Third Technical Workshop in the context of the revision of the Prüm decisions

Part II: Including the exchange of (pseudonymised) biographical data in the Prüm framework

Federal Criminal Police Office Germany

20th April 2021
Contents

- Preliminary remarks
- The scope of EPRIS and ADEP
- The concept of EPRIS (current and future state of processes)
- Analysis
- Conclusions

Annex
- The EU project (general information, objectives and achievements, way forward)
- Where to find additional information
This is **not a status report** about the EU-funded project EPRIS-ADEP as one action of working party IXIM, but

→ the solutions of the project will be reflected.

This is **not a concrete** functional, technical or legal **proposal**, but

→ the findings of the feasibility study of 2012 as well as the current considerations of the Commission will be reflected.

Over the years **EPRIS was occasionally sceptically perceived**, but

→ please envision law enforcement in 5 to 10 years time and ask yourself what should be standard by then – designing this future has started!
EPRIS – European Police Records Index System

Is a system, currently in the second pilot phase, providing the law enforcement end user a smooth and efficient workflow in order to locate and retrieve certain biographical data contained in national police records EU-wide. EPRIS follows the ADEP approach.

ADEP – Automation of Data Exchange Processes

Is a general approach following the principle of availability, while automation is understood as a management concept by which a process or procedure is performed without or with minimal human assistance. It can be applied to several use cases within the law enforcement community where data and information have to be exchanged.
Nowadays Requesting information

Do you have any law enforcement information about John Doe?

**Germany**: about 50 to 70% of initial requests to another EU MS aiming at gaining knowledge if data is available contain data about a person

**Europol**: about 80% of initial SIENA requests are such general requests
Nowadays Receiving information

Do you have any law enforcement information about John Doe?

France: about 70% of such general requests result in a: ‘No information available’

Inefficient! Spreading suspicion against a person! Risk of disclosing confidential investigative information!
Do you have any law enforcement information about John Doe?

Searching for a match with xh98h408njghg

Future

Step 1: Locating information - Search

Search for xh98h408njghg

Search for xh98h408njghg

Search for xh98h408njghg

Search for xh98h408njghg
Step 1: Locating information – Result

No (clear) personal data disclosed!
Addressing all potential data owners in one go!
Step 2: Requesting information

According to an EPRIS search you are supposed to have information about John Doe...

According to our records John Doe has...

Well targeted request with high probability to receiving meaningful information!

Prerequisites for more process automation given!
Matching approach with pseudonymised data

1. Query
   - Surname: Doe
   - First Name: John
   - Date of Birth: 31/12/1987
   - Gender: Male

2. Pseudonymisation
   - Surname: xh98h408njghg
   - First Name: 5L1*$h&0fg=?($^2;
   - Date of Birth: 31/12/1987
   - Gender: Male

3. Search

4. Matching Surname
   - Strict hit: xh98h408njghg
   - Tolerant hit: E%483h408njghg

EPRIS Index

Surname | First Name
--------|------------
.xh98h408njghg | ...
.E%483h408njghg | ...
?(§^2;Z%29 | ...
.§^2;Z%29Hb9 | ...
... | ...

This project was funded by the European Union’s Internal Security Fund - Police
Key questions

1. What is a police record?
2. What are the search criteria?
3. Who should be using EPRIS?
4. What is the system architecture?
5. Why to pseudonymise data?
6. How to cope with false positives?
7. How to cope with expected high number of requests?
8. Shall additional information be provided with a hit?
9. Is a new system really necessary?
1. What is a police record?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A common definition may be needed for functional and legal reasons.</td>
<td>• All data source(s), currently (conventionally) queried in order to</td>
<td>• Consider current state of play of conventional requests.</td>
</tr>
<tr>
<td>• A broad definition with limitations specified in system specific legislation was recommended.</td>
<td>answer a standard request of a MS – ‘Please provide any police information available regarding this person.’ – should be considered as generally applicable for the creation of the respective national index.</td>
<td>• Principle of sovereignty vs. Prescribing EU-wide standard.</td>
</tr>
<tr>
<td>• Definition proposed: A ‘Police Record’ shall mean any information available in the national register or registers recording data of competent authorities, for the prevention, detection, investigation and prosecution of criminal offences.</td>
<td>• Aim: national indices include all data sources currently used to answer a standard request.</td>
<td>• Consider future design of EU architecture and its impact on (national) legacy processes.</td>
</tr>
</tbody>
</table>
| • The name ‘European Criminal Investigation Records Index System (ECIRIS)’ was suggested. | • Decision of each MS in compliance with national law (including question about classified data). | • Consider FD 2006/960/JHA (Art. 2, Definitions and Art. 3 [3]: MS shall ensure that conditions not stricter than those applicable at national level for providing and requesting information [...] are applied for providing information [...] to competent LEA of other MS [...] )
|                                                                                  | • Transparency: documentation/publishing at a central point of data sources used to create individual national index. | • Consider Benelux solution (Article 14 of Police Treaty).                                       |
Article 14 – Match querying of police data banks

1. The Contracting Parties may give their respective competent services, insofar as their national law does not expressly forbid it, the possibility of direct automated querying of the police data banks referred to in Article 5(2) of this Treaty, for the purposes set out in Article 2 of this Treaty.

