BLOCKSTAND

A Proposal for a Blockchain-based Academic Certification System that Complies with the GDPR

Luis Anido-Rifón

Funded by
the European Union

INDEX

1. Introduction

The process of issuing and registering academic outcomes, both in vocational/informal educational programs promoted by companies and other training institutions, and in formal programs from regulated academic organizations, is currently a process carried out within each institution's proprietary systems and largely isolated from other organizations' record-keeping procedures. In cases where certificates are officially recognized by the public administrations, some standardized information must be sent to the governmental body in charge of their registration (e.g., the Ministry of Education in Spain or their equivalents in other countries). For the rest of accreditations, both vocational and informal [1], institutions or companies use their own custom model that normally cannot be accessed easily. This situation directly impacts the verification of students' educational data by a third party since, in many cases, the authentication of a transcript or certificate can only be performed manually, which is very costly in terms of resources and time. Due to this, the apparition of forged or false academic accreditations is not uncommon. On the other hand, there are other problems, such as dependence on third parties for access to academic certificates, and the difficulty or impossibility of verifying their validity in the event of the disappearance of the issuing entity.

To further study the broad range of data models used by education providers to store and generate certifications, this report also emphasizes that these models vary among formal institutions (e.g., higher education institutions); vocational training and completely informal settings (e.g., skills developed by a worker from their colleagues).

Formal education is defined as the educational experience that takes place in an official institution, such as accredited schools or universities, and that it is structured according to an official itinerary or program. Vocational education refers to educational experiences that may or may not take place in a registered educational institution, also structured according to some program or study plan. Finally, informal education takes place outside of any planned schedule or place, being the abilities and expertise gained from informal education providers [1], [2].

Table 1 offers a comparison among the characteristics defining each type of education [1].

Table 1. Broad education modalities.

Formal	Vocational	Informal
Usually at school	At institution out of school.	Everywhere
May be repressive	Usually supportive.	Supportive
Structured	Structured.	Unstructured
Usually prearranged	Usually prearranged	Spontaneous
Motivation is typically more extrinsic	Motivation may be extrinsic, but it is typically more intrinsic.	Motivation is mainly intrinsic
Compulsory	Usually voluntary	Voluntary
Teacher-driven	Can be learner-driven or teacher-driven.	Usually learner-driven
Learning is evaluated	Learning is usually not evaluated.	Learning is not evaluated
Sequential	Typically, non-sequential	Non-sequential

The relevance of vocational and informal certifications is due to the exponential growth that these educational approaches are experiencing. Consequently, there is a need to create a system where their quality and traceability can be guaranteed [3].

The primary challenge in validating knowledge acquired through vocational or informal learning stems from a significant lack of data management within these education providers, as insufficient data registries and information systems hinder the verification of the corresponding educational certificates. In other words, no convenient data sources are available to confirm whether an individual possesses the knowledge they assert. The absence of a standardized data model encompassing both formal and vocational makes this situation even more challenging, as evidence often relies on subjective factors like recommendations, statements, or claims from other individuals whose own knowledge may not be independently verified.

This proposal can be seen as a comprehensive approach to managing academic credentials using blockchain technology while ensuring compliance with the General Data Protection Regulation (GDPR). Section 2 explores the role of blockchain in academic certification, highlighting its advantages and challenges. Sections 3 and 4 provide an overview of the regulatory landscape, focusing on GDPR and the Electronic Identification, Trust Services, and Electronic Documents Regulation (eIDAS), both of which are instrumental for ensuring data protection and trust in digital transactions.

Section 5 details the proposed blockchain-based academic certification management system, designed to be GDPR-compliant. It covers key aspects such as academic information characterization, data management strategies, blockchain-based access control, stakeholder identification, privacy and data protection measures, eIDAS integration, and the process workflow for issuing and verifying credentials. Section 6 outlines business use cases, demonstrating the practical applications of the system in real-world scenarios.

To further assess its regulatory adherence, Section 7 evaluates the system's compliance with GDPR principles. Section 8 provides a comparative analysis between the proposed system and EBSI's Verifiable Credentials framework, emphasizing their similarities and differences in privacy, access control, scalability, and implementation. Finally, Section 9 presents the conclusions, summarizing key findings and future considerations.

2. Blockchain and Academic Credentials

Blockchain is the technology behind Bitcoin, the cryptocurrency initially proposed by Satoshi Nakamoto. [4]. A blockchain is a set of temporally ordered data blocks; that is, each block is cryptographically linked to the previous one, forming a chain ("blockchain"). All blocks are stored in a decentralized and distributed ledger and become trusted digital records that are virtually unmodifiable, but easy to verify. There is no centralized or hierarchical structure governing a blockchain, and information is shared through a peer-to-peer network.

Each block contains a trusted record of one or more transactions, created and exchanged among the participants (peers) of the blockchain network, which may modify the blockchain's status. To add new information to the chain, a consensus on its veracity must be reached among the peers in the network using different mechanisms.

Some blockchains support the use of smart contracts, [5] which are computer programs deployed on the blockchain and executed autonomously by all active participants in the network. These smart contracts can read and write information on the blockchain and collect information from other sources using oracles [6], process information, store transaction results, and even trigger actions in off-chain systems. All transactions are recorded tamper-proof on the blockchain, which adds trust.

Blockchain has matured from its first theoretical characterization by David Chaum [7], and its first implementation with Bitcoin by Satoshi Nakamoto [4] as a solution to implement digital currency; to the added layer of privacy, the development of smart contracts with Ethereum by Vitalik Buterin [8] and the incorporation of decentralized applications (dApps), as well as the expansion to different markets and fields, such as health, education or, more recently, artificial intelligence.

There exist various types of blockchains, ranging from public blockchains like Bitcoin and Ethereum to private ones, like those based on the Hyperledger framework [9] and consortium blockchains. Public blockchains operate openly without permission (i.e., they are known as permissionless blockchains), allowing anyone to join, read, send, and add information. This transparency might be suitable for ensuring transparency in academic certificate issuance. However, the lack of trust among network members in public blockchains necessitates stricter security mechanisms, which can hinder transaction speed and volume compared to other blockchain types.

Conversely, access to information in private blockchains requires permission, issued by the institution that created, deployed, and maintains the blockchain (i.e., they are known as permissioned blockchains). While offering increased trust among network teams, the security mechanisms need not be as stringent, enabling higher transaction speeds and volumes compared to public blockchains. Hyperledger is presently the most popular framework to create such blockchains.

Consortium blockchains combine elements of both public and private blockchains. They require permission to join but may be configured to allow any user to access stored information, depending on the group of institutions maintaining the blockchain.

Initially, private and consortium blockchains might seem preferable in educational settings, where greater data privacy management and control are necessary. However, this is not always the case [10], [11], due to other data protection aspects outlined in the General Data Protection Regulation, one of the world's strictest data protection legislations.

In this context, the educational sector addresses the blockchain technology and how it could be applied to efficiently protect academic information with relevant initiatives and projects since 2013 [12]. Although there is a growing number of blockchain-based educational applications, most of them are in an early stage of practical development and only a few of them can be used by academic institutions and their stakeholders [13].

Among the most illustrative examples of the introduction of blockchain for certificate management, the University of Nicosia was the first higher education institution to use Bitcoin to register such certificates [12]. As an additional example, the MIT Media Lab designed, prototyped and developed an open-source platform called Blockcerts [14] on top of the Bitcoin and the Ethereum blockchains, to issue academic diplomas and let them be verified easily by third parties [12]. This open-source platform is being utilized by several academic institutions. Both projects basically register in the corresponding blockchain the hash digest of academic certificates (i.e., not the complete data) to safeguard the information. As will be discussed below, immutably storing hashes of data does not guarantee compliance with data protection regulations.

While blockchain in education could revolutionize certificate issuance and management by creating immutable and accessible records for validating academic achievements, its adoption must be carefully considered concerning personal data protection and compliance with regulations like the GDPR. Existing literature on blockchain's educational applications [13], [15]–[33] commonly highlight certificate issuance and validation as the most frequent use case [15], [30], with other applications also present [31], [34]. However, these proposals do not predominantly focus on solutions compliant with the GDPR in their entirety—except for [35][36] from the proposers of this system. [36] and specially [35] outline a detailed model that will serve as the foundation for the proposal introduced in this report. This proposal will be compatible in design with the GDPR and its requirements.

3. The General Data Protection Regulation

As discussed in the previous section, while existing blockchain-based initiatives for the management of academic certifications represent inspiring proposals in educational credential management, such proposals do not explicitly integrate the principles and requirements established by the General Data Protection Regulation (GDPR). This poses significant challenges because, as discussed below, the GDPR not only defines rigorous standards for privacy protection - such as informed consent, data minimization and the right to be forgotten - but also requires transparency in the processing of personal information, critical aspects in systems that handle sensitive student and professional data. Addressing the GDPR framework is indispensable to ensure that technological solutions, including those based on blockchain, are legally sound, ethically responsible and compatible with the European regulatory ecosystem, thus ensuring user trust and the long-term sustainability of these tools in a global context.

The General Data Protection Regulation (GDPR) [37] is a regulation of the European Parliament (EU), 2016/679, concerning the protection of individuals regarding the processing of their personal data and the free movement of such information. This regulation applies not only to personal information like names, surnames, etc., but also to any other information that, in combination with other means, could be used to identify a person. Infringements of the GDPR may result in significant fines. As a consequence, initiatives and projects like this that require the management of personal data are conceived according to the principle “privacy, secure, and legal by design”, that is, in determining the means for data processing and during the processing itself, it is necessary to implement appropriate technical and organizational measures (General Data Protection Regulation, Article 25, paragraph 1).

Article 5 of the GDPR establishes the fundamentals and principles regarding data processing with which the proposed model must comply:

1. Lawfulness, fairness, and transparency: personal data will be processed lawfully, fairly, and transparently concerning the data subject (GDPR, art. 5.1.a).
2. Purpose limitation: Personal data will be collected for specified, explicit, and legitimate purposes and not processed in a manner incompatible with those purposes. Subsequent processing for archiving purposes in the public interest, scientific or historical research, or statistical purposes, according to Article 89.1, is not considered incompatible with those initial purposes (GDPR, art. 5.1.b).
3. Data minimization: personal data will be adequate, relevant, and limited to what is necessary for the purposes they are processed (GDPR, art. 5.1.c).
4. Accuracy: personal data will be accurate and, when necessary, kept up to date. All reasonable measures will be taken to ensure that inaccurate personal data, considering the purposes for which they are processed, are erased or rectified without delay (GDPR, art. 5.1.d).
5. Storage limitation: Personal data will be stored in a way that allows for the identifying of data subjects for the necessary processing purposes. Personal data can be stored for longer periods, subject to implementing the appropriate technical and organizational measures required by the regulation, to safeguard the rights and freedoms of the data subject (GDPR, art. 5.1.e).
6. Integrity and confidentiality: personal data will be processed securely, ensuring the appropriate security of personal data against unauthorized or unlawful processing and accidental loss, destruction, or damage, using suitable technical or organizational measures (GDPR, art. 5.1.f).

7. Accountability: the data controller will be responsible and able to demonstrate compliance with points 1-7 above (GDPR, art. 5.2).
8. Data modification (GDPR, art. 16): the data subject has the right to rectify personal data without undue delay by the data controller if it is inaccurate or incomplete.
9. Right to erasure: the user can request the deletion of inaccurately or unlawfully processed data as specified in the GDPR (GDPR, art. 17).
10. Data portability, granting individuals control over their personal data, allowing them to request and receive their provided data from a controller in a structured, commonly used, machine-readable format (GDPR, art. 20).
11. Right not to be subject to a decision based solely on automated processing, including profiling, that produces legal effects concerning them or significantly affects them: if a profile generated based on personal information has legal consequences, its management is entirely prohibited, a situation that is highly probable when storing academic information that might help create a profile (GDPR, art. 22).
12. Data transfer to countries outside the European Union: personal information can only be transferred to a destination country if it is deemed to have a data protection standard equivalent to that of the GDPR (GDPR, art. 45).

How this proposal addresses compliance with the principles above is discussed in Sect. 7.

