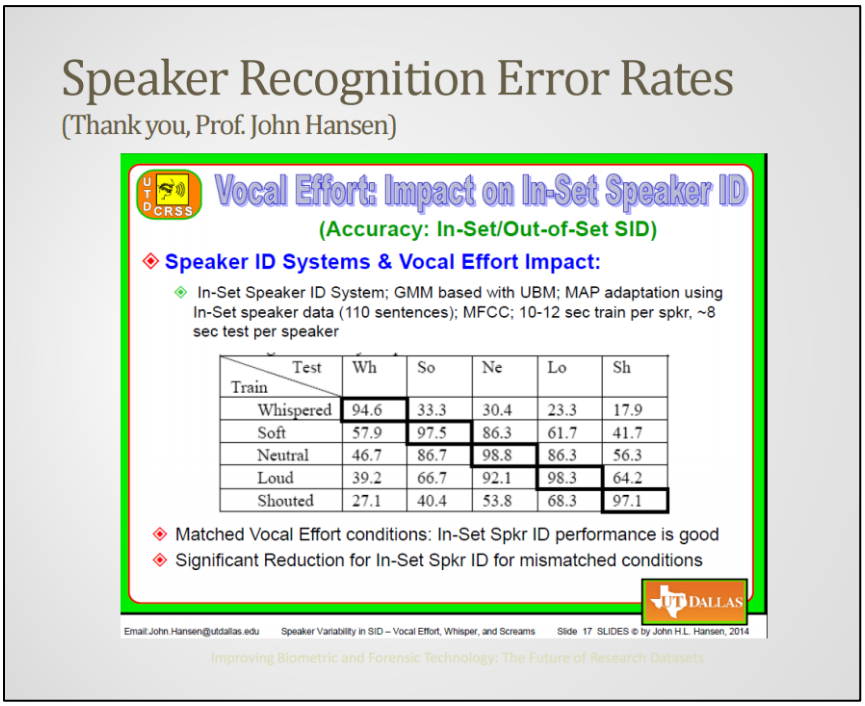
9 Modality Panel: Voice

Peter Higgins, Higgins & Associates, presented material on the voice, which was co-authored by Jim Wayman, Research Administrator in the Office of Graduate Studies and Research at San Jose State University. Other panel members could not attend.

9.1 Summary of Presentations

Higgins' presentation, "Speaker Datasets for Research or FRE 702(b), Daubert and the Problem of Mismatched Conditions," described the discipline's move into the forensic world. Little or no voice data is presented as evidence in court. Current research focuses on determining under what conditions voice data might be admissible as forensic evidence.

Research databases are often not a good predictor of identification performance in an operational environment. Error rates are highly dependent on the dataset used. In the case of voice, such features as the manner and tone of speech, e.g. normal speech, whispered, or shouted, vary widely across the population. Currently it is almost impossible to get data on all of the speech characteristics for the same subject. However, normal-to-normal speaker recognition under optimal conditions is very high. Depending on the speech sample characteristics speaker recognition accuracy can vary from 17.1% to 98.8% as shown in Figure 9-1.



Train \ Test	Wh	So	Ne	Lo	Sh
Whispered	94.6	33.3	30.4	23.3	17.9
Soft	57.9	97.5	86.3	61.7	41.7
Neutral	46.7	86.7	98.8	86.3	56.3
Loud	39.2	66.7	92.1	98.3	64.2
Shouted	27.1	40.4	53.8	68.3	97.1

Figure 9-1: Speaker recognition error rates.

Determining error rates by testing for each situation is difficult. "Sufficient" testing requires "sufficient" data in all combinations of influence variables, most of which are unknown. There is also the problem of comparing data samples of different quality; degrading good quality data to create problems within the speech sample for use in speaker recognition testing (by humans or automated systems) is ineffective.

Continuous parallel recording has been found to be very useful for voice recognition. It was also noted that voice recognition is highly dependent on the type of operational data that is available.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

9.2 Discussion

The following summarizes the recommendations, challenges, and data collection issues identified during the plenary session:

9.2.1 Recommendations

- Guidelines should be developed for large data collection programs and for using sensitive data from real forensic cases.
- Determine how to make operational voice samples from 911 call centers, court ordered wire taps, and other sources available to researchers.

9.2.2 Research Challenges

- Guidelines on dataset collection are needed to address the first amendment issues associated with speech samples.
- Methods need to be developed for distinguishing multiple speakers (interrogator / interviewer and subject -maybe a translator) within single and multiple channel sources.
- Methods need to be developed for detecting language changes within a sample set.
- Assess how the performance of mentally or physically demanding tasks while speaking affects voice recognition.

9.2.3 Data Collection and Implementation Issues

The recommendations for speaker data collection include the development of datasets that include:

- Operational voice samples from 911 call centers, court ordered wire taps, etc.
- Different types of speech (interview, read text, recitation of known material).
- Multiple variations in the speaker age, native language fluency, sex, socioeconomic status and education, and dialectal factors.
- Many types of microphones (cellphones, land lines, radios, etc.), and acoustic and other environmental conditions (indoors / outdoors, distance to microphone, recording devices, channels, etc.).
- Different speech modes: such as whisper, shout, etc.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

10 Modality Panel: Handwriting

The handwriting panel included Ted Burkes, Questioned Documents Unit, FBI; Christopher Saunders, South Dakota State University; and Peter Tytell, Private Forensic Document Examiner. The discussions were moderated by Danice Ommen (FBI Laboratory Division).

10.1 Summary of Presentations

In Saunders' briefing, "Scientific Approaches to Statistical Analysis and Collection of Handwriting Databases," he described the process used in comparing handwriting samples. The comparison process, known as biometric individuality (of a population with respect to a comparison methodology), is illustrated in Figure 10-1.

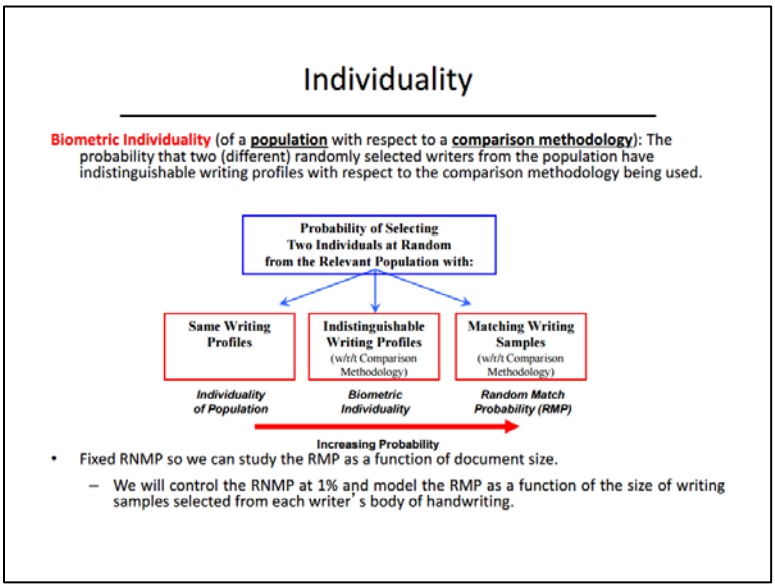


Figure 10-1: Biometric individuality.

The algorithm computes the probability that two randomly selected writers have indistinguishable writing profiles with respect to the comparison methodology being used. Two metrics are computed to determine individuality:

- Random Match Probability (RMP) is the chance of randomly selecting two subjects from the population and then randomly selecting a writing sample (of a given size) from each subject that is declared a match by the biometric.
- Random Non-Match Probability (RNMP) is the chance of randomly selecting a single subject and then sampling two documents from the selected subject's body of handwriting that are declared a non-match with respect to the biometric.

