

ID 2020

Program Monitoring and Evaluation *Pilot Metric Development*

The ID2020 Alliance is committed to rapidly scaling up portable, privacy-protecting and user-centric digital identity, and recognizes that meaningful evidence and learning is significant to realizing that aim.

In order to generate the requisite base of knowledge, it is critical that a set of common learning objectives, metrics and monitoring practices be woven into the design and implementation of digital identity pilots. If well selected and well implemented, these common learning objectives will generate the program data needed to compare learnings, facilitate scalability and promote accountability.

To this effect, ID2020 has identified three overarching assessment areas:

1. Effectiveness, efficiency and impact of pilot projects
2. Replicability and scalability of the digital identity programs
3. Privacy and data protection of the digital identity solutions

With the support of the Rockefeller Foundation, ID2020 partnered with IDinsight to identify metrics which would capture feasible, actionable and generalizable data on digital identity programs. This partnership resulted in the following report which is intended to not only guide internal decisions but also assist external stakeholders with designing insightful monitoring and evaluation protocols.





ID2020 Pilot Metric Development

Recommendations

Prepared for the ID2020 Alliance

September 6, 2018

CONTENTS

| | |
|---|----|
| Executive Summary | 2 |
| Context | 3 |
| Engagement Objectives | 3 |
| Methods | 3 |
| Recommendations | 4 |
| Looking Ahead..... | 16 |
| Appendix A: Detailed Theory of Change | 17 |
| Appendix B: Stakeholders | 18 |
| Appendix C: Interview Guide..... | 19 |
| Appendix D: Metric Grading Definitions..... | 22 |
| Appendix E: Excluded Metrics..... | 23 |
| Appendix F: Interview Insights | 25 |
| Appendix G: Data Quality Guidelines | 27 |
| Appendix H: Metrics Tables | 29 |
| Appendix I: Secondary metrics..... | 36 |

EXECUTIVE SUMMARY

Background

ID2020 is a global partnership of governments, non-governmental organizations (NGOs), and the private sector committed to improving lives through digital identity (ID). The partnership coordinates and channels funds toward high impact projects. ID2020 is currently identifying promising digital ID pilots to receive catalytic funding.

ID2020 partnered with IDinsight to establish a set of universal metrics across ID2020-supported pilots that will support learning and contribute to global evidence in the digital ID sector. More broadly, these metrics will enable a range of actors to build responsible digital ID systems globally.

Engagement Overview

To establish recommended metrics, IDinsight conducted 26 expert interviews across 23 stakeholders involving governments, NGOs, the private sector, and funding organizations. These interviews focused on stakeholder priorities, activities, and capacities relating to digital ID and measurement. Interviewee input, combined with internal measurement and digital ID expertise was used to create a bank of potential metrics. Each metric was then ranked according to six criteria (feasible, actionable, generalizable, accurate, precise, and perceived importance) to generate a list of candidate metrics. Universal metrics were selected from among this list to minimize data collection burden on pilots and facilitate meaningful comparisons.

Recommendations

To maximize learning from the pilot phase of ID2020's strategic roadmap, we recommend the following:

- 1. Conduct a pre-pilot exercise with implementers to gain a deeper understanding of pilots' initial program design.** This pre-pilot exercise can generate a deeper understanding of a pilot's initial program design, its available capacity at the start of the pilot, and other factors that are likely to signal the pilot's potential for success and alignment with ID2020's vision of a high-impact digital ID system.
- 2. Collect seven universal metrics from each ID2020 pilot to facilitate cross-pilot comparisons.** All pilots should be required to collect and report on the universal metrics to maximize learning across pilots.
- 3. Support data collection for 11 user-based universal metrics from each pilot.** User-based metrics should be collected directly from end-users of the digital ID system. Therefore, ID2020 should only require these metrics from pilots if ID2020 is willing to support data collection.
- 4. Work with pilots to determine the feasibility and applicability of collecting up to 12 secondary metrics.** Secondary metrics will complement universal metrics to allow ID2020 to make a more comprehensive assessment about the likelihood of a given pilot's success. They will not necessarily be used to compare across pilots.
- 5. Select a subset of pilots for deep measurement and evaluation to generate more rigorous evidence for the digital ID sector.** ID2020 has an opportunity to generate deeper evidence for the digital ID sector through impact evaluations of a subset of pilots that represent high potential use cases.

