AAMVA DECISION SUPPORT CONTRACT

STRUCTURED DECISION-MAKING ROADMAP
for the
EVALUATION OF BIOMETRIC TECHNOLOGIES
in a
DRIVER LICENSE ENVIRONMENT

December 2003

Prepared by Fischer Consulting Inc

EXECUTIVE SUMMARY

The purpose of this document is to set out a roadmap for evaluating biometric technologies for use in a driver license environment.

The roadmap is the result of a Project conducted by AAMVA's UID9 Task Group to investigate the options to ensure (1) that every driver in North America has one and only one driver license and driver license record, (2) that the bearer of a driver license card is the individual to whom the card was issued, and (3) that a driver record only contains information pertaining to the bearer. In order to determine if biometric technology could be used to achieve the above, AAMVA appointed two Contractors (Fischer Consulting Inc and International Biometric Group) to respectively assist in the decision process and technical evaluation of biometric technologies.

AAMVA split the Project into two phases, the first considering technical performance, and the second considering the broader feasibility of a biometric technology implementation. During the Phase I evaluation, it became evident that insufficient data existed to continue with the evaluation. AAMVA is now soliciting the needed information to allow them to continue with the evaluation.

This document is illustrated with the work that was performed on the Project up to the point where the unavailability of data dictated the suspension of evaluation activities. The remainder of the document walks through the process to be followed to establish a decision context and continue with the evaluation. The roadmap will serve to guide AAMVA in its decision making process once they have received complete performance data.

The roadmap suggests and discusses the following activities:

- Define a decision context. The decision context describes the setting within which the evaluation is taking place
- Identify objectives. Differentiate between fundamental objectives (representing decision-maker values) and means objectives (representing causal relationships). The combination of the decision context and the objectives should be limited to those issues that are relevant to the decision in terms of scope, ownership, time and money
- Identify or design measurement scales for the fundamental objectives.
- Use performance targets to filter out technologies that will not be able to deliver the required functionality.
- Compare various alternatives against each other, using utility functions. Utility functions reflect a decision-maker's preferences and risk attitude under uncertainty.
- Perform a sensitivity analysis.

The document concludes with a discussion of aspects of the roadmap that need to be considered before application to jurisdictional level.

TABLE OF CONTENTS

ABBREVIATIONS / GLOSSARY

AAMVA	American Association of Motor Vehicle Administrators
Conversion period	The period over which all drivers are scheduled to apply for a new DL that is based on a biometric technology.
DL	Driver License
DSP	Decision Support Project
False Match	A traditional metric for expressing the performance of a biometric technology. Used in connection with 1:1 matching in this document. A false match results when a technology reports a match when the biometric templates of different persons are compared.
False Non-Match	A traditional metric for expressing the performance of a biometric technology. Used in connection with 1:1 matching in this document. A false non-match results when a technology reports no match when two biometric templates of the same person are compared.
FMR	False Match Rate. The rate/percentage of false matches.
FNMR	False Non-Match Rate. The rate/percentage of false non-matches.
FTE	Failure To Enroll. Refers to an instance of a person whose biometric information cannot be acquired, or is of insufficient quality to be used.
FTER	Failure To Enroll Rate. The rate/percentage of people who fail to enroll.
OSIPM	Open Set Identification Performance Metrics. A set of metrics used to express the performance of a biometric technology, explicitly allowing for the scenario where the person whose biometric information is being compared against a database is not already contained in the database. In this document, the OSIPM are used in connection with 1:n searching.
UID9	The Unique Identifier Task Group, Task Group 9, of AAMVA's Uniform Identification Subcommittee. UID9 consists of representatives of the four AAMVA regions, each representative being an expert in the field of driver licensing.
UID	AAMVA's Uniform Identification Subcommittee

1. INTRODUCTION

On the identification front, the challenge faced by any driver license (DL) system is to uniquely identify an individual such that:

- One person will have only one driver record on the database and only one driver license (DL)/ identification (ID) card (also known as the “one driver one driver license on driver license record” concept).
- Authorized users can verify that the holder of a DL/ID card is the individual to whom the card was issued.

On October 24, 2001, AAMVA's Executive Committee passed a resolution creating a Special Task Force on Identification Security to develop a strategy on enhancing the issuance of secure identification credentials for driver licensing and photo ID purposes. The Uniform ID Subcommittee, a component of the Driver License & Control (DL&C) Committee, addressed the bulk of the work that was created by the Special Task Force. The Uniform ID Subcommittee created a host of Task Groups to address components of the identification security strategy. Task Group UID9 was charged with addressing "unique identifiers" for driver license and identification card issuance. UID9 explored the issues of unique identifiers, formulated and proposed recommendations to the larger Uniform ID Subcommittee[1].

UID9 believed that biometric technology might provide a solution to the unique identification requirements. However, as biometric technology had not yet been implemented or tested in a manner that would support a decision to implement a solution across North America, UID9 initiated a Decision Support Project (DSP). The purpose of the Project was to determine the feasibility of implementing a biometric technology (or biometric technologies) as part of a comprehensive North American system to assist in the unique identification of drivers. Fischer Consulting Inc was appointed to facilitate a structured decision support process, and International Biometric Group (IBG) was appointed to provide expert technical advice on biometrics.

The purpose of this document is to set out a roadmap for evaluating biometric technologies for use in a driver license environment. The roadmap is illustrated with the work that was performed on the DSP up to the point where the unavailability of data dictated the suspension of evaluation activities. The subsequent part of the roadmap suggests the process to be followed when the needed performance data becomes available.

Although the roadmap is focused on the use of biometric technology on a continent-wide basis, it also provides a framework to individual jurisdictions for evaluating biometric technology for use at the jurisdictional level.

This document is structured as follows:

- An overview is provided of the steps in a generic decision-making process. Brief reference is made to the manner in which the process was implemented on the DSP.

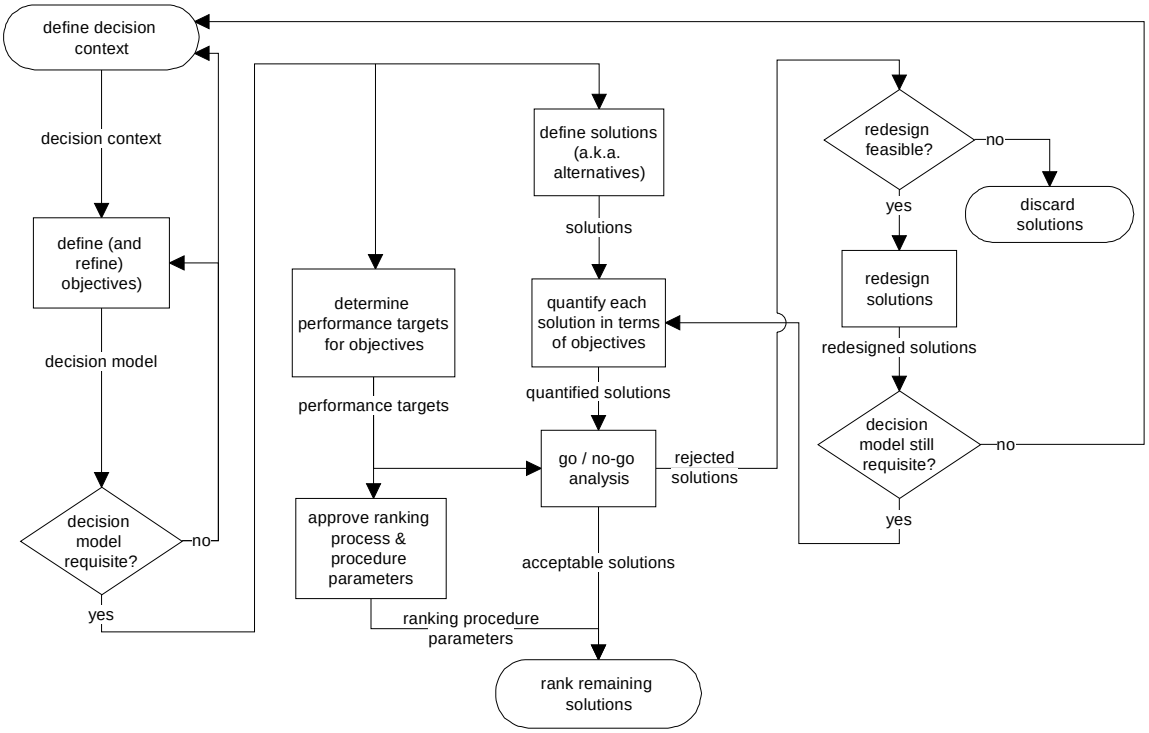
[1] Normally, a subcommittee would forward recommendations through the parent committee - the DL&C Committee in this case. However, a special procedure was established to streamline the approval process for AAMVA's secure ID effort. Recommendations of the UID Subcommittee flowed to the AAMVA Board, through the AAMVA President & CEO, for approval. The AAMVA Board is the highest level of approval within the AAMVA approval hierarchy

- Using the decision-making process as a framework, the progress and results obtained are discussed.
- Continuing with the decision-making process, the activities that still need to be performed are identified, explained and illustrated by way of example data.
- A few notes are provided on customizing the roadmap for use at individual jurisdictional level.

2. THE DECISION-MAKING PROCESS

2.1 OVERVIEW

The diagram below illustrates a generic evaluation process:



In essence, the diagram can be summarized as follows:

- Define the decision at hand.
- Identify the objectives. That is, state the goals that need to be achieved.
- Create/design solutions that would satisfy the objectives.
- Quantify each solution in terms of the stated objectives.
- Determine performance targets for each objective. That is, identify the required performance of any solution with regards to each objective. Use these performance targets to weed out any solutions that are not in compliance.
- Design a process for ranking the remaining solutions, and rank accordingly.
- Allow for feedback loops to revisit previous steps in the process.

2.2 DECISION CONTEXT

The decision context is a description of a decision (or in this case an evaluation) and the setting within which it needs to me made. The description needs to be consistent with the objectives, and needs to reflect the variables that define and influence the measurement of the objectives as well as any other issues that define the environment within which the decision is made and within which the outcome will be assessed.

Change to a decision context is possible, and is not indicative of poor decision-making; instead, it indicates that the decision situation is being taken seriously, and increases the quality of the end result.

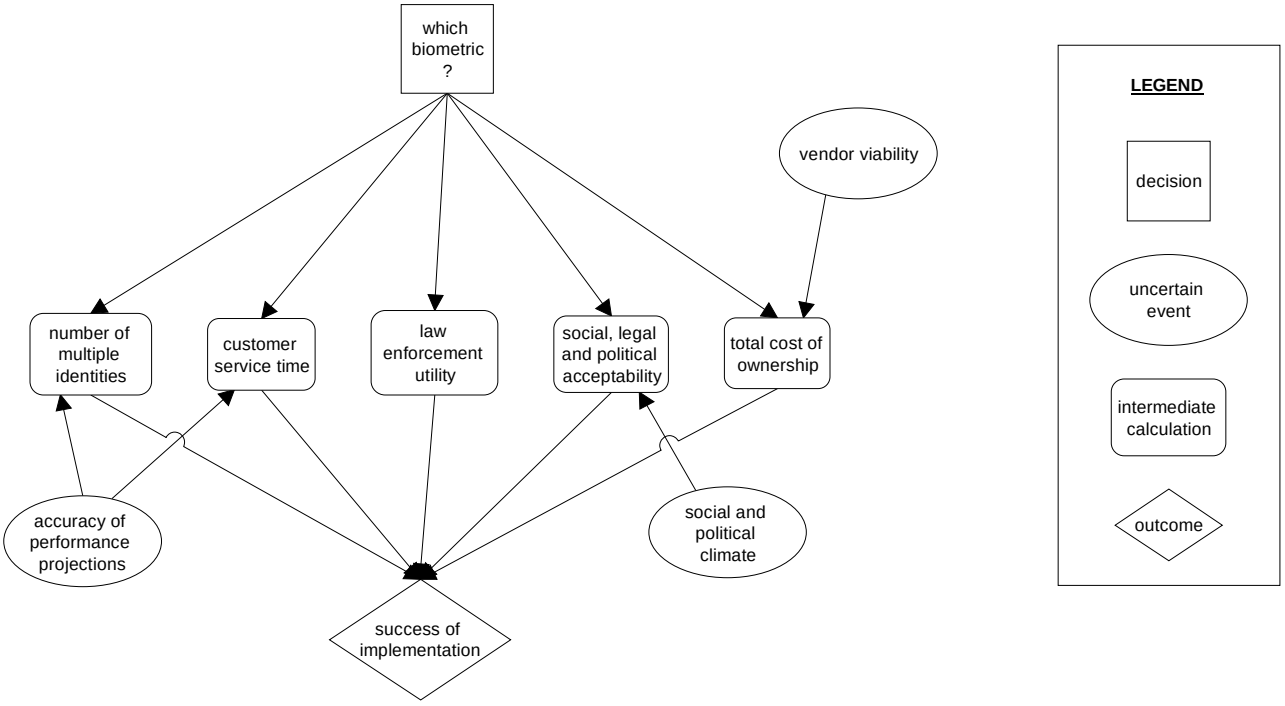
In the case of biometric technology, the decision context includes the processes, equipment, infrastructure, resources, etc., that apply to the use of a specific biometric technology or combination of technologies. For example, the decision context for the DSP included all objectives and all variables applicable to the implementation of a biometric technology throughout North America, inclusive of the business requirements and the environment within which the implementation is to take place.

2.3 DECISION MODEL

A decision model consists of a decision context and a set of associated objectives. A requisite decision model limits the decision context and objectives to those issues that are relevant to the decision, in terms of scope, ownership, time and money. A requisite decision model thus is a decision context and a set of associated objectives that:

- Reflects the true problem at hand, i.e. contains only those issues relevant to the decision at hand. For example, UID9 decided to initially limit the decision context to the technical capabilities of biometric technologies, to determine if a technology that can actually work, exists (see Paragraph 3). In this case, the requisite decision model excluded other considerations such as cost, privacy, etc., which, although important, would not contribute value to the question being considered.
- Matches the decision ownership.
- Is feasible in terms of time and money.

A decision model can also graphically be represented in a diagram such as the following:



2.4 OBJECTIVES

Objectives are those issues that define the goals of a project. A project can have many different objectives, some of which may be conflicting. A distinction is made between fundamental objectives representing the values of the decision-makers, and means objectives that represent the means by which to achieve the fundamental objectives. Not making this distinction often leads to double counting during evaluation. In order to facilitate their administration within the decision-making process, fundamental objectives can be structured into a fundamental objective hierarchy, and means objectives can be structured into a means objective network.

Fundamental objectives rather than means objectives are used as criteria for the evaluation of different alternatives. To facilitate this, a measurement scale is identified or designed for each fundamental objective. The measurement scale can be any of the following:

- A natural scale, such as money, time, distance etc.
- A proxy scale, where one scale is used as representative of another, e.g. money spent on research as an indication of the robustness of a technology.
- An attribute scale, which defines levels of compliance (e.g. Best, Better, Satisfactory, Worse, Worst), and fully describes each level.

2.5 PERFORMANCE TARGETS

The purpose of setting performance targets is to create a filter to eliminate solutions that do not meet certain minimum requirements. The performance targets depend only on the fundamental objectives forming part of the decision model, and thus can be set in parallel with the drafting of the solutions.

Note that performance targets (specifically those influenced by the availability of money and human resources) are often determined by the political and economic climate of the day. As a result, the duration of an evaluation may suggest that performance targets be reviewed during the process.

In general, performance targets are set either directly, or derived from an influence for which a performance target can be set.

2.6 SOLUTION IDENTIFICATION

Once the decision model has been drafted, various alternatives or solutions that would satisfy the stated objectives are designed. The identification of alternatives that will provide a solution to the objectives is a vast field and has been studied extensively. On the DSP, however, the decision context limited the possible alternatives to biometric modalities and technologies currently in use with large-scale applications. In particular, facial recognition, iris recognition, and fingerprint recognition were identified as the contending modalities.

2.7 SOLUTION QUANTIFICATION

An assessment is made of each identified solution in terms of the fundamental objectives.

On the DSP, the evaluation process did not progress beyond the quantification of the Phase I objectives (see Paragraph 3 below). UID9 determined that the available performance data (from tests and operational systems) applied only to environments that were too small or operationally too different from the environment UID9 foresaw, and consequently could not be used with sufficient confidence to project performance. In order to proceed, more applicable performance data was needed.

At the time of writing this Report, actions were under way to obtain the required performance data. Specifically, the biometric industry was approached with the data requirements[2], with a request that the necessary tests be performed and/or verified by independent entities. As the time scales for the delivery for the data were unknown and outside the control of UID9, the DSP concluded with this Report.

2.8 RANKING PROCEDURE

The ranking procedure allows for:

- Evaluating trade-off between conflicting objectives.
- Assessing the influence of uncertainty and chance events on expected outcomes.
- Incorporating subjective judgments.

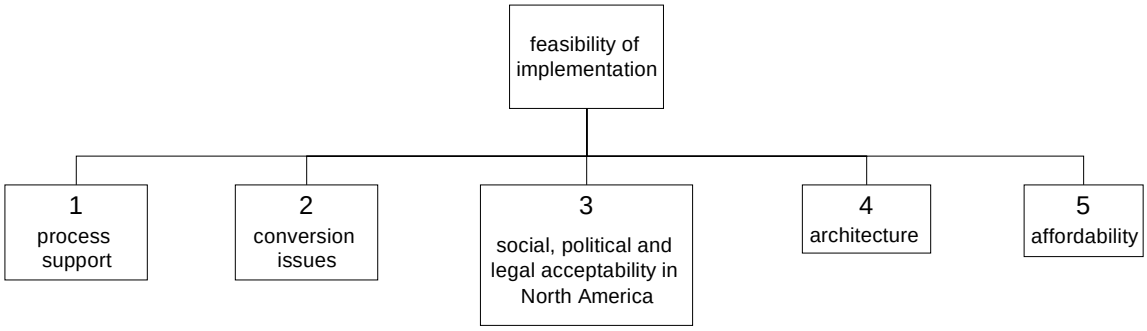
The basic steps in the ranking process are:

- Construct an overall utility function that captures the relative importance of the fundamental objectives.
- Construct a utility function for each lowest level fundamental objective.
- Perform a sensitivity analysis.

[2] "Biometric Technology Information Needs" B004.REPT.002 dated December 2003

3. RESULTS TO DATE

As was mentioned above, the initial question faced by UID9 was if a biometric technology could be used to ensure that each driver has only one driver license and only one driver record. Within this context, the DSP started off by listing relevant objectives, and categorizing these objectives according to the categories reflected in the diagram below.

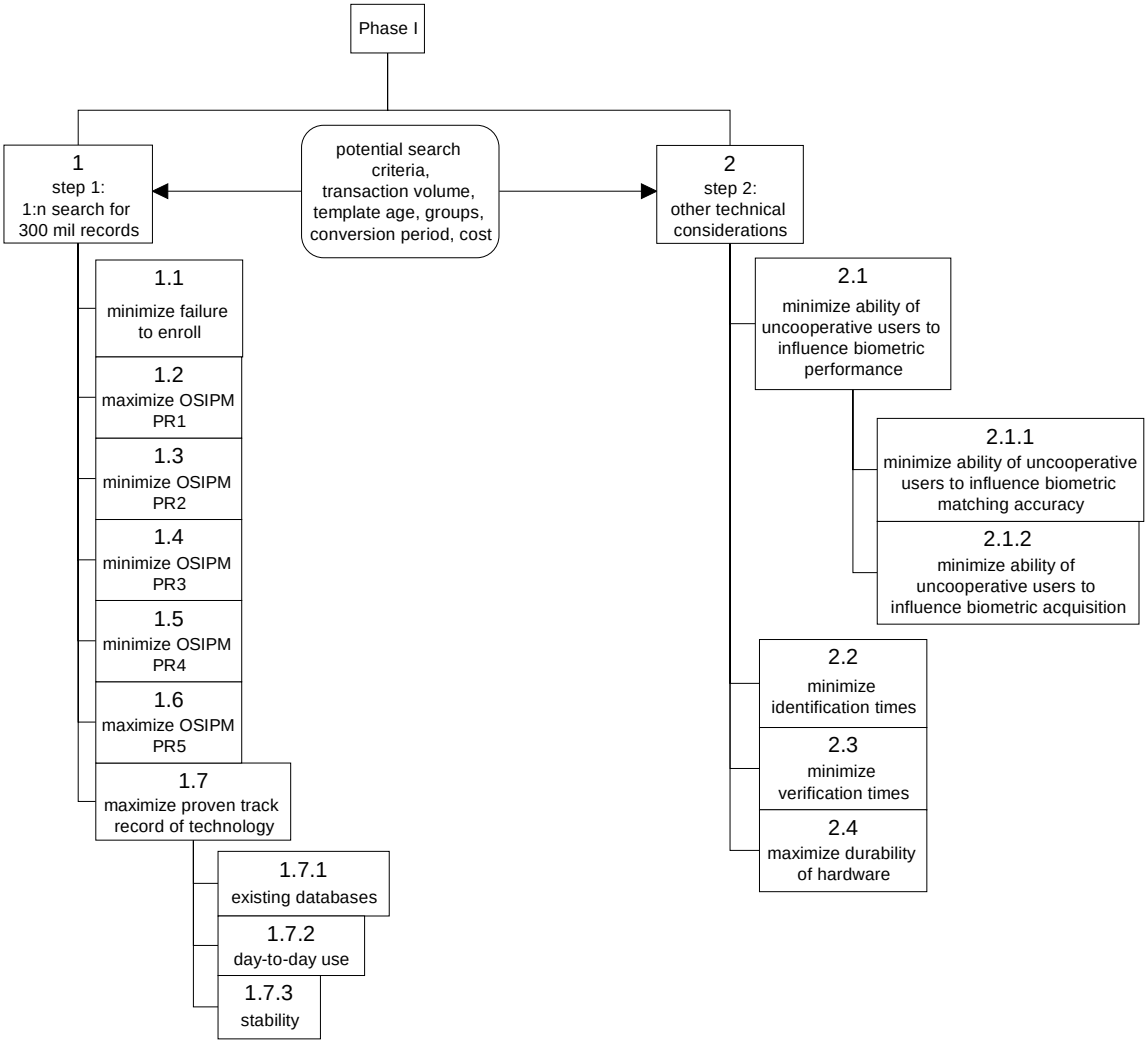


A complete list of the objectives identified can be found in Appendix A.