The process studies the RMP as a function of document size for a fixed RNMP. The RNMP is typically set at 1% and RMPs are computed as a function of the size of writing samples selected from each writer's body of handwriting. Figure 10-2 describes how the data was processed in the studies conducted by Saunders. The application of these types of statistics is the norm for handwriting analysis and is representative of evidentiary statistical presentations.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

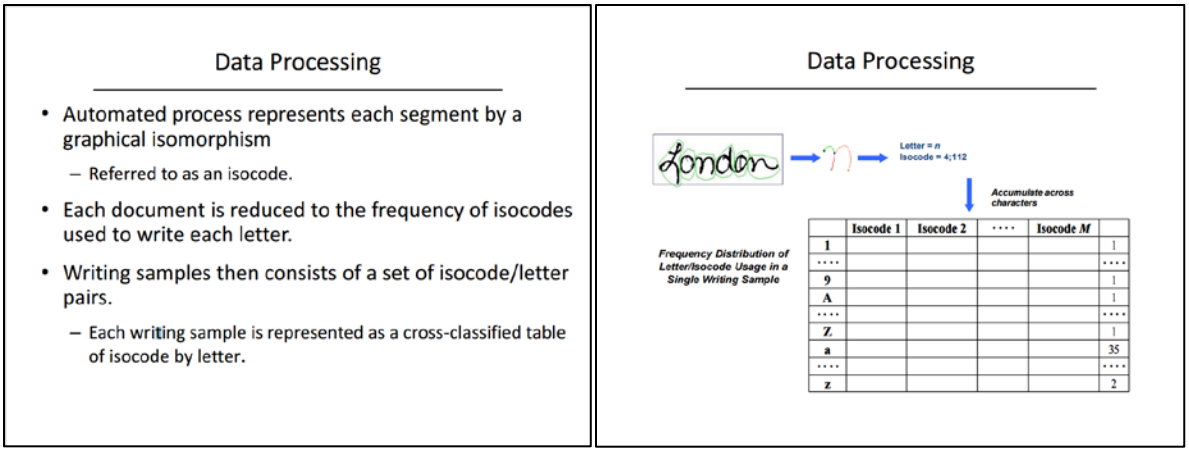


Figure 10-2: Handwriting image data processing.

Further, Saunders described a modeling approach to statistical analysis of handwriting comparisons. The model approach required the development of a subsampling algorithm for comparing two writers and the development of sufficient sizing conditions. The study demonstrated that the model can provide guidance on the relationship between the size of a writing sample collected and the RMP.

The model also provided an estimate of the behavior of upper confidence bounds on the RMP — given a combination of writing sample size and number of sampled writers. The ideal setting is a sample of documents from a large number of people in which no matches are observed when the collected documents are combined. Unlike other biometrics, the samples are not independent of each other, thus variance is hard to model. The model was used to calculate confidence curves for the number of writers compared, and the number of letters needed to determine a match at a given level of confidence.

Handwriting analysis does not involve open set searching of databases. Burkes observed that handwriting examinations are typically conducted using a side by side comparison, either of questioned writing to known writing or questioned writing to questioned writing. The inability to ensure a controlled environment for data capture makes data collection a difficult issue. Saunders is currently addressing this issue by constructing handwriting datasets for research purposes.

10.2 Discussion

The following summarizes the recommendations, research challenges, and various data collection issues identified during the handwriting panel session:

10.2.1 Recommendations

- Consider publicizing and possibly standardizing the Saunders' Individuality study as a performance standard for handwriting comparison.
- Develop guidelines for the minimal accepted sample size (for a dataset of forged signatures, for example)
- Develop a recidivist database.

10.2.2 Research Challenges

- Determine how age-related changes in handwriting affect handwriting recognition.
- Determine how intentionally distorted writing, simulations, forgeries, and disguised handwriting affect handwriting recognition.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Determine how changes in handwriting due to Intoxication, old age, or disease affect handwriting recognition.
- Develop method for categorizing different complexities in signatures.
- Determine how the choice of writing instrument (pen/paper, digital signatures) affect handwriting recognition.

10.2.3 Data Collection Issues

- Need to have more datasets that are representative of writing styles and a variety of environments and devices.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

11 Face and Person at a Distance

The panel included Mark Burge, Office of Smart Collection, Intelligence Advanced Research Projects Activity (IARPA), Office of the Director of National Intelligence (ODNI); Brendan Klare, Noblis; Marcia Kimura and Rebecca Erikson, Department of Energy, Pacific Northwest Laboratory (PNNL); and Jonathon Phillips, NIST ITL. Nick Lombardo, a co-author of the PNNL briefing, was not able to attend. The panel discussion was moderated by Richard Vorder Bruegge, FBI Laboratories Division.

11.1 Summary of Presentations

Burge described the Janus project: a technology development effort that will identify unconstrained images obtained in the wild. The project has multiple teams and substantial funding for its implementation. Janus started a year ago, with the creation of an annotated database. This is critical because identification in the wild involves large amounts of unconstrained data in which ground truth is elusive. To conduct research, algorithm developers need controls. The data must be collected under realistic conditions, typically using a well-crafted scenario. The collection will consist of very large datasets that can be partitioned in accordance with the developers' needs. This, of course, is likely to raise privacy issues because this process has little or no precedence, Burge noted.

To obtain "ground truth," the metadata on all of the images must be annotated, including their sources, Burge observed. All of the "unconstrained" data have some limitations. For instance, automated image detectors, which are inherent to most image capture devices, played a role in some of the data collection. Humans are needed to annotate images not captured using automated means and to verify data captured by automation. The data will eventually be made publicly available, possibly through NIST. This will give a wider set of players an opportunity to do independent research, Burge observed.

Burge noted that image comparison obtained without constraints of the capture environment ("identification in the wild") primarily involves images captured by video. This emerging modality raises many new issues; such as how to define a match in a video sequence. The project has purposefully varied demographics — gender, age, and ethnicity. Up to 20 contractors are expected to be working on the project, which is why dataset development has received such a high priority. It must be validated prior to system development, Burge said.

Klare described the development of a dataset for the support of Face Recognition at a Distance (FRAD). To better understand biometric capabilities as they apply in an FRAD operational settings, data is needed that best replicate the intended scenario. Data are also needed to train and evaluate the algorithms being developed. FRAD focuses on two operational paradigms: watch list identification and re-identification. (Re-identification is the determination of whether data from the same object or person lies in different samples or sample segments). FRAD typically compares images captured under non-controlled conditions against a large gallery of frontal facial images (See Figure 11-1).

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

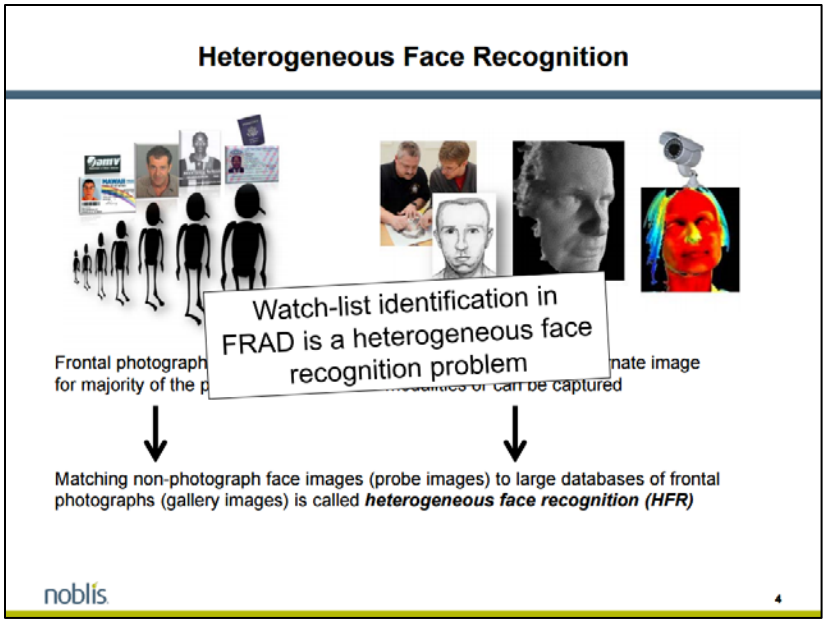


Figure 11-1: Heterogeneous face recognition.

Klare described several databases available to researchers, including:

- Korea University's Face-recognition database used in research by WVU and partially collected by the West Virginia High Technology Consortium (WVHTC) Foundation. It contains images captured at a distance.
- University of Colorado at Colorado Springs (UCCS Large Scale Unconstrained Open Set Face Database is a fully covert FRAD database (just approved by IRB), with images captured in public environments without subject cooperation (to be released soon).
- University of Maryland (UMD) Remote Face Database contains images captured at a distance in an unconstrained environment from cooperating subjects.
- Pinellas County Sheriff's Office (PCSO) has provided the research community with many unprecedented datasets.

