SUMMARY

INTRODUCTION

REQUEST – The University of Twente has been invited by the Research and Documentation Centre – WODC on 12 July 2018 to conduct a study entitled "Biometrics in the aliens' identity chain – a literature study" (project number 2965).

In the Netherlands, the aliens' identity chain is an identity management process, which is part of the migration chain. It relies on two types of personal data: biographic, such as date and place of birth, name and nationality, and biometric, such as fingerprints, face and irides. Biometric technology is used in the aliens' identity chain to (quickly) achieve automated identity verification and identification of aliens (immigrant or resident, legal or illegal).

AIM – This literature study focuses exclusively on the role of biometric data in the aliens' identity chain. At present the fingerprint is the single biometric mode automated for this purpose. It aimed at understanding the capabilities and limits of the fingerprint made in the current operational biometric process and to make an inventory of what is known about the possibilities to improve and/or combine the use of fingerprints and other biometric modes in a multimodal approach. Eventually, the purpose is to improve the verification of identity and identification processes in the aliens' identity chain as applied by the Netherlands.

DESIGN

QUESTIONS – In order to answer to the request, the issues have been operationalised in 5 research questions:

1. What are the current operational biometric processes (border control, alien surveillance and asylum application scenarios)?
2. What is the role of the fingerprint mode in the current operational biometric processes?
3. What are the current reliability (effectiveness and efficiency) issues of the fingerprint mode in the current biometric processes?
4. Which other biometric modes are fit for the intended purposes and what is their reliability when implemented in automated and human-based biometric processes?
5. What is the potential of combining the fingerprint mode with other biometric modes?

METHOD – The study has been organised in two phases: (1) the questions 1, 2 and 3 have been answered on basis of a set of interviews conducted with the stakeholders that are identifying partners in the migration chain:

- Immigration and Naturalisation Agency – IND,
- Royal Netherlands Marechaussee – KMAR,
- National Police – NP,
- Central Organ for Asylum seekers – COA,
- Ministry of Foreign Affairs – BUZA,
- Repatriation and Departure Service – DT&V,
- Custodial Institutions Agency – DJI,

and (2) the questions 4 and 5 have been answered on basis of a review of the scientific literature focusing on human identification, biometrics and their applications. Special attention has been given to the experience gained in the practical implementation of the fingerprint mode and/or its combination with other biometric modes for large-scale identity verification and identification, such as in the Aadhaar Biometric Project, the world's largest biometric ID system developed by the Unique Identification Authority of India (UIDAI).

STRUCTURE – Chapter 2 "Current implementations" reports on the phase 1 of the project. It answers questions 1, 2 and 3 on the categorisation of the different biometric processes existing in the aliens' identity chain and in the description of their specificities.

Chapter 3 "Properties of the biometric modes" reports on the phase 2 of the project. It answers question 4 describing the biometric modes, their strengths and weaknesses and a selection will be made of the

ones that can be included in a multimodal biometric system for the Dutch aliens' identity chain (hereafter alien's identity chain).

Chapter 4 "Discussion" answers question 5, discussing the advantages and drawbacks of the different biometric modes considered for the different biometric processes existing in the aliens' identity chain and their combination.

Chapter 5 "Conclusion" summarises the answers to the 5 research questions and concludes this study.

BIOMETRIC DATA MANAGEMENT

In the Netherlands, the aliens' identity chain is a decentralized infrastructure involving several identifying partners: IND, KMAR, NP, COA, BUZA, DT&V and DJI. These partners exercise dedicated areas of responsibility in the migration chain: admission (*toelating*), border control (*toegang*), monitoring (*toezicht*), housing (*opvang*), detention (*bewaring*), repatriation (*terugkeer*) and naturalisation (*naturalisatie*). These areas of responsibility include both identification and identity verification tasks.

The organisation that registers the biometric data is responsible for their quality. In practice, the biometric data are registered anyway if after at least 3 attempts the result still does not comply with the quality requirements. In the aliens' identity chain, the biometric identity check has primacy over the biographic identity and covers two distinct processes: biometric identification and biometric verification of identity.

All the partners of the aliens' identity chain (IND, KMAR, NP, COA, BUZA, DT&V and DJI) are performing biometric identity check in one of two ways: 1) in form of identity verification (1:1, hereafter biometric verification check) or 2) in form of open-set identification (1:N+1, hereafter biometric identification check). Closed-set identification is almost never applicable in practice because the assumption that the biometric reference of the person checked is in the database cannot be demonstrated.

