V. 14.10.2021

DRAFT MINUTES

of the 2nd meetings of the Working Group on Artificial Intelligence

Video conference meeting
14 September 2021

A points	**Agenda points for: information** 1. Documents/presentations shall be provided prior to the meeting 2. Speaking time shall be provided only upon request to MB Secretariat, prior to the meeting
B points	**Agenda points for: opinion / discussion / information** • Including estimated timeslots

B	1.	**Introduction and adoption of the agenda**	Adoption

The Chair of WG on Artificial Intelligence, [redacted] opened the 2nd WG meeting.

- Chair welcomed the Slovenian Presidency as an observer to the meeting.
- AI in JHA is constantly increasing its footprint and it becomes more crucial to find proper use cases. It is very important to look at the scope of the implementation overall and this is the main reason having this WG.
- The agenda was adopted with 2 AOB.

Summary: *MS took note of the adoption of the Agenda with 2 AOB points.*

B	2.	**Adoption of the minutes of the previous meeting**	Adoption

The minutes of the 1st WGAI meeting were adopted. By the deadline, comments were received from COM and SE. All of the proposed amendments were incorporated to the text.

Summary: *MS took note that the minutes of the 1st meeting were adopted*

B	3.	**Shared Biometric Matching Service (eu-LISA)**	Information

On behalf of eu-LISA, [redacted], Senior Business IT Consultant introduced the state of play on AI implementation in biometrics and the sBMS.

Landscape

- Machine learning and deep learning are currently used for the research of several biometric modalities such as: face, 3D face, fingerprints, palmprints, veins, iris, voice and handwritten signature.
- AI is applied at different levels, from algorithm development, quality assessment, presentation attack detection to synthetic database generation.
- Face recognition used before the AI was mainly based on facial landmarks mapping, however, the error rates achieved were not good enough for large scale systems. For facial images, the shift to deep learning has increased accuracy substantially. The increase is more relevant for low quality data, supporting more seamless means of acquisition. Another important aspect is the application of the AI for the assessment of the quality of the facial image.
- Fingerprints have also benefited from the AI technology. But the current AI implementations are not mature enough to substitute standard image analysis techniques. AI techniques are used to enhance the standard image analysis techniques and not to replace them. AI have been successfully used for fingerprint image quality assessment algorithms, being the NIST NFIQ algorithm the most known example.
- Two main approaches are used for biometric algorithm development: machine learning and deep learning.
- In machine learning different types of algorithms are used. Machine learning approach has been used for feature selection and in comparison engines. The features are defined by engineers with the deep knowledge in the domain (face or fingerprint). Machine learning techniques are used to decide which of those features are most useful and which may be introduced in the matching systems. After filtering those features, machine learning can also be used to combine the features in order to decide if the biometric sample belongs to a specific user or not.
- Deep learning approach takes biometrics sample raw data as input and uses it through all the different steps, from feature extractions, feature vector, feature selection to comparison engine. It specially simplifies the features extraction process, letting the deep neural network decide which biometrics features will be used and how they should be combined.

AI in the sBMS

- sBMS provide biometric functionalities to several systems at eu-LISA, among those functionalities are enrolment, verification and identification. For providing these functionalities, the sBMS has deployed different AI algorithms.
    - For enrolment, the AI algorithm takes as input the raw biometric data and creates the user's template. These AI algorithms are normally treated as trade secret.
    - For the verification, it can be done in two different ways: a) there is a possibility to compare 2 different sample in order to decide if they belong to the same person or; b) it is also possible to compare one biometric sample against a user's template to decide if it belongs to that specific user. For those 2 possibilities of verification the comparison engine may or may not be the same.
- Finally, in the case of identification, the biometric sample is compared against the whole data base of user's templates. A list of hits is returned with the user IDs of potential matches.
- All these algorithms are developed, provided and maintained by the sBMS contractor.

- The development of those algorithms, as well as the datasets used for training and testing datasets are considered trade secrets.
- Even though biometric algorithms are generally provided as a black box, they are evaluated at public competitions. Those public competitions allow to compare and rank them across the different vendors and research institutions. They also allow a very open discussion in the field of biometrics.

Quality assessment

- Biometric quality can be divided into three concepts:
  - Physical trait: quality of the physical features of the individual (source);
  - Fidelity: the degree of similarity between a biometric sample and its source (of use for law enforcement, human inspection);
  - Utility: the impact of the individual biometric sample on the overall performance of a biometric system.
- The idea behind this concept is to use the biometric quality value as a predictor of the utility (expected automatic comparison score or human inspection) of a biometric sample.
- Biometric quality is used for both enrolment and verification/identification.
- Assessment of the quality of biometrics uses different approaches, ML and DL, in a very similar way than in the development of biometric algorithms.
  - In machine learning it is used to filter biometric features and define which way to combine them to obtain a single final value.
  - In deep learning, the deep neural network performs all the tasks from the feature extraction to feature selection, to the quality assessment, to provide a final single value.
- Those approaches are currently used for fingerprint (NFIQ). AI approaches are also being researched for facial images quality algorithm.