2. The querying referred to in paragraph 1 shall be limited to determining whether data relating to a specific person or a specific object is contained in those data banks.

3. The querying may be carried out only in individual cases and in accordance with the national law of the requesting Contracting Party.

4. If the querying referred to in paragraph 1 reveals that data relating to a specific person or a specific object are contained in the file, the content of the data may be obtained under Article 4 of this Treaty.

5. If the Contracting Parties make use of the possibility referred to in paragraph 1, it shall be the subject of an implementing agreement between those Contracting Parties. The implementing agreement shall lay down the necessary conditions, provisions, and operational and technical arrangements for such querying. The implementing agreement shall determine which data banks referred to in Annex 6 to this Treaty may be queried. The competent services may not carry out the querying referred to in this article until the implementing agreement has been concluded and has entered into force.

Annex 5: Data banks from which information and personal data may be transferred

From the following data banks, the competent services may transfer information and personal data to the competent services of another Contracting Party on the basis of Article 4 of this Treaty.

For the Kingdom of Belgium:

The data banks directly available or directly accessible to the Belgian police services.

‘Directly available’ means the data banks already available to the police services.

‘Directly accessible’ means the data banks available to other public or private authorities, services or persons and to which the Belgian police services have access under the law.

For the Grand Duchy of Luxembourg:

The data banks of the Grand Ducal Police in accordance with national law.

For the Kingdom of the Netherlands:

- The system called 'Basisvoorziening Handhaving' [basic facility - enforcement] (BVH);
- The system called 'Basisvoorziening Informatie' [basic facility - information] (BVI).
2. What are the search criteria?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• First name</td>
<td>• First name</td>
<td>• First name</td>
</tr>
<tr>
<td>• Surname</td>
<td>• Surname</td>
<td>• Surname</td>
</tr>
<tr>
<td>• Date of birth</td>
<td>• Date of birth</td>
<td>• Date of birth</td>
</tr>
<tr>
<td>• Gender</td>
<td>• Gender</td>
<td>• Gender</td>
</tr>
<tr>
<td>• Alias</td>
<td>• Any other names/aliases</td>
<td>• Any other names/aliases</td>
</tr>
<tr>
<td>• Residence or known adress</td>
<td>• ./.</td>
<td>• ./.</td>
</tr>
<tr>
<td>• Nationality</td>
<td>• ./</td>
<td>• ./</td>
</tr>
<tr>
<td>• Place of birth</td>
<td>• Place of birth (currently not in use)</td>
<td>• Place of birth (geo-identification number)</td>
</tr>
</tbody>
</table>
<pre><code>                                                                               | • Minimum search attributes for a successful query are defined. |
                                                                               | • The pilot system has a strict and a tolerant search mode. Additionally, it offers a fuzzy name comparison to match not only exactly identical names, but also sufficiently similar ones. |
                                                                               | • The devil is in the details: consider implementing/delegated act |
</code></pre>
3. Who should be using EPRIS?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Police authorities should have direct access to the system from the onset.</td>
<td>• Individual decision of participating MS which LEA get access to EPRIS and what conditions apply.</td>
<td>• Consider different organisational structures in MS, thus several implementation models for the access to EPRIS should be facilitated.</td>
</tr>
<tr>
<td>• Other competent authorities working with a criminal justice finality should have indirect access to the system (via the criminal investigation liaison authority) - they may obtain direct access at a later stage.</td>
<td>• Prior to creating a query the user has to be satisfied that all preconditions for the bilateral information exchange are fulfilled (e.g. principle of proportionality, respective legal requirements) thus, as a rule of thumb all LEA staff currently allowed to initiate or execute international LEA correspondence are considered as future EPRIS users.</td>
<td>• Consider different use cases and conditions applying.</td>
</tr>
<tr>
<td>• Europol should have direct access to the system, other EU agencies may apply.</td>
<td>• No measure concerning the alleged person(s) of interest must be taken based only on an EPRIS hit before it has been confirmed by the data owner – thus, users with a need of prompt action are not to be included or with certain precautions taken.</td>
<td>• Broad access vs. Keeping confidentiality of data (see pseudonymisation).</td>
</tr>
</tbody>
</table>

