EDPS
EUROPEAN DATA PROTECTION SUPERVISOR

EDPS OPINION ON eu-LISA DPIA on the Shared Biometric Matching Service (Case 2021-0757)

1. PROCEEDINGS

On 28 July 2021, the EDPS received a request from eu-LISA on the high risk stemming from the use of biometric matching technologies in the Entry Exist System ('EES') and the Shared Biometric Matching Service ('sBMS') and related measures to mitigate it.

The consultation sent by eu-LISA contains the sBMS DPIA and the sBMS DPIA for Accuracy Measurement. The following additional documents were provided from a dedicated eu-LISA SharePoint side:

- Annex_III_EES Data Protection Impact Assessment
- Annex_IV_Evaluation and Improvement of eu-LISA Synthetic Biometric Dataset
- Annex_V_VIS Face Image Quality Assessment
- Annex_VI_EES Security Risk Assessment and Annexes from A to I
- Annex_VII Decision of the Management Board concerning the EES PMB recommendation on the EES technical specifications
- Annex VIII eu-LISA Single Programming Document 2020 - 2022
- Annex IX eu-LISA Single Programming Document 2021- 2023.

Eu-LISA consults the EDPS as it is of the opinion that:

- the high risks resulting from the specific processing related to the biometric data engine cannot be mitigated without a proper testing of the biometric matching engine acquired from an external vendor and,
- that only testing with representative real biometric data will offer enough assurance on the validity of this accuracy testing.

The EDPS notes that eu-LISA does not specify the legal basis for its consultation. To the extent that the sBMS DPIA identifies high risks, which after applying mitigation measures remain high, the EDPS considers this consultation as a prior consultation under Article 40 (1) of Regulation 2018/1725[1] ('the Regulation').

[1] Regulation (EU) 2018/1725 of the European Parliament and the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC, OJ, L 295, 21.11.2018, pp. 39-98.

According to Article 40(2) of the Regulation, the EDPS has to issue his Opinion within a period of up to eight weeks of receipt of the request for consultation, with a possible extension by six weeks.

Given the complexity of the intended processing due to the use of innovative technology (machine learning) applied to sensitive data (biometric data) on a large scale and illustrated by the lengthy documentation attached to the consultation, the EDPS decided to extend the deadline with 6 weeks and informed eu-LISA on 27 August 2021.[2]

The initial deadline for the EDPS to provide written advice was 'up to 8 weeks of receipt of the request for consultation', namely 24 September. Following the deadline extension, the final deadline for the EDPS to provide his advice is the **4th of November 2021**.

2. DESCRIPTION OF THE PROCESSING

Pursuant to Article 1 (3) and (4) of Regulation 2018/2226[3] (the 'eu-LISA Regulation'), eu-LISA is responsible for the operational management of the Visa Information System (the 'VIS') and the preparation, development or operational management of the Entry/Exit System (the 'EES').
Regulations 2019/817[4] and 2019/818[5] (the 'Interoperability Regulations) further establish a framework to ensure interoperability between the EES, the Visa Information System (VIS) as well as other EU large-scale systems[6].

The shared Biometric matching Service (the 'sBMS') is an interoperability component prescribed by the Interoperability Regulations and to be developed under the responsibility of eu-LISA.[7] It aims at facilitating the identification of an individual who is registered in different databases by using a single technological component to match that individual's biometric data across different systems.[8] It will store templates of the biometric data (fingerprints and facial images) contained in the EU centralised information systems to enable on one hand, a simultaneously search biometric data stored in the different systems and on the other hand, a comparison of these data.[9]

[2] Email of 27.08.21 addressed to Mr Ramat, our reference D(2021) 1781

[3] Regulation 2018/1726 of the European Parliament and of the Council of 14 November 20018 on the European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice (eu-LISA), and amending Regulation (EC) No 1987/2006 and Council Decision 2007/533/JHA and repealing Regulation (EU) No 1077/2011, OJ, L295,21.11.2018.

[4] Regulation 2019/817 of the European Parliament and of the Council of 20 May 2019 on establishing a framework for interoperability between EU information systems in the field of borders and visa and amending Regulations (EC) No 767/2008, (EU) 2016/399, (EU) 2017/2226, (EU) 2018/1240, (EU) 2018/1726 and (EU) 2018/1861 of the European Parliament and of the Council and Council Decisions 2004/512/EC and 2008/633/JHA, OJ, L135, 22.05.2019.

[5] Regulation 2019/818 of the European Parliament and of the Council of 20 May 2019 on establishing a framework for interoperability between EU information systems in the field of police and judicial cooperation, asylum and migration and amending Regulations (EU) 2018/1726, (EU) 2018/1862 and (EU) 2019/816, OJ, L135, 22.05.2019.

[6] The European Travel Authorisation Information System (ETIAS), Eurodac, the Schengen Information System (SIS) and the European Criminal Records Information System for third country nationals (ECRIS-TCN)

[7] Article 12 (3) of the interoperability Regulations.

[8] Recital 18 of the Interoperability Regulations.

[9] Chapter III of the Interoperability Regulations.

3. PRIOR CONSULTATION UNDER ARTICLE 40

3.1. The threshold assessment and the DPIA

Under Article 39 of the Regulation, where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller must - prior to the processing - carry out an assessment of the impact of the envisaged processing operations on the protection of personal data.

