Data Protection Law and Multi-Party Computation: Applications to Information Exchange between Law Enforcement Agencies

Amos Treiber
TU Darmstadt
treiber@encrypto.cs.tu-darmstadt.de

Dirk Müllmann
University Frankfurt
muellmann@jur.uni-frankfurt.de

Thomas Schneider
TU Darmstadt
schneider@encrypto.cs.tu-darmstadt.de

Indra Spiecker genannt Döhmann
University Frankfurt
spiecker@jur.uni-frankfurt.de

ABSTRACT

Pushes for increased power of Law Enforcement (LE) for data retention and centralized storage result in legal challenges with data protection law and courts—and possible violations of the right to privacy. This is motivated by a desire for better cooperation and exchange between LE Agencies (LEAs), which is difficult due to data protection regulations, was identified as a main factor of major public security failures, and is a frequent criticism of LE.

Secure Multi-Party Computation (MPC) is often seen as a technological means to solve privacy conflicts where actors want to exchange and analyze data that needs to be protected due to data protection laws. In this interdisciplinary work, we investigate the problem of private information exchange between LEAs from both a legal and technical angle. We give a legal analysis of secret-sharing based MPC techniques in general and, as a particular application scenario, consider the case of matching LE databases for lawful information exchange between LEAs. We propose a system for lawful information exchange between LEAs using MPC and private set intersection and show its feasibility by giving a legal analysis for data protection and a technical analysis for workload complexity. Towards practicality, we present insights from qualitative feedback gathered within exchanges with a major European LEA.

CCS CONCEPTS

• **Security and privacy** → **Privacy-preserving protocols**; **Privacy protections**; • **Social and professional topics** → **Database protection laws**; • **Applied computing** → **Law**.

KEYWORDS

data protection law; secure computation; law enforcement

ACM Reference Format:
Amos Treiber, Dirk Müllmann, Thomas Schneider, and Indra Spiecker genannt Döhmann. 2022. Data Protection Law and Multi-Party Computation: Applications to Information Exchange between Law Enforcement Agencies. In *Proceedings of the 21st Workshop on Privacy in the Electronic Society (WPES '22), November 7, 2022, Los Angeles, CA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3559613.3563192

WPES '22, November 7, 2022, Los Angeles, CA, USA

ACM ISBN 978-1-4503-9873-2/22/11…$15.00
https://doi.org/10.1145/3559613.3563192

1 INTRODUCTION

Law Enforcement Agencies (LEAs) are charged with Law Enforcement (LE), i.e., carrying out the rules that govern societies. Public security is a LE responsibility with a lot of attention. Consequently, this is also an area where LE has come under a lot of scrutiny, e.g., if criminal mass casualty events like terrorist incidents were not prevented by LEAs, for instance due to a lack of coordination between them across different domains, government levels, and jurisdictions. Even further, in the modern electronic society, criminal actors may be perceived as more powerful because of easy and secure communication for organizing across jurisdictions.

Therefore, LE is sometimes portrayed to be hindered by complex data protection rules to share data of subjects between different LEAs. As a result, there is a political push for less strict rules and increasingly centralized storage for better coordination and information exchange between LEAs, possibly violating citizens' right to privacy for security. For instance, in Germany, LEAs are currently in the process of developing and will soon deploy a centralized storage system of all of Germany's LEA databases in order to facilitate an effective information exchange between them [26]—despite the fact that the European Court of Justice (ECJ) has identified several legal problems concerning such central storage.[1]

Privacy-enhancing technologies relying on techniques from applied cryptography like Secure Multi-Party Computation (MPC) [20, 54] and Private Set Intersection (PSI) [25, 38] have long been argued and assumed to solve such issues stemming from data that cannot be analyzed because it cannot be shared for privacy reasons. Fundamentally, they allow computation on private data without the need to share or centralize plaintext information. Hence, they can be used to obtain information on data that *remains decentralized*, thereby offering a lot of potential to alleviate the above conflicts between LE and the right to privacy. Such a secure computation approach could provide more trust in how LEAs operate on private data: Ideally, this could enable them to adequately protect society while simultaneously ensuring citizens' privacy. However, research on this topic (surveyed in [13, 14]) largely focuses on enabling surveillance and not LEA information exchange. Furthermore, works often do not include perspectives of experts from data protection law and LE.

[1] ECJ, Judg. of 08.04.14 - C-293/12, C-594/12 (Digital Rights Ireland), Rn. 27; Judg. of 21.12.2016 - C-203/15, C-698/15 (Tele 2 Svergie AB), Rn. 93; Judg. of 06.10.2020, C-511/18, C-512/18, 520/18 (La Quadrature du Net).

For moving towards actually deploying such solutions to alleviate these problems, establishing the compliance to data protection law and feasibility for real-world LE is crucial.

In this work, we focus on a potential mitigation of pushes for privacy-invasive centralized LEA data storage by proposing and analyzing a decentralized system for lawful and private information exchange between LEAs via MPC/PSI techniques. Importantly, our goal is not to use the above problems as abstract motivation to develop sophisticated cryptographic techniques. Instead, our approach is interdisciplinary and oriented in feasibility throughout, developing our architecture according to the normative requirements of German and European data protection law and taking into account practical feedback in collaboration with a major European LEA situated in Germany. As such, we offer both extensive technical and legal analyses of MPC as well as LE information exchange and an architecture for it, and report on a qualitative study with a LEA to establish its feasibility and degree of practicality. Due to the high standards of European data protection, our approach may also be transferable to other legal systems.

Our Contributions. In particular, we contribute the following:

(1) Given that most works on MPC for data protection miss a view by legal experts, we lay out the fundamentals of European data protection law as a primer for a legal perspective on how to use secure computation techniques for data protection, giving an analysis of secret sharing-based MPC.
(2) As the application focus of our paper, we investigate the case of data protection and information exchange between LEAs in Europe. We perform a legal and a technical analysis of information exchange between LEAs as well as of the concrete systems currently proposed in Europe and Germany.
(3) Using the above analyses, we propose a system for lawful and private information exchange between LEAs relying on MPC techniques and PSI. We conduct a legal examination of the system to ensure data protection and a technical examination of a proof of concept implementation to establish feasibility.
(4) Throughout our paper, we not only offer technical and legal perspectives but also practical LE feedback. We obtain this via a qualitative exchange in a collaboration with a major European LEA, the police of the city state of Hamburg, Germany, taking into account their assessment if a system like ours can actually be deployed. Our intention here is that our ideas may not remain academic but also may lay foundations in LEAs towards effective LE that *does not push against the right to privacy*.

Our results show that our system would be feasible to develop and operate but that many practical challenges would still lie ahead.

We hope our work may help in decreasing persistent and reoccurring efforts towards centralized LE data storage, similar to repeatedly ruled unconstitutional systems for data retention still being pushed for[2], in particular in light of recent decisions to unify registries despite heavy criticism and data protection and IT security friendly alternatives[3]. As we combine efforts from legal studies, cryptography, and LE, we think that ideally research like ours could establish—also in the LE community—that privacy-preserving and lawful information exchange between LEAs is achievable and feasible, possibly subsiding the push for centralized storage.

Citations in this Paper. Our interdisciplinary approach necessitates two different citation styles. Good scientific practice in law requires very specific and highly variable source citations. Without them, adequate identification of sources is just as impossible as their retrieval. To meet these standards, we present the *legal citations* in our work in the form of footnotes. Furthermore, because legal publications are often specific to individual countries and published in their languages without adequate equivalent publications in English, citations in this work can refer to non-English texts.

2 RELATED WORK AND PRELIMINARIES

Here, we review related academic work. For an overview over data protection law, see Sect. 3. We survey proposed real-world systems for LEA information exchange in Sect. 4.3.

2.1 Secure Multi-Party Computation (MPC)

First proposed by Yao in 1982 [53], MPC allows multiple parties to interactively compute the output of any function in a secure manner, i.e., without revealing additional information about the inputs. Popular protocols include Yao's Garbled Circuits (GCs) [33, 54] and the protocol of Goldreich-Micali-Wigderson (GMW) [20].

Generally, MPC protocols enable the secure evaluation of a circuit representation of the function by providing composable ways to securely compute fundamental gate types. A useful abstraction is to see MPC protocols as running on *secret shares* of a secret that cannot be recovered by just one secret share alone. For two parties for simplicity, a secret x can be shared as $(\langle x\rangle_0, \langle x\rangle_1)$, where each $\langle x\rangle_0$ or $\langle x\rangle_1$ provably does not give any information about x. E.g., in the GMW protocol, one can set $\langle x\rangle_0 = r$ and $\langle x\rangle_1 = x \oplus r$ for a bit r chosen uniformly at random. MPC protocols are composable and operate on secret shares at each gate $z = f(x, y)$ in the circuit by yielding $\langle z\rangle_i$ to the party $i \in \{0, 1\}$ given just $\langle x\rangle_i$ and $\langle y\rangle_i$ as inputs, without leaking any information. The final result can be revealed to one or both parties by sending the secret shares of a party to the other. Thereby, MPC protocols securely yield an output, where security is understood as not leaking more knowledge about the data than the willingly revealed output.

MPC is considered in several security models. The main models are *semi-honest* and *malicious* security. In the semi-honest security model, the parties are assumed to not deviate from the protocol specification but they still want to find out additional information about the secret data by observing the transcript of the protocol. In the malicious security model, parties additionally are allowed to deviate from the protocol. We focus on semi-honest security due to the results of our legal analysis of semi-honest MPC (cf. Sect. 3.2).