Due to the critical significance of the technical feasibility of a biometric implementation (in other words, is there a technology that can do it?), UID9 split the Project into two Phases, by first considering a core collection of technical objectives, before evaluating the full spectrum of objectives. The purpose of Phase I was to establish if a biometric technology that complies with the technical performance requirements does in fact exist. Upon completion of Phase I, if compliant biometric technologies have been identified, these technologies were to be submitted to the Phase II evaluation.

3.1 OBJECTIVES (PHASE I)

Using the initial list of objectives (see Appendix A), the following set of fundamental objectives and variables describing the Phase I decision model was identified (the block with the rounded corners represents the variables; all other blocks represent objectives).



In order to further focus UID9's time and other resources, Phase I was further divided into Step 1 and Step 2 as indicated above. Step 1 defined a decision context involving only a few core technical objectives, and Step 2 defined a decision context considering a small number of additional technical objectives. UID9 also identified a collection of variables applicable to both contexts, and left itself the option to add these to either Step 1 or Step 2 when appropriate as more information about the feasibility of a biometric implementation became available.

A detailed discussion of the Phase I objectives and their associated measurement scales can be found in Appendix B.

3.2 PERFORMANCE TARGETS (PHASE I)

The performance targets for the Phase I objectives defined the level of performance that a biometric technology should demonstrate in order to be considered for evaluation in Phase II of the Project. These

performance targets are neither necessarily the same that will apply to the decision context defined for Phase II, nor intended for use as operational requirements (i.e. performance targets for a live system). The performance targets were identified only as a threshold to use for identifying technologies that merited further consideration.

Considerable discussion preceded the determination of the Phase I performance targets. This discussion (and the resulting performance targets) is reflected in Appendix C.

The following are some of the performance targets specified for Phase I:

- The rate at which the sum of Potential Result 2 and Potential Result 4[3] occurs must be equal to or lower than 10%.
- The rate at which the sum of Potential Result 2 and Potential Result 3 occurs must be equal to or lower than 10%.
- Failure To Enroll rate (FTER): Equal to or lower than 5%. The failure to enroll rate is the percentage of individuals whose biometric information is either unobtainable or of insufficient quality to be used for identification in a biometric system.

It is very important that the figures quoted above are not used in isolation, but that cognizance is taken of their purpose within the DSP, as well as of the reasoning leading to their specification (as set out in Appendix C).

3.3 SOLUTION QUANTIFICATION: STEP 1 (PHASE I)

UID9 decided to conduct Step 1 of Phase I whilst considering only the transaction volume and conversion period variables. These variables were fixed and did not cover the ranges suggested in Appendix B.

IBG compiled a technical report pertaining specifically to Step 1 of Phase I. The report covered existing deployments and tests, as well as feedback from various vendors in response to a questionnaire.

The report evaluated the three identified alternative biometric technologies, namely facial recognition, iris recognition and fingerprint recognition. The report concluded the following:

- Facial recognition is unlikely to perform at the required levels in a 300 million record environment.
- Insufficient information exists to determine if either fingerprint recognition or iris recognition will perform as required in a 300 million record environment. This can be attributed to several reasons:
    - No one has yet implemented a system of the same size. Although several large fingerprint databases (of up to 55 million persons) have been deployed, the objectives and environment differ from that of AAMVA.
    - The AAMVA concept of operations has not been afforded due attention yet. Most technology tests (for facial, iris, and fingerprint biometrics) used a closed set approach, where the assumption is that a person whose biometric is being matched against a database, is already in the database. This is clearly inconsistent with the AAMVA environment. For those tests that did use an open set approach (i.e. the assumption is that a person whose biometric is being matched against a database may or may not already be

[3] The Open Set Identification Performance Metrics (OSIPM) consist of 5 Potential Results. Refer to Appendix C for a definition of each.

on the database), a false match resulted in a “failure”, whereas a false non-match could be handled with an exception procedure. In the AAMVA environment, exactly the opposite is true – a false non-match results in a “failure” (allowing a person to obtain multiple identities), and a false match is handled by way of an exception procedure.

- No vendor independent testing has been performed on iris recognition.
- Most biometric testing does not investigate the influence of template ageing.

- None of the available test or deployment information was directly applicable to the AAMVA environment.

The current state of affairs thus was summarized as follows:

1. Step 1 (of Phase I) requires the evaluation of a small set of objectives, given a database of 300 million records (the target system). The goal is to determine if a biometric technology will perform as required in respect of each of the objectives (i.e. if it can meet the performance targets).
2. In an ideal scenario, this determination is based on the following:
    - Performance of existing biometric systems similar to the target system, i.e. similar size, within a similar environment (e.g. civil identification), and with similar purpose (e.g. 1:n identification). In the remainder of this Paragraph, such systems are referred to as reference systems.
    - Results of tests designed specifically to investigate the performance of biometric technologies in the target system environment.
    - Optionally, vendor claims/data supporting the required performance targets.
3. The current situation regarding each of the above was as follows:
    - It was considered unlikely that a biometric system of similar scope/size as the target system will be implemented anywhere in the world prior to the implementation of a driver license system in North America.
    - Although the published results of some biometric tests provided insight into the possible behavior of biometric technologies within the target system environment, no tests had to date been specifically designed or conducted to investigate the performance of the system in question.
    - Some vendors claimed compliance with the performance requirements of the target system.
4. Without confirmation by an independent 3rd party, vendor claims and/or data were deemed inadequate for decision-making, and could at most be used to confirm if a particular technology does not comply with the performance requirements.
5. When looking at the performance of existing systems, the bigger the difference in environment between the target system and the reference system, the less useful the reference system’s performance is in estimating the target system’s performance. Due to the multitude of variables that may influence the performance of an operational system, extrapolating from an operational system (e.g. a system half the size of the target system) to project performance of the target system is likely to introduce too much uncertainty for the results of such extrapolation to be meaningful. Thus,

unless a reference system is acceptably similar[4] to the target system, the reference system's performance is unlikely to be of much use in estimating the target system's performance.

6. Tests specifically designed to evaluate biometric technologies for the target system may be able to estimate the performance of specific parts of the target system.

Given the above situation, UID9 accepted that any determination of whether or not a specific biometric technology (or combination of biometric technologies) does meet the specified performance targets would be made under some uncertainty (i.e. it will require a "leap of faith" to accept). However, due to the complexity of the environment, UID9 did not specify the maximum size of the "leap" (i.e. the maximum acceptable difference between the reference and target environments). What was specified, was that at least the following have to hold:

- Test results (or results from operational deployments) for 1:n searches need to be expressed in terms of the OSIPM[5].
- Tests have to be conducted by an independent entity.
- Test data used should be statistically representative of the 1:300 million environment in terms of demographic distribution, template age, and database size.

Once test data complying with the above requirements becomes available, AAMVA will consider whether or not it allows for the evaluation of Phase I Step 1 to be resumed.

4. THE ROAD AHEAD

As mentioned before in this document, at the time of writing this Report, actions were under way to obtain the performance data that would allow AAMVA to continue with the evaluation of biometric technologies. Assuming that such data will become available in due course, the remainder of this document identifies and discusses the actions required to complete the evaluation process.

4.1 PHASE I

4.1.1 GO / NO-GO ANALYSIS: STEP 1

Once the objectives within a specific decision context have been quantified, they are compared against the performance targets. This is a simple and mechanical go / no-go comparison.

[4] Note that even though the difference between the reference and target system environments is negligible, the reference system information will still be limited to the particular technology of the vendor that supplied the reference system. It is also highly unlikely that accurate information for the OSIPM Potential Result 2 and Potential Result 3 will be available.

[5] These metrics are fully described in the Department of Defense's "Face Recognition at a Chokepoint - Scenario Evaluation Results" dated 14 November 2002, issued by the Department of Defense Counterdrug Technology Development Program Office. Refer to Appendix C for additional information.

For example, suppose the evaluation of Phase I Step 1 yields the following results (See Appendix B for a discussion of the Phase I Objectives):

	Objectives								
	PR1 (%)	**PR2 (%)**	**PR3 (%)**	**PR4 (%)**	**PR5 (%)**	**FTER (%)**	**Existing database size (records)**	**Day-to-day use (years)**	**Stability**
Technology 1	94	2	4	4	96	4	60 million	8	Most satisfactory
Technology 2	97	1	2	1	99	2	2 million	2	Satisfactory
Technology 3	79	9	4	8	92	0.01	4 million	2	More satisfactory

Note that for Technology 3, PR2 + PR4 > 10%, and PR2 + PR3 > 10%. Consequently, Technology 3 does not pass the performance targets (as identified in Paragraph 3.2), whilst Technology 1 and Technology 2 do.

Adding some sophistication to the process, the degree of confidence in the stated performances can be incorporated. Such confidence intervals should be supplied by the data source. AAMVA may decide that for the various rates, the outcome must for example have a 90% probability of meeting or exceeding the stated performance target.

4.1.2 STEP 2

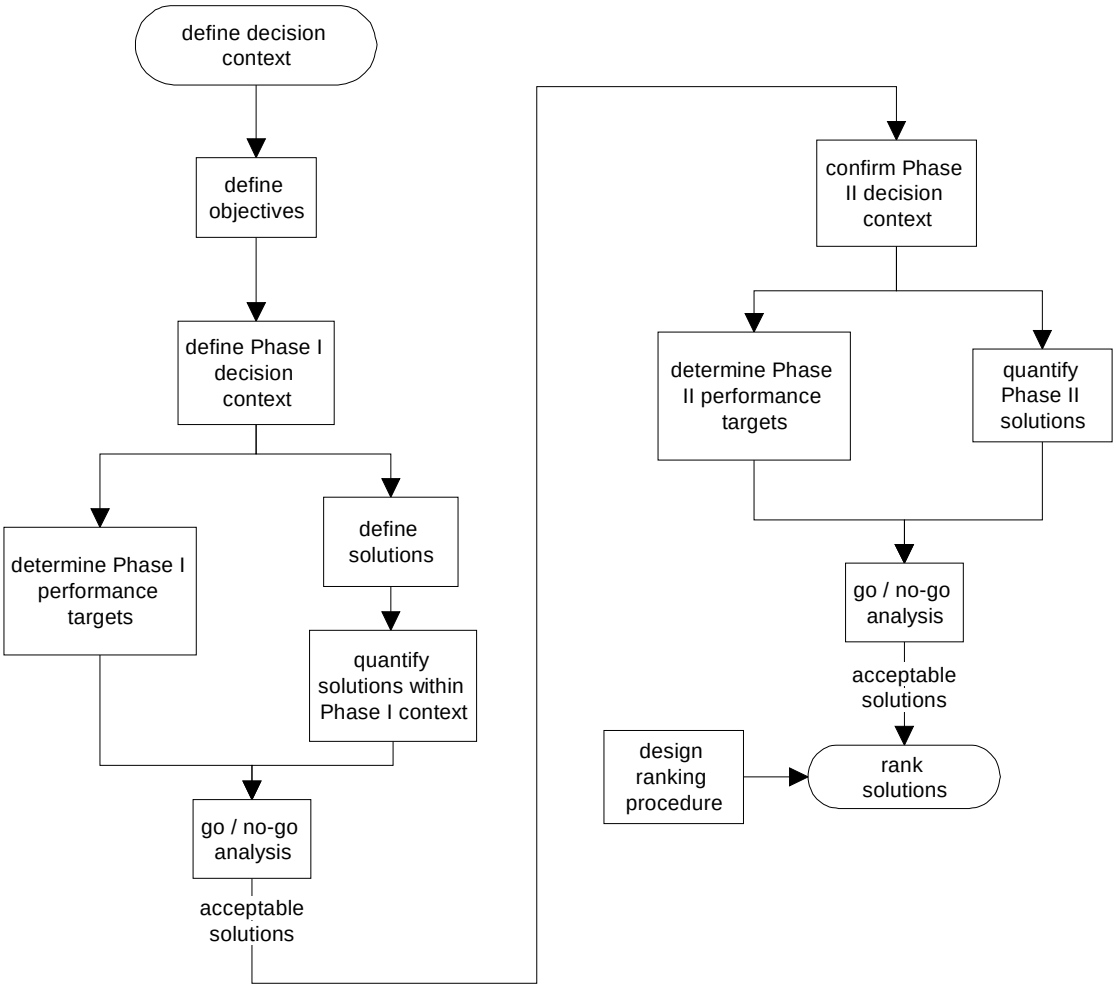
Five additional objectives must be evaluated:

- Minimize ability of uncooperative users to influence biometric acquisition
- Minimize ability of uncooperative users to influence biometric matching accuracy
- Minimize identification times
- Minimize verification times
- Maximize hardware durability

In addition, as mentioned earlier AAMVA has the option to add any or all of the variables not considered during Step 1, and to change and/or determine the values of these variables (i.e. potential search criteria, template age, etc.). In order to proceed with the quantification, AAMVA thus has to decide which variables will apply.

4.2 PHASE II

The diagram below shows a simplified version of the process as adapted for the two Phases of the DSP.



4.2.1 REVIEW DECISION CONTEXT

By default, the Phase II decision context consists of all the objectives identified in Appendix A. However, AAMVA should review the Phase II decision context to ensure that only those objectives that are applicable to the decision faced at that time are included. The review is prompted by the complexity of the quantification (see below) as well as by the differentiation between fundamental and means objectives (see Appendix D)

The quantification of a solution (in terms of the objectives) within the full Phase II decision context will be complex, and may impact the decision context. In general, the following approaches are available:

- Break down the decision context into more than one evaluation. The purpose[6] of such compartmentalizing[7] would be to allow certain aspects of each solution to be quantified independently of the rest of the solution. The advantages of this approach are:
    - Entities that are subject matter experts on only a particular portion of biometric technology (e.g. acquisition devices) can participate in the quantification, as the quantification does not require one entity (which may be knowledgeable in all aspects but not necessarily a subject mater expert in all) to perform a full quantification of a solution.
    - It is now possible to focus on smaller (and more manageable) "chunks" of the solution, decreasing the complexity (and cost) of the quantification.

  The disadvantages of this approach are:
    - Any influence between such compartmentalized quantification areas (each with its own decision context) has to be approximated, thus adding additional input variables.
    - Apart from modeling the variables with insufficient granularity or inappropriately defined ranges, some of the more subtle interdependencies between such quantification areas may be missed.
- Require a quantification of the full decision context by one entity. As pointed out above, the disadvantage of this approach is that the quantification entity may not be a subject matter expert in all areas of quantification, and that the complexity of the quantification may increase the cost thereof. The advantage is that the quantification party takes into account all dependencies between various objectives – the quantification entity is supplied with the solutions and the ranking procedure, and has to determine the best score for each solution.

A good example of how the above approaches differ is when looking at system cost vs. (1:n) search time. Under a compartmentalized approach, one entity would quantify the search time for each solution given the money available (or maybe more than one amount), and another entity would quantify the cost of each solution given certain required search times. Under a full evaluation approach, one entity would be given the solutions and the ranking procedure, and after playing off the cost and performance, will be expected to produce the maximum score for each solution.

The extent of compartmentalization to apply will depend amongst others on the data received from the industry (see Paragraph 2.7). The structure of the objective hierarchy and objective networks (see Paragraph 4.2.2) can also provide valuable insight into this matter. Note that the determination of the decision context(s) and the associated objectives is an iterative process.

4.2.2 OBJECTIVE HIERARCHY

As discussed above, the objectives (identified initially) need to be reviewed in order to ensure support for the decision context(s). A clean distinction should also be made between means objectives and fundamental

[6] Note that this would be different from the limited decision context considered in Phase I. The purpose of Phase I was to filter out non-complying solutions; the purpose of compartmentalizing the decision context in Phase II would be to decrease the evaluation effort.

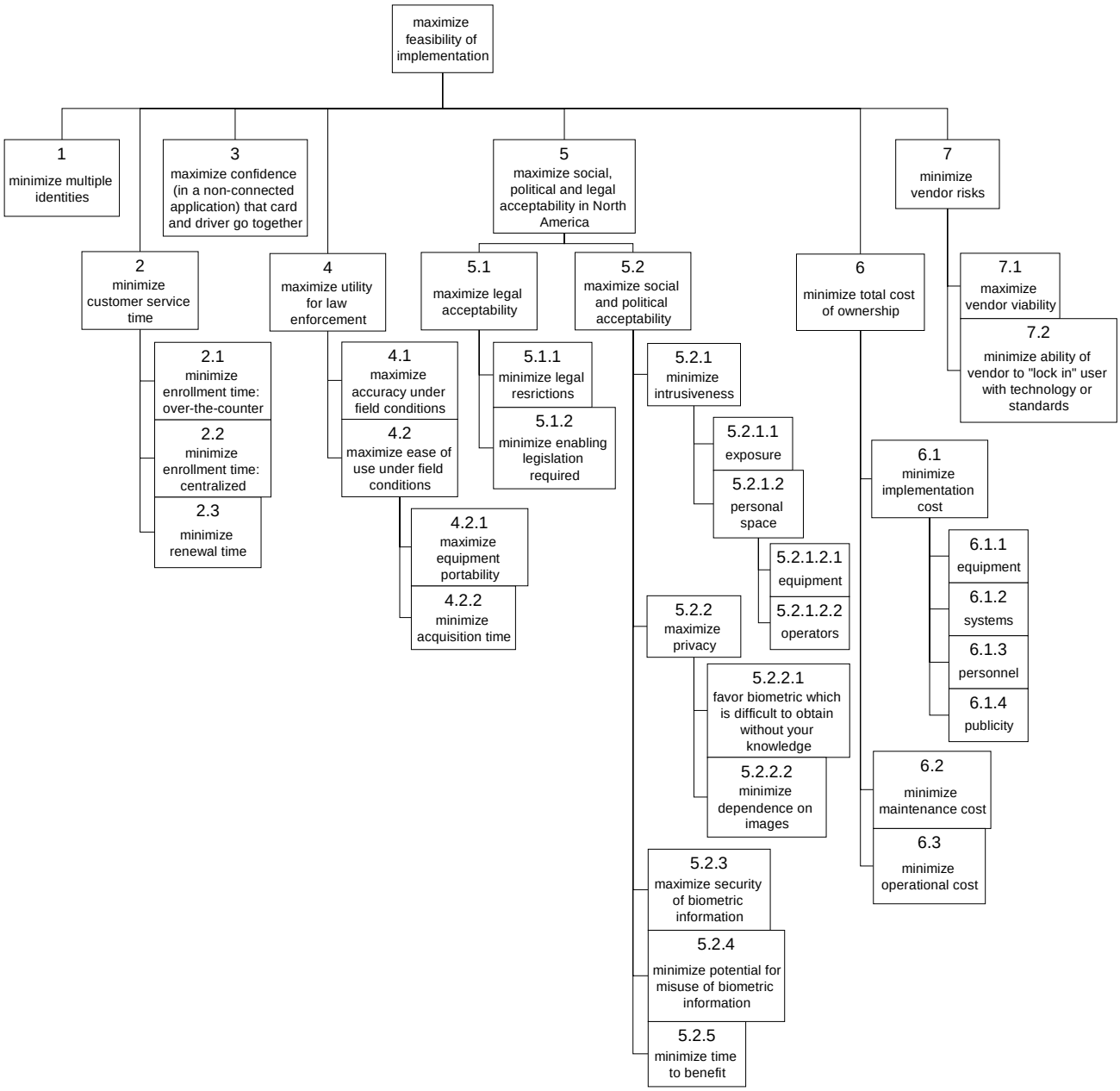
[7] Note that compartmentalization is not necessarily related to the various forms of dependence between fundamental objectives (see Paragraph 4.2.5.1). Dependence considers the value of the decision-maker; compartmentalization considers causal relationships pertaining to the evaluation of a solution.

objectives. In a nutshell, means objectives represent factual knowledge of aspects that are important to the decision (as in “a means to an end”), whereas fundamental objectives represent the values of the group (the “this is important just because this is important” reasons for doing the evaluation). The fundamental objectives are the objectives that are used in the evaluation; means objectives are used to generate alternatives.

Fundamental objectives can be presented in a fundamental objective hierarchy. In a fundamental objective hierarchy, each sub-objective can link to only one higher-level objective. Any higher-level objective that is further clarified has to have at least two lower level objectives.