Under Article 39(4) of the Regulation, the EDPS has established a positive list of processing operations that prima facie require a DPIA as well as a list of criteria for assessing whether a processing operation is likely to result in high risks and should therefore be subject to a DPIA.[10]

In Chapter 4 of the sBMS DPIA, eu-LISA has identified one processing operation from the EDPS' positive list (i.e. the large scale processing of special categories of personal data) justifying the need for a DPIA.[11]

Although Chapter 4 does not further explain the reasons to carry out a DPIA, the EDPS notes that the executive summary of the DPIA also considered the following criteria triggering the need for a DPIA:

- the high sensitivity of the data,
- the involvement of vulnerable data subjects,
- the large scale of processing operations,
- the use of innovative technology and automated decision making processes producing legal or similar significant effects on data subjects and,
- the fact that a machine learning based solution may perform unanticipated personal data combinations or process data for different purposes other than those which are expected.

3.2. Need for a prior consultation under Article 40 of the Regulation

Under Article 40(1) of the Regulation, the controller must consult the EDPS - prior to processing - where a DPIA under Article 39 indicates that the processing would, in the absence of safeguards, security measures and mechanisms to mitigate the risk, result in a high risk to the rights and freedoms of natural persons and the controller is of the opinion that the risk cannot be mitigated by reasonable means in view of the available technologies and costs of implementation.

Eu-LISA has identified in the sBMS DPIA a set of high risks to the rights and freedoms of natural persons that remain high (significant or maximum) even after application of proposed mitigation measures. Since eu-LISA is of the opinion that it could not or not sufficiently mitigate the identified high risks by reasonable means in view of the available

[10] Decision of the European Data Protection Supervisor of 16 July 2019 on DPIA lists issued under Articles 39(4) and 39(5) of Regulation (EU) 2018/1725.

[11] See Chapter 4 Reasons for this DPIA on page 19-20 of the sBMS DPIA.

technologies and costs of implementation, it was obliged to consult the EDPS in accordance with Article 40 of the Regulation.

3.3. Scope of the Opinion

This Opinion **focuses on the high risks generated by the data processing related to the sBMS in the context of the EES and the VIS as described in the notification** of eu-LISA and the accompanying documentation.[12] It analyses key aspects in this respect that raise issues of compliance with the applicable data protection legal framework or otherwise merit further analysis.

In addition, the Opinion provides some recommendations on lower risks. The EDPS stresses that these recommendations aim at improving the DPIA but should not be considered as exhaustive to ensure compliance with the Regulation.

The EDPS expects to be consulted on any significant update of the DPIA, as a result of a substantial modification of the personal data processing operations at stake.

4. ANALYSIS

4.1. General remarks on sBMS DPIA

While the focus of the EDPS in this Opinion is on residual high risks (significant, maximum) identified from the sBMS DPIA (prior consultation of article 40 of the Regulation), this section provides some general information and recommendations for improvement of the DPIA. These recommendations concern the scope of the sBMS DPIA and elements that need to be clarified or improved and should not be considered as exhaustive recommendations on the DPIA.

4.1.1. Scope of sBMS DPIA - Assignment of risks to different systems

Eu-LISA has identified 100 (one hundred) data protection risks to data subjects (and the organization) in the sBMS DPIA. The EDPS is of the view that some of the risks attributed to sBMS are not in fact stemming from it, but from the systems using it (e.g. EES, VIS, SIS II). Examples include the risk of producing automated decisions with an impact to data subjects, without ensuring the right of human intervention and the use of sBMS services for secondary purposes (abuse of rights).

The EDPS notes that the risk attribution should be done to the system where the mitigation measures can and will be applied, to ensure thorough mitigation. For instance, audit logs for

[12] The sBMS DPIA (p.10-11) indicates that its scope 'is strictly limited to the biometric handled by the EES and the VIS and transmitted to the sBMS' and that it ' assesses the risks of data processing in regard to the use of biometric templates resulting from real EES and VIS biometric data in the sBMS. It assesses the specific data protection risks inferred from EES or from VIS, which impact the sBMS but does not pretend to assess the specific risks inferred from the above-mentioned systems and the three other EU information systems such as Eurodac, SISII and ECRIS-TCN'.

users' extensive and abnormal use of sBMS services could better be audited by establishing logs and audit procedures on the systems using the sBMS, such as the EES. If mitigation measures can also be applied to sBMS to complete such measures, these should be welcome.

Recommendation 01:

The EDPS recommends eu-LISA to review the list of risks, distinguish the risks explicitly related to the sBMS and ensure that risks related to its use are included in the DPIAs of the systems/processes using sBMS.

4.1.1.1. Use of personal data for unspecified purposes or for purposes which are incompatible with those originally declared (Risks 52 and 57)

Eu-LISA identifies a risk of using personal data in sBMS for secondary purposes, incompatible with the purposes provided for in the legal framework. Examples include behavior monitoring and use for decision making in different contexts such as financial, social or employment. For the mitigation of this risk, eu-LISA has proposed measures such as provision of clear and transparent information about the purpose for which data are processed. They also propose to provide information in the criteria used in decision making and allow the data subject to challenge the decision, set up regular quality assurance checks of the system to ensure it does not lead to discrimination and accuracy measurement.

