SMARTEN—A Sample-Based Approach towards Privacy-Friendly Data Refinement

Subjects: Computer Science, Information Systems

Contributor: Christoph Stach, Michael Behringer, Julia Bräcker, Clémentine Gritti, Bernhard Mitschang

SMARTEN applies a revised data refinement process that fully involves domain experts in data pre-processing but does not expose any sensitive data to them or any other third-party.

Keywords: privacy ; data refinement ; data cleansing ; data transformation ; human-in-the-loop

1. Data Preparation Phase

Due to the IoT and decreasing storage costs, data are constantly being captured and stored persistently. Data processors and especially data stewards are therefore faced with a flood of raw data that they cannot handle manually. Therefore, there are approaches towards data pre-processing powered by artificial intelligence. However, it cannot substitute human knowledge in data preparation. Human data wrangling experts, such as data stewards, achieve better results in data preparation because they can rely on broad domain knowledge and their experience. In the data preparation phase, it must be ensured to keep the human in the loop [1].

Researchers therefore adopt a dynamic sampling approach to meet this premise. Here, the data steward operates on a representative yet manageable sample which is extracted from the bulk of raw data. In doing so, the data steward defines data preparation rules that can be generalized and applied to the entire base data. While such an approach is primarily intended to relieve the experts, it also represents a privacy preserving measure [2]. Since the data steward never has access to the entire dataset, he or she can only gain as much insight as required for the fulfillment of his or her task. In addition, particularly sensitive data can be concealed entirely. However, by applying afterwards the defined rules to the entire base data, a comprehensive data preparation can still be ensured. **Figure 1** illustrates the workflow that is followed in the approach.

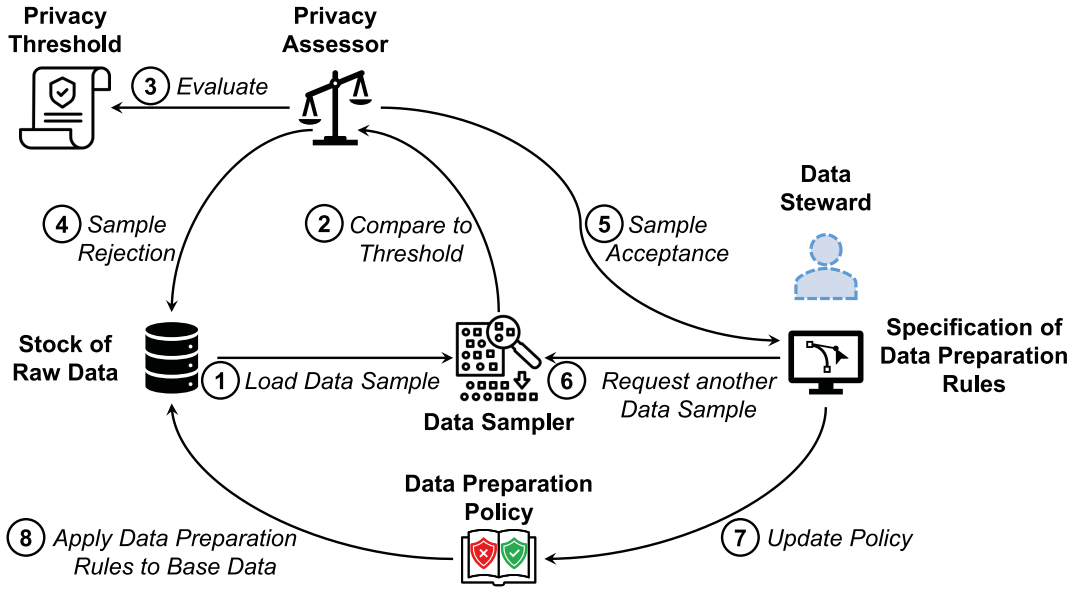

Figure 1. Workflow of the sample-based approach for data preparation applied in SMARTEN.