Kimura and Erikson of the PNNL presented an overview of the dataset collection process using a process flow diagram (Figure 11-2). They focus on scenario-based testing in which the subjects are known and the capture process conditions can be set to simulate various operational conditions. This process creates a multi-use dataset.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

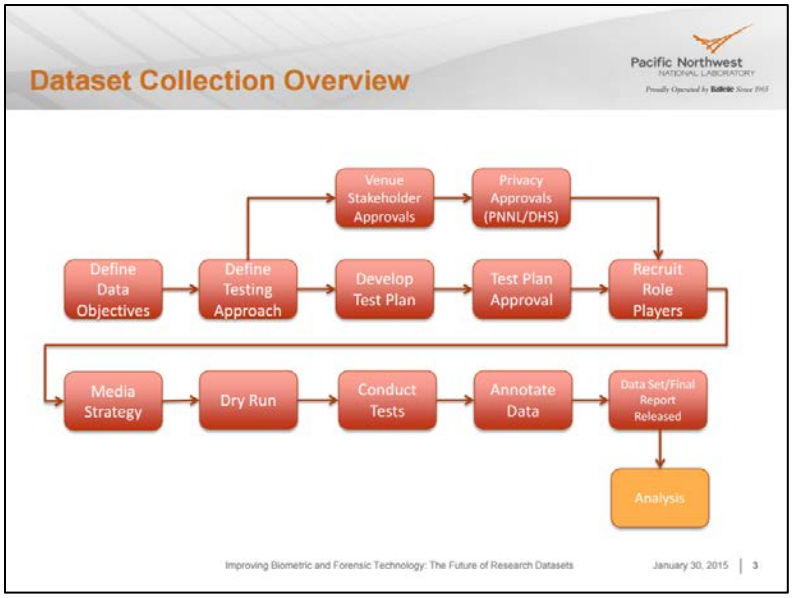


Figure 11-2: Dataset collection overview.

Three scenarios, defined with the cooperation of the stakeholders, have been simulated and datasets created for each of them. To manage the privacy issues, PNNL held sessions with an IRB early in the project to smooth the approval process. Different approvals were required for each scenario, which consumed significant time and resources. Some of the IRB requirements were developed with clear instructions and permissions; this helped expedite the approval process. Scenario parameters that were controlled include:

- Use of multiple cameras
- Variable camera heights
- Variable field of view (FOV)
- Collect extra data
- Control background conditions
- Capture mugshots at multiple angles
- Collect candid images
- Annotate extra information as time allows

PNNL has created a realistic corpus of video data to support the evaluation and enhancement of facial recognition systems. The dataset simulates “non-cooperative behavior.” Dry runs were found to be very useful. In one of the databases, PNNL role players interact with public crowds in five indoor live events in three flow patterns:

- One-way crowd flow
- Two-way crowd flow
- Linear and serpentine queues

PNNL is now building a threat video dataset called Resilient Buildings. It is intended to be used in developing video analytics systems and for human operator evaluation for enhanced building security during threat events. This scenario has up to 12 role players, including an active shooter (single and multiple weapons and persons). Actions simulated include:

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Unauthorized entry
- Unauthorized parking
- Left-behind packages
- Continuous video of background that contains no events
- Sparse video that contains occasional events
- Variations in crowd density, dynamics, and clothing contrast
- Variations in lighting conditions

PNNL has access to a venue that can accommodate a wide range of video dataset needs, including those with live crowds. A repeatable process has been developed to enable a cost-effective collection of video data for forensic video analytic needs. Extensive experience in the annotation of datasets ensures that datasets can be used for benchmarking AND enhancement of video analytics technologies, Kimura and Erickson noted.

Highly annotated resilient building and standoff tracking datasets obtained without active cooperation of the subjects are available for government partners subject to Department of Homeland Security, Science & Technology approval (Figure 11-3).

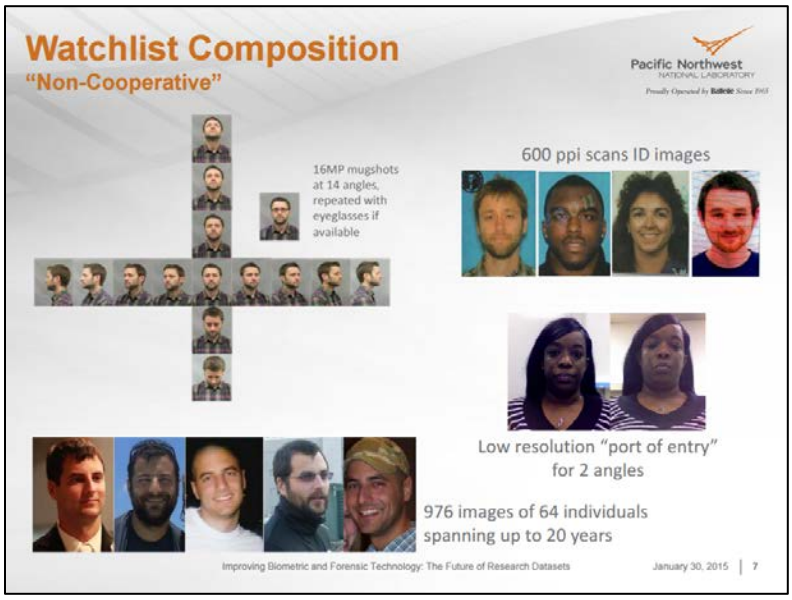


Figure 11-3: Non-cooperative watchlist composition data.

Phillips provided an overview of previous facial dataset collection efforts and the impact of performance evaluation programs on the state of the art. He is now working on the release of a large facial dataset (over 1 million images) collected with Notre Dame University. The size of the database will enable researchers to evaluate the impact of multiple factors on performance. Phillips pointed out that dataset collection is a complex effort that can take a year to complete.

He noted that the comparison of machine versus human performance has emerged as a new field of research. Continuing improvement in machine recognition has enabled machines to outperform humans in some controlled cases. Nonetheless, depending on automated recognition to make correct identifications carries considerable risk; incorrect decisions can result in a bad reputation, commercial losses, security lapses, and can undermine judicial decision validity.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Phillips reiterated the need to test performance using operational data. Data quality is a major determinant of performance. Interpreting results using laboratory data is risky and should be restricted to proof of concept testing.

Many existing databases can test human performance, panelists agreed. However, facial recognition is subject to many variables that must be reflected in the datasets, they cautioned. Researchers need to build experiments that test human performance similar to the Black Box and White Box studies for friction ridge comparisons (see Measuring What Latent Fingerprint Examiners Consider Sufficient Information for Individualization Determinations, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4221158/), These studies should develop metrics that assess the difficulty of comparison. The metrics could also play a role in the triage of comparisons sent for human comparison.

Researchers may want to consider developing a model of how facial features change over time — second to second (expressions) and year to year.

Evaluating facial image and iris data that are captured simultaneously warrants further study.

11.2 Discussion

During the Q&A, Mark Stolorow (NIST) suggested that researchers should explore ways to measure uncertainty. Klare answered that he did not think this would be done any time soon, but Phillips said that progress is being made in this area. While uncertainty measurement is a strong point of DNA comparisons, which are well defined and understood, pattern evidence is not as simple. Unlike DNA, pattern (fingerprints, facial images, etc.) comparison uncertainty is not readily defined. Human performance testing is the key to determining uncertainty for pattern comparisons. Richard Vorder Bruegge called for more funding to test human identification performance.

The following summarizes the recommendations, challenges, and data collection issues identified during the plenary session:

11.2.1 Recommendations

- More datasets are needed in the emerging field of face recognition at a distance using images obtained under unconstrained conditions. Many new datasets have been developed but are not yet publicly available.
- More data is needed that best replicates the intended scenario.

11.2.2 Research Challenges

Several research challenges are expected to impact the creation of new datasets. There is a need to:

- Determine how to qualify a match when video data is used
- Develop and document a ground truth process for scenario testing
- The need to create datasets that address problems that are unique to this modality
- Determine how to address matching of images:
    - captured at different resolutions (impact on identification rate);
    - blurred by motion; and
    - obtained under different means of illumination

11.2.3 Data Collection and Implementation Issues

Datasets need to address problems unique to this modality

- One or more of the images has low resolution
- Some images have motion blur

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Datasets for heterogeneous face recognition (HFR) matching are needed -- matching non-photograph face images (probe images) to large databases of frontal photographs recognition
- Different means of illumination need to be represented in the datasets
- Variety of sensors need to be represented
- Datasets with environmental noise are needed for certain applications
- Data with pose, illumination, and expression (PIE) variations are needed

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

12 Statistical Significance in Biometric and Forensic Datasets

The panel members included Chris Saunders, South Dakota State; Mike Thieme, Novetta; Hariharan Iyer, Statistical Engineering Division, NIST; and Karen Kafadar, University of Virginia. Kafadar moderated the panel and Brad Wing, NIST-ITL, presented in place of Thieme.