Recommendation 02:

The EDPS understands that this risk is related to use of sBMS services, by the controllers in the Member States and expects eu-LISA to provide access only to necessary services to each controller, in alignment with the legal framework allowing for such access and to simultaneously log and audit the use of sBMS by other systems. At the same time, the EDPS recommends eu-LISA to reassess these risks in the DPIAs of systems using sBMS.

4.1.1.2. Risks related to the right of human intervention (Risks 54 and 59)

Eu-LISA has identified a risk of "Automated decision-making with possible relevant consequences for individuals". The EDPS understands this risk concerns the sBMS use of machine learning models to assign an identity to an individual without any human intervention. As described above, since sBMS will always be used in conjunction with other IT systems (initially with the EES), eu-LISA should assess this risks in the DPIA of the IT systems using sBMS.

The EDPS notes that in the EES DPIA[13], eu-LISA has identified the risk of lack of transparency for automated individual decision (R8), as well as the risk of technical errors that may result in incorrect decisions for the data subjects (R14). However, the EES DPIA is missing the risk of data subjects not being able to exercise their right to human intervention (Article 24 of the Regulation), and the risk of not having meaningful human intervention, i.e. the low likelihood of the end users (of the EES) to challenge the result of the algorithm (automation bias).

[13] EES Core Project - EES Data Protection Impact Assessment, v. 05_01_00, 23/06/2021.

Recommendation 03:
The EDPS recommends that:
-Eu-LISA includes, assesses and treats the the risk of data subjects not being able to exercise their right to human intervention (Article 24 of the Regulation), and the risk of not having meaningful human intervention, due to automation bias.
-Eu-LISA ensures controllers of other systems/processes using the sBMS are adequately informed of these risks so they can include them in their DPIAs and take adequate mitigation measures, such as relevant user manuals for any steps of human intervention (e.g. second line of border checks). In case the design of sBMS allows, the confidence of the sBMS match could be displayed to the end users of the other systems, along with the matching result.
- Eu-LISA provides training material to the controllers so that end users learn the capacities and limitations of the sBMS and can critically challenge its outcome.

4.1.2. Unclear Description of risks and/or mitigating measures

4.1.2.1. Decrease in the likelihood that people exercise their fundamental rights (Risks 38, 53, 58, 63, and 68)

Eu-LISA has identified the risk of the sBMS processing operations to decrease the likelihood that people exercise their fundamental rights, such as respect for private and family life, protection of personal data and non-discrimination.

While the EDPS understands that this risk covers the whole range of risks the DPIA needs to assess (risks to the freedom and fundamental rights of data subjects introduced by this processing), it is not clear how biometric identification or verification process will reduce the likelihood of data subjects exercising their fundamental rights.

The risk lies in the processes in which the sBMS services will be used, which is the subject of these processes or systems' DPIAs.

Recommendation 04:
The EDPS recommends eu-LISA revisit, clarify and reassess this risk.

4.1.2.2. Reliance on low confidence outputs (Risk 92)

Eu-LISA has identified a risk of 'Reliance on low confidence outputs.' From the provided description of the risk and of the processing, it is not clear if the sBMS will provide outputs (verification, identification) regardless of potentially low reliability scope on the matching results and whether the end users (of EES) will be provided with a confidence or reliability score on the matching results.

Recommendation 05:
The EDPS therefore asks eu-LISA to clarify and reassess this risk.

4.1.3. Unclear impact of mitigation measures

In the sBMS DPIA, although eu-LISA proposes several mitigation measures, these measures proposed have no effect in reducing the risk to the data subjects. An example is the risk of malfunction of the matching system causing false positives or false negatives, with an initial risk assessed as Maximum. While eu-LISA has proposed two mitigating measures (mechanisms and procedures to enable quick and efficient error fixing and periodic audit of the algorithm along with accuracy measurements), neither the likelihood or the impact of the risk has been assessed as lower. This could lead to wrong assumptions on the inefficiency of the proposed measures.

Recommendation 06:

The EDPS recommends for such cases thorough re-assessment of the likelihood and impact of the risk and where the residual risk remains the same even after the application of proposed mitigation measures, to explain verbally the effect of such measures.

4.2. Overview of residual high risks

The EDPS notes 6 risks of sBMS initially identified as high (i.e. maximum or significant) still remain high after the application of proposed mitigation measures (i.e. Risks 56, 61, 66, 71, 76, 83). These are the risks subject to Article 40 of the Regulation.

4.2.1. Malfunction of the matching system causing false positives or false negatives (Risks 56, 61, 66, 71, 76)

This risk concerns the possible low performance of the matching engine (false positives and false negatives[14]), due to either errors, or unintended behavior producing erroneous or unjustified results. Eu-LISA plans to mitigate the risks by establishing mechanisms to quickly fix errors and a process to periodically test the performance of the algorithms (accuracy measurement). The latter is based on the accuracy measurement process explained in annex II[15] which was shared with EDPS, in the context of prior consultation to address low performance risks prior to the entry into operation of the sBMS. The EDPS further analyses the risks of this procedure in section 4.4.