① Initially, the *data sampler* picks a representative data sample from the stock of raw data. Descriptive analyses are performed on the base data to identify data items that require special attention, e.g., null values, extreme values, and outliers. Metadata provided by the data producers can further facilitate the sample selection, for instance, to identify attributes that are either particularly relevant for data preparation or have little or no influence on data refinement. ② If a

reasonable sample is found, the data sampler forwards it to the *privacy assessor* ③. The privacy assessor evaluates whether the sample meets the necessary privacy requirements. These are specified by the data producer by means of a *privacy threshold*.

Different metrics can be used for this purpose. For instance, *k-anonymity* describes how well an individual can disappear in a crowd. The *k* indicates that there are at least k−1 other individuals (respectively, data items) with similar attribute values, so that they cannot be distinguished unambiguously. Thus, information obtained from the data cannot be associated with a single individual, but only with a group of size *k*. Another metric is the *entropy* which evaluates the information content of a dataset. In simple terms, it assesses how much a data item deviates from the expected norm. For instance, if it is known that the value of an attribute A is almost in all cases *x*, then a data item where A has the value *y* has a particularly high information content. For more information on these and other technical privacy metrics that can be used to evaluate the samples, please refer to literature, e.g., the work by Wagner and Eckhoff [3]. With regard to the quality of the data refinement, however, the principle of fairness must also be considered in the sampling. For instance, no bias must be introduced due to the sampling [4]. For instance, if the value of the attribute A is predominantly *x* in the sample, while the attribute values *x* and *y* are equally distributed in the base data, the data steward might estimate the significance A incorrectly. This would have a negative impact on the quality of the outcome of the data preparation phase. There are also a wide range of technical fairness metrics that can be used in the privacy assessor of SMARTEN. For details on such metrics, please refer to literature, e.g., the work by Lässig et al. [5].

④ If the privacy assessor rejects the sample because it violates the privacy threshold, the data sampler must prepare a new sample and the assessment cycle restarts. If no sample can be found that satisfies the privacy threshold, e.g., since the threshold is too restrictive or the base data contain too sensitive information, automated data pre-processing can be applied. Here, an artificial intelligence approach performs both the data cleansing [6] and the data transformation [7]. However, the quality of the data preparation is expected to be impaired, as the data steward's domain knowledge cannot be taken into account [8]. In addition, such fully automated data processing is not permissible in certain application areas, e.g., when dealing with medical data [9]. Alternatively, synthetic data with characteristics similar to those of the real base data can be generated. Such an approach is used for data refinement when only very little base data are available [10]. The data steward can define preparation rules for these synthetic data, without having access to the real base data. However, this also impairs quality of data preparation, as *overfitting*—i.e., the generated data imply that certain issues are more significant than they are—or *underfitting*—i.e., certain data issues are not reflected in the generated data at all—may occur in the synthetic data.

⑤ If a valid sample is found, it is forwarded to the data steward. Only at this stage does a human gain insight into the base data excerpt. The data steward identifies data issues in the sample and specifies which data cleansing steps and transformation tasks are necessary as part of the data preparation. ⑥ If he or she needs further insights into the base data, he or she can request another sample from the data sampler until he or she is satisfied with the data quality. He or she can also retrieve statistical information about the base data, such as minima, maxima, or average values, e.g., to fill missing attribute values. ⑦ These data cleansing steps and transformation tasks for the sample can then be mapped to three general transformation operators, namely *filter operators*, *map operators*, and *reduce operators*. A filter operator filters out data items from the base data for which certain properties hold, a map operator transforms data items or modifies certain attributes, while a reduce operator groups and aggregates data items. ⑧ These generalized transformation operators can then be applied to the entire base data [11].

The approach therefore combines a sample-based, an expert-based, and a rule-based approach. This way, the best possible data preparation quality can be achieved [12]. In technical terms, selection, projection, and aggregation, are used to ensure privacy. In sampling, selection is primarily used to reduce the number of data items. Projection can enhance privacy, e.g., by removing identifying attributes. When specifying the preparation rules, the data steward can use aggregations to deal with data issues such as missing attribute values.