2.2 Private Record Linkage

The application to use MPC to link records of databases according to specific properties in a privacy-preserving manner is crucial in many domains, e.g., it has been recently used for studying of rare diseases across patient databases in real-world medical environments [32, 46]. Further applications and techniques for private record linkage are surveyed in [50]. Secure linking of databases

[2]ECJ, Judg. of 08.04.14 - C-293/12, C-594/12 (Digital Rights Ireland), Rn. 27; Judg. of 21.12.2016 - C-203/15, C-698/15 (Tele 2 Svergie AB), Rn. 93; Judg. of 06.10.2020, C-511/18, C-512/18, 520/18 (La Quadrature du Net).
[3]Cf. Sorge/Spiecker/von Lucke, Registermodernisierung - Datenschutzkonforme und umsetzbare Alternativen, BT-Ausschussdrucksache 19(4) 667 C.

between different entities is closely related to Private Set Intersection (PSI), which is the secure computation of the specific functionality $f(X, Y) = X \cap Y$ for private input sets X and Y. It can be instantiated via public-key cryptography [11, 34] or Oblivious Transfer (OT) [18, 30, 37, 40, 41] using mostly symmetric cryptography due to OT extensions [3, 27]. PSI instantiations relying on securely evaluating a circuit with MPC are called *circuit-based* PSI. Although a naive circuit would require quadratic costs in the input size, subsequent works [9, 10, 24, 25, 37–39, 42] have put down the complexity to an amount of gates linear in the input size [9, 38, 42]. Extensions for *fuzzy* set intersections are given by [19, 49] and for *multi-party* circuit-based PSI are given by [8]. Due to the composability of MPC, this approach allows to compute another function $g(X \cap Y)$ with the intersection as an input. Furthermore, some protocols allow for payloads associated with the input sets, i.e., securely computing some functionality

$$g\left(X \cap Y, \left\{\left(P^X(k), P^Y(k')\right) \mid X(k) = Y(k') \wedge X(k) \in X \cap Y\right\}\right)$$

on ordered sets P^X and P^Y associated with the ordered sets X and Y.

2.3 LE and Data Protection via Cryptography

Feigenbaum and Weitzner [15] describe a tension between LE and cryptography that could possibly be alleviated by appropriate deployments of secure computation techniques like MPC/PSI with clear legal rules for lawful use that do not violate privacy. Using such techniques within LE was overwhelmingly considered for LE access or *surveillance*, i.e., methods to accountably compare private data to those of LEAs related to a suspect given that the LEA has an appropriate warrant. Overviews can be found in [13, 14]. Segal et al. [45] use PSI for warrants of cell-tower dumps, which is extended by [44] as a system for privacy-preserving contact chaining. Kroll et al. [31] propose secure protocols to execute warrants for private data and Kamara [28] proposes a privacy-preserving alternative to the NSA metadata program. A system for privacy-preserving and accountable electronic surveillance is given by Frankle et al. [17]. Enabling LE access in end-to-end encrypted systems is considered by Green et al. [22]. Another approach [52] enables decryption by the government with arbitrary costs to limit abuse. The case where the law itself is secret has been covered by Goldwasser and Park [21]. Furthermore, [4] investigate legal dilemmas where facts need to be verified while being kept confidential and the relation of zero-knowledge proofs to them, and [6] use zero-knowledge proofs to provide a mechanism where LEAs can prove, in a court, claims obtained by LE software while keeping the software hidden. To the best of our knowledge, lawful exchange of data *between* LEAs has not been explicitly considered so far.

Furthermore, most of the above works are concerned with the legal context of the United States, whereas we provide a European perspective. Here, Helminger and Rechberger [23] provide an understanding of the role of MPC in the General Data Protection Regulation (GDPR). They focus on the role that MPC can have in meeting the requirements of the GDPR as a privacy-enhancing technology. They consider on the one hand the understanding of the concept of personal data in view of the dispute as to whether an absolute or relative view should be applied when determining whether data can be related to a person, which has been decided by the European Court of Justice (ECJ)[4]. On the other hand, they examine whether the requirements of Art. 25 GDPR *can be met by means of MPC*. Schreibner at al [43] look at possible applications of privacy-enhancing technologies for the exchange of health data in medical research and clinical practice under GDPR. Apart from the different fields of application and the associated different legal data protection requirements, they primarily consider the aspect of the anonymity or pseudonymity of the data exchanged. In contrast, the legal object of our work is to determine *which data protection requirements* the use of MPC is subject to, whether the use of the technology in the context of data exchange between LEAs is at all possible under data protection law, and what legal advantages it offers over other technologies used to date—also for the preservation of the fundamental rights of citizens and society, thereby taking into account all the main data protection requirements.

3 DATA PROTECTION LAW

In this section, we give a brief primer on European data protection regulations and the interrelation with MPC. These regulations are internationally regarded as the gold standard[5] of data protection law, which other countries follow in new legislation[6] and which is assumed that if adhered to, the product can also be distributed in all other countries, as they are the strictest possible standard.

3.1 Fundamentals

The process of European unification of data protection law began in the 1990s with the Data Protection Directive[7], on which national data protection laws were based from then on. Afterwards, European integration was completed in 2015 by the General Data Protection Regulation (GDPR)[8], which entered into force in May 2018. Within its broad scope, it applies directly throughout the Union[9]. Only at points where the GDPR explicitly grants national legislators regulatory scope, particularly in the form of so-called *opening clauses*[10], they are permitted to legislate and continue to have room for autonomous national regulations. An important area that is completely outside the scope of the GDPR is the area of LE

[4] ECJ, Judg. of 19.10.2016 - C 582/14 (Breyer/Deutschland, NJW 2016, 3579.

[5] Cf. Jan Philipp Albrecht, Der europäische Datenschutz als Goldstandard, in: VATM (Ed.), VATMN-Jahrbuch 2012/2013, 2013, pp. 34 et seqq.; Schünemann / Windwehr, Towards a „gold standard for the world"? The european General Data Protection Regulation between supranational and national norm entrepreneurship, Journal of European Integration, issue 7, Vol. 43, 2021, pp. 859 et seqq.; Taylor, Data Protection: threat to GDPR's status as "gold standard", 25th August 2020, https://www.ibanet.org/article/A2AA6532-B5C0-4CCE-86F7-1EAA679ED532.

[6] As to California cf. Hoeren/Pinelli, MMR 2018, pp. 711 et seqq.; Spies, ZD-Aktuell 2018, 06156; Botta, Der California Consumer Privacy Act und die DSGVO: Ein transatlantisches Zwillingspaar? in: Taeger (Ed.), Die Macht der Daten und der Algorithmen, 2019, pp. 567 et seqq.; Brasil Laubach/Dräger, ZD-aktuell 2018, 06254; Hoeren/Pinelli, ZD 2020, pp. 351 et seqq. Japan and Brasil cf. Hoeren/Wada, ZD 2018, pp. 3 et seqq.; Fujiwara/Geminn/Roßnagel, ZD 2019,pp. 204 et seqq.; Sarlet/Mendes in: Spiecker gen. Döhmann/Bretthauer, Dokumentation zum Datenschutz, D.2.25.0, sec. 19 et seqq.

[7] Directive 95/46/EC of 24th October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281/31, 23.11.1995.

[8] Regulation (EU) 2016/679 of 27th April 2016 on the protection of natural persons with regard to the processing of personal data on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119/1, 04.05.2016.

[9] Cf. Art. 288 par. 2 TFEU.

[10] Hornung/Spiecker gen. Döhmann in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, Art. 1, sec. 13; Laue, ZD 2016, pp.463 et seqq.

and police security[11] (Cf. Art. 2 par. 2 lit. d GDPR). In these areas, the GDPR is not applicable and national legislation continues to be permissible. However, this is in turn backed up by the so-called Justice and Home Affairs (JHA) Directive[12], which ensures a *common minimum level of protection* in this area throughout the EU. Although the GDPR is therefore not applicable to LEAs, an intensive examination of its content and principles is nevertheless indispensable. On the one hand, this is due to the fact that it represents the central legal framework of European data protection law, without which a fundamental understanding of European data protection law is impossible. On the other hand, the principles and standards of the GDPR are fundamental decisions and requirements that have also found their way into other data protection laws, especially the JHA Directive, which is relevant to our work.

Applicability. Data protection law is factually applicable only if personal data are processed. These are, according to Art. 4 No. 1 GDPR, any information relating to an identified or identifiable natural person. The limitation of the scope of application to personal data is *based on the functioning of fundamental rights*. They are defensive rights of the individual, i.e., the citizen, against the state, so that in principle only human beings can be bearers of fundamental rights.[13] Although many fundamental rights are also extended to legal entities[14] like companies, this is not possible for the fundamental right to data protection. This is due to the origin of the fundamental right to informational self-determination in German law, which is derived from the general right of personality and its rooting in human dignity, Article 2(1) of the German Basic Law (GG) in conjunction with Article 1(1) of the Basic Law Art. 1 para. 1 GG.[15] Human dignity can however still be entitled to persons.[16] However, if trade secrets, *non-personal data* are nevertheless subject to legal protection by national laws, which are also protected under European law by the Trade Secrets Directive[17].

Within the personal data, the so-called *special categories* of personal data in the sense of Art. 9 GDPR are placed under special protection. These are particularly sensitive data concerning personal views or characteristics with a high potential for discrimination and are closely related to the exercise of activities particularly protected by fundamental rights.[18] Their processing poses a particular risk of violating a person's privacy or his or her fundamental rights and freedoms.[19] For this reason, their legal level of protection is even higher compared to that of "simple" personal data.

Data Processing. The lawfulness of processing data under European data protection law, even by means of MPC, cannot be determined in a blanket manner, but only on the basis of the respective, concrete use case. To keep pace with technological progress and avoid circumvention of data protection regulations by new technologies, European data protection law is technology-neutral.[20] Instead, data processing must always meet basic requirements. The law initially assumes a so-called *prohibition with a reservation of permission*, meaning that data processing is prohibited unless the law explicitly permits it.[21] As a consequence, any processing of personal data must be based on a *processing basis*, which may result from the GDPR itself, in particular Art. 6 GDPR, or special laws. In view of the comprehensive standardization, national standards can only permit data processing within the scope of the GDPR if an opening clause is provided for this purpose. While, especially in the commercial context, the permissibility of data processing in law otherwise often depends on a balancing of interests[22], the most important basis for permission for data processing in practice is probably the consent of the person affected by the data processing[23]. In order to be effective, however, it must be given voluntarily, for a specific case, in an informed manner and unambiguously, as defined in Art. 4 No. 11 GDPR. In particular, according to Art. 7 sec. 4 GDPR, it is subject to a so-called *prohibition of tying*, according to which consent is not deemed to have been given voluntarily if the performance of a contract is made dependent on consent to the processing of personal data that is not necessary for the performance of the contract. Unlike in the context of contractual situations, obtaining consent to processing in a government context is regularly not possible due to the coercive situation that is to be assumed, as it therefore lacks voluntariness. Despite these problems in obtaining effective consent, carrying out data processing on the basis of consent offers the possibility of extensive processing operations. In addition, consent represents a means that is known not only within the GDPR, but also to all special laws, which enables the unification of the legal bases of data processing operations within a company.

