Measuring Strength of Identity Proofing

Workshop: Applying Measurement Science to the Identity Ecosystem
Version: 1, December 16, 2015

Information Technology Laboratory, NIST

1 INTRODUCTION

This document serves as a primer for discussions to be held at the "Advanced Identity Workshop" at the National Institute of Standards and Technology (NIST) in Gaithersburg, Maryland, on January 12 and 13, 2016. That workshop will convene federal agencies, commercial relying parties, and identity solution providers to collaborate on improving standards, guidance, and practices related to identity management.

1.1 PURPOSE

Identity proofing is used to establish the uniqueness and validity of an individual's identity to facilitate the provision of an entitlement or service, and may rely upon various factors such as identity documents, biographic information, biometric information, and knowledge of personally-relevant information or events. With the increased ability of malicious actors to gain access to data that was once considered private, the efficacy of proofing methods that rely on private data declines. Simultaneously, the emergence of technologies such as biometric sensors and high-resolution cameras in smartphones provides an opportunity for evolved identity proofing methods.

This document introduces some of the technical, policy, and implementation considerations associated with the development of a schema by which federal agencies and other organizations could measure the strength of identity proofing methods to improve risk-based decision-making.

Specifically, this document explores the following topics:

1. The current approaches to identity proofing and the tradeoffs associated with those approaches,
2. New proofing techniques currently not addressed by existing guidance, and
3. Determination of whether "strength of proofing" can be evaluated, normalized, asserted, and consumed by relying parties in a way that appropriately aligns with the risk tolerance of accepting proofed individuals.

1.2 SCOPE

This document and the forthcoming workshop will focus on new approaches to apply measurement science to the evaluation of identity solution performance.

NIST is responsible for establishing requirements that can be used to proof individuals who will access online government services. Federal agencies can accept identity proofing assertions from non-government sources, provided an appropriate federation and trust framework authority, such as the Federal Identity, Credential and Access Management (FICAM) Trust Framework Solutions (TFS), can provide assurances that individuals are appropriately proofed before they are bound to

the credential that they use to access an online transaction. For this reason, this document focuses on a framework that can serve both government and commercial solution providers.

NIST anticipates that the end products resulting from these efforts could be applied to existing guidance, such as Special Publication 800-63-2, Electronic Authentication Guideline.[1] In addition, scores and thresholds can be used within trust framework definitions and by relying parties to make real-time access control decisions about a remote user.

Finally, this effort is focused on those use cases in which the relying party has a business requirement for some level of knowledge of an individual's physical identity. This is only a subset of digital transactions; many applications exist in both the public and private sectors that should be absent identity proofing, maintaining pseudonymity or anonymity for the user. However, for those use cases that need some level of identity proofing, improved measurement science can facilitate relying parties' trust in that proofing.

2 WORKSHOP FOCUS AREAS

The content that follows introduces considerations for a performance measurement framework for identity proofing. NIST requests that readers consider the following questions as they review this document and prepare for January's workshop:

Scoring Identity Proofing Processes:

1. Should identity resolution specifics, such as the set of attributes required to unique resolve to an identity, as well as the target percentile, be considered in the overall score?
2. Do datasets exist that can be used to evaluate proofing outcomes?
3. Do open-source or publicly available risk analytic datasets exist that could continuously inform the efficacy of a proofing process, and allow for real-time adjustments?
4. Is there a framework that can be used to determine the ability for knowledge-based questions to be guessed correctly? Or inversely, not be able to be guessed?
5. Can false positives and false negatives be adequately tested and measured such that they are viable performance metrics for identity proofing?
6. Due to the ability for knowledge-based approaches to be vulnerable to data compromise or guess-ability, should they be excluded as valid?
7. Should each major element of the proofing process be scored and conveyed, or should an aggregate be the norm?
8. Does the score of an individual proofing transaction matter if the designed process has a score and is adhered to?
9. Should outcomes such as civil or criminal penalties be included in strength measurements?

Potential Impacts to Trust and Interoperability:

1. If a process scores lower than the asserted designed range, is that indicative of failure or just a downgrade in identity assurance level?