4. Electronic Identification, Trust Services and Electronic Documents Regulation

Besides the General Data Protection Regulation (GDPR), a fundamental pillar in the protection of privacy and the ethical management of personal information in the European Union, another key component in the European digital architecture is the Electronic Identification, Trust Services and Electronic Documents Regulation (eIDAS) [Regulation (EU) No. 910/2014]. While the GDPR lays the groundwork for safeguarding personal data, eIDAS provides secure digital identification mechanisms and trust services, which are instrumental for ensuring reliable and transparent electronic transactions. This regulation not only reinforces security in the digital environment, but also facilitates interoperability between member states, thus creating an ecosystem where privacy, authenticity and technological integrity converge to strengthen citizens' rights in the digital age, facilitating secure and legally valid electronic transactions.

The eIDAS Regulation does not specifically address the issuance and verification of academic diplomas, but its framework facilitates the creation of digital solutions that enable the secure issuance and verification of electronic documents, including diplomas. Indeed, it is worth mentioning the Verifiable Credentials initiative in combination with the European Blockchain Services Infrastructure (EBSI[1]) as a promising solution, presently under development, that is supported by several institutions in the European Union [38].

Verifiable Credentials are authenticated digital certificates defined according to the W3C Verifiable Credentials[2] specification that provide verifiable evidence of the identity, competencies, achievements or qualifications pertaining to the entity or person to whom they are issued. These credentials are stored in a personal digital wallet that is expected to be compatible with eIDAS, allowing individuals to manage and share their information in a secure and controlled manner and with the guarantee of knowing who each intervening party is. The service provided on a a peer-to-peer network of interconnected nodes, promoted by the European Union, that utilizes blockchain technology to support different applications focused on various use cases, including the issuance and verification of verifiable credentials.

The combination of Verifiable Credentials with EBSI provides the EBSI's Verifiable Credentials Framework, a full framework for expressing, exchanging, and verifying information, using Verifiable Credentials, digital wallets, supported by blockchain technologies to provide trust[3].

An educational institution, at the time of issuing an academic certificate, issues a verifiable credential containing information about that diploma, such as the holder's name, the degree obtained and the graduation date. This credential is digitally signed to ensure its authenticity and is given to the student to store in their digital wallet to securely manage and share their credentials. When the holder needs to prove their qualification, for example when applying for a job, the verifiable credential is shared with the employer or institution concerned from the digital wallet.

The recipient of the credential uses the EBSI infrastructure to verify its authenticity. This process involves checking the issuer's digital signature and ensuring that the credential has not been altered or voided. EBSI's blockchain technology provides a trusted, decentralized registry that facilitates this verification without the need to contact the issuing institution directly.

[1] https://ec.europa.eu/digital-building-blocks/sites/display/EBSI/Home
[2] https://www.w3.org/TR/vc-overview/
[3] https://ec.europa.eu/digital-building-blocks/sites/display/EBSI/EBSI+Verifiable+Credentials

With eIDAS evolution towards eIDAS 2.0, emerging technologies, such as verifiable credentials and the European Blockchain Services Infrastructure (EBSI), are planned to be integrated to promote digital identity services in Europe.

Thus, the aim of this solution, in line with the proposal discussed in this report, is to reduce academic fraud, simplify the verification processes of academic certificates and allow users to manage their credentials according to their needs, in the context of an already available public blockchain ecosystem (i.e., EBSI).

5. A Proposal for a GDPR-compliant, Blockchain-based Academic Certification Management

The proposed model for the standardization of a blockchain-based academic certification system that complies with the GDPR and supports eIDAS-based identity accreditation will be described in this section, indicating its main components, operation and use cases, as well as its main advantages and disadvantages.

This proposal builds on the conceptual model for the issuance and validation of academic information, without content or format restrictions, supporting formal, vocational and informal training introduced in [35]. As pointed out above, presently there is no universally accepted standard for the representation of academic certifications. This proposal extends the original conceptual model with additional functionalities, namely the trusted identification of the intervening parties to guarantee the proper identification of all interacting partners by means of eIDAS.

Figure 1 depicts the proposed solution for a GDPR-compliant, blockchain-based academic certification management system.

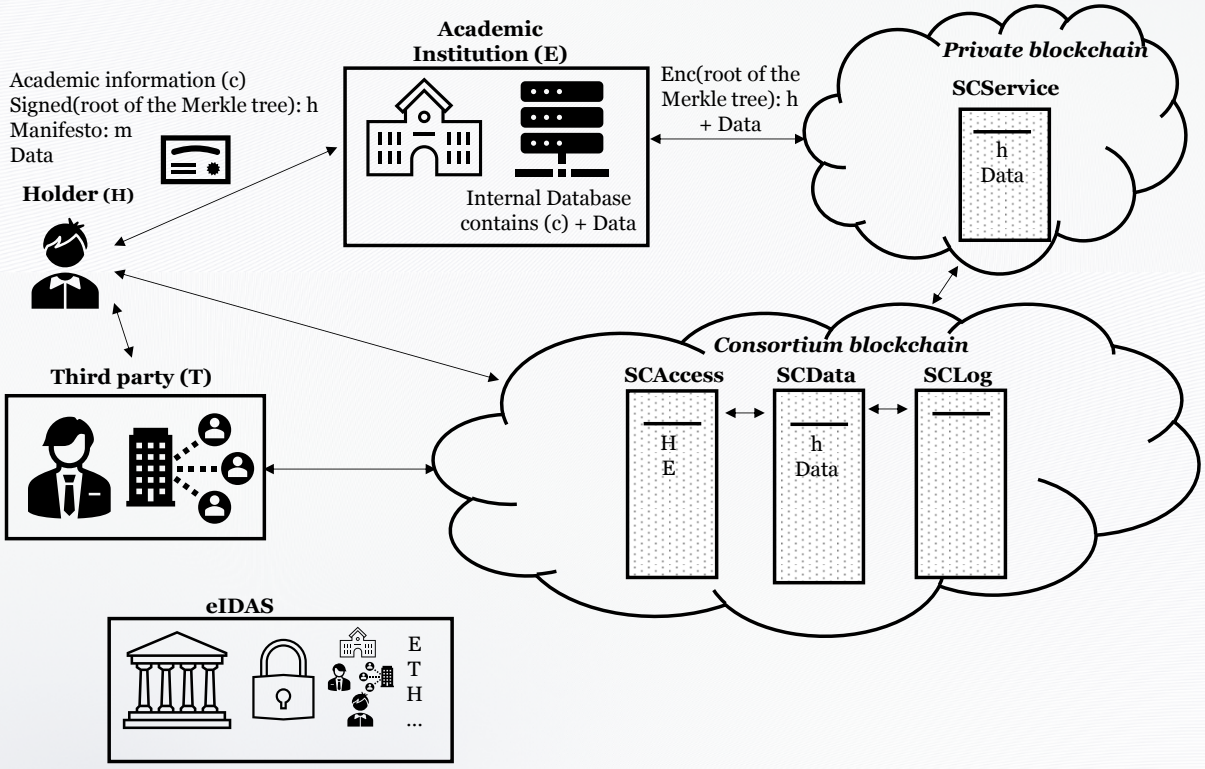

Figure 1 A GDPR-compliant, blockchain-based academic certification management system. General overview.

The solution is composed of the elements below:

a) Stakeholders:
   - Educational/Academic Institutions (with address E). They issue certificates -academic information (c) - and credentials for the education and training that they provide.
   - Data Holders (with address H). They receive certificates and credentials issued by academic institutions (E) certifying the education and training acquired.

- Third Parties (with address T). Entities (e.g., employers, public bodies, insurance companies, etc.) that need to verify or validate academic credentials and certificates issued to data holders (H).

b) Information:

- Academic Credentials or Certificates (c): is issued by the Educational/Academic Institution.
- Information to verify Academic Credentials (h): as it will be later explained (c.f. Sect. 5.7) h represents the data necessary to check the authenticity and integrity of the Academic Certificate.
- Manifesto (m): is a document designed by the Educational Institution that defines the structure of the Academic Credentials, specifying their fields and their structure. Its purpose is to provide a detailed schema that allows for the correct interpretation of the academic information.
- Data: represents additional data items intended to complement the academic information, but not part of neither academic credentials or the manifesto. This data is stored on the blockchain together with verification information (h) (cf. Sect. 5.7).

c) Infrastructure:

- Private blockchains: store anonymized verification data (h) for an educational institution (E) in an immutable and trusted manner, by means of smart contracts (cf. SCService below).
- Consortium blockchain: collects verification data (h) from several private blockchains. It serves as the endpoint for the verification of educational credentials, that is, third parties (T) interact with the consortium blockchain to verify data holders' credentials as discussed below (cf. Sect. 5.7). Its functionality is also provided by means of smart contracts (i.e., SCService, SCAccess, SCLog below) and contains only non-personal information.
- eIDAS: eIDAS identity accreditation service. Provide reliable and trusted identification of stakeholders E, H and T.

d) System's logic. The smart contracts discussed above are deployed on the private and consortium blockchains to address specific tasks in this proposal:

- SCService: in the private blockchains, there is a smart contract where the issuing entity stores the academic certificate verification data (h) and other data (Data, for example type of course for statistical purposes, information about how to retrieve the certificate by the holder directly from the issuing institution externally, without including it in the certificate information, allowing it to be modified without changing the certificate, etc.) that may be necessary, which will be eventually transferred to the consortium blockchain.
- SCAccess: smart contract responsible for receiving access requests to certain academic certificate verification information. When the certificate is registered in the system, the only accounts authorized to access this data are those of the certificate holder and the institution that issued the certificate (i.e., E and H in Fig. 1). If someone wants to access a particular certificate verification data stored in the system (i.e., in SCData), they must first check in SCAccess whether they have permission, since H can grant or revoke permissions at any time.
- SCData: smart contract in charge of storing the verification information of the academic data, (h) together with any other data (Data) defined by the issuing institution, as long as the latter does not include personal data.

- SCLog: generic smart contract in charge of recording access to academic verification information, so that the holder (H) can know whom and when accessed it, as required by the GDPR. The very nature of the blockchain logs any change in the state of the information stored, but the introduction of SCLog enables logging of read accesses too, even if they do not modify existing data. This structure is designed to cope with periods of high academic data issuance, such as at the end of a course or semester, by distributing processing across multiple private blockchains. This parallelization ensures efficient data management, while information is transferred to the consortium blockchain as soon as possible. However, some delay may occur as the consortium blockchain gradually absorbs all incoming data from the private chains. While Fig. 1 illustrates a single private blockchain, the proposed system is designed to accommodate multiple private chains within the consortium.

The next sections provide a detailed characterization of academic information within the system, emphasizing its flexibility in handling different types of educational credentials. Next, Section 5.2 outlines the data management approach, ensuring compliance with GDPR by storing personal information off-chain while leveraging blockchain for verification data. Compliance with GDPR is further discussed in Sect. 7. Section 5.3 describes the blockchain-based access control mechanism, allowing educational certificate holders to authorize and revoke third-party accesses securely. Section 5.4 focuses on stakeholder identification, highlighting the role of eIDAS in ensuring trust and authentication, while Section 5.5 discusses privacy and data protection measures embedded in the system's design. Section 5.6 briefly discusses the integration with eIDAS, reinforcing the legal validity and cross-border interoperability of academic credentials. Section 5.7 discusses how the data that is used to verify academic credentials is generated. Finally, Section 5.8 details the process workflow for registering and verifying academic credentials, demonstrating the system's efficiency, security, and practical applicability.

5.1 Academic Information Characterization

The academic information defined in this proposed system can be in any format, without limiting the number of fields or data contained in them. This is aligned with the requirement that it must be able to accommodate academic information of any type (formal, vocational, informal) and does not limit in any way the format, structure or type of data to be stored. In addition, this flexibility makes it easier to adapt existing models.

The issuing institution defines in a manifest document what data will be issued (e.g., name, surname, grades, etc.) and the structure of the manifest and this information, together with the academic data issued, is shared with the owner of the academic certificate. It is not necessary to follow any established format by consensus, but you can do it freely, although in case that adherence to some existing format is required, it can be seamlessly adopted.

5.2 Data Management

The institution issuing the academic certificate creates it, registers it in its own internal systems (i.e., its own backend) and sends the information to the data owner, who stores it in their digital wallet so that it can be shared with third parties, which in turn will be able to validate the academic information contained in the certificate.