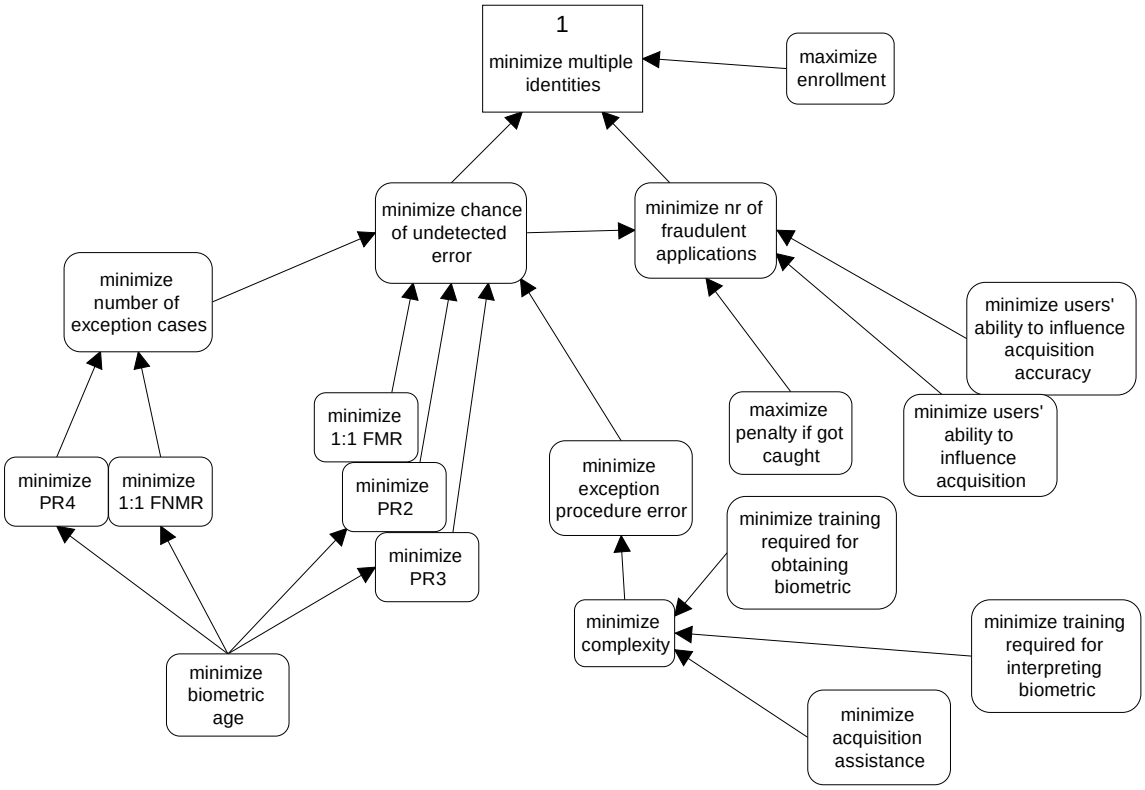
Means objectives can be presented in a means objective network. In a means objective network, each link represents a causal relationship. A means objective can be caused by only one other objective; a means objective can also cause more than one other objective.

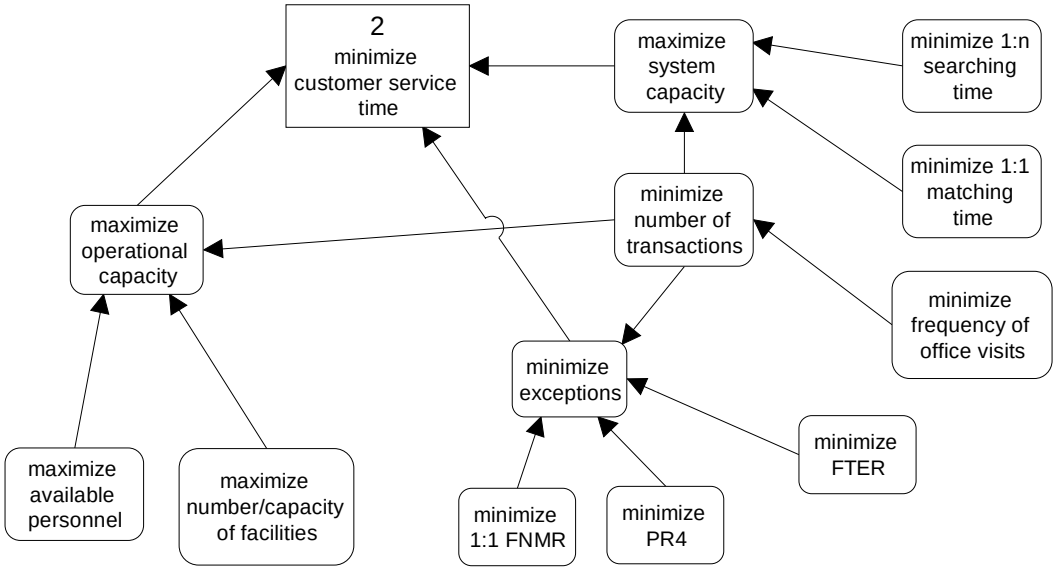
One possible manner in which the fundamental objectives can be represented in a hierarchy is provided below. Note that this representation, as well as the discussion that follows, is applicable only to a "full decision context quantification" as discussed in Paragraph 4.2.1.

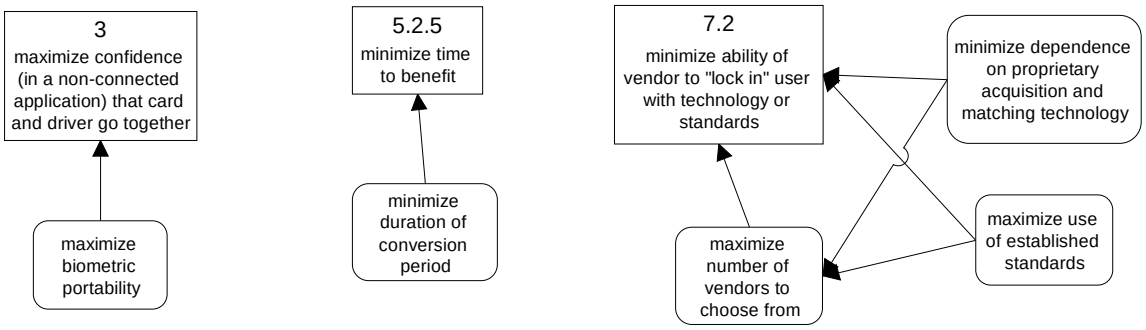


Note that in the fundamental objective hierarchy above, all objectives are (per definition) fundamental objectives.

The means objective networks associated with some of the fundamental objectives are provided below.







For each fundamental objective, a measurement scale needs to be identified or designed (see Paragraph 2.4). The table below provides additional information in respect of the fundamental objectives, and includes possible measurement scales.

Number	Description	Notes	Measurement scale
1	Minimize multiple identities		Number of expected multiple identities.[8]
2	Minimize customer service time	The customer service time is determined by 1:1 match time, 1:n search time, and acquisition time[9].	
2.1	Minimize enrollment time: over the counter		Transaction time
2.2	Minimize enrollment time: centralized[10]		Transaction time
2.3	Minimize renewal time		Transaction time

[8] This is a complex calculation, which requires a number of estimates to be made. As can be seen from the means objectives diagram, the number of multiple identities is influenced by a number of factors, and the relative contribution of each is difficult to estimate. What is needed is a model that estimates the number of multiple identities (and an associated confidence interval), given the concept of operations (including enrollment rate, population size, etc), and the performance of the biometric technology. This is one of the items included in the industry data request (see Paragraph 2.7).

Ideally, the model would be used during evaluation to estimate the number of multiple identities on a database given a number of input parameters or variables that define both the environment and the technology being evaluated. The model boils down to establishing the relationship between the number of multiple identities, and the various issues that influence this number as set out in the means objective network. The model requires knowledge of both the technical workings of a technology, and of the social dynamics that determine the number of fraudulent applications that can be expected.

In the absence of such a model, the various performance metrics can be used as a number of proxy scales (e.g. individual OSIPM, FMR, FNMR, FTER, training requirements, susceptibility to user interference, etc.) However, the challenge now becomes weighting the various scales (which essentially is what a large part of the model discussed above is all about).

[9] Where transaction times are obtained from technology testing, such times per definition exclude acquisition times. Given the size of the environment, it is considered unlikely that scenario testing will be used. Consequently, the results of technology testing can be used as a proxy for the full transaction time.

[10] Depending on whether the evaluation of centralized vs. decentralized issuing needs to be kept explicitly apart, Objectives 2.1 and 2.2 can actually be merged into one Objective, “Minimize enrollment time”. Whether or not a centralized or decentralized approach is used then becomes different alternatives to be evaluated and compared against each other.

Number	Description	Notes	Measurement scale
3	Maximize confidence (in a non-connected application) that card and driver go together	Maximize the extent to which the biometric can be stored on a DL card for purposes of off-line verification	Most satisfactory – the biometric can visually be reproduced[11] on a driver license card, and the data associated with the biometric is less than 600 bytes (capacity currently available on a PDF417 bar code after provision for non-biometric data) More satisfactory – the biometric can visually be reproduced on a driver license card, and the data associated with the biometric is less than 1200 bytes (current PDF417 storage limit) Satisfactory – the biometric can visually be reproduced on a driver license card, and the data associated with the biometric is less than x bytes (expected storage limit of technologies currently being developed) Unsatisfactory – the biometric cannot be visually reproduced on a driver license card, or the data associated with the biometric is more than x bytes.
4	Maximize utility for law enforcement		
4.1	Maximize accuracy under field conditions	Consider time of day (day vs. night), temperature, humidity, visibility, rain, snow, etc. It is suggested that more than one field condition be defined (e.g. Average, Cold & dark & wet, Hot & warm & humid & daylight), representing an "average" environment, and two extreme environments.	Average success rate for biometric acquisition for each field condition.

[11] This is used for two reasons: (1) to comply with jurisdictions who require all information stored in machine readable form, to also be human readable, and (2) to increase the capability for offline verification by way of visual inspection.

Number	Description	Notes	Measurement scale
4.2	Maximize ease of use under field conditions		
4.2.1	Maximize equipment portability		Most Satisfactory – Acquisition equipment can easily be carried with one hand, can independently (i.e. without connecting to other equipment) confirm quality of biometric acquisition, contains internal power supply, and can internally store acquired biometric. More Satisfactory – Acquisition equipment can be carried with one hand, needs to connect to external equipment to confirm quality of biometric acquisition, contains internal power supply, and can internally store acquired biometric. Satisfactory – Acquisition equipment can be carried with one hand, needs to connect to external equipment to confirm quality of biometric acquisition, relies on external power supply, and cannot internally store acquired biometric. Less Satisfactory – Acquisition equipment can be carried with two hands, needs to connect to external equipment to confirm quality of biometric acquisition, relies on external power supply, and cannot internally store acquired biometric. Least satisfactory – Can only be moved on a trolley, needs to be connected to external equipment to confirm quality of biometric acquisition, relies on external power supply, and cannot internally store acquired biometric.
4.2.2	Minimize acquisition time	Average time required to obtain a biometric of acceptable quality under field conditions.	Duration.

Number	Description	Notes	Measurement scale
5	Maximize social, legal and political acceptability in North America		
5.1	Maximize legal acceptability		
5.1.1	Minimize legal restrictions	Minimize existing legislation restricting the use of the biometric	Number of jurisdictions that have legislation restricting the use of the biometric.
5.1.2	Minimize enabling legislation required	Minimize new legislation to be drafted and tabled to enable implementation.	Number of jurisdictions that require new legislation
5.2	Maximize social and political acceptability		
5.2.1	Minimize intrusiveness	Minimize the physical intrusiveness of obtaining a person's biometric.	
5.2.1.1	Exposure	Physical contact with or exposure to biometric reading device.	Duration
5.2.1.2	Personal space	Minimize extent to which both equipment and operators intrude "personal space".	Distance of operators or equipment to the long axis of a user's body.
5.2.2	Maximize privacy		
5.2.2.1	Favor biometric which is difficult to obtain without your knowledge	Consider acquisition in the normal course of business (i.e. do not include extreme cases such as abduction etc.), with the aim of identifying the "owner" of the biometric. This objective does not consider the unnoticed acquisition of a person's biometric for purposes of falsely representing the "owner" of the biometric.	Satisfactory: Biometric cannot be obtained without your knowledge. Unsatisfactory: Biometric can be obtained without your knowledge.

Number	Description	Notes	Measurement scale
5.2.2.2	Minimize dependence on images	Images of biometric features that are stored / input into a database are more subject to theft and misuse than biometric records that are templates (algorithms that represent the biometric data, but from which the original biometric image cannot be recreated).	Satisfactory: Images are not required. Unsatisfactory: Images are required.
5.2.3	Maximize security of biometric information		Constructed scale
5.2.4	Minimize potential for misuse of biometric information		Constructed scale
5.2.5	Minimize time to benefit	The quicker the benefits of the system can be seen, the more politically acceptable the system is. True benefits will only be obtained when all drivers have been enrolled.	Duration until complete enrollment.
6	Minimize total cost of ownership		
6.1	Minimize implementation cost		
6.1.1	Minimize cost of equipment		$
6.1.2	Minimize cost of systems		$
6.1.3	Minimize cost of personnel		$
6.1.4	Minimize cost of publicity		$
6.2	Minimize maintenance cost		$
6.3	Minimize operational cost		$
7	Minimize vendor risks		

Number	Description	Notes	Measurement scale
7.1	Maximize vendor viability		Various proxy scales can be used, e.g. financial strength, track record, etc.
7.2	Minimize ability of vendor to “lock in” user with technology or standards.		Two proxy scales are applicable, i.e. the use of established standards, and the dependence on proprietary acquisition and matching technology. Note that both these also serve to maximize the number of vendors to choose from.

4.2.3 PERFORMANCE TARGETS

The purpose of the Phase II performance targets is to identify the absolute minimum performance that any solution must comply with. The performance targets are identified for those objectives that are considered crucial on their own, regardless of the value of other objectives. For example, it may be determined that no solution may result in more than 100 multiple identities in any year of operation, regardless of how inexpensive the system is.

Because of the difference between the decision contexts of Phases I and II, some of the fundamental objectives of Phase I became means objectives in Phase II. Phase II also considers objectives that were not at issue in Phase I. Consequently, as evaluation is conducted only in respect of fundamental objectives, new performance targets may have to be set.

The following fundamental objectives are contenders for review or for setting of performance targets:

- 1: Minimize multiple identities
- 2.1: Minimize enrollment time: Over-the-counter
- 2.2: Minimize enrollment time: Centralized
- 2.3: Minimize renewal time
- 6.1: Minimize implementation cost
- 6.2: Minimize maintenance cost
- 6.3: Minimize operational cost

4.2.4 SOLUTION QUANTIFICATION

The information needed to evaluate each objective can be categorized according to the source of the information as follows:

- Information to be obtained via structured testing, performed by independent 3rd parties, or from other biometric technology subject matter experts.
- Information that can be obtained by AAMVA due to its close relationship with the jurisdictions.
- Information that is dependant on current hardware, which, due to the rapid pace of progress in this area, is best obtained at bid evaluation stage. That is, some information changes so quickly that it would be unwise to base decisions on information that will have been superceded at the time when an implementation is started.

The extent to which information is available thus places an additional limitation on the evaluation of objectives. The objectives most impacted by this are those objectives whose measurement are influenced by rapidly improving technology (the third bullet above). In the case of biometric technology, acquisition equipment especially is susceptible to rapid change. If the acquisition of biometric data is not included in the evaluation, some of the measurement scales listed in Paragraph 4.2.2 need to change, or some of the objectives need to be left for later evaluation. The table below discusses these changes.

Number	Description	Information availability impact
1	Minimize multiple identities	As illustrated in the means objective diagram, the number of multiple identities is determined amongst others by the ability of users to influence the accuracy of biometric acquisition. This in turn is highly dependent on the properties of the acquisition equipment used. If an “Expected Multiple Identities” model is designed, this model thus needs to be adapted to recognize the uncertainty of the influence of acquisition equipment. If an “Expected Multiple Identities” model is not available, individual performance metrics (excluding those influenced by acquisition equipment) can be used as proxy scales instead. The challenge remains to determine the relative weight of each.
2	Minimize customer service time	In addition to the unlikely availability of scenario test information, acquisition time is directly impacted by acquisition equipment. This is another reason for using technology test results as a proxy for full transaction time.
4	Maximize utility for law enforcement	All the measurements identified are highly dependent on acquisition equipment. Consequently, this objective (and the associated sub-objectives) should only be evaluated at bid stage.
5	Maximize social, legal and political acceptability in North America	
5.2	Maximize social and political acceptability	
5.2.1	Minimize intrusiveness	This is highly dependent on the acquisition equipment, and should only be evaluated at bid stage.
6	Minimize total cost of ownership	Implementation cost can easily change over time, and hence it could be argued that the implementation cost should only be calculated at bid stage. However, it is important to at least determine if there is any difference in cost magnitude between various technologies. It is therefore suggested that the cost of an implementation be estimated based on current equipment and prices.

4.2.5 RANKING

The ranking process can more eloquently be described as determining value preferences applicable to decisions with multiple (conflicting) objectives. The value preferences of the decision-maker (AAMVA in this case) are expressed in the form of utility functions. Once the alternatives are ranked according to their utility, a sensitivity analysis is performed.

The sub-paragraphs that follow are not intended as a theoretical study in decisions with multiple objectives. Rather, the theoretical component is only mentioned where necessary. A detailed explanation of the assumptions and conditions underlying the techniques discussed here can be found in literature dealing with Decision Analysis.

4.2.5.1 UTILITY FUNCTIONS

The goal is to calculate the relative utility (a number) of each alternative, and then to use this utility to compare the alternatives against each other. The utility of an alternative is calculated using the overall utility function. This utility function represents the contribution of each objective to the overall utility of a solution. Ideally, the overall utility function can be expressed in terms of utility functions for each of the objective attributes (i.e. measurement scales)[12].

The level of independence between objective attributes is crucial in establishing if and how the utility function can be expressed in terms of utility functions for each of the objectives[13]. Different forms of independence are:

- Utility independence
- Preferential independence
- Additive independence

The type of independence between objectives is also dependent on the uncertainty associated with the measurement of each objective for each alternative. If the uncertainty is so small that it is not considered meaningful, the outcome of the alternative with respect to that particular objective can be regarded as a certainty. If however the outcome is clearly uncertain, a probability distribution of the possible outcomes needs to be constructed.

Provided that additive independence (the most restrictive type of independence) holds, the utility function becomes a linearly weighted function[14]. Slightly more complex forms of the overall utility function apply when other forms of independence hold instead.

---

[12] That is,

$U(x_1, x_2,\ldots,x_n) = f[U_1(x_1), U_2(x_2),\ldots, U_n(x_n)]$,

where

f is a scalar-valued function,

x_i is a specific amount of X_i

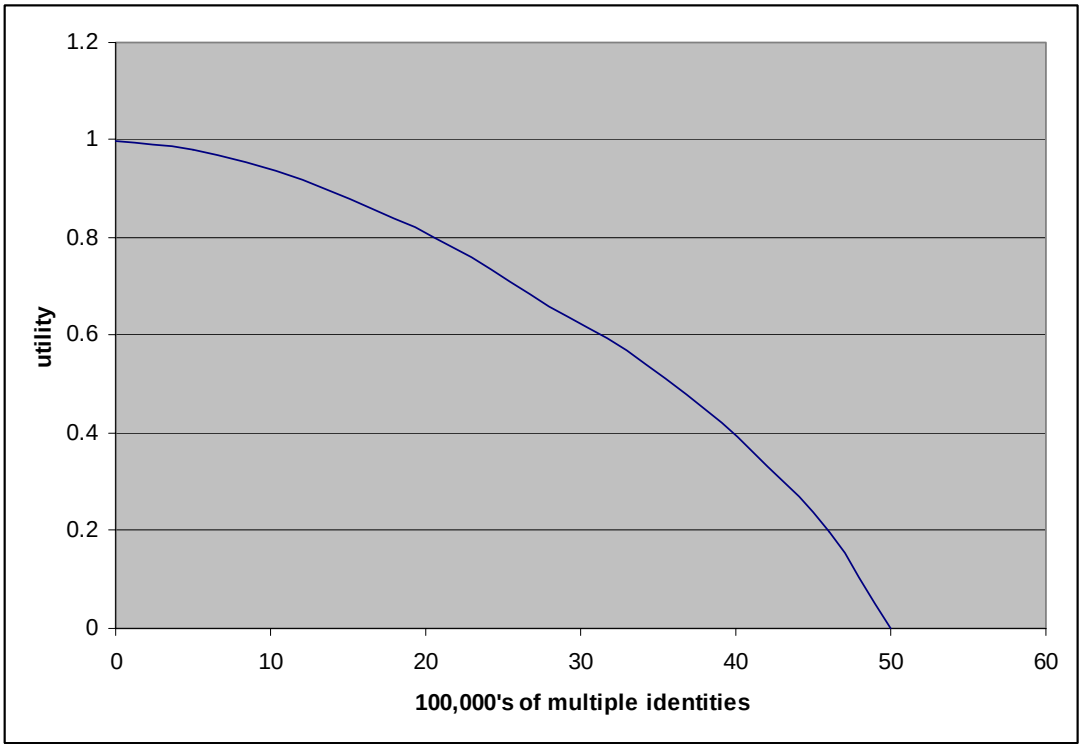
X_i, $i = 1,\ldots,n$ are the objective attributes

U_i is a utility function over X_i

[13] That is, if and how $U(x_1, x_2,\ldots,x_n)$ can be expressed in terms of $U_i(x_i)$

Assuming that the independence conditions allow the utility function to be expressed in terms of utility functions for each of the objectives[15], the next step is to construct the utility functions for individual objective attributes. These utility functions essentially describe the importance of the different levels of a measurement scale as they relate to one another. Its aim is to capture the decision-maker's values, and specifically the decision-maker's attitude towards risk.

A utility function is normally defined over the probable values a particular objective attribute can reasonably assume, and takes on values over an arbitrary range, often between 0 and 1. The following is an example of a utility function for the number of multiple identities allowed onto the system during initial enrollment.



The above graph represents the utility of various levels of multiple identities. For 0 multiple identities, the utility value is 1. For 5,000,000 multiple identities, the utility value is 0. Note that a decrease of 1,000,000 multiple identities is regarded as more valuable if it decreases the total number of multiple identities from 5,000,000 to 4,000,000 than when it decreases the total number of multiple identities from 2,000,000 to 1,000,000. The above is just an example, and the actual form of UID9's utility function for the indicated objective attribute will have to be established.