[1] National Institute of Standards and Technology, Special Publication (SP) 800-63-2, *Electronic Authentication Guideline*. August 2013. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63-2.pdf.

2. Will relying parties accept values within a range, or will they demand the highest score for each individual?
3. Will relying parties segment a system based on an individual's score? In other words, would a relying party to provide less access than an individual with a higher score, even if the score remains within designed range and applicable assurance level?
4. Will granularity diminish the value of a finite set of assurance levels, therefore undermining any possibility for true interoperability and federated identity?

3 METHODS OF IDENTITY PROOFING

Generally, there are two options for identity proofing: in-person and remote (performed over an online, networked session). These seemingly simple approaches vary significantly, often in nuanced ways that can strengthen or weaken the expected proofing results. Moreover, the market is blurring the distinctions; in-person processes often conduct digital checks while remote processes sometimes employ human checks.

Identity proofing to support the issuance of physical credentials, such as drivers' licenses and passports, is an established field that is typically based on an in-person registration event where "breeder documents" are: Presented, verified as authentic, and were legitimately issued to the person claiming the documents belong to them. Identity proofing is also an essential element for those online transactions that require association of a user's online persona with his or her real-world identity. While not necessary for all online applications, consistently proving that an individual is who they claim to be remains a focal point in efforts to limit impersonation and ensure that only the correct users receive services and entitlements.

Consumers and implementers of identity proofing technologies, such as federal and state agencies and commercial relying parties, have choices regarding the selection and configuration of available solutions. There is a lack of normalized metrics available to determine the efficacy of such processes and configurations and to compare the effects of configuration choices of specific solutions to each other. This white paper explores mechanisms to quantify and evaluate the strength of identity proofing processes and configurations. By properly understanding the strength of an identity proofing process, relying parties (RPs) can make more informed decisions about the risk of accepting a digital identity to access their service.

3.1 IN-PERSON IDENTITY PROOFING

Users conduct in-person identity proofing when they are required to present themselves and their documentation directly to a person. Historically, in the absence of strong remote identity proofing techniques, in-person proofing has been required for proofing associated with high-risk transactions such as obtaining a government-issued identity credential, gaining access to trusted traveler programs, or applying for a large financial loan. While in-person proofing is often considered the 'gold-standard' for proofing, it is hampered by criticisms that it is too costly for providers and too inconvenient for individuals. Although in-person proofing is generally considered straightforward and secure, several factors may influence its efficacy, such as the:

- Level of training for manual and visual validation and verification of presented documentation,

- Availability of fraudulent or counterfeit documents to attackers,
- Complexities of understanding and detecting the nuances of many valid documentation types issued by a wide range of jurisdictional elements,
- Use of automated validation and verification of documentation by specialized proprietary equipment, and
- Training of the person verifying the proofing evidence.

3.2 REMOTE IDENTITY PROOFING

Remote identity proofing is used when the user is not expected to present themselves or their documents at a physical location. Remote proofing is executed online, and traditionally involves validating and verifying presented data against one or more corroborating authoritative sources of data. The amount and types of confirmed data influences the performance of remote proofing systems and commonly includes attributes that can uniquely resolve a person in a given population, such as name, address, date of birth, and place of birth, as well data considered private such as account numbers or the amount of a loan or tax payment. Some instantiations of remote proofing also include a virtual session where a user may digitally present documents for verification.

Increasingly, malicious actors can easily obtain data once considered private via mechanisms such as credit reports or targeted online searches. In this current environment, remote proofers can reduce risk by assuming that knowledge of any given datum, even if intended to remain private, has been compromised.

Whether in-person or remote proofing is used, organizations that implement identity proofing generally seek to balance cost, convenience, and security for both the provider and the individual. Examples of these tradeoffs include:

- Reducing the complexity of a remote proofing experience to improve the online experience can result in an increased risk of false acceptance,
- Increased complexity to reduce false acceptance can result in increased abandonment and false rejections rates that are unacceptable to some service providers, and
- Users that do not share characteristics with the expected user population (e.g., national origin, country of residency) can lead to persons that are unable to complete proofing.