Datasets

- Training process: Training is a task of the provider and proprietary datasets are used. These datasets are considered trade secrets and its collections require expensive curation. At the moment, the use of synthetic datasets for training (or evaluation) is not recommended.
- Once the algorithm has been developed, it is essential to test it in operational environments. Performance evaluation (throughputs, memory use, cpu loads) requires a lot of data and in those cases, where biometric accuracy is out of the scope, the use of synthetic datasets is a possibility and is a real convenience.
- These datasets are generated by making use of AI techniques (i.e. GAN). It is possible to generate large synthetic datasets without any legal constraints. These databases are able to simulate the conditions the system will face when they are deployed in production. It is a very good tool to test the system for performance evaluations.
- eu-LISA, as a client, needs to evaluate the biometric accuracy of the product. To do so, eu-LISA needs to acquire real datasets. This is a challenge in light of the General Data Protection Regulation (GDPR).
- eu-LISA is looking at the possibility of using real data in order to assess the accuracy before entry into operation (EiO). This task is complex and it is done with the approval of the European Data Protection Supervisor to define how and what data can be used for this purpose.

- The possibility to use synthetic dataset to assess biometric accuracy before the EiO of systems is being researched. It will be assessed at a given time after EiO if the synthetic data evaluation results were in line with the real data evaluations. Depending on the delta, new synthetic data will then be regenerated. At the moment it is not recommended.
- In summary, when algorithms are developed, the contractor uses datasets for both training and testing the solutions. eu-LISA needs to evaluate the solutions and follow the GDPR. To do so, eu-LISA uses its own data sets, and for performance evaluation those datasets could be made of synthetic data. Finally, once the solution is deployed our systems will start filling the production database with real samples and this could be also used for biometric accuracy evaluation.

QA:

- **COM** asked what kind of accuracy eu-LISA is aiming at, when the system goes live. If the system would not use the comparison, the identification under controlled circumstances but so called real time remote biometric identification which would be possible under the legal bases of the Schengen Information System, would these data sets, which are in the system, be suitable. What would be the accuracy and are this kind of use cases foreseen now under the development of the system.
  [redacted] replied that eu-LISA is looking at all of those aspects. For instance, in the Entry Exit System (EES), the accuracy targets are written down in the regulations. eu-LISA needs to make sure that the system deployed meets those targets. Before EES EiO data provided by the contractor is used for evaluations. Once the EES will be running, real data will be used to assess the system accuracy. That will give the real accuracy for the system and this accuracy has to meet the targets defined by the regulation. **Chair** added that the accuracy target is defined in the Implementing Act, not in the regulation.
- **COM** asked if there is a biometric identification in uncontrolled circumstances, is this use case foreseen while deploying the system.
  [redacted] replied that eu-LISA is trying to assess the biometric accuracy under similar capture conditions/scenarios as the one of the systems will be deployed. For example, for the EES, for some scenarios, like airports, the capture conditions are well controlled such as illumination, light, biometric sample presentation. However, there other type of border, like road border controls, where the capture conditions are not as good and not as controlled. eu-LISA will also assess the biometric accuracy of the system for these different type of scenarios.
- **DK** is currently working on deploying a national biometric matching system and wanted to know if it would be possible to run systems in parallel at the central level to be able to assess the central data. [redacted] replied that eu-LISA is setting up something similar. When the EES will EiO, real biometric data will be sent to the central system. Fraction of that data will be used to assess the biometric accuracy of the system. It will be collected and labelled to make sure it is good for evaluation and eu-LISA will run in parallel the evaluation of that data to measure the accuracy of the system. However, it is not certain if it can be done in other systems.
  DK asked what does eu-LISA see as obstacles to run other systems in the same way. DK would like to have the data and try to deploy an algorithm to assess it. There is risk of differences in performances between local matching versus central matching.
  [redacted] replied that the biometric matching takes place at the central system, and

therefore, the biometric accuracy can be measured. The biometric samples quality checks will be done at both national level and central system. Also question from **NL** was answered with this reply.