This project was funded by the European Union’s Internal Security Fund - Police
### 4. What is the system architecture?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Hub-and-Spoke-System.</td>
<td>• Hub-and-Spoke-System.</td>
<td>• Hub-and-Spoke-Model in discussion.</td>
</tr>
<tr>
<td>• Decentralised data storage at national level.</td>
<td>• Decentralised data storage at national level.</td>
<td>• Central component would be more than a mere router – different central services need to be implemented.</td>
</tr>
<tr>
<td>• Central component: Central Forwarding System.</td>
<td>• Central component: Central Services (proxy and essential central services).</td>
<td></td>
</tr>
<tr>
<td>• Re-usable (uniform) solution at each MS for essential services (creating index, querying, etc.).</td>
<td>• Re-usable (uniform) solution at each MS for essential services (creating index, querying, etc.).</td>
<td></td>
</tr>
<tr>
<td>• Data exchanged in EU format.</td>
<td>• EU standards used as far as applicable.</td>
<td></td>
</tr>
</tbody>
</table>
5. Why to pseudonymise data?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Focusing on usual data protection measures.</td>
<td>• Privacy by design approach as fundamental principle of the technical solution piloted.</td>
<td>• Privacy by the design is one of the key principles of the EU data protection regime.</td>
</tr>
<tr>
<td>• Privacy by design not a specific subject (European Data Protection Regulation and Directive as of 2016).</td>
<td>• Strengthens the rights of data protection subject.</td>
<td>• Consider position of the EDPS and the development of a Code of conduct for pseudonymisation.</td>
</tr>
<tr>
<td>• Need to protect confidentiality of LEA data mentioned.</td>
<td>• Supports all parties involved to keeping confidentiality of respective data.</td>
<td>• Consider impact on data owning parties regarding facilitation of keeping confidentiality of LEA data.</td>
</tr>
<tr>
<td></td>
<td>• Pseudonymisation transforms directly identifying personal data to the level of biometric data (FP or DNA pattern), which still is personal data but not directly identifying = leveling identifying biographic data used in EPRIS to the state of the current PRÜM approach.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Solution is economical and meets user requirements (incl. different search features).</td>
<td></td>
</tr>
</tbody>
</table>
### 6. How to cope with false positives?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• False positives only mentioned as a risk.</td>
<td>• A hit indicates a <strong>strong probability</strong> that data is available.</td>
<td>• Consider impact analysis; challenge: valid and representative statistics hardly available.</td>
</tr>
<tr>
<td></td>
<td>• Rule: users have to enter all available data to increase the quality of the matching.</td>
<td>• Interpretation of the data vary and can be biased because of current organisational and resource frame of LEAs affected.</td>
</tr>
<tr>
<td></td>
<td>• An indicator of the quality of the hits is being tested; too many hits (meaning: no useful results to be expected) are indicated; consequence: conventional request if justified.</td>
<td>• Effects on the positive side are regularly not weighed adequately, namely the requests not sent (any more).</td>
</tr>
<tr>
<td></td>
<td>• Results of the requested party are identifiable (only) for them by a specific ID – allows replicating the former matching results with direct access to the origin of the hits (ADEP+), thus facilitating process automation on data owner’s side.</td>
<td>• Take future developments into account, notably AI, in order to further improve the matching results.</td>
</tr>
<tr>
<td></td>
<td>• Aim: to significantly improve the success rate of (manual) requests for information in comparison to today.</td>
<td>• Still: shortcomings of a very limited set of biographic data can never reach the quality level of matching biometric data. Significant improvements are only possible by adding additional information (context information).</td>
</tr>
</tbody>
</table>
### 7. How to cope with expected high number of requests?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only mentioned as 'additional workload' in the Costs-Benefits-Assessment, as far as automation is not implemented.</td>
<td>Assumption: by default all MS will be queried, resulting in hits which would not be found nowadays (see Art. 5 [1] FD 2006/960/JHA: ‘factual reasons to believe that relevant data is available in another MS’). This will lead to more MS affected with false positives (see question 6) but foremost to more meaningful information of MS which were not considered as holding relevant information based on the events of life under investigation = dramatic increase of effectiveness. Success stories will lead to an increase of using EPRIS, becoming an initial routine before any international LEA request. Automation and probably re-allocation of resources will be necessary.</td>
<td>Consider setting common prerequisites for automation of processes in the legal frame (such as defining a data format standard [presumably UMF], defining minimum data fields and metadata elements, etc.). Consider supporting MS in implementing solutions for process automation (funding, designing common solutions to be implemented at national level). Take future developments into account, notably AI.</td>
</tr>
</tbody>
</table>
8. Shall additional information be provided with a hit?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
</table>
| • The consolidated query response should provide a list of “hits”, enlisting Member States where information on the queried individual is available together with contact data of the national contact point and categorised information concerning the offence type for which an individual has been suspected and the date of offence. (18/27 MS) | • Assumption that certain information in addition to hit/no-hit information would support assessment of the hit-information received.  
• However, additional information is not included in the pilots due to legal limitations (no specific frame existing) and in order to reduce complexity of solution piloted. | • Limited amount of additional information (such as the offence type and date of offence) does not suffice to take measures but it may suffice to take a better informed decision for further process steps incl. prioritization.  
• The feature adds complexity to the solution and to the requirements of MS data sources.  
• The feature increases data protection issues.  
• Consider incremental approach in implementation (start small without added information and grow over time). |