4 MEASURING STRENGTH IN IDENTITY PROOFING

The lack of measurement science around identity proofing serves as an impediment to the market. The difficulty of measuring the performance of different solutions limits organizations' ability to determine if they are properly mitigating risk. This uncertainty in performance reduces both providers' incentive to innovate strong proofing solutions and RPs incentive to accept identities proofed by other entities. Unable to reliably measure the degree of risk mitigation, the market has difficulty developing models to efficiently assign liability. Standardizing measurements for identity proofing strength could convert uncertainty to measurable risk, properly aligning incentives on both the demand and supply side of identity proofing solutions. This will allow organizations to mitigate risk in identity proofing more effectively as part of their overall effort to enhance the system security and usability. To that end, improving measurement science in identity proofing will

achieve another step toward the identity ecosystem envisioned in the National Strategy for Trusted Identities in Cyberspace (NSTIC). [2]

Grant recipients under the NSTIC Pilots Cooperative Agreement Program identified the need for performance standards, specifically in remote identity proofing schemes that leverage knowledge-based authentication (KBA), or for the purposes of this document, Knowledge-Based Proofing (KBP). They encountered challenges in identifying standard configurations to meet minimum proofing requirements, and had difficulty determining important metrics such as false rejections, false acceptances, and failures to enroll. Some pilots realized low acceptance rates to enroll, but had a limited basis for determining whether the solution effectively mitigated the risk of fraudulently proofed identities making it through the system, or indicated a poor user experience. Additionally, no indicators suggested whether successful individuals were of the claimed identity or were a successful fraudster.[3]

4.1 Measuring Performance

Identity proofing related guidance from NIST allows for significant flexibility in the processes implemented by the entity responsible for proofing (the proofer). This is by design, as the proofer should have room to innovate and offer comparable processes that are marketable and effective for the population they serve. In addition, specifying an overly prescriptive set of requirements reduces the ability for relying parties to adjust for differing risk tolerance.

This flexibility has tradeoffs, including a lack of interoperability, and interpretive variants in proofing rigor that can lead to vulnerabilities. With a variety of approaches, an RP needs to make determinations as to what approaches perform sufficiently well to meet its needs. Measurements of performance have thus far proved elusive as proofers have trouble quantifying the outcomes of slight—or significant—variations in processes. Consider a hypothetical scenario in which one agency performs in-person proofing with trained individuals inspecting source documentation, while another performs in-person proofing with document inspection conducted by both trained individuals and specialized equipment. It stands to reason that the addition of the specialized equipment reduces the risk of false acceptance, but such a judgment is difficult to quantify.

4.2 Establishing a Measurement Framework

No standard mechanism exists to assert strength of a designed and implemented proofing process, or the result of an individual completing a proofing process. Defining a performance score could follow one of two approaches: building the score from the ground up by combining scores for the underlying components or determining the degradation of a "perfect" system by analyzing the types of threats to which it would be subject. Several approaches could be used to develop these methods. The score could consist of a scalar value that provides reliable information on the performance of a proofing process, based on variations in approaches. Without mature sets of performance standards for different processes, the measurability of any given process might initially exist as a set of ordinal values or intervals. NIST Special Publication 800-63-2, for instance,

---

[2] Executive Office of the President, *National Strategy for Trusted Identities in Cyberspace*. April 2011. https://www.whitehouse.gov/sites/default/files/rss_viewer/NSTICstrategy_041511.pdf.

[3] National Institute of Standards and Technology, *NSTIC Pilots: Catalyzing the Identity Ecosystem*. April 2015. http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8054.pdf.

describes processes by which an identity proofer could attain each of four ordinal values called levels of assurance (LOA). For any measurement approach, however, users can typically derive more information from scalars and, where necessary, can map scalar values back to ordinal values, such as a specific LOA.

Based on the desire to build a tractable scoring framework based on performance that allows for the combination of different technologies to improve performance, this document and the forthcoming workshop will focus on developing a scalar scoring framework, understanding that specific communities of interest may map that scalar framework into an ordinal approach that meets their needs. Specifically, NIST will proceed with this work in hopes of developing a scalar-based framework for identity proofing performance, and then map those scores to LOAs that comport with OMB Memorandum 04-04, *E-Authentication Guidance for Federal Agencies*.[4]

As a starting point for establishing this framework, NIST believes it would be useful to consider each function that supports identity proofing, such as: identity resolution; identity validation; identity verification, and activity history. Table 1 provides examples of scoring approaches within the framework for each function.[5,6]

Table 1: Identity proofing functions and sample scoring factors.