The EDPS notes the severity of this risk and welcomes the proposed measures, which will enable to identify low performance of the matching algorithms and arrange for fixing errors. The EDPS also agrees with the assessment of high impact from false positives or false negatives in the matching of data subjects' biometrics that could lead to wrong identification. At the same time, the EDPS is aware that any algorithm will not be able to reach 100% accuracy[16] and thus agrees that the risk to data subjects' fundamental rights and freedoms will remain high even after the mitigation measures have been applied.

[14] Acceptable levels of accuracy of the matching algorithms are defined in Commission implementing Decision (EU) 2019/329 of 25 February 2019 laying down the specifications for the quality, resolution and use of fingerprints and facial image for biometric verification and identification in the Entry/Exit System (EES), OJ, L57, 26.02.2019. The term false positives in the DPIA refers to false matching rates (FMR) and false negatives to false non-matching rates (FNMR).

[15] Shared Biometric Matching service (sBMS) - Data Protection Impact assessment for Accuracy Measurement, v. 06_00_40, 05/07/2021.

[16] Acceptable levels of accuracy of the matching algorithms are defined in Commission implementing Decision (EU) 2019/329 of 25 February 2019 laying down the specifications for the quality, resolution and use of fingerprints and facial image for biometric verification and identification in the Entry/Exit System (EES), OJ, L57, 26.02.2019.

Recommendation 07:

The EDPS recommends that eu-LISA:

- Introduce in all processes / systems using the sBMS mitigation measures to ensure alternative means of identification to challenge any false result (e.g. manual identification by a border guard) and minimize the impact to the data subject and
- Introduce procedures and mechanisms to provide feedback on the matching algorithms' false outputs in order to trigger the procedures of fixing errors in sBMS.
- Implement recommendations on Accuracy measurement Procedure (section 4.4)

4.2.2. Whitebox Inference of training data from examination of model state (e.g. examination of weights (Risk 83)

Eu-LISA has identified a risk of whitebox inference[17] for individuals that were part of the training dataset, since model data structures are available to third parties. Despite the proposal of a set of mitigation measures, such as restriction to access the neural network, access restriction to only necessary services of the system, and application of differential privacy, the risk was not reduced.

The EDPS recognises that by employing a commercial matching algorithm, the model can be accessible by third-parties and the risk of whitebox inference can materialise. The EDPS also understands that for this risk to materialize, the model (AI matching algorithm) will store training data. The whitebox inference risk can have major impact on data subjects, if the membership to the training database reveals further personal information. For instance, if an algorithm is trained with a dataset of convicted persons, or people having a specific medical or financial condition, the inference data a data subject was part of the training dataset would reveal information on his/her conviction or medical condition or financial condition.

In the context of sBMS, third-parties are considered the entities that have access to its services, authorised EUIs and Member State Authorities. Any other part would be denied access, by the security measures already described in risk 26 "Lack of control in EU information systems access management allowing to access to sBMS" (safeguarding transmission in the backend - using TLS that also authenticates the systems, install Intrusion Prevention and Intrusion Detection systems, apply network separation-zoning).

In the context of the sBMS use and the fact the model is trained by the company that developed it, with a dataset not provided by eu-LISA, the EDPS sees no high risk of third-parties (authorised authorities) to identify if a data subject's biometrics were part of the

[17] The term white box inference refers to an attack where the attacker has knowledge of the model's parameters and tries to identify whether an element was included in the dataset used to train the algorithm. By this, the attacker can make inferences of an individual's personal data. For instance, if the attacker knows that an algorithm was trained with real personal data of patients with a specific disease and has access to the parameters of the model, by employing white box inference attacks, he/she could identify if a person was part of the training dataset and as a result would infer health data of this person.

commercial training dataset. This knowledge by the authorised Member States Authorities would not result in high impact to the data subjects.

Recommendation 08:
As a result, the EDPS considers that this risk can be reduced to limited, provided that mitigation proposed mitigation measures are successfully implemented.

4.3. Risks and measures missing from the sBMS DPIA

This section provides a set of elements missing from the SBMS DPIA, including overlooked risks.

4.3.1. Risks related to bias

Considering the high impact for data subjects from the sBMS and its use in the context of large scale IT systems, it is important to ensure high performance (high level of accuracy) of the matching algorithms. However, the employment of machine learning technology creates new risks that should be taken into account. The EDPS welcomes the fact that eu-LISA identifies specific risks to the technology to be applied, such as Whitebox inference of training dataset, model evasion and model inversion (section 9.2 of Annex II[18]).

Although the risk of not selecting representative data on measuring the accuracy of the algorithm, has been identified in the Accuracy measurement DPIA (annex III)[19], which is proposed as a mitigation measure for low performance risks of the algorithms, the EDPS notes the lack of reference to bias related risks, such as gender, ethnic origin and age bias in the sBMS DPIA (Annex II).

Even if eu-LISA manages to select a representative data set for the accuracy measurement exercise, an overall accuracy measure will not be enough. Many facial recognition algorithms perform worse on certain demographics.[20] Fingerprint recognition does not equally perform in all age ranges.[21] It could be perfectly possible that the overall accuracy for the system complies with the legal threshold but that the accuracy for certain groups such as dark skinned females is far below that threshold. It will be necessary to measure not only the overall system performance, but also the performance for groups to avoid bias. It is important for eu-LISA to measure the performance of the algorithms, but at the same time to ensure that these results do not differ significantly across different groups of people sharing different characteristics (e.g. age).