Further general requirements can be found in particular in Art 5 GDPR, which establishes the principles of *lawfulness* (Art. 5 sec. 1 lit. a GDPR), *purpose limitation* (lit. b), *data minimization* (lit. c), *accuracy* (lit. d), *storage limitation* (lit. e), and *integrity and confidentiality* (lit. f). In particular, purpose limitation, which requires that the collection of data takes place only for a purpose that has already been defined in advance and that the data are not subsequently used in a manner incompatible with this purpose, has proven in practice to be a limiting factor for data processing. This is all the more true as the collection is limited to a necessary extent by the principle of data minimization and a storage of the data is limited to the necessary period of time in view of storage limitation.

[11] For the meaning of this term cf. Roßnagel, in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, Art. 2, sec. 40; Müllmann in: Kießling (Ed.), Infektionsschutzgesetz, 2021, Vor. §§ 6 ff., sec. 8.

[12] Directive (EU) 2016/680 0f 27th April 2016, on the protection of natural persons with regard to the processing of personal data by competent authorities for the purpose of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA; OJ L 119/89, 04.05.2016.

[13] Hufen, Staatsrecht II, 2. Ed, 2009, §1, par. 6; §6, par. 31.

[14] Enders in: Epping/Hillgruber (Eds.), BeckOK Grundgesetz, 50th Ed., Art. 19, par. 39; Hufen, Staatsrecht II, 2. Ed, 2009, §6, par. 36.

[15] Hufen, Staatsrecht II, 2. Ed, 2009, §12, par. 6; Enders in: Epping/Hillgruber (Eds.), BeckOK Grundgesetz, 50th Ed., Art. 19, par. 40.

[16] Enders in: Epping/Hillgruber (Eds.), BeckOK Grundgesetz, 50th Ed., Art. 19, par. 40; Hufen, Staatsrecht II, 2. Ed, 2009, §12, par. 6.

[17] Directive (EU) 2016/943 of 8th June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157/1, 15.06.2016.

[18] Kühling/Buchner in: Kühling/Buchner (Eds.), Datenschutz-Grundverordnung/Bundesdatenschutzgesetz, Art. 9, par. 1; Franzen in: Franzen/Gallner/Oetker (Eds.), Kommentar zum europäischen Arbeitsrecht, 4. Ed., Art. 9 DSGVO, Rn. 1.

[19] Recital 51 GDPR; Kühling/Buchner in: Kühling/Buchner (Eds.), Datenschutz-Grundverordnung/Bundesdatenschutzgesetz, Art. 9, par. 1; Franzen in: Franzen/Gallner/Oetker (Eds.), Kommentar zum europäischen Arbeitsrecht, 4. Ed., Art. 9 DSGVO, sec. 1.

[20] Recital 15 GDPR.

[21] Instead of many: Albrecht in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, Art. 6, sec. 1.

[22] Cf. in particular Art. 6 sec. 1 p. 1 lit. f GDPR.

[23] Cf. Art. 6 sec. 1 p. 1 lit. a GDPR.

Rights and Obligations. The GDPR also provides for comprehensive rights for persons affected by data collection[24], e.g., to information, correction, deletion, and even transfer of data. It obliges data processors to provide information about the collection and further processing of data[25] and to implement further technical and organizational measures. These include, among others, compliance with the principles of privacy by design and privacy by default (Art. 25 GDPR), maintaining an adequate level of IT security (Art. 32 GDPR), keeping an inventory of data processing operations in a company (Art. 30 GDPR), and, under certain circumstances, conducting a data protection impact assessment (Art. 35 GDPR).

3.2 MPC and Data Protection

The requirements described in Sect. 3.1 must be met by processing using MPC if used for personal data. According to Art. 4 No. 2 GDPR, processing is "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction". Since even the structuring or structured storage of data, which is also performed for data input to MPC, constitutes processing, *data protection law always applies initially to applications using MPC*, insofar as it involves personal data[26]. However, in terms of data protection law, the whole process can subsequently be split into several individual processing steps.

Applicability to Secret Shares. In further course, data protection law and requirements are only applicable for the respective processing stage if personal data continue to exist (cf. Art. 2 sec. 1 GDPR). When using MPC on secret shares (here for the case of two parties), this appears doubtful at any rate from the creation of the first secret share for its respective recipient. Despite the change in the information content, the *sender* of secret shares of an input in standard MPC can continue to assume that the information embodied in the secret share $\langle x \rangle_1$ (cf. Sect. 2.1) is personally identifiable, since $\langle x \rangle_0$ remains stored on the sender's machine and the sender thus has the possibility of restoring the personally identifiable nature of the data x[27]. The *recipient* of a secret input share does not have this possibility, so that she cannot restore the personal reference of the secret share and it is an anonymous date for her. Secret shares generated subsequently in MPC are, hence, anonymous data, too. *Data protection law does not apply to such data.*[28]

Applicability to other Parts. The applicability of data protection law to the other parts of the processing of personal data by MPC (like the structuring of the inputs) does not mean that the execution of these processing steps is per se impossible or prohibited. Rather, when they are carried out, the requirements imposed by law must be observed. In this respect, non-applicability merely facilitates data processing from both a technical and legal perspective. The requirements to be met include, in particular, *the existence of a legal basis for processing* for each processing step subject to data protection law (Art. 6 (1) GDPR). *Whether there is a legal basis for carrying out a processing step is determined on the basis of the specific case in which and for which MPC is to be used, and cannot be determined across the board*[29]. In the private sector in particular, however, this will regularly be obtained by granting consent to the person whose data is to be processed. As described above, strict requirements must be placed on the granting of consent under data protection law, whereby the requirements of informedness and voluntariness of consent (cf. Art. 4 No. 11 GDPR), but also the so-called prohibition of tying of Art. 7 (4) GDPR are of particular importance[30]. The other requirements listed in Sect. 3.1, in particular the principles of Art. 5 GDPR or the data subject rights of Art. 12 et seq. GDPR, must always be observed where parts of using MPC constitutes processing of personal data. Similarly, a legal basis of *revealing of the output* computed via MPC must be ensured.

Security Models. Insofar as MPC is not carried out in the malicious security model (cf. Sect. 2.1), in which a violation of the rules inherent in the system becomes apparent, there is a risk that the protocol and thus also the assumptions and requirements under data protection law can be maliciously circumvented by parties assumed to be semi-honest in the security model. There, compliance should be ensured through the implementation of technical and organizational measures or neutral certification (cf. Art. 32 GDPR).

General Assessment. The legal requirements described in the above concern the different parts and aspects of using MPC: The preparation of inputs, the secure computation on secret shares, the revealing of the output, and the used security model. If they can be met, from a data protection perspective, MPC then even represents a particularly data protection-friendly way of computing on personal data, as it prevents the exchange and knowledge of data that is not relevant (not part of the function output) for an exchange partner or may even not be obtained by her. MPC is thus a particularly data-saving and less intrusive way of computing between parties also in applications that are possible under the GDPR. This especially enables data to be used in situations in which there were previously data protection concerns. *MPC can therefore be a real game changer for the relationship between data privacy and technology.*

4 LEA INFORMATION EXCHANGE

Particularly after terrorist incidents, LE has been scrutinized for not exchanging information among different agencies across different jurisdictions and areas of responsibilities. E.g., within the investigation report [48] on the Berlin Christmas market attack of December 19th, 2016 by a commission appointed by the German parliament, a deficiency has been noted in [48, p. 1101, Sect. V]: The perpetrator became conspicuous to several LEAs in different areas

[24] Cf. Artt. 15 et seqq. GDPR.
[25] Cf. Artt. 12 et seqq. GDPR.
[26] Cf. Roßnagel in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, 2019, Art. 4 Nr. 2 DSGVO, sec. 17 et seq.; Schild in: Wolff/Brink (Eds.), BeckOK Datenschutzrecht, 40th Ed., 2022, Art. 4 DSGVO, sec. 43; Herbst in: Kühling/Buchner (Eds.), DS-GVO/BDSG, 3rd Ed., 2020, Art. 4 DSGVO, sec. 23.
[27] Cf. principles of the judgement ECJ, Judg. of 19.10.2016 - C - 582/14 (Breyer/Deutschland), NJW 2016, 3579.
[28] Hansen in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, 2018, Art. 4 Nr. 5 DSGVO, sec. 23; Zierbarth in: Sydow (Ed.), Europäische Datenschutzgrundverordnung, 2. Ed., 2018, Art. 4 GDPR, se. 24; Klar/Kühling in: Kühling/Buchner (Eds.), DS-GVO/BDSG, 3rd Ed., 2020, Art. 4 Nr. 1 DSGVO, sec. 31.
[29] Albrecht in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, 2018, Einführung zu Art. 6 DSGVO, sec. 1 et seqq.
[30] Golland, MMR 2018, 130; Krohm/Müller-Peltzer, ZD 2017, 551; critical Engeler, ZD 2018, 55; on the problems of consent in practice in general cf. Uecker, ZD 2019, 248.

of responsibilities early on. These individual instances in individual states initially did not warrant a "particular importance" or arrest[31], but in their entirety would have shown this.

As a response to such terrorist incidents, different LE entities within and across states initiated closer collaborations. Consequently, data protection law in Europe has increasingly been applied to LE compared to a prior focus on market actors [5]. However, such events are often also used politically to push for more powers and less data protection scrutiny for LE to enable more information exchange. The motivation is to prevent incidents by finding individuals with well-founded potential across LEAs for particular LE attention that do not receive particular attention by the individual agencies because their individual knowledge does not warrant it.

Therefore, the setting we are concerned with in this work is multiple LEAs wanting to match their existing databases containing lawfully obtained and stored data in a private manner. In order to enable this, a system must *satisfy the legal data protection requirements*, be *feasible to construct and run*, and *be effective* in finding instances as mentioned above. In the following, we describe the legal and technical challenges for such systems within Europe.