Function	Expected Outcome	Possible Process Steps
Identity Resolution	A set of identity factors that uniquely resolve an individual within a stated population	Attribute Selection
Identity Validation	Assurance that the presented identity information is valid	Self-asserted data collection Physical document data collection Data Validation
Identity Verification	Confirm valid identity is associated with the correct individual. Confirm existence of claimed identity over a period of time	Manual Verification Automated Verification Historical Activity Verification Out-of-Band Verification

Aggregating each function into a total proofing score could be misleading, as one function may score very high whereas another function could score very low, therefore weakening the entire process. NIST believes that, if individual functions are interrogated for score, the lowest function score would drive the overall acceptance, based on risk, of the identity proofing an individual underwent. This approach mitigates the possibility that a designed process degrades the overall score by placing too much emphasis on a given function. This approach is currently used in the United Kingdom, per *Good Practice Guide Number 45, Identity Proofing and Verification of an Individual*.

Tables 2-4 detail possible considerations, variances, and complexities with applying a score to proofing functions and the individual processes an organization can select to complete each function.

---

[4] Office of Management and Budget, OMB Memorandum M-04-04, *E-Authentication Guidance for Federal Agencies*. December 16, 2003. https://www.whitehouse.gov/sites/default/files/omb/memoranda/fy04/m04-04.pdf.

[5] Treasury Board of Canada, *Directive on Identity Management*. August 1, 2011. http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=16577.

[6] Cabinet Office, *Good Practice Guide No. 45: Identity Proofing and Verification of an Individual*. July 2014. https://www.gov.uk/government/publications/identity-proofing-and-verification-of-an-individual.

Table 2: Identity resolution elements and scoring considerations.

Function 1: Identity Resolution	
Elements of Proofing Functions	**Scoring Considerations**
Attribute Selection	• Scoring this function could ensure that organizations take the appropriate measures to establish the baseline attributes required to obtain the highest possible percentage of coverage. Too much variance in the required attributes could translate to vulnerabilities in the overall process and offer a window for falsely proofing fraudulent identities.

Table 3: Identity validation elements and scoring considerations.

Function 2: Identity Validation	
Elements of Proofing Functions	**Scoring Considerations**
Overall Considerations	• The source of identity data, or evidence, could be considered part of the score for identity validation. For example, a score could factor identity evidence that was issued by an entity bound by regulation, such as Know Your Customer in the United States,[7] by a government authority with or without digital signatures verifying the source, or some other formal document with identifying information. • Additional consideration is the strength of the proofing process to obtain the identity evidence. It is possible that the documents used to prove identity have a weak trust anchor and could reduce the rigor of proofing.
Self-asserted data collection	• Self-asserted data might only be used where no assurance of identity is required. Yet, in remote proofing schemes for higher LOAs, self-asserted data typically initiates the transaction. While this may be a form of KBP in a strict sense, in that providing an identifier on a government-issued ID or of a financial account equates to knowledge of ownership, some question the market tolerance for this type of proofing in today's market. • Furthermore, the digital and information ages have tradeoffs, and today's private and protected data could become tomorrow's public data. Shifting from requesting social security number (which is cheaply and almost universally attainable on black markets) to driver's license number (which is somewhat less attainable and typically revocable) will only drive criminals to obtain a massive cache of DL's in order to spoof an identity. Such a move is, at best, a stopgap. • If self-asserted data is acceptable, should the asserted attributes, especially if they are sourced from a personal data store, carry attribute metadata so the relying party or proofer can determine how much confidence they have in the data element? • Self-asserted data, or any form of KBP, suffers from an inability to effectively test for false-positives, false-negatives, and other similar metrics. Without a proven testing methodology, there is no way to test "knowledge". An organization cannot predict what an individual may or may not know about themselves. Worse yet, organizations cannot test for the ability for the right or wrong person to guess the correct answer. And in any scenario when data are aggregated, even in the form of KBP quizzes, it is difficult to determine what information an individual (good or bad) can glean from one question, to correctly answer another.