Various techniques exist to assess the parameters of the utility functions. For example, if additive independence holds for the overall utility function, swing weighting can be used to determine the weight contribution for each objective. This method is sensitive for the range of values an attribute takes on. Another often used method of assessing parameters is lottery weighting. Various forms of lottery weighting can be employed, the basic approach being to find a value *p* that would make the decision-maker indifferent between:

[14] That is, a function of the form $U(x_1, x_2,\ldots,x_n)=c_1U_1(x_1) + c_2U_2(x_2) + \ldots + c_nU_n(x_n)$

[15] That is, assuming that $U(x_1, x_2,\ldots,x_n)$ can be expressed in terms of $U_i(x_i)$

1. Achieving outcome **B** for certain, and
2. A lottery with probability *p* on outcome **A** and probability *1-p* on outcome **C**.

An example of a questionnaire that uses lottery weights to assist decision-makers in thinking hard about their preferences for various biometric error rates can be found in Appendix D.

The following is an example of an overall utility function:

$$U(x_1,x_2,x_3,x_4,x_5,x_6,x_7) = .31U_1(x_1) + .11U_2(x_2) + .09U_3(x_3) - .18U_4(x_4) + .16U_5(x_5) + .07U_6(x_6) + .08U_7(x_7)$$

Where

$$U_1(x_1) = .976[1-e^{.13x_1}]$$

$$U_2(x_2) = .102[1-.3e^{.07x_2} - .3e^{.12x_2}]$$

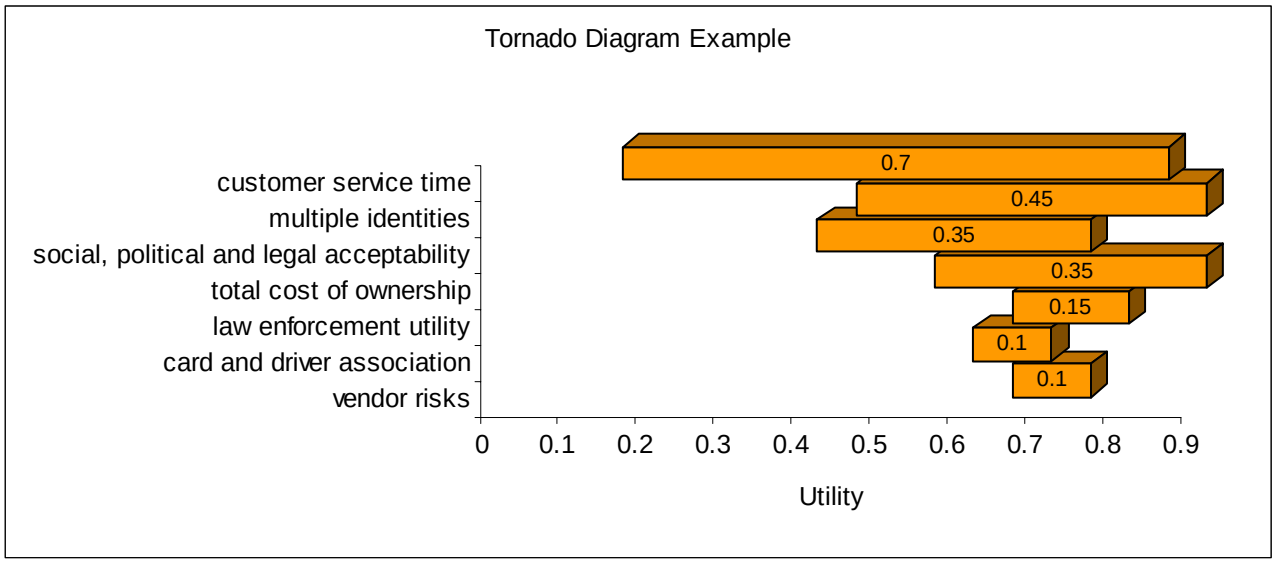
…

$$U_7(x_7) = .233[1-e^{.09x_7}]$$

4.2.5.2 SENSITIVITY ANALYSIS

Once an overall utility has been calculated for each alternative, it is prudent to conduct a sensitivity analysis on the assessment. The purpose of a sensitivity analysis is to identify aspects of the calculations that may be particularly sensitive to changes in the underlying parameters and/or input values. Depending on the extent of sensitivity, it may be necessary to give additional attention to these areas.

One way in which to conduct a sensitivity analysis, regardless of the form of the overall utility function, is to use so-called tornado diagrams. Tornado diagrams, while not error-proof, generally provide a good representation of those issues for which a utility function is sensitive to. It is a simple process whereby each attribute in turn is varied between the largest and smallest values, whilst all the other attributes are kept fixed. The result on the overall utility is then plotted, as in the following example.



In the example, customer service time is the one objective to which the overall utility is most sensitive. Depending on the granularity of the customer service time measurement scale used in the objective's utility function, additional attention to customer service time assessment may be beneficial.

The sensitivity analysis can also be refined with the use of two-way sensitivity graphs. This is especially useful to visualize the sensitivity to uncertain inputs. The drawback of this technique is that it is limited to the consideration of two variables at a time.

5. ROADMAP USE BY JURISDICTIONS

As mentioned earlier, the roadmap put forth in this document can serve as a framework for jurisdictions to evaluate the use of biometric technology at a jurisdictional level. The following aspects of the roadmap needs to be reviewed:

- Re-frame the decision context. The current over-arching decision context considers the use of biometric technology on a continent-wide basis to ensure "one driver, one driver license, one driver record." However, this is only feasible if all jurisdictions are involved. Whilst the use of biometric technology at jurisdictional level may have the same purpose, it can also be used as an investigative tool or a tool to assist in cleaning up a database, rather than as a true unique identifier. The scale of the implementation is also different (from the scale assumed in this document). Hence, the decision context has to be adapted to reflect the jurisdiction's environment.
- Once the decision context has been drafted, it is necessary to ensure that the associated fundamental objective hierarchy and means objective networks support the decision context.

APPENDIX A: INITIAL OBJECTIVES

1 INTRODUCTION

The stated goal of AAMVA's UID9 Task Group is to develop a way to uniquely identify an individual such that:

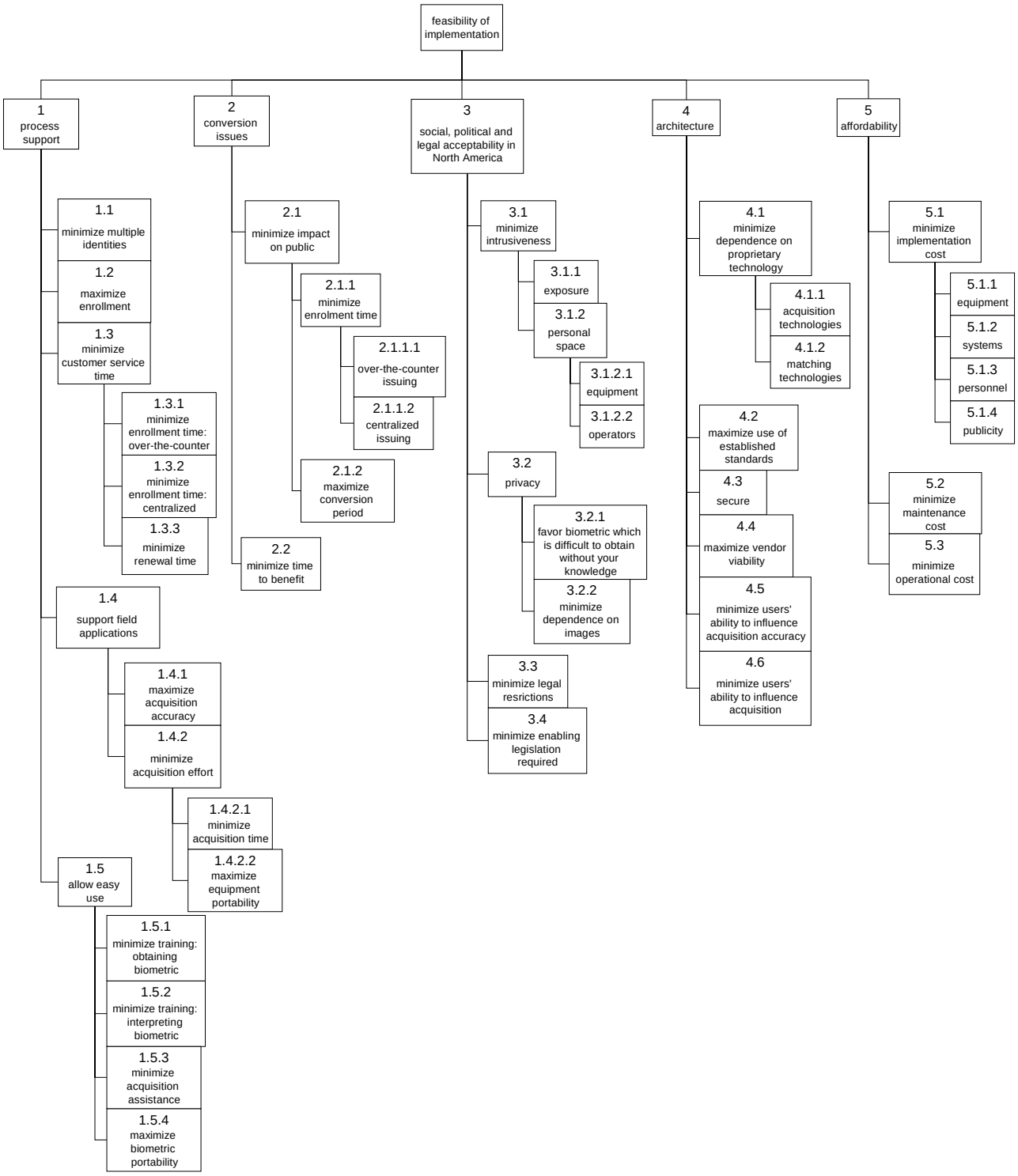
- One person will only have one driver record on the database and only one driver license (DL)/ identification (ID) card in North America.
- Authorized users can verify that the holder of a DL/ID card is the individual to whom the card was issued.

This Appendix considers the issues to be considered when deciding whether or not it is feasible to use a biometric technology to meet these requirements.

2 OBJECTIVES AND UNITS OF MEASUREMENT

When investigating the feasibility of implementing a biometric or combination of biometrics to achieve the goal stated above, various objectives (some of which may be conflicting) need to be considered.

The diagram that follows reflects the initial categorization of the objectives identified by UID9. Although specifically designed for application in a 300 million record environment, most of the objectives are also applicable to smaller environments such as are encountered by individual jurisdictions. A table discussing the objectives follows the diagram.



Objective #	Objective description	Notes
	Confirm feasible implementation	"Feasible" as used here is qualified by the next level objectives.
1	Process support	Operational processes required for the function of issuing and using driver licenses need to be supported.
1.1	Minimize multiple identities	See Appendix C for a discussion of this objective
1.2	Maximize enrollment	See Appendix C for a discussion of this objective
1.3	Minimize customer service time	Both over-the-counter and centralized issuing is to be considered. Jurisdictions follow different processes, and hence the time an applicant spends in an office to obtain a DL varies. As a result, this objective will be measured by the duration added to an applicant's stay in the DL office. As some jurisdictions already perform some checks against "outside" databases, and these checks are to be supplemented by the biometric check, the measured duration thus becomes an upper limit on the duration added to an applicant's stay in the DL office (i.e. at least part of the time required for the 1:n search may occur in parallel with existing processes). This objective considers customer service time under steady state conditions.
1.3.1	Minimize enrollment time: Over-the-counter issuing	Typical process assumed to consist of reading biometric, performing 1-n verification, and processing of the results in real time. The time involved is influenced by various error rates, enrollee application rate, the process followed to resolve false matches, and the availability of resources (both human resources and system resources)(see Appendix C).
1.3.2	Minimize enrollment time: Central issuing	Typical process assumed to consist of reading biometric, and issuing some type of proof of application. Assuming that the issuing of the proof of application will be independent of the type of biometric technology used, the time required to do this becomes irrelevant for the decision at hand. In the mean time, the 1-n verification, and the processing of the subsequent results are performed at a central facility. The work performed at the central facility does not add to the customer service time.
1.3.3	Minimize renewal time	Typical process assumed to consist of reading biometric, performing 1-1 authentication, and processing of the results in real time. The customer service time is assumed to be the same for centralized and over-the-counter issuing.
1.4	Support field applications	Field applications essentially consist of two aspects, i.e. biometric acquisition, and the verification process. As the verification process is already addressed above, this objective aims to maximize the effectiveness and efficiency of the biometric acquisition.

Objective #	Objective description	Notes
1.4.1	Maximize acquisition accuracy under field conditions.	Consider time of day (day vs. night), temperature, humidity, visibility, rain, snow, etc. It is suggested that more than one field condition be defined (e.g. Average, Cold & dark & wet, Hot & warm & humid & daylight), representing an “average” environment, and two extreme environments.
1.4.2	Minimize effort of acquiring biometric under field conditions.	
1.4.2.1	Minimize acquisition time	Average time required for a successful biometric acquisition for each field condition defined.
1.4.2.2	Maximize equipment portability	Portability of acquisition equipment.
1.5	Allow easy use	
1.5.1	Minimize amount of training required for operator to become proficient in obtaining a person’s biometric	
1.5.2	Minimize amount of training required for operator to become proficient in interpreting biometric search results.	Consider 1:1 and 1:n training separately if training for these two cases is presented separately.
1.5.3	Minimize the extent of assistance and supervision required to acquire biometric.	“Trained staff” as used here implies rudimentary knowledge by frontline staff of the biometric acquisition process and minor training in "subversive tactics" to look for – methods commonly employed to dupe a biometric reader, or common user mistakes resulting in poor acquisition. “Little supervision” as used here means that enrollment is done within plain sight of a staff person. “Light supervision” as used here means that enrollment is done within plain sight of, and at the verbal direction of, a staff person. “Direct supervision” as used here means that enrollment is done under physical supervision by a staff person, i.e. there is physical contact between the enrollee and the staff person.
1.5.4	Maximize biometric portability	Maximize the extent to which the biometric can be stored on a DL card for purposes of off-line verification. Another portability issue concerns the ability to read and use biometric data across different platforms. Biometric information on driver license card must be readable by other vendors’ equipment and applications. This assists in wide use, as well as in allowing jurisdictions a choice of vendors at contract renewal time. Cross-platform use essentially depends on the extent to which standards are open, and is addressed as part of Objective 4.1.
2	Conversion issues	

Objective #	Objective description	Notes
2.1	Minimize impact on public	
2.1.1	Minimize enrollment time	Additional time an individual has to spend in the office in order to become enrolled. This objective considers the time during the initial enrollment period, and thus implies a higher enrollee application rate than is applicable to Objective 2.1.1. Both over-the-counter and centralized issuing is considered.
2.1.1.1	Minimize enrollment time: Over-the-counter issuing	Typical process assumed to consist of reading biometric, performing 1-n verification, and processing of the results in real time. The time involved is influenced by various error rates, enrollee application rate, the process followed to resolve false matches, and the availability of resources (both human resources and system resources)(see Appendix C).
2.1.1.2	Minimize enrollment time: Central issuing	Typical process assumed to consist of reading biometric, and issuing some type of proof of application. In the mean time, the 1-n verification, and the processing of the subsequent results are performed at a central facility. The work performed at the central facility does not add to the customer service time.
2.1.2	Maximize period of enrollment	Time required for full population to be enrolled. The longer this period is extended, the more time the public will have to enroll. Resource requirements in DMV offices will be lower, as the load is spread out over a longer time.
2.2	Minimize time to benefit.	As actual benefit (i.e. fraud prevented) is not currently measured, the time to full implementation is to be measured instead. Note that this objective is in direct opposition to the preceding objective. The evaluation process will assign a relative importance to each objective.
3	Social, political and legal acceptability in North America	
3.1	Minimize intrusiveness	Minimize the physical intrusiveness of obtaining a person's biometric.
3.1.1	Exposure	Physical contact with or exposure to biometric reading device.
3.1.2	Personal space	Minimize extent to which both equipment and operators intrude "personal space".
3.1.2.1	Equipment	Minimize proximity of equipment to the long axis of a user's body.
3.1.2.2	Operators	Minimize proximity of operators to the long axis of a user's body
3.2	Privacy	Note that the privacy objectives included here only consider the difference between various biometric technologies. A decision on whether or not to use any biometric technology (versus not using a biometric technology at all) will be influenced by a slightly different set of objectives.

Objective #	Objective description	Notes
3.2.1	Favor a biometric technology which is difficult to obtain without your knowledge	Consider acquisition in the normal course of business (i.e. do not include extreme cases such as abduction etc.), with the aim of identifying the "owner" of the biometric. This objective does not consider the unnoticed acquisition of a person's biometric for purposes of falsely representing the "owner" of the biometric.
3.2.2	Minimize dependence on biometric images	Images of biometric features that are stored / input into a database are more subject to theft and misuse than biometric records that are templates – algorithms that represent the biometric data, but from which the original biometric image cannot be recreated. Storage of the original biometric image for manual verification purposes is an operational issue, and thus independent from the type of technology used.
3.3	Minimize legal restrictions	Existing legislation restricting the use of the biometric.
3.4	Minimize enabling legislation required	New legislation may need to be drafted and tabled to enable implementation
4	Architecture	
4.1	Minimize dependence on proprietary technology	The ability to read and use biometric data across different platforms (i.e. by various vendors' equipment and applications) facilitates wide use, and allows jurisdictions a choice of vendors at contract renewal time.
4.1.1	Acquisition technologies	Minimize the dependency on a particular technology for biometric acquisition
4.1.2	Matching technologies	Minimize the extent to which proprietary standards are required to interpret biometric data.
4.2	Maximize use of established standards; minimize use of non-established standards.	Standardization can be viewed a few ways: at the API level, the data format level, the raw sample level, and the template level. Each of these standards levels provides a progressively greater degree of interoperability, while raising new challenges (e.g. privacy, performance). A challenge in this area is that large-scale 1:N systems are among the least amenable to many types of standardization.
4.3	Secure	The more proprietary a technology is, the more secure it is (often referred to as "security through obscurity"), because there is less expertise available. However, the advantage of an open standard is that it has been tested thoroughly, which is not necessarily the case with a proprietary standard.
4.4	Maximize vendor viability	Favor vendors with longer experience in supporting the biometric technology
4.5	Minimize ability of uncooperative users to influence biometric matching accuracy	
4.6	Minimize ability of uncooperative users to influence biometric acquisition	
5	Affordability	

Objective #	Objective description	Notes
5.1	Minimize implementation cost	
5.1.1	Equipment	Cost of equipment to be procured.
5.1.2	Systems	Cost of systems required for the use of the biometric.
5.1.3	Personnel	Cost of additional personnel and/or training required.
5.1.4	Publicity	Cost of publicity informing customers
5.2	Minimize maintenance cost	Include information on hardware durability – see Objective 1.2.4.
5.3	Minimize operational cost	Specifically include the cost associated with jurisdictional resources required to work through suspect lists generated by a 1:n search.

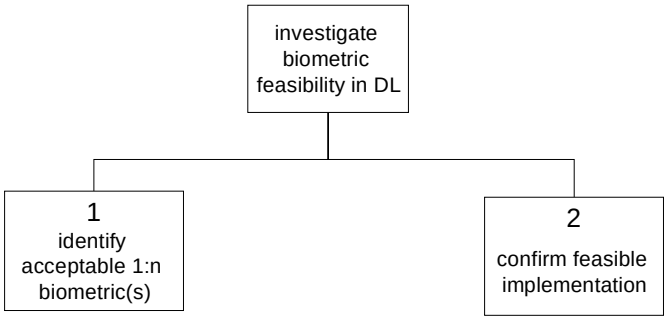
APPENDIX B: PHASE I OBJECTIVES AND VARIABLES

To avoid spending time and resources on the evaluation of the functionality and feasibility of implementing biometric technologies that may not meet the minimum technical requirements, the Project was divided into two Phases as follows:

Phase I: Identify a biometric or combination of biometric technologies that will be able to perform a 1 to many comparison in a database of 300 million records (Scenario 1); and

Phase II: Once step 1 has been completed, evaluate the functionality and feasibility of implementing the identified biometric(s) (Scenario 2).