12.1 Summary of Presentations

Saunders briefed the workshop on developing statistics for estimating random match probabilities. His talk focused on evidentiary presentation based on his fingerprint and handwriting individuality studies. Saunders described different methodologies for open set and closed set identification. The statistical analysis requires that at least two probability models be developed:

- How each of the specific watch list sources generates data for evaluating the evidence.
- How population from a background (reference data) generates data for evaluating the evidence.

The two methods are radically different, including the definition of error rates and the data and reference data. Different methods are therefore needed in each case to answer questions about the use of the data.

Wing presented an overview of International Organization for Standards (ISO) standards for biometric performance testing and reporting. The briefing addressed the ISO/IEC 19795-N set of standards, which provides dataset guidance for scenario and operational testing. The standards also provide guidance on the type of datasets required to test emerging modalities and the use of datasets comprised of or including artifacts (e.g., fake fingerprints). A standard for video surveillance applications is under development.

The Forensics Science Standards Board (FSSB) has provided greater recognition of the role of statisticians, giving them a stronger presence, Wing noted. Involving statisticians in data collection and reporting will also improve the data collection and analysis processes. ISO also recognizes the strong need for statistical analysis.

Iyer provided an overview of the “Rule of 3” and “Rule of 30,” noting that he had never heard of either before investigating them. He found that the rules yielded results that closely resembled those obtained using classical formulas. He explained when these rules can be used and the expected error in their application. Results of the estimation were compared with results obtained using exact formulae.

Kafadar’s briefing, “Purposes of Databases,” described the theoretical basis of datasets and their proper application. She cited numerous misuses of statistics, using examples of bullet identification and mistakes made during the anthrax scare. She covered the following topics:

- Database development requirements
- Validation requirements
- Implement methods: Reference Database (Exemplars)
- Information sharing
- Identify shortcomings
- Improve methods

Kafadar cited the need for data that is both representative of the target population and complete with respect to the expected characteristics to be analyzed. Statisticians need to be involved in the process so that the important statistical measures relating to uncertainty, experimental design, and selection of representative data are properly executed, she said.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

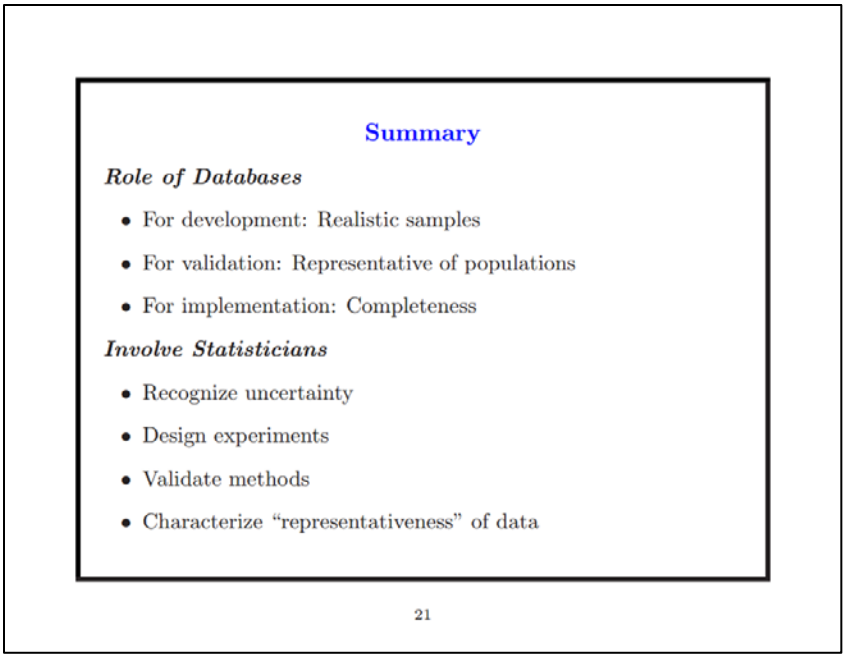


Figure 12-1: Summary of presentation on purposes of databases.

12.2 Discussion

The following summarizes the recommendations, challenges, and various data collection issues identified during the plenary session:

12.2.1 Recommendations

- Identify potential gaps in existing performance standards.

12.2.2 Research Challenges

Several research challenges are expected to impact the creation of new datasets. These include the need for:

- Defining guidelines for statistical analysis for forensic and evidentiary applications.
- Developing statistics on reference datasets in order to make forensic judgments

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

13 Findings

The specific recommendations, research challenges and data collection issues identified during the workshop sessions were collated and analyzed. Workshop discussions revealed agreement on a number of key issues:

Finding 1: Establish an Institutional Review Board (IRB) Toolkit Working Group

A variety of legal issues are involved with the collection, storage, and dissemination of biometric and forensic data for use in research. Workshop members widely acknowledged the difficulties in obtaining approvals for data collection due to the complicated and frequently unclear process for protecting the privacy of individuals and their personal data. To facilitate data collection, all interested parties need to understand the process and issues and arrive at a common interpretation and consensus.

General guidance and streamlining of the IRB process is needed for all stakeholders (Federal, state, and local law agencies; academic and research institutions; as well as corporations).

Specific guidance is needed to understand and implement the policies defined by Common Rule. One issue is that legal counsel for different organizations may provide inconsistent interpretations about how to apply the Common Rule to biometric and forensic data. The problem is compounded when data collected by one organization is shared with another.

Finding 2: Enhance and Maintain the Biometric and Forensic Research Database Catalog

NIST has completed the initial Biometric and Forensic Research Database Catalog. Taxonomy for the Database Catalog should be published and a plan to maintain the Catalog should be created.

- Develop and publish Database Catalog Taxonomy - The taxonomy used in the Database Catalog should be published to encourage its further use, adoption, and revision.
- Establish a plan to enhance and maintain the Database Catalog. A comprehensive plan to maintain the Catalog is needed to identify additional datasets for inclusion, to review and update existing listings, and to include additional modalities.

Finding 3: Identify and Develop Datasets for Rapid Public Release

Workshop participants repeatedly cited the need for more and larger publicly available datasets for use in algorithm development. Collection of new datasets is a time consuming and expensive process. It needs a two-pronged approach:

1. Identify and develop datasets for rapid public release. This effort focuses on identifying the types of data and subjects that could have expedited approvals, dataset collection processes that can be expedited, or datasets that are readily available. Because the collection of new public datasets, is a major undertaking, this task seeks ways to expedite the process by:
    a. Identifying types of data or subjects or collection processes that could have expedited approvals (e.g., at the Workshop, several speakers suggested that datasets from deceased subjects may have fewer privacy concerns associated with the approval process). This effort should be done in collaboration with the IRB Toolkit Working Group.
    b. Investigating whether any Federal/state/local agencies already have datasets that they would be interested in making publicly available.
    c. Investigating which organizations would have legal approval to share data with NIST.
    d. Coordinating with other organizations to make data public (e.g. encourage universities working with local/state law enforcement agencies to make datasets public).

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

e. Determining the greatest needs for new collections and identifying those that may take longest to develop.

2. A separate effort should focus on data about unusual physical conditions (pathological cases). Some biometric modalities include unusual physical conditions that are difficult to process using the standard assumptions for the modality. “Keyhole” irises and scars or dysplasia in friction ridge skin are two examples. These special cases might be handled using special processing but these conditions are not encountered often enough to develop specific algorithms.

Finding 4: Coordinate Data Gap Analysis

Because the development of biometric and forensic datasets is so costly, a data gap analysis should be conducted before a major new dataset is created. The analysis should identify the key areas of research, development, and system acquisition that are hindered by lack of data.

In recognition of emerging technologies and applications and the length of time necessary for data collection, long-range needs should be identified. Data collection coordinated by NIST or NIJ should prioritize and balance broad-based needs against the costs of collection. For example, the lack of public (or large sequestered) latent palm print datasets is impeding the advancement of an entire modality, whereas the lack of datasets for a new acquisition device might be only of niche interest to a few vendors or researchers.

