Article

GDPR-Compliant Academic Certification via Blockchain: Legal and Technical Validation of the GAVIN Project

Alvaro Gomez * , Diego Estévez García , Christian Von-Eitzen-Delgado

Posted Date: 11 July 2025

doi: 10.20944/preprints202507.0926.v1

Keywords: blockchain; GDPR; academic certification; data protection; GAVIN project; privacy; decentralized systems

Article

GDPR-Compliant Academic Certification via Blockchain: Legal and Technical Validation of the GAVIN Project

Álvaro Gómez-Vieites [1,*], Diego Estévez-García [1] and Christian Delgado-von-Eitzen [2]

[1] Universidad Intercontinental de la Empresa, 15704 Santiago de CompostelaUE 1, Spain
[2] University of Vigo, 36310 Vigo, Spain
* Correspondence: alvaro.gomez@uie.edu

Abstract

Combining the immutability associated with blockchain technology with the European Union's General Data Protection Regulation (GDPR) has been considered for years a practically unsolvable conflict due to the very nature of blockchain and the GDPR. This article presents the GAVIN project (GDPR-Compliant Blockchain-Based Architecture for Universal Learning, Education and Training Information Management), a pioneering initiative that overcomes this challenge through an innovative technical and legal approach to trusted digital academic certification. Developed by atlanTTic (University of Vigo) and funded by the European Union, GAVIN proposes a scalable architecture that combines off-chain storage, encrypted HMAC anonymization, access notarization, and blockchain-based access control. Through exhaustive legal validation and the implementation of a working prototype, it is demonstrated that blockchain decentralization and GDPR compliance are not incompatible concepts. The model is presented as a replicable reference for institutions wishing to leverage distributed ledger technologies without compromising personal data protection. This paper details the legal design, technical architecture, and compliance mechanisms, offering a practical framework for implementing decentralized systems with privacy by design.

Keywords: Blockchain; GDPR; academic certification; data protection; GAVIN project; privacy; decentralized systems

1. Introduction

The digitization of educational records has amplified the demand for secure and verifiable systems capable of long-term data retention [1], fraud prevention [2], and global interoperability [3]. Traditional centralized systems are vulnerable to institutional discontinuity, technological obsolescence, and malicious manipulation. Blockchain [4] is a technology characterized by decentralization, transparency, and cryptographic security, and due to these features, it offers potential solutions to these challenges. Nevertheless, the integration of blockchain into personal data processing domains, particularly education, raises significant regulatory concerns, especially regarding compliance with the General Data Protection Regulation (Regulation (EU) 2016/679, hereafter GDPR) [5].

This paper investigates the GAVIN project (GDPR-Compliant Blockchain-Based Architecture for Universal Learning, Education and Training Information Management) a pioneering initiative that overcomes the challenge of issuing and verifying academic information through an innovative technical and legal approach to trusted digital academic certification. It was developed by atlanTTic (University of Vigo) and funded by the European Union, as a real-world implementation of a GDPR-compliant academic certification model [6] using blockchain and seeks to answer whether it is legally and technically feasible to align blockchain-based certification with GDPR principles. We aim to assess its legal soundness and technical viability, proposing it as a referential model for institutions seeking to leverage distributed ledger technologies in compliance with data protection standards.

1.1. General Data Protection Regulation (GDPR)

GDPR [7] is a comprehensive legal framework adopted by the European Union in 2016 to strengthen and harmonize data protection across member states. Its primary purpose is to safeguard individuals' fundamental rights and freedoms, particularly their right to privacy, in the context of personal data processing.

The GDPR establishes key principles that govern data handling, including lawfulness, fairness, and transparency; purpose limitation; data minimization; accuracy; storage limitation; integrity and confidentiality; and accountability.

These key principles aim to ensure that personal data is processed in a responsible, secure, and transparent manner, empowering individuals with greater control over their information while imposing obligations on organizations that collect and manage such data.

Table 1. GDPRs main articles.