2. Data Processing Phase

Researchers designed a new data management architecture for the data processing phase—or, more precisely, researchers added privacy features to an existing data management concept suitable for the data refinement process. To this end, researchers leverage the *zone architecture for data lakes* introduced by Sharma [13]. Data lakes are an architecture for managing big data. For this purpose, the data are stored not only as raw format, but also in various processing stages. In a zone architecture, data in the same processing stage are organized in a common zone. Apart from zones in which data are stored persistently, there are also zones in which data are only kept in a volatile form, e.g., in order to transform them and forward them to downstream zones. Data consumers are granted access to one or more data

zones via dedicated interfaces according to predefined access policies. In SMARTEN, researchers extend this generic core concept by additional privacy features. The resulting SMARTEN architecture is shown in **Figure 2**.

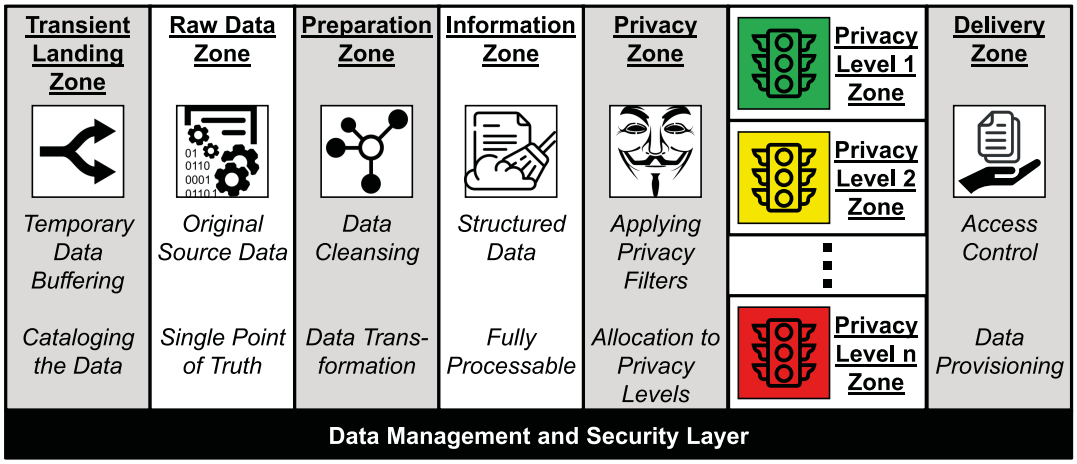

Figure 2. Extended date lake zone architecture to ensure privacy in the data processing phase.

This zone architecture reflects the data refinement process. The zones depicted in gray are processing zones in which data are only temporarily held, whereas the zones depicted in white are storage zones in which data are kept persistently. There are storage zones for raw data, information, and knowledge.

Raw data enter the data processor's scope via the *Transient Landing Zone*. This zone is a temporary buffer to be able to enrich the data with metadata relevant for processing, as well as to build access structures, such as data catalogs that facilitate the handling of the data. Then, these enriched data are stored in the *Raw Data Zone*. The SMARTEN workflow presented in **Figure 1** operates on this zone in the data preparation phase. The specified data preparation rules are applied in the *Preparation Zone*. For this purpose, a processing engine reads the raw data affected by the rules into a processable data structure and applies the required transformations, namely a sequence of filter operators, map operators, and reduce operators. This zone represents a key extension to the original zone architecture.

The outcomes of the data preparation phase are stored in the *Information Zone*. The data items present in this zone are therefore in a fully processable structure. Thus, these data items can be made more privacy-friendly in a target-oriented manner by the *Privacy Zone*. This zone represents another second key extension, since it initiates the data processing phase in SMARTEN. In this zone, a repository is maintained in which privacy scripts are stored that implement different dedicated privacy techniques for specific types of data. The appropriate scripts are selected by the Privacy Zone and applied to the prepared data [14].