This can graphically be represented as follows:

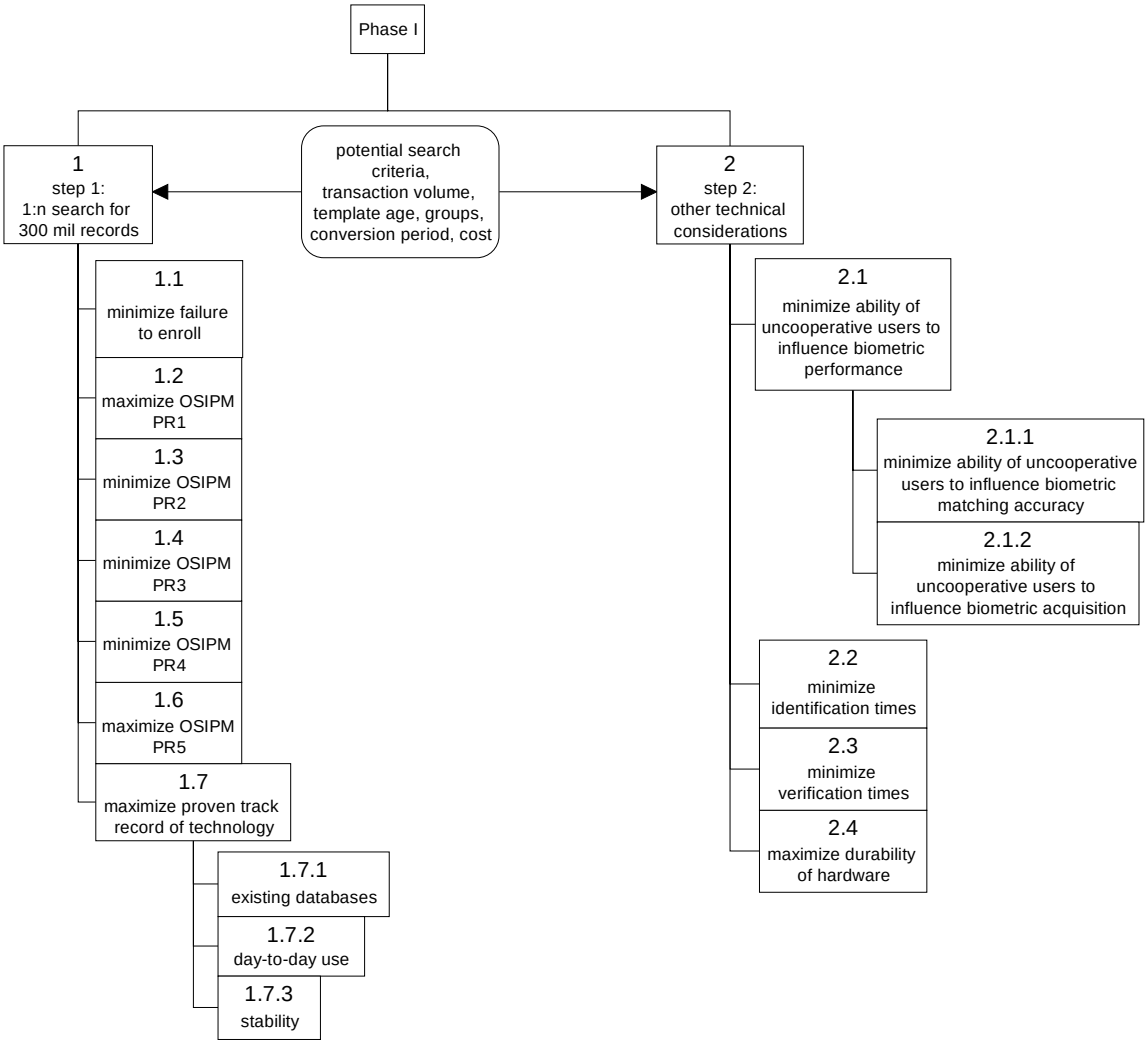


To further manage the resources available for the Project, Phase I was divided into two steps as follow:

Step 1: Confirm ability to perform a one to many (1:n) search (Scenario 1).

Step 2: Confirm other technical considerations (Scenario 2).

The objective hierarchy for the two scenarios can graphically be represented as follows:



In the diagram above, the block with the rounded corner represents input variables, and the other blocks represent objectives.

The table on the following pages provides additional detail in respect of the above objectives.

Table B.1: Objectives			
Objective #	**Objective description**	**Notes**	**Measurement scale**
	Identify acceptable 1:n biometric(s)	Limited to certain crucial technical properties of the biometric(s). Any biometric or combination of biometrics can be included Objective 2 is to be addressed only once the evaluation of Objective 1 has provided a satisfactory outcome.	N.A.
1	1:n search for 300 million records	Four factors inherent to a biometric characteristic contribute to a biometric system's failure to enroll (FTE), false match rate (FMR), and false non-match rate (FNMR)(although the FMR and FNMR are used here, the information is easily applied to the OSIPM – refer to Refer to Appendix C for an in depth discussion). These factors are availability (the percentage of individuals who have such a characteristic), stability (a characteristic's resistance to temporal or intentional changes), uniqueness (a characteristic's distinctiveness relative to the population), and measurability (the ability of a characteristic to be measured). • Availability impacts FTE and FNMR directly, FMR indirectly • Stability impacts FMR and FNMR directly, FTE indirectly • Uniqueness impacts FMR directly, FNMR indirectly • Measurability impacts FTE and FNMR directly, FMR indirectly The fifth factor that contributes to a biometric system's FTE, FMR, and FNMR is the state of the technology: how well sensors and algorithms can leverage a characteristic's availability, stability, uniqueness, and measurability. The FTE, FMR, and FNMR percentages thus are reflective of current technology capabilities and of the inherent nature of the biometric characteristics. The evaluation results for the sub-objectives of Objective 1 represent collective performance at a single security / enrollment threshold – these metrics cannot be defined independently of each other.	N.A.
1.1	Minimize failure to enroll	Refer to Appendix C for an in depth discussion of failure to enroll.	Expressed as a percentage of all people in North America older than 16 years of age

Table B.1: Objectives			
Objective #	**Objective description**	**Notes**	**Measurement scale**
1.2	Maximize OSIPM Potential Result 1	Refer to Appendix C for a detailed discussion of the OSIPM.	%
1.3	Minimize OSIPM Potential Result 2		%
1.4	Minimize OSIPM Potential Result 3		%
1.5	Minimize OSIPM Potential Result 4		%
1.6	Maximize OSIPM Potential Result 5		%
1.7	Maximize proven track record of technology	The technology must be **in practical use** and have a track record, **although not at the scale required** for DMV's. The focus is on identifying technologies that work in the field today. However, sight should not be lost of the experimental improvements to such technologies that are being developed today.	
1.7.1	Favor technologies with larger existing databases	Must be used for 1:n verification. Actual accuracy will not be considered – the physical existence and intended use of a database will be sufficient for inclusion.	Number of enrollee records in database
1.7.2	Favor technologies that have been implemented for day-to-day use	Not a "pilot study" or 1-month trial period. The purpose is to have technologies that have been in operation in the "real world" and hopefully have developed a record of lessons learned, established predictability in terms of accuracies and inaccuracies, and that may have existing proven exceptions processes that may be adopted by jurisdictions.	Number of implementations; number of implementations in environments / for purposes similar to jurisdictional functions; collective age (in years) of number of implementations.

Table B.1: Objectives			
Objective #	**Objective description**	**Notes**	**Measurement scale**
1.4.3	Maximize stability of core acquisition and matching technologies	Maximize the degree to which the basic principles of extraction and matching logic are implemented in a standardized fashion for a given technology. If additional definition of each measurement level is required, the inclusion of the difference between the "version" of the system/enrolment devices in operation vs. being shipped for new installations can be considered. A large difference may imply an unstable system/device; however, it may also imply rapid development. Note that comparison between different vendors may also be difficult due to different version numbering practices.	Most Satisfactory - Highly stable core technology, supported by leading technology providers for over five years, that leverages highly mature imaging and/or matching methods and standards More Satisfactory - Stable core technology, supported by leading technology providers for under five years, that leverages reasonably mature imaging and/or matching methods and standards Satisfactory - Fairly stable core technology, supported by leading technology providers for under two years, that leverages emerging imaging and/or matching methods and standards Unsatisfactory – Fairly unstable or only recently established core technology, with few well-defined or standardized approaches to biometric acquisition and/or matching.
2	Other technical considerations	This objective will evaluate certain very important properties of the biometric solution that passed Objective 1.1. The importance of these objectives stems from the fact that non-compliance will render any further evaluation moot. Objective 1.2 is to be addressed only once the evaluation of Objective 1.1 has provided a satisfactory outcome.	N.A.
2.1	Minimize ability of uncooperative users to influence biometric performance	How successful is the biometric solution in identifying users in spite of uncooperative behavior? An "uncooperative" user is a user who is on purpose trying to influence the biometric, and excludes users who may influence the biometric due to conditions outside their control, e.g. physiological limitations. Depending on the technologies being evaluated, consideration can be given to combining objectives 2.1.1 and 2.1.2.	

Table B.1: Objectives			
Objective #	**Objective description**	**Notes**	**Measurement scale**
2.1.1	Minimize ability of uncooperative users to influence biometric matching accuracy	The distinction between Objectives 1.2.1.1 and 1.2.1.2 is necessary because the action taken by a jurisdiction subsequent to a non-acquisition may differ from the action taken subsequent to a non-match. “Trained staff” as used here implies rudimentary knowledge by frontline staff of the biometric acquisition process and minor training in "subversive tactics" to look for – methods commonly employed to dupe a biometric reader, or common user mistakes resulting in poor acquisition. “Little supervision” as used here means that enrollment is done within plain sight of a staff person. “Light supervision” as used here means that enrollment is done within plain sight of, and at the verbal direction of, a staff person. “Direct supervision” as used here means that enrollment is done under physical supervision by a staff person, i.e. there is physical contact between the enrollee and the staff person.	Most Satisfactory - Matching accuracy cannot be influenced by uncooperative users; no supervision required to prevent influencing More Satisfactory - Matching accuracy can only be influenced by a user highly familiar with the operations of the biometric technology; little to no supervision required by trained staff to prevent influencing Satisfactory - Matching accuracy cannot be easily influenced by a user, even one with some familiarity with the operations of the biometric technology; light supervision required by trained staff to prevent efforts to influence accuracy Unsatisfactory – Matching accuracy can be influenced by a typical user with little or no effort regardless of technical background of the technology; direct supervision required by trained staff to prevent efforts to influence accuracy

Table B.1: Objectives			
Objective #	**Objective description**	**Notes**	**Measurement scale**
2.1.2	Minimize ability of uncooperative users to influence biometric acquisition		Most Satisfactory - Acquisition cannot be influenced by uncooperative users; no supervision required to prevent influencing More Satisfactory - Acquisition can only be influenced by a user highly familiar with the operations of the biometric technology; little to no supervision required by trained staff to prevent influencing Satisfactory - Acquisition cannot be easily influenced by a user, even one with some familiarity with the operations of the biometric technology; light supervision required by trained staff to prevent efforts to influence accuracy Unsatisfactory – Acquisition can be influenced by a typical user with little or no effort regardless of technical background of the technology; direct supervision required by trained staff to prevent efforts to influence accuracy
2.2	Minimize identification times	Time required to perform a search for one biometric record on a database of 300 million records. This objective allows for the use of multiple biometric technologies within one solution.	Duration
2.3	Minimize verification times	Time required to perform a 1:1 match on a database of 300 million records.	Duration
2.4	Maximize durability of hardware	Expected life of equipment, assuming normal office conditions.	Time

Table 2 defines the input variables applicable to Phase I.

Table B.2: Input variables				
Variable #	**Input variable description**	**Notes**	**Variable values**	**Objectives influenced**
1.1	Potential search criteria			
1.1.1	Gender-based search functions		% of database that can be reliably excluded from a 1:N search based on exogenous characteristics. Gender-based searches should reliably reduce by approximately 50% the size of the target database a 1:N search. This type of filtering assumes that the operator classifies the person correctly.	1.1.2, 1.1.3, 1.2.2
1.1.2	Age range search functions		% of database that can be reliably excluded from a 1:N search based on exogenous characteristics. Age range search functions may be able to reduce by approximately 50% the size of the target database a 1:N search, but this needs to be evaluated (it may not be viable). If we assume that all imposters will lie about their age, then all that can be established are ranges (e.g. someone who appears to be 20 cannot enroll as a 60 year old but could as a 30 year old with fraudulent documentation). It is uncertain whether the potential for classification errors that occur here would reduce overall accuracy.	1.1.2, 1.1.3, 1.2.2

Table B.2: Input variables				
Variable #	**Input variable description**	**Notes**	**Variable values**	**Objectives influenced**
1.1.3	Geography search functions		% of database that can be reliably excluded from a 1:N search based on exogenous characteristics. The range of geographical search reduction is unknown. Clearly an enrollee in California does not need to be searched against the same day's enrollees in Maine. However at some point it seems that the system may need to be capable of comparing everyone against everyone	1.1.2, 1.1.3, 1.2.2
1.1.4	Temporal search functions		% of database that can be reliably excluded from a 1:N search based on exogenous characteristics. The range of temporal search reductions is unknown - the dynamic is similar to that of Variable 1.1.3 above (time and geography may be part of the same consideration).	1.1.2, 1.1.3, 1.2.2
1.2	Transaction volumes			
1.2.1	Enrollees per day	Experience has shown that people generally put off such an enrolment until the last possible moment, implying a much larger than average number of enrollees per day towards the end of the enrolment period. The rate is also dependent on the period of time allowed for all persons to be enrolled in the database (See Variable 1.5), and on the success with which this is communicated to the public. • For purposes of this evaluation, the peak daily enrollment rate is estimated to be 1.3 times the average enrollment rate. The assumptions upon which this figure is based, are discussed in detail in Appendix C.	# of daily enrollees (peak)	1.1.2, 1.1.3, 1.2.3

Table B.2: Input variables				
Variable #	**Input variable description**	**Notes**	**Variable values**	**Objectives influenced**
1.2.2	Size of database	The database size increases as more individuals are enrolled. Assume that no legacy data will be used.	# of enrollees: 75 million, 150 million, 300 million	1.1.2, 1.1.3, 1.2.3
1.3	Groups susceptible to increased enrollment and matching error rates for certain biometric technologies	Demographic, race, age and other groups for which it is known that biometrics are sensitive. Note that this variable will be used for evaluation purposes only, to consider the inherent capabilities of the biometric technology. It is not the intention to utilize such groupings as search criteria or as the basis for data segmenting (i.e. “profiling”). The following are generally known sensitivities of the leading biometric technologies: • Fingerprint: suspected susceptibility to increased enrollment (FTE) and matching (FNMR) error rates attributable to manual laborers; race (primarily east Asian); gender (females more susceptible than males); age (elderly more susceptible than younger demographics) • Facial recognition: suspected susceptibility to increased enrollment (FTE) and matching (FNMR) error rates attributable to gender (females more susceptible than males), age (younger demographics more susceptible than elderly); ethnicity (various reports of increased error rates for populations who comprise a small percentage of a main database, e.g. Black, Asian); individuals with eyeglasses and/or facial hair • Iris recognition: suspected susceptibility to increased enrollment (FTE) and matching (FNMR) error rates attributable to individuals with eyeglasses; also suspected susceptibility for individuals with very dark or very light eyes.	Occupation (manual laborer or not) Race (east Asian, Asian, Black, Other) Gender (male, female) Age (younger than 50, 50 and older) Eyesight (wearing glasses or not) Eye color (very light, very dark, other)	1.1.2, 1.1.3, 1.1.1

Table B.2: Input variables				
Variable #	**Input variable description**	**Notes**	**Variable values**	**Objectives influenced**
1.4	Age of biometric template		One month, One year, Three years	1.1.3
1.5	Enrollment period	Time allowed for full population to be enrolled on the system.	4 years, 8 years, 12 years	1.2.1
1.6	Central processing capabilities cost	Inclusive of hardware and software to execute large-scale matching. Does not include costs associated with acquisition devices (scanners, etc.) at the DMV level or cost associated with infrastructure upgrades necessary to get data from points A, B, and C to point D.	$25m, $50m, $75m	1.2.2

APPENDIX C: PHASE I PERFORMANCE TARGETS

1 INTRODUCTION

The purpose of this Appendix is to discuss the process and reasoning whereby UID9 arrived at the Phase I performance targets. The objectives involved are the following:

- The Open Set Identification Performance Metrics (OSIPM) (5 objectives)
- Proven track record of a technology (subdivided into 3 objectives)
- The ability of uncooperative users to influence biometric performance (subdivided into 2 objectives)
- Identification and verification times (2 objectives)
- Durability of hardware

Note that the performance targets apply to Phase I only, i.e. they are to be used only to identify biometric technologies (or combinations of biometric technologies) that should be submitted to the Phase II evaluation. These performance targets thus are specified for purposes of the decision making process only, and do not necessarily reflect the performance that will eventually be required of a technology under operational conditions.

This Appendix covers the following:

- Process overview.
- Setting OSIPM performance targets
- Setting performance targets for other objectives

2 PROCESS OVERVIEW

2.1 OSIPM OBJECTIVES

UID9 has determined that the OSIPM are better suited than the traditional FMR and FNMR to describe the required performance of biometric technologies within a DL environment. However, this determination was only made after performance targets were set for the FMR and FNMR. In setting performance targets for the OSIPM, two options were available:

- Do the whole process all over again; or
- Derive performance targets for the OSIMP based on the performance targets set for the FMR and FNMR. Due to the cost and effort involved in setting performance targets, this option was followed.

Setting performance targets for the traditional FNMR and FMR as well as for the FTER commenced with documenting the interrelationship between these rates and other issues in the DL environment. This was illustrated by way of an influence diagram.

The influence diagram identified resource requirements (including personnel and equipment) as one of the main determinants for setting the rates. That is, the extent of resources available places certain limits on the rates. Determining the resource availability and working back from there thus is one way of motivating performance targets for the rates. However, the availability of resources is critically dependent on the economic and political environment at the time when an implementation decision is made. Given that a

current environment generally is not a good estimate of a future environment, this approach could not be used to determine performance targets.

Alternatively, the impact of a range of values for each rate on the DL environment can be estimated. For each value, an expert opinion of the probability that the impact will be politically acceptable is determined. These probabilities, and the risk profile of the decision makers, are then used to identify rates that will be acceptable performance targets. However, this approach proved to be unsuitable. Questions formulated to allow a relatively comprehensible evaluation did not fully address the objective at hand. Once the question was formulated to adequately address the objective, the evaluation became too complex to remain focused, and resulted in widely divergent assessments.

Given the above history, and the fact that the performance targets would apply to Phase I only, the performance targets were set by way of a consensus decision. That is, rates were suggested, discussed, and unanimously accepted by UID9. Even though this approach resulted in performance targets that may be difficult to justify to decision-makers, UID9 agreed that the consensus performance targets should fulfill the intended purpose.

As explained above, UID9 determined that the OSIPM are better suited to UID9's needs. In order to convert the traditional FMR and FNMR into OSIPM, a new influence diagram, replacing the FMR and FNMR with the OSIPM, was constructed. A comparison of the two diagrams then resulted in performance targets for the OSIPM.

2.2 OTHER OBJECTIVES

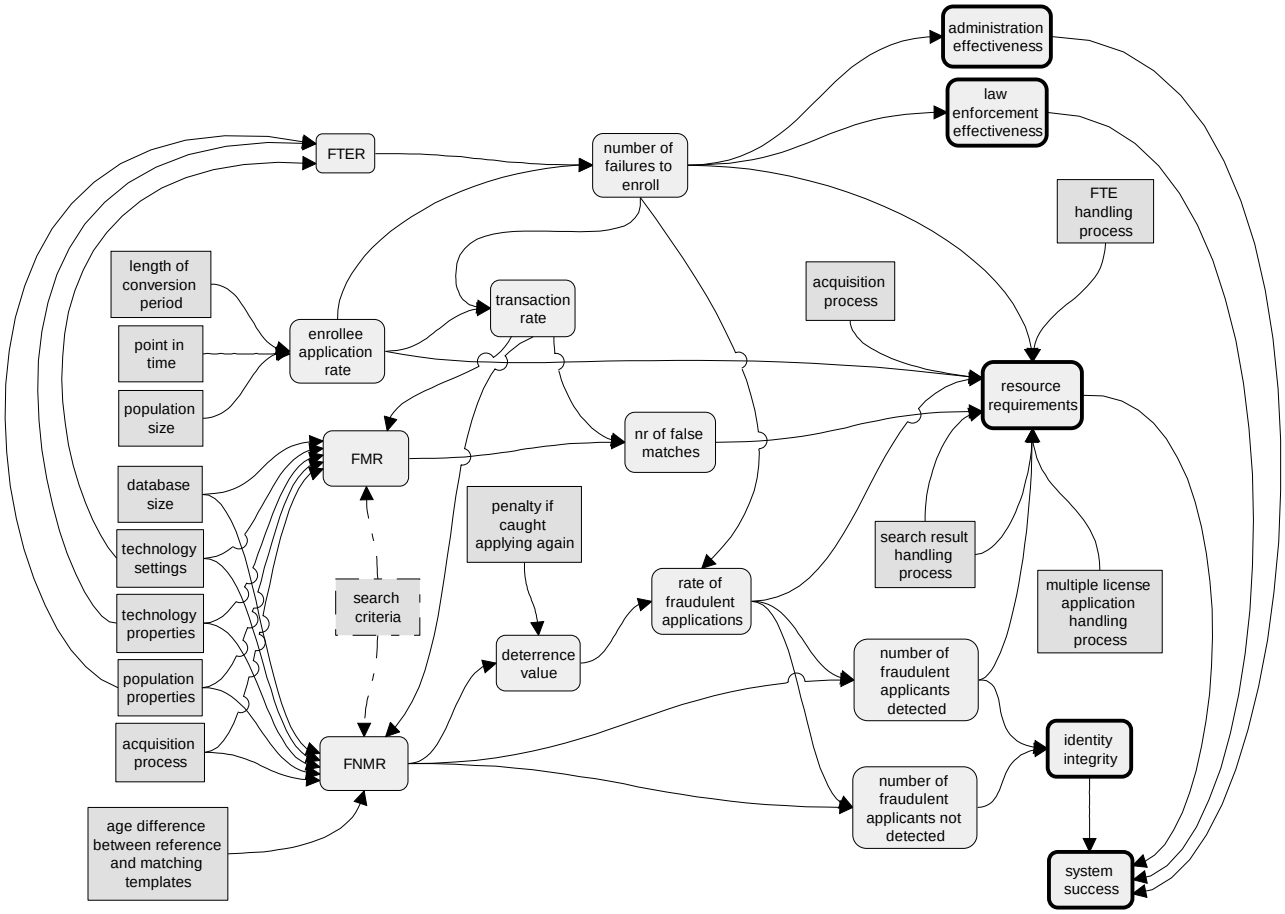
Performance targets for the remaining Phase I objectives were determined by submitting suggested targets to UID9 for perusal, and discussing comments until consensus was reached.