[18] Shared Biometric Matching service (sBMS) - Data Protection Impact Assessment, CD044-002, 15/01/2020.

[19] Shared Biometric Matching service (sBMS) - Data Protection Impact assessment for Accuracy Measurement, v. 06_00_40, 05/07/2021.

[20] Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[21] Galbally, Javier & Haraksim, Rudolf & Beslay, Laurent. (2018). A Study of Age and Ageing in Fingerprint Biometrics. IEEE Transactions on Information Forensics and Security. PP. 1-1. 10.1109/TIFS.2018.2878160. https://www.researchgate.net/ publication/328526153_A_Study_of_Age_and_Ageing_in_Fingerprint_Biometrics

Recommendation 09:
The EDPS recommends eu-LISA to include in the sBMS DPIA the risks stemming from potential bias, at the very least gender, ethnic origin and age bias.

4.3.2. Risks related to low performance of the matching algorithms, due to inadequate data used in accuracy measurements.

The EDPS notes that eu-LISA has identified risks from malfunction of the matching system causing false positives or false negatives (Risks 56, 61, 66, 71, 76). One of the mitigation measures proposed, to mitigate such risks prior to the entry into operation of the sBMS and the EES, is the accuracy measurement procedure, further analysed in section 4.4. The EDPS notes that in eu-LISA's proposal for the accuracy measurement of the matching engine for facial images, due to lack of EES data, images from VISA applications will be used. The EDPS understands that such data may differ in quality, angle and lighting from the images that will be captures in the EES e-gates. Given this difference of the data, there is a risk that the performance of the matching algorithm is lower than experienced during the accuracy measurements.

Recommendation 10:
The EDPS recommends eu-LISA to include this risk in the sBMS DPIA and adopt mitigation measures, such as re-measuring the performance when the sBMS goes live.

4.3.3. Risks to synchronization of sBMS and other large scale IT system databases

Eu-LISA considers risks with respect to data deletion (50 and 51) and data retention (78 and 79), concerning deletion of data from the sBMS when they are deleted from the other systems. However, it is not clear whether eu-LISA has assessed the risks of non-synchronicity of the databases, meaning data not being deleted or updated at the same time in the original databases and in the sBMS.

Recommendation 11:
The EDPS recommends eu-LISA to include this risk in the sBMS DPIA.

4.4. Analysis of the sBMS Accuracy Measurement Procedure

The sBMS Accuracy Measurement procedure is introduced by eu-LISA as one of the mitigation measures for the risk of low performance of sBMS (risks 56, 61, 66, 71 and 76[22]). The Accuracy Measurement procedure consists of two flows with similar steps, one for testing the accuracy of fingerprint matching and one for facial images matching.

Eu-LISA plans to apply this procedure at least once before the entry into operation of the sBMS, to ensure that the contractual clauses regarding the performance (accuracy) of the matching algorithms are met by the contractor. Since no third-country national's data from

[22] Shared Biometric Matching service (sBMS) - Data Protection Impact assessment for Accuracy Measurement, v. 06_00_40, 05/07/2021.

EES will be available, eu-LISA announced that they plan to use converted VIS data to perform the tests. After the entry into operation of the sBMS, eu-LISA plans to apply a similar process, to periodically measure and monitor the performance of the matching algorithms.

The EDPS notes that the accuracy testing procedure will be performed in a logically separate environment, in the production environment[23], with the same, but undersized matching engines. Eu-LISA has foreseen measures to prevent any risks to existing data, such as the build of dedicated databases and datasets from real data from VIS and security measures as regards the access control to the environment and the provided services. Also, the EDPS notes that eu-LISA has defined the general steps of the process to be followed, while the measurement protocols will be defined in a future step by the MS biometric Experts and JRC experts.

4.4.1. Lawfulness

In this prior consultation, eu-LISA proposes to use personal data sampled from VIS production database, in order to measure the accuracy of the matching algorithms of the sBMS, in the context of EES operations.

In that regard, the EDPS recalls its general position against the use of production data for testing purposes in the absence of a clear legal basis. Where possible, artificially created test data should be used, or test data which is derived from real data so that its structure is preserved but no actual personal data is contained in it. [24] The risks are particularly high in here because of the nature of the production data at stake, namely biometric data.

The EDPS notes that neither the eu-LISA Regulation nor the Interoperability Regulations provide for specific rules on the use of real (production) biometric data for the purpose of accuracy measurements of matching algorithms for the purposes of the development of the sBMS.

This processing operation constitutes a further processing of the data for a purpose other than that for which the personal data have been collected. As such, it should comply with the conditions defined under Article 6 of the Regulation The EDPS notes that eu-LISA considers that such further processing of VIS data meets the criteria of reasonable expectations of data subjects as “the data subjects may reasonably expect that, where necessary, their personal data may be processed to ensure the correct operation of these large-scale IT systems, and that by doing so, the appropriate safeguards will be implemented into the systems”. The EDPS however considers it doubtful that data subjects may have such expectations.