The concealed data are then transformed into knowledge by the data analyst. However, since different data consumers may have different privacy requirements, the data processor generates *n* different versions of the knowledge rather than just one. For each of these variants, a different combination of privacy scripts has been applied prior to processing. Thus, they represent different levels of privacy. For each level, there is a zone labeled *Privacy Level x Zone*. As a result, there is an appropriate variant of the knowledge for any purpose [15]. To determine the appropriate number of such privacy levels, privacy experts first analyze the available data sources and assess the potential threats posed to data subjects, i.e., the severity of the impact that a disclosure would entail. Mindermann et al. [16] demonstrate how a method called *STPA-Priv* [17], which is based on *System Theoretic Process Analysis*, can be used for this purpose. Here, experts are systematically guided by tools to uncover existing data privacy risks. The experts can document their findings regarding the privacy risks in a processable model that describes which knowledge can be derived from which data sources [18]. The derivable knowledge is thereby represented as knowledge patterns. Additionally, explanatory keywords are assigned to each of these patterns. By means of these keywords, data subjects can easily identify exactly those patterns that are relevant for them, i.e., patterns that have to be concealed in their data. In addition, a *collaborative filtering* approach can be used to recommend further patterns to data subjects that might be relevant to them as well [19]. These patterns are then translated into non-disclosure requirements, which are used for the configuration of SMARTEN.

The *Delivery Zone* regulates which data consumer has access to which privacy level. For each access, this zone checks which knowledge may be disclosed and shares only that variant with the data consumer. Third parties are not aware of any other variants. An access control policy defines which data consumer is allowed to access which privacy level. Please refer to literature for details on how such an access control policy can be designed, e.g., the work by Stach and Mitschang [20].

Orthogonal to these processing and storage zones there is a *Data Management and Security Layer*. Such a layer is needed in every data lake for data governance tasks, e.g., to assist data retrieval, to support efficient data access, or to implement data security measures. In SMARTEN, this layer has another relevant function. Permanent storage of the raw data and all their processing stages is necessary so that the data processor can fulfill its obligation to prove to the data producer that, e.g., the requested privacy techniques have been applied prior to processing. For this purpose, *Provable Data Possession* can be applied. Here, a data producer can verify whether the data are at rest in the respective privacy level. In the approach, a third-party auditor, e.g., a data protection authority, can perform this task on behalf of the data producer. In doing so, the third-party auditor requires neither insight into the data of the data processor nor into the data of the data producer [21].

To this end, the Data Management and Security Layer encodes and encrypts all data at rest in SMARTEN. The third-party auditor challenges the data processor regularly to ensure that the data are present in all required privacy levels. For this purpose, there is a puzzle mechanism that ensures that solving a puzzle always takes a certain amount of time. This puzzle is related to the encrypted data at rest. If the data processor takes too long to provide an answer to a challenge, it implies that he or she needs to generate a solution to the puzzle. Therefore, it can be assumed that the data processor has to generate a version of the data in the intended privacy level on the fly during the challenge. However, if the data processor responds quickly to the challenge, it means that he or she already knows the solution to the puzzle, i.e., the data are at rest in the requested privacy level [22].

3. Elicitation of Non-Disclosure Requirements

For the configuration of SMARTEN, i.e., the specification of privacy thresholds and privacy techniques that have to be applied, researchers adopt a two-stage approach in order to address the two separate phases of the data refinement process. For this purpose, researchers rely on digital signatures that can be inseparably attached to the data. These signatures specify the non-disclosure requirements of the data producers. Such an approach ensures that an attacker cannot alter the requirements retrospectively in order to gain deeper insights into the data of the data producer [23].