- **SE** asked how this type of facial recognition technology might be labelled as high risk within the new AI regulation and if eu-LISA foresees any problems using this bought proprietary software where you might need to be able to spot, test, understand and correct biases, where underneath that it can be seen as a black box and you don't get full access to it.
  [redacted] replied that in the context of EES, the use of the biometrics is well justified and there are lot of security and legal checks. For example, when someone wants to enter Schengen area, somehow users are aware that the biometric data is going to be captured for their entry to be checked, manually or automatically. It is a different situation when it comes to the facial recognition at public spaces (surveillance), then it is done without the user's awareness and consent from people passing by. Regarding bias, it was a big topic when it was first discovered. There was bias in the system based on gender, skin colour, age. Since then, a lot of work and improvements have been done. Looking at the last reports on this matter (NIST FRVT), most of those issues have been solved. eu-LISA does not have the insight of the algorithms, the companies have and it is trade secret. However, eu-LISA can run its own evaluation to see if the bias is there. [redacted] added two critical points. Firstly, there is always human decision at the end. In case of system controlling, like EES, third country nationals (TCN), matching system supports the decision of the border guard. Secondly, the proprietor of the matching systems is a valid concern, however, even when not being able to look inside the source code (as they are proprietary solutions), their behaviour can be properly measured and they can be intervened if issues are found. Regulation foresee well defined way how this should be measured in regular basis and it gives enough tools to continually check if the systems are performing correctly. Especially EES regulation, where there is a very strict accuracy bias. Potential bias is aligned by the flexibility of the regulation. Should there be any deviation, there are tools and ways, technical and contractual, to intervene if necessary.
- **Ms** [redacted] proposed to have a presentation, which was given to the Management Board by Mr [redacted] on the topic of bias, also at the WGAI meeting.
- **Chair** confirmed that it is a good idea and it is also related to the question from COM in the chat if the bias could be tested. The gender bias, racial basis (better/worse level of accuracy according to skin colour). It would be useful to know to what extent is the bias resolved and what is tested, in order to be well aware and answer people who are worried about it. [redacted] added that the more transparency can be provided to the citizens, the better. The evaluation if there is bias in the systems is an ongoing process at international evaluations.
- [redacted] added that as eu-LISA has a black box approach, where the algorithms are provided by vendors, it is surrounded by secrecy. However, the fact that most of the algorithms go through public evaluations (e.g. NIST FRVT) gives the confidence about the outcome. The more representative data is used, the better results we have. So why are the systems really working in production, sBMS for example, where simple learning algorithms will not readjust itself based on the actual data. Do we have the control of this process. [redacted] replied that in the EES the

algorithms are provided by the contactor and eu-LISA is required by the EES regulation to check the results in terms of accuracy, bias, etc. If there is a problem with that, the contractor is legally responsible for addressing those issues. It will be a constant evaluation with real data but the algorithm will not learn from that data. [redacted] added that training is a complex process and therefore it cannot happen automatically in production. In order to make sure that bias is limited as much as possible over time it has to perform automatic measurements as described by the biometric Implementing Act of the EES. Something similar will be put in place for the other systems, when their biometric data will be migrated to sBMS. Based on the outcome, the training of the initial sBMS will be done when possible. However, there is one problematic point, the question, to which extent the live data will be accessible. This is a complex process and it will probably require the specific DPIA and the approval of eu-LISA DPO and probably from the EDPS. But technically speaking the training could be performed, if required, but not directly in production. It will be properly tested and the system accuracy will be measured before going live with a new version. [redacted] added that the main point here is that the adjustment has to be done in the controlled way for better monitoring results. eu-LISA and all the actors who need to have the oversight would need to get involved in this training with the vendors.

- **COM** added that for biometric systems, the identification functionality will be a high-risk under the proposed AI regulation. As a result, it will require the conformity assessment and the authorisation from the supervisory authority. It is a grey zone because the responsibilities of the providers and users are mixed. Basically already the vendor, you are buying this product from, will be obliged to go through all the conformity assessment and authorisation procedure before eu-LISA actually buys the product. The vendor will be already obliged to provide all the documentation to fulfil the accuracy requirement before even this product is tested under the concrete context. Once eu-LISA will require this product together with vendor, it has to go through the testing procedure again. eu-LISA will become a provider with its own documentation and gone through the conformity procedure again. The user and the buyer are intertwined in very tight manner. This situation is foreseen in the regulation whether the user will become a provider and go through the entire authorisation process from the beginning. It would then double the competences of the authorities. This point needs to be clarified during the legislation process and COM needs to make a very strong point.
- **FR** asked if deep learning is used for fingerprint matching. [redacted] replied that it is used not yet for finger prints matching alone (it is used combined with standard techniques), but it is used for face recognition algorithms. For fingerprint, the results from standard image analysis techniques were really good. What has been done with matching learning and fingerprint is to enhance the results of standard techniques.

Summary: *eu-LISA presented the state of paly on AI implementation in biometrics and how it is used in the context of sBMS. Additionally, the quality and datasets were also explained.*

	4.	**Presentation of the results of the questionnaire on national Initiatives**	Information

On behalf of eu-LISA, Vice Chair [redacted] presented the summary of the results of the questionnaire on national AI initiatives.