The EDPS notes that pursuant to Article 36 (1) (a) of eu-LISA Regulation allows eu-LISA to process personal data where necessary for the performance of its tasks related to the operational management of large scale IT systems. Pursuant to Article 1 (3) and (4) of the eu-

[23] Shared Biometric Matching service (sBMS) - Data Protection Impact assessment for Accuracy Measurement, v. 06_00_40, 05/07/2021., page 190.

[24] EDPS guidelines on the protection of personal data in IT governance and IT management of EU institutions (para. 80-81), available at https://edps.europa.eu/sites/default/files/publication/it_governance_management_en.pdf .

LISA Regulation, eu-LISA is responsible for the operational management of the VIS and the preparation, development or operational management of the EES.

The Interoperability Regulations further tasks eu-LISA to develop the sBMS as interoperability component between (among others) the EES and the VIS that will store biometric templates obtained from the biometric data of each system, and will enable searching and comparing biometric data stored in each system[25]. Eu-LISA is also legally required to ensure compliance with the accuracy and quality thresholds set by Commission implementing Decision (EU) 2019/329 of 25 February 2019[26]. This requirement aims at mitigating risks for data subjects by setting maximums on the matching engine false non-matching rate and to eu-LISA by reducing the false setting maximums on the matching engine false matching rate.

The EDPS understands that even if the vendors' product accuracy would be perfectly measured, the matching engine could still be subject to representation bias, i.e. the vendor training dataset not being representative of the EES dataset. To avoid it, eu-LISA should compare the distribution of the training dataset demographic features with the ones eu-LISA expects to have in EES (the EES is not in production yet). This comparison also requires that both the vendor and eu-LISA use common values for the demographics that need to be compared, which is not always easy (e.g. there is no ethnic origin encoding standard). It is unclear to the EDPS to what extent this representativeness test is feasible. Eu-LISA could of course just trust the vendor's accuracy measures and monitor the sBMS accuracy once the EES is put in production. However, there is a risk that the sBMS accuracy does not meet the legal requirements, with the related impact on data subjects being wrongly identified,

The EDPS further notes that eu-LISA has assessed the use of synthetic fingerprints as an alternative to using real personal data in the context of biometric matching performance tests. The outcome of such assessment is that the suitability of synthetic data as substitutes to real fingerprints is questionable[27]. eu-LISA concludes that the risks deriving from the use of biometric matching algorithms cannot be mitigated by using artificially created synthetic data and that the use of real biometric data from the representative data subjects is strictly necessary to ensure the adequate testing of the matching algorithms to obtain reasonable assurance concerning their accuracy and quality. The EDPS is missing an equivalent assessment regarding the validity of synthetic facial image data for matching engine performance validation.

The EDPS finally notes that the accuracy testing procedure will be performed in a logically separate environment, in the production environment[28], with the same, but undersized matching engines. Eu-LISA has foreseen measures to prevent any risks to existing data, such as the build of dedicated databases and datasets from real data from VIS and security

[25] Chapter III of Regulation (EU) 2019/817 of the European Parliament and of the Council, of 20 May 2019, on establishing a framework for interoperability between EU information systems in the field of borders and visa and amending Regulations (EC) No 767/2008, (EU) 2016/399, (EU) 2017/2226, (EU) 2018/1240, (EU) 2018/1726 and (EU) 2018/1861 of the European Parliament and of the Council and Council Decisions 2004/512/EC and 2008/633/JHA.

[26] Commission implementing Decision (EU) 2019/329 of 25 February 2019 laying down the specifications for the quality, resolution and use of fingerprints and facial image for biometric verification and identification in the Entry/Exit System (EES), OJ, L57, 26.02.2019.

[27] Evaluation and Improvement of eu-LISA Synthetic Biometric Dataset, v0.0.1,23/02/2021

[28] Shared Biometric Matching service (sBMS) - Data Protection Impact assessment for Accuracy Measurement, v. 06_00_40, 05/07/2021., page 190.

measures as regards the access control to the environment and the provided services. Also, the EDPS notes that eu-LISA has defined the general steps of the process to be followed, while the measurement protocols will be defined in a future step by the MS biometric Experts and JRC experts.

In light of the above, the EDPS understands that the use of fingerprints sampled from VIS production database is strictly necessary in order to measure the accuracy of the matching algorithms of the sBMS, in the context of EES operations. No alternative measures, in particular the use of synthetic data, could provide assurance that the sBMS reaches the level of performance required in Commission implementing Decision (EU) 2019/329 of 25 February 2019. Failing to do so would create higher risks for the data subjects due to wrong identification. Finally, the EDPS notes that eu-LISA will implement adequate safeguards to limit the processing of the sampled data to what is strictly necessary and to prevent risks of misuses. For all these reasons, the further processing of sampled fingerprints from VIS to develop the sBMS and meet the related legal requirement of accuracy and quality threshold could be deemed compatible with the original purpose of collection.

Recommendation 12:
Given eu-LISA's necessity to comply with its legal tasks to develop the sBMS including the related legal accuracy requirements, the risks for data subjects and eu-LISA of not doing so and the unsuitability of fingerprint synthetic data for ensuring the matching engine's accuracy, the EDPS considers justified eu-LISA's use of sampled VIS production data with the purpose of ensuring the fingerprint matching engine's legal compliance with respect to the ESS accuracy requirements. However, eu-LISA should report to the EDPS about the relevant details (e.g. dates, number of fingerprints used, retention period) and the outcome of the measurement.