Digital signatures are used to prove the authenticity of a data object, e.g., a message, a file, or a data item. For this purpose, asymmetric cryptography is used, i.e., a key pair is generated that consists of a *signing key* and a *verification key*. The signing key is kept secret by the signer, while the verification key is shared with the public. To sign a data object, the signer encrypts it with his or her signing key. The data object can only be verified with the corresponding verification key. If the verification is successful, it is ensured on the one hand that the data object has not been tampered with and on the other hand that it originates from the alleged source, since only this source has access to that signing key. An *attribute-based signature* uses keys that are composed of a set of attributes. A signed object can only be successfully verified if the key used for this purpose has a certain subset of these attributes. In SMARTEN, researchers also use attribute-based signatures, except that the non-disclosure requirements of the data producers are used instead of attributes. However, attribute-based encryption algorithms require a lot of computational power. Therefore, researchers adopt an approach introduced by Gritti et al. [24], in which the computations are distributed and can be partially outsourced. The adapted approach is shown in **Figure 3**.

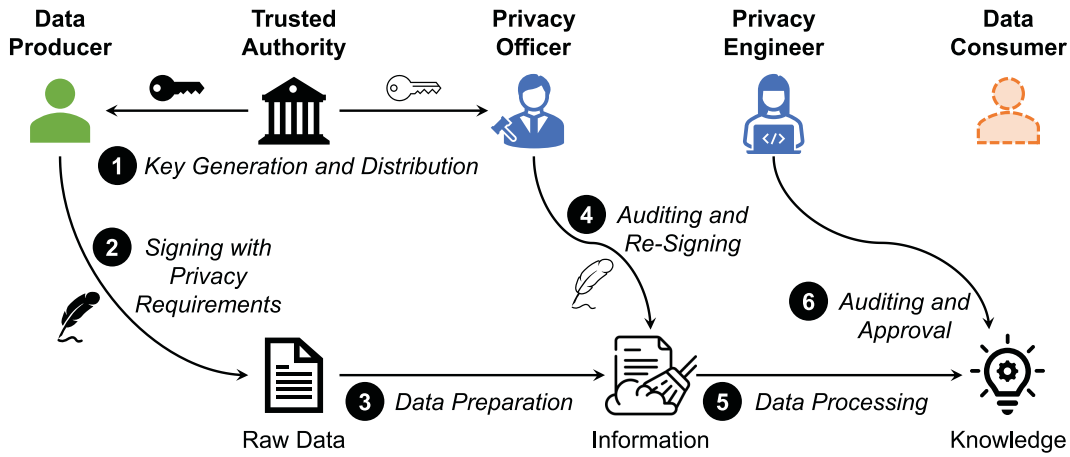

Figure 3. Two-stage approach to specify non-disclosure requirements applied in SMARTEN.

In addition to the actual two stages that are used to specify the privacy thresholds in the data preparation phase and the privacy requirements in the data processing phase, the approach also includes a preliminary phase for organizational purposes.

- **Key Generation and Deployment.** Initially, a data producer announces his or her non-disclosure requirements to a trusted authority. When dealing with sensitive data, regulations such as the GDPR mandate that there has to be an independent supervisory authority (Art. 51 and Art. 52). This can, for instance, serve as the trusted authority. ❶ This authority then generates two key pairs for the signature. Let τ be the set of privacy thresholds and λ be the set of requested privacy levels for the knowledge gained from processing the data. Then, *full keys* (depicted in black) reflect the union of these two sets $\varphi=\tau \cup \lambda$, while *delegated keys* (depicted in white) reflect only λ, i.e., a true subset of φ. The full keys are provided to the data producer, while the delegated keys are provided to the data processor.

- **Full Authentication.** ❷ To ensure that the data are not tampered with and that the non-disclosure requirements are not lost during transmission, the data producer signs the raw data with his or her full key. ❸ In the data preparation phase, the data processor verifies the signature against his or her privacy policy $\rho 1$. This policy describes which privacy thresholds are applied by the privacy assessor. Only if the requirements in τ are satisfied by $\rho 1$ is the applied privacy policy valid and the raw data in question can be preprocessed. ❹ This is monitored on behalf of the data processors by their privacy officer. A privacy officer requires no technical knowledge, as the privacy thresholds sufficiently specify how the data may be processed. If the data preparation is executed in compliance with the non-disclosure requirements, the privacy officer re-signs the data with the delegated key, i.e., τ is removed from the signature. In the data processing phase, the thresholds contained in τ are no longer relevant. However, non-disclosure requirements might indicate what a data producer wants to conceal. Thus, this filtering is necessary due to data minimization [25].