- eu-LISA received responses from 20 MS. Majority of respondents are from the Ministry of Interior and Police/Law Enforcement, however, there are some responses also from Immigration authorities and one Ministry of Foreign Affairs.
- 79% of the respondents are aware of some national level strategies and road maps on AI.
- Frome the literature of the subject and in the context of the couple of studies eu-LISA has done, it can be said that the majority of the MS, either have in place broad national level strategies for AI or are working on those.

National strategies for AI, use-cases, approaches

- Some of the MS indicated that there are specific strategies for AI in the area of Justice and Home Affairs (JHA). While majority indicated that there are broad strategies for AI but no specific strategies or roadmaps for AI in the area of JHA.
- From the questionnaire it can also be seen that 79% of the respondents are aware of specific AI initiatives and projects.
- Number of respondent indicated that there are various solutions in place or are being developed on national level of processing, automated translation tools, named entity recognition and chatbot tools.
- Other prominent use cases are biometric recognition, licence plate recognition and video analyses.
- Next prominently featured topic for AI is big data analytics, like passenger name (PNR) data, passenger profiling and money laundering investigations. eu-LISA is also working at the use case of passenger profiling in the context of ETIAS and application of the AI with in CRRS.
- Following group of use cases is the analyses of unstructured data and improvement of internal processes, e.g. focusing on including allocation of resources/personnel in disaster management, etc.
- Some MS deploy different kind of experimental development (PoCs, pilots, prototypes). Some MS rely on in-house development, some outsourced development of AI solutions.

Use of open source and the possibility of sharing.

- BE, CZ, DK, HU, IT, NL, NO, SE and SI reported the use of open source technologies to some extent. The use of open source ranges from the open source programming languages, open source libraries, GAN approaches for synthesising age progression and tools for text and media analysis.
- Some MS have indicated the reliance on proprietary software. In the majority of cases you can make a link between in-house developments and the use of open source vs. contacting out and the reliance on proprietary technologies.
- Very few MS are open to sharing code. Limitations mentioned were lack of a general code of ethics for AI/ML endorsed by all parties, lack of common legal framework that would allow sharing of the solutions, contractual limitations related to the proprietary software and data protection issues.
- Some MS are open to sharing code, however, on a case by case basis, upon bilateral agreement and pending the approval of the legal services at national level.

Legal barriers at national level to deploying AI, using real data sets for AI training and exchange of data between organisations & MS

- 58% of the MS indicated that there are barriers to implement AI in JHA and 42% indicated no barriers.
- Regarding the real datasets of the AI training 53% see that there are barriers vs. 47% see no restrictions.
- When it comes to the barriers to exchange of data between organisations/MS most of the MS, 73%, indicated that there are barriers, particularly at law enforcement and police, where sharing the data is prohibited by law.

Areas where deployment of AI is more/less feasible

- More feasible for MS are the areas under human supervision, analysis of internal/ seized data, systems that don't use personal data, biometric recognition in investigations and AI for data analytics for investigations.
- Less feasible for MS are real-time biometric recognition and automated decision-making which are not supervised by human being, use of surveillance/intelligence gathering and where explaining the outcome is difficult.

Legal restrictions on the use of real data and how can those be addressed

- There is a lack of a clearly defined legal framework for using real data sets and general data protection regulation (GDPR) and similar legal acts focused on data protection.
- To address the issue, it is necessary to provide a clear legal basis for specific use cases and provide data anonymization.

Use-cases for the implementation of AI within the scope for large-scale IT systems:

- AI in EES/ETIAS, for example passenger profiling which eu-LISA is already exploring;
- Different kinds of predicting peak loads at Border Crossing Points (BCPs);
- Tools for victim identification using various media (voice/image/video);
- Big data analysis tools e.g. in the context of SIS alerts;
- eu-LISA as a SaaS and HaaS service provider with the focus on big data and IA. Discussions are already ongoing on EU dataspace for the security innovation;
- Data quality assessment (e.g. biometric data);
- Fraud detection in visa submission applications;
- Use of AI for eu-LISA internal processes.

Limitations on the use of AI in large-scale IT systems include:

- Restrictions on automated decision-making;
- Restrictions related to the training of AI and the use of real data sets;
- Frequent upgrades and changes in the systems;
- Protection of fundamental rights, privacy;
- Heterogeneous regulations what need to be aligned;
- Data fragmentation and data quality, issues with sharing data across MS;
- Supervision of AI systems and ensuring explainability;
- Large number and complexity of ongoing projects.
- MS were asked to confirm if eu-LISA could share the summary of more detailed responses of the questionnaire in the form of word document.
- **NL** commented that use case regarding the supervision under humans was indicated as more feasible. However, in NL such supervised AI system have caused lot of problems. Humans are easily fooled by AI. Human supervision is often just a formal system and not proper supervision. **Chair** added that this might be one of the points eu-LISA would discuss at the future meetings. [redacted] confirmed that this could be a topic to have a dedicated session.