Therefore, the proposed solution assumes that personal data will be stored in off-chain systems on the issuing institutions' facilities, since storing personal data on the blockchain is not compatible with the GDPR (not even a hash summary of the information, since hash functions are not considered valid for generating anonymized information [10], [39]). This is a key point of the proposed system to guarantee its compatibility with legacy systems with minimal adaptations. Note that it is unrealistic to expect every institution issuing academic certificates to overhaul their existing systems to adapt to an eventual standard. In turn, the proposed system would communicate through APIs or other widely consolidated technical solutions with the existing systems of the institutions. Obviously, depending on the technological level of the issuing entity, it may opt for a more or less automated solution, but in any case, there will be technical solutions within the reach of any organization.

5.3 Blockchain-based Access Control

The data owner keeps their academic certificates and credentials in their digital wallet. By means of a blockchain-based access control mechanism, they can grant access to third parties to their academic information, which can be obtained even from off-chain systems once access is granted. Third parties can also verify the information provided by the data owner reliably and privately.

The procedure for granting access involves the data owner expressly authorizing a third party, identified by their account in the blockchain system linked to their eIDAS trusted digital identity, to access the data needed to verify academic information. This anonymized verification data is stored in the blockchain in a trusted manner using smart contracts, so that it cannot be accessed without the appropriate permission. In addition, the data would be conveniently protected to comply with the GDPR, as described in more detail below.

5.4 Stakeholders' Identification

The proposed system ensures that issuing entities, third parties, and data owners are authenticated through eIDAS-trusted service providers, guaranteeing that all involved actors are verified and legitimate. The assurance that issuing entities, verifiers, and holders of academic certifications cannot be impersonated—thanks to authentication via eIDAS-trusted service providers—combined with the security of blockchain and autonomously executed smart contracts, enables the creation of a fully trusted, tamper-proof system. This system allows academic information holders to securely share their data, either fully or partially, with third parties while retaining complete control over granting and revoking access at any time.

Since third parties requesting certificate verification are reliably identified through eIDAS, the academic information holder always knows exactly who is receiving their data. This prevents phishing attacks and personal data theft, which can occur in other solutions where the identities of the involved actors are not fully trusted.

5.5 Privacy and Data Protection

As further discussed in detail below, the proposed system complies with the GDPR and is designed according to the philosophy of "privacy by design" (Art. 25 of the GDPR). Blockchain technology ensures that academic certificate verification data remains protected while still guaranteeing authenticity, as certificates can be reliably traced back to the issuing institution via eIDAS. This approach aligns with GDPR requirements, as data owners retain full control over their information, allowing them to selectively share portions of their academic records and revoke

access at any time. If deletion is legally permissible, data can be erased from off-chain systems, and since no personal data is stored directly on the blockchain, the system remains fully compliant with GDPR regulations.

This proposal also considers an access log system, so that the data owner can verify with whom educational information has been shared and by whom and when data verification information has been accessed.

5.6 Integration with eIDAS

As can be inferred from the discussion above, eIDAS has a relevant role in this proposal. Note that the eIDAS framework enhances the proposal through qualified digital certificates, which ensure the legal validity of digital signatures applied to academic credentials while guaranteeing the authenticated identity of the issuing institution (E). These certificates serve as a cornerstone for trust, as they comply with stringent EU standards for electronic identification and trust services.

Facilitating interoperability is another critical contribution, enabling academic credentials issued under this system to be seamlessly verified across all European Union member states. This eliminates jurisdictional barriers and ensures universal recognition of certifications, fostering cross-border collaboration and mobility.

Security and trust in this proposal are further reinforced by eIDAS, which enhances the legitimacy of the academic credential issuance and verification process. For instance, any verifying entity can validate the authenticity of an academic credential by using the public key of the issuing institution (E), which is securely registered with an eIDAS-recognized trust service provider. This mechanism ensures transparency and reduces the risk of fraud.

Additionally, eIDAS mandates the identification of third parties (T) involved in the system, ensuring that data holders (e.g., students or professionals) can trust the entities with whom they share their personal information. This safeguards against unauthorized data handling and strengthens compliance with privacy regulations.

Finally, eIDAS establishes mutual trust among issuing institutions, third parties, and data holders (H). It ensures that academic credential holders are unequivocally verified as their claimed identities, mitigating risks such as identity theft, document manipulation, or impersonation. By streamlining the sharing and verification of academic certifications in a private, reliable, and user-friendly manner, eIDAS lays the foundation for a secure and harmonized digital ecosystem across Europe.

5.7 On-Chain Verification Data

To allow the data holder (H) to share only a part of the data, a Merkle tree is used [40]. A Merkle Tree is a cryptographic data structure that organizes data into a hierarchical tree of hashes, ensuring data integrity and efficient verification. Each leaf node in the tree represents a hashed piece of data, while parent nodes store the hash of their child nodes, ultimately leading to a single Merkle root at the top. This root serves as a compact, tamper-proof representation of all the underlying data. Merkle Trees are widely used in blockchain technology, enabling fast and secure verification of large datasets without requiring access to the entire dataset, thereby enhancing scalability and security.

In our case, each branch of the tree represents an information node according to the institution's data model in which the different academic information elements are modeled, as well as additional metadata, such as whether the certificate is valid or not, among many other flexible possibilities (e.g., an expiration date, the record where it is stored in the entity's database, a pointer to it, etc.). A non-trivial secret key (i.e., key) is added to each piece of data in the Merkle tree, its hash-based message authentication code (HMAC [41]) is computed, and the root of the Merkle tree is calculated. In other words, each data item includes not only the information, but also a different key that E sends to H together with the academic information. This is done to further anonymize the information. Figure 2 illustrates the construction of a Merkle tree from four different data items in an educational credential.

HMAC ensures data integrity and guarantees authenticity, as only those with the secret key can generate the correct HMAC. Additionally, it offers stronger protection against preimage and length-extension attacks, making it more suitable for scenarios, such as the one considered in this proposal where sensitive data needs to be selectively shared (cf. Sect. 6). In other words, each data item contains both academic information and a unique key that entity E sends to the data holder H, ensuring that the shared information is anonymized and protected. This facilitates partial sharing of the academic data, as by providing only the desired information from the Merkle tree, a third party can reliably verify just the information required in a controlled way.

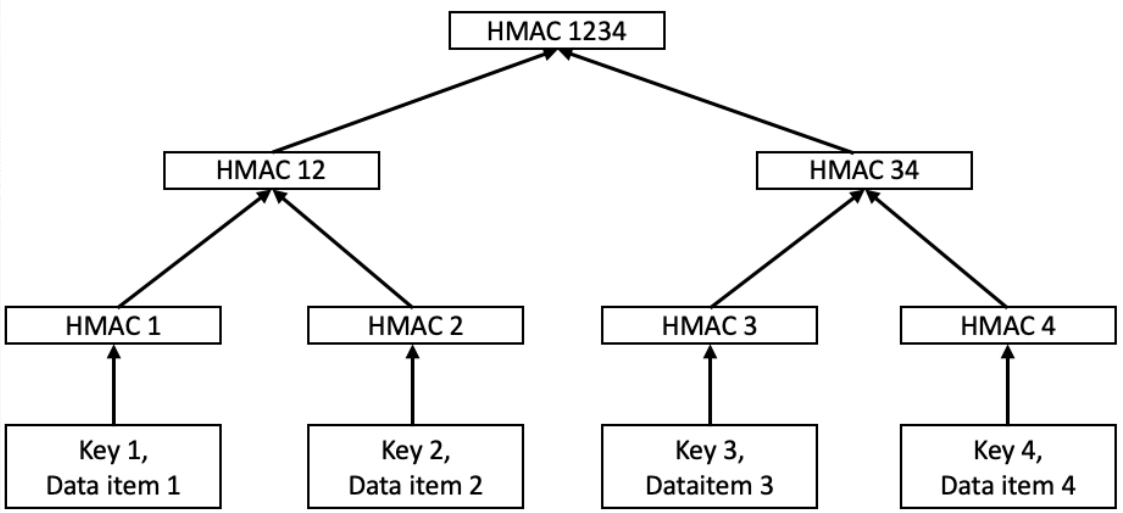

Figure 2 Example of a Merkle tree constructed from four data items in an academic credential.

Thus, the system is designed to accommodate any type of academic data with a completely flexible number of fields. These data are structured within a Merkle Tree and undergoes a HMAC process using unique keys for each piece of information. The resulting root of the nodes after calculating the HMAC is then signed and encrypted with the issuing institution's private key (cf. below), ensuring that the stored data is anonymized [39]. This proposal follows the recommendations of the Spanish Data Protection Agency (Agencia Española de Protección de Datos) [39] a key reference in data protection legislation, and suggests using an HMAC not only for a personal datum with a robust key but also for each and every datum in the Merkle tree, each with a different key, further strengthening the data anonymization process. This non-personal data is then recorded in a smart contract (i.e., SCService) within a dedicated private blockchain, to which the issuing organization connects via an API or other equivalent technologies, depending on its technological capabilities.

This integration allows institutions to retain their existing information systems without modifications. The data is subsequently transferred to SCData in the consortium blockchain, ensuring that high activity periods do not disrupt the registration process. As previously discussed in relation to scalability, multiple private blockchains operating in parallel further enhance system efficiency.

5.8 Detailed Process Workflow to Describe the Proposal for a GDPR-compliant Blockchain-based Certification Management

Figure 3 and the next paragraphs discuss the step-by-step process by which an academic institution registers verification information on its private blockchain, for third parties to be able to verify the original educational certificates with the holder's authorization, leveraging blockchain technology and eIDAS identification for trust and security. This approach guarantees privacy, transparency, and compliance with regulatory requirements while preventing fraud and unauthorized access.

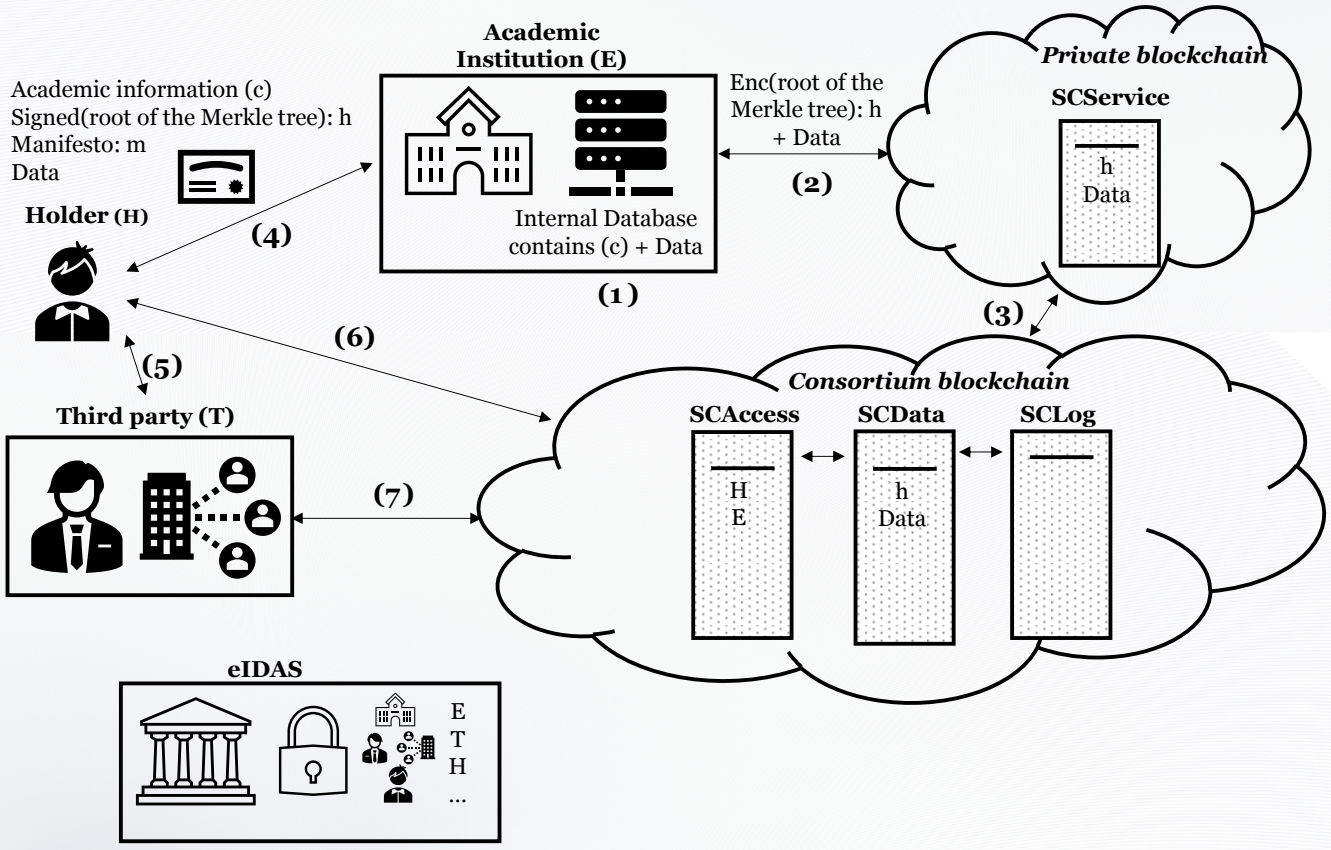

Figure 3 Detailed Process Workflow of the Proposal.