CONTEXT

ID2020 is an alliance that brings together and coordinates funding for stakeholders who are working to improve lives through digital identity (ID). ID2020 is set to begin funding promising pilots working in digital ID to help them scale. This requires a set of universal metrics by which ID2020-funded pilots can assess and improve their processes and make a case for support to scale up. ID2020 will use the metrics for advocacy and program prioritization. More broadly, these metrics will be a useful tool for a range of actors in the sector to help them build healthy digital ID systems globally.

While some stakeholders have begun to develop policies and guidelines for digital ID systems, there is limited guidance as to how digital ID programs can measure and track their fulfilment of these guidelines. Therefore, ID2020 has partnered with IDinsight to interview digital ID experts and define a set of metrics to contribute to cross-pilot monitoring and evaluation frameworks.

ENGAGEMENT OBJECTIVES

This engagement had two objectives:

- Identify measurement priorities and capacities of current digital ID stakeholders.
- Define a set of universal metrics that can be commonly measured and would capture pilots' performance and effectiveness.

METHODS

IDinsight drew from three sources in developing priority metrics: (i) a digital ID pilot theory of change;¹ (ii) stakeholder interviews; and (iii) IDinsight's experiences leading the State of Aadhaar initiative.²

IDinsight and ID2020 built a list of potential interviewees based on ID2020's network and IDinsight's State of Aadhaar project expertise. The key consideration when building this list was to ensure that it represented the different types of actors (funders, governments, technology developers, etc.) as well as different sectors (blockchain, health, financial inclusion, etc.) that currently work in digital ID. We reached out to 38 stakeholders and conducted 26 interviews across 23 organizations and companies.³ The interviews investigated three areas:⁴

- **Current and future state of the digital ID sector.** Themes discussed in this area highlighted existing debates within the digital ID sector where more data could be transformational and inform priority measurement areas and related metrics.

¹ ID2020 provided an initial high-level theory of change that depicted how digital ID systems can lead to improved lives. IDinsight broke that theory of change down into more detailed nodes and pathways to better unpack how digital ID inputs lead to outputs, outcomes, and impact. This detailed theory of change (Appendix A) was used to identify priority nodes and pathways for measurement.

² The State of Aadhaar initiative, hosted by IDinsight, aims to catalyze data-driven discourse and decision-making in the Aadhaar ecosystem. The goals of the initiative include: (1) generating unique insights on Aadhaar through primary and secondary research, (2) catalyzing a data-driven discourse on Aadhaar, and (3) informing policy on the appropriate use of Aadhaar. More information can be found here: <https://stateofaadhaar.in>.

³ The breakdown of the interviews conducted by type of stakeholder and sector can be found Appendix B.

⁴ The full interview guide is in Appendix C.

- **Role that the interviewee’s organization plays in the sector.** The goal of this section was to understand what type of programs are being implemented in order to develop and refine the universal metrics.
- **Monitoring and evaluation capacity of the organization with regards to digital identity.** This section aimed to understand what type of data collection methods would be feasible to recommend for the universal metrics. This section also included questions about the metrics that interviewees perceive to be the most important to collect.

From these interviews, we derived a list of measurement areas and related metrics that emerged as important, addressed a range of issues (i.e. access to services, privacy, user-centricity, etc.), were applicable across all/most sectors, and represented a mix of process and outcome metrics. We then linked each metric to the theory of change and ranked it according to six criteria:

1. **Feasible** – The ease of collecting data for the metric and amount of additional resources that need to be invested for data collection
2. **Actionable** – The likelihood that the metric will inform decision-making
3. **Generalizable** – The relevance of the metric to a wide range of pilots
4. **Accurate** – The metric’s relationship to the measurement construct
5. **Precise** – The reliability and verifiability of the data collection method
6. **Perceived importance** – How often the metric was emphasized in interviews

We graded each proposed metric according to these criteria with a “3” (high), “2” (medium), or “1” (low).⁵ The grading of these criteria was the main tool we used to filter between the recommended metrics (universal, user-based, and secondary) and metrics that we do not recommend to collect.^{6,7}

RECOMMENDATIONS

The primary goal of the ID2020 engagement was to define a focused set of metrics to apply across ID2020 digital pilot sites. These universal metrics were selected to facilitate comparison of pilots to guide ID2020’s strategic decision-making, including funding decisions.