The most general and critical issue related to biometric data management concerns the identification process. If during the enrolment process no fingerprints are available for automatic deduplication, then the process only relies on the human-based comparison of available biographic data and sometimes a frontal face image. Such situation potentially compromises all the subsequent tasks along the aliens' identity chain. The reason is that the initial absence of biometric data prevents to demonstrate an unequivocal link between the person initially registered and the person interacting later with the aliens' identity chain.

PROPERTIES OF THE BIOMETRIC MODES

In addition to fingerprints that are already used in the aliens' identity chain, there are (many) other biometric modes that allow to distinguish individuals and that can contribute to their identity verification or their identification. Only the following biometric modes have been selected in the request and initially considered for a comparison with the fingerprint mode: face, iris, palm, hand (geometry), ear (shape), retina (vein pattern), speech (speaker) and DNA.

The relevance of one or several biometric modes for a specific application depends on properties of the biometric mode and of the underlying biometric technology. For this study, 14 properties structured in four categories have been considered: intrinsic properties of a biometric mode (universality, permanence, distinctiveness); properties of the underlying biometric technology (maturity, performance, real-time verification, identification speed); properties related to the interaction of the user with the biometric technology (collectability, acceptability, cooperation, invulnerability); and properties related to the integration of a biometric mode within an ID management infrastructure (database availability, scalability and interoperability).

Because of the diverse nature of biometric applications, no single biometric trait is likely to be optimal and satisfy the requirements of all applications. In many cases, a multimodal biometric system combining or fusing multiple biometric traits may be required to attain the desired level of performance. One such example is the very large-scale Aadhaar Biometric Project in India ($>10^9$ individuals), that uses all 10 fingerprints and both irides of subjects for deduplication purpose and identification purpose.

The analysis of the literature reveals some limits in each of the biometric modes selected for this study. For various and specific reasons, the majority of the biometric modes are not fit for an implementation in the aliens' identity chain: implementation costs (palmprint), limited distinctiveness (hand geometry, ear shape), contamination, sensitivity, speed, data protection and privacy issues (DNA), user

acceptance, data protection and privacy issues (retina), limited distinctiveness and sensitivity to environmental conditions (speech). A minority of them can be considered as fit for an implementation in the aliens' identity chain. The fingerprint mode is distinctive, permanent, scalable, interoperable and widely implemented but not universal. The face mode is universal, easy to collect, accepted and interoperable, but its scalability is limited, preventing deduplication in large datasets. The iris mode is distinctive, permanent, scalable and almost universal but it is not interoperable due to the lack of available databases. The fact that none of them is ideal and fail-safe and that each of them encounter specific shortcomings and limits advocates for a multimodal approach.

The fingerprint face and iris modes are largely deployed in a various type of administrative/civil, travel, law enforcement and security applications, due to their properties. The fingerprint mode is already implemented within the different applications the aliens' identity chain. Therefore only the face and iris mode will be considered in this review for replacement or combination with the fingerprint mode. The deployment of the other biometric mode is limited to one type of applications, due to the limitations of some of their properties. DNA and palmprint are exploited primarily in law enforcement for identification, speech and hand geometry have been deployed in commercial applications, mostly for identity verification when ear and retinal vein pattern have been proposed by researchers for biometric recognition in niche applications, but are yet to attain sufficient level of technological maturity and acceptance.

MULTIBIOMETRIC APPROACH INCLUDING FINGERPRINT

Any multimodal biometric approach faces implementation challenges in terms of (1) training and performance – the level of investment is high and the users' acceptance is low (2) design – the number of biometric modes (multi-mode), sensors within each biometric mode (multi-sensor) and number of instance to collect data of a biometric mode (multi-instance/multi-sample) as well as (3) architecture and level of fusion of the results.

COMBINATION OF FINGERPRINT AND FACE MODE – In all studies considered, the performance of the fingerprint and face modes in combination systematically supersedes the performance of each mode when used independently. The same phenomenon is observed for invulnerability: the resistance to the presentation attack of the combination of the fingerprint and face mode supersedes the performance of each mode considered for itself.

The unbalance between the universality and collectability of fingerprints (being limited at 95% of the individuals) and the face mode (being almost universal) advocates for a serial architecture rather than parallel for the combination of these two modes. Organising the identity check in sequence, starting with the face mode and, if necessary, adding the fingerprint mode improves the convenience of the process.