A comprehensive strategic plan for research across the multiple modalities would be desirable but was beyond the scope of the Workshop. Instead we propose that a person or small team coordinate, request, and collate information from a variety of sources to identify needs and priorities for data among these modalities:

- Face
- Person at a distance
- Friction ridge (finger and palm)
- Iris
- Voice
- Handwriting

The attendees identified a large number of datasets in their possession as well as additional dataset needs. (These are identified in the Issues and Challenges summary for each of the sessions). The gap analysis should address how data collection could be shared among users to minimize collection costs. All of the datasets should be categorized using the taxonomy in the Data Catalog.

Finding 5: Develop Best Practices in the Collection, Dissemination, and Use of Biometric and Forensic Datasets

The Workshop produced many recommendations about data collection practices, pitfalls, and distribution issues. These have been compiled in a document, Best Practices in the Collection, Dissemination, and Use of Biometric and Forensic Datasets (to be published shortly).

The process of collecting biometric and forensic datasets is complex. To collect a dataset that is appropriate for use in testing or research, a variety of factors should be considered. The Best Practices document discusses issues that arise in the collection, dissemination, and use of biometric and forensic datasets and provides recommendations. The types of problems — and recommendations — discussed generally fall into these categories:

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

- Collection: Problems include ad hoc or sloppy data collection; biased procedures; inadequate quality control; removing or filtering data. These can be addressed through planning, quality assurance procedures, and documentation.
- Documentation: Inadequate documentation means that the consumer cannot fully understand the data and limitations. This can be addressed by verifying that datasets and any revisions are accompanied by complete documentation. Standardized ways of characterizing datasets (e.g. taxonomy and metadata) may be of benefit.
- Use: Not all datasets are appropriate for a given use. Care must be taken to make certain that datasets are not misused or misrepresented, and that conclusions are not overstated.
- Complying with standards: Make sure all data collection meets regulatory requirements along with applicable standards.
- Resources: Include provisions for ongoing data collection process

Finding 6: Develop Sequestered Data Testing Plan

The sparse amount of data for performance testing can be mitigated by using sequestered data handled by trusted and authorized testing laboratories. This practice allows algorithm developers to submit their technology to the testing laboratory for evaluation without making the data public. This model is used by NIST, where algorithms are submitted for testing on large holdings of protected, government-owned biometric data. To maintain relevance, the data must be continually updated to reflect new modalities, sources of information, new collection technologies, and applications.

To provide algorithm developers with easy access to test results and timely evaluations of revised algorithms, these tests should be available on an ongoing basis, similar to the distribution of NIST's Proprietary Fingerprint Template (PFT) Evaluations. The PFT evaluations have been invaluable to both the developers and users of fingerprint matching algorithms, in part because vendors can submit algorithms at any time. This strategy allows the developers, their users, and their competitors a clear and up-to-date understanding of current capabilities. In the ideal situation, test results would be automated, so that algorithms could be tested in accordance with the developers' schedule and needs.

Additional funding is needed to support expanding repositories of sequestered data and evaluate new technologies.

Finding 7: Develop a Path Forward

Implementing the various recommendations and suggested actions listed above will require a plan. The Path Forward will provide a set of prioritized steps for their implementation. It will also identify potential early successes and the necessary prerequisites for their implementation.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

14 References

14.1 Presentations at the workshop (with hyperlinks):

1. Mark Greene, NIJ, *The Role of Research Datasets in Improving Biometric and Forensic Technology*
2. Austin Hicklin, Noblis, Issues in the Collection and Use of Biometric and Forensic Datasets
3. Brad Wing (NIST) and Austin Hicklin, *More about Biometric Datasets*
4. Mike Garris, NIST, Protect Subject's Privacy, workshop briefing
5. Shannan Williams, NIST, Biometric and Forensics Research Database Catalog, Project Briefing
6. Jeremy Dawson, WVU, WVU Biometric Data Collection Projects
7. Dorothy Glancy, Santa Clara University, Biometric Research Databases and the Privacy Act of 1974
8. Anil Jain, Michigan State University, Fingerprints and Palmprints: Research Datasets
9. Craig Watson, NIST, NIST Fingerprint Data
10. Elham Tabassi, NIST, Usefulness of Current Biometric Datasets
11. Chris Saunders, South Dakota State University, Introduction to Forensic Source ID Problems
12. Mike Thieme, IBG, Discussions of Datasets in Biometric Standards
13. Hariharan Iyer, NIST, Improving Biometric and Forensic Technology: The Future of Research Datasets
14. Karen Kafadar, UVA, The Role of Databases in Forensic Science
15. Stephanie Schuckers, Clarkson University, 2015 Iris—State of Industry
16. George W. Quinn, NIST, Usefulness of Existing Iris Databases and Future Priorities
17. Arun Ross, Michigan State University, Iris as a Forensic Modality: The Path Forward
18. Peter Higgins, Higgins & Associates, Speaker Datasets for Research or FRE 702(b), Daubert and the Problem of Mismatched Conditions
19. Ted Burkes, FBI, State of the Industry Handwriting
20. Christopher Saunders, South Dakota State University, Scientific Approaches to Statistical Analysis and Collection of Handwriting Databases
21. Brendan Klare, Noblis, Datasets for Face Recognition at a Distance (FRAD)
22. Marcia Kimura and Rebecca Erikson and Nick Lombardo, Pacific Northwest National Laboratories, Visual Analytics Dataset Generation

14.2 Related documents:

1. Biometric and Forensic Research Database Catalog at https://tsapps.nist.gov/bdbc
2. Home page for NIST Forensics, Forensic Science @ NIST Portal
3. Home page for NIST Biometrics Portal, Biometrics Portal
4. NIST Biometric Special Databases and Software , Related reference
5. Participant consent sample form from Cornell University, at http://www.irb.cornell.edu/forms/sample.htm
6. Sample database use form (License Agreement) from Notre Dame University at http://www3.nd.edu/~cvrl/LicenseAgreements/UNDLicenseAgreementF.pdf
7. Office of the Privacy Commissioner of Canada reference document at www.priv.gc.ca

14.3 Speaker Biographies

Sue Ballou, Manager, Forensic Science Research Program, NIST, Special Programs Office

Susan Ballou is the program manager for forensic science research in the Special Programs Office (SPO) at NIST. Since 2000, she has managed this program, which targets the needs of the forensic science practitioner by identifying and funding research at NIST in such areas as latent print analysis, burn patterns, computer forensics, and material standards. Her forensic crime laboratory experience spans over 27 years and includes working on case samples in the areas of toxicology, illicit drugs, serology,

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

hairs, fibers, and DNA. She is a member of several forensic science associations. She holds fellow status with the American Academy of Forensic Science (AAFS), was past chair of the criminalistics section, had the opportunity to represent this section on the AAFS Board of Directors, and is a recipient of the AAFS Criminalistics Section's Mary E. Cowan Outstanding Service Award. She accepted the offer to join the Forensic Sciences Foundation Board of Trustees and chairs its Student Travel Grant Award Subcommittee.

Mark Burge, Program Manager, Office of Smart Collection, Intelligence Advanced Research Projects Activity, Office of the Director of National Intelligence

Dr. Mark Burge is a program manager at the Intelligence Advanced Research Projects Activity in the Office of the Director of National Intelligence, where he currently leads the Janus and BEST programs. His broad focus areas are computer vision and pattern recognition, especially as applied to biometrics. He previously served at the National Science Foundation as a program manager and as a research scientist at the Swiss Federal Institute of Science, the Johannes Kepler University in Austria, and the MITRE Corporation. His publications include the three-volume set, "Principles of Digital Image Processing," which has been translated into both German and Chinese, and "The Handbook for Iris Recognition." He holds a Visiting Professor appointment at the United States Naval Academy.

Ted Burkes, Questioned Documents Unit, Federal Bureau of Investigation

Ted Burkes works in the Questioned Documents Unit of the Federal Bureau of Investigation (FBI). He obtained his B.S. in Criminal Justice from the University of Southern Mississippi. He began training in Forensic Document Examination in 1991 at Mississippi Crime Lab at Jackson, MS and joined FBI Questioned Documents Unit in 1998. He is a member of QD section of the American Academy of Forensic Sciences and currently holds office as vice president of the Southeastern Association of Forensic Document Examiners. He also serves as Chairman of Scientific Working Group for Forensic Document Examination, member of the Organization of Scientific Area Committees (OSAC) Questioned Documents subcommittee, and diplomate of the American Board of Forensic Document Examiners.