- **DK** added that in the previous presentation there was a distinction between the machine trained systems and more adaptable systems, learn on the go. Is there any indications in the questionnaire that any MS is using adaptive systems. [redacted] replied that there was no mention of these kind of specific systems. **DK** added that the example given by DK is actually an adaptive system what is supposed to be learning on the go, and that brings up the use case about monitoring what is not mentioned here. The idea of monitoring the development of the adaptive model is something what is very important when having adaptive systems.

Summary: *eu-LISA presented the summary of the results of the questionnaire on national AI initiatives. MS were asked to confirm if eu-LISA could share the summary of more detailed responses of the questionnaire in the form of word document.*

B	5.	**Presentations from the Member States and Agencies** - Presentation from Norway on AI initiatives at national level; - Presentation from CEPOL	Information

On behalf of the Norwegian Directorate of Immigration, Head of Statistics and Analytics Ms Musk, gave a presentation about automation and Artificial Intelligence

- Norway handles about 90.000 cases a year at the immigration authorities with approximately 1000 employees. It also runs asylum reception centres and is responsible for IT-systems for the immigration administration in Norway.
- The immigration administration in Norway has a strategic priority to reduce the waiting time. At the moment the waiting times are long, e.g. citizenship 251 days, permanent residence 335 days, protection 171 days and family reunification 315 days.
- However, to go from complex processes to streamlined digital processes is not straight forward road. NO is looking at the areas where they can automate the process and have started with citizenship applications. The acceptance rate is high for people applying for citizenship and it is a good place to start implementing automation process, given that a lot of information on the applicants is already available to immigration authorities, making processing of their applications more straight-forward.
- The route to the use of AI in the Norwegian Directorate of Immigration (UDI) is divided into four areas.
  1. RPA – Robotics: It is very easy to apply with low cost.
  2. Rule based automation: It is used to automate the citizenship processes. It is costly IT development, however, with significant effects.
  3. Standard software with embedded AI: Here the usage is widespread. An example is Splunk, where an application running is checking that the employees do not do things that they are not supposed to do, e.g. checking at the applications they are not supposed to check. Many people have also experimented with chatbot and the process mining is used as well.
  4. Business Intelligence and data-warehouse solutions: NO has a big data-warehouse with lot of data and an extensive business intelligence platform. Using big data analytics in AI is a very important part of the journey. NO is using their own data, building skills and building platform that can then be developed to more advanced later. It also interacts with other areas e.g. with robotics application and it interact with business intelligence platform. If it is decided an area of old cases can be

automatically closed, it will be identified and closed. Having a very extensive business intelligence platform with lot of operational reports means that the employees get hands on experience with their own data. It also helps to notice when the issues with the data quality appear.

- The biggest effect has been through the automation of the citizen application, which is all rule based. In July 30% of the applications were processed automatically. It has reduced process time, has been cost effective, has improved user experience and its scalable.
- NO had a change of regulation in 2020 to open dual nationalities and overnight the volume of the applications doubled. Automation of the citizen application made handling the applications much easier.
- At the Immigration Authorities most managers have the background in law, not in analytics or sciences, however it is changing. It is important to change the understanding of the organisation what to do with the data and technology, it also changes the top managements understanding of both what technologies can do and also what needs to be done to make it work.
- Data quality is an issue what needs to be resolved on top of the regulations with privacy.
- IT-architecture is complex and difficult issue. Once more advanced analytical solutions will be embedded into the running processes the IT update is very important.
- Regulation around migration is not developed so that it is easy to be digitalised, this field is very political. However, the dialog with the law department is ongoing to simplify the process and build up an understanding that digitalisation is needed when creating laws.
- **DK** representative has background in health, where there is very strict separation of daily use for primary concerns like case handling and secondary which is research, statistics, reporting. It seems like that you are actually peering back into this single case, is that correct. NO answered that they are not applying business analytics on individual cases. The automation of the citizenship is purely rule based process. DK added that, the presentation mentioned automatic closing individual cases. **NO** replied that they have a data warehouse which is a copy of production data bases. But no other environment where they are anonymised. **DK** continued that they have multiple case handling systems, and there is a risk that the business analytics environment and the case handling environment disagree because of multiple interpretations of raw data. When to start thinking of primary, secondary use of data DK also has some analytics use cases what need to be built into the systems which have different tasks. **NO** added that it is always a challenge when you transform data to understand it. We transform data from raw data and merge, however NO works very closely with case handlers and the production line in creating those applications. The data in the case handling system will not be designed well for most analytical purposes. The data needs to be moved to the environment where it is transformed and enriched, which is a complex process. That is where the big challenge for public sector accrues. To have that deep understanding you have to work with that long time and be close to operations. Consultants and the external IT companies will never be able to achieve that deep understanding. **DK** added an example from previous workplace where the analyses was done large amount of data from a data analytic platform which was there to get to know their user. In the scope of one-year data analytic platform turned out to be so preformat that it gained production data. Instead of having split environments with different data bases and purposes they unified the data

search, to have data in only one place, which is used form different places. So the separated environments were not needed