4.1 Legal Challenges

LEAs being able to freely match their databases (which we assume to be collected legally) provides potential of abuse as established by various courts[32] on the occasion of decisions in different cases concerning data exchange and data storage for LEAs. Relevant in this respect is the violation of fundamental rights of citizens[33], which can also occur if the legislator in this respect has created a legal basis for the implementation of data collection or exchange, but has not sufficiently taken into account fundamental rights. In addition, the courts also include observations on an overall monitoring calculation in its considerations and evaluate measures and their interaction also against this background.[34] In case of the violation of these fundamental rights and principles, the legislative decision, i.e., the law, is unconstitutional or contrary to European law.

Risks of Centralized Storage. In the case law of the German Federal Constitutional Court and the European Court of Justice (ECJ), central databases pose the risk of creating comprehensive profiles of citizens[35], which could affect the core of the fundamental right to data protection and impair a person's free and self-determined development[36]. This is a basic prerequisite for the functioning of a pluralistic and democratic society, which can only function without the feeling of constant surveillance and evaluation of knowledge about the individual[37]. Equally problematic is the high "spread width"[38] of a general collection of data[39] and problems in limiting the possibilities of access and verifying the legitimacy of access to the data[40]. In such a system, it is difficult to maintain the principles of purpose limitation under data protection law, according to which data may only be processed for the purpose for which it was collected, and proportionality[41], which must be ensured by means of the question of *hypothetical data recollection*[42]. Against this background, *cross-authority databases are only possible in very few cases and to a very limited extent*[43]. In addition, the storage of data in clouds that are not set up and managed by the authorities themselves poses additional problems.

Risks of Decentralized Data Exchange. Decentralized databases in which information can be retrieved or transmitted on an ad hoc basis also harbor problems and dangers. First of all, data exchange cannot take place here without cause, so important information may be disregarded because there is a lack of awareness of its existence. On the other hand, unregulated disclosure or regular exchange without cause again harbors the danger of information being obtained too comprehensively or even of profiling. It is also difficult to ensure compliance with the purpose limitation principle and proportionality when querying decentralized databases[44]. In the course of automated procedures, it is also problematic that spatial and temporal barriers are almost completely removed and, as with a central database, the data is potentially available to all recipients at all times[45]. This is all the more true since, in practice, there is often no verification of the legitimacy of the retrieval or a transmission, which means that the recipient unilaterally determines the process[46]. When there are several parties involved, it is more difficult to safeguard data subjects' rights. Against the background of these dangers, the *ECJ always requires a minimum significance of the crimes or dangers* to be averted with the transfer, so that their processing appears justified against this background[47].

4.2 Technical Challenges

The following approaches can enable LEAs to match individuals without exchanging plaintext data:

- **Deterministic hashing.** With this technique, records can be checked for equality by comparing their hashes. E.g., one could compare cryptographic hashes of the names and birth data of individuals to see if they co-occur. That way, no plaintext information is transferred. However, this just covers which individuals occur in multiple databases and not if the information present in those databases warrants further exchange of information about matching subjects. Furthermore, deterministic hashes could be reversed due to the low entropy of name and birth data information

[31] The perpetrator was at times viewed as a "small drug dealer" [48, p. 1091, Sect. IV].
[32] Since a well-founded understanding of the legal system is necessary for scientifically well-founded statements on legal criticism and dangers, but the legal systems, especially in the area of security/LE law, continue to be almost exclusively national in character, a restriction is made here to German law and the law of the European Union.
[33] The relevant sources and legal standards are the German Basic Law and the European Charter of Fundamental Rights.
[34] BVerfG, Judg. of 02.03.2010, 1 BvR 256/08, et al., sec. 218.
[35] ECJ, Judg. of 08.04.14 - C-293/12, C-594/12 (Digital Rights Ireland), Rn. 27; Judg. of 21.12.2016 - C-203/15, C-698/15 (Tele 2 Svergie AB), Rn. 93; Judg. of 06.10.2020, C-511/18, C-512/18, 520/18 (La Quadrature du Net).
[36] BVerfG, Judg. of 15.12.1983, 1 BvR 209/83 u.a. (Volkszählungsurteil), BVerfGE 65, 1.
[37] Müllmann, Brauchen wir ein Recht auf digitalen Herdenschutz? in: Greve et al. (Eds.), Der digitalisierte Staat, 2020, 129, 148.
[38] Meaning a measure affects many people not relevant to it.
[39] BVerfG, Judg. of 02.03.2010, 1 BvR 256/08 u.a.; BVerfGE 109, 279, 307 et seq.; BVerfGE 115, 320; BVerfGE 120, 378.
[40] BVerfG, Decision of 27.05.2020, 1 BvR 1873/13 (Bestandsdatenauskunft II); Judg. of 20.04.2016 - 1 BvR 966/09, 1 BvR 1140/09 (BKA Gesetz).
[41] Cf. ECJ, Opinion of 26.07.17 - Avis 1/15, Rn. 124, 137 et seqq.
[42] BVerfG, Judg. of 20.04.2016, 1 BvR 966/09, 1 BvR 1140/09 (BKA-Gesetz), NJW 2016, 1781; Müllmann, NVwZ 2016, 1692, 1693 et seq.
[43] BVerfG, Judg. of 24.04.2013, 1 BvR 1215/07 (Antiterrordatei).
[44] BVerfG, Decision of 27.05.2020, 1 BvR 1873/13 (Bestandsdatenauskunft II); Judg. of 20.04.2016 - 1 BvR 966/09, 1 BvR 1140/09 (BKA Gesetz).
[45] Petri in: Lisken/Denninger, Handbuch des Polizeirechts, 6. Ed., 2018, G. 955.
[46] Petri in: Lisken/Denninger, Handbuch des Polizeirechts, 6. Ed., 2018, G. 955.
[47] Petri in: Lisken/Denninger, Handbuch des Polizeirechts, 6. Ed., 2018, G. 1019.

and, due to their deterministic outputs, are a pseudonymization technique. As such, these hashes cannot just be exchanged to check for matches because, like the plaintext case, safeguards to satisfy legal demands for data protection would still be needed. In this respect, the hashed information is pseudonymized[48] and therefore continues to be personal data (Cf. Art. 4 No. 5 GDPR) to which data protection law and its requirements apply.

- **Private record linkage.** This approach enables linking subjects across the different databases in a secure manner. Hashing can be used as identifiers for the matching operation but an additional protection layer using, e.g., provably secure MPC securely computes these matches, thereby not even revealing the hashes to other parties. The advantage of using MPC or circuit-based PSI here is that secure computations can be built on top of the matching result. Since lawful information exchange between LEAs requires more properties to be present than just the mere fact that an individual appears in multiple databases, these legally required properties can be securely computed such that at the end only the matches fulfilling the requirements are revealed at the end of the secure computation. Because of this and the advantages of MPC w.r.t. data protection law (cf. Sect. 3.2), this approach can satisfy data protection requirements to allow for lawful information exchange between LEAs (as we show in Sect. 5.2).

While other privacy technologies that manipulate the data, like anonymization, are also frequently used to solve computations on data from multiple sources without harming privacy, these inherently remove information that uniquely determine individuals. Since the task here is to uniquely determine individuals according to specific legal requirements without learning information about other individuals that do not, they are not applicable here.

LE Feedback. Data protection safeguards are also seen as a major technical challenge to address in our exchange with the police of Hamburg, Germany, however, further aspects were identified as technical challenges. These are more general problems of LE information management in federal Germany, as the different LEAs use different systems and standards that need to be made compatible. As a response, a program as part of a national strategy to modernize and standardize the information management of LEAs called "Polizei 2020" ("Police 2020") [26] was initiated. It contains efforts to harmonize the systems and standardize the data encountered in LE information management. Furthermore, while new standards may capture categorical data like elements of criminal offenses committed, we found that the impression in this LEA is that this data is often impractical to collect as officers would spend too much time noting or selecting the appropriate elements when entering data. Thus, in addition to ensuring data protection, practical challenges like these remain in order to facilitate lawful and effective LEA information exchange.

4.3 Previously Proposed Systems

The following systems have been proposed:

Europe. The European Police Records Index System - Automation of Data Exchange Processes (EPRIS-ADEP) allows members to check other members' databases for individuals by passing on a pseudonymized query to the other member, who matches it with their own pseudonymized database. If a match is found, the plaintexts can be exchanged after manual verification. Not many technical details can be found about this system, but according to the evaluation report [47], it seems like matching *and* pseudonymization is performed via Bloom Filters. We argue that this may not be sufficient protection for the legal requirements of Sect. 4.1, as members can still infer information about individuals subject to their received queries that are not in their databases because Bloom Filter queries can be seen as deterministic hashes, corresponding to the deterministic hashing approach discussed in Sect. 4.2. Thus, a queried individual's data is exchanged via a pseudonym in this system and has to be treated as such, requiring further data protection measures. This also only allows to check for equality or closeness of personally identifiable data and not associated criminal elements or other factors, which given the challenges outlined in Sect. 4.1, would often not be enough to warrant an exchange of data.

Germany. In the Federal Republic of Germany, within "Polizei 2020", a centralized *data house* is proposed where according to the white paper [26] data of the different LEAs is stored. To comply with data protection, an access control mechanism should ensure only actors with appropriate clearance can access the data belonging to specific LEAs. However, this may not satisfy the data protection demands of Sect. 4.1: It seems unclear to what extent the problems and dangers discussed in Sect. 4.1 for central databases can be eliminated merely by granting different access rights. Either, *no exchange* takes place in view of the correct compartmentalization of the individual databases against each other. In this case, it offers no added value for the exchange and evaluation of data. Or otherwise, compartmentalization of the databases against each other is merely insufficiently done, resulting in a large common data pool. In this case, the dangers and problems of a central database mentioned in Sect. 4.1 would fully come through.

5 OUR SYSTEM FOR PRIVATE LEA INFORMATION SHARING

Currently proposed and to-be-deployed systems have data protection and privacy drawbacks (cf. Sect. 4.3). To mitigate these, we propose a distributed system based on MPC and analyze it with regards to legality, workload performance, and practicality next.