Joanne Buscalgia, FBI

Dr. JoAnn Buscaglia is a Research Chemist with the FBI Laboratory in the Counterterrorism & Forensic Science Research Unit. Her research focus includes microscopy and microanalysis of trace evidence, interpretation of analytical data using statistical methods, assessment of the scientific basis for the premises underlying the impression and pattern evidence disciplines, and the evaluation of the accuracy and reliability of examiners' decisions. Prior to joining the FBI Laboratory in 1998, Dr. Buscaglia worked for 10 years in academia and as a consultant for both private- and public-sector forensic, environmental, and industrial hygiene laboratories. In 2013, she was honored for her scientific contributions to forensic science with the Paul L. Kirk Award, the highest honor given by the Criminalistics Section of the American Academy of Forensic Sciences.

Jeremy Dawson, Assistant Research Professor, West Virginia University

Jeremy M. Dawson, Ph.D. received B.S. M.S. and Ph.D. degrees from the Lane Department of Computer Science and Electrical Engineering (LCSEE) in 1997, 1999, and 2002 respectively. Dr. Dawson returned to West Virginia University (WVU) in 2007, and is currently a Research Assistant Professor within LCSEE. He has over 10 years of research experience in the fields of integrated optics and nanophotonics gained in both academic and industrial research arenas, including advanced research projects that involve the

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

design, fabrication, and integration of complex, multi-domain sensor systems and devices. He is currently the principal investigator and project leader on DNA signal processing and photonic crystal device development efforts undertaken within the Center for Identification Technology Research (CITeR) and WVNano. He is also involved in research involving emerging biometric sensor technologies and new device architectures for solid state lighting.

Rebecca Erikson, Applied Physicist, Pacific Northwest National Laboratories, Department of Energy

Rebecca Erikson is a senior scientist at the Department of Energy's (DOE) Pacific Northwest National Laboratory (PNNL) with a significant focus in test and evaluation of systems for national security applications. She has been principal investigator in the testing and evaluation of visual analytic systems for DHS and the Department of Defense (DoD). She has been involved with the design, execution and ground-truthing of three large datasets targeting applications including automated human tracking at a distance, non-cooperative facial recognition, and identifying security threats from behavioral and physical markers. She also has experience evaluating algorithms and systems against performance metrics and standards.

Michael Garris, Biometrics Senior Scientist, ITL, NIST

Michael Garris is a biometrics senior scientist at the NIST where he has worked for the past 29 years. In his career, Mr. Garris has primarily worked in the area of image processing and pattern recognition including biometrics and optical character / handwriting recognition. He was privileged to manage for 7 years the world-class research, test, and evaluation Image Group in the Information Technology Laboratory (ITL) focused on promoting and advancing biometrics standards and technologies at NIST. Mr. Garris continues to provide senior leadership and coordination within the Image Group. His focus includes standardization and adoption of new contactless fingerprint technologies, for which he recently established a Contactless Fingerprint Capture Device Measurement Research CRADA Program facilitating open collaboration between NIST and industry. He has a BS in Computer Science from Clarion University of Pennsylvania, and a MS in Computer Science from Johns Hopkins. In 2003, Mr. Garris was part of a biometrics team which received the Department of Commerce Gold Medal Award.

Dorothy Glancy, Santa Clara University, Professor of Law

Dorothy Glancy is Professor of Law at Santa Clara University School of Law. A graduate of Wellesley College (B.A.) and Harvard Law School (J.D.), Professor Glancy studies law and technology issues related to privacy and the rights of individuals. A member the California and the District of Columbia bars, she is a Life Member of the American Law Institute. She serves as a Research Fellow of the Gruter Institute for Law and Behavioral Research and is a member of the California Judicial Council's Court Technology Advisory Committee and of the OSAC Speaker Recognition Subcommittee. Her interest in Big Data includes service on the NAS Federal Advisory Committee that recommended a structure for long-term research use of the 4 terabyte SHRP2 naturalistic driver behavior database.

Mark Greene, Division Director, Office of Science and Technology, NIJ

Mark Greene is the Policy, Standards, and Grants Management Division Director in the Office of Science and Technology at NIJ. His unit covers issues of technology policy, standards and conformity assessment for equipment used by criminal justice practitioners, and oversee administration of extramural grants and cooperative agreements. Before this, Mark was a program manager at NIJ where he managed research,

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

development, test, and evaluation projects in emerging technology for criminal justice applications, such as facial recognition, video analytics, and noncontact fingerprint devices. The program also included the NIJ Sensors, Surveillance, and Biometrics Center of Excellence which focused in part on testing and evaluation of completed R&D projects supported by NIJ in biometrics, sensors, and surveillance technologies. He holds a Doctorate and Bachelor degree in Materials Science and Engineering and a Bachelor degree in Communications from Northwestern University.

Patrick Grother, Computer Scientist, ITL, NIST

Patrick Grother is a staff scientist at NIST responsible for biometric standards development, algorithm testing and analysis. He currently serves as editor of four ISO standards in the areas of interchange of facial imagery, performance testing and modular fusion processes. He leads the ongoing Minutia Interoperability Exchange (MINEX), Iris Exchange (IREX) and Multiple Biometric Evaluation - Still image (MBE-STILL) evaluations of fingerprint, iris, and face recognition technologies at NIST. These support biometrics in national scale identity management. He was the recipient, in 2003 and 2007, of Department of Commerce Gold Medals for, respectively, tests of border management technologies and for the specification of biometric data elements and procedures needed for the US Government's Personal Identity Verification program. He advises US Government agencies on a number of biometric projects.

Austin Hicklin, Research Scientist, Noblis

Austin Hicklin is a Fellow at Noblis and has been involved in a broad range of biometric/forensic projects for various government agencies since 1995, including biometric and forensic standards, evaluations of latent print examiners, AFIS engineering and interoperability, video analytics, fingerprint quality metrics, and evaluations of biometric identification systems. He designed and developed the Universal Latent Workstation for the FBI. He serves on the OSAC Forensic Science Standards Board, and is the chair of the Scientific Area Committee for Physics and Pattern Evidence. He has a B.A. from the University of Virginia, a M.S. from Virginia Tech, and is completing a doctorate in Forensic Science from the University of Lausanne.

Peter Higgins, Higgins & Associates

Peter Higgins retired in 1995 from the FBI as the Deputy Assistant Director in charge of the IAFIS program. He operates Higgins & Associates, International, a small-by-design biometric consultancy, which has supported 14 US states and the Royal Canadian Mounted Police (RCMP) in the procurement of the AFIS and the DoD in procuring biometric systems for the governments of Iraq and Afghanistan. Mr. Higgins led a small team in modeling the Next Generation Identification (NGI) system for the FBI and served on the UK Biometrics Assurance Group. Mr. Higgins has been appointed to the OSAC's Speaker Recognition Subcommittee, having served as the SWG-Speaker previously.

Hariharan Iyer, Mathematical Statistician, Statistical Engineering Division, ITL, NIST

Dr. Iyer is a Mathematical Statistician in the Statistical Engineering Division at NIST. He also is a fellow of the American Statistical Association and is Professor Emeritus in the Department of Statistics at Colorado State University. His research areas include statistical modeling, design of experiments, statistical inference, machine learning applications in science, forensic statistics, and general problem solving.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Anil Jain, University Distinguished Professor, Department of Computer Science and Engineering, Michigan State University

Anil K. Jain is a University Distinguished Professor in the Department of Computer Science and Engineering at Michigan State University. His research interests include pattern recognition, computer vision, and biometric authentication. He has received Guggenheim, Alexander von Humboldt, and Fulbright awards. He served as a member of the Defense Science Board, National Research Council (NRC) studies on Whither Biometrics and Improvised Explosive Devices, and currently serves on the Forensic Science Standards Board.

Karen Kafadar, Professor of Law, UVA

Karen Kafadar is Commonwealth Professor & Chair of Statistics at University of Virginia. She received her Ph.D. in Statistics from Princeton University, and previously held positions at NIST, Hewlett Packard's RF/Microwave R&D Division, the University of Colorado-Denver, and Indiana University. Her research focuses on robust methods, exploratory data analysis, characterization of uncertainty in the physical, chemical, biological, and engineering sciences, and methodology for the analysis of screening trials. She has served on several National Academy of Science committees, including the ones that led to the reports Weighing Bullet Lead Evidence (2004), Strengthening Forensic Science in the United States: A Path Forward (2009), Evaluating Testing, Costs, and Benefits of Advanced Spectroscopic Portals (2011), and Identifying the Culprit: Assessing Eyewitness Identification (2014). She has served on the governing board of several prominent organizations is a member of the newly established OSAC Forensic Science Standards Board. She is a Fellow of American Statistical Association (ASA), American Association for the Advancement of Science (AAAS), and International Statistics Institute (ISI), and has authored over 100 journal articles and book chapters, and has advised numerous M.S. and Ph.D. students.