1. The data holder H completes a course and obtains an academic certification that is stored in the internal database of the issuing institution E (1), together with the corresponding academic records as well as other data (e.g., whether the certificate is valid or not, expiration date, etc.).
2. The institution (E) connects to a private blockchain (2) and submits verification information (h) using its blockchain account, which is associated to the official institution's credentials thanks to eIDAS, as well as other data that may be required.

3. This information is stored in the private blockchain and, when possible, sent to a consortium blockchain (3).
4. The academic institution sends the certificate holder (H), who is also identified through eIDAS, the certificate information (4), including personal details (such as name and surname), along with the necessary verification data (h) recorded in both the private blockchain and later in the consortium blockchain. H securely stores this information in their digital wallet. Additionally, a manifest (m) is included, the purpose of which will be explained later when detailing the specific contents of the verification data (h) stored on the blockchain. Extra data (Data) can be added by the issuing educational institution, containing information that will be stored on the blockchain outside the certificate itself but associated with it. These may include, for example, indications of the certificate type or other non-personal data that need to be recorded.

The issuance of an academic credential is illustrated in Figure 4.

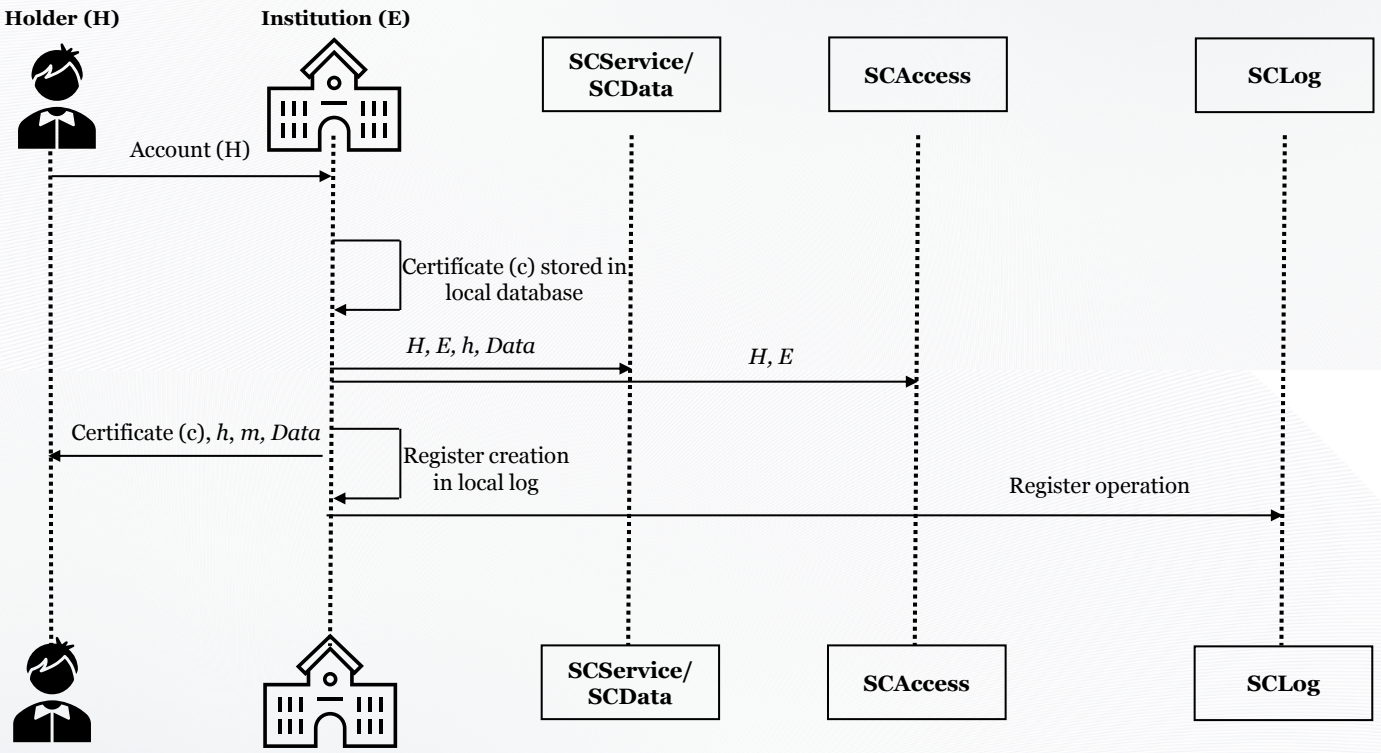

Figure 4 Issuance of an academic credential or certificate.

Educational credential verification is provided as discussed below (numbers refer to Fig. 3):

5. A third party (T), successfully identified by eIDAS to the data holder (H) (e.g., a recruitment company, a company's human resources department, a headhunter, etc.) contacts H and requests certain academic credentials from them. H sends the requested educational credential information (5) to T.
6. H adds the third party (T) to the SCAccess smart contract (6) so that they can obtain the verification data for the provided credentials.
7. Once enabled, T can now check that the information sent by H is correct and valid by simply making the corresponding query to the consortium blockchain (7). At any time, H

can revoke T's access to verification data in SCAccess and request T the deletion of the educational information received, in accordance with the GDPR.

With respect to the structure of the verification data (h), which enables full or partial validation of an academic certificate or credential, the academic records (c) remain stored in the issuing institution's databases, following their standard practices, as some open data structure. The system would employ an open data structure capable of representing various types of information, including names, email addresses, grades, and subjects studied. This data is organized into a file, with its structure detailed in a manifest document (m), which specifies what each data field represents (e.g., the first field corresponds to the first name, the second to the last name, and so on). This flexible approach supports formal, non-formal, and informal education without imposing any constraints on the types of academic credentials that can be represented.

Taking into consideration all the aspects discussed above, the general procedure is formally described as follows:

1. **Issuance of academic credentials by educational institution E**

- Institution E creates an academic certificate with the required H's data (e.g., name, surname, year, courses, credits, final grade, ...).
    - A summary is generated as the root of a Merkle Tree constructed from anonymized academic data items using HMAC with unique keys for each item.
- E signs the Merkle Tree root using its eIDAS qualified digital certificate, which guarantees that:
    - The root was generated by E.
    - The signature is legal and verifiable.
- E sends to H the academic certificate, the keys and the signed root of the Merkle tree. Note that the signed root of the Merkle tree can be eventually used to check the identity of E, both by H and any third-party T.

2. **Blockchain registration of verification data h**

- E submits to its private blockchain, by means of SCService smart contract the signed Merkle Tree root (computed from H's academic and personal information).
    - The transaction is registered as issued from E's blockchain account to H's blockchain account.
    - Verification data is eventually transferred to the consortium blockchain.
    - Transactions are logged by means of SCLog.

3. **Access authorization by H**

- H wants to share educational information with a third-party T (e.g., an employer).
  - Thanks to eIDAS, H can check T's identity.
- H grants access to blockchain's verification data to T and sends to them the complete educational credentials or parts thereof.

4. **Verification of educational credentials by T**

- T access the consortium blockchain, by means of the SCAccess smart contract, to obtain the signed root of the Merkle Tree.
  - Using eIDAS infrastructure, T can check that verification data was issued by E, and that E's blockchain address corresponds to E's identity.
  - The transaction is logged by means of SCLog.
- T verifies the data received from H.
  - Using the received data and keys, it computes the Merkle Tree root of the received data. Then, it compares the root generated with the root downloaded from the blockchain. If both values match, it can be guaranteed that:
    - H's personal and educational data was not tampered with.
    - E (as identified by eIDAS) is the actual issuer.

6. Business Use Cases

The proposed blockchain-based system for academic credential verification offers a secure, privacy-preserving and efficient way for individuals to manage and share their educational records. To illustrate its functionality, four key use cases are introduced below, each addressing a different scenario in which a credential holder interacts with third parties for verification purposes. In the **first use case**, the data holder shares their complete academic credentials with a third party, who then verifies them through the blockchain to ensure authenticity. The **second use case** introduces selective data sharing, where the credential holder only discloses specific information while keeping other details private. The **third use case** involves a scenario where the third party obtains academic credentials directly from the issuing institution, ensuring the highest level of trust and legitimacy. Finally, the **fourth use case** demonstrates how a credential holder can revoke previously granted access, ensuring full control over their academic data.

These use cases collectively highlight the system's flexibility, security, and compliance with GDPR regulations, enabling trusted verification while preserving user privacy. In each of these cases, we assume a context where a third party (T) has been granted access—through the SCAccess smart contract—to the verification data (h) associated with the data holder (H).

Case1: Data holder H shares their complete educational credentials received from an educational institution E.

Emma, a recent graduate, is applying for a postgraduate program at an international university. The institution requires a complete academic record, including her degree details, coursework, and grades. To streamline the verification process, Emma decides to share her full academic credentials securely through the blockchain-based system.

Emma accesses her eIDAS-compatible digital wallet, where her academic credentials are securely stored. She selects the university (T) as a recipient and grants authorization for verification through SCAccess. The complete set of credentials (c), including her academic data, corresponding keys for the HMAC calculations, and a manifest (m) defining the data structure, is securely transmitted off-chain to the university.

To ensure the authenticity of Emma's credentials, the university submits a verification request to SCAccess. SCAccess checks Emma's educational credentials' verification permissions and retrieves the verification data (h) from SCData on the blockchain.

The university receives the verification data and decrypts h to extract the Merkle Tree root corresponding to Emma's academic information. Using the received data, keys, and Merkle Tree structure, the university reconstructs the Merkle root and compares it with the one retrieved from SCData to ensure the information has not been altered. If the values match, the credentials are confirmed as authentic and issued by the original academic institution (E).

All accesses are recorded in SCLog, ensuring an auditable history of who accessed the credentials and when. If Emma later decides to revoke access, any changes in her permissions through SCAccess are also logged and, if necessary, recorded in the consortium blockchain.

Eventually, Emma successfully shares her full academic credentials with the university in a secure and verifiable way. The university can confidently confirm the authenticity of her records without the need for intermediaries or manual verification processes. The blockchain-based system ensures data integrity, privacy, and compliance with GDPR regulations while allowing Emma to maintain control over her credentials.

Figure 5 illustrates this case.