5.1 Architecture of Our System

We propose to use private record linkage with MPC and PSI to securely compute the identities of individuals where the data present in all databases legally warrant lawful exchange of information relating to those identities (according to rules stricter than mere presence in multiple databases). In this work, we restrict ourselves to the case of two parties, i.e., two LEAs. Our general approach is to use circuit-based PSI to securely compute the intersection of subjects occurring in both databases and invoking MPC to securely compute for which of these subjects the combined LE data known about them (given to the protocol as associated payload) warrants information exchange. Only the identities of these subjects are revealed, as only exchanging their information is lawful.

[48] Hladjk in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2nd Ed., 2018, Art. 32, sec. 7.

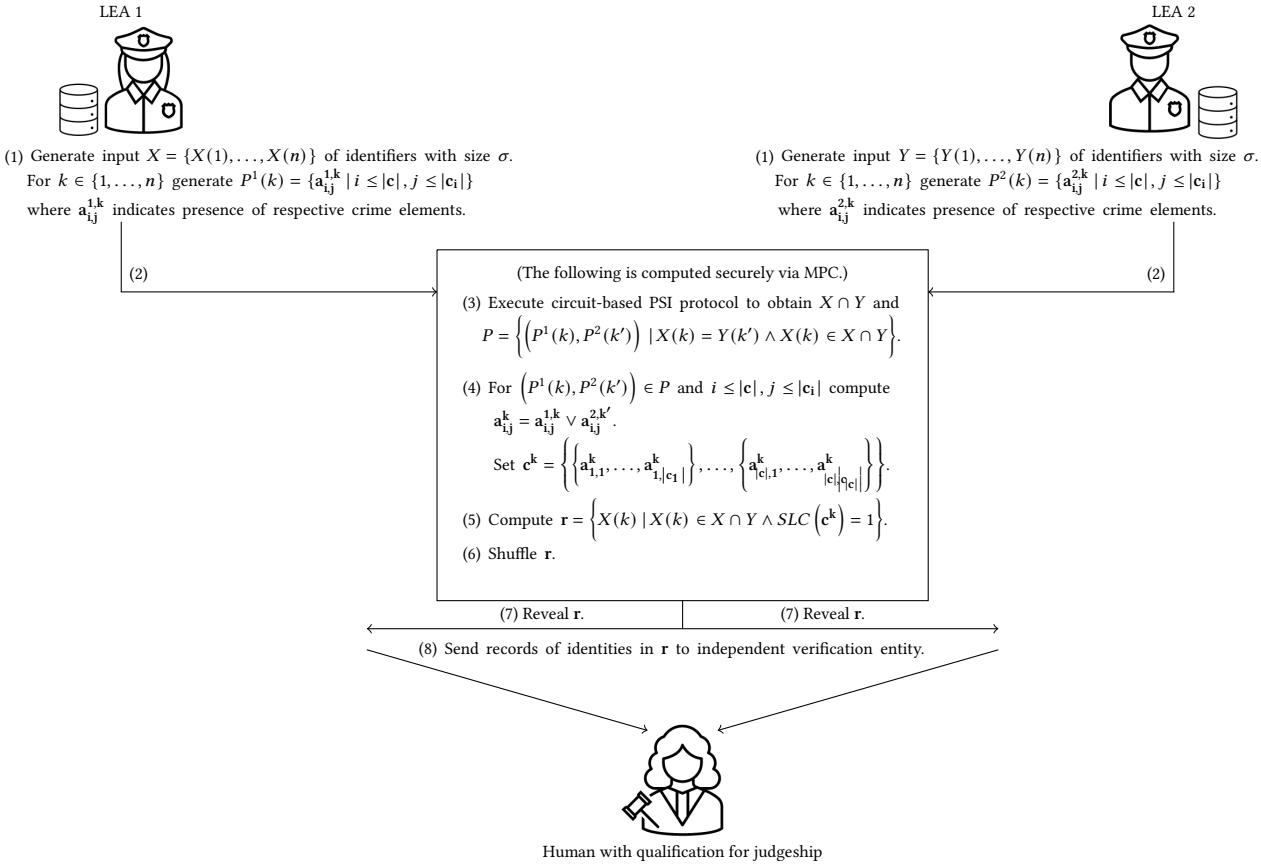

Figure 1: Overview of the architecture of our system. Two LEAs want to privately match their databases such that only lawfully exchangeable data is actually exchanged. This is based on a secure lawfulness computation algorithm $SLC(\mathbf{c})$ that checks lawfulness based on input (crime) elements given in form of a vector $\mathbf{c} = \{\mathbf{c_1}, \ldots, \mathbf{c_{|c|}}\}$ for each record.

5.1.1 Secure Lawfulness Computation. We require a well-specified algorithm whose binary output corresponds to whether laws regarding information exchange are satisfied, as these laws are given as legal texts and are subject to legal interpretation. This heavily depends on the jurisdiction and requires a manual transcription of laws into an algorithm. In this work, we conduct this on the exemplifying case of § 100a(2) StPO (German Code of Criminal Procedure), which specifies "serious crimes" that warrant surveillance/exchange of data. In their legal definitions, these serious crimes are made up of several elements, *all of which* have to be given to fulfill this crime (although one element might have several *alternative elements* where *only one of them* needs to be present). We conducted a legal analysis of the crimes listed in § 100a(2) StPO and transcribed all crimes and their respective elements into categories of *crimes* and *crime elements*. These are 215 crimes and 1 478 crime elements (1 010 of which are alternatives). Thus, our algorithm's output corresponds to § 100a(2) StPO compliance if for *any* crime, *all* of its crime elements (some of which consist of several alternatives where a *minimum of one* is required) have been recorded.

Formally, $SLC(\mathbf{c} = \{\mathbf{c_1}, \ldots, \mathbf{c_{|c|}}\})$, our secure lawfulness computation, takes as input vectors $\mathbf{c_i} = \{\mathbf{e_1}, \ldots, \mathbf{e_{|c_i|}}\}$, where each $\mathbf{e_j}$ indicates whether the j-th element of the i-th crime is present. For § 100a(2) StPO, this would be $|\mathbf{c}| = 215$, so the total amount of crime elements (including alternatives) $|\mathbf{e}| = \sum_{i \leq |\mathbf{c}|} \left(\sum_{j \leq |\mathbf{c_i}|} \left|\mathbf{a_{i,j}}\right|\right) = 1\,478$, where the bitvector $\mathbf{a_{i,j}}$ indicates the presence of each alternative element of the j-th element of the i-th crime. If no alternative exists, we set $\mathbf{a_{i,j}}$ to the single bit whether the crime element is present. It then computes $SLC(\mathbf{c}) = \bigvee_{i \leq |\mathbf{c}|} \left(\bigwedge_{j \leq |\mathbf{c_i}|} \left(\bigvee \mathbf{a}_{i,j}\right)\right)$.

We stress that our lawfulness computation here is an exemplifying case to obtain the general regulatory complexity of legal elements that define whether LE data can lawfully be exchanged, as in this work we are concerned with establishing feasibility. Other algorithms, e.g., for other jurisdictions than Germany or other elements (e.g., purchases of hazardous substances), are possible to be used in our system due to the black-box nature of MPC. In a deployed system, this algorithm would need further practical refinement and each crime element would need to be specified exactly.

5.1.2 Complete overview. Our system works according to the following steps, as visualized in Fig. 1. In step (1), both agencies prepare the input vectors $\mathbf{a_{i,j}}$ as specified in Sect. 5.1.1 for each subject in

their input sets X or Y. Subjects are represented as unique identifiers, e.g., via a hash value computed on the basis of names as well as date and place of birth. Then, in step (2), both parties execute an MPC protocol performing the following secure computation. At step (3), a circuit-based PSI protocol with associated payloads is invoked. The LEAs input the set of their subjects and as associated payload all $\mathbf{a_{i,j}}$ of the subject. The output is then used in step (4) to aggregate the associated payloads, i.e., which $\mathbf{a_{i,j}}$ occurs in either data. Then, in step (5), the LEAs securely compute *SLC* on the aggregated payloads, i.e., whether information exchange of the respective individual is lawful. This gives us our desired outputs, namely those elements where the output of step (5) is true. However, these cannot be revealed because their position would leak information about individuals not in the intersection. Thus, the identifiers of the matched subjects need to be obliviously shuffled in step (6). The shuffled identifiers are then revealed to the LEAs in step (7). As a final step (8), all data associated with these subjects are sent to an independent verification entity where a human with qualification for judgeship manually verifies for each hit whether the subjects actually match and whether the legal requirements are met to exchange this data. If yes, the original data for the subject in question can be sent to the appropriate LE entity for investigation.

The final step is motivated by the intention of Art. 22(1) of the GDPR, according to which measures that affect a person should be based on a human decision and cannot be based on a solely automatic decision. Our *SLC* can therefore be seen as a basic, privacy-preserving algorithm indicating that data exchange would be lawful and significantly lowering privacy infringements. To ensure complete lawfulness, a human with qualification for judgeship has the final decision power over whether full data can be exchanged. Technical solutions auditing matching criteria [29] could potentially also be used to reduce human effort.

5.2 Legal Analysis of Our System

The data protection assessment of a data comparison between LEAs is not governed by the GDPR, as the subject area of law enforcement is an area not covered by the scope of the GDPR (Art. 2(2)(d) GDPR). To this extent, however, the so-called Justice and Home Affairs (JHA) Directive (cf. Sect. 3.1) acts as the background under European law to the relevant national rules in this area and ensures minimum harmonization[49] under data protection law. Since national law is based on this directive, which is to be implemented in all EU member states, the considerations made in the following also have significance for the understanding of the legal situation concerning MPC solutions in all countries of the European Union.

German Norms. Unlike criminal prosecution, for which the federal government has regulatory competence in Germany (cf. Art. 74 Sec. 1 No. 1 GG), the comparison of data to prevent future criminal acts or "threats to public safety or order" is legally a matter for individual state legislation. The data protection treatment of MPC techniques for this purpose therefore arises for each of the 16 federal states from the respective police law. Given the novelty—at least from a legal perspective—of such techniques, there are no regulations that explicitly address data comparison using these methods. The attempt to fall back on existing norms faces various problems. E.g., the application of the standards as a legal and processing basis for data comparison with MPC in the application cases considered by us *fails* in part due to the existence of essential factual prerequisites, such as a *requirement for concrete suspicion in the individual case* by § 25a of the Hessian Law on Public Safety (HSOG). The application of other norms, e.g., §§ 20 ff. HSOG, would mean splitting up a contiguous data processing operation, which would violate the purpose of the data protection standards[50].