Stakeholder interviews highlighted that the digital ID sector can benefit from more evidence and measurement to unlock deeper understanding of both the potential pathways and magnitude of impact that digital ID systems

⁵ Definitions of these grades are included in Appendix D.

⁶ The full list of metrics that were not selected as potential candidates for universal or secondary metrics can be found in Appendix E.

⁷ In addition to our metric development process, at ID2020’s request, we noted key takeaways from each interview that were then used to generate high-level insights. Through reflection on the key takeaways from each interview, we identified consistencies and contrasts in opinion on key topics, including privacy, security, and interoperability. These insights and related recommendations are included in Appendix F.

have on improving lives through digital identity. Interviews also revealed that digital ID implementers' monitoring and evaluation capacity varied widely.

With this framing in mind, we selected a set of universal metrics from an initial list of candidate metrics according to two principles:

1. **Minimize data collection burden on pilot implementers.** Given the likelihood of limited M&E capacity and data collection capabilities, we recognize that a few data points could translate into a large ask. Therefore, we limited the number of universal metrics and prioritized the “Feasibility” rating. Allowing pilots to focus on collecting a small number of metrics well will likely increase data quality and minimize diversion of capacity and resources away from pilot implementation.⁸
2. **Select metrics that are applicable to a variety of pilots and span critical nodes on the theory of change.** The metrics will best be positioned to facilitate comparisons across pilots if they apply to a variety of pilots. Therefore, we also prioritized the “Generalizability” rating. Further, since there are a limited number of universal metrics, they should cover various nodes on the theory of change to provide a more comprehensive overview of the pilot. The metrics are highlighted on our detailed theory of change in Appendix A.

While the universal metrics will facilitate comparisons across pilots, ID2020 may need additional data to more fully understand the likelihood of success for a given pilot. In many cases, these additional data points will need to be interpreted with the context of the specific pilot in mind. Therefore, we have five recommendations for ID2020 to maximize learning from the pilot phase of their strategic roadmap:

1. Conduct a **pre-pilot exercise** with implementers to gain a deeper understanding of pilots' initial program design
2. Collect **seven universal metrics**⁹ from each ID2020 pilot to facilitate cross-pilot comparisons
3. Support data collection for **11 user-based metrics** from each pilot
4. Work with pilots to determine the feasibility and applicability of collecting **up to 12 secondary metrics**
5. Select a subset of pilots for **deep measurement and evaluation** to generate more rigorous evidence for the digital ID sector

Recommendation 1: Conduct a pre-pilot exercise with implementers to gain a deeper understanding of pilots' initial program design

The universal metrics will help ID2020 assess how pilots are performing, whereas the pre-pilot exercise will generate a deeper understanding of a pilot's initial program design, its available capacity at the start of the pilot, and other factors that are likely to provide evidence of the pilot's potential for success. For any initial pilots that ID2020 has already selected, this pre-pilot exercise is a way to build upon current thinking and gather more information to improve the chances of success. Further, a pre-pilot exercise could help inform which secondary metrics to prioritize for a given pilot (see Recommendation 4). As ID2020 gains better familiarity of which

⁸ Data quality guidelines are included in Appendix G.

⁹ Each metric will not be applicable to every category of pilot. There are five to seven metrics that are applicable to ID providers (depending on if they are linked to specific services) and four metrics that are applicable to service providers.

factors are most correlated with success, this pre-pilot exercise could become a formal part of the pilot application process to inform pilot selection.