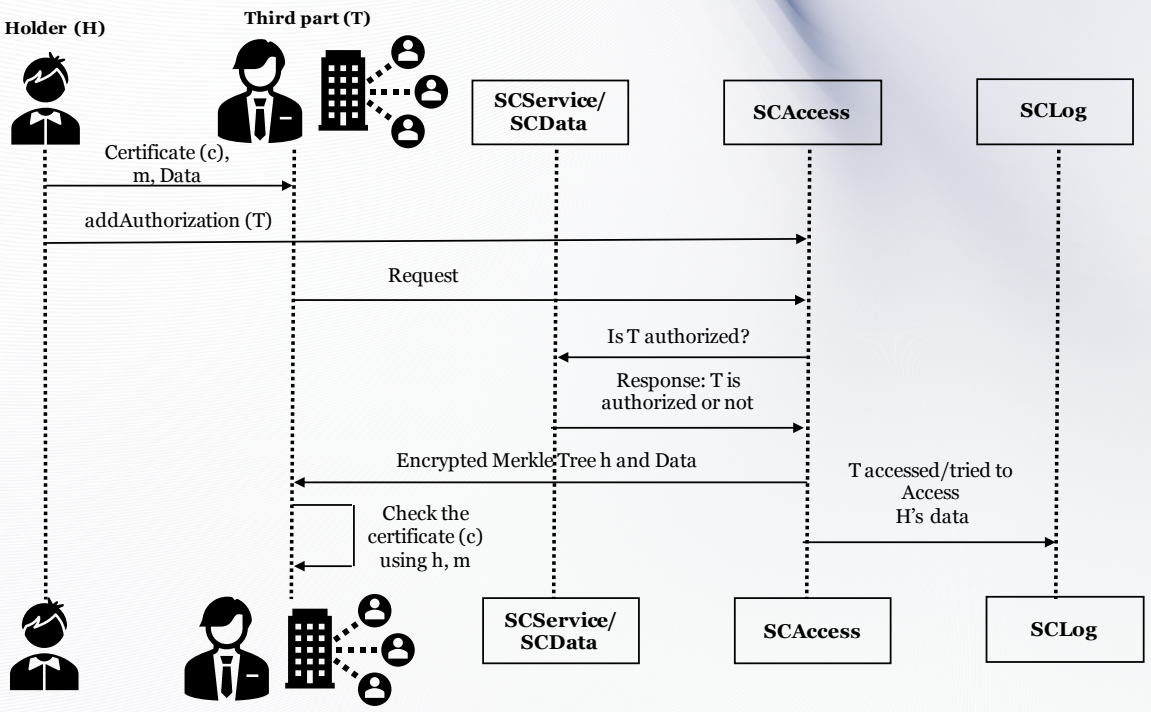

Figure 5 Case 1: complete educational credentials are shared by the data holder.

Case 2: The data holder shares just specific elements from their educational credentials.

Liam, a recent university graduate, is applying for a research fellowship. The selection committee requires proof of his master's degree and thesis title, but Liam prefers not to disclose his full academic transcript, including grades and other coursework details, to maintain privacy.

Liam accesses his digital wallet, where his academic credentials are securely stored. He selects the specific data elements he wants to share, such as his degree title and thesis topic, while omitting details like grades. Using a secure communication channel, he sends the selected data, their associated keys, the corresponding Merkle Tree paths, and the manifest (m) to the fellowship committee (T).

To verify the information, the committee (T) submits a request to SCAccess for the encrypted Merkle Tree root (h) associated with Liam's credentials. SCAccess first checks whether Liam has granted the committee permission to access the verification data.

Once permissions are verified, SCAccess retrieves the necessary verification data from SCData and records the transaction in SCLog, ensuring an auditable record of the access. The committee retrieves the verification data and authenticates it using the university's eIDAS public key to confirm that the data was indeed issued by the institution.

The committee then reconstructs the Merkle Tree using the received data, keys, and paths. They compare the computed Merkle Tree root with the one retrieved from SCData to ensure that the provided data is valid and has not been altered.

If the committee requires additional details—such as the issuance date of a language certificate to verify its validity over time—they must directly request Liam to share more data. The system does

not enforce specific requirements on what must be shared; it simply verifies the authenticity of what is provided.

Eventually, Liam successfully shares only the necessary academic details while keeping the rest of his transcript private. The fellowship committee can confidently verify the authenticity of the information, ensuring transparency and trust without unnecessary exposure of personal data. The blockchain-based system guarantees security, compliance with privacy regulations, and full user control over shared credentials.

Figure 6 illustrates this case.

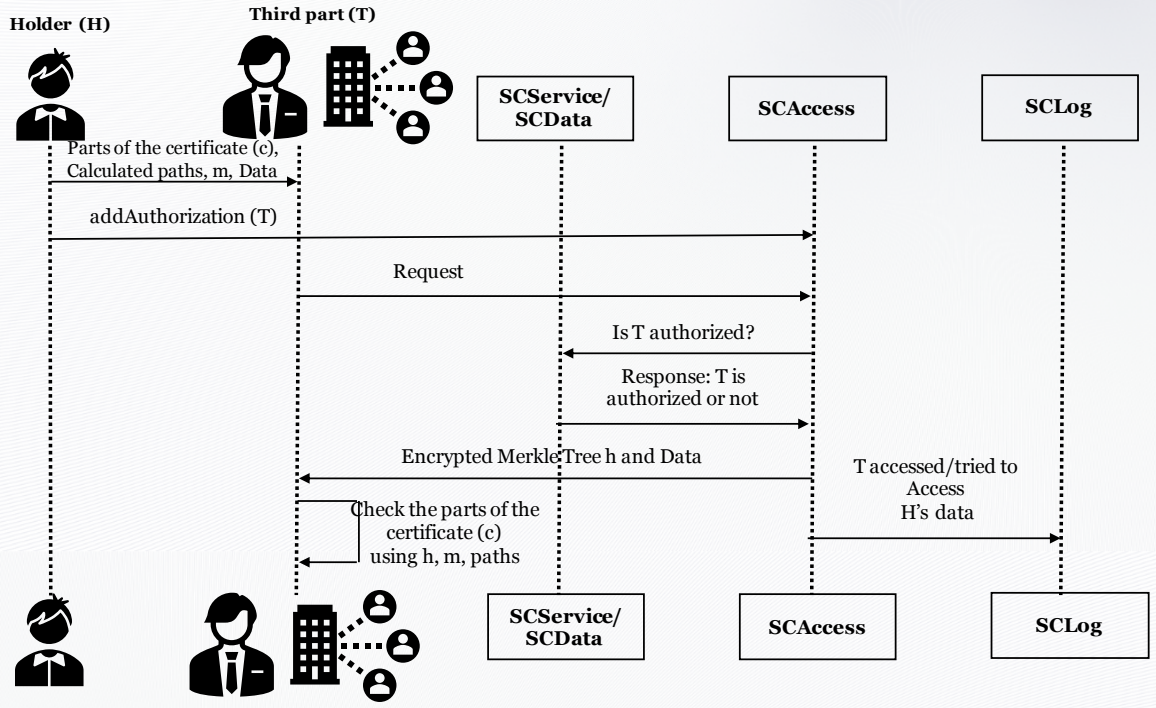

Figure 6 Case 2: specific data items in the academic credentials are shared.

Case 3: A third-party receives H's educational credentials directly from the issuing educational institution.

In this case (cf. Fig. 7), Sophia, a recent graduate in computer science from SpringfieldTech, is applying for a position at a multinational company, GlobalTel. She has just completed all the paperwork, but Sophia has not yet received an authenticated copy of her graduation certificate, but still she needs to provide evidence of her formal accreditation as a computer scientist on time to be considered for that position. Thus, she will send her resume to GlobalTel together with and authorization for the company to verify Sophia's education directly from SpringfieldTech, ensuring credibility and authenticity.

Sophia accesses her digital wallet. She selects GlobalTel (T) from the list of third parties and authorizes it to request her credentials directly fromSpringfieldTech (E). This authorization is recorded in SCData, and an authorization token is generated for GlobalTel to verify Sophia's accreditation directly from the university's information systems (e.g., access tokens, one-time passwords, API keys, OAuth codes, etc.).

Using the access token stored in SCData, GlobalTel sends a verification request to SpringfieldTech's academic database. Depending on SpringfieldTech technological infrastructure, this request is either processed automatically through an API or manually by university staff.

Before responding, SpringfieldTech verifies GlobalTel's permissions by consulting the SCAccess smart contract. Once verified, SpringfieldTech retrieves Sophia's academic records from its database and transmits them securely to GlobalTel through an off-chain channel.

Upon receiving the academic certificate, GlobalTel checks its authenticity by querying the SCData smart contract on the consortium blockchain. The verification process involves comparing the final certificate's HMAC with the Merkle Tree root stored on the blockchain. If the values match, the system confirms that the certificate is valid, was issued by SpringfieldTech, and has not been tampered with or revoked.

All access attempts, verification queries, and successful data retrievals are recorded in SCLog, ensuring a transparent and auditable process. Since GlobalTel retrieved the certificate directly from SpringfieldTech, the hiring process gains an additional layer of trust, minimizing the risk of credential fraud.

Sophia successfully verifies her academic credentials without having to manually send any documents. The employer gains confidence in the authenticity of her degree, while the blockchain-based system ensures privacy, security, and compliance with GDPR regulations.

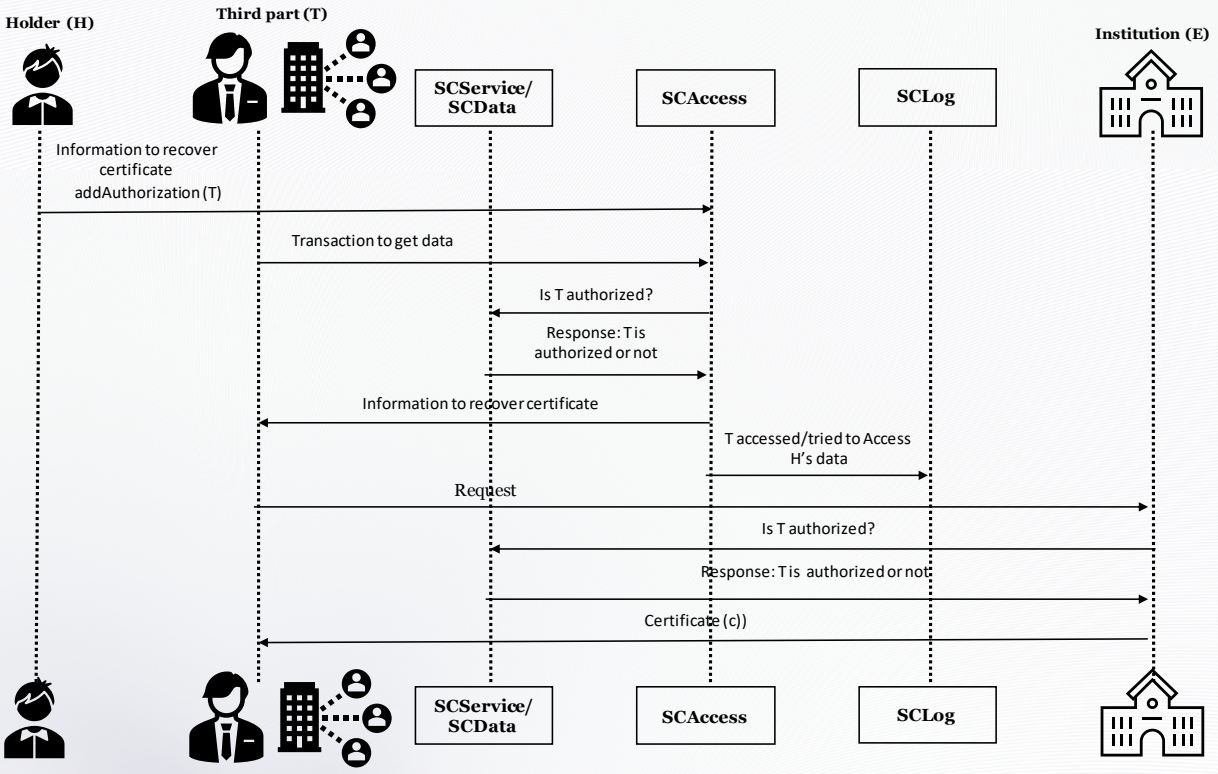

Figure 7 Case 3: T directly requests E educational certificates.

Note that Case 3 is also applicable in the case that Sophia loses her academic certificates. By means of this procedure, Sophia can act as a third party to obtain a fresh copy of her educational credentials.

Cases 1 and 2 support a situation where the entity issuing educational credentials ceases to operate. In case the issuing organization no longer exists, the system will continue to operate normally, if the holder still has their academic certificates, they can continue to share them, and the system will keep confirming their validity. Note that only in case 3 is the existence of the issuing institution necessary, as it assumes the role of sending academic certificates directly to third parties.

By utilizing Merkle Trees, the holder of academic credentials can selectively grant access to specific portions of the certificate data for validation. As previously mentioned, the data is structured within a Merkle Tree, where individual academic items and metadata are organized. This metadata includes information such as certificate validity, the location from which the certificate can be retrieved, its registration number in the internal database, the issuance date, and other relevant details. The flexible structure of the Merkle Tree allows for the inclusion of various types of information while maintaining data integrity and security.

To enable partial verification, H must specify which data is being shared, its corresponding location within the Merkle Tree, and the associated key values required for recalculating each information. This ensures that the algorithm correctly pairs the provided HMAC values and keys, allowing it to reconstruct the Merkle Tree and compare the final root with the one stored in SCData. By doing so, the system verifies both the authenticity of the certificate and the identity of the issuing institution.