Processing Basis. Due to the above points, a central legal problem for LEA data comparison via MPC is first of all the lack of a data protection law processing basis, which would, however, be mandatory due to the principle of a processing prohibition with permission reservation (cf. Sect. 3.1). On the other hand, compliance with other processing requirements identical to the requirements of Article 5 of the GDPR (cf. Sect. 3.1), which are anchored in all police laws in view of their incorporation in the JHA Directive, does not in principle prevent the use of MPC techniques. In order to take advantage of the opportunities offered by MPC for the comparison of personal data between security authorities, *a corresponding basis for processing must be created by the legislature.* In view of the data protection-friendly design of the technology (cf. Sect. 3.2), *the creation of such a legal basis also appears constitutionally possible.*

(Low) Risks of Our Approach. Such a legal processing basis seems possible because the problems outlined in Sect. 4.1, which are addressed by case law in connection with data comparisons or databases in the context of state security legislation, do not arise in the context of MPC. In view of the decentralized data storage and the possibility of linking the exchange of data to content-related prerequisites via the secure lawfulness computation, there is neither the danger of profiling, nor of an exchange without the necessary prerequisites. This also includes linking the possibility of exchange to such high requirements, e.g., risk of committing other serious crimes, that any further use of the data that may occur for a different purpose is possible. This is in view of the fulfillment of the requirements for a hypothetical new collection of data (mentioned in Sect. 4.1) and thus the principle of proportionality is also safeguarded. The dangers associated with the establishment of cloud solutions are also not realized here.

Balance of Legal Interests. Since the dangers otherwise feared do not materialize when our system is used, the exchange is tied to high hurdles. The matching can take place without the possibility of third parties becoming aware of the information, to whom the data remain unknown as long as the conditions for detection do not exist, so the depth of infringement on the fundamental right to data protection or informational self-determination of those affected by the processing remains very low. Against this backdrop, using the technology is favorable, given the balance between the legal interests of the data subjects and the dangers to society as a whole from the use of this technology and the benefits in the form of the possible prevention of serious crime or security threats.

[49] Roßnagel in: Roßnagel (Ed.), Hessisches Datenschutz- und Informationsfreiheitsgesetz, 2021, Einleitung, Rn. 51.

[50] It is not for nothing that the definition of the processing of personal data also provides for the existence of interrelated series of operations; cf. Roßnagel in: Simitis/Hornung/Spiecker gen. Döhmann (Eds.), Datenschutzrecht, 2019, Art. 4 Nr. 2 DS-GVO, sec. 11; Herbst in: Kühling/Buchner (Eds.), DS-GVO/BDSG, 3rd Ed., 2020, Art. 4 Nr. 2, sec. 15.

Data Quality. Towards practice, sufficient data quality is an essential basic requirement for the use of a technology from a data protection perspective. One aspect here is that bias and discrimination already present in the data could be reinforced by algorithmic decisions on it.[51] The principle of accuracy of data (cf. only Art. 5(1)(d) GDPR) is not only a prerequisite for processing, but an *excessive* number of incorrect results also makes a processing method appear very intervention-intensive, since its use would be without any practical benefit that could justify data processing. Sufficient data quality thus is also a prerequisite for our solution.

5.3 Technical Analysis of Our System

As a proof of concept, we implement our system and analyze it.

5.3.1 Implementation Details. We use ABY [12] for semi-honest secure two-party computation in C++ with the Boolean GMW protocol [20] as it is well-suited for parallelization. For oblivious shuffling we use the implementation in ABY of [25] using Waksman permutation networks [51]. As the circuit-based PSI protocol, we use the implementation of [38] due to its linear complexity, open-source availability, and compatibility with the ABY framework for use within a broader secure computation architecture. While the protocol of [38] has since been improved upon [9, 42] and better performance results can be expected from them, this proof of concept study is mainly concerned with establishing the feasibility of our general approach (and we will see that the PSI is not a significant cost factor). We add associated payloads, which was not yet implemented. Additionally, these payloads are rather large in our system: We found 1 478 crime elements in § 100a(2) StPO (cf. Sect. 5.1), so for each subject (set element), the associated payload consists of 1 478 bits. We therefore also enable the case where the payload size is much larger than the size of the set elements, which we implement using GMP [16] for multi-precision arithmetic.

5.3.2 Experimental Setting. We run our experiments on two machines with 128 GB memory and Intel Core i9-7960X CPUs. To simulate LEAs separated by physical distance, we restrict the network between the machines to 100Mbit/s bandwidth and 100ms round-trip time. The results are the mean communication and computation times of both parties over 10 invocations.

Concerning the choice of PSI parameters, we stick to the ones used in [38]. The length of the identifiers σ is set to 40, which yields a false positive probability that can be analyzed via the birthday problem. Our choice here will introduce a probability of 50% for 2^{20} subjects that a false-positive match occurs, which given the analysis in Sect. 5.2 we think is appropriate given that a human with qualification for judgeship performs a final check before data exchange. To obtain a negligible false positive probability, one would need to set $\sigma \geq 40 + 2\log_2 n - 1$, where n is the number of individuals.

We vary two dimensions of input parameters: The number of elements (individuals) in the set n and the bit-length of the associated payloads, which is equivalent to the number of crime elements $|\mathbf{e}|$. For the latter, we consider a length of 61 bits[52], i.e., using only 61 (alternative) crime elements, or the full 1 478 bits requiring multi-precision arithmetic. Note here that for simplicity, we do not implement our precise transcript of the secure lawfulness computation according to § 100a(2) StPO (cf. Sect. 5.1). Instead we use the average of 7 alternatives, totalling 211 crimes, each with 7 alternative elements, for covering the whole complexity of § 100a(2) StPO. For 61-bit payloads, we consider 12 crimes, each with 5 alternative elements. This case could be used if only a selection of most serious crimes would be considered as a trade-off for less workload.

5.3.3 Result Analysis. Our results are given in Tab. 1. The circuit-based PSI protocol of [38] relies on hashing the inputs into a table and then creating and evaluating an oblivious programmable pseudorandom function (OPPRF) via polynomial interpolation. The outputs are then securely compared in MPC/ABY, and the remaining secure computation of our system builds on top of that. The runtime of the hashing step is several orders of magnitude smaller than any other measurement and we therefore ignore it here. Thus, we distinguish between two phases in our analysis: OPPRF and ABY.

Overall, the results show that our system is feasible: For databases with millions of individuals, expected runtimes are in the order of hours and expected communication is in the order of dozens to possibly hundreds of gigabytes. Given the resources put into public security and LE, we think costs of these magnitudes are acceptable because the other ways to facilitate LEA information exchange currently proposed may violate data protection requirements (cf. Sect. 4.3). Furthermore, the purpose of this system is not to allow for real-time data analysis but to feasibly enable—in principle—a lawful information exchange between LEAs.

Expectedly, the bulk of the costs is carried by the secure computation within ABY. The smallest costs stem from the PSI comparison circuit itself, whereas the shuffling and lawfulness computation are responsible for most of the ABY costs. Asymptotically, this is clear for the oblivious shuffle on the PSI outputs, as costs of $O(n \log n)$ are expected there, where n is the number of individuals. In comparison, the lawfulness computation only requires $O(n|\mathbf{e}|)$ gates in the circuit that is evaluated securely. It follows that shuffling is the main asymptotic performance drawback in our system. However, a look at Tab. 1 reveals that, in concrete terms, the secure lawfulness computation can actually heavily influence the costs: For 2^{16} individuals, the costs rise from about 2 to about 10 GB and from about 2 to about 27 min if the full 1 478-bit payloads are used in the lawfulness computation compared to only using 61-bit payloads. This can be explained by the fact that for the evaluated parameters, the concrete costs of the lawfulness computation greatly surpass those of the shuffling if the full amount of crime elements is used, i.e., if $|\mathbf{e}| = 1\,478$. For instance, for $n = 2^{16}$, the shuffling results in about 1.6 GB of communication, whereas the lawfulness computation results in about 0.3 GB for $|\mathbf{e}| = 61$ and 7.94 GB for $|\mathbf{e}| = 1\,478$, respectively. In the former case, shuffling dominates the costs, whereas lawfulness computation does so in the latter case. Increased lawfulness computation complexity also introduces noticeable overhead in the OPPRF computation time due to using multi-precision arithmetic, e.g., for 2^{16} individuals and $|\mathbf{e}| = 61$, the OPPRF time is 0.1 min but is increased to 13.52 min for $|\mathbf{e}| = 1\,478$. Thus, not just input set size but also the complexity of computing the lawfulness of a data exchange is an important factor influencing the workload of our system.

[51] Cf. the example at Fröhlich/Spiecker gen. Döhmann, Können Algorithmen diskriminieren?, On Matters Constitutional, 26.12.2018, https://verfassungsblog.de/koennen-algorithmen-diskriminieren/.

[52] This number stems from the output size used in the implementation of [38].

Table 1: Total communication and runtime of our private LEA data matching. The amount of individuals n and crime elements $|\mathbf{e}|$ (including alternative elements) are varied. Two phases are considered: The oblivious programmable pseudorandom function for the PSI of [38] (OPPRF) and the secure computation within the ABY framework [12] for MPC. We indicate that our machines ran out of memory via "-" and in those cases give a theoretical estimation (denoted via *) for the expected communication. We omit the PSI hashing step as its runtime is several orders of magnitude smaller than any other step.

Number of individuals n	2^{12} = 4 096		2^{16} = 65 536		2^{20} = 1 048 576	
Number of crime elements $\lvert\mathbf{e}\rvert$	61	1 478	61	1 478	61	1 478
Communication in	(MB)	(MB)	(GB)	(GB)	(GB)	(GB)
For OPPRF	1.17	17.89	0.02	0.29	0.30	1.30*
For ABY	115.05	595.07	2.27	9.94	43.30	162.93*
Timing in	(s)	(s)	(min)	(min)	(min)	(min)
For OPPRF	2.36	50.45	0.10	13.52	1.26	-
For ABY	15.51	60.78	2.16	27.33	39.72	-

5.4 Possible Extensions of Our System

Our system relies on circuit-based PSI, which usually performs exact matching. However, in the application scenario at hand one can expect a large amount of inexact information like, e.g., misspellings of names. Thus, incorporating *fuzzy* matching may be a promising idea. This could be done via fuzzy PSI [19, 49] or other techniques, e.g., securely computing the similarity via Bloom filters, as done in private record linkage of medical records [46].