- **COM** asked if NO used cahtbot and if it had any impact or not since NO had very long waiting times. **NO** answered that experiments with chatbot have been done but they have not yet gone live. Using chatbot is challenging because it needs lot of training data and the great variety questions are asked. Very often the questions are specific to the case and not a general question. Chatbots work well with general questions. Lot of the users are also not native speakers and the base of users is not as big as for banks for example, which is required to develop a good chatbot.
- **SE** commented that Swedish Migration Agency has the same approach to chatbots as UDI.

On behalf of CEPOL, Senior Training Officer [redacted] gave a presentation on the current work on AI capability building in law enforcement

Training needs assessment

- Law enforcement is asked about their training needs and CEPOL conducts an annual training needs assessment, focusing on specific prioritised topics formulated during EU strategic training needs assessment. It is implemented by CEPOL every 4 years.
- 2018 CEPOL published EU strategic training needs assessment. This assessment emphasised the importance of the topics in the context of law enforcement work.
- In 2020 training need analyses was conducted on COVID-19 pandemic, to assess the challenges and impact what this situation has had on different types of crime and on the law enforcement.
- The aim of the EU strategic training assessment is to identify strategic EU level training needs for law enforcement officials, for four years' period, currently 2022-2025. This is a multi-step process, coordinated by CEPOL and multiple stakeholders.
- The first step, is desk research, to extract all capability challenges from current policy documents and group them into the thematic horizontal groups.
- The next step is to organise focus group discussions with experts for each thematic horizontal area. The focus group decides which capacity challenge can be addressed by EU level training and this way the list of EU level training need is compiled by the priority order by MS.
- By reviewing and approving the policy documents, the capability challenges are extracted from the documents. As a result, three main training areas related directly to the domain of AI regarding the horizontal issues were established:
    1. In the context of the key enabled technologies/new technologies there is a need for training to adapt to new technology by building capacity in digital investigations and learn to exploit AI for law enforcement purposes;
    2. In the context of the digital skills there is also a need for exploiting AI for law enforcement purposes;
    3. In the context of fundamental rights there is a need for ethical guidelines for use of AI and also the use of AI by law enforcement.

AI related training activities in 2021

- In 2021 Interpol in cooperation with CEPOL has already organised, three steps virtual training sessions about the introduction of AI to law enforcement.

- Interpol underlined that AI has a potential to be a permanent part of criminal justice ecosystem, to provide investigative assistance to professionals for better maintenance of public safety.
- The virtual training sessions were organised to provide an overview what is the impact of AI of law enforcement and how it can influence police models and activities, used as a tool and source of evidence.
- The main objective of the training was to introduce AI concepts and explain the relevance in the context of policy.
- The first session included introduction to AI and machine learning and the opportunity to ask questions. Definitions, history and fundamental concepts were covered together with modern samples of AI systems. Putting AI into practise was also covered, including what can and cannot be done currently with AI. Last part of the first session was dedicated to AI in law enforcement covering the topics related to ethics, transparency and use cases.
- The second session covered the information, which part of law enforcement AI could be useful and a brainstorming session about the possible use of AI in policy.
- The third session included how to start thinking about AI from the ethical point of view and why AI rises the concerns regarding ethics.
- The training was recorded and is available on CEPOL lead platform.

CEPOL KNOWLEDGE CENTRE (CKC) on law enforcement cooperation, information exchange and interoperability

- CEPOL knowledge centre on law enforcement cooperation, information exchange and interoperability is cooperating with the MS, JHA agencies and Commission.
- The main aim with the CKC is to bring together professionals who are experts on the given thematic area, in this case law enforcement cooperation, information exchange and interoperability. To elaborate the holistic training portfolio as well as follow up activities.
- The new CKS, established in the beginning of 2021, is composed of 8 representatives from MS, as well as representative from COM, Frontex, Europol, FRA and eu-LISA.

CKC INT Portfolio 2022

- In cooperation, CEPOL portfolio for 2022 with multiple topics was prepared.
- Topics directly related to AI were introduced (see slide no 54 for more info).
- CEPOL also organises ad-hoc webinars when training on any other AI related topic is necessary. MS were encouraged to keep in touch with CEPOL if any other training is required regarding AI.

Participants of training activities

- Year by year participation of training activities was presented (see side no 55 for more info)
- **Chair** stressed the importance of AI capacity building and the use of the available training tools to be able to benefit from the cooperation.