Apart from the technical and methodological challenges directly related to the development and implementation of multimodal biometric systems, other aspects inherent to their complexity need to be considered in order to translate them into operational advancements: higher competence of the personnel, ergonomics, productivity, scalability to large heterogeneous datasets; compliance with security, data protection and privacy requirements; robustness to changes such as the environment, the population or the sensors.

COMBINATION OF FINGERPRINT AND IRIS MODE – Even if a multimodal biometric system combining the face and the iris mode falls beyond the scope of this study, it should be noted that a system which combines these two modes would make sense. The biometric data of both modes could be collected in the same instance and by the same sensor, namely a camera. The face could then be used for identity verification and irides for deduplication and identification purposes, if iris databases become available. Nevertheless, it would also mean investing in a second biometric technology and database with as a result a more complex and more expensive system, requiring more qualified personnel to develop, maintain and operate it.

ANSWERS

ANSWER TO QUESTION 1 – What are the current operational biometric processes (border control, alien surveillance and asylum application scenarios?

The identifying partners in the migration chain operate two biometric processes: (1) identification process and (2) identity verification process. The identification process consists of comparing the biometric data of a requester with the reference data of a biometric database. It is used during the

enrolment process of a requester to avoid the duplication of the same person in the database. In the aliens' migration chain, this process is performed exclusively for the fingerprint mode using Automatic Fingerprint Identification Systems (AFIS). The databases checked depend on the nature of the request; they can contain Dutch or international civil (*Basisvoorziening Vreemdelingen* – BVV, European Visa Information System – EU-VIS) or law enforcement (*Het Automatisch Vinger Afdrukkensysteem Nederlandse Kollektie* – HAVANK, *Strafrechtsketendatabank* – SKDB, European Dactylography System – EuroDAC) data. The identifying partners operating the identification process are IND, KMAR, NP and DJI.

The identity verification process consists of comparing the biometric data of a requester with their own reference data present in an identity document or in a database such as BVV. This process is performed automatically for the fingerprint mode, and by a human being for the face mode. The identifying partners operating the identity verification process are IND, KMAR, NP, COA, DT&V and DJI.

ANSWER TO QUESTION 2 – What is the role of the fingerprint mode in the current operational biometric processes?

The fingerprint mode plays a central role in the current operational biometric processes of the aliens' identity chain. The biometric data has primacy over the biographic data in both identification and identity verification processes. On the one hand, the fingerprint data are distinctive and permanent, the fingerprint mode is reasonably accepted by the users and very large civil and law enforcement fingerprint reference databases exist. The fingerprint biometric technology is mature, performant, fast and it can be integrated due to advanced scalability and interoperability properties.

On the other hand, the universality and collectability of the fingerprint data is limited at about 95% of the individuals, its use requires cooperation and the fingerprint mode can be vulnerable to presentation attacks. The controversies related to data protection and privacy of fingerprint data limit the development of civil fingerprint databases. It also explain why, at international level, the civil identity chain depends on law enforcement fingerprint databases for their identification and deduplication processes.

ANSWER TO QUESTION 3 – What are the current reliability (effectiveness and efficiency) issues of the fingerprint mode in the current biometric processes?

The latest NIST Evaluation of Fingerprint Matching Algorithms shows the potential of the fingerprint mode: the most accurate submission achieved a False Non Identification Rate of 1.97% for the left index finger and 1.9% for the right index finger when searched against an enrolment set of 100 000 subjects (1 million fingerprints), for a False Positive Identification Rate of 10^{-3}. In the Dutch Identity Chain, the Guideline for the Biometric Verification within the Partners of the Migration Chain aims at standardising the collection and at controlling the quality of the biometric data, as well as at monitoring the effectiveness, efficiency and transparency of the biometric processes of the aliens' identity chain.

But apart from this Guideline, there is currently no quality assurance system and accreditation program in place or in development. The parameters influencing the quality of the fingerprint data (age, occupation, location of collection) and face images (facial expression, pose, illumination, occlusion and image resolution) are known, but no documented procedures and quantitative figures are available to control the quality of the data. In the same way, the design and the technology of the biometric processes are constantly improved, but no documented procedures are available to assess quantitatively the performance of the automatic (fingerprint) and human-based (face) identification and identity verification processes.