If the institution deems it appropriate, it can include additional data elements alongside the previously explained verification information h. This additional data, while also suitable to be structured within the Merkle Tree, may also be stored directly on the blockchain, despite its limited flexibility as a general storage solution. Examples of such data might include a pointer to the database for querying information in case 3, or encoded details about the type of academic certificate issued (e.g., metadata useful for anonymous statistical queries by consortium managers via SCData). Note that this additional data should not include personal information to guarantee compliance with the GDPR. The system is designed with complete flexibility, allowing the institution to decide what is stored on the blockchain, and enabling the holder of educational credentials to choose whether to share this information or not.

Case 4: Revoking Permission for Academic Credential Verification

A graduate, Alice, previously granted permission to a recruitment agency (T) to verify her academic credentials while applying for a job. However, after securing a position elsewhere, she decides to revoke this access to prevent the agency from continuously verifying her data.

Alice accesses her digital wallet, where her academic credentials and associated permissions are managed. She selects the recruitment agency (T) from the list of authorized third parties and chooses the option to revoke access.

The revocation request triggers an update in the SCAccess smart contract. Alice's blockchain account removes the recruitment agency's access rights, ensuring that T can no longer retrieve verification data from the SCData smart contract on the consortium blockchain.

Once the smart contract is updated, the system logs the action in SCLog, ensuring an auditable record of the revocation. The recruitment agency (T) attempts to verify Alice’s credentials again but receives an "Access Denied" response, confirming that their permission has been successfully revoked.

Since GDPR requires that personal data be handled according to the owner’s consent, Alice also contacts the recruitment agency, requesting that they delete any locally stored copies of her

credentials. The agency is legally obligated to comply and remove the data from their internal systems.

The revocation process ensures that Alice retains full control over her academic credentials. The blockchain mechanism enforces access restrictions in real-time, while GDPR compliance ensures that third parties delete unauthorized copies. This guarantees privacy, security, and compliance with regulatory standards.

7. Compliance with the GDPR

The issuance of certificates within the educational sector is continually evolving to guarantee authenticity and information readiness. Additionally, the present-day job market requires people to keep track of educational achievements not originating from an academic institution, such as informally acquired skill sets or professional experience.

The innovative system proposed in this document is based on issuing academic certificates supported by blockchain technology and eIDAS to represent academic information uniquely and dynamically, ensuring compatibility with the GDPR requirements at the same time.

The core concept behind this model is a reformulation of the model proposed in [35] combined with eIDAS to add confidence not only in relation to the information recorded, but also among the actors who interact with it.

Generating a record containing personal information (e.g., names, surnames, etc.) and storing it directly in a blockchain or decentralized storage (e.g., IPFS[4], Swarm[5], or an equivalent storage model), does not comply with the GDPR [10]. Additionally, erasing this information from the mentioned decentralized infrastructure is impossible, breaching the right to be forgotten and the right to update personal data. The GDPR directly prohibits storing data in a permanent storage medium like a blockchain, even if encrypted. For instance, using a public blockchain like Ethereum and encrypting data with a robust algorithm would permanently store the encrypted information in the blockchain. This poses a risk, as advancements in computing power or algorithm vulnerabilities could expose this personal information in the future. As already explained, some widely used projects like Blockcerts [12] store a hash resume of academic information. However, the GDPR considers the hashed summary an invalid pseudonymization technique [10]. Although it may seem that information on a blockchain can be modified or deleted—since a new value can overwrite a previous one, effectively changing or nullifying it—the original value remains permanently recorded in a past transaction within a previous block of the blockchain. This ensures that all historical data is immutable and can always be traced back, even if updates occur.

With the proposed system, the data required to verify educational credentials is anonymized by means of the different operations performed (i.e., calculating the HMAC with different keys of each information and constructing a Merkle tree to extract its root), and only the data necessary to verify the authenticity of the original credentials is recorded, promoting data minimization, with the original information always being kept off-chain by the issuing institution and the credential holder in their digital wallet. On the other hand, anonymized verification data is not exposed to anyone but protected by a smart contract that provides strict access control, so that only users duly authorized by the credential holder may consult it. The system also guarantees that all interactions with this academic verification data will be reliably and securely notarized, thanks to a smart contract, and that data is available to the certificate holder.

One of the most complex GDPR requirements with which to comply in a blockchain-based system is that of data modification and deletion. Figure 8 illustrates the data modification process in the proposed system.

[4] https://ipfs.tech/
[5] https://www.ethswarm.org/

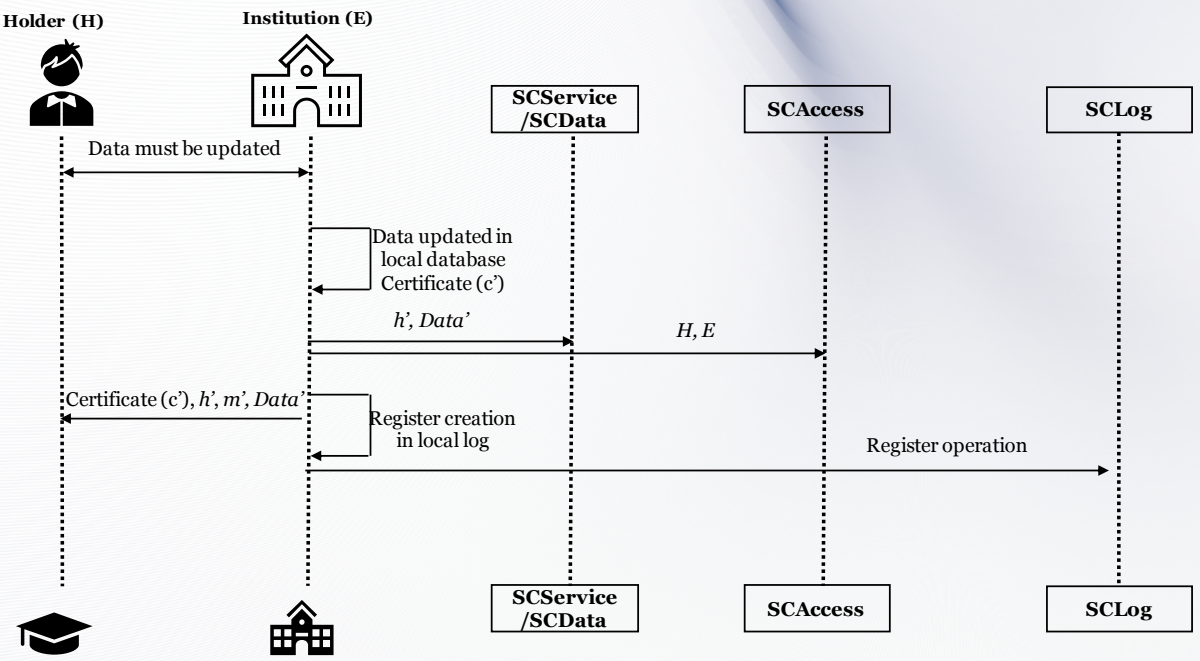

Figure 8 Data modification.

The data deletion is a particular case of data modification and is illustrated in Figure 9.

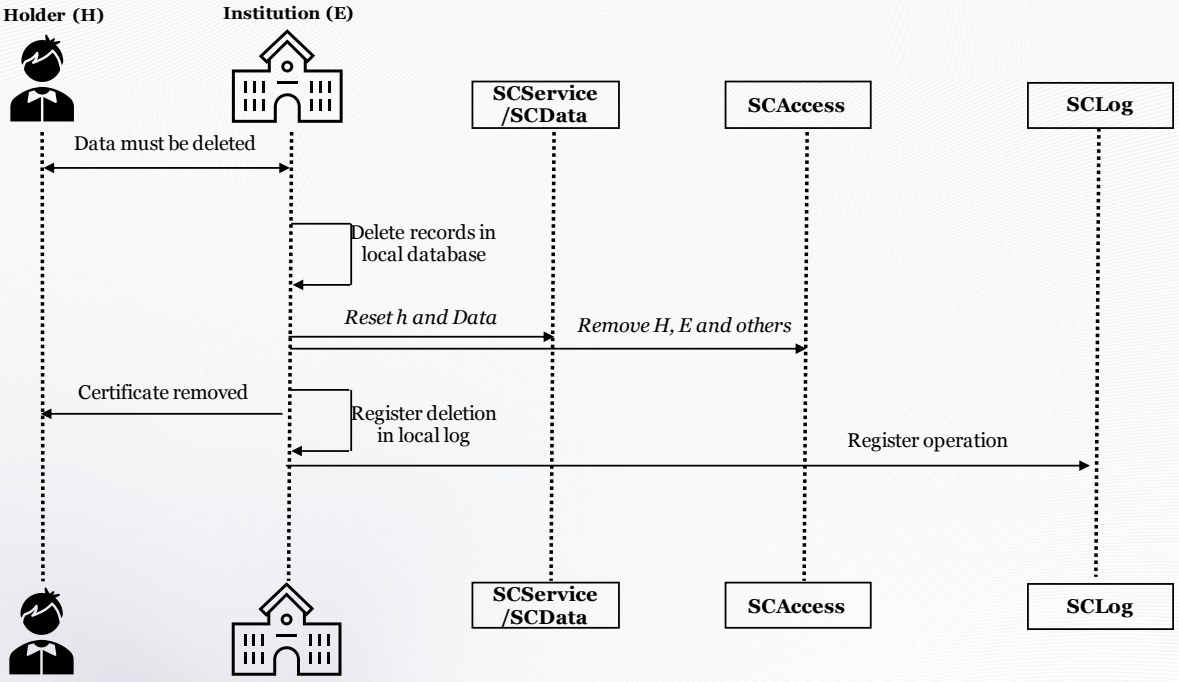

Figure 9 Data deletion.

The proposed system offers complete flexibility in verifying a variable amount of information, allowing the data holder to selectively share only specific data items. Additionally, the issuing institution can include certain verification details that it deems relevant, while ensuring that unauthorized users cannot access them. Furthermore, the system enables the recording of

anonymized data that can contribute to overall statistical insights, such as the number of credentials issued by a particular institution for a given qualification type. This functionality would be implemented in accordance with the preferences of issuing institutions and in full compliance with GDPR regulations.

Table 2 summarizes the main articles of the GDPR, as presented in Section 3, and how the proposed system complies with them.

Table 2. Compliance with GDPR principles.

Requirement	Compliance justification
Principle of lawfulness, fairness, and transparency. Personal data shall be processed lawfully, fairly, and transparently in relation to the data subject. (GDPR, Art. 5.1.a)	Explicit consent from the data subject is obtained, as requested by the institution and third parties seeking access to the data. Each issuer using this system must clearly specify to its users what data will be processed and handled. When users start interacting with the system through their wallets or an API, they should agree to a set of terms that clearly and concisely explain the processing of their data. This will make it clear that they have actively agreed to the terms under which their data will be processed and will be held as evidence of this.
Principle of purpose limitation (GDPR, Art. 5.1.b; Art. 7.3)	The issuing institution and authorized third parties must only use them for specific purposes. Data collected will only be used by the issuing entities to generate the certificates, and this will be done outside the system, when students register for degrees, courses, etc. Data within the model is not collected per se, only data elements' HMAC output (MAC), that is, anonymized data, which is the only information required for data integrity validation purposes. Service providers cannot send the data to other third parties; the user has full control over who can access their information and can revoke access permissions at any time.
Principle of data minimization (GDPR, Art. 5.1.c)	The model is entirely flexible regarding the set of fields used, ensuring that they are only the minimum necessary for the intended purpose. Moreover, the holder can select which fields third parties may access when interacting certificate. Besides, only the HMAC-output of educational certifications, signed by the issuer, will be stored for validation and integrity checks.
Principle of accuracy (GDPR, Art. 5.1.d)	The personal data is kept accurate, and in case of any inaccuracy, the issuing institution can easily remove or rectify them, as they are stored within their own databases. In other words, data holders can delete or modify data, if necessary, by modifying the information in the databases off-chain, and updating the verification