[7] Public Law 107–56. October 26, 2001. http://www.gpo.gov/fdsys/pkg/PLAW-107publ56/pdf/PLAW-107publ56.pdf.

Function 2: Identity Validation	
Elements of Proofing Functions	**Scoring Considerations**
Physical Document Data Collection	• Presenting and collecting information from physical documents has long been used for in-person proofing, and is now becoming available for remote proofing by inspecting documents using video or photos taken with users' cameras. To date, guidance is limited in how a proofer should validate authenticity of the document and the data asserted upon the document. Strength of validation for physical document can range from: visual inspection; the use of trained staff or specialized equipment; and knowledge of proprietary methods to fabricate and secure source documents. Some solutions may only be able to validate that the encoded bar code matches human readable text. • Security features can also be detected and validated, but lighting conditions or resolution of equipment can vary. Is better lighting or a higher resolution camera any better than commodity equipment on modern computing platforms, or from visual inspection of a trained workforce? Where does scoring need to stop, as the residual benefit from additional checks is not worth the cost, time, and potential privacy implications of viewing more information than is necessary? • In addition, as physical documents begin to carry cryptographic modules, validation of signatures and encrypted data can be used to validate asserted information. While this document considers a range of possibilities acceptable to proofing without passing judgment on how to conduct checks that are more intrusive, proofers must also take into consideration users' comfort with an approach. • In making a determination that cryptographic validation is equal to, or greater than, equipment that validates every physical security feature of a document, what empirical data must be collected, tested, weighted, and scored to make that determination? Is the presence of a private cryptographic key more important than the presence of physical security? To determine this, an approach must take account of potential vulnerabilities in the supply chain used to establish these security features. How protected is the private key? Who owns it? Who issued it? Is the material to produce a document difficult to come by? Are the security features resistant to reverse engineering if a fraudster inspected an authentic document? • While this section has waded into the electronic possibilities available today, there exist trusted 3rd parties, such as notaries, that are bound by law, standards of care, and business relationships to validate identity. How should NIST consider comparing these processes to anything done in the digital domain? • Finally, much as the introduction of multiple-factors of authentication increases security, or multi-modal biometrics increase the ability for distinctness, how many documents should be checked to attain confidence? Are there diminishing returns in the number of checked documents? What if in the valid range of allowable documents, some had more security features than another?
Data Validation	• Online validation can occur in both remote and in-person proofing processes, and could be considered subordinate to how the data that is being validated was captured. Did the data come from a physical document that underwent every physical and electronic security validation step, or was it self-asserted? • While it may be more secure to validate against authoritative databases rather than aggregated datasets, if the integrity of the data collection is questioned, does dataset "authoritative-ness" matter? This implies that attribute metadata may be valuable in the process of validating claimed identity data. • In addition, exact matching of the supplied data to online repositories is often not possible. Given that, how is matching strength calculated? Fuzzy matching algorithms are typically proprietary and may require human in-the-loop to validate a match or set of matches. Should the number of attributes matched be considered? For example, if five individual datum need to be matched, can all data be fuzzy matched? Should a minimum be exact match? How many can be allowed to not match? If authoritative-ness matters, must each datum be matched to an authoritative source, or can some matches occur in aggregated, non-authoritative datasets?

Table 4: Identity verification elements and scoring considerations.

Function 3: Identity Verification	
Elements of Proofing Functions	**Scoring Considerations**
Manual Verification	• Historical verification of identity is typically done via comparison of the picture on a photo ID to the live person claiming ownership of the ID and the information contained therein. Other forms of manual verification could include one person vouching for another, specifically if the "voucher" was proofed by the organization prior to them vouching for someone else. • Manual verification becomes increasingly more complicated based on the number of photos an individual must compare. Variations in appearance and photographic conditions can be significant enough where it is difficult to visually determine if the IDs correlate to the same person, let alone that person presenting the IDs is the same. Regardless of the strength of validation, should additional controls be applied to manual verification, such as legal implications to the "voucher" should the "vouchee" create harm in the system? If so, harm must be well defined, as the "vouchee" may be performing fraudulent transactions as themselves, meaning the "voucher" did not do anything wrong. A person can vouch that someone is who they say they are, but should not be required to be accountable should that valid identity do something bad in the system. • Also, with commodity camera and voice equipment on personal computing devices, manual verification is not necessarily assumed to occur in-person. Virtual sessions can be established and leveraged to perform manual verification. The vulnerabilities and environmental considerations with a virtual session should be considered when scoring verification performed over this channel. How would this affect confidence in a proofing process?