Article	Principle	Description
GDPR, art. 5.1.a	Lawfulness, fairness and transparency.	Personal data shall be processed lawfully, fairly and in a transparent manner in relation to individuals.
GDPR, art. 5.1.b	Purpose limitation.	Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall not be considered to be incompatible with the initial purposes.
GDPR, art. 5.1.c	Data minimization.	Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.
GDPR, art. 5.1.d	Accuracy.	Personal data shall be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.
GDPR, art. 5.1.e	Storage limitation.	Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes subject to implementation of the appropriate technical and organisational measures required by the GDPR in order to safeguard the rights and freedoms of individuals.
GDPR, art. 5.1.f	Integrity and confidentiality (security).	Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.

doi:10.20944/preprints202507.0926.v1

GDPR, art. 5.2	Accountability.	The controller shall be responsible for, and be able to demonstrate compliance with.
GDPR, art. 16	Right to rectification.	The data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Taking into account the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement.
GDPR, art. 17	Right to erasure (right to be forgotten).	The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies: (a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed; (b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing; (c) the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2); (d) the personal data have been unlawfully processed;
GDPR, art. 20	Right to data portability.	The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided
GDPR, art. 22	Right not be subject to a decision based solely on automated processing.	The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.
GDPR, art. 25.	Data protection by design and by default.	Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures (…) in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

Additionally, it is important to note that the GDPR establishes a series of key figures in the processing of personal data, each with well-defined responsibilities:

- Data Subject: This is the natural person who owns the personal data. They have the right to control how their information is collected, used, and protected.
- Data Controller: This is the entity, which can be either an organization or an individual, that decides what personal data is collected, for what purpose it is processed, and what means are used. Although it may outsource processing, it remains the primary guarantor of compliance with the GDPR and must ensure that data processing is legitimate, transparent, and secure.
- Data Processor: This acts on behalf of the Data Controller, performing operations on personal data in accordance with their instructions. They are required to implement appropriate technical and organizational measures to ensure data protection and comply with the provisions of the GDPR. We should point out that personal data processing is any operation or set of operations performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction, in accordance with Article 4 of the GDPR.

1.2. Blockchain

Blockchain technology emerged in 2008 with Satoshi Nakamoto's [4] publication as the basis for the Bitcoin cryptocurrency. Since then, its use has expanded beyond the financial sector with cryptocurrencies such as supply chain [8], the medical sector [9], and Decentralized Finance (DeFi) [10], to name a few examples, establishing itself as a decentralized data structure that distributes information among nodes belonging to a peer-to-peer network. This technology allows data to be grouped into blocks linked together using cryptographic techniques, thus forming an unalterable chain. Each block contains a set of transactions (or other data depending on the use case), as well as a hash pointer that references the previous block, ensuring the integrity of the recorded information.

One of the most relevant features of blockchain is its ability to offer immutable and verifiable records without relying on a trusted third party. This makes it an attractive option for systems where traceability, transparency, and tamper resistance are priorities.

There are different types of blockchain, each with different technical and legal implications:

- Public blockchain: These are completely open networks where any user can participate by sending transactions, operating nodes, or even mining new blocks. Although this model guarantees high transparency and censorship resistance, it has significant limitations in terms of scalability and privacy. All stored information is visible to all participants, which implies total data exposure. Although encryption can be used, its use in immutable environments poses future risks due to the obsolescence of cryptographic algorithms, especially with the emergence of technologies such as quantum computing. Representative examples of this model are Bitcoin and Ethereum [11].
- Private blockchain: Unlike the previous model, private blockchains are managed by a centralized organization that restricts access to authorized nodes. This type of implementation allows for greater control over the network, improving both efficiency and privacy. Frameworks such as the Linux Foundation's Hyperledger Fabric [12] enable the development of these types of solutions, adapted to contexts where confidentiality and performance are essential, such as in the financial or corporate sectors.
- Consortium or federated blockchain: This approach combines elements of public and private networks. They are maintained by a group of organizations that share governance of the system. Although participation is restricted, certain levels of openness for public consultation can be enabled. This balance between controlled decentralization and privacy enables its use in inter-institutional collaborative contexts, such as higher education.

Generally speaking, an inherent limitation of blockchain networks is their processing capacity. Distributed verification, which guarantees the security and integrity of the system, leads to lower performance compared to centralized architectures.

Another key component of the blockchain ecosystem is smart contracts. Initially conceived by Nick Szabo [13], these smart contracts are programs that execute automatically when certain conditions are met. Although not exclusive to blockchain, their implementation within these networks ensures autonomous, unalterable execution without the need for intermediaries. Ethereum, one of the most well-known platforms, has popularized their use since its launch in 2015.

The synergy between blockchain and smart contracts has given rise to multiple innovative applications. In the educational field, they allow for the automation of processes such as issuing certificates, monitoring academic achievements, and verifying credentials, among other use cases, guaranteeing their validity without human intervention. In short, blockchain technology, with its ability to offer a secure, immutable, and distributed ledger, along with smart contracts, provides a solid foundation for revolutionizing the management of formal, non-formal, and informal academic information. This article will explore how these technologies can be effectively implemented in the education sector, offering a GDPR-compliant solution for the issuance, storage, and verification of academic data.