Requirement	Compliance justification
	information stored on-chain with a new MAC and with keys referring to the modified data.
Principle of storage limitation (GDPR, Art. 5.1.e)	The institution issuing personal academic information has the flexibility to delete it, when necessary, thanks to the proposed model. Due to data's academic nature, it will be available for as long as the user holds such academic credentials. Access to them by third parties will be managed by data holders, who can choose when to provide and revoke access rights. Credentials with expiration dates would be destroyed at issuers' premises. On-chain verification information does not contain personal data.
Principle of integrity and confidentiality (GDPR, Art. 5.1.f)	The issuing institution is responsible for safeguarding the data's integrity and confidentiality. The proposed blockchain-based model does not compromise this requirement. Data is anonymized using HMAC as recommended [39] with robust secret non-trivial keys forming a Merkle tree) using secure cryptographic algorithms. The MAC generated in this way cannot be exposed even knowing the original data due to the lack of these keys and a smart contract (SCAccess) prevents non-authorized users to query this information as an additional measure. In case of loss or destruction of the keys, if the original data remains in the user's account or wallet, allowing the verification of the certificate. If a data holder deletes the educational certificate data, they can request a new copy from the original issuer. However, if the issuer no longer exists and does not maintain a copy of the educational credentials, regeneration would not be possible.
Accountability principle (GDPR, Art. 5.2)	Access logs are maintained by system, which can be reviewed by the holder, as outlined in the system. The GDPR's data controller will be the issuer of educational credentials. In the same way that they provide educational certifications to their students in other formats, they shall be responsible for storing on-chain only the verification data of the credentials indicated by their data holders. Every access to the data will be logged. Data controllers, i.e., all participants with write rights on the chain who send (validation) data to miners/stakers for validation, must be informed about the data processing that they perform. Therefore, the model uses a more controlled environment, such as a private or consortium chain (although no personal data is ever exchanged). Issuing entities generate verification information based on data that they already possess

Requirement	Compliance justification
	independently of the model. Verifiers should only process anonymized data or proof of possession.
Data modification principle / right to rectification (GDPR, Art. 16)	The system allows for easy data modification, ensuring that third parties accessing the information always see the updated, accurate, complete, and correct version. Removal of the keys used during the HMAC process means that the data stored in the chain is arbitrary. The new verification data will then be added to the chain after going through a HMAC process with keys and encryption, as described in the proposal.
The user has the capacity to exercise the right to be forgotten (GDPR, art. 17).	If the academic information needs to be deleted, it is a matter of removing it from the issuing entity's internal centralized databases, ensuring no personal data remains on the blockchain. The destruction of the keys used during the calculation of the verification information causes the encrypted and data stored in the chain to become random/arbitrary data.
Data portability (GDPR, Art. 20)	The model enables data to be moved from one place to another by requesting it from the institution. The data holder has the right to have their personal data in a format that can be transferred to another (off-chain) environment.
The right not to be subject to a decision based solely on automated processing (GDPR, Art. 22)	Even if automated tools are employed to detect academic information on-chain, no such information is stored on-chain.
Principle of data protection by design and by default (GDPR, Art. 25)	The system was designed with a strong emphasis on GDPR compliance. From its inception, privacy and data protection principles were fundamental design guidelines, ensuring that all aspects of the system align with regulatory requirements.
Control of data transfer to countries outside the European Union (GDPR, Art. 45)	The model does not require this control, which remains in the hands of data processing entities, as no personal data are stored in the blockchain. In any case, this proposal utilizes consortium and private chains. Therefore, validators and nodes can be required to be physically located in the territory of the European Union or in countries with analogous data protection regulations already accepted by the Commission and the GDPR.

8. The Proposed System in the context of EBSI's Verifiable Credentials

The proposed blockchain-based academic certification system, designed to be GDPR-compliant and integrated with eIDAS, represents an innovative and disruptive solution to existing challenges in the issuance and verification of academic credentials. It addresses a critical need by ensuring the interoperability, integrity, and reliability of academic qualifications within a legal framework that prioritizes personal data privacy and security. At the same time, it meets the demands of an increasingly globalized world where individuals must prove their professional skills when seeking employment or career advancement, often competing with candidates from different educational backgrounds. Additionally, as digital interactions become more prevalent, reliable tools are essential to facilitate secure yet user-friendly transactions.

A key advantage of this system is its ability to support formal, non-formal, and informal certifications, setting it apart from existing solutions that may not fully comply with GDPR requirements. The integration of eIDAS further enhances trust among participants while ensuring the legal validity of certifications across Europe—an essential feature in today's interconnected educational and professional landscape.

From a technical perspective, the proposed model offers significant flexibility and scalability. By storing sensitive data off-chain and limiting on-chain information to anonymized and encrypted MACs, the system aligns with GDPR's data minimization principle. Furthermore, the use of Merkle trees enables selective data sharing, granting users granular control over which pieces of information they disclose. This represents a major advancement over previous models, which lacked such a high level of granularity, customization and data protection. The proposal also empowers users by allowing them to explicitly authorize third-party verification of their credentials while retaining the ability to revoke access at any time.

Despite these advantages, certain challenges remain. Although the system's architecture addresses scalability through the division into private and consortium blockchains, its adoption may be constrained by the technological infrastructure of issuing institutions. Some organizations with limited IT resources may find implementation challenging, although the solution does not require advanced technological investments—only the use of widely available APIs or equivalent integration methods.

When compared to initiatives like the European Blockchain Services Infrastructure (EBSI) Verifiable Credentials, it is evident that both share common goals, such as combating academic fraud and streamlining verification processes. The Verifiable Credentials Framework was developed by EBSI to enhance trust and security in digital information exchange across Europe. This framework utilizes the W3C Verifiable Credentials standard, W3C Decentralized IDs (DID), digital wallets, and blockchain technology to create a decentralized trust model. In this model, issuers can provide tamper-evident credentials to holders, who store them in their eIDAS compatible digital wallets. Holders can then present these credentials to verifiers, who can cryptographically confirm their authenticity without relying on centralized authorities. DIDs allow individuals, organizations, and institutions to create and control their own digital identities, which can be used to issue, hold, and verify Verifiable Credentials (VCs). These identifiers are registered on the blockchain, ensuring their authenticity and immutability while preserving privacy. In the case of holders' DIDs, they are anonymized to preserve personal data. This approach empowers individuals with control over their personal data and facilitates seamless, trustworthy interactions across various sectors within the European Union.

While this proposal and EBSI align in several key aspects, the proposed system introduces distinctive features that enhance privacy, scalability, and control over data sharing, reinforcing its potential as a secure and user-centric approach to academic credential management:

Data management and storage

The proposed system utilizes a hybrid approach, where sensitive academic data remains stored off-chain on issuing institutions' backend, while only anonymized information (by means of recursively calculating the HMACs of each data item with different keys, structured as a Merkle tree) is recorded on-chain and access control is performed using this technology. This approach ensures strict GDPR compliance, allowing modification and deletion of personal data. Similarly, with the DID-based verifiable credential model, only whether the credential is valid or not is recorded in a public blockchain.

Access control and data sharing

One of the strengths of the proposed system, which it shares with the EBSI's VC proposal, is its ability to allow holders to share their academic certifications in whole or in part by selecting which specific data they wish to disclose. This granular control is enabled by Merkle tree-based structures in the proposed system and in EBSI's VC, as W3C VCs also support this option.

Stakeholder identification

Both EBSI's VC system and this proposal are integrated with eIDAS, which ensures not only the legal validity of digital signatures but also the reliable identification of all parties involved, including issuers, verifiers and holders. This additional level of trust strengthens the security of the system against possible phishing or fraud attacks.

Scalability and decentralization

EBSI's VC relies on a single public blockchain, which may limit its ability to handle high transaction volumes, especially at peak times, such as the end of academic terms, although it is true that amount of information exchanged in EBSI's VC is smaller in comparison with this proposal, which is also more complex as it involves different smart contracts, interconnections and associated services. In contrast, the proposed system implements a modular architecture composed of multiple private blockchains interconnected with a consortium blockchain, which significantly improves scalability and distributes the workload more efficiently. To address the scalability challenge, considering that blockchain technologies are not particularly scalable or efficient as are centralized solutions, this proposal distributes the global infrastructure into different subsystems. According to geographical, academic, commercial or other criteria, different private blockchains are created and maintained by groups of institutions. The organizations sharing a private blockchain will use it to register verification data for their credentials. In this way, the global process of academic data verification is distributed into several private blockchains. Private blockchains will send the data to a consortium blockchain by means of secure connection mechanism. This consortium blockchain acts as the endpoint where data holders and third parties will be able to query verification information to confirm that both credential issuers' identities and credential content are correct and true. All actors will be eventually identifiable according to eIDAS.

GDPR compliance

Although EBSI's VC is designed to comply with European regulations, the practical implementation of some GDPR principles, such as the right to be forgotten or the withdrawal of access to data is less clear (note than once a VC is issued and stackeholders' DIDs are stored on-chain, they can be voided but their previous versions remain on-chain in some previous block).

Deployment and flexibility

EBSI's VC's data model utilizes W3C standards, namely Verifiable Credentials and Decentralized IDs. The proposed system does not specify the format or structure of academic data and ID management is done completely off-chain, making it highly adaptable for integration with existing educational management systems used by issuing institutions. Institutions implementing this proposal may continue issuing certifications in the form of EBSI verifiable credentials if desired, any other educational data standard (e.g., IMS Learning Information Package, IEEE Personal and Private Information, etc.), or even custom data models. Note that in the case of utilizing W3C Verifiable Credentials or other auto-verifiable credential model, the data verification component of the proposed system would no longer be necessary.

Permission granting and withdrawal

The proposed model allows the data holder to grant and withdraw permissions to the verification data and check the status of authorisations issued at any time. In the EBSI model, once a CV has been shared, the holder can revoke access, but the third party is responsible for deleting the data. If the third party fails to do so and retains a copy, they can still verify its validity indefinitely. In contrast, the proposed system offers stronger protection for the certificate holder by ensuring that access control is enforced directly at the verification level, preventing unauthorized validation after permission has been revoked. Besides, in the proposed model verification permissions can be granted and revoked selectively to individual third parties.

Access logging for accountability

The GDPR requires all accesses to personal information to be recorded. Through the proposed system, changes in access permissions are immutably recorded on-chain, so that any granting or withdrawal must be expressly authorized by the data holder, regardless of academic credentials being shared off-chain for verification. Additionally, the proposal includes a module that record accesses to the academic verification data, duly protected in SCData. Besides, thanks to the registration system, the data holder can check who has permissions to verify their data and who and when they have done it, something that is not possible with VC and EBSI.

Implementation technologies

Compared to EBSI's CV (which follows W3C standards such as VC and DID and utilizes a public blockchain already deployed in Europe), deploying the proposed model is relatively more complex. This is because it requires setting up private blockchains, connecting them to a

consortium blockchain, which is a relevant technical challenge), and implementing dynamic access control mechanisms for each account—an area where the necessary technology is still evolving. However, despite this added complexity, the trust and security provided by the proposed approach, as outlined earlier, offers advantages over the CV model with EBSI, particularly in terms of flexibility, access control, and data notarization.

In conclusion, this system represents a comprehensive and advanced solution that outperforms current initiatives in terms of privacy, flexibility and reliability. However, to ensure its widespread adoption, it will be key to address the challenges related to technological implementation and training of users and issuers.

Table 3 summarizes the four presently most advanced initiatives related to blockchain and education extracted from the academic literature, including this proposal and EBSI's VC, from the perspective of data storage infrastructure and compliance with the GDPR. More information about these additional initiatives and others can be found in [42]. Table 3. Comparison of this proposal and other related initiatives on to how information is stored and processed, and GDPR compliance.

Initiative	Scalability Nr. of Blockchains	Storage	Reg. Data	Verif	Right to forget	Modif	Revoke	ort	Data Acc.
Blockcerts [14]	1 blockchain Bitcoin/Eth	Off. + hash	✗	✓	-	✗	✗	-	✗
Nicosia [12]	1 bc Bitcoin	Off. + hash	✗	✗	-	-	-	-	-
Verifiable Credentials and EBSI	1 blockchain (EBSI)	Verifiable Credentials	✓	✓	✓	✓	✓ (per credential)	✓	✓
This solution	1 consortium and *n* private bcs	Off. + encrypted MAC	✓	✓	✓	✓	✓ (per third party)	✓	✓

Legend: BC = blockchain; Off. = off-chain; On. = on-chain

9. Conclusions

This proposal represents a significant advance in the issuance and verification of academic certifications in a digital environment, offering a solution that meets both the legal requirements of the GDPR and the practical needs of interoperability and trust. This system not only addresses the limitations of current solutions but also sets a higher standard in terms of flexibility, security and adaptability.