Another aspect is that we confined ourselves to two LEAs. While this is already a large step forward (especially compared to the existing efforts described in Sect. 4.3 that may violate data protection), realistic deployments would want to involve more LEAs. As an instance, in Germany, information of the state LEAs of each of the 16 states and the federal LEA cannot be exchanged among each other. Thus, a system with 17 LEAs may be desirable here. For this, one could use multi-party circuit-based PSI [8]. However, increased workloads (or confining to smaller input sets) are expected here. Using [8] with 17 parties, we expect feasible communication of 58.46 GB for $n = 2^{12}$ and $|\mathbf{e}| = 1\,478$. But for $n = 2^{20}$, we expect communication of 3.92 TB and 16.37 TB for $|\mathbf{e}| = 61$ and $|\mathbf{e}| = 1\,478$, respectively. Here, again, the bottleneck is the secure computation. Future research, e.g., by outsourcing the computation (including the initial set intersection) to a small number of non-colluding computing parties via secret sharing may be of benefit here.

Lastly, a next step would be a more sophisticated analysis of distributed LEA databases, e.g., via privacy-preserving machine learning (PPML) [7, 55]. However, here one needs to consider that decisions made by algorithms need to be explainable and should be justified by a human according to Art. 22(1) of the GDPR, making this step non-trivial and even privacy-invasive, which defeats the goal of our work in protecting citizens' privacy.

5.5 LE Feedback on Our System

Given our legal and technical analyses, we presented our system and our analysis results within our exchange with the Hamburg police. In this section, we describe the perspective and feedback we got from them, starting with general feasibility before going into practical challenges that would lie ahead. In summary, our system is theoretically feasible and could be deployed with some further more precise legal and technical specification. However, issues of data quality need to be solved before deployment would be sensible.

5.5.1 Feasibility. The general feedback is that both developing and operating such a system is feasible. Though seen as a complex and resource-intense system, other existing and completed IT projects within LE are regarded as more complex and expensive.

Development. For development, a more concrete concept would need to be created, including a more well-founded lawfulness computation algorithm than our proof of concept, which mainly served to obtain a general sense of complexity for establishing lawfulness. The LE feedback is that development is feasible with this and in collaboration with legal and cryptography experts if the legality of the system is established. One aspect here is that MPC and PSI techniques are not standardized, requiring close involvement of cryptographic experts for the development. Thus, standardization processes, which have just now started [1, 2, 36], are crucial for the costs of developing such privacy-preserving systems. However, according to the LEA in Hamburg, this would not be the main challenge. Instead, due to the interjurisdictional nature of the system, the organizational coordination and agreement overhead within the structure of public services and LE in a federal system like Germany could become a significant factor that could present hurdles.

Operation. For operating the system, the necessary infrastructure that can handle the costs estimated in Sect. 5.3 and Sect. 5.4 could be obtained and the costs of running it would be feasible given the goal of the system, even though our evaluation covered smaller database sizes than expected. In the state of Hamburg alone, there are about 11 Million unique names registered in the LEA's database. However, for many records, no relevant crime elements may have occurred and thus these records would not be used as input—making it hard to quantify the actually expected input sizes. The police additionally stated that a more complex lawfulness computation algorithm, involving at least the full 1 478 crime elements, would be preferable and the resulting overhead tolerable. The system should

be operated between all states, i.e., costs in the order of dozen terabytes and runtimes of possibly days could be expected (cf. Sect. 5.4). Again, this is seen as tolerable given the goal of the system. It could run during low-demand times like at night, as other systems in their infrastructure do, and an output after some days is regarded as adequate for this application scenario.

5.5.2 Practical Challenges. While the subject of this work is feasibility, we obtained a lot of practical feedback and concerns that would need to be addressed for any possible deployment. This mainly concerns an aspect already mentioned within the LE perspective given in Sect. 4: *Currently, LE databases do not contain data clean enough to be used in our system (and other proposed systems).* A lot of data is just stored as strings (even birth dates) and data is frequently entered incorrectly or omitted (in many interactions, only names and no other personal data is obtained). Thus, it is often hard to match individuals due to missing and imprecise data, possibly only yielding no matches or too many false positive matches. Ensuring better data quality would help, but a wish here is also to use unique identifiers of persons like in other countries (it is legally debated in Germany whether this is possible for privacy reasons).

Furthermore, while the crime elements used in our system are already categorized in new police IT standards, these are not really used (a system where this needs to be entered from a large catalog of crime elements is seen as lacking usability). These are practical problems not possible to be addressed within our work here but are relevant to other systems and that are already in the process of being addressed (cf. Sect. 4's LE perspective). Therefore, the feedback is that these will be practical prerequisites before a system like ours could function in the real world.

6 CONCLUSIONS AND OUTLOOK

In this work, we propose to use Secure Multi-Party Computation (MPC) to satisfy data protection regulations with a focus on the application of information exchange between Law Enforcement Agencies (LEAs). We show that currently proposed systems for this lack data protection guarantees, so we propose a new system based on MPC and Private Set Intersection (PSI). In an interdisciplinary effort, we provide legal and, via a proof of concept implementation, technical analyses of our system. In a qualitative exchange with the LEA of the police of the state of Hamburg, Germany, we obtain feedback on the system and our analyses to move towards practicality. In conclusion, we establish the legal and technical feasibility of our system and identify further practical challenges ahead as well as the need for a concrete regulatory framework for a privacy-preserving and lawful data exchange between LEAs.

It is our hope that with our combination of legal and cryptographic techniques with feedback of a real-world LEA, we can establish that there are feasible, lawful, and privacy-preserving alternatives to the privacy-invasive solutions for LEA information exchange currently proposed. Ideally, moving towards our approach could strengthen trust in LE that has been weakened due to efforts for public security undermining encryption and privacy in general.

Associated Risks and Potential for Abuse. Still, we also have to consider that our system comes with associated risks. It opens up LEAs performing (secure) computations on functionally centralized data, which introduces abusive potential for *mission creep* (see, e.g., [35] on mission creep in US data-sharing "fusion centers"). Extensions that we specifically do not consider may hence be pushed for that perform privacy-invasive computations, like, e.g., including bias via machine learning. Despite us assuming the input data to our system has been lawfully collected, in practice the existence of our system may further *incentivize excessive or unlawful data collection* by LEAs. *Wrong inputs* or *collusion between the LEAs* may also break the system. We also understand that the existing input data and crafted algorithm may reflect discrimination such as racial bias and that, hence, our system may *perpetuate biases of LE in practice* to a cross-jurisdictional degree not possible before. Furthermore, the technology is used to process data relating to individuals who have committed no significant wrongdoing at the time of use. It also has a high "spread width"[53], so that there is a risk of *effects on society as a whole*. When designing our system, we thus consider and make clear that any such system and all used algorithms need to be verified independently and their lawfulness and *necessity* be justified. To mitigate these dangers, technical and organizational measures must also be taken to rule out misuse of the method as far as possible in the current state of affairs. The guarantee that the legal requirements for data processing, both at the level of simple and constitutional law, are complied with at all times is an indispensable prerequisite for this technology.

But because the mentioned associated risks could remain for actual deployments, a certain amount of resistance to the use of MPC in LE must be expected, particularly from the direction of data protection and civil rights. We believe such resistance is justified: We have established the principal legality of our system and shown that its risks are far less than current systems, but it remains open if its risks are preferable to other approaches that are more restrictive on data exchange. It ultimately must be up to society to make an *informed decision* if our approach is the better trade-off regarding public security. To enable this informed decision, the opportunities, *risks*, and, ideally, even technical foundations of the use of MPC techniques in the LE sector should be discussed widely in society before the necessary permission standard for the use of the technology would be created by the legislature.

ACKNOWLEDGMENTS

For great assistance in conducting the technical analysis of the law, we would like to thank Julia Mahnken. We thank our partners from LE for our joint collaboration and their continuous feedback to our ideas and analyses: The Polizei Hamburg and their Netzwerk Digitale Polizei (NetDigPol) of the Akademie der Polizei Hamburg, particularly Prof. Eike Richter and Fee Weinberger, for coordination of the collaboration and general as well as practical feedback, and Programm Polizei 2020, particularly Nadja Oppermann and Karsten Büttner, for practical feedback. This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 850990 PSOTI). It was co-funded by the Deutsche Forschungsgemeinschaft (DFG) within SFB 1119 CROSSING/236615297 and GRK 2050 Privacy & Trust/251805230, and by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within ATHENE.

[53] In the sense that the measure affects many people not relevant to it.

REFERENCES

[1] ISO/IEC JTC 1/SC 27 Information security, cybersecurity and privacy protection. 2022. *ISO/IEC CD 4922-2 Information security — Secure multiparty computation — Part 2: Mechanisms based on secret sharing.* International Organization for Standardization.

[2] ISO/IEC JTC 1/SC 27 Information security, cybersecurity and privacy protection. 2022. *ISO/IEC DIS 4922-1 Information security — Secure multiparty computation — Part 1: General.* International Organization for Standardization.

[3] Gilad Asharov, Yehuda Lindell, Thomas Schneider, and Michael Zohner. 2013. More efficient oblivious transfer and extensions for faster secure computation. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.

[4] Kenneth A Bamberger, Ran Canetti, Shafi Goldwasser, Rebecca Wexler, and Evan J Zimmerman. 2022. Verification Dilemmas in Law and the Promise of Zero-Knowledge Proofs. In *Berkeley Technology Law Journal*, Vol. 37.

[5] Francesca Bignami. 2007. Privacy and law enforcement in the European union: the data retention directive. In *Chicago Journal of International Law*, Vol. 8.

[6] Dor Bitan, Ran Canetti, Shafi Goldwasser, and Rebecca Wexler. 2022. Using Zero-Knowledge to Reconcile Law Enforcement Secrecy and Fair Trial Rights in Criminal Cases. In *SSRN Preprint 4074315.* Available at https://ssrn.com/abstract=4074315.