Marcia Kimura, Pacific Northwest National Laboratories, Department of Energy

Marcia Kimura is an engineer on the process engineering team at DOE's Pacific Northwest National Laboratory. She has been a researcher at PNNL since 2008 and is a certified systems engineer. Ms. Kimura has managed and participated in the testing and evaluation of visual analytics systems for DHS and DoD. She directed the design, execution, and ground-truthing of three large datasets targeting applications including automated human tracking at a distance, non-cooperative facial recognition, and identifying security threats from behavioral and physical markers.

Brendan Klare, (Former) Research Scientist, Noblis

Brendan Klare is a co-founder of Rank One Computing. Prior to this, Brendan founded the Visual Analytics Lab at Noblis. Brendan has published over 25 peer reviewed papers on the topic of automated face recognition. Brendan's professional interests are in accurate algorithm design and developing novel applications of face recognition technology. Brendan received his Ph.D. in Computer Science from Michigan State University.

Greg Michaud, Michigan State Police

Captain Greg Michaud is in his third year as Director of the Michigan State Police Forensic Science Division. He leads one of the nation's largest publically funded laboratory systems with eight regional laboratories and over 240 personnel. Forensic science services are provided to approximately 600 law enforcement agencies across Michigan at the local, state, and federal levels. During his tenure, he has

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

been instrumental in establishing a long-standing partnership with Michigan State University, entering into several research opportunities that include face recognition, fingerprint matching, biometric fusion, and sexual assault evidence collection.

Jonathon Phillips, Research Scientist, NIST-ITL

Dr. Jonathon Phillips is a leading technologist in the fields of computer vision, biometrics, and face recognition. He is at National Institute of Standards and Technology (NIST), where he runs challenge problems and evaluations to advance biometric technology. From 2000-2004, Dr. Phillips was assigned to DARPA. In an Essential Science Indicators analysis of face recognition publication over the past decade, Jonathon's work ranks at #2 by total citations and #1 by cites per paper. He was awarded the Dept. of Commerce Gold Medal. His work has been reported in the New York Times, BBC Radio 4, and the Economist. He has appeared on NPR's Science Friday show. He is a fellow of the Institute of Electrical and Electronics Engineers (IEEE) and International Association for Pattern Recognition (IAPR). In 2013, he won the inaugural Mark Everingham Prize.

George Quinn, Research Engineer, NIST

Research Interests

Mechanical testing. Reliability analysis. Fractography. Standard test methods.

Awards and Honors

ASTM Committee C-28 "Advanced Ceramic" Award, 2007

Fellow, American Ceramic Society, 2001

U. S. Department of Commerce Bronze Medal, 2000

ASTM Certificates of Appreciation, (5), 1997-1999

Fellow, American Society of Testing and Materials, 1994

American Society of Testing and Materials Award of Merit, 1994

William J. Alcott Award (Best overall student in Northeastern University), 1973

Standards Outputs:

ASTM C 1683 (2008) "Standard Practice for Size Scaling of Tensile Strengths Using Weibull Statistics for Advanced Ceramics"

ASTM C 1684 (2008) "Standard Test Method for Flexural Strength of Advanced Ceramics at Ambient Temperature – Cylindrical Rod Strength"

ASTM C 1323 (2009 rev) "Standard Test Method for Ultimate Strength of Advanced Ceramics With Diametrally Compressed C-Ring Specimens at Ambient Temperature"

Arun Ross, Michigan State University Associate Professor in the Department of Computer Science and Engineering and Director of the i-PRoBe Lab

Arun Ross is an Associate Professor in the Department of Computer Science and Engineering at Michigan State University. He was in the faculty of West Virginia University from 2003 to 2012, and served as the Assistant Site Director of the NSF CITeR between 2010 and 2012. He is the coauthor of the books "Introduction to Biometrics" and "Handbook of Multibiometrics". Arun is a recipient of the JK Aggarwal

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Prize, the International Association for Pattern Recognition Young Biometrics Investigator Award, the NSF CAREER Award, and was designated a Kavli Frontier Fellow by the National Academy of Sciences in 2006.

Chris Saunders, Assistant Professor of Mathematics and Statistics, South Dakota State University

Christopher Saunders is an assistant professor of Statistics at South Dakota State University. He received his Ph.D. from the University of Kentucky and was awarded an Intelligence Community Post-Doctoral Research Fellowship under which he was trained in forensic evidence interpretation (by Dr. Buscaglia of the FBI Labs) and statistical pattern recognition. Since the completion of his fellowship his research focus has been on the interpretation and presentation of forensic evidence for impression and pattern evidence-specifically handwriting evidence. He has also served as a visiting scientist to the FBI Labs, and on the NIJ standing scientific review panel as well as the American Statistical Association's ad-hoc advisory committee on forensic science led by Professor Kafadar.

Stephanie Schuckers, Professor, Clarkson University

Dr. Stephanie Schuckers is a Professor in the Department of Electrical and Computer Engineering at Clarkson University and serves as the Director of CITeR, a NSF Industry/University Cooperative Research Center. She received her doctoral degree in Electrical Engineering from The University of Michigan. Professor Schuckers research focuses on processing and interpreting signals which arise from the human body. Her work is funded from various sources, including NSF, DHS, and private industry, among others. She has started her own business, and has over 40 journal publications as well as over 60 other academic publications.

Elham Tabassi, Electronics Engineer, ITL, NIST

Elham Tabassi is a researcher at NIST working on various biometric research projects including biometric sample quality, fusion, and performance assessment. She is the principal architect of NIST Fingerprint Image Quality (NFIQ). She is involved in the development of Government, US and International standards for biometric sample quality, data interchange formats for biometric data, and conformance to data interchange formats. She received the Department of Commerce Gold Medal in 2003 for her work on biometric system performance assessment, and the Department of Commerce Bronze Medal in 2007 for her work on biometric sample quality. Her research interests are in biometrics algorithms, biometric sample quality, biometric fusion, pattern recognition, data mining, and signal processing.

Melissa Taylor, Senior Forensic Science Research Manager, Special Programs Office, NIST

Melissa Taylor is a senior forensic science research manager with the Special Programs Office at the National Institute of Standards and Technology. Her work within the Forensic Science Research Program focuses primarily on improving forensic science management practices and integrating human-factors principles into impression and pattern evidence disciplines. Ms. Taylor has more than 10 years of scientific and program management experience including positions with Booz Allen Hamilton and Lockheed Martin. She is a member of the American Society for Quality, the Human Factors and Ergonomic Society, and the International Association of Identification. She currently serves as a member of the Human Factors Subcommittee of the National Commission on Forensic Science.

Mike Thieme, Vice President, Novetta

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Mr. Michael Thieme currently serves as Director of Special Projects at International Biometric Group, LLC Vice President, Novetta. Mr. Thieme is an authority on biometric testing, system integration, and next-generation identity technologies. Mr. Thieme leads IBG teams delivering integrated biometric systems and conducting biometric research for customers including DoD, DHS, TSA, White House OSTP, California DMV, NIJ, and dozens of high-tech firms and systems integrators. He is the Chair of INCITS M1.5 and editor of US and ISO biometric performance testing standards. Mr. Thieme is the primary author of IBG's performance test reports and Biometric Market and Industry Report series.

Peter Tytell, Private Forensic Document Examiner

Peter Tytell has been a private practice document examiner in New York City for over forty years. He trained with his parents, Pearl and Martin, and took relevant courses at Georgetown, the Institute of Paper Chemistry, Rochester Institute of Technology, as well as many other specialized seminars and workshops. He has made numerous presentations and has taught and lectured in the US and other countries. He is a diplomate of the American Board of Forensic Document Examiners, a member of the American Society of Questioned Document Examiners, and the Questioned Document section of the American Academy of Forensic Sciences; he is a recipient of the section's Ordway Hilton Award.