Summary: *WGAI members were given a presentation about the automation and Artificial Intelligence at the Norwegian Directorate of Immigration and about CEPOL's training options in the field of AI for 2021 and 2022.*

B	6.	**Presentation from the [redacted] project**	Information
[redacted]			

Past (Analyses systems for gathered raw data)

- In 2013 the first conversation on a possible "[redacted] Law Enforcement Agencies (LEAs)" [redacted] started.
- 2014 & 2015 The Hague Workshops and Consultation with Europol took place.
- In August 2015 [redacted] project as a proposal was submitted.
- In September 2016 [redacted] began as a project.
- In June 2017 the **Month 9 Milestone and Hackathons** with LEAs was achieved.
- During 2018 the project began to become more public.
- During 2019 the sustainability goal of the project had to be addressed and [redacted] was created.
- In 2020 maximising the uptake of [redacted] and preparing [redacted] and connecting projects took place.
- In July 2021 [redacted] begins formally as a project.
- In October 2021 [redacted] will begin.
- [redacted] project includes currently 12 LEAs, 15 Research and Technological Organisations (RTO) and 6 Small and Medium Enterprises/ Industrial Partners.
- Main objectives:
  - To improve the efficiency of LEA, Industry and Researchers collaboration in Security Research projects;
  - To deliver an active and sustainable community of security research practitioners (LEA, industry, research);
  - To help LEAs improve their technological autonomy;
  - With fluid, frequent, and fruitful collaboration between all stakeholders focused on short full-development cycles and face-to face "Hackathons" every 6 months;
  - What Began as a Big Data Analytics project;
  - Forensics = Hindsight;
  - Intelligence = Insight;
  - Horizon Scanning = Foresight;
  - With an emergence of a new wave in Artificial Intelligence Advanced the project gained more relevance in AI, Machine Learning, Need for Training Data;
  - "Big Intelligence and Artificial Data";
  - Complementary to other existing and future initiatives.
- An Overview of the initial plan of [redacted] methodology was introduced together with the autonomy of the [redacted] which is open for closed community. (For more information, see slide no 6 and 7)
- 65 different tools were developed with many new tools, but also many wrappers or improvements to pre-existing ones. All of them are free and well documented. (For more information, see side no 8)
- Lessons Learned:
  - The Human Element: capacity of people spending time, churn, opportunity to staff will get you every time if not considered;
  - Accept Risk, build it in to the project;
  - Multi-disciplinarity in practice, not theory, has huge benefits but needs time and energy to pay off;
  - Identify what is a reasonable expectation of your partners early and revise it continuously;

- Rhythm: Keep things continuous, even if your partners run at different paces;
- Adapt, learn, evolve, change, repeat;
- Fast failure, if it isn't working, stop it sooner rather than later;
- Don't confuse something that is a tough challenge as something that is a failure either;
- Technical challenges in Security Research are not unique. Learn from other domains;
- Circles of Trust can be extended. Leads to more cooperation and innovation at the margin of a project;
- Rigidness in the rules can be actually used to create agility in the project by guaranteeing minimums.

Present [redacted]

- [redacted] is open for membership for industrial partners, research organisations and LEAs. At the moment 50% of the members are LEAs.
- [redacted] has recently received 2 million of funding and that would help to fund some of that work, for example when digital Europe programme comes alive and to have some specific funding for LEAs and data spaces.
- A *sui generis* license has been created, inspired by open-source principles but adapted to the special characteristics of its intended end-users. No license costs for LEAs.
- A vast majority of [redacted] results have been transferred (licensed) to [redacted] so that it can foster their sustainability and, when needed, evolution towards higher Technology Readiness Levels (TRLs).
- [redacted] will primarily focus on development of technological solutions for EU LEAs and on being the "custodian" of a repository of technological solutions.
- The association conducts software development to deliver tools free of license costs to help EU public security practitioners in their fight against cybercrime and it manages a repository of tools that are "under development" but at high maturity level (both its own and of third parties).
- [redacted] is a piece of a larger puzzle, and it hence will collaborate with multiple other relevant actors: Europol, ECTEG, ENLETS, third party owners of tools of interest to EU public security practitioners.

Future [redacted]

- [redacted]: secure and resilient AI technologies, tools and solutions in support of law enforcement and citizen protection, cybersecurity operations and prevention and protection against adversarial AI. It has 53 partners and a stakeholder advisory group. (For more information, see slide no 17.)
- Strategic goals of the project include the understanding of AI based technologies, the protection of public using the AI technologies, creation of the datasets for AI, exploitation of the results and also combat the misuse of the AI. (For more information, see slide no 18.)
- An overview of the [redacted] Project Structure was given. (For more information, see slide no 19.)
- Expertise and resources are needed for the capable community to make use of the AI technology in the future. It is also important to look into how the AI technologies can be transformational.
- The methodology of the project is similar to the [redacted] project. It has six full development cycles of six months each. It has three releases:

- The "first look" includes the finalisation of the AI platform, early versions with empty boxes or mock-ups and training and testing.
- Alpha includes fine tuning of the collaboration processes, delivery of tools and data sets and AI community building.
- Beta includes scalability and maturity of tools and solution and the finalisation of the AI community.