[7] José Cabrero-Holgueras and Sergio Pastrana. 2021. SoK: Privacy-Preserving Computation Techniques for Deep Learning. In *Proceedings on Privacy Enhancing Technologies (PoPETs)*, Vol. 4.

[8] Nishanth Chandran, Nishka Dasgupta, Divya Gupta, Sai Lakshmi Bhavana Obbattu, Sruthi Sekar, and Akash Shah. 2021. Efficient Linear Multiparty PSI and Extensions to Circuit/Quorum PSI. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.

[9] Nishanth Chandran, Divya Gupta, and Akash Shah. 2022. Circuit-PSI with linear complexity via relaxed batch OPPRF. In *Proceedings on Privacy Enhancing Technologies (PoPETs)*.

[10] Michele Ciampi and Claudio Orlandi. 2018. Combining private set-intersection with secure two-party computation. In *International Conference on Security and Cryptography for Networks (SCN)*.

[11] Emiliano De Cristofaro and Gene Tsudik. 2010. Practical private set intersection protocols with linear complexity. In *International Conference on Financial Cryptography and Data Security (FC)*.

[12] Daniel Demmler, Thomas Schneider, and Michael Zohner. 2015. ABY-A framework for efficient mixed-protocol secure two-party computation.. In *Network and Distributed System Security Symposium (NDSS)*.

[13] Joan Feigenbaum. 2017. Multiple objectives of lawful-surveillance protocols (transcript of discussion). In *Cambridge International Workshop on Security Protocols*.

[14] Joan Feigenbaum. 2019. Encryption and surveillance. In *Communications of the ACM*, Vol. 62.

[15] Joan Feigenbaum and Daniel J Weitzner. 2018. On the incommensurability of laws and technical mechanisms: Or, what cryptography can't do. In *Cambridge International Workshop on Security Protocols*.

[16] Free Software Foundation. 1991. The GNU Multiple Precision Arithmetic Library. https://gmplib.org/.

[17] Jonathan Frankle, Sunoo Park, Daniel Shaar, Shafi Goldwasser, and Daniel Weitzner. 2018. Practical Accountability of Secret Processes. In *USENIX Security Symposium (USENIX Security)*.

[18] Michael J Freedman, Kobbi Nissim, and Benny Pinkas. 2004. Efficient private matching and set intersection. In *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*.

[19] Gayathri Garimella, Mike Rosulek, and Jaspal Singh. 2022. Structure-Aware Private Set Intersection, With Applications to Fuzzy Matching. In *Annual International Cryptology Conference (CRYPTO)*.

[20] Oded Goldreich, Silvio Micali, and Avi Wigderson. 1987. How to play any mental game. In *ACM Symposium on Theory of Computing (STOC)*.

[21] Shafi Goldwasser and Sunoo Park. 2017. Public accountability vs. secret laws: can they coexist? A cryptographic proposal. In *Workshop on Privacy in the Electronic Society (WPES)*.

[22] Matthew Green, Gabriel Kaptchuk, and Gijs Van Laer. 2021. Abuse resistant law enforcement access systems. In *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*.

[23] Lukas Helminger and Christian Rechberger. 2022. Multi-Party Computation in the GDPR. In *Privacy Symposium - Data Protection Law International Convergence and Compliance with Innovative Technologies (DPLICIT)*.

[24] Brett Hemenway Falk, Daniel Noble, and Rafail Ostrovsky. 2019. Private set intersection with linear communication from general assumptions. In *Workshop on Privacy in the Electronic Society (WPES)*.

[25] Yan Huang, David Evans, and Jonathan Katz. 2012. Private set intersection: Are garbled circuits better than custom protocols?. In *Network and Distributed Systems Security Symposium (NDSS)*.

[26] Bundesministerium des Innern. 2018. Polizei 2020 - White Paper -. Available at https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2018/polizei-2020-white-paper.html.

[27] Yuval Ishai, Joe Kilian, Kobbi Nissim, and Erez Petrank. 2003. Extending oblivious transfers efficiently. In *Annual International Cryptology Conference (CRYPTO)*.

[28] Seny Kamara. 2014. Restructuring the NSA metadata program. In *International Conference on Financial Cryptography and Data Security (FC)*.

[29] Murat Kantarcioğlu and Chris Clifton. 2003. Assuring privacy when big brother is watching. In *ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD)*.

[30] Vladimir Kolesnikov, Ranjit Kumaresan, Mike Rosulek, and Ni Trieu. 2016. Efficient batched oblivious PRF with applications to private set intersection. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*.

[31] Joshua Kroll, Edward Felten, and Dan Boneh. 2014. Secure protocols for accountable warrant execution. In *White Paper.* Available at https://www.cs.princeton.edu/felten/warrant-paper.pdf.

[32] Tobias Kussel, Torben Brenner, Galina Tremper, Josef Schepers, Martin Lablans, and Kay Hamacher. 2022. Record Linkage based Patient Intersection Cardinality for Rare Disease Studies using Mainzelliste and Secure Multi-Party Computation. In *Research Square Preprint 1486673.* Available at https://europepmc.org/article/ppr/ppr476493.

[33] Yehuda Lindell and Benny Pinkas. 2009. A proof of security of Yao's protocol for two-party computation. In *Journal of Cryptology (JoC)*, Vol. 22.

[34] Catherine Meadows. 1986. A more efficient cryptographic matchmaking protocol for use in the absence of a continuously available third party. In *IEEE Symposium on Security and Privacy (S&P)*.

[35] Torin Monahan. 2009. The murky world of 'Fusion Centres' Torin Monahan critiques the emergence of data-sharing 'Fusion Centres' intended to reduce crime and prevent terrorism. In *Criminal Justice Matters*, Vol. 75.

[36] NIST. 2021. Toward a PEC use-case suite (preliminary draft). NIST White Paper (Draft). Available at https://csrc.nist.gov/publications/detail/white-paper/2021/01/21/toward-a-pec-use-case-suite-preliminary-draft/draft.

[37] Benny Pinkas, Thomas Schneider, Gil Segev, and Michael Zohner. 2015. Phasing: Private set intersection using permutation-based hashing. In *USENIX Security Symposium (USENIX Security)*.

[38] Benny Pinkas, Thomas Schneider, Oleksandr Tkachenko, and Avishay Yanai. 2019. Efficient circuit-based PSI with linear communication. In *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*.

[39] Benny Pinkas, Thomas Schneider, Christian Weinert, and Udi Wieder. 2018. Efficient circuit-based PSI via cuckoo hashing. In *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*.

[40] Benny Pinkas, Thomas Schneider, and Michael Zohner. 2014. Faster private set intersection based on OT extension. In *USENIX Security Symposium (USENIX Security)*.

[41] Benny Pinkas, Thomas Schneider, and Michael Zohner. 2018. Scalable private set intersection based on OT extension. In *ACM Transactions on Privacy and Security (TOPS)*, Vol. 21.

[42] Peter Rindal and Phillipp Schoppmann. 2021. VOLE-PSI: Fast OPRF and circuit-PSI from vector-OLE. In *International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*.

[43] James Scheibner, Jean Louis Raisaro, Juan Ramón Troncoso-Pastoriza, Marcello Ienca, Jacques Fellay, Effy Vayena, Jean-Pierre Hubaux, et al. 2021. Revolutionizing medical data sharing using advanced privacy-enhancing technologies: Technical, legal, and ethical synthesis. In *Journal of Medical Internet Research (JMIR)*, Vol. 23.

[44] Aaron Segal, Joan Feigenbaum, and Bryan Ford. 2016. Privacy-Preserving Lawful Contact Chaining: [Preliminary Report]. In *Workshop on Privacy in the Electronic Society (WPES)*.

[45] Aaron Segal, Bryan Ford, and Joan Feigenbaum. 2014. Catching Bandits and Only Bandits: Privacy-Preserving Intersection Warrants for Lawful Surveillance. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*.

[46] Sebastian Stammler, Tobias Kussel, Phillipp Schoppmann, Florian Stampe, Galina Tremper, Stefan Katzenbeisser, Kay Hamacher, and Martin Lablans. 2022. Mainzelliste SecureEpiLinker (MainSEL): Privacy-preserving record linkage using secure multi-party computation. In *Bioinformatics*, Vol. 38.

[47] Council of the European Union. 2019. Information Management Strategy (IMS) action No 2 – Action EPRIS-ADEP - final evaluation report – ADEP Technology - Services and Applications. Note 7886/19.

[48] Volker Ullrich, Fritz Felgentreu, Stefan Keuter, Benjamin Strasser, Martina Renner, and Irene Mihalic. 2021. *Beschlussempfehlung und Bericht des 1. Untersuchungsausschusses der 19. Wahlperiode gemäß Artikel 44 des Grundgesetzes.* Deutscher Bundestag. Available at https://dip.bundestag.de/drucksache/beschlussempfehlung-und-bericht-des-1-untersuchungsausschusses-der-19-wahlperiode-gem%C3%A4%C3%9F/255728.

[49] Erkam Uzun, Simon P Chung, Vladimir Kolesnikov, Alexandra Boldyreva, and Wenke Lee. 2021. Fuzzy Labeled Private Set Intersection with Applications to Private Real-Time Biometric Search. In *USENIX Security Symposium (USENIX Security)*.

[50] Dinusha Vatsalan, Ziad Sehili, Peter Christen, and Erhard Rahm. 2017. Privacy-preserving record linkage for big data: Current approaches and research challenges. In *Handbook of Big Data Technologies*.

[51] Abraham Waksman. 1968. A permutation network. In *Journal of the ACM (JACM)*, Vol. 15.
[52] Charles V Wright and Mayank Varia. 2018. Crypto crumple zones: Enabling limited access without mass surveillance. In *IEEE European Symposium on Security and Privacy (EuroS&P)*.
[53] Andrew Chi-Chih Yao. 1982. Protocols for secure computations. In *Symposium on Foundations of Computer Science (SFCS)*.
[54] Andrew Chi-Chih Yao. 1986. How to generate and exchange secrets. In *Symposium on Foundations of Computer Science (SFCS)*.
[55] Qiao Zhang, Chunsheng Xin, and Hongyi Wu. 2021. Privacy-Preserving Deep Learning Based on Multiparty Secure Computation: A Survey. In *IEEE Internet of Things Journal*, Vol. 8.