The pre-pilot exercise should involve an interview with pilot representatives that focuses on:

- **Definition of the “end user” and enrollment targets.** Pilots should describe the eligibility criteria for the end users of the pilot. They should also provide an initial target number of beneficiaries that they aim to reach over the course of the pilot.
- **Current protocols and processes for recruitment, enrollment, and access of services.** Pilots should describe their envisioned protocols for identifying and enrolling individuals, including consent options and procedures. These protocols should also articulate how users will use the digital ID to access services (if applicable). The data collected during the pilot can then be used to assess if protocols are being followed and processes are occurring according to plan.
- **Characteristics of the digital ID system and linked services.** ID2020 can assess and explore the current and expected levels of interoperability and portability of the digital ID system.
- **Existing capacity and infrastructure to support the pilot.** Pilots should describe their existing personnel, infrastructure, and additional resources that they plan to dedicate to implementation during the pilot.
- **Activities that the pilot or partner organizations will do to strengthen implementation.** Examples include trainings, awareness campaigns, and registration drives.
- **The pilot-specific theory of change.** Pilots should articulate the specific pathways through which they believe they improve lives through digital identity.

Recommendation 2: Collect seven universal metrics from each pilot funded by ID2020 to facilitate cross-pilot comparisons

We recommend that ID2020 require all relevant pilots to collect and report on the universal metrics outlined in this section. Five to seven metrics are applicable to ID providers (depending on whether they are linked to specific services), and four metrics are applicable to service providers.¹⁰ Pilot applicants should outline which metrics are already routinely collected, by whom, how, and with what frequency. For those metrics that are not currently collected, pilot applicants should outline their plan for how they will begin collecting these metrics during the pilot.

We recommend that ID2020 pilot these recommended universal metrics with early pilots to get a better understanding of how feasible these metrics are to collect and the level of learnings that they are able to generate. Our list of universal metrics and type of pilot for which the metrics are relevant are outlined in

¹⁰ “ID provider” refers to a stakeholder that is developing the digital ID technology and program. “Service provider” refers to implementers that are providing digital ID-linked services. Some stakeholders may play both roles.

Table 1 and described in greater detail below.¹¹

¹¹ More detailed tables of the recommended metrics are included in Appendix H.

Table 1. Recommended universal metrics

| Measurement area | Metric | Relevant pilot | |
|--------------------------------------|---|--|------------------|
| | | ID provider | Service provider |
| [1] Feasibility of Adoption/Scale-up | [1a] Projected cost per enrolled user over the span of five years | Yes | Yes |
| [2] Enrollment | [2a] Number of users enrolled in the digital ID system | Yes | No |
| | [2b] Proportion of targeted users that enrolled in the digital ID system during the pilot | Yes | Yes |
| [3] Privacy | [3a] Proportion of non-user initiated usages of data to which users are asked to consent | Yes | No |
| | [3b] Number of personal data points that are accessed by a digital ID-linked service/right that are not needed to provide it | Yes | No |
| [4] Access to Services | [4a] Proportion of people who have successfully used the digital ID platform to access a service during the pilot among those enrolled in the digital ID system | Yes (if linked to a specified service) | Yes |
| | [4b] Proportion of people who have experienced a failed authentication when attempting to access a digital ID-linked service during the pilot | Yes (if linked to a specified service) | Yes |

Metric 1a: Projected cost per enrolled user over the span of five years¹²

ID2020 is interested in the cost-benefit ratios for pilots in the digital ID sector. A pilot’s projected cost per enrolled user will allow ID2020 both to compare the cost-benefit ratios across pilots and to assess the financial sustainability of each pilot. This metric is relevant for all pilots that have direct digital ID costs, though ID2020 should focus on comparisons within a specific category of stakeholders.¹³

While ID2020 could ask pilots to provide initial projections in the pilot application, pilots should collect real cost data to update inputs. Important cost categories include:

- Upfront investment costs (e.g., hardware, design of technology, cost of initial connections to digital ID system, etc.)
- Operational costs for running and maintaining the digital ID system (e.g., additional personnel, maintenance fees, publicity, training, etc.)

¹² Five years was chosen to allow for smoothing of investment costs over time and to capture expected uptake beyond the pilot stage.

¹³ For example, the expected cost-benefit of a digital ID provider who may be developing digital ID technology will likely be different from a service provider who will link to existing digital ID technology.

- Cost-savings as a result of the digital ID system (e.g., reduced time to provide a service, faster identity verification, etc.)
- Revenues as a result of more customers / service users
- Service costs/cost-savings if the cost of service delivery changes as a result of the digital ID¹⁴ (e.g., reduced time to link across service instances, reduced time to gather customer information, etc.)