- [redacted]
- NL asked for the explanation of the different abbreviations used during the presentation. A slide with the abbreviations was provided after the meeting (for more information, see slide no 24).
- **DE** is working on the AI act and wanted to know if there are any challenges dealing with the AI regulations by the European Commission regarding mentioned projects. In some fields, like facial recognition, MS are confronted with very high regulations and it is really hard to go forward with AI. [redacted] replied that [redacted] project will be self-contained with own work packages and experts. There is no simple answer to it but there are few important things. There needs to be a common understanding what AI is and what it is not. It is no panacea. There needs to be society outreach, where the conditions are communicated to pubic and explained what is done and what the AI technology is used for. These values had to be overcome before the AI regulation as well, however, these projects are research projects so there is certain difference what is allowed. It all comes down to have an honest debate with the experts. The interpretation of the regulation and what is and what is not allowed does not mean it should not be explored scientifically. It does not mean that technologically the experiments should not be done, it is important to make sure it is compliant and then have a conversation if morally it should be done. Our experts created a set of recommendations which were included as an annex for pubic results. Agency who takes the results to deployment needs to ensure that the technology used is for certain ethical and legal purposes or within the frame of what was agreed in the project. When it is finalised we go step by step and work within the limits. Different MS have different cultures, different society and perception what should be done with AI, there is a different level of sensitivity or flexibility form the perception perspective. If to have proper technology what can be used, what is the trade off and what is the legal limit what can be do with it.

Summary: WGAI took note of the presentation on [redacted] and [redacted] projects.

B	7.	**AOB**	Information

3h10.50 [redacted] (DG HOME) gave an update on the state of negotiations on AI legislation and EU Security Dataspace for Innovation.

- The AI act proposal was published on 21 April 2021 and the discussions have started in the Council, in the competent council working party, which holds only physical meetings, where MS and COM participation is limited. There is no online version of that working party and that's why it is especially important that all of the opinions and the national positions are generalised to delegate to the person participating, who is not the law enforcement or home affairs experts. The Presidency organised already two workshops

which will be also followed by two other workshops on different topics for law enforcement community. There was a Presidency workshop about the high-risk systems consequences in the system and on 30 September there will be a full- day workshop on the impact of the AI Proposal on law enforcement. COM received six pages of questions from the MS which COM will give answers during this full day workshop. There are more questions than replies in the discussion. MS are trying to understand the consequences of the provisions and most probably the regulation will not be adopted in the near future. However, AI related developments should not be frozen because of the outcome of this legislative process since there is already strong framework applicable to law enforcement. For data protection there are number of warranties coming from procedural law which are already framing technological development. It seems there will not be prohibited practices for the law enforcement at the moment. If possible, the aim is to generate this procedure into the existing ones, so it does not change so much on the substance. Substantial discussions in the European Parliament have not started yet as the rapporteur has not yet been confirmed.

- Dataspace initiative is stepping into the second phase, the development of the national infrastructure. The funding would be provided by the Digital Europe programme, the first work programme which should be adopted in the coming weeks. It contains funding for national components of the security data space. On the 7 July a workshop was organised where the draft of substantial requirement of this call. MS were asked to give feedback and so far feedback was received only form EE. COM also wants to launch a study, to study in depth the use cases for security dataspace and all of the technical components including the interaction with the central components and the central services.

On behalf of [redacted] (DG HOME), [redacted] gave an update on the status of the chatbot project.

- The project was formally kicked off on 29 July. Three MS (CH, NL, SI) participate in the project with the visa authorities and FR with a “ministerial delegation” promoting AI in in the administration. More MS are still welcome to join the project. If interested were asked to get in touch with eu-LISA or [redacted].

Summary: *WGAI took note of 2 AOB points.*

B	**8.**	**Meeting recap and next steps**	Information

Series of very interesting presentations took place. Covering general aspects of how sBMS works to quality evaluation measures and then a quick summary of the responses to questionnaire provided by MS. MS were kindly asked if they object sharing the detailed results of the questionnaire with the Working Group. Without any objections the document will be paced in WGAI SharePoint space. Presentation from Norway was given on the use of automation and AI by the immigration authority and presentation form CEPOL on their training initiatives. Presentation on [redacted] project and [redacted] project was also introduced.
Next meeting will take palace as a web meeting on 23 November.