3. Methodology

3.1. Legal-Technical Evaluation Process

The methodology followed in the GAVIN project included:

- Initial Consultation: Meetings with technical developers and legal experts to understand the architectural blueprint.
- Document Review: Evaluation of system specifications, privacy policies, and technical diagrams.
- Regulatory Mapping: Analysis of GDPR articles, relevant case law, and guidance from the European Data Protection Board (EDPB)[1], The Spanish Data Protection Agency (Agencia Española de Protección de datos – AEPD)[2], and European Data Protection Supervisor (EDPS)[3].
- Iterative Feedback: Drafting of a preliminary legal report, integration of stakeholder feedback, and generation of a final report.

3.2. Analytical Framework

The assessment was guided by GDPR principles (Art. 5), rights of data subjects (Chapter III), controller and processor responsibilities (Chapters IV and V), and special attention to data minimization, pseudonymization, international transfers, and joint controllership.

2. Related Works

The implementation of blockchain in education has attracted growing academic attention, particularly in relation to its potential for enhancing trust, transparency, and portability of credentials. However, the regulatory implications (especially compliance with the GDPR) remain unresolved. This section reviews relevant contributions in two main areas: (1) the use of blockchain for educational credentialing, and (2) the legal challenges of aligning such systems with European data protection standards. A final subsection highlights the limitations of existing models and positions the GAVIN project within this research gap.

2.1. Blockchain in Education

Numerous academic contributions have explored the application of blockchain to educational certification. Grech and Camilleri [14] argued that blockchain can enhance the trustworthiness and portability of academic records. Sharples and Domingue [15] envisioned lifelong learning passports

[1] https://www.edpb.europa.eu/edpb_en

[2] https://www.aepd.es/en

[3] https://www.edps.europa.eu/_en

stored on immutable ledgers. More recently, Turkanović et al. [16] presented EduCTX, a decentralized platform for credit exchange. These works primarily emphasize technical feasibility and educational innovation. After reviewing relevant systematic reviews of the blockchain literature in the education sector, [17]–[31], it is concluded that blockchain in this field is applied with very varied approaches, although the issuance and validation of academic certifications are its most frequent use case [23], [27] but in the initiatives little attention is paid to something as relevant as the protection of personal data in general and compliance with the GDPR in particular. Most models lack robust legal compliance and verification, which undermines their applicability in real-world educational contexts [27]. While there are numerous initiatives aimed at applying blockchain technology in the educational sector very few of them except [6], [32] consider aspects as crucial as compliance with data protection regulations such as the GDPR. The only developed initiative is based on [6], which is precisely the model considered for the implementation of GAVIN and a previous work using Non-Fungible-Tokens (NFTs) to issue and validate academic accreditations [33].

2.2. GDPR Challenges in Blockchain Systems

The GDPR enshrines principles such as the right to erasure ("right to be forgotten", Art. 17), purpose limitation (Art. 5.1.b), and data minimization (Art. 5.1.c), which are difficult to implement in immutable, decentralized environments. Finck [34] categorizes blockchain as "technology at odds with the GDPR," noting irreconcilable conflicts between inalterable ledgers and dynamic regulatory rights. The European Data Protection Board [5] stresses that data controllers remain responsible, regardless of technological constraints.

Additionally, as noted in section 2.1, there are very few academic initiatives specifically focused on resolving the conflict between issuing and verifying academic credentials and strict GDPR compliance. Some proposals choose to record only the hash digest of academic information, rather than storing the full data on the blockchain.

The first institution to issue certificates using blockchain is considered to be the University of Nicosia [7]. In 2014, they recorded the hash digest [18] of academic information (not the full data) on the Bitcoin network, so that anyone could verify its authenticity simply by comparing the hash digest of the data with the information stored on the Bitcoin blockchain. It was an innovative proposal for its time, but very manual. The MIT Media Lab initiative, which created Blockcerts [19], is somewhat more elaborate. This is an open-source development that can be used in Bitcoin and Ethereum, storing the hash digest of the information in the corresponding blockchain for later verification.

In these cases, the certificate data is not stored, but only its hash digest. This means that if the holder of the academic information loses it and the institution disappears, it can no longer be used.

A slightly different approach is that of Alumnichain [21], which stores academic credentials in a private blockchain based on Hyperledger Fabric, something that the GDPR, as already mentioned, considers impossible since it is a non-erasable storage medium.