This project was funded by the European Union’s Internal Security Fund - Police
9. Is a new system really necessary?

<table>
<thead>
<tr>
<th>Feasibility Study</th>
<th>Project</th>
<th>Considerations re Prüm Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The study results revealed that the majority of MS confirmed the need for an</td>
<td>• Two pilots funded by the European Commission were conducted, the second</td>
<td>• Consider complimentarity of existing systems/tools incl. EPRIS in the entire EU law enforcement information ecosystem.</td>
</tr>
<tr>
<td>improved use of several currently existing law enforcement instruments.</td>
<td>one is still ongoing.</td>
<td>• Experience of the two pilots should be taken into account, without pre-dominating the future solution.</td>
</tr>
<tr>
<td>• The low amount of data inputted, the vagueness of data types and inadequate</td>
<td>• Six MS and Europol are active partners in the project and two MS are</td>
<td>• Considerations to evolve PRÜM frame (notably architectural considerations) already facilitate the implementation of EPRIS similar to the pilots and in line with the recommendation of the feasibility study, since central services in a hub-and-spoke-model might be established anyway.</td>
</tr>
<tr>
<td>data quality are examples of reported shortcomings</td>
<td>are observing.</td>
<td></td>
</tr>
<tr>
<td>• Four information systems were in focus: EIS, SIENA, SIS II, PRÜM.</td>
<td>• The feasibility of a new system has been proven.</td>
<td></td>
</tr>
<tr>
<td>• Recommendation: Maximize the use of existing systems; if not successful within</td>
<td>• It could be shown how EPRIS would even strengthen central systems, notably EIS,</td>
<td></td>
</tr>
<tr>
<td>three years (2015), then a pilot project should be initiated with the aim to</td>
<td>since EPRIS can establish unknown cross-border links to a larger extent than</td>
<td></td>
</tr>
<tr>
<td>evaluate the technical feasibility and impact of a new, specific EPRIS system.</td>
<td>presently.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Recommended rule: In case of a confirmed hit the EPRIS participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>affected shall provide the data to the most possible extent to a central</td>
<td></td>
</tr>
<tr>
<td></td>
<td>information system, thus making the information available for other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>interested parties.</td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

- The need for a law enforcement capability to search for a limited set of biographical data in relevant data stored at national level has been proven; such functionality is complimentary to central systems.

- The feasibility of such a system has been proven by the two pilots conducted; the scope and the mode of operation of such a system were substantiated and requirements have been described and tested.

- The solution piloted is in line with the recommendations of the feasibility study and would fit to the considerations of the Commission in terms of the revision of the Prüm decisions.

- Several aspects, notably the impact on existing (national) processes incl. the potential need for automation in order to adequately cope with the expected high amount of data to be processed, need closer attention.

- An EU legal basis is necessary to gain legal certainty and a binding character for using EPRIS; the revision of the Prüm decisions is the best opportunity to establish a proper legal basis for the EPRIS functionalities.
Thank you for your interest
Example for similarity search

- Decomposition of the search terms into N-grams
  N=2: Baggins à {" B", "Be", "eu", "ut", "tl", "li", "in", "n " }

- Apply several different cryptographic hash functions to each element of an N-gram.

- Mapping the hash values to a bit vector of given length

- Probabilistic wildcard search with Bloom filters

- Fuzzy search via Bloom filters is resource-intensive and dependent on the chosen hash algorithms and parameters

Source: Fraunhofer FOKUS (07/07/2016)
EPRIS Webfrontend Demonstrator Example for search results

Traffic light as indicator for the quality of the EPRIS hit
- Illustrative example only
- Thresholds for green, yellow and red are currently being defined
- The upcoming new software version 2.2 will include yellow

Strong hit (0.84)

No hit: 0.00 - 0.30 (Exclusive max.)
Strong hit: 0.30 - 1.00 (inclusive max.)