However he notes, contrary to the use of fingerprints, for facial images, eu-LISA has failed to demonstrate the necessity and proportionality of this processing, as alternatives such as the use of synthetic data were not assessed. Given the sensitive nature of biometric data, additional justification for the use of real facial images should be provided. Therefore the EDPS recommends a study on alternatives of using real facial images in this processing.

4.4.2. Risks

4.4.2.1. Erroneous assessment by experts leading to wrong accuracy measurements (Risks 44 and 59)

Eu-LISA has identified the risk of "erroneous - and potentially- harmful inferences or conclusions on specific individuals through the use of artificial intelligence techniques - in particular data mining - facial recognition or biometric analysis of any kind". Since this risk concerns the accuracy measurement process, the EDPS understands that this risk does not directly affect data subjects, as the result of the processing is only to measure the performance of the matching algorithms in a controlled environment. Indirectly it could lead to erroneous results of the matching and eventually to false positives or false negatives when the sBMS will operate. The EDPS understands this risk as the risk of erroneous assessment of the results by experts, leading to a wrong accuracy measurement.

As a result, the EDPS does not understand how measures aiming to human intervention by the establishment of alternative channels and offer of remedies to data subjects could be applied in the context of this processing. However, the EDPS notes that in the described process, eu-LISA plans to involve more than one experts on biometrics from MS and JRC.

Recommendation 13:
The EDPS recommends eu-LISA to apply the 4-eyes principle to any step including verification from an MS biometric expert.

4.4.2.2. Whitebox inference of training data (Risks 45 and 60)

Eu-LISA has identified a risk of whitebox inference for individuals that were part of the training dataset, from the MS biometric experts that took part in the accuracy measurement procedure. The EDPS understands that the MS biometric experts will sign a binding agreement addressing their exact activities.

Recommendation 14:
The EDPS recommends to reduce this risk to Negligible, provided that:

- access is restricted only to necessary services in the accuracy testing dedicated environment, and
- a confidentiality (non-disclosure) clause is included in the binding agreement of the MS biometric experts

4.4.2.3. Automated decision making for third country nationals (Risks 43 and 58)

Eu-LISA has defined a risk of automated decision making, with possible relevant consequences for individuals. The EDPS understands that this risk cannot materialize in the accuracy testing procedure, as any selected biometric data will not be used in a real operational context and that no direct impact on the data subjects will instantiate.

4.4.2.4. Injection of inadequate VIS biometric data into the test process - Low accuracy in the performance metrics.

Eu-LISA has identified a high risk (Significant) from the Injection of inadequate VIS biometric data into the AMDS and/or the BGDB, leading to malfunction or inappropriate results during the accuracy measurement. The EDPS understands that this risk refers to both quality and representativeness of the extracted data to be used in the accuracy measurement procedure and can lead to erroneous accuracy measurements.

Indeed, there is a risk related to the quality of data selected to represent EES data that do not yet exist. Eu-LISA plans to convert VIS data to the EES format, however since VIS biometric data were not captured by e-gates as it will be the case for the operations of EES, the angle and the lighting of the samples may differ in reality. At the same time, by selecting samples from VIS, there is a risk of excluding biometric data for TCNs that do not require a visa to enter the EU from the accuracy measurement procedure. At the end, the accuracy of the algorithms on VIS data is calculated, however the risk that the algorithm presents a high number of false positives for TCNs that do not require a visa to enter the EU is overlooked.

At the same time, risk stemming from selection of representative data from a random part of the VIS dataset, to be used as input to the accuracy measurement process, could materialize, given the fact that the representativeness of the selected part is inverse analogous to the parts' size.

Recommendation 15:

The EDPS recommends eu-LISA to perform a study on the necessary/optimal size of the random VIS dataset to be used to initiate the process and to introduce it as a requirement in the first step of the process, where representativeness is analyzed.

Also, since measurements related to bias measurement are not described for the accuracy measurement process, the EDPS recommends eu-LISA to include additional separate metrics at least per gender, age and ethnicity as part of the measurement protocols to be defined (by MS and JRC).

The EDPS understands that due to lack of actual EES data, the risk of resulting to erroneous accuracy measurement cannot be further mitigated. The EDPS notes that this risk should be clearly described in order to be taken into account in the sBMS DPIA (as proposed in section 4.3.2).

5. CONCLUSIONS

Eu-LISA has provided all necessary elements foreseen in article 40(3) of the Regulation, for consulting the EDPS, namely the respective responsibilities of the controller, joint controllers and processors involved in the processing, the purposes and means of the intended processing, the measures and safeguards provided to protect the rights an freedoms of data subjects, the contact details of the data protection officer and the data protection impact assessment provided for in article 39 of the Regulation.

Eu-LISA has performed a DPIA on the operation of sBMS, in the context of EES operation that requires biometric matching with VIS and has concluded in some residual high risks, such as risks from low performance of the matching engine and whitebox inference of the training datasets. In addition, for the most prominent risk related to low performance, eu-LISA proposes an accuracy measurement procedure to be applied by using real VIS data. As such a procedure also introduced risks, a separate DPIA was performed.