3 SETTING OSIPM PERFORMANCE TARGETS

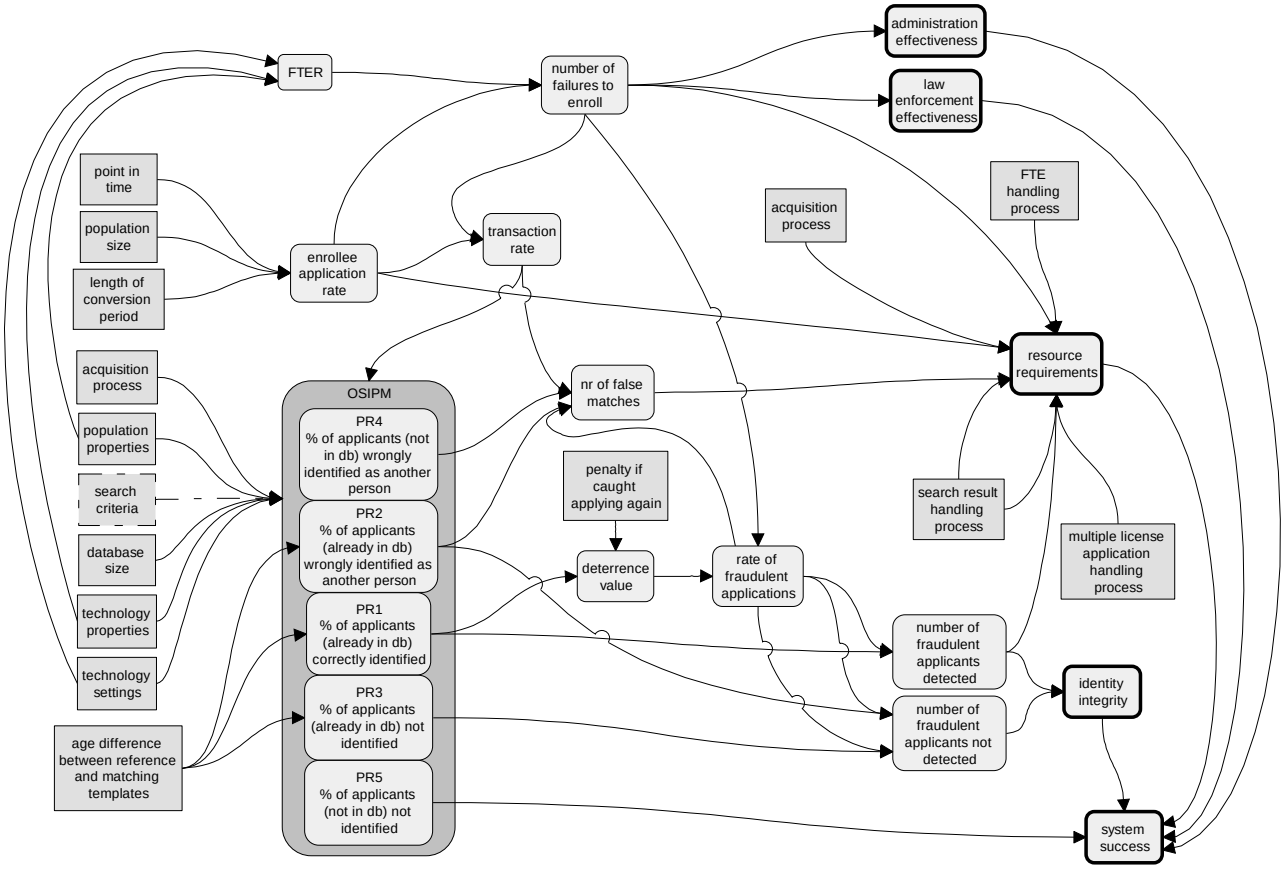
3.1 ENVIRONMENT OVERVIEW

The issues that influence the traditional FTER, FNMR and FMR, as well as the impact of these rates on other issues relevant to this Project, are illustrated in the influence diagram below.

In the diagram, each arrow indicates that the originating block influences the destination block. In those cases where arrows arrive together at a block, the originating influences work together. Blocks with squared corners represent inputs to the diagram (note that these blocks may be influenced by other issues outside the scope of the immediate question, and thus not shown). Thick lined blocks (with rounded corners) represent the outputs of the diagram.



Replacing the FMR and FNMR with the OSIPM yields the following influence diagram.



The following notes apply to the diagram (references to block titles are in *italic* typeface):

- The diagram is an influence diagram, which is not the same as a process flow diagram. A simplified process flow diagram is shown below later in this Paragraph. The influence diagram is not concerned with the sequence of events per se, but has as aim to show the interrelationships between the various rates, processes, input parameters, results, etc.
- The number of applicants who cannot be enrolled because an acceptable biometric template cannot be obtained (*number of failures to enroll*) may influence the rate of drivers who apply for more than one license (*rate of fraudulent applications*). Depending on the process used to capture unenrollable applicants onto the system, these applicants may find it easier than enrolled drivers to obtain multiple licenses. This influence is however expected to be quite small[16].
- The *transaction rate* is the number of drivers (per time period) for whom a 1:n search is performed.

[16] In the absence of supporting "evidence," this statement is based on the reasoning that the *penalty if caught applying again* can potentially have a much bigger influence on the *rate of fraudulent applications* than will the FTER. Also, it is only that portion of fraudsters (i.e. people who are contemplating trying to enroll more than once) who actually fail to enroll, who may consequently try to do so again. That is, fraudsters who do successfully enroll the first time will not be influencing the *rate of fraudulent applications* because of the FTER. Given these assumptions, it could be expected that (within the bigger picture) the influence of the FTER on the rate of fraudulent applications will be quite small.

- The number of applicants who cannot be enrolled (*number of failures to enroll*) is subtracted from the number of drivers applying for enrollment for the first time (*enrollee application rate*, Variable 1.2.1 in B004.OBJT.001), in order to arrive at the *transaction rate* used to calculate the expected *number of false matches*.
- A consequence of the number of applicants whose biometrics cannot be enrolled (*number of failures to enroll*) is to decrease *law enforcement effectiveness*[17]. Every person who cannot be enrolled using his/her biometric means one more person whose identity cannot be verified by law enforcement using a biometric.
- Driver license *administration effectiveness* is also negatively influenced by the number of applicants whose biometrics cannot be enrolled (*number of failures to enroll*). Every driver that cannot enroll implies that biometric technology cannot be used to identify such a driver at any jurisdiction for any future transaction such a driver may apply for. This places a limit on the possible driver license *administration effectiveness*.
- Another consequence of the number of applicants whose biometrics cannot be enrolled (*number of failures to enroll*) is the increased resources required to verify that the applicant has not already been enrolled, before capturing the person's details onto the system. Note that those drivers who are difficult to enroll, and thus require additional resource time, are not considered separately in the diagram, but are included in the *resource requirements* applicable to the *acquisition process*.
- The *number of false matches* is not the same as the number of possible false matches. The *number of false matches* represents the actual false matches, i.e. those cases where a person is actually wrongly matched with someone else. A possible false match is a person whom the biometric technology identifies as matching with someone else, before the search result handling process confirms if this is true or not.
- The OSIPM can be defined as follows:
    - PR1: The percentage of applicants who are in the database, and who fraudulently apply for a 2nd identity, who are identified
    - PR2: The percentage of applicants who are in the database, and who fraudulently apply for a 2nd identity, who are identified as someone else.
    - PR3: The percentage of applicants who are in the database and who apply for a 2nd identity, who are not identified as someone else
    - PR4: The percentage of applicants who are not in the database, who are identified as someone else
    - PR5: The percentage of applicants who are not in the database, who are not identified as someone else.
- The number of multiple licenses that enter system (*number of fraudulent applicants not detected*) is determined by both *PR2* and *PR3*, together with the *rate of fraudulent applications*. It is important to note the *PR2* and *PR3* are rates that illustrate an aspect of the technology's effectiveness – a

[17] This statement assumes that Law Enforcement has the infrastructure to perform biometric verification. Absent this infrastructure, every non-enrollable driver is one more driver that Law Enforcement will not be able to verify once the necessary infrastructure has been established.

combined rate of 1% does not mean that 1 out of every 100 records will be false, but that 1 out of every 100 attempts to enroll under a new identity is expected to succeed.

- A high *PR1* has a *deterrence value*, which in turn influences the rate of drivers who apply for more than one license (*rate of fraudulent applications*).
- The rate of fraudulent applications for driver licenses can significantly be influenced by the *penalty if caught applying again*. A severe enough penalty is expected to substantially decrease the relative importance of *PR2* and *PR3*.
- Although not indicated in the diagram, other variables that may potentially impact the *rate of fraudulent applications* include the *acquisition process*, the *FTE handling process*, and the *multiple license application handling* process. Each of these processes have the potential of introducing "holes" in the overall license application process that may encourage fraudulent applications.
- The *system success* as a whole is influenced by the other output blocks in the diagram, i.e. *resource requirements*, *law enforcement effectiveness*, *DLA effectiveness* and *identity integrity*.
- The *number of false matches* is influenced by both PR2 and PR4. The confirmation of false matches in turn consumes resources (*resource requirements*).
- The *search result handling process* is the process whereby possible matches (estimated by *PR1*, *PR2* and *PR4*) are evaluated to determine if there are any true matches. The *search result handling process* is the same for all cases.
- The potential *search criteria* (see Variable 1.1 in Appendix B) influence all of the OSIPM by effectively reducing the number of records to be searched (i.e. the *database size*). Potential *search criteria* are not to be considered during the initial evaluation.
- The *technology properties* include both the hardware and software involved with a particular biometric technology.
- As the *transaction rate* increases, many biometric technologies will dynamically adapt the matching thresholds in order to alleviate the load on the system. These changes are most often outside the control of the system operator. An increasing *transaction rate* thus can negatively influence the integrity of search results.
- The term *identity integrity* as used in the diagram is a measure of the number of multiple licenses allowed onto the system database. Each fraudulent applicant detected means one less multiple license on the database; each fraudulent applicant not detected means one more multiple license on the database.

The input parameters to the above diagram are the following:

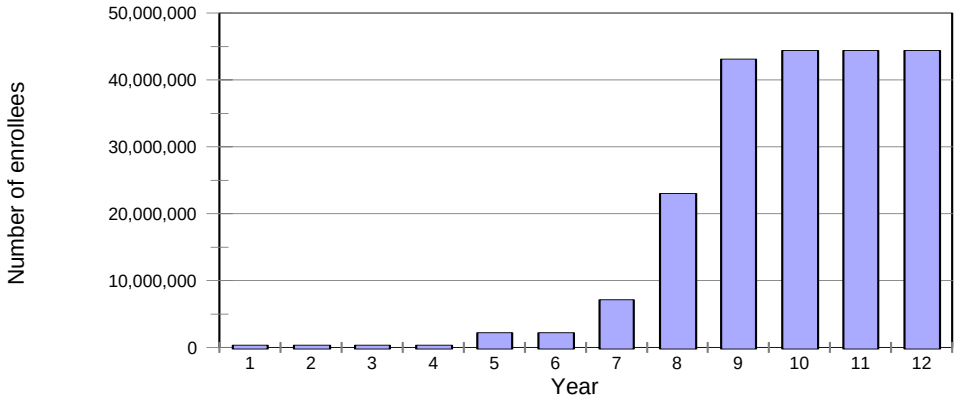
- Population properties: Size: 300 million driver records. Demographic: All drivers in North America.
- Technology properties: The unique properties of each biometric technology, as manifested in the performance of the technology.
- Technology settings: Input and setup parameters that can be changed by an operator.
- Length of conversion period: 4, 8 or 12 years (as per Variable 1.5 in Appendix B). According to 1999 data, all states and provinces except for 9 have a renewal period of 5 years or less. The only exceptions are five states with a 6-year period, three states with an 8-year period, and one state (AZ)

with an effective 12-year period. In practice, provided that Arizona can commence enrollment at the start of the conversion period, a 12 year conversion period makes the most sense: All states and provinces except for 3 will be able to fit at least two renewal cycles into the 12 year period. This allows some leeway in managing the transaction rate during the conversion period. Although a shorter conversion period would lead to the benefits of the system being observed sooner, a shorter conversion period implies that some jurisdictions will have to amend their renewal cycles, and also increases the responsibility of jurisdictions to commence the conversion period at a particular time. Past experience in similar matters has shown that such compliance is improbable.

- Point in time: This refers to the period of time during the conversion period that the enrollee application rate will be at its highest. The evolution of this parameter is discussed below.

  An initial estimate was that roughly 50% of drivers would have to be enrolled onto the system during the last 25% of the conversion period. That is, the actual transaction rate during the last 25% of the conversion period comes to twice the average transaction rate.

  This initial estimate can be tested by looking at the worst case scenario: If each jurisdiction is to schedule the end of their respective conversion periods to coincide with the end of the (12 year) continent-wide conversion period, and taking into account each jurisdiction's current conversion period and number of licensed drivers[18], the actual transaction rate during the last 3 years comes to approximately 2.5 times the average rate. The graph below shows the annual number of enrollees



  under this scenario.

  However, realistically speaking, it would not be prudent to discard a biometric technology just because it cannot accommodate the worst-case transaction rate based on the absence of coordinated scheduling of the various jurisdictions' conversion periods. For purposes of calculating a "best case" actual transaction rate, assume the following scenario:

  - A conversion period of 12 years.

[18] The noted figures are based on US and aggregate Canadian licensed driver numbers (see Appendix C1). Although a more accurate calculation can be performed using driver figures for individual Canadian provinces, it is expected that such figures will not perceptibly change the results of the calculations as used in this document.

- Each jurisdiction will be enrolling over at least a 7-year period. That is, a jurisdiction with a 5-year renewal cycle will be enrolling over 10 years, a jurisdiction with a 3-year cycle will be enrolling over 9 years, a jurisdiction with a 4-year cycle will be enrolling over a period of 8 years etc.
- All jurisdictions will be busy enrolling at the same time during some point in the 12 year conversion period

Under the above assumptions, the actual transaction rate comes to 1.3 times the average transaction rate (see Appendix C1).

Given that the actual transaction rate specified in this document is used only to describe the "entry level" environment, the actual transaction rate was set at 1.3 times the average transaction rate. Note that this may not necessarily be the same actual transaction rate that will eventually apply to operational conditions.

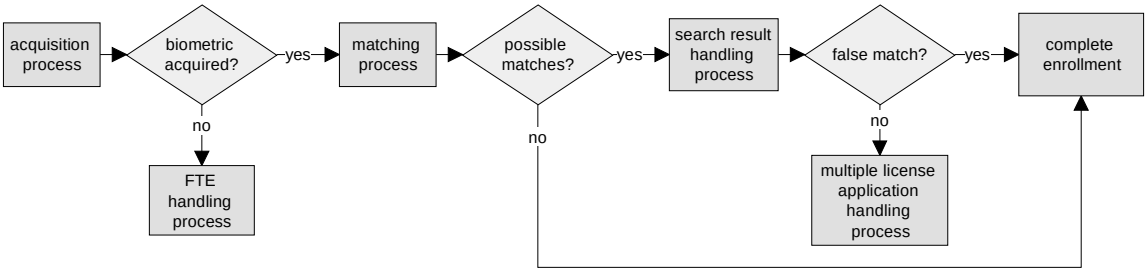
- Database size: Although the peak demand period (highest transaction rate and largest database size) can be expected sometime just before all drivers have been enrolled, a conservative approach would be to use a database size of 300 million records (less the estimated number of drivers who will "fail to enroll").

  Given the above assumptions, the following holds:

  - The actual transaction rate will be the highest towards the end of the enrollment period.
  - Database growth relative to existing size will be most during the beginning phases of the enrollment period.
  - Database size increases steadily over time.

- Processing times. In order to estimate the resource requirements for processing enrollees, the various processes involved need to be specified. Although difficult to do so at this time, broad outlines of the process detail is provided below.

  The relationships between the various processes are illustrated in the following diagram.



  Detail in respect of each process is as follows:

  - Acquisition process: The process of capturing raw biometric data and extracting a biometric template for use in an automated biometric matching algorithm. In general, this requires a person to present him/herself to an acquisition device, with or without assistance from an operator. The detail of the actual process (i.e. duration, operator involvement) depends on technology and equipment used.
  - Failure to enroll (FTE) handling process: This process allows a person, whose biometric template could not be obtained, to be checked against the existing database (for possible

multiple licenses) using alternative identifiers. Due to the importance of ensuring "one record one driver," this process requires the jurisdiction to perform additional checks over and above those performed during a "normal" enrollment. Depending on the checks performed during normal enrollment, the additional checks could include verifying a person's name, social security number and date of birth from approved documentation, and possibly verifying this information with other databases. A different UID Task Group is investigating additional verification procedures.

- Matching process: This is the process during which the acquired biometric template is submitted to the database of existing templates, to determine if it belongs to a person who has already enrolled onto the system. This is an automated process.
- Search result handling process[19]: The matching process may return no matches, or one or more possible matches. Depending on the technology (or combination of technologies) used, the process for verifying if a possible match is a true match may differ. Note that a possibility exists that, due to an imperfect search result handling process, a possible match can be declared a false match whilst in actual fact it is a true match. The search result handling process thus has the potential to either improve or decrease the accuracy of the solution.
- Multiple license application handling: If a possible match is confirmed as a true match, it means that the applicant has tried to obtain an additional identity on the system. The time and resource requirements for this process will depend amongst others on the penalty imposed on such applicants.

3.2 DETAILED DISCUSSION OF OBJECTIVES

The sub-paragraphs that follow comprise the following:

- For each rate (i.e. the FTER, FMR and FNMR), a definition and a summary of the rate's impact (from the preceding discussion and influence diagram in Paragraph 3.1)
- A discussion of approaches followed in trying to determine performance targets for the above rates, i.e. resource requirements, rate justifications, and expert matter opinion.
- The performance targets that were set for the above rates.
- A discussion on how the performance targets for the FMR and FNMR were converted into performance targets for the OSIPM.

3.2.1 RATE DEFINITION AND IMPACT SUMMARY

3.2.1.1 FTER

Definition

The FTER is defined as the proportion of applicants whose biometric cannot be enrolled. Applicants who are difficult to enroll (i.e. need more than one attempt to successfully enroll) are not regarded as failures to enroll.

[19] Note that under test conditions, the search result handling process is optional. The outcome of the matching process can immediately be categorized according to the OSIPM because it is known beforehand if an applicant is already on the database.

A FTE only results after repeated unsuccessful attempts at enrolling an applicant. Applicants who cannot be enrolled as a result of improper use of equipment or insufficiently trained operators are, for purposes of the definition provided here, not regarded as a FTE. In a multi-modal environment, the definition of a FTE (i.e. failure to enroll in one or in all technologies deployed) will depend on the technologies used and the manner in which they are integrated.

Impact summary

A FTE implies that the jurisdiction has to follow an "exceptions process" to enroll the customer (without a biometric template) into the local DL/ID system. This will negatively impact customer service in the following ways:

- Increased time in the office for the customer.
- Increased time spent by staff to enroll the customer.
- Increased time / suspicion of the customer upon interaction with law enforcement personnel at later date.

The impact of a FTER on driver license administration is considered minor in comparison to the impact of the FMR. It should however be noted that, depending on the exceptions process for non-enrollees, and assuming that drivers who fail to enroll are added to the system without a biometric, non-enrollees may find it easier (than drivers who do enroll) to establish additional identities.

The FTER also has an impact on law enforcement. Each "failure to enroll" implies one more driver without a biometric, and thus for these drivers it cannot be verified (using biometrics) if they are who they say they are (1:1 match). For example, for a population of 300 million registered drivers, a FTER of 5% implies that 15 million drivers will not be able to enroll onto the system with their biometrics. Stated differently, for each 1 million registered drivers in a jurisdiction, it is expected that 50,000 would fail to enroll using their biometrics.

3.2.1.2 FMR

Definition

Technically, the false match rate represents the probability of a 1:1 comparison of biometric records from two different people resulting in a false match. In a system within which 300m users are to be enrolled over a 12 year period, and assuming a peak demand transaction rate of 1.3 times the average transaction rate, 130,000 1:n checks (i.e. up to 3.9×10^{13} (39 trillion) individual comparisons) will be performed per day. As a result, the expression of the required false match rate can become difficult to interpret. For example, if a false match is allowed for 1 in 10 applicants, the corresponding false match rate (given the details above) is 1 in 10*300,000,000, or 3.33×10^{-8}%.

Instead, for purposes of this evaluation, an "effective false match rate," which is easier to relate to, is defined. An effective false match rate is the probability that a customer is falsely matched against a database. Given the size of the database, an effective FMR can be converted to a technical FMR. Note however that for large databases, the transaction rate (and not just the database size) may also influence the relationship between the effective FMR and technical FMR.

For biometric technologies returning a "candidate list" of possible matches in answer to a 1:n search, the search is counted as one false match (provided of course that none of the records in the candidate list is determined to be a true match), regardless of the number of records in the candidate list. The motivation for this approach is twofold: If a candidate list is returned after a search, the average time required to work through the candidate list can be estimated within reasonable limits of variance, and due to the process followed to determine if any record in a candidate list is a true match, the time to process ten galleries

consisting of one record each is expected to be considerably more than the time required to process one candidate list containing ten records.

Note that a driver who has already enrolled, and who is (upon trying to enroll for a 2nd time) falsely matched, contributes to both the FMR and the FNMR.

Impact summary

The time spent on checking if possible matches are false matches is the biggest single factor determining an acceptable false match rate. A FMR of 5% would imply that for 1 in every 20 applicants (whose biometrics were acquired successfully), additional time will have to be spent confirming that the match is not true.