Enrollment data should be collected from administrative records. Costs and the expected number of users should be modeled over a five-year time horizon. To ease the burden of these calculations and to facilitate standardization across pilots, ID2020 could include a cost-benefit template for pilot applicants.

Metric 2a: Number of users enrolled in the digital ID system

User enrollment is a critical step on the digital ID theory of change. This metric should be used to compare the absolute number of beneficiaries included in the digital ID system across ID provider pilots. It is particularly useful when tracked across time as an indication of the demand for the digital ID as well as the capacity of the pilot to enroll users.

This metric should come from administrative data.¹⁵

Metric 2b: Proportion of targeted users that enrolled in the digital ID system during the pilot

Expanding on the previous metric, this metric presents the proportion of the targeted users who were enrolled in the digital ID system over the course of the pilot. This metric should be used to compare how quickly and effectively pilots enroll users according to expectation.

The numerator – number of users enrolled in the system over the course of the pilot– should be tracked by the number of total users who are enrolled or registered in a digital ID system.

The denominator – number of potential users targeted during the pilot – should be based on the enrollment targets set at the start of the pilot.

Metric 3a: Proportion of non-user initiated usages of data to which users are asked to consent

User consent is an important component of privacy. Users should know how their personal information is being used, and they should have the opportunity to consent (or refuse consent) to this usage. This metric can be used to compare completeness of consent across pilots and, therefore, the degree of transparency that pilots have with their users.

This metric can be self-reported by the ID provider. A single usage of data refers to a specified explicit purpose by a specific stakeholder (i.e. a government agency, a private company, etc.) for which the identified data is used. It could be within the same body or by additional actors / third parties. Consent refers to an action in which the user agrees to each specific usage separately. There may be cases where separate consent is not possible, but that should be publicly justified and documented.

¹⁴ If the digital ID system will not affect the cost of delivering the service beyond identity verification, then service costs do not need to be included.

¹⁵ In some cases, digital ID systems may not be able to distinguish which IDs belong to unique individuals (e.g., one person can have multiple email addresses). In this case, the ID provider will need to make some assumptions about the average number of IDs a unique individual has in order to calculate the estimated number of unique enrollees.

Note that this metric may be vulnerable to misreporting if providers either do not accurately report on the various usages of user data or on the usages to which they solicit consent. If ID2020 does pursue user-based metrics (see Recommendation 3), consent could be explored further during those surveys. Additionally, ID2020 could pose as a “mystery shopper” (if possible) to confirm what is made clear during the consent process.

Metric 3b: Number of personal data points that are accessed by a digital ID-linked service/right that are not needed to provide it
Digital ID-linked services or rights should only have access to the data that are needed to provide these services or rights. Service providers should not request additional data from ID providers, and ID providers should be responsible gatekeepers for their users’ personal data. This metric can be used to compare how well systems protect user privacy across service provider pilots.

This data can be collected from self-report or administrative data from service providers, though this metric is relevant for ID providers who should ultimately be responsible for safeguarding user data. As with the previous metric, this metric may be vulnerable to misreporting. When possible, ID2020 should observe the raw data points that a service provider can access.

Metric 4a: Proportion of people who have successfully used the digital ID platform to access a service during the pilot among those enrolled in the digital ID system

Enrolling the digital ID system is only one step to improved lives. This should be followed by usage of the digital ID to access services. This metric will allow ID2020 to understand the percentage of enrollees who are actively using the digital ID system to access a service. It is an indication of the magnitude of impact that a pilot could achieve or is achieving.

This metric should be collected by both service-linked ID providers and service providers. The numerator is the number of unique IDs (which we assume map to individuals) that have been successfully used during the pilot to access a service (i.e, authentication was successful). The denominator is the number of individuals enrolled in the system who would theoretically access this service or would have a genuine claim to accessing that service.¹⁶

This metric should also be collected using administrative data. If this metric is low, it does not necessarily mean that people are not accessing the service. However, low usage could suggest that the digital ID component to accessing the service has smaller value-add than one might hope.