Automated Verification	• The predominant method online today is to perform KBP. This should not be confused with the validation of biographic data submitted by the claimant to records that the organization may have access to. KBP is intended to confidently bind the real identity to the claimed data. This document has identified some deficiencies in the KBP scheme and posits its appropriateness for identity proofing in today's environment. However, should KBP remain as a viable proofing alternative at higher assurance levels, what elements of KBP can be evaluated and scored? Since KBP is typically proprietary, to include the risk engines that underpin the aggregated identity data, and the subsequent quiz that a user participates in, it is difficult to surmise a normalized set of criteria that can be set to achieve a particular strength, or trustworthiness, in KBP results. Examples of variables in a strength metric could include, among others:
	  ◦ Number of questions presented and answered correctly
	  ◦ Multiple choice or freeform answers
	  ◦ Time to complete quiz
	  ◦ Number of failed attempts
	  ◦ Number of diversionary questions
	  ◦ Dissonance among questions
	  ◦ Authoritative-ness of the data the KBP is performed against
	• Biometric verification against documents that have local biometrics stored onboard may be effective to bind the identity to the claimed data, as is verification against a central store of biometric data. This assumes the individual was proofed by the organization that issued the document containing a template or maintains the template centrally. For this method to be adequate, the organization that is proofing the individual again, based on antecedent proofing to capture the biometrics, has determined that the prior proofing event occurred in a manner consistent with their requirements. This should be included in the proofing strength score, and if the initial proofing transaction is weaker than organizational requirements allow, additional controls should be put in place to increase the trustworthiness of the proofing event.

Function 3: Identity Verification	
Elements of Proofing Functions	**Scoring Considerations**
	• Finally, proving account ownership by validating activity can be used to perform identity verification. Similar, if not the same, to KBP, this verification technique uses knowledge of transaction history to verify the individual. This technique has precedent, as some organizations test ownership of bank accounts by depositing a random small amount, confirming the individual can assert the correct amount, and then backing out the transaction. Organizations would need additional information to properly score this approach. For example, how far back in time should transactions be validated? How recent transactions are verified to avoid the risk that a fraudulent identity has just occurred (e.g., someone steals a credit card, performs 3 transactions, and then attempts to prove their identity based on those transactions)? Should account activity such as last call to helpdesk or last postmark of mailed correspondence be checked? And similar to biometrics, the overall assurance level of the account being used to verify should be in alignment with proofing requirements, since it is possible for accounts protected with weak credentials have been breached.
Out-of-band verification (also known as "Closing The Loop")	• Performing round-trip verification of an identity is typically not required with in-person verification, but is necessary with remote proofing to provide reasonable protection against impersonation. Existing alternatives, such as sending postal mail, email, or SMS to an address of record assist organizations in verifying they proofed the correct individual. Potentially not a factor in scoring, but a benefit of this factor nonetheless, is the ability to alert someone that they have been proofed prior to providing access. • Each potential channel available to close the loop will have vulnerabilities, and should factor into scoring this element. For example, SMS has weakness when sharing secrets. Email is readily phish-able and commonly protected with weak credentials. Postal mail takes time and is vulnerable to theft. Any out-of-band channel that satisfies this element needs to prevent, to the best extent possible, a scalable attack. While postal mail has vulnerabilities, attacks scale poorly—if at all—when compared to email. In addition, scoring should consider the ability for others to have access to the selected channel. Spouses may share an email address, while postal mail is available (regardless of addressee) to any member of the household.
Historical Activity Verification	• Many organizations validate the existence of an identity over time. Credit bureaus and data aggregators look for timeline anomalies to determine if identities are fraudulent or synthetically generated. However, much of this information, and the associated risk models, are proprietary and challenging to compare. While activity history may add value to proofing confidence, questions remain as to the overall efficacy of activity history. • Over how long of a period of time does evidence of existence provide meaningful information to proofing? • Should activity history be evaluated before or after initial proofing is complete? • Does account activity, like financial transactions or help-center call logs, instill additional confidence or do they fall into the same category as KBP, and are therefore limited in establishing additional proofing confidence? • Do social networks provide valuable insight into an identity by allowing individuals to vouch for one another? • Building off social and online reputation, do signals about account activity provide value, particularly whether any of an individual's accounts are associated with fraudulent behavior? • What other fraud indicators should be used to identify historical anomalies? • Device authentication, behavioral patterns, and environmental context are often used by organizations performing adaptive authentication based on risk. Can and should this data be part of the proofing process? • The privacy implications of some of the concepts are significant and should not be ignored. But if implemented and accepted, is scoring of these techniques beneficial and possible? If individuals opt-in to these services, can they be shared securely and effectively in a vast ecosystem, rather than being stove-piped or offered by a single social provider?