Although it has been mentioned that false matches may also lead to increased time for a customer in the office, it is generally assumed that for the scale involved (i.e. a 300 million record database), performing a 1:n query within the time a customer normally spends in an office is not feasible[20]. Hence, the FMR is not expected to impact the applicant during his/her visit to the driver license office. However, the procedure prescribed to resolve possible matches may in exceptional cases require an applicant to revisit the DMV.

As law enforcement is for the moment interested primarily in 1:1 matching, the effective FMR (as defined above) does not have a significant impact on law enforcement. It is however foreseen that 1:n matching may be performed by law enforcement at a future date, albeit in an office environment and not from the roadside.

3.2.1.3 FNMR

Definition

A false non-match occurs when a 1:n search is performed for an existing record, and the query returns no match or a false match.

Note that the FNMR applies only to applicants who on purpose try to obtain a second identity on the system[21]. For drivers who apply for a renewal after having been enrolled with their biometric information, a 1:1 search is performed, and is supported by existing documentation.

Impact summary

A false non-match implies that an enrollee with an existing identity on the DL/ID system has successfully registered an additional identity into the system. This has no immediate impact upon customer service, because the applicant's ruse was successful. However, if identity theft was involved, customer service will definitely be impacted when the true owner of the stolen identity applies for enrollment. Note that the actual false non-match rate for field applications is difficult to determine once the system has been implemented, even if a jurisdiction regularly schedules controlled tests of the system.

3.2.2 RATE ASSESSMENT

Several approaches were followed to arrive at performance targets for the rates discussed above. These approaches and the associated results are discussed below.

---

[20] This is due to the astronomical cost that would be associated with a system capable of the required response times. Should in future an over-the-counter solution become viable, the impact of the FMR on a customer's waiting time should again be considered.

[21] Unintended re-enrollments (e.g. when an applicant is incorrectly instructed by an operator to enroll) may also contribute to the FNMR.

3.2.2.1 RESOURCE REQUIREMENTS

From the influence diagram in Paragraph 3.1, it is clear that one way of arriving at minimum performance targets for the various rates is to quantify acceptable levels/outcomes for the consequences of those targets, and to work back from there. For example, an acceptable FMR can be calculated by estimating the resources that will be available (personnel, computing power, etc.) as well as the amount of effort that will be required to process each false match.

Given that all the rates under discussion have an impact on the resource requirements, it thus follows that the resources that will be available are crucial for calculating these rates in the manner contemplated above. However, funding these resources is a political issue, involving the values each jurisdiction and the respective Governments assign to the project at the time when the decision to implement is made. Given that it is almost impossible to estimate the resources that will be available in advance, it follows that the resource requirements cannot be used to calculate a maximum value for the FMR for this project.

3.2.2.2 RATE JUSTIFICATIONS

Several performance targets were justified based on the general perceived impact and importance of the rates. These explanations are provided below for each of the rates in question. However, the explanations tended to only partially address the impact of each rate, and essentially relied more on opinion than upon reasoning that would hold up in a legislature.

FTER

In response to an initially proposed target FTER of 5% (i.e. 1 out of every 20 drivers stopped will not have a biometric associated with their driver license), the following arguments have been raised in support of a smaller FTER:

- The ability to “sell” a system based on a biometric technology to decision makers rests on the improvement of the system beyond the perceived effectiveness of the current system. It was questioned if support for the technology can justifiably be obtained if the system is only 95% effective in terms of enrollment, much less the next step which is the ability to match the records that can be enrolled.
- From the perspective of fraudsters trying to “beat the system,” a 5% FTER was considered to be too high. If it is known that 5% of the time in normal circumstances the system will NOT enroll even the honest person, it can be expected that a dishonest person will try harder to thwart the enrollment process, and the unsuspecting employee will also be fooled because they expect that one out of 20 will not be enrolled under normal circumstances. This argues for making a failure to enroll the very small exception.

The counter argument is that it can be questioned whether discarding a technology based on a FTER more stringent than 5% is prudent at this time.

FMR

The AAMVA participated with the Federal Highway Administration's Office of Motor Carriers (now the Federal Motor Carrier Safety Administration (FMCSA)) to produce a report on Biometric Identification Standards Research, published December 1997. A series of functional requirements for a biometric identification system were established including a requirement for false match error rates. The final report indicated, "A recognition system false match should occur with no more frequency than one applicant in one million when the system is at the target enrollment level of 8.5 million." (The report also allowed 72 hours for such a matching process.) This equates to an error rate of 0.0001%. The report does however not provide the reasoning behind this number.

As explained above, the resources that will be available in the future to process false matches cannot be estimated now, and thus cannot be used to calculate a meaningful maximum FMR.

Given the inverse relationship between the FMR and FNMR, it is conceivable that a FMR can be derived from the maximum FNMR. However, such a FMR would essentially be based on the specific technology, and not on business requirements.

One way that could be used to determine a minimum acceptable false match rate is to assess an expert opinion on the maximum amount of additional time that can be spent resolving false matches. However, the additional time involved is dependent on amongst others the false match resolution process, and any estimate of this time would be quite uncertain, given the current knowledge of this process.

FNMR

The FMR and FNMR are roughly inversely proportional, i.e. an improvement in the one leads to a deterioration in the other. The properties of the relationship are determined by the technology settings, the technology properties, and the database size. In this case, it essentially means that the resources available to process false matches is played off against the probability of not catching a driver applying for a second license.

The FMCSA's 1997 report established "less than 10% false non-match error rate in the recognition mode" (i.e. 1 to many search) as a requirement.

The actual fraudulent application rate is very difficult to estimate, and hence cannot reliably be used to calculate a maximum FNMR.

An argument has been made that because the ultimate goal of the system is to ensure “one driver one record,” the FNMR should take precedence over the FMR. That is, establish an acceptable FNMR, and determine the corresponding FMR.

The counter-argument is that a substantial penalty for getting caught trying to apply for a second time, even if the FNMR is not that low, will act as an effective deterrent, thus decreasing the rate of fraudulent applications for enrollment. Consequently, a less stringent FNMR is warranted.

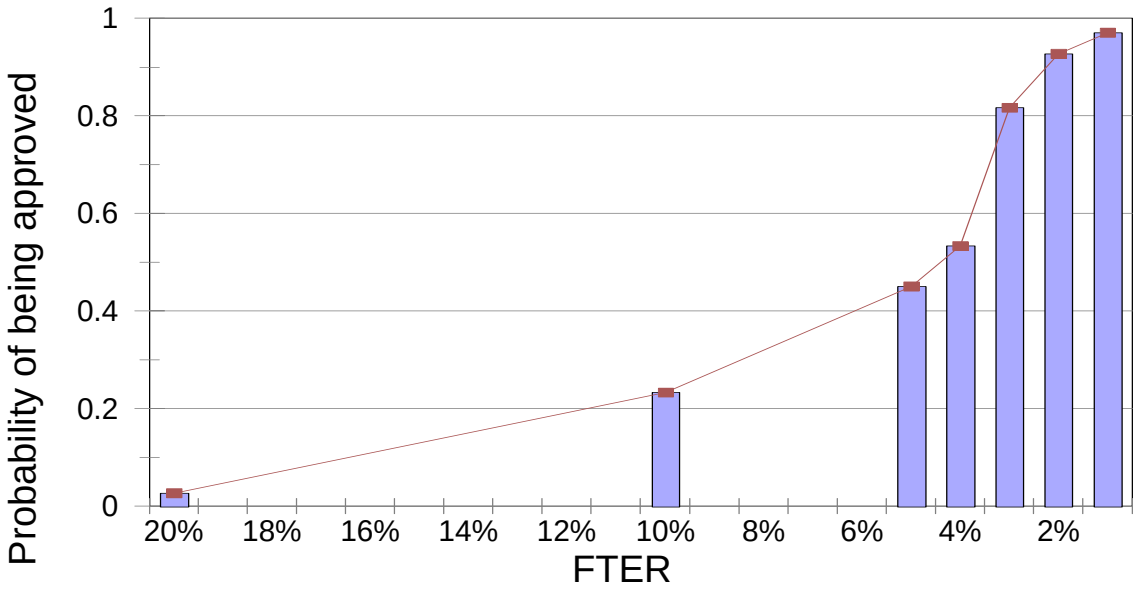
An alternative way of looking at the FNMR is to argue that drivers who try to obtain an additional identity, although they may be few (given a severe enough penalty if getting caught), are the people whom the system is really designed to prevent from being successful. What would be the minimum acceptable FNMR for these applicants? That is, how many of the hard-core criminals who try to circumvent the system are we prepared to allow getting away with it? Note however that the FNMR so derived does not (on its own) give an indication of the actual number of multiple identities that can be expected.

3.2.2.3 EXPERT MATTER OPINION

As an alternative (essentially a structured refinement of the Rate Motivations discussed above), the impact of a range of values for each rate can be estimated. For each value, an expert opinion of the probability that the value will be politically acceptable is determined. These probabilities, and the risk profile of the decision makers, are then used to identify rate values that will be acceptable performance targets.

For example, given the impact of a FTER (as discussed in Paragraph 0), the following graph shows the expected probability (as could be assessed by a group of subject matter experts, i.e. UID9) that the relevant decision makers will approve a biometric technology solution with a particular FTER.

EXAMPLE



Note that the FTER is enumerated only for those values where a significant change in the probability of acceptance from one rate to the next is expected.

After assessing the minimum acceptable probability of a solution being approved, the minimum FTER could be read from the graph. For example, if UID9 determined that no proposal will be submitted unless the probability of approval is at least 80%, the minimum acceptable FTER (from the graph above) would be 3%.

However, this approach proved to be unsuitable. Questions formulated to allow a relatively comprehensible evaluation (such as in the above example) did not fully address the rate at hand. Once the question was formulated to adequately address the rate, the evaluation became too complex to remain focused, and resulted in widely divergent assessments.

3.2.3 FINAL RATE ASSESSMENT

Considering the complexity of the rates and the lack of information (no one has implemented a biometric system of this scale before, and it is essentially necessary to estimate the impact on an environment that has not yet been specified to an extent that would allow this), and given the fact that the performance targets are to be used only as an initial “filter,” a consensus decision was made, and the following performance targets were set:

- FTE: 5%
- FMR: 10% (maximum candidate list size: 10 records)
- FNMR: 10%

Upon proceeding to Phase II of this evaluation, additional research will have to be performed in this area.

3.2.4 CONVERSION OF TRADITIONAL RATES INTO OSIPM

In the influence diagram for the traditional rates, the FMR influences, together with the transaction rate, the number of false matches that can be expected. If the transaction rate stays the same, it follows that when replacing the FMR with the OSIMP, the collective influence of the OSIPM on the number of false matches should stay the same. That is, the sum of Potential Result 2 and Potential Result 4 must be equal to or lower than 10%.

Following the same argument, the sum of Potential Result 2 and Potential Result 3 must be equal to or lower than 10%.

In addition, the following relationships hold:

- Potential Result 1 plus Potential Result 2 plus Potential Result 3 equals 1.
- Potential Result 4 plus Potential Result 5 equals 1.

Unfortunately, as the above set of equations have an infinite set of possible solutions (4 equations, 5 variables), exact values for the OSIPM cannot be set at this time. The limits placed on the OSIPM should however be sufficient for Phase I purposes.

4 SETTING PERFORMANCE TARGETS FOR OTHER OBJECTIVES

4.1 PROVEN TRACK RECORD

Objectives 1.1.4.1, 1.1.4.2, and 1.1.4.3 focus upon the proven track record of biometric technologies, as reflected by implementation aspects of the technologies in the world at large.

4.1.1 OBJECTIVE 1.1.4.1 – FAVOR TECHNOLOGIES WITH LARGER EXISTING DATABASES

The ability of a biometric technology to handle large volumes of persons is an important consideration. Taking into account that the system is to expand to 300 million records, it was decided that an arbitrary limit of 0.5 million enrollee records in a database be set for a biometric technology to be considered.

4.1.2 OBJECTIVE 1.1.4.2 – FAVOR TECHNOLOGIES THAT HAVE BEEN IMPLEMENTED FOR DAY-TO-DAY USE

It was decided that each biometric technology being considered should have been implemented in at least one live environment.

Although it will be beneficial for a biometric technology if it has been implemented in a DMV environment, this is not a prerequisite.

The collective age (in years) of all known implementations should equal at least two. That is, the technology should have been operationally used for at least 2 years. This period should allow enough time for major deficiencies in the technology to become known.

4.1.3 OBJECTIVE 1.1.4.3 – STABILITY OF CORE ACQUISITION AND MATCHING TECHNOLOGIES

The range of values have been defined as follows:

- Most Satisfactory - Highly stable core technology, supported by leading technology providers for over five years, that leverages highly mature imaging and/or matching methods and standards

- More Satisfactory - Stable core technology, supported by leading technology providers for under five years, that leverages reasonably mature imaging and/or matching methods and standards
- Satisfactory - Fairly stable core technology, supported by leading technology providers for under two years, that leverages emerging imaging and/or matching methods and standards
- Unsatisfactory – Fairly unstable or only recently established core technology, with few well-defined or standardized approaches to biometric acquisition and/or matching

The minimum acceptable value is “Satisfactory.” Anything less than this value would focus upon emerging technologies rather than mature technologies. UID9 participants have stressed the need for proven technologies with an existing track record of performance.

4.2 ABILITY TO BE INFLUENCED BY USERS

4.2.1 OBJECTIVE 1.2.1.1 – MINIMIZE ABILITY OF UNCOOPERATIVE USERS TO INFLUENCE BIOMETRIC MATCHING ACCURACY

The range of values have been defined as follows:

- Most Satisfactory - Matching accuracy cannot be influenced by uncooperative users; no supervision required to prevent influencing
- More Satisfactory - Matching accuracy can only be influenced by a user highly familiar with the operations of the biometric technology; little to no supervision required by trained staff to prevent influencing
- Satisfactory - Matching accuracy cannot be easily influenced by a user, even one with some familiarity with the operations of the biometric technology; light supervision required by trained staff to prevent efforts to influence accuracy
- Unsatisfactory – Matching accuracy can be influenced by a typical user with little or no effort regardless of technical background of the technology; direct supervision required by trained staff to prevent efforts to influence accuracy

The following definitions apply:

- “Trained staff” as used here implies rudimentary knowledge by frontline staff of the biometric acquisition process and minor training in "subversive tactics" to look for – methods commonly employed to dupe a biometric reader, or common user mistakes resulting in poor acquisition.
- “Little supervision” as used here means that enrollment is done within plain sight of a staff person.
- “Light supervision” as used here means that enrollment is done within plain sight of, and at the verbal direction of, a staff person.
- “Direct supervision” as used here means that enrollment is done under physical supervision by a staff person, i.e. there is physical contact between the enrollee and the staff person.

The minimum acceptable value is Satisfactory. Matching accuracy must not be easily influenced by user behavior, and it is anticipated that trained staff will supervise every instance of interaction with customers where biometric verification or identification is required.

4.2.2 OBJECTIVE 1.2.1.2 – MINIMIZE ABILITY OF UNCOOPERATIVE USERS TO INFLUENCE BIOMETRIC ACQUISITION

The range of values have been defined as follows:

- Most Satisfactory - Acquisition cannot be influenced by uncooperative users; no supervision required to prevent influencing
- More Satisfactory - Acquisition can only be influenced by a user highly familiar with the operations of the biometric technology; little to no supervision required by trained staff to prevent influencing
- Satisfactory - Acquisition cannot be easily influenced by a user, even one with some familiarity with the operations of the biometric technology; light supervision required by trained staff to prevent efforts to influence accuracy
- Unsatisfactory – Acquisition can be influenced by a typical user with little or no effort regardless of technical background of the technology; direct supervision required by trained staff to prevent efforts to influence accuracy

The minimum acceptable value is Satisfactory. Biometric acquisition must not be easily influenced by user behavior, and it is anticipated that trained staff will supervise every instance of interaction with customers where biometric measurements are captured.

4.3 IDENTIFICATION AND VERIFICATION TIMES - 1:1 SEARCH (VERIFICATION)

4.3.1 IMPACT

A 1:1 search is conducted for every renewal transaction after the applicant has had his/her biometric enrolled in the system using the 1:300m identification process. It is assumed that the 1:1 search is conducted in real time, regardless of whether the actual document issuing is performed over-the-counter or in a centralized environment. The time required for this search thus directly impacts the renewal process. However, depending on the operational sequence of events, the time for the search will not necessarily prolong the applicant's stay in the driving license office.

4.3.2 RATE ASSESSMENT

The FMCSA's Biometric Identification Standards Report indicated "the verification processes [i.e. 1:1 match, typically performed in the case of renewal applicants] must be accomplished with 5 seconds of receipt of the inquiry."

The motivation for a 5 second response time for a 1:1 match was not provided. A generally accepted goal of 7 seconds has been used in industry for systems in similar environments, and will be used for purposes of this Project.

4.4 IDENTIFICATION AND VERIFICATION TIMES - 1:N SEARCH (IDENTIFICATION)

4.4.1 CENTRALIZED ISSUING

4.4.1.1 IMPACT

In a centralized environment, the turnaround time for a driver license is not as important as in an over-the-counter environment. The main impact is on the administration – the longer the search time, the greater the number of driver license applications that are "in process" at any given time.

Note that the wait time for a 1:n search result in itself does not influence the buildup of backlog. Whether or not backlog is created depends on the “arrival rate” of search results, and the rate at which search results can be processed by a jurisdiction. As long as the arrival rate is smaller than the processing rate (the acceptable difference being determined by the variation in the arrival and processing rates), the actual maximum time allowed to produce a search result will not influence the queue (backlog) in front of the search results processing station. For example, if each search takes say 72h, search results are forthcoming at a rate of 100/h, and the jurisdiction has the capacity to process 200 search results per hour, it is highly unlikely that there will be any backlog.

4.4.1.2 TIME ASSESSMENT

The FMCSA's Biometric Identification Standards Report indicated "(t)he recognition process [i.e. 1:n search] must be accomplished within 72 hours after receipt of the inquiry".

For a number of reasons, it has been argued that a 24-hour response target for a 1:n search simplifies the administrative requirements for large-scale biometric systems. It was decided that this duration will be used within this Project.

4.4.2 OVER-THE-COUNTER ISSUING

4.4.2.1 IMPACT

In an over-the-counter environment, the 1:n search time can have a significant impact on the duration of an applicant’s visit to the driver license office. However, the duration of the visit is not necessarily extended with the total time of the search. It is conceivable that the operational processes can be adapted in such a manner that the 1:n search is performed in parallel with other processes. For example, an applicant’s biometric could be read when he/she enters the office, and the 1:n search conducted while the person is waiting to be served. When being served, the person’s biometric can again be read to make the link to the initial reading when the person entered the office.

4.4.2.2 TIME ASSESSMENT

Assuming that operational procedures in a driver license office can be adapted to accommodate parallel processing as contemplated above, it has been decided that a 1:n search time of 15 minutes is acceptable.

4.5 OBJECTIVE 1.2.3: MAXIMIZE DURABILITY OF HARDWARE

The replacement cycle duration should be at least equal to the standard industry depreciation cycle of 3 to 5 years used for laptops and similar computing equipment.

APPENDIX C1

The purpose of this table is to show how long it takes jurisdictions to replace an old license with a new license. In some cases, a jurisdiction's initial validity period may be less than the period indicated in the table, where subsequent renewal periods are longer (e.g., the first license may be valid for 5 years, but subsequent renewals may be valid for 8 years). The majority of these jurisdictions' drivers at any given time will be subject to the lengthier validity period. Calculations of this table were simplified by using the lengthier renewal periods for these jurisdictions.

Jurisdiction	Renewal cycle period (years)	Driving population	Enrollment period (years)	Annual number of enrollees
Alabama	4	3,559,897	8	444,987
Alaska	5	472,211	10	47,221
Arizona	12	3,550,367	12	295,864
Arkansas	4	1,961,883	8	245,235
California	5	21,623,793	10	2,162,379
Colorado	5	3,287,922	10	328,792
Connecticut	4	2,650,374	8	331,297
Delaware	5	564,099	10	56,410
District of Columbia	***5****	328,094	10	32,809
Florida	6	12,743,403	12	1,061,950
Georgia	4	5,833,802	8	729,225
Hawaii	6	787,820	12	65,652
Idaho	4	896,666	8	112,083
Illinois	4	7,809,500	8	976,188
Indiana	4	4,116,924	8	514,616
Iowa	4	1,978,748	8	247,344
Kansas	6	1,871,301	12	155,942
Kentucky	4	2,756,634	8	344,579
Louisiana	4	2,718,209	8	339,776
Maine	6	942,556	12	78,546
Maryland	5	3,451,966	10	345,197
Massachusetts	***5****	4,610,666	10	461,067
Michigan	4	6,976,982	8	872,123
Minnesota	4	2,961,236	8	370,155
Mississippi	4	1,859,487	8	232,436
Missouri	3	3,862,300	9	429,144
Montana	8	683,351	8	85,419
Nebraska	5	1,267,284	10	126,728
Nevada	4	1,420,714	8	177,589
New Hampshire	4	941,829	8	117,729
New Jersey	4	5,715,089	8	714,386
New Mexico	***5****	1,231,701	10	123,170
New York	8	11,014,805	8	1,376,851
North Carolina	5	5,884,651	10	588,465
North Dakota	4	455,921	8	56,990
Ohio	4	7,736,115	8	967,014
Oklahoma	4	2,172,394	8	271,549
Oregon	4	2,534,464	8	316,808
Pennsylvania	4	8,226,202	8	1,028,275
Puerto Rico	4	0	8	0
Rhode Island	5	660,435	10	66,044
South Carolina	5	2,849,885	10	284,989
South Dakota	5	544,997	10	54,500

Tennessee	5	4,188,317	10	418,832
Texas	6	13,045,727	12	1,087,144
Utah	5	1,495,887	10	149,589
Vermont	4	515,348	8	64,419
Virginia	5	4,920,753	10	492,075
Washington	4	4,237,845	8	529,731
West Virginia	***5****	1,316,955	10	131,696
Wisconsin	8	3,667,497	8	458,437
Wyoming	4	370,713	8	46,339
Canada	***5****	20,879,000	10	2,087,900
		212,154,719		23,103,683

Enrollment period (years):	12
Average no of enrollees per year:	15,939,643
Average enrollment rate / effective enrollment rate:	1.3

* Estimated; data not available in source used.