Metric 4b: Proportion of people who have experienced a failed authentication when attempting to access a digital ID-linked service during the pilot

Many pathways to impact for digital ID are connected to utilizing services and rights. This metric largely attempts to capture the degree to which the digital ID limits people from accessing services. Authentication could fail if there is a glitch in the software, the hardware is broken, or there are errors in the system.

This metric should be captured using administrative data from either the service provider or the ID provider, depending on which is better positioned to capture this data.¹⁷ The numerator should be the reported number

¹⁶ In most cases, this will be the same as the number of individuals enrolled on the system. However, in some contexts, services may only apply to subsets of the enrollees. For example, people over retirement age may use a digital ID to access social security benefits, even though people of all ages may have that digital ID. The denominator should be restricted to the number of enrollees over retirement age.

¹⁷ ID2020 should work with pilots ahead of time to understand what administrative data on authentication failures their system can automatically capture and the format of that data. Authentication failures might not always align with the number of individuals who cannot access services due to authentication failures. For example, if one person attempts to

of people who have experienced a failed authentication, while the denominator should be the total number of people who attempt to use the digital ID to access the service.

Many services have some form of backup, so it is possible that this metric does not fully capture exclusion. While individuals still may ultimately access the service, this metric is still a helpful proxy in most, if not all, cases.

Recommendation 3: Support data collection for 11 user-based metrics from each pilot

There are an additional 11 metrics that are also universal since they facilitate comparisons across pilots in key areas of the digital ID theory of change. However, these metrics require user-based data collection, which can be resource-intensive. We encourage ID2020 to support data collection for these 11 metrics either by providing an external measurement team or by financing data collection by the pilot organization.¹⁸ If ID2020 is not in a position to support data collection for these metrics, we do not believe that ID2020 should request these metrics from the pilots (unless the pilots are already collecting them).

While some of these metrics may have proxy metrics that could be measured through administrative data from ID or service providers, these sources are likely to be less rigorous than user-based data either because administrative data does not exactly capture the measurement construct or because it may be less accurate. Therefore, we recommend that these metrics are collected directly from users for the pilot phase. Once pilots move into a scale-up phase, they could move to administrative proxy measures for ongoing monitoring.

Our list of user-based metrics and the type of pilot for which they are relevant are outlined in

Table 2 and described in greater detail below. The phrasing of the questions for each of these metrics should be piloted with potential users to refine the exact language. To the extent possible, the framing should be consistent across pilots.

Table 2. Recommended user-based metrics

| Measurement area | Metric | Relevant pilot | |
|------------------------|---|----------------|------------------|
| | | ID provider | Service provider |
| [4] Access to services | [4c] Proportion of people without a digital ID who have been denied access to a digital ID-linked service | Yes | Yes |
| [5] User-centricity | [5a] Proportion of eligible users who receive information on the consent process | Yes | Yes |
| | [5b] Proportion of eligible users who understand the consent process | Yes | Yes |
| | [5c] Proportion of eligible users who receive information on the enrollment process | Yes | Yes |

access a service four times with her digital ID, the system may capture four authentication failures, even though this is one individual. ID2020 should work with pilots to understand how this data is captured and how best to approximate authentication failure of a single individual.

¹⁸ Pilots should only collect their own user data if they have adequate measurement capacity within their organization.

| | | | |
|-----------------------------------|---|-----|-----|
| | [5d] Proportion of eligible users who understand the enrollment process | Yes | Yes |
| | [5e] Proportion of eligible users who receive information on how their data will be used | Yes | Yes |
| | [5f] Proportion of eligible users who understand how their data will be used | Yes | Yes |
| [6] User attitudes | [6a] Proportion of users who have a grievance related to the digital ID during the pilot | Yes | Yes |
| | [6b] Proportion of eligible users who feel that the digital ID system is secure | Yes | Yes |
| | [6c] Proportion of eligible users who feel that the digital ID system adequately protects their privacy | Yes | Yes |
| [7] Attribution and Credentialing | [7a] Average number of credential documents added per user over the pilot period | Yes | Yes |

Metrics 4c: Proportion of people without a digital ID who have been denied access to a digital ID-linked service

While administrative data can be used to measure the degree to which services are accessed through digital ID, it cannot be used to capture the degree with which access is denied as a result of not having a digital ID. This metric can be used to compare ease of access to services across pilots.