4.3 Scoring Individual Identity Proofing Transactions

It is possible that variances can be introduced during an individual transaction that could be scored as well. Table 5 presents a hypothetical use case of how a score could vary depending on individual transactions, which assumes for illustrative purposes that the score for each function is totaled into a final score.

Table 5: Hypothetical identity proofing processes and performance scores.

	Process	Hypothetical Score (max 100)
Process Definition	In-Person; two forms of government issued picture ID, maximum of one ID can be expired within one year; IDs electronically validated; data verified with issuing entity	70-85
Specific instantiations of the defined process	Alice arrives with a Real ID compliant state license and a U.S. State Department issued e-passport. Proofer has validated all security controls of the state license as well as verified the data against the issuing DMV. All physical security elements of the e-passport are validated and the information is verified via cryptographic checks.	85
	Parvarti arrives with a state driver's license issued by a state that has not implemented Real ID. She also has an expired passport that does not have modern electronic enhancements. The proofer can validate the physical security features of both forms of ID but must validate driver's license information through a data aggregation service.	75
	Adnan arrives with a Real ID compliant state license, a valid green card issued by USCIS and a valid passport from the United Kingdom. The proofer can validate all security features of the driver's license and performs data verification via the issuing state. The green card and foreign passport are visually inspected, data on these documents match the information verified on the license.	80
	Jacob arrives with a state driver's license and an expired Transportation Worker Identification Credential (TWIC) card. Some of the physical security features can be verified but the card has been damaged. Jacob also changed his address but does not have the DMV issued change of address card. Records check with a credit bureau validates the remaining biographic information. The TWIC card cannot be electronically validated but the picture and on-card biographic information matches records.	70

As seen in Table 5, many variants can be introduced into a seemingly simple process. The generic process allows a proofing entity to assert its ability to achieve a certain threshold of proofing strength, which would likely map to an assurance level. However, each individual that underwent the actual proofing transaction would be scored significantly differently based on allowable variations of the process.

5 Conclusion

Identity proofing is a complex process with significant implementation variance. This is compounded by an ever-increasing threat landscape associated with both physical source documentation and the availability of historically private data. However, as a trust anchor to secure online transactions, and ideally, a process that an individual need only go through very few times,

industry and government alike need improved processes for understanding the strength of identity proofing processes.

The introduction of strength of identity proofing scores may have secondary effects, such as the restriction of services for users that do not achieve the appropriate score, or a limitation of interoperability between identity ecosystem participants based on the score granularity. Regardless, NIST believes that the current level of uncertainty in the performance of identity proofing solutions necessitates collaboration in public and private fora to develop a quantified process to determine the efficacy of identity proofing processes, such that relying parties can fully understand their system risk. NIST hopes this approach will also provide additional incentive for new and continually evolving innovations that allow more individuals access to services, and enable RPs, including the federal government, to adopt these innovations more quickly.

This paper will seed one of three topics during January's Advanced Identity Workshop. NIST intends to leverage the outputs to produce a report. Such a report would address the considerations of applying a scoring framework for identity proofing and a proposed implementation approach. The content of that report relies on feedback to this paper and discussion of this topic at the January workshop.