The most general and critical issue concerns the identification process. If during the enrolment process no fingerprints are available for automatic deduplication, then the process only relies on the human-based comparison of available biographic data and sometimes of a frontal face image. Such situation potentially compromises all the subsequent tasks along the aliens' identity chain preventing the demonstration of an unequivocal link between the person initially registered and the person interacting later with the aliens' identity chain.

ANSWER TO QUESTION 4 – Which other biometric modes are fit for the intended purposes and what is their reliability when implemented in automated and human-based biometric processes?

FACE – The collectability of a face specimen in visible light as 2D representation is straight forward. It simply requires a commercially available camera that can be on almost any mobile device. Deep learning face recognition algorithms have improved rapidly during the last five years, becoming more robust with non-ideal collection conditions, and are yet to reach their maturity. Face recognition

technology moves further towards the strongest modes, even if it is yet to reach the level of performance of the fingerprint or iris modes yet. One critical challenge that remains for deep-learning methods in general and for face recognition methods in particular, is the stability and uniformity of performance, in unconstrained conditions, between all the individuals of a reference database. The face databases (*Centrale Automatische TeChnology voor de Herkenning van personen* – CATCH, EuroDAC, Schengen Information System – SISII, Entry Exit system – EES, European Visa Information System – EU-VIS and European Criminal Records Information System for Third Country Nationals – ECRIS-TCN) and their interoperability are progressing, but they are still yet to reach the level of the fingerprint mode. Therefore, the scalability of the face mode will remain limited in situations whereby face data is collected in unconstrained conditions and because of the intrinsic limitation of the distinctiveness of the face due to genetic factors, inducing physical resemblances (ethnical background and family relationships). The prevalence of monozygotic twins in the human population in particular, has significant implications on the performance of face recognition systems. Finally, the face mode is also vulnerable to presentation and morphing attacks, especially if the identification check process is unsupervised.

IRIS – The universality of the iris ensures a theoretical collectability for almost 100% of the individuals even if, in practice, the user acceptance is lower than for the face. The features of the iris are more permanent than the ones of the face, limiting the necessity of regular re-enrolment. The distinctiveness of the iris is uniform due to the epigenetic nature of the data and the technology is highly performing. The iris is also less vulnerable to attacks than the face and fingerprint. But the implementation of the iris mode has been delayed for various reasons, including a limited acceptability by the users and the cost of proprietary and patented technology.

The potential of the combination of the fingerprint and iris modes has been demonstrated for identification in India by the Aadhaar Biometric Project. However, the development of any other multimodal large-scale identification system including the iris is currently prevented by the absence of any other large-scale iris database, preventing any deduplication process based on the iris mode.

ANSWER TO QUESTION 5 – What is the potential of combining the fingerprint mode with other biometric modes?

For identification processes not only taking place within the BVV database, but that are extended to national (BVV, HAVANK, SKDB) and international (EuroDAC, EU-VIS, SISII) databases, there is currently no alternative to the fingerprint mode. As already mentioned, even if the iris mode is technically a credible alternative, it is currently not a practical option for identification, as iris databases are non-existent. Face databases are still fragmentary and only partially interoperable and the face recognition technology is yet to reach large scalability. Even if the face mode continuously improves, the limit of distinctiveness of the face and of face data collected in non-ideal conditions will be remaining obstacles.

For identification and identity verification processes only taking place within the BVV (standalone), a serial multimodal biometric system combining the face and fingerprint modes would offer several advantages: primarily universality and user acceptance of the face mode, and secondarily performance of the fingerprint mode, used in case on failure of the face mode only. It has to be noted that such a multimodal system would improve the integrity of the BVV database.

Multimodal biometric systems combining the fingerprint and face modes have been mainly developed for identity verification, with performance systematically superseding performance of each mode when used independently. The same phenomenon is observed for invulnerability. Resistance to presentation attack of the combination of fingerprint and face modes supersedes performance of each mode considered for itself. Such results can be explained by the complementarity of the two modes. A serial multimodal biometric system, exploiting primarily the face mode and secondarily the fingerprint mode would offer the advantages of universality and user acceptance of the face mode; and only in case of failure, the distinctiveness, permanence and performance of the fingerprint mode would be exploited.

Finally, it has to be noted that investing in a second biometric technology and database to develop a multimodal approach will result in a more complex and more expensive system, requiring more qualified personnel to develop, maintain and operate it.