Eu-LISA has demonstrated the necessity to use real data in their accuracy measurement procedure, as a result of their study on the replaceability of real fingerprints with synthetic data.

The EDPS understands that for facial images such an analysis has not been made. This is not in line with the requirements of article 39 of the Regulation, in which the controller needs to include in the DPIA an assessment of the necessity and proportionality of the processing operations in relation to the purposes. Therefore, the EDPS asks eu-LISA to demonstrate the necessity and proportionality of the use of real facial images for achieving reasonable assurance on the facial images matching algorithms' performance (Recommendation 12).

Furthermore, it is not demonstrated that the facial images from VIS will resemble the facial images the EES e-gates will capture. However, given the lack of existing data and the fact that eu-LISA has to comply with a legal obligation concerning high performance metrics, the EDPS understands that the proposed simulation of EES facial images in the accuracy measurement procedure is the only measure available before the start of operations and that eu-LISA takes a best effort approach in avoiding materialization of low performance risks.

In addition to this, the EDPS has made recommendations to ensure compliance with the Regulation. In particular, the EDPS recommends:

- All processes / systems using the sBMS introduce mitigation measures to ensure alternative means of identification to challenge any false result (e.g. manual identification by a board guard) and minimize the impact to the data subject. (Recommendation 01).
- Eu-LISA and the controllers of the processes/systems using the sBMS Introduce procedures and mechanisms to provide feedback on the matching algorithms' false outputs, so as to trigger the procedures of fixing errors in sBMS (Recommendation 07).
- Eu-LISA includes, assesses and treats the risk of data subjects not being able to exercise their right to human intervention (article 24 of the Regulation), and the risk of not having meaningful human intervention, due to automation bias. For this, eu-LISA must ensure controllers of other systems/processes using the sBMS are adequately informed of these residual risks, so they can include them in their DPIAs and take adequate mitigation measures, such as relevant user manuals for any steps of human intervention (e.g. second line of border checks). In this context, eu-LISA must also provide training material to the controllers, so that end users learn the capacities and limitations of the sBMS and can critically challenge its outcome and provide if possible the confidence rates of the sBMS results to end users.(Recommendation 03)
- Eu-LISA includes, assesses and treats risks from biased algorithms, in the sBMS DPIA (Recommendation 09).
- Eu-LISA includes, assesses and treats risk of concluding to erroneous accuracy measurements prior to the EES operation, including setting milestones to re-measure accuracy in the first period of EES operation (Recommendation 10).
- Eu-LISA includes, assesses and treats risks to synchronization of sBMS and other large scale IT system databases (Recommendation 11).
- Eu-LISA clearly described risks from reliance on low confidence outputs (Recommendation 05).
- Eu-LISA reassesses risk of whitebox inference of training data in the use of sBMS (Recommendation 08).

- Eu-LISA implements the following measures on the accuracy measurement procedure:
  - Apply the 4-eyes principle to any step including verification from an MS biometric expert (Recommendation 13).
  - Include a confidentiality (non-disclosure) clause in the binding agreement of the MS biometric experts (Recommendation 14).
  - Perform a study on the necessary/optimal size of the random VIS dataset to be used to initiate the process and introduce it as a requirement in the first step of the process, where representativeness is analyzed (Recommendation 15).
  - Include additional separate metrics at least per gender, age and ethnicity as part of the measurement protocols to be defined (by MS and JRC) (Recommendation 15).
  - Revisit, clarify and reassess the risk for whitebox inference of training data in the context of the accuracy measurement procedure (Recommendation 14).
  - Describe clearly the residual risk of resulting to erroneous accuracy measurement, in order to be taken into account in assessing the relevant risk to be introduced in the sBMS DPIA (Recommendation 15).

- Eu-LISA reports to the EDPS about the relevant details (e.g. dates, number of fingerprints used, retention period) and the outcome of the accuracy measurements (Recommendation 12).

- Eu-LISA revisit the DPIA, and thoroughly reassess the likelihood and impact of the risks where mitigation measures are proposed (Risks 56, 61, 66, 71, 76, 83). For this, Eu-LISA needs to explain and document how proposed mitigation measures affect the residual risks' likelihood and impact, in order to advocate for their efficacy, even in cases where residual risk remains the same (Recommendation 06).

Last but not least, the EDPS recommends eu-LISA revisit the sBMS DPIA, clearly describe the risks and attribute them to the DPIAs of the relevant systems where the risks could materialize and where the mitigation measures should be taken (e.g. risks related to exercise of data subjects' fundamental rights (risks 38, 53, 58, 63 and 68), secondary uses of sBMS data (52 and 57)) (Recommendations 01, 02, 04)

The EDPS expects that eu-LISA implements the above-mentioned recommendations (including demonstrating necessity and proportionality for the use of real facial images from VIS) and provides documentary evidence of this implementation within three months of the date of this Opinion.

Finally, the EDPS would like to recall that this is an opinion for sBMS in the context of EES and VIS biometric data comparison. Any biometrics from other systems added to sBMS, would imply a different processing and eu-LISA is not entitled to apply similar accuracy measurement procedures without a new DPIA and without assessing alternative solutions.

Done at Brussels on 4 November 2021

po

Leonardo CERVERA NAVAS

[e-signed]

Wojciech Rafał WIEWIÓROWSKI