- **Delegated Authentication.** ❺ The re-signing initiates the data processing phase. In this phase, the data processor verifies the modified signature against his or her privacy policy $\rho 2$. This policy describes the privacy measures for which privacy scripts are available in the Privacy Zone. Only if the scripts comply with the requirements described in the signature can the prepared data be further processed. ❻ However, as such an auditing is by no means trivial, it has to be handled by a privacy engineer. A privacy engineer represents an intermediary between legal experts and IT experts. He or she is able to evaluate the means by which the non-disclosure requirements of the data producers can be met without rendering the quality of the processing results useless. Only if the privacy engineer approves the applied measures is the gained knowledge offered in the respective Privacy Level *x* Zone to data consumers. Since a semi-honest-but-curious data processor can be assumed, this approach is a reliable way to enforce the non-disclosure requirements of the data producers.

As always with cryptographic approaches of this kind, security stands and falls with the authority that issues the keys. If this authority is malicious or becomes compromised, no assurances can be made regarding the security of the signatures. In this case, all keys issued by this authority would have to be invalidated and replaced. However, assuming that the data processor is at least semi-honest-but-curious and therefore complies with applicable law such as the GDPR, the associated authority can also be trusted. Since this supervisory authority has to be independent according to GDPR, no involved party has any influence on it. Thus, even if the data processor is assumed to be somehow compromised, the key authority remains trustworthy. Furthermore, government agencies, such as the German Federal Office for Information Security (see https://www.bsi.bund.de/EN/Home/home_node.html, accessed on 31 July 2022), can also provide such key generation services, i.e., operate as such this trusted authority. However, if no authority can be found that all parties involved can fully trust, it is also possible to have several authorities responsible for key generation and key escrow instead of a single central authority. In this case, if some of these authorities become compromised, the security of the keys would not be affected, as long as not all of them are exposed at the same time. Such a distributed approach could be achieved for instance by means of *multi-party computation* [26] or *secret sharing* [27].

Protective measures must also be taken with regard to the keys themselves. Only if the methods used to generate them are secure are the digital signatures generated with them trustworthy. For the purpose in SMARTEN, researchers recommend the use of an *elliptic curve cryptography*, with which researchers aim for 224-bit/256-bit keys, giving researchers a minimum strength of 112 bits/128 bits of security [28]. Technological progress, such as quantum computing, may in the future render this encryption method, which is considered sufficiently secure today, ineffective [29]. In this case, *post-quantum cryptography* [30] would have to be used for digital signatures in SMARTEN. However, this would not affect the underlying concepts and procedures presented, since they are inherently technology-independent.

References

1. Zagalsky, A.; Te'eni, D.; Yahav, I.; Schwartz, D.G.; Silverman, G.; Cohen, D.; Mann, Y.; Lewinsky, D. The Design of Reciprocal Learning Between Human and Artificial Intelligence. Proc. ACM Hum.-Comput. Interact. 2021, 5, 443.

2. Arcolezi, H.H.; Couchot, J.F.; Al Bouna, B.; Xiao, X. Random Sampling Plus Fake Data: Multidimensional Frequency Estimates With Local Differential Privacy. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM), Gold Coast, QLD, Australia, 1–5 November 2021; ACM: New York, NY, USA, 2021; pp. 47–57.
3. Wagner, I.; Eckhoff, D. Technical Privacy Metrics: A Systematic Survey. ACM Comput. Surv. 2018, 51, 57.
4. Oppold, S.; Herschel, M. A System Framework for Personalized and Transparent Data-Driven Decisions. In Proceedings of the 32nd International Conference on Advanced Information Systems Engineering (CAiSE), Grenoble, France, 8–12 June 2020; Springer: Cham, Switzerland, 2020; pp. 153–168.
5. Lässig, N.; Oppold, S.; Herschel, M. Metrics and Algorithms for Locally Fair and Accurate Classifications using Ensembles. Datenbank Spektrum 2022, 22, 23–43.
6. Gemp, I.; Theocharous, G.; Ghavamzadeh, M. Automated Data Cleansing through Meta-Learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI), San Francisco, CA, USA, 4–9 February 2017; AAAI Press: Palo Alto, CA, USA, 2017; pp. 4760–4761.
7. Dutta, A.; Deb, T.; Pathak, S. Automated Data Harmonization (ADH) using Artificial Intelligence (AI). OPSEARCH 2021, 58, 257–275.
8. Behringer, M.; Hirmer, P.; Mitschang, B. A Human-Centered Approach for Interactive Data Processing and Analytics. In Proceedings of the Enterprise Information Systems: 19th International Conference, ICEIS 2017, Porto, Portugal, 26–29 April 2017; Revised Selected Papers. Hammoudi, S., Śmiałek, M., Camp, O., Filipe, J., Eds.; Springer: Cham, Switzerland, 2018; pp. 498–514.
9. Stöger, K.; Schneeberger, D.; Kieseberg, P.; Holzinger, A. Legal aspects of data cleansing in medical AI. Comput. Law Secur. Rev. 2021, 42, 105587.
10. El Emam, K.; Mosquera, L.; Hoptroff, R. Practical Synthetic Data Generation; O'Reilly: Sebastopol, CA, USA, 2020.
11. Stach, C.; Bräcker, J.; Eichler, R.; Giebler, C.; Mitschang, B. Demand-Driven Data Provisioning in Data Lakes: BARENTS—A Tailorable Data Preparation Zone. In Proceedings of the 23rd International Conference on Information Integration and Web Intelligence (iiWAS), Linz, Austria, 29 November–1 December 2021; ACM: New York, NY, USA, 2021; pp. 187–198.
12. Hosseinzadeh, M.; Azhir, E.; Ahmed, O.H.; Ghafour, M.Y.; Ahmed, S.H.; Rahmani, A.M.; Vo, B. Data cleansing mechanisms and approaches for big data analytics: A systematic study. J. Ambient. Intell. Humaniz. Comput. 2021, 1–13.
13. Sharma, B. Architecting Data Lakes: Data Management Architectures for Advanced Business Use Cases, 2nd ed.; O'Reilly: Sebastopol, CA, USA, 2018.
14. Stach, C.; Bräcker, J.; Eichler, R.; Giebler, C.; Gritti, C. How to Provide High-Utility Time Series Data in a Privacy-Aware Manner: A VAULT to Manage Time Series Data. Int. J. Adv. Secur. 2020, 13, 88–108.
15. Stach, C.; Giebler, C.; Wagner, M.; Weber, C.; Mitschang, B. AMNESIA: A Technical Solution towards GDPR-compliant Machine Learning. In Proceedings of the 6th International Conference on Information Systems Security and Privacy (ICISSP), Valletta, Malta, 25–27 February 2020; SciTePress: Setúbal, Portugal, 2020; pp. 21–32.
16. Mindermann, K.; Riedel, F.; Abdulkhaleq, A.; Stach, C.; Wagner, S. Exploratory Study of the Privacy Extension for System Theoretic Process Analysis (STPA-Priv) to elicit Privacy Risks in eHealth. In Proceedings of the 2017 IEEE 25th International Requirements Engineering Conference Workshops, 4th International Workshop on Evolving Security & Privacy Requirements Engineering (REW/ESPRE), Lisbon, Portugal, 4–8 September 2017; IEEE: Manhattan, NY, USA, 2017; pp. 90–96.
17. Shapiro, S.S. Privacy Risk Analysis Based on System Control Structures: Adapting System-Theoretic Process Analysis for Privacy Engineering. In Proceedings of the 2016 IEEE Security and Privacy Workshops (SPW), San Jose, CA, USA, 22–26 May 2016; IEEE: Manhattan, NY, USA, 2016; pp. 17–24.
18. Stach, C.; Mitschang, B. ACCESSORS: A Data-Centric Permission Model for the Internet of Things. In Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP), Funchal, Portugal, 22–24 January 2018; SciTePress: Setúbal, Portugal, 2018; pp. 30–40.
19. Stach, C.; Steimle, F. Recommender-based Privacy Requirements Elicitation—EPICUREAN: An Approach to Simplify Privacy Settings in IoT Applications with Respect to the GDPR. In Proceedings of the 34th ACM/SIGAPP Symposium On Applied Computing (SAC), Limassol, Cyprus, 8–12 April 2019; ACM: New York, NY, USA, 2019; pp. 1500–1507.
20. Stach, C.; Mitschang, B. Elicitation of Privacy Requirements for the Internet of Things Using ACCESSORS. In Proceedings of the Information Systems Security and Privacy: 4th International Conference, ICISSP 2018, Funchal,

Portugal, 22–24 January 2018; Revised Selected Papers. Mori, P., Furnell, S., Camp, O., Eds.; Springer: Cham, Switzerland, 2019; pp. 40–65.

21. Gritti, C.; Chen, R.; Susilo, W.; Plantard, T. Dynamic Provable Data Possession Protocols with Public Verifiability and Data Privacy. In Proceedings of the 13th International Conference on Information Security Practice and Experience (ISPEC), Melbourne, VIC, Australia, 13–15 December 2017; Springer: Cham, Switzerland, 2017; pp. 485–505.
22. Gritti, C. Publicly Verifiable Proofs of Data Replication and Retrievability for Cloud Storage. In Proceedings of the 2020 International Computer Symposium (ICS), Tainan, Taiwan, 17–19 December 2020; IEEE: Manhattan, NY, USA, 2020; pp. 431–436.
23. Stach, C.; Gritti, C.; Mitschang, B. Bringing Privacy Control Back to Citizens: DISPEL—A Distributed Privacy Management Platform for the Internet of Things. In Proceedings of the 35th ACM/SIGAPP Symposium on Applied Computing (SAC), Brno, Czech Republic, 30 March–3 April 2020; ACM: New York, NY, USA, 2020; pp. 1272–1279.
24. Gritti, C.; Önen, M.; Molva, R. CHARIOT: Cloud-Assisted Access Control for the Internet of Things. In Proceedings of the 2018 16th Annual Conference on Privacy, Security and Trust (PST), Belfast, Ireland, 28–30 August 2018; IEEE: Manhattan, NY, USA, 2018; pp. 1–6.
25. Gritti, C.; Önen, M.; Molva, R. Privacy-Preserving Delegable Authentication in the Internet of Things. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC), Limassol, Cyprus, 8–12 April 2019; ACM: New York, NY, USA, 2019; pp. 861–869.
26. Chaum, D.; Damgård, I.B.; van de Graaf, J. Multiparty Computations Ensuring Privacy of Each Party's Input and Correctness of the Result. In Proceedings of the 7th Annual International Cryptology Conference (CRYPTO), Santa Barbara, CA, USA, 16–20 August 1988; Springer: Berlin/Heidelberg, Germany, 1988; pp. 87–119.
27. Shamir, A. How to Share a Secret. Commun. ACM 1979, 22, 612–613.
28. Barker, E. Recommendation for Key Management: Part 1—General; NIST Special Publication 800-57 Part 1, Revision 5; National Institute of Standards and Technology, Technology Administration: Gaithersburg, MD, USA, 2020; pp. 1–158.
29. Mavroeidis, V.; Vishi, K.; Zych, M.D.; Jøsang, A. The Impact of Quantum Computing on Present Cryptography. Int. J. Adv. Comput. Sci. Appl. 2018, 9, 405–414.
30. Borges, F.; Reis, P.R.; Pereira, D. A Comparison of Security and its Performance for Key Agreements in Post-Quantum Cryptography. IEEE Access 2020, 8, 142413–142422.