This metric should be collected among a representative sample of the population who could be served by a potential service¹⁹ but who do not have a digital ID.

Metrics 5a-f: Eligible user knowledge of the digital ID consent, enrollment, and data use processes

We recommend six metrics to capture availability of information and potential user knowledge of the digital ID system. These include:

- 5a. The proportion of eligible users who receive information on the consent process
- 5b. The proportion of eligible users who understand the consent process
- 5c. The proportion of eligible users who receive information on the enrollment process
- 5d. The proportion of eligible users who understand the enrollment process
- 5e. The proportion of eligible users who receive information on how their data will be used
- 5f. The proportion of eligible users who understand how their data will be used

¹⁹ The lack of digital ID may be the only factor limiting access to the service.

These six metrics should be collected among a representative sample of the target population – those that have the ID already and those who do not.²⁰ For each step and the relevant binary sub-metric, users should be asked knowledge-based questions to ascertain their level of understanding by comparing user responses to what is actually known from the pilot’s description of its processes and procedures.

Metric 6a: Proportion of users who have a grievance related to the digital ID during the pilot

User satisfaction with the digital ID system is critical for the system’s sustainability and uptake. This metric is a binary metric for whether users have a grievance related to the digital ID. Grievances can relate to any aspect of the digital ID system, including (but not limited to) ease of use, links to services, hardware, ID requirements, etc.²¹

This metric should be collected among a representative sample of users who have or have attempted to enroll in the digital ID system to ensure that this mechanism is capturing user grievances rather than perceived grievances. While this metric focuses on whether a user has a grievance, the opportunity to survey ID users should be used to explore the factors behind these grievances. Learnings could help pilots understand why some aspects of digital ID might not be working and identify strategies for improvement.

Metric 6b: Proportion of eligible users who feel that the digital ID system is secure

This metric is intended to understand user perceptions of security. It can be used to compare one dimension of user trust across pilots. This metric should not be interpreted as a measure of whether a system *is* secure, as user perceptions may not align with reality.

This metric should be collected among a representative sample of the target population who have heard about the digital ID system. As with the previous metric, while this metric is binary, follow-up questions could be added to explore what factors drive user perceptions to help pilots improve their brand and improve user trust.

Metric 6c: Proportion of eligible users who feel that digital ID system adequately protects their privacy

Another dimension of user trust is users’ perception that a system protects their privacy. As with the previous metric, this should not be interpreted as a measure of whether a system *does* protect user privacy.

This metric should also be collected among a representative sample of the target population who have heard about the digital ID system. Again, follow-up questions could be added to understand why eligible users feel the way that they do.

Metric 7a: Average number of credential documents added per user to digital ID over the pilot period

ID2020 prioritizes systems that provide ease of adding credentials. This metric will help ID2020 compare this dimension of digital ID systems across pilots. Since ID providers and service providers may be unlikely to know the full extent to which credential documents are added to digital IDs, this metric is best measured by asking users how many credential documents they have added over the pilot period. This metric should be collected

²⁰ The sample should be limited to those who are aware that the digital ID system exists. It may not need to be limited to those who have had the chance to enroll, as individuals could still know this information but choose not to enroll.

²¹ We recommend that this metric be collected through a user survey rather than data captured through a grievance mechanism because 1) not all systems will have a grievance mechanism and 2) some users may not know about or be able to use the grievance mechanism if it does exist. Therefore, relying on administrative data may capture the ease of issuing a grievance rather than the extent to which users have grievances.

for all ID systems. While user responses will be “0” for those systems that do not allow for adding credentials, this is still meaningful given ID2020’s interest in this aspect.²²

This metric should be collected among a representative sample of enrolled users.