The main conclusions derived from this analysis are as follows:

- **Compliance and privacy by design**. The solution strictly respects the principles of the GDPR, including data minimization, the right to be forgotten and information portability. The use of anonymized data, the verification data access control mechanism and the off-chain storage of personal data ensure that personal data is protected at all times.
- **Integration with eIDAS**. The incorporation of eIDAS strengthens trust among parties, validating both issuers and holders and verifiers. This ensures that all interactions are secure and verifiable, reducing the risk of tampering and identity fraud.
- **Flexibility and scalability**. The ability to handle different types of academic certifications and credentials (formal, vocational and informal) and compatibility with existing infrastructures make the model adaptable to a wide variety of use cases. In addition, the distributed architecture improves scalability by distributing the load between private and consortium blockchains.
- **Comparison with other solutions**. While initiatives such as Blockcerts or EBSI's Verifiable Credentials are relevant steps towards the digitization of academic certifications, this proposal supports selective credential revocation and access notarization.
- **Challenges and opportunities**. Although the system is sound in its conceptual design, its implementation may require significant investment in infrastructure and training, especially in the case of institutions with limited resources or informal training settings. It will also be essential to ensure uniform adoption at the European level to maximize its impact.

In summary, this proposal for standardization not only addresses existing deficiencies in the management of academic certifications but also establishes a robust and scalable framework compatible with the most demanding standards. With the necessary developments for its technical implementation, its integration with European digital wallets and its dissemination, this proposal has the potential to become a global reference in the management of digital academic certifications.

10. Acknowledgement

The author would like to thank Manuel J. Fernández-Iglesias and Christian von Eitzen Delgado for their active contributions to this work.

11. References

[1] H. Eshach, "Bridging In-school and Out-of-school Learning: Formal, Non-Formal, and Informal Education," *J. Sci. Educ. Technol.*, vol. 16, no. 2, pp. 171–190, 2007, doi: 10.1007/s10956-006-9027-1.

[2] D. Colardyn and J. Bjornavold, "Validation of Formal, Non-Formal and Informal Learning: policy and practices in EU Member States1," *Eur. J. Educ.*, vol. 39, no. 1, pp. 69–89, 2004, doi: 10.1111/j.0141-8211.2004.00167.x.

[3] H. Cha and H. J. So, *Radical Solutions and Open Science: An Open Approach to Boost Higher Education*. 2020. [Online]. Available: https://doi.org/10.1007/978-981-15-4276-3

[4] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008, [Online]. Available: https://git.dhimmel.com/bitcoin-whitepaper/

[5] M. Swan, "Blockchain: Blueprint for a new economy," *Clim. Chang. 2013 - Phys. Sci. Basis*, no. 2014, pp. 1–30, 2015, [Online]. Available: https://books.google.com/books/about/Blockchain.html?id=RHJmBgAAQBAJ&pgis=1

[6] I. Bashir, "Mastering Blockchain," *Packt*, vol. 161, no. 4, pp. 335–339, 2017.

[7] D. L. Chaum, "Computer Systems Established, Maintained, And Trusted By Mutually Suspicious Groups," 1979.

[8] G. W. et Al., "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum Proj. yellow Pap.*, vol. 151, pp. 1–32, 2014.

[9] HYPERLEDGER, "Whitepaper Introduction Hyperledger," *July 2018*, 2018. https://www.hyperledger.org/learn/white-papers

[10] T. Lyons, L. Courcelas, and K. Timsit, "Blockchain and the GDPR," 2018. [Online]. Available: https://www.eublockchainforum.eu/sites/default/files/reports/20181016_report_gdpr.pdf

[11] T. Lyons and L. Courcelas, "Blockchain and cyber security," 2020. [Online]. Available: https://www.eublockchainforum.eu/sites/default/files/reports/report_security_v1.0.pdf

[12] A. Grech and A. F. Camilleri, "Blockchain in Education," Publications Office of the European Union, Luxembourg, 2017. doi: 10.2760/60649.

[13] A. Alammary, S. Alhazmi, M. Almasri, and S. Gillani, "Blockchain-Based Applications in Education: A Systematic Review," *Appl. Sci.*, vol. 9, no. 12, p. 2400, 2019, doi: doi.org/10.3390/app9122400.

[14] M. Baldi, F. Chiaraluce, M. Kodra, and L. Spalazzi, "Security analysis of a blockchain-based protocol for the certification of academic credentials," *arXiv Prepr. arXiv1910.04622*, pp. 1–12, 2019.

[15] H. Yumna, M. M. Khan, M. Ikram, and S. Ilyas, "Use of Blockchain in Education: A Systematic Literature Review," N. T. Nguyen, F. L. Gaol, T.-P. Hong, and B. Trawiński, Eds., in Intelligent Information and Database Systems, vol. 11432. Springer, Cham, 2019, pp. 191–202. doi: 10.1007/978-3-030-14802-7_17.

[16] R. Q. Castro and M. Au-Yong-Oliveira, "Blockchain and Higher Education Diplomas," *Eur. J. Investig. Heal. Psychol. Educ.*, vol. 11, no. 1, pp. 154–167, 2021, doi: 10.3390/ejihpe11010013.

[17] M. K. Dash, G. Panda, A. Kumar, and S. Luthra, "Applications of blockchain in government education sector: a comprehensive review and future research potentials," *J. Glob. Oper. Strateg. Sourc.*, vol. 15, no. 3, pp. 449–472, Jan. 2022, doi: 10.1108/JGOSS-09-2021-0076.

[18] C. Reis-Marques, R. Figueiredo, and M. de Castro Neto, "Applications of Blockchain Technology to Higher Education Arena: A Bibliometric Analysis," *European Journal of Investigation in Health, Psychology and Education*, vol. 11, no. 4. pp. 1406–1421, 2021. doi: 10.3390/ejihpe11040101.

[19] RushabhBalpande and K. Patil, "Usability of Blockchain Technology in Higher Education: A systematic review identifying the current issues in the education system," *J. Phys. Conf. Ser.*, vol. 1964, no. 4, p. 42017, 2021, doi: 10.1088/1742-6596/1964/4/042017.

[20] V. Aulia and S. Yazid, "Review of Blockchain Application in Education Data Management," in *2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, 2021, pp. 95–101. doi: 10.1109/ICSCEE50312.2021.9497997.

[21] B. Razia and B. Awwad, "A Comprehensive Review of Blockchain Technology and Its Related Aspects in Higher Education BT - Technologies, Artificial Intelligence and the Future of Learning Post-COVID-19: The Crucial Role of International Accreditation," A. Hamdan, A. E. Hassanien, T. Mescon, and B. Alareeni, Eds., Cham: Springer International Publishing, 2022, pp. 553–571. doi: 10.1007/978-3-030-93921-2_29.

[22] F. A. Sunny *et al.*, "A Systematic Review of Blockchain Applications," *IEEE Access*, vol. 10, pp. 59155–59177, 2022, doi: 10.1109/ACCESS.2022.3179690.

[23] M. Talat, S. Riaz, and M. S. Farooq, "Effect of Blockchain on Education: A Systemic Literature Review," *VFAST Trans. Softw. Eng.*, vol. 10, no. 2 SE-Articles, pp. 116–124, Jun. 2022, doi: 10.21015/vtse.v10i2.941.

[24] P. Ocheja, F. J. Agbo, S. S. Oyelere, B. Flanagan, and H. Ogata, "Blockchain in Education: A Systematic Review and Practical Case Studies," *IEEE Access*, vol. 10, pp. 99525–99540, 2022, doi: 10.1109/ACCESS.2022.3206791.

[25] F. Kabashi, H. Snopce, A. Aliu, A. Luma, and L. Shkurti, "A Systematic Literature Review of Blockchain for Higher Education," in *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 2023, pp. 1–6. doi: 10.1109/ITIKD56332.2023.10100049.

[26] N. A. Malibari, "A Survey on Blockchain-based Applications in Education," in *2020 7th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2020, pp. 266–270. doi: 10.23919/INDIACom49435.2020.9083714.

[27] P. Bhaskar, C. K. Tiwari, and A. Joshi, "Blockchain in education management: present and future applications," *Interactive Technology and Smart Education*, vol. ahead-of-p, no. ahead-of-print. 2020. doi: 10.1108/ITSE-07-2020-0102.

[28] G. Caldarelli and J. Ellul, "Trusted Academic Transcripts on the Blockchain: A Systematic Literature Review," *Appl. Sci.*, vol. 11, no. 4, 2021, doi: 10.3390/app11041842.

[29] R. Raimundo and A. Rosário, "Blockchain System in the Higher Education," *Eur. J. Investig. Heal. Psychol. Educ.*, vol. 11, no. 1, pp. 276–293, 2021, doi: 10.3390/ejihpe11010021.

[30] C. Delgado-von-Eitzen, L. Anido-Rifón, and M. J. Fernández-Iglesias, "Blockchain Applications in Education: A Systematic Literature Review," *Appl. Sci.*, vol. 11, no. 24, 2021, doi: 10.3390/app112411811.

[31] B. Hameed *et al.*, "A Review of Blockchain based Educational Projects," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 10, 2019, doi: 10.14569/IJACSA.2019.0101065.

[32] F. Loukil, M. Abed, and K. Boukadi, "Blockchain adoption in education: a systematic literature review," *Educ. Inf. Technol.*, vol. 26, no. 5, pp. 5779–5797, 2021, doi: 10.1007/s10639-021-10481-8.

[33] B. Awaji, E. Solaiman, and A. Albshri, "Blockchain-Based Applications in Higher Education: A Systematic Mapping Study," in *Proceedings of the 5th International*

Conference on Information and Education Innovations, in ICIEI 2020. New York, NY, USA: Association for Computing Machinery, 2020, pp. 96–104. doi: 10.1145/3411681.3411688.

[34] T. M. Fernández-Caramés and P. Fraga-Lamas, "Towards next generation teaching, learning, and context-aware applications for higher education: A review on blockchain, IoT, Fog and edge computing enabled smart campuses and universities," *Appl. Sci.*, vol. 9, no. 21, 2019, doi: 10.3390/app9214479.

[35] C. Delgado-von-Eitzen, L. Anido-Rifón, and M. J. Fernández-Iglesias, "Application of Blockchain in Education: GDPR-Compliant and Scalable Certification and Verification of Academic Information," *Appl. Sci.*, vol. 11, no. 10, 2021, doi: 10.3390/app11104537.

[36] C. Delgado-von-Eitzen, L. Anido-Rifón, and M. J. Fernández-Iglesias, "Blockchain for the Scalable Issuance and Verification of Private Academic Information," in *2021 International Conference on Advanced Learning Technologies (ICALT)*, 2021, pp. 436–438. doi: 10.1109/ICALT52272.2021.00138.

[37] P. Voigt and A. Von dem Bussche, *The EU General Data Protection Regulation (GDPR)*. Springer Cham, 2017. doi: https://doi.org/10.1007/978-3-319-57959-7.

[38] E. Tan, E. Lerouge, J. Du Caju, and D. Du Seuil, "Verification of Education Credentials on European Blockchain Services Infrastructure (EBSI): Action Research in a Cross-Border Use Case between Belgium and Italy," *Big Data Cogn. Comput.*, vol. 7, no. 2, 2023, doi: 10.3390/bdcc7020079.

[39] Agencia española de protección de datos (AEPD) and European Data Protection Supervisor (EDPS), "Introduction to the hash function as a personal data pseudonymisation technique," 2019. [Online]. Available: https://edps.europa.eu/data-protection/our-work/publications/papers/introduction-hash-function-personal-data_en

[40] H. Liu, X. Luo, H. Liu, and X. Xia, "Merkle Tree: A Fundamental Component of Blockchains," in *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS)*, 2021, pp. 556–561. doi: 10.1109/EIECS53707.2021.9588047.

[41] H. Krawczyk, M. Bellare, and R. Canetti, "HMAC: Keyed-hashing for message authentication," 1997.

[42] C. Delgado-von-Eitzen, L. Anido-Rifón, and M. J. Fernández-Iglesias, "NFTs for the Issuance and Validation of Academic Information That Complies with the GDPR," *Appl. Sci.*, vol. 14, no. 2, 2024, doi: 10.3390/app14020706.