Detailed View

Response from France

25.03.21, 19:25
Query UUID: DE 8acfd3f6674daa646f8b7c173337a
One hit

Strong hit (0.84)

Next steps: A Hit has been found. A Follow-Up process should be initialized.

First name: 0.07
Family name: 1.00
Name Union: No value
Gender: Not requested
Day of birth: Not requested
Month of birth: Not requested
Year of birth: 1.00
<table>
<thead>
<tr>
<th>Project ID:</th>
<th>952740</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project title:</td>
<td>European Police Records Index System (EPRIS)</td>
</tr>
<tr>
<td>Project manager:</td>
<td></td>
</tr>
<tr>
<td>Project duration according to the grant agreement:</td>
<td>1st of February 2020 to 31st of May 2021 (16 months)</td>
</tr>
<tr>
<td>A project extension until 31 January 2022 has been requested</td>
<td></td>
</tr>
<tr>
<td>EU funding:</td>
<td>Internal Security Fund, 1 million Euros</td>
</tr>
<tr>
<td>Project partners:</td>
<td>Belgium (Police Federale Belge), Finland (Poliisihallitus), France (Ministere de L’Interieur), Germany (Bundeskriminalamt and Bundesministerium des Innern, für Bau und Heimat), Ireland (An Garda Siochana) and Spain (Ministerio del Interior) – as beneficiaries</td>
</tr>
<tr>
<td></td>
<td>Europol (no beneficiary)</td>
</tr>
<tr>
<td>Observers:</td>
<td>Hungary (Belügyminisztérium)</td>
</tr>
<tr>
<td></td>
<td>The Netherlands (Nationale Politie)</td>
</tr>
<tr>
<td>External project partners:</td>
<td>Fraunhofer Institute for Open Communication Systems (short: Fraunhofer FOKUS)</td>
</tr>
<tr>
<td></td>
<td>PD – Berater der öffentlichen Hand GmbH (short: PD)</td>
</tr>
</tbody>
</table>
### Current EU-project objectives and achievements

#### 1. Stabilization of the software solution and processes
- Implementation software versions 1.2, 2.0 and 2.1
- Development software version 2.2
- Implementation of a new EPRIS client

#### 2. Definition & implementation improved matching behaviour
- Implementation of SIS II name transliteration
- Definition and alignment enhanced matching rules
- Development of a matching score (traffic lights)

#### 3. Building trust into the system
- Exhaustive system integration tests and local testing
- Endorsement of a master test plan
- Use of pseudonymisation component by RKI related to COVID 19

#### 4. Planning transition of central services from project to operation
- Central Roll-out Coordinator established
- Analysis of central services and effort estimation with Europol
- Draft transition plan of central services

#### 5. Process improvement towards more automation
- Analysis of process steps where automation could apply
- Development of ADEP+ functionality

#### 6. Involvement of more partners
- Belgium takes part as project partner
- The Netherlands welcomed as new observer
Ultimate goal:
an operating EPRIS with a proper legal basis
probably coming into existence with Prüm Recast (~2025)

Key objective until then:
Preserving the EPRIS community and the system, as well as the knowledge and experience gained during the two EU-funded pilot projects

Keep community

Maximum extension of the EU-funded period of the current project
After: keep continuing project work while causing minimum efforts
Examining all legal possibilities to continue with final BAT-phase
Explore every possibility for another EU-funding

Keep technology

Defining data protection concept
Identifying other use cases
Support any other use case with an appropriate legal basis at hand
Explore potential activities within the European Innovation Hub
Key features of the ADEP technology

- Practitioner-centred approach
- Strong process-orientation incl. integration in existing systems
- Making available more sources of distributed information
- Cost-efficient while using existing infrastructure and rather low technical efforts
- Confidentiality of personal data is kept from both operational point of view and data subject‘s point of view
- Advanced privacy-by design solution, using pseudonymisation and micro services architecture, adaptable to different use cases
- Support of exact and tolerant search using pseudonymised data
Further information

- General information: see documents to be disseminated after the workshop (project dossier, current project status, FAQ)

- About EPRIS and the privacy-by-design solution of the ADEP-technology:

  https://www.bka.de/DE/UnsereAufgaben/Ermittlungsunterstuetzung/ElektronischeFahndungsInformationssysteme/Polizei2020/EPRIS_ADEP/EPRIS.html

  or search for BKA + EPRIS on the internet

- Prüm Recast: