Regulatory Compliance Verification: A Privacy Preserving Approach

Mohammed B. M. Kamel*†‡ , Massimo Morello§, Petri Sainio¶
*Department of Computer Algebra, Eotvos Lorand University, Budapest, Hungary
†Department of Computer Science, University of Kufa, Najaf, Iraq
‡Institute for Data Science, Cloud Computing and IT Security,
Furtwangen University of Applied Sciences, Furtwangen im Schwarzwald, Germany
§European Central Bank, Frankfurt, Germany
¶Department of Computing, University of Turku, Finland
mkamel@{inf.elte.hu, uokufa.edu.iq, hs-furtwangen.de}, morelloo1997@gmail.com, petri.sainio@utu.fi

***Abstract*—During the regulatory compliance verification, the verifier may need to gain access to private information that can present risks to the privacy of the entities being verified. Therefore, while ensuring that entities are compliant with the regulations, such as GDPR, the regulatory compliance verification process need to safeguard the privacy of those entities. This paper proposes a privacy preserving regulatory compliance verification protocol, which has been integrated and implemented in a use case to verify the compliance with the article 32 of the GDPR. It provides a regulatory verification protocol, based on the attribute verification protocol, that reveals no private information of the entity being verified, other than the fact that it is compliant. Our results showed that the proposed protocol can efficiently verify the regulatory compliance of an entity by an external verifier.**

***Index Terms*—attribute verification, regulatory compliance, privacy preserving protocol.**

I. Introduction

To safeguard systems, applications, data, and infrastructure against unauthorized access, theft, damage, or data loss, a collection of rules, technologies, and controls is used [1]. This can include data validation [2], identity and access management [3], data protection [4], data communication [5], threat detection and response [6], network security [7], and compliance management [8]. The domain of regulatory compliance represents, by definition, the first line of defense against threats. Ensuring up-to-date compliance is sometimes a legal obligation (like in the case of Data Protection Regulations), and many other times a desired goal. But compliance verification comes always with two drawbacks: the need for manual effort, and infringed confidentiality during the auditing process.

As companies depend on the scalability and accessibility of cloud-based services to store and manage their sensitive data, the cloud environment has grown more and more common [9]–[11]. Nevertheless, because of the broad usage, there are a number of security and privacy issues [12], many of them related to data security and regulatory compliance, as the data must be protected and kept private, especially in accordance with various regulations such as GDPR [13], SOC2 [14], and NIST [15].

The fundamental driver behind this paper is the necessity for privacy-preserving technologies that guarantee confidentiality for the company that is being audited for regulatory compliance checks, while still allowing a smooth assessment of the company assets. The contribution of this paper lies in providing an in-depth privacy-preserving attribute verification mechanism. We utilized the recent verification protocol [16] to overcome these issues.

II. Attribute Verification Protocol

The Attribute Verification Protocol [16] is a cryptographic protocol used to verify the possession of certain attributes by a user without disclosing the actual attribute values. This protocol is particularly useful for privacy-preserving applications where revealing the user's attributes might lead to privacy breaches. The protocol generally involves an attribute holder and an attribute Verifier. The Attribute Verification Protocol aims to achieve the following security properties:

- **Completeness**: a legitimate *Prover* who possesses the required attributes can always convince the *Verifier*.
- **Soundness**: a malicious *Prover* who does not possess the required attributes cannot convince the *Verifier*.
- **Zero-Knowledge**: the *Verifier* learns nothing about the user's attributes except that they satisfy the verification requirements.
- **Privacy**: the *Verifier* cannot link the user's verification requests to their actual identity, ensuring unlinkability and anonymity.

Several cryptographic primitives can be used to build Attribute Verification Protocols, such as ZKPs, group signatures, and anonymous credentials. The choice of the underlying primitive depends on the specific requirements and security assumptions of the application. An Attribute Verification protocol can be built using a combination of Decentralized Attribute-based Encryption (DABE) and Zero-Knowledge Proofs (ZKP): DABE provides fine-grained access control

This research was supported by Project no. TKP2021-NVA-29 implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

based on attributes, while ZKP allows users to prove possession of attributes without revealing the actual values.

III. Regulatory Compliance Verification Protocol

The proposed protocol in this paper leverages the Attribute Verification Protocol [16] to enhance the privacy and security of sensitive data. There are three main participants in the proposed protocol: *Prover*, which is the entity aims to prove its compliance with the regulations, *Verifier*, which is an independent auditor aims to verify the regulatory compliance of the prover, and *Issuer*, which is a trusted independent participant providing a specific service (e.g., a Cloud provider). Specifically, Attribute Verification Protocol will allow a verifier (i.e., an auditor) to verify the attributes of a Prover without revealing the actual attributes and grant access to encrypted data only if the verification is successful [17]. The adaptation process introduced a ticketing system that managed to fit perfectly in a decentralized logic.

The proposed protocol includes five main procedures as follows:

1. Global Setup This step is performed once during the system setup, and it is done by getting the security parameter λ as input, and performing the global setup with which will generate the global parameters that include: Two cyclic groups G, G_T, a generator g in G, a bilinear mapping $e : G \times G \rightarrow G_T$, a hash function $\mathcal{H} : \{0,1\}^* \rightarrow G$, a hash function $H : \{0,1\}^* \rightarrow \{0,1\}^d$

2. Issuer Setup An Issuer joins the system, and takes the global parameters as input, in order to produce its private key pair $\alpha_i, \beta_i \in Z_p$, which will be kept private. Then it will compute its public key pair as follows:

$$e(g,g)^{\alpha_i}, g^{\beta_i}$$

3. Tickets Generation A Prover with identity I_u and a set of attributes, requesting a ticket of compliance sk for each attribute i, contacts the relevant Issuers. The relevant Issuers, in turn, generate the user's corresponding tickets, i.e. the corresponding secret keys related to those attributes, leveraging (1) for each attribute i:

$$ticket(i,u) = g^{\alpha_i}\mathcal{H}(I_u)^{\beta_i} \tag{1}$$

A ticket related to a certain attribute will be a proof of possession of that attribute. The integrated ticketing system inside the Attribute Verification Protocol shapes the role of a secret key issued for a certain attribute.

The decentralized nature of the protocol will make an Issuer node be responsible for the issuance of a certain number of tickets, while other Issuers are responsible for others, distributing the power or authority across multiple nodes and decreasing significantly the issues revolving around the centralized scenarios.

4. Challenge Generation Prior to challenge generation, the Verifier needs to perform some setup steps: first of all, it defines the set of target attributes $T_v = \{t_0, t_1, \ldots, t_n\}$ for verification. Subsequently, it also randomly generates a challenge key $R \in G_T$, and prepares a challenge ciphertext for the Prover, by first hashing the challenge key R that will be used as the key to the symmetrically encrypted challenge, and right after that, encrypting the challenge by using the key $k = H(R)$, which includes a nonce $r \in Z_p$, the timestamp ts, and the public key of the Verifier PK_v to be used later to secure the returned response, where $\|$ defines the concatenation:

$$\text{challenge} = \text{Enc}_{H(R)}(r \| ts \| PK_v)$$

The verifier then generates a random number $s \in Z_p$, and converts the compliance policy Γ to the equivalent linear secret sharing scheme (LSSS) matrix $M(\Gamma)$ for encryption. The compliance policy in the proposed protocol requires the possession of all the attributes in T_v (Boolean operator AND). Therefore, it enforces an n-out-of-n policy [16] to prove the possession of all required attributes.

The Verifier then gets from the Issuers the public keys (pair of $e(g,g)^{\alpha_i}$ and g^{β_i}). These public keys will be based on the target attributes in T_v. Based on the number of columns in the LSSS matrix, two vectors γ and ω are generated, where their first elements are set to s and 0, respectively, and the remaining elements are randomly chosen from Z_p. The randomly generated challenge key R will be encrypted using DABE [18].

In order to perform the DABE encryption algorithm, the Verifier generates three parameters r_i, γ_i, and ω_i, based on the number of rows in the LSSS matrix $M(\Gamma)$, and for each of the attributes in T_v. r_i, taken as a parameter from the algorithm, is a random value that is chosen from Z_p, and γ_i and ω_i are computed using (2), where $M(\Gamma)_i$ indicates the ith row in $M(\Gamma)$.

$$\begin{aligned} \gamma_i &= M(\Gamma)_i \gamma, \\ \omega_i &= M(\Gamma)_i \omega \end{aligned} \tag{2}$$

The challenge key R will be then encrypted using (3):

$$C_0 = Re(g,g)^s \tag{3}$$

Additionally, three components C_{i1}, C_{i2} and C_{i3} are computed for each attribute i in T_v, using the (4):

$$\begin{aligned} C_{i1} &= e(g,g)^{\gamma_i} e(g,g)^{\alpha_i r_i}, \\ C_{i2} &= g^{r_i}, \\ C_{i3} &= g^{\beta_i r_i} g^{\omega_i} \end{aligned} \tag{4}$$

5. Response Generation The Prover can prove the regulatory compliance if the defined compliance policy Γ returns true, which means the Prover needs all the secret keys for each attribute in T_v in order to get R. In order to decrypt C_0, the Prover computes an intermediate value for attribute i using its secret key $sk(i,u)$ and parameter $C_i = (C_{i1}, C_{i2}, C_{i3})$.

Intermediate values are computed using (5), and the final step of getting the challenge key is performed using (6).

$$\frac{C_{i1} \cdot e(H(I_u), C_{i3})}{e(sk(i,u), C_{i2})} = e(g,g)^{\gamma_i} e(H(I_u), g)^{\omega_i} \quad (5)$$

$$e(g,g)^s = \prod_{i=1}^{|\Gamma|} e(g,g)^{\gamma_i} e(H(I_u), g)^{\omega_i} \quad (6)$$

The challenge key R is recovered from C_0 as $R = \frac{C_0}{e(g,g)^s}$. By recovering R, the prover can decrypt the challenge and sends back the response encrypted with the verifier's public key PK_v.

IV. Implementation and Evaluation

The proposed protocol has been integrated into the following use case: a Branch Entity (Prover) requesting a certain resource or service from the Main Entity (Verifier) and in order to access that resource/service, the prerequisite or necessary condition to be met, is being compliant with Article 32 of the GDPR [19] (See Figure 1). The branch entity uses services from Cloud providers (Issuers).

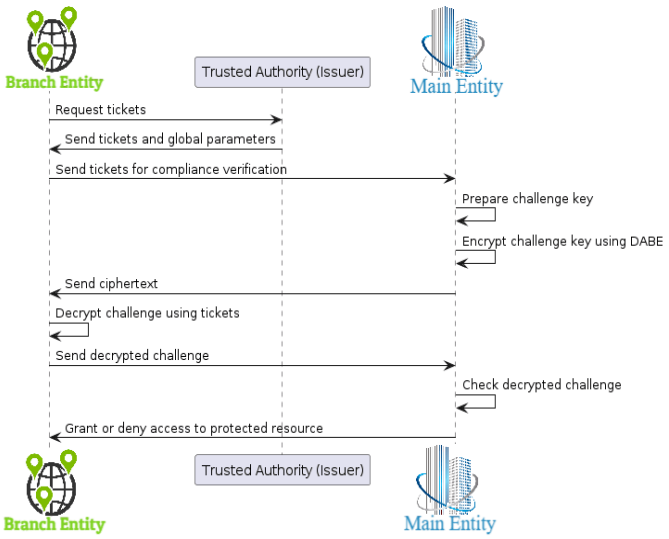

Fig. 1. Proposed Protocol applied to a cloud environment use case

The addressed requirements of Article 32, transposed then into attributes, are four:

- **Attribute 1**: Pseudonymization and encryption of personal data.
- **Attribute 2**: Ensuring the ongoing confidentiality, integrity, availability, and resilience of processing systems and services.
- **Attribute 3**: The ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident.
- **Attribute 4**: A process for regularly testing, assessing, and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.

Starting from these four requirements enforced by Article 32, a logic to assess compliance with them, seen as attributes, needed to be established, in order to assess whether a Prover (i.e., the Branch Entity) fulfills each of them. Moreover, in order to be compliant with the whole Article 32, the Prover would need to pass the check of all four of them (n-out-of-n compliance policy), so the necessary condition would be receiving four tickets (secret keys associated with each attribute) from the Issuers (Cloud providers).

The developed logic to assess compliance with each of the four attributes, is the following:

Pseudonymization and encryption of personal data (**Attribute 1**) needs to cover the following: Pseudonymization technique, Encryption algorithm and Key length. To be compliant with Attribute 1, the pseudonymization technique must be either "tokenization" or "masking"; The encryption algorithm must be one of the approved algorithms, e.g., "AES", "RSA", "ChaCha20";and, The encryption key length must be equal to or greater than a certain value, e.g., 128 bits for symmetric encryption or 2048 bits for asymmetric encryption.

Ensuring the ongoing confidentiality, integrity, availability, and resilience of processing systems and services (**Attribute 2**) needs to cover the following: security measures, assessment frequency, vulnerabilities count, and vulnerabilities severity. To be compliant with Attribute 2, the security measures must include certain mandatory items (e.g., a firewall);the frequency of security assessments should be at least quarterly; the number of vulnerabilities found in the most recent assessment must be below a certain threshold (e.g., 5);and, the severity of vulnerabilities found should not exceed a certain level (e.g., "low" or "medium").

The ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident (**Attribute 3**) needs to cover the following: backup type, backup frequency, Recovery Time Objective (RTO) in hours, and Recovery Point Objective (RPO) in hours. To be compliant with Attribute 3, the backup type must be either "incremental" or "full"; the backup frequency must be either "daily" or "weekly";the RTO should be less than or equal to a maximum value, e.g., 8 hours; and, the RPO should be less than or equal to a maximum value, e.g., 4 hours.

A process for regularly testing, assessing, and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing (**Attribute 4**) needs to cover the following: assessment frequency, assessment type, assessment status, and last assessment date. To be compliant with Attribute 4, The assessment frequency should be one of the valid options, e.g., "monthly", "quarterly", "yearly";the assessment type must be one of the valid types, e.g., "vulnerability scanning", "penetration testing", "security audit";the assessment status must be "passed"; and, the time elapsed since the last assessment date must be less than or equal to a maximum value, e.g., 18 months.

V. Evaluation

Evaluating the proposed system involved successfully integrating the modifications to the existing Attribute Verification protocol, transposing the protocol into code, assessing the successful implementation of all the participants in the Cloud Environment, their automation, the automation of regulatory compliance verification, and the system's overall performance in achieving privacy-preserving regulatory compliance. The ticketing system for regulatory compliance check was successfully integrated into the skeleton of the Attribute Verification Protocol, and so the transposition of it into code, by having a Lambda function written in Python for all the three participants, plus another one orchestrating chronologically their interactions and the whole workflow.

In terms of performance, the Lambda function of the Issuer was taken as a sample for assessing the overall efficiency of the artifact, being the most complex one and so the most computationally expensive. During the assessment, the Duration time metric have been retrieved from CloudWatch, which provides information about the execution time of the Lambda function during different invocations.Considering the operations performed by the Lambda function, which involves also cryptographic operations that can be computationally expensive, the execution times (See Figure 2) seemed to be even better than the most optimistic expectations:

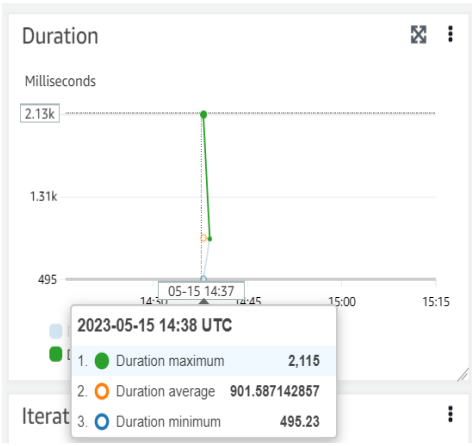

Fig. 2. Maximum, Average, and Minimum Duration

While Regulatory Compliance check contexts might not require high-speed and optimal verification time in comparison to other scenarios like electronic payments, high-frequency trading (HFT), real-time gaming, and video streaming, the achievement in terms of performances that were reached, established the foundations for creating awareness about this possibility. The lightweight protocols [20], [21] can be utilized to perform an efficient challenge generation phase.

The chart drawn by CloudWatch based on the logs, confirmed what has already been tested: the Issuer, no matter if the output was a compliant or not compliant record, always succeeded to reach the end of its lifecycle, resulting in a Success Rate of 100% (See Figure 3).

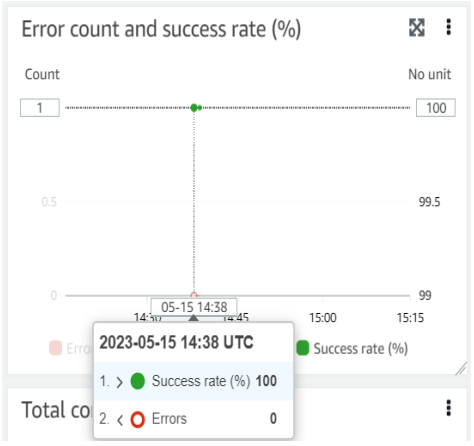

Fig. 3. Success Rate of the Issuer's Lambda function

VI. Related Works

Cloud computing is an abstraction of computing, storage, and network infrastructure put together as a platform that allows for speedy application deployment and dynamic scaling. A Key feature of cloud computing is its way of being self-service: the end user just fills in a form and it will already be up and running. Some clear advantages of this technology are: scalability, cost-savings, accessibility, flexibility, disaster recovery, environmental sustainability. The largest part of cloud users uses cloud computing services, which are housed in sizable, distant data centers that are kept up by the providers, through the internet [22].

Adam et al. [23] illustrated a solution for streamlining the process of analyzing and enforcing regulatory compliance in the cloud, by relying on text classification methodologies. The authors also discussed the use of multi-label Machine Learning techniques to automate compliance verification. Barati et al. [24] proposed a Blockchain-based architecture designed for the verification of GDPR compliance in service chains. The architecture introduced smart contracts for user consent, compliance, container management, and verification. The latter concerned enabling the identification of GDPR violators by employing smart contracts, and a trusted third-party verifier. In a related work [25] the same authors also illustrated an encoding of GDPR rules, used to record and verify operations, and smart contracts to translate the rules into code executable in Blockchain networks.

A classification scheme, supporting the implementation of GDPR rules as smart contracts, was defined. The scheme defined the types of operations carried out on user data by providers and the purposes of data processing. Barati et al. [26] proposed a framework relying on a container-based architecture to provide an isolated environment for executing cloud services, which helps to ensure GDPR compliance. The system was enhanced by a GDPR-priority contract that allowed users to specify their preferences for verifying compliance with specific obligations. Later on, the feasibility of the aforementioned framework was demonstrated by the same

authors in a cloud pharmacy scenario [27]. Cambronero et al. [28] developed "GDPRValidator": an automation tool, which was validated through a TAM model [29], and it assists SMEs in achieving GDPR compliance and auditing when managing data in the cloud. It implements a process to generate a series of documents with recommendations, which do not guarantee full GDPR compliance, but they help companies to better understand the regulation.

VII. Conclusion

The proposed Attribute Verification Protocol provides a novel solution for privacy-preserving regulatory compliance in Cloud Environment. The design and implementation of the suggested protocol aim to enhance the privacy of sensitive data stored in the cloud, while ensuring adherence to regulatory compliance. Even though the covered use case was developed around GDPR and its Article 32, the work has shown the versatility of the proposed solution, which will be able to be adopted for other regulations and multiple security objectives.

In terms of attribute verification, the Issuer in the future works can be tailored to a specific security control or article, like in the considered use case, with Article 32, or issuing tickets for multiple articles (like in the case of Data Protection Regulations) or security objectives (like in the case of standards, e.g. NIST). With the latter, the Prover will be able to request proof of compliance (to send to the Attribute Verifier) for multiple articles/securty objectives, with just a single interaction with the Issuer.

Acknowledgment

The authors thank Ammar Shareiyat from Orange Cyberdefense, Sweden for his insightful discussions.

References

[1] N. Thillaiarasu, S. Chenthur Pandian, G. Naveen Balaji, R. Benitha Shierly, A. Divya, and G. Divya Prabha, "Enforcing confidentiality and authentication over public cloud using hybrid cryptosystems," in *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018*, pp. 1495–1503, Springer, 2019.

[2] M. B. Kamel, P. Ligeti, and C. Reich, "D3vn: Decentralized abe-based distributed data validation network," in *Proceedings of Seventh International Congress on Information and Communication Technology: ICICT 2022, London, Volume 4*, pp. 653–661, Springer, 2022.

[3] Z. Georgiopoulou, "Trust management, privacy, authorization, and authentication in cloud computing environments," 2022.

[4] N. D. Dewani, Z. A. Khan, A. Agarwal, M. Sharma, and S. A. Khan, *Handbook of Research on Cyber Law, Data Protection, and Privacy*. IGI Global, 2022.

[5] M. B. M. Kamel and L. E. George, "Secure model for sms exchange over gsm," *International Journal of Computer Network and Information Security*, vol. 8, no. 1, p. 1, 2016.

[6] R. Brown and R. M. Lee, "The evolution of cyber threat intelligence (cti): 2019 sans cti survey," *SANS Institute. Available online: https://www. sans. org/white-papers/38790/(accessed on 12 July 2021)*, 2019.