Recommendation 4: Work with pilots to determine the feasibility and applicability of collecting up to 12 secondary metrics

The universal metrics alone do not capture the complete story of the pilot. Therefore, we have included a list of secondary metrics that can add relevant information about the likelihood of success of a given pilot. These metrics are secondary metrics rather than universal because:

1. The feasibility of data collection depends highly on the availability of existing data, which may vary widely across pilots
2. The metric is less generalizable across pilots
3. The interpretation of the metric relies on the context of the pilot, making it less comparable across pilots
4. The metric is meant to generate additional data that will help the pilot diagnose underlying challenges, but is not a performance metric

We recommend that ID2020 work with pilots to understand which secondary metrics pilots could feasibly collect and how they would collect them. Pilots should collect those metrics that are feasible to collect and are relevant to that pilot’s context. ID2020 may also select two to three additional metrics that pilots should find a way to collect if they are particularly important to that pilot’s theory of change. This can be determined by mapping the pilot’s theory of change to the detailed digital ID theory of change and identifying the critical nodes for that pilot.²³

The starting list of secondary metrics is included in Table 3 and described in greater detail in Appendix I.

Table 3. Recommended secondary metrics

| Measurement area | Metric | Relevant pilot | |
|--|--|----------------|------------------|
| | | ID provider | Service provider |
| [1] Feasibility of adoption / scale-up | <i>Fully captured in Table 1</i> | | |
| [2] Enrollment | [2c] Proportion of people who have not been able to enroll in the digital ID system when attempting to enroll during the pilot | Yes | Yes |
| [3] Privacy | [3c] Number of third parties involved in processing data during enrollment and service provision | Yes | Yes |

²² There may be digital ID systems that are in early stages or that serve specific populations, where no ability to add credential documents is reasonable.

²³ IDinsight could be a thought partner in this work.

| | | | |
|-----------------------------------|--|-----|-----|
| [4] Access to services | [4d] Proportion of authentication events during which a manual override is required during the pilot | Yes | Yes |
| | [4e] Average staff time required for direct and indirect service delivery that is linked to digital ID | No | Yes |
| | [4f] Proportion of digital identities that are reported as having incorrect information | Yes | Yes |
| [5] User-centricity | [5g] Did the pilot consult users on their needs during the initial design of the digital ID system? [Y/N] | Yes | Yes |
| | [5h] Number of changes in design of digital ID system as a direct result of user feedback during the first year of the pilot | Yes | Yes |
| | [5i] Score on approval rating system | Yes | Yes |
| [6] User attitudes | [6d] Number of times a person accesses a service with a digital ID | Yes | Yes |
| [7] Attribution and Credentialing | <i>Fully captured in Table 2</i> | | |
| [8] Security | [8a] Number of security breaches during the pilot | Yes | Yes |
| [9] Data quality | [9a] Number of instances users correct data once it has been entered into the digital ID database | Yes | No |
| [10] Data portability | [10a] Number of reported service providers that have accepted the digital ID platform | Yes | No |

Recommendation 5: Select a subset of pilots for deep measurement and evaluation to generate more rigorous evidence for the digital ID sector

The majority of the proposed metrics are related to tracking inputs, implementation, and outputs. Few test the linkage between outputs and outcomes (and, by extension, impact). Further, several of the interviewees highlighted that there is not yet evidence that digital ID actually improves lives beyond the ID systems that already exist. There also is a lack of evidence that tests which use cases are most impactful and cost-effective.

While we generally recommend that ID2020 calibrate selected metrics across pilots so that they are feasible given the pilots' capacities, this focus may limit the rigor with which pilots can measure their impact. For this reason, we encourage ID2020 to invest more deeply in specific pilots that can 1) rigorously test assumptions that relate to the social impact of digital ID and 2) generate deeper data-driven insights into the efficiencies and challenges of digital ID systems.

Pilots that should be considered for deeper investigation are those that:

- Represent likely future use cases in the digital ID space
- Currently lack evidence for their broader use case

- Have the capacity to work with an external measurement team
- Buy into the idea of deeper investigation and are committed to acting on results
- Implement interventions that are replicable at a large scale

LOOKING AHEAD

ID2020 is in a unique position to convene an array of stakeholders working in the digital ID sector united behind a common vision of responsible and impactful digital ID systems. Generating data and evidence will help ID2020 optimize the impact of its funding and prioritize the most promising digital ID pilots. Early pilots should generate valuable learnings for ID2020. We recommend that this document be considered a “living document.” ID2020 should continue to refine the metrics it collects based on the types of pilots that they support and the learnings that are most valuable to their strategic decisions and to advancing the goals of the digital ID sector.