APPENDIX D: USE OF LOTTERIES TO ELICIT DECISION-MAKER PREFERENCES

1 INTRODUCTION

As is illustrated in Appendix C, it is almost impossible to arrive at a calculated value for the FTER, FMR and FNMR that can unequivocally be defended against any criticism. Arriving at "acceptable" target rates thus requires the subjective opinion of subject matter experts, in this case the members of UID9. This Appendix contains some thought experiments, which were designed to help UID9 members express their expert opinions specifically regarding the FTER.

For these thought experiments to be successful, it was crucial that the relevant background information be perused before doing each experiment. Hence, each participant had to check this requirement.

The definition of the lottery is crucial. In the questionnaire, the lottery considered submitting a proposal for a biometric technology with a particular error rate to the participant's decision makers. The participant gets a huge reward if the proposal is accepted, and a trivial reward if the proposal is not accepted. As such, it essentially assessed each participant's estimation of his of her superior's values. An alternative formulation can be used to assess each participant's own values, which most likely will differ from his or her superior's estimated values. For example, it could be set up so that each participant is responsible for approving or not approving a biometric technology with a particular error rate. If the technology eventually results in a successful implementation, the participant is rewarded, otherwise not.

It needs to be noted that these thought experiments were valuable in pointing out the difficulty in relating the error rates of biometric technologies to the measures of success of implementing a technology. This contributed to the recommendation to reconsider the objective structure, and to divide objectives into means and fundamental objectives. It was realized that the various error rates are actually means objectives that contribute to the fundamental objective "minimize the number of multiple identities". Consequently, it is not necessary for UID9 to consider the error rates directly in the evaluation – it becomes the responsibility of a technical expert to relate error rates to the number of multiple identities.

2 PREPARATION

Read the discussion of the FTER provided.

Check the appropriate box:

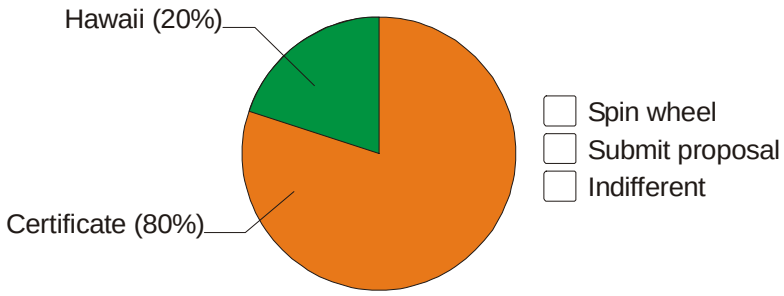
- ☐ I have read the background information
- ☐ I have not read the background information

3 EXPERIMENT: FTER = 5%

Suppose that you can enter a competition to win either a trip to Hawaii, or a $10 gift certificate. You can enter the competition in either of the following manners:

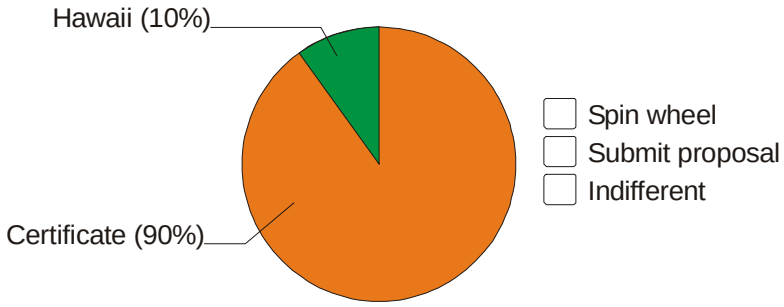
- Spin a "wheel of fortune", or
- Submit a biometric proposal with a FTER of 5% to your decision makers. If the proposal is accepted, you win the trip to Hawaii. If the proposal is not accepted, you win the $10 gift certificate.

Suppose the wheel of fortune looks as follows:



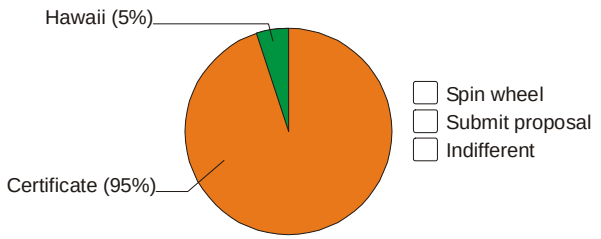
Would you rather spin the wheel, or submit the proposal? Or are you indifferent between the two options?

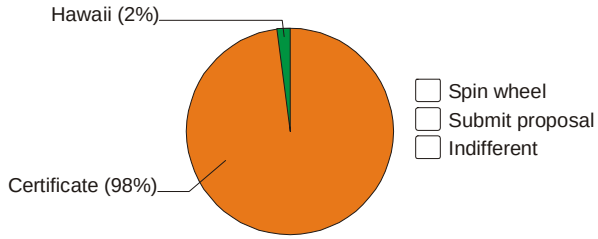
Now suppose a different wheel is used, which looks as follows:

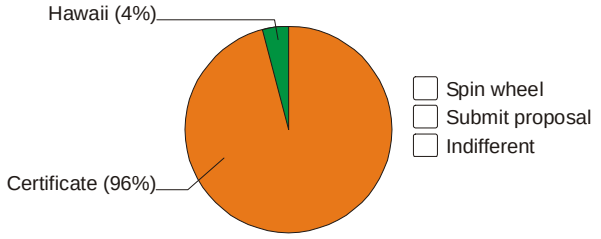


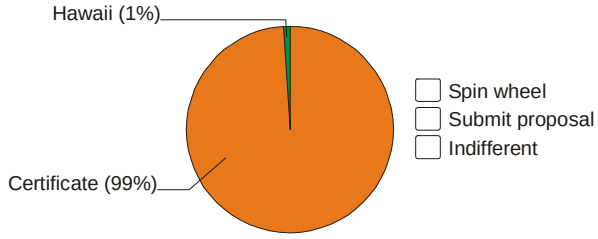
Would you spin the wheel, would you submit the proposal, or are you indifferent between the two options?

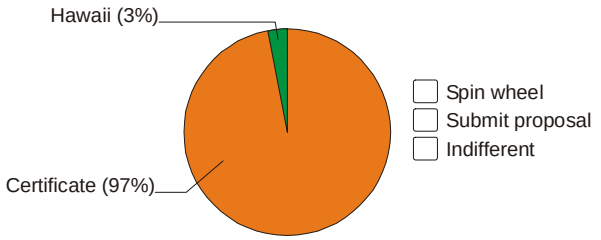
Continuing the above process, for each of the wheels below, indicate if you would (1) spin the wheel, (2) submit the proposal, or (3) you are indifferent between spinning the wheel and submitting the proposal.

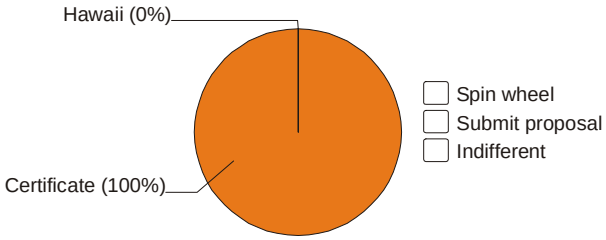




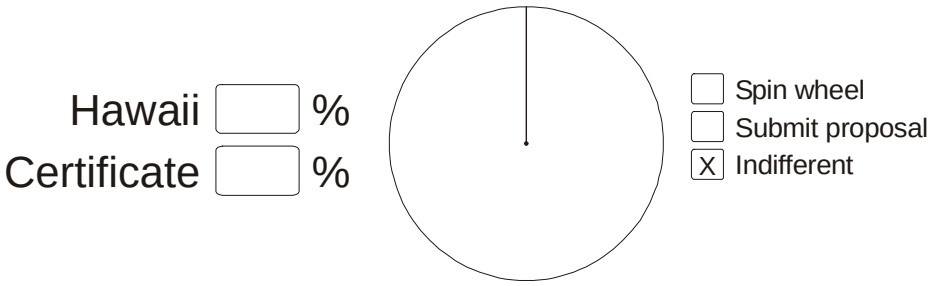








If you have not marked “Indifferent” for any of the above wheels, please complete (and quantify) below the wheel that would make you indifferent to the two ways of entering the competition. That is, what does the wheel have to look like to make you indifferent to entering the competition by spinning the wheel, or entering the competition by submitting a biometric proposal with a FTER of 5%?

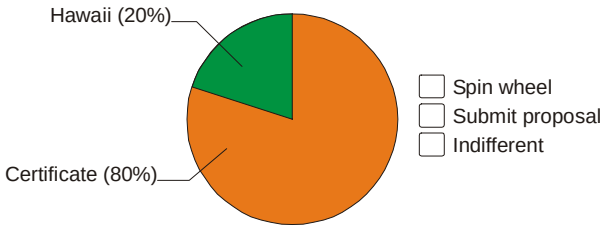


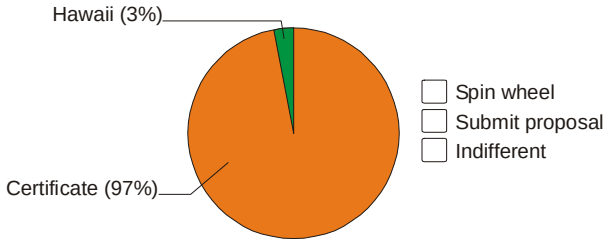
4 EXPERIMENT: FTER = 1%

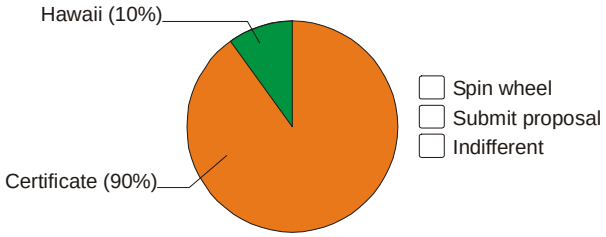
Now, suppose that you can enter the same competition (and win either a trip to Hawaii, or a $10 gift certificate) by doing any of the following:

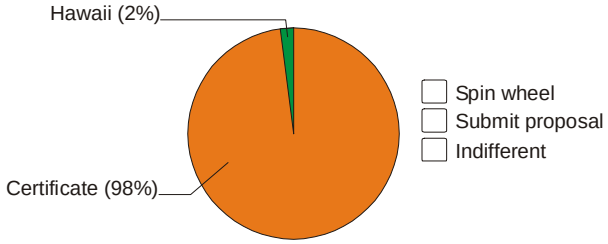
- Spin a “wheel of fortune”, or
- Submit a biometric proposal with a FTER of **1%** to your decision makers. That is, 1 in 100 of applicants cannot be enrolled on the system. If the proposal is accepted, you win the trip to Hawaii. If the proposal is not accepted, you win the $10 gift certificate.

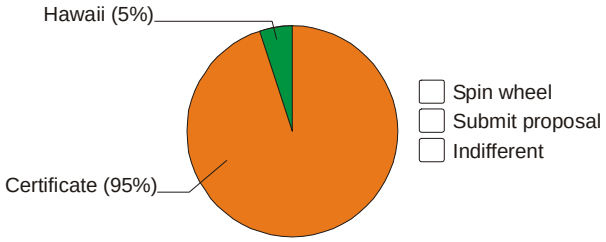
For each of the wheels below, indicate if you would (1) spin the wheel, (2) submit the proposal, or (3) you are indifferent between spinning the wheel and submitting the proposal.

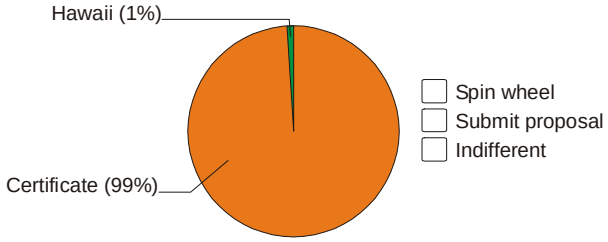


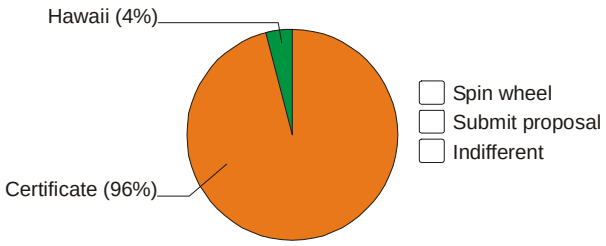


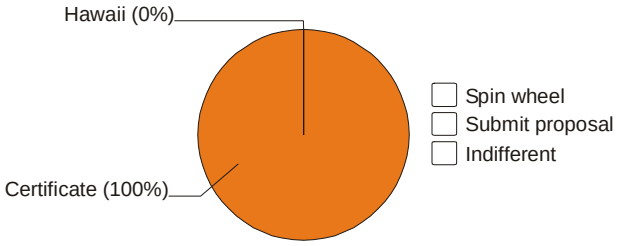




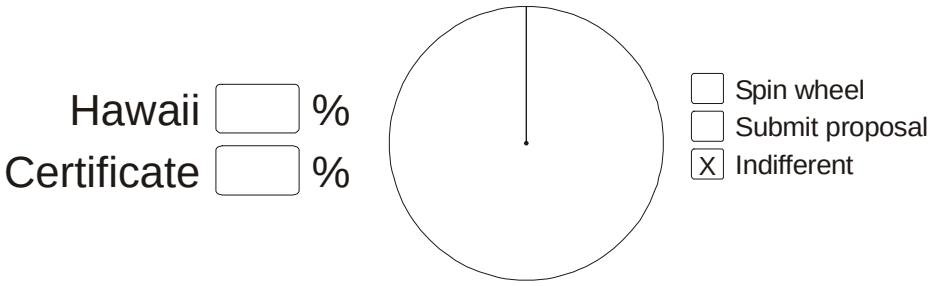








If you have not marked “Indifferent” for any of the above wheels, please complete (and quantify) below the wheel that would make you indifferent to the two ways of entering the competition. That is, what does the wheel have to look like to make you indifferent to entering the competition by spinning the wheel, or entering the competition by submitting a biometric proposal with a FTER of 1%?

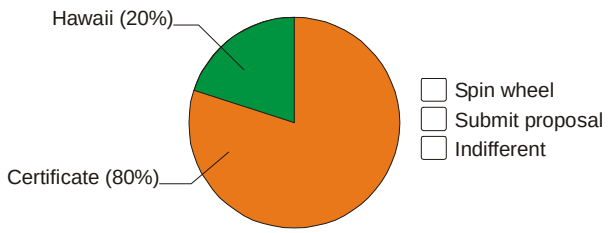


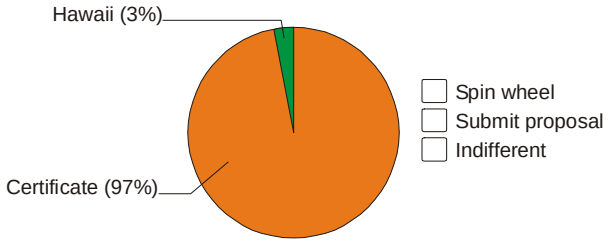
5 EXPERIMENT: FTER = 0.1%

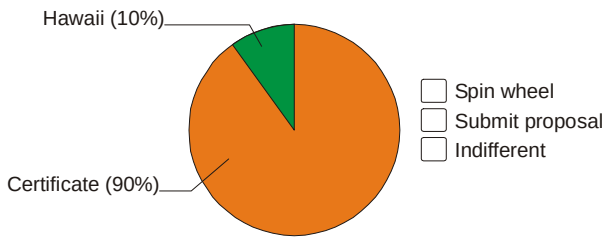
Now, suppose that you can enter the same competition (and win either a trip to Hawaii, or a $10 gift certificate) by doing any of the following:

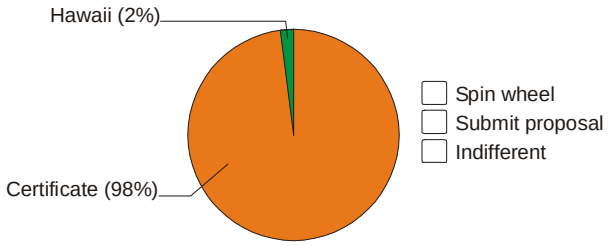
- Spin a “wheel of fortune”, or
- Submit a biometric proposal with a FTER of **0.1%** to your decision makers. That is, 1 in 1000 of applicants cannot be enrolled on the system. If the proposal is accepted, you win the trip to Hawaii. If the proposal is not accepted, you win the $10 gift certificate.

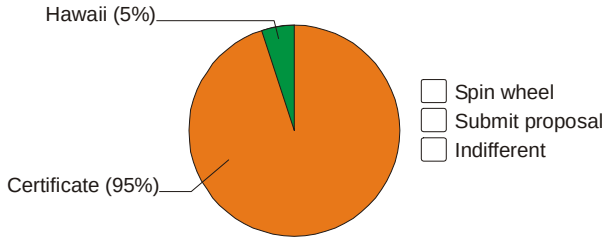
For each of the wheels below, indicate if you would (1) spin the wheel, (2) submit the proposal, or (3) you are indifferent between spinning the wheel and submitting the proposal.

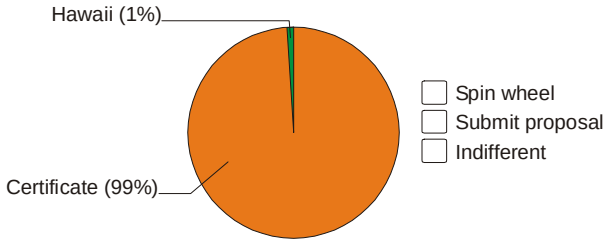


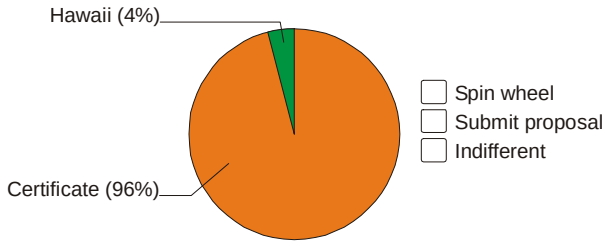


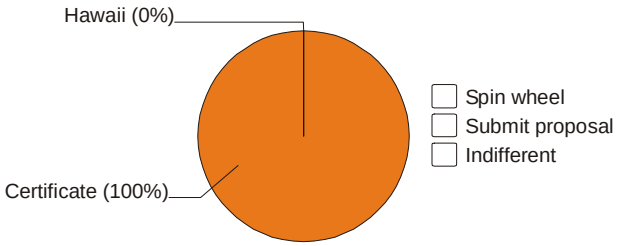




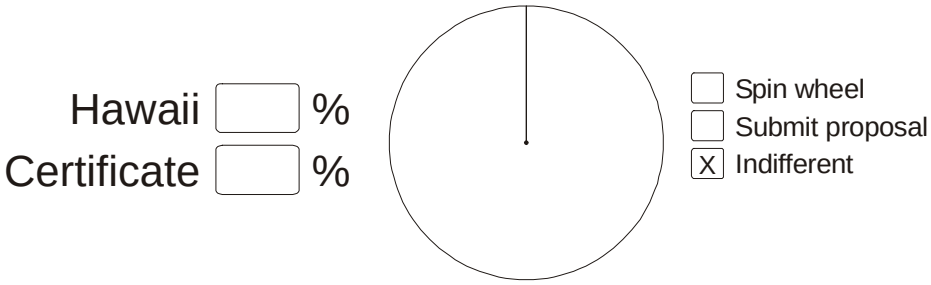








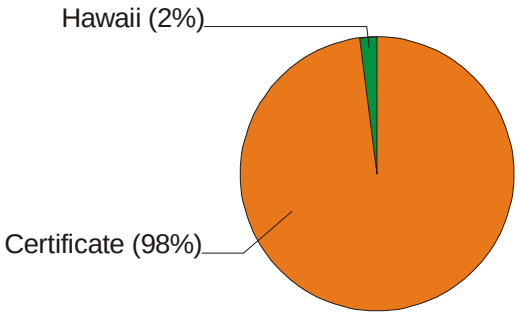
If you have not marked “Indifferent” for any of the above wheels, please complete (and quantify) below the wheel that would make you indifferent to the two ways of entering the competition. That is, what does the wheel have to look like to make you indifferent to entering the competition by spinning the wheel, or entering the competition by submitting a biometric proposal with a FTER of 0.1%?



6 EXPERIMENT: FTER

Now, suppose that you can enter the same competition (and win either a trip to Hawaii, or a $10 gift certificate) by doing any of the following:

- Spin the following “wheel of fortune”:



, or

- Submit a biometric proposal. If the proposal is accepted, you win the trip to Hawaii. If the proposal is not accepted, you win the $10 gift certificate.

If the proposal contains a FTER of 0%, it is assumed that you would choose to submit a proposal rather than spin the wheel of fortune. Now, consider increasing the FTER. What is the highest FTER that you would submit to your decision makers, before you would rather spin the wheel of fortune? Hint: Think of the FTER in terms of a scale that you can relate to, e.g. number of FTEs per 1000 drivers.

Highest FTER: _____