Richard Vorder-Bruegge, Senior Photographic Technologist, FBI Laboratory Division

Richard W. Vorder Bruegge is a senior photographic technologist at the Federal Bureau of Investigation where he is responsible for overseeing science and technology developments in the imaging sciences. He has an Sc.B. in engineering, and an Sc.M. and Ph.D. in geological sciences from Brown University. He has been with the FBI since 1995, where he has performed forensic analysis of image and video evidence, testifying in state, federal and international courts as an expert witness over 60 times. Vorder Bruegge was chair of the Scientific Working Group on Imaging Technology from 2000 to 2006 and chair of the Facial Identification Scientific Working Group from 2009 to the present. He is a fellow of the American Academy of Forensic Sciences in the Digital and Multimedia Sciences Section. In 2010 he was named a Director of National Intelligence Science and Technology Fellow for his work in facial recognition. In addition to several publications on forensic image analysis, he has also co-authored multiple peer-reviewed articles on facial recognition and identification and was co-editor of Computer-Aided Forensic Facial Comparison (CRC Press, 2010).He was appointed to the FSSB in 2014.

Craig Watson, IT Specialist, NIST

Craig Watson is currently an IT specialist at the National Institute of Standards and Technology (NIST). He has invested 24 years with NIST, and currently serves as the Project Lead for the Image Group's biometrics research lab. During his time with NIST, he has been engaged in numerous initiatives to improve fingerprint biometric technology. This includes extensive work to develop fingerprint testing datasets, fingerprint image compression, and slap fingerprint segmentation. He has expertise with biometric technology evaluations, including the Fingerprint Vendor Technology Evaluation (FpVTE), Proprietary Fingerprint Testing (PFT), Minutiae Interoperability Exchange Test (MINEX), and Slap Fingerprint Segmentation Evaluation. He has been part of a team that received a Department of Commerce Gold Medal Award (2003) and a Department of Commerce Silver Medal Award (2014) for their work related to biometric technologies.

Jim Wayman, Research Administrator in the Office of Graduate Studies and Research, San Jose State University

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

Dr. James L. Wayman Biography Dr. Jim Wayman is a research administrator in the Office of Graduate Studies and Research at San Jose State University He received the Ph.D. degree in engineering in 1980 from the University of California, Santa Barbara. In the 1980s, under contract to the U.S. DoD, he invented and developed a biometric authentication technology based on the acoustic resonances of the human head. He joined San Jose State University in 1995 to direct the Biometric Identification Research Program, serving as Director of the U.S. National Biometric Test Center at San Jose State from 1997-2000. He is co-editor of J.Wayman, A. Jain, D. Maltoni and D.Maio (eds) Biometric Systems (Springer, London, 2005), a Fellow of the British Institution of Engineering and Technology, U.K. "Head of Delegation Working Group 1" on the ISO/IEC JTC1 SC37 standards committee on biometrics, a "core member" of the U.K. Biometrics Working Group, and a member Biometrics Experts Committee of the U.K. Home Office. He is a member of the U.S. National Academies of Science/National Research Council Committee "Whither Biometrics?", the NRC Panel on Information Technology, and previously served on the NRC "Authentication Technologies and their Implications for Privacy" committee. He is an honorary Professor of Electronics at the University of Kent, holds 4 patents in speech processing and has served as a paid biometrics advisor to 8 national governments.

Shannan Williams, Project Manager, Forensic Science Research Program, SPO, NIST

Shannan Williams is a project manager in the Forensic Science Research Program in the Office of Special Programs at NIST. She manages research projects in two new forensic research focus areas at NIST: drugs/toxins and trace evidence analysis. Prior to officially joining NIST in 2015, Shannan was a management consultant providing facilitation, organizational strategy, outreach, and change management support to various forensic science related initiatives occurring at NIST for over 5 years. In her support of the *Technical Working Group on Biological Evidence Preservation*, she led and significantly contributed to three published reports providing multidisciplinary best practices and guidance to the forensic science community. She is a member of ASTM E30 and an affiliate member of the Quality Infrastructure Committee of the Organization for Scientific Area Committees. Shannan graduated *magna cum laude* from the University of Maryland Baltimore County and obtained her Master of Public Policy from Harvard University.

Bradford Wing, (Former) Biometrics Standards Coordinator, ITL, NIST

Brad Wing has just retired from the National Institute of Standards and Technology (NIST), where he was the Biometrics Standards Coordinator and head of the NIST-ITL standards development organization, as well as being the chair of the Interagency Forum on Biometrics and Identity Management. Prior to joining NIST, Mr. Wing worked as Chief Biometrics Engineer at the US Visitor and Immigrant Status Indicator Technology (US-VISIT) Program of DHS, where he founded the DHS Biometrics Coordination Group, which brings together the various components of DHS to ensure a consistent approach to biometrics across the Department. He served as editor of the American National Standards Institute (ANSI) / NIST-ITL standard “Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information and co-Chair of the Interagency Forum on Biometrics and Identity Management. He currently serves as Secretariat for the AAFS Standards Board.

Henry Wixon, Office of Chief Counsel, NIST

Henry Wixon became Chief Counsel for the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce in March 2009. Prior to joining NIST, he was a partner with the law firm of Wilmer Cutler Pickering Hale and Dorr, LLP, in Washington, D.C. In more than 20 years in private practice as a registered patent attorney and intellectual property specialist, Mr. Wixon developed and implemented

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

worldwide patent defense and enforcement strategies for domestic and international clients, coordinating the efforts of counsel in multiple European and Asian countries. In addition to extensive patent litigation, he conducted intellectual property due diligence and valuation analyses in connection with venture capital funding, mergers and acquisitions, and initial public offerings. He counseled research consortia on intellectual property matters, including the drafting and negotiation of consortia agreements, patent landscape analyses, and identification of essential patent claims. With extensive experience in technology transfer, Mr. Wixon worked cooperatively with and between private and public laboratories, universities, corporations and government agencies, including the negotiation of patent and know-how licenses, Cooperative Research and Development Agreements, employment contracts, and confidentiality agreements. Mr. Wixon received his J.D. from George Washington University, where he was a member of The George Washington Law Review. He holds a B.S. in zoology with a minor in chemistry and an M.S. in neuropharmacology from the University of Maryland, College Park.

14.4 Participant List

The following is a list of registered attendees who indicated interest in being added to a public participant list.

	Registrant Name	Organization
1	Kristen Allen	Noblis
2	William Arnold	Houston Forensic Science Center
3	John Barnett	U.S. DHS, HSI Forensic Laboratory
4	Edward Bartick	George Washington University
5	Mike Beachler	Noblis
6	Kevin Bowyer	University of Notre Dame
7	Michael Brockly	FBI
8	John Butler	NIST
9	Leonard Butt	Maryland State Police
10	Jon Stacy Byrd	Ron Smith & Associates, Inc.
11	Stephen Dunbar	
12	Gregory Dutton	National Institute of Justice
13	Rebecca Erikson	Pacific Northwest National Laboratory
14	Michael Floersh	FBI
15	William Ford	US DOJ/NIJ
16	Donald Gantz	George Mason University
17	Samantha Heise	Collaborative Testing Services, Inc.
18	Peter Higgins	Higgins & Associates, International
19	Kodiak Hill-Davis	
20	Robert Michael Mccabe	IDTP
21	Nicole Mccleary	RTI International
22	Patrick Mcconlogue	GoodData
23	John Morgan	RTI

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156

24	Kristin Murrock	Washington DC Department Of Forensic Sciences
25	Larry Nadel	Accenture
26	Lina Nardecchia	Noblis
27	Danica Ommen	South Dakota State University
28	Lance Orr	Scitor
29	David Plocher	U. S. Government Accountability Office
30	Daniel Rappaport	CACI, Inc.
31	Jeri Ropero Miller	RTI International
32	Eileen Sexton	ImageWare Systems, Inc.
33	Edgar Ronald Smith	Ron Smith & Associates, Inc.
34	Mark Stolorow	NIST
35	Janet Temko-Blinder	U.S. Government Accountability Office
36	Laura Tierney	US Immigration and Customs Enforcement
37	Ryan Triplett	
38	Peter Tytell	Forensic Research, Llc
39	Richard Vorder Bruegge	FBI
40	Mark Walch	Sciometrics
41	Garold Warner	Defense Forensic Science Center (Anser, Inc.)
42	Victor Weedn	George Washington Univ Dept Of Forensic Sciences
43	James Yen	NIST

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8156