[7] F. Al-Turjman, H. Zahmatkesh, and R. Shahroze, "An overview of security and privacy in smart cities' iot communications," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, p. e3677, 2022.

[8] F. Salguero-Caparrós, M. d. C. Pardo-Ferreira, M. Martínez-Rojas, and J. C. Rubio-Romero, "Management of legal compliance in occupational health and safety. a literature review," *Safety science*, vol. 121, pp. 111–118, 2020.

[9] Y. A. Qasem, R. Abdullah, Y. Y. Jusoh, R. Atan, and S. Asadi, "Cloud computing adoption in higher education institutions: A systematic review," *Ieee access*, vol. 7, pp. 63722–63744, 2019.

[10] A. Khayer, N. Jahan, M. N. Hossain, and M. Y. Hossain, "The adoption of cloud computing in small and medium enterprises: a developing country perspective," *VINE Journal of Information and Knowledge Management Systems*, vol. 51, no. 1, pp. 64–91, 2021.

[11] O. Ali, A. Shrestha, V. Osmanaj, and S. Muhammed, "Cloud computing technology adoption: an evaluation of key factors in local governments," *Information Technology & People*, vol. 34, no. 2, pp. 666–703, 2021.

[12] Y. S. Abdulsalam and M. Hedabou, "Security and privacy in cloud computing: technical review," *Future Internet*, vol. 14, no. 1, p. 11, 2022.

[13] "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," 2016. Accessed: 2023-02-26.

[14] "Soc 2 description criteria," 2018. Accessed: 2023-05-20.

[15] "Nist security and privacy controls for information systems and organizations," 2020. Accessed: 2023-05-21.

[16] M. B. M. Kamel, Y. Yan, P. Ligeti, and C. Reich, "Attribute verifier for internet of things," in *2022 32nd International Telecommunication Networks and Applications Conference (ITNAC)*, IEEE, 2022.

[17] A. Sahai, B. Waters, and H. Wee, "Attribute-based encryption with verifiable outsourced decryption," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 129–140, ACM, 2014.

[18] A. Lewko and B. Waters, "Decentralizing attribute-based encryption," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 568–588, Springer, 2011.

[19] "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," 2016. Accessed: 2023-02-26.

[20] M. B. Kamel, P. Ligeti, and C. Reich, "Odabe: Outsourced decentralized cp-abe in internet of things," in *Applied Cryptography and Network Security Workshops*, Springer, 2022.

[21] M. B. Kamel, P. Ligeti, and C. Reich, "Sdabe: Efficient encryption in decentralized cp-abe using secret sharing," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pp. 1–6, 2022.

[22] InfoWorld, "What is cloud computing?." https://www.infoworld.com/article/2683784/what-is-cloud-computing.html, 2014. [Online; accessed 9-March-2023].

[23] C. Adam, M. F. Bulut, M. Hernandez, and M. Vukovic, "Cognitive compliance: Analyze, monitor and enforce compliance in the cloud," in *2019 IEEE International Conference on Cloud Computing (CLOUD)*, IEEE, 2019.

[24] M. Barati and O. Rana, "Checking gdpr compliance for cloud-based services," in *2021 IEEE World Congress on Services (SERVICES)*, IEEE, 2021.

[25] M. Barati and O. Rana, "Tracking gdpr compliance in cloud-based service delivery," *IEEE TRANSACTIONS ON SERVICES COMPUTING*, vol. 15, no. 3, p. 1498, 2022.

[26] M. Barati, G. S. Aujla, J. T. Llanos, K. A. Duodu, O. F. Rana, M. Carr, and R. Ranjan, "Privacy-aware cloud auditing for gdpr compliance verification in online healthcare," *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, vol. 18, no. 7, p. 4808, 2022.

[27] M. Barati, K. Adu-Duodu, O. Rana, G. Aujla, and R. Ranjan, "Compliance checking of cloud providers: Design and implementation," *Distributed Ledger Technologies: Research and Practice*, 02 2023.

[28] E. Cambronero, M. Martínez, J. L. de la Vara, D. Cebrián, and V. Valero, "Gdprvalidator: a tool to enable companies using cloud services to be gdpr compliant," *PeerJ Computer Science*, vol. 8, p. e1171, 12 2022.

[29] F. Davis, "A technology acceptance model for empirically testing new end-user information systems," 01 1985.