While this approach of storing only the hash digest may initially seem compatible, given that hash functions are unidirectional, the GDPR itself and regulatory bodies such as the Spanish Data Protection Agency (AEPD) and the European Data Protection Committee (EDP) warn that the use of hashes alone is not a valid anonymization technique. Consequently, this practice does not exempt data controllers from legal obligations relating to personal data since, according to [5], [20], a hash digest of data is not a valid anonymization technique under the GDPR perspective.

In a traditional centralized information storage system, there is a central node that determines what information, when and where it will be stored, as well as who will have access to it and with what level of privileges, as well as the relationships with third parties. Therefore, in a traditional centralized system, the central node will be the data controller, and the third parties that collaborate (subcontractors, suppliers, etc.) will be the data processors. A personal data controller, in accordance with Article 4 of the GDPR, is the natural or legal person, public authority, agency, or other body that, alone or jointly with others, determines the purposes and means of processing. If Union or Member State law determines the purposes and means of processing, the controller or the specific criteria for its nomination may be laid down by Union or Member State law. A data processor is a natural or legal person, public authority, service, or other body that processes personal data on behalf of the data controller.

However, blockchain is a distributed information storage technique, in which there is no centralized or hierarchical management of the information (decentralization). This technique uses consensus policies (algorithms) to validate the information included by each node and policies to detect the integrity of the originally recorded data and check whether the recorded information has been altered. Therefore, to understand the specific blockchain model, we must understand its constituent elements, which are basically:

- A participation policy that establishes the conditions under which and with what role one participates in the network.
- A distributed information storage policy.
- A data sharing policy.
- A consensus policy to validate new information in the system.
- An information integrity management policy.

In addition, blockchain governance is necessary to define:

- Who can access or participate in the network,
- With what permissions,
- What levels of service are provided,
- Block validation strategies,
- Traceability of information, etc.

The implementation of blockchain technology itself does not constitute the processing of personal data. Rather, it will be the use or various uses made of this blockchain that will give rise to one or more processing operations of personal data.

However, it must be taken into account that blockchain technology, regardless of the number and type of processing that can be implemented on it, will require a series of processing operations for its own management, such as:

- Recording transactions on the blockchain, which requires transforming the data into operations valid for the blockchain.
- Validating transaction blocks, proceeding to verify them according to the selected consensus mechanism.
- Implementing other additional processing operations that a node can perform on transactions, such as managing pending transactions, reordering them, storing historical data, processing events associated with smart contracts, providing data and services to applications, etc. In this sense, the use of a specific blockchain infrastructure as an element for the processing of personal data could generate specific non-compliance and risks to the rights and freedoms of data subjects. One of its key aspects is ensuring compliance with data protection principles, as well as allowing the exercise of data subjects' rights, and particularly the principle of accuracy, the principle of retention, and the rights of rectification and erasure.

Furthermore, taking into account the provisions of Recital 15 of the GDPR: "In order to avoid a serious risk of circumvention, the protection of natural persons should be technologically neutral and should not depend on the techniques used." Compliance with this regulation must be guaranteed. In this regard, the European Data Protection Board has stated that technical impossibility cannot be invoked to justify non-compliance with the requirements of the GDPR, especially considering that Article 25(1) of the GDPR establishes that data protection by design shall be taken into account at the time of determining the means of processing and at the time of the processing itself.

As established by the Spanish Data Protection Agency (AEPD) [35]:

“Compliance with the GDPR in the processing of personal data using one or more blockchain infrastructures requires two phases:

1) ensuring regulatory compliance and, once achieved,

2) evaluating and assessing the risks, some of which can be mitigated with legal, organizational, and technical measures. In cases of high risk, the assessment of the suitability, necessity, and proportionality of the processing must be completed."

Another factor to consider is that due to the relocation of nodes, international transfers of personal data may be taking place, which must comply with the GDPR provisions in this regard.

4. Technical Architecture of the GAVIN Model

The GAVIN project aims to build an academic certification model based on blockchain architecture to overcome the shortcomings of current certification systems. These include the lack of standards for managing long-term academic certificates worldwide. This is due both to a lack of persistence, which occurs when the issuer ceases operations or ceases to exist, and to the potential circulation of counterfeit certificates, which represents a legal challenge for society.

Academic certification should be understood as any type of learning accreditation, from regulated educational programs to the validation of professional competencies and informal training, professional training in companies, online courses, continuing education courses, letters of endorsement, etc. Therefore, it should be understood in a broad sense. For academic certification, the theoretical model must meet the following conditions or premises:

1. It must be possible to use this system without modifying the existing information systems of academic/training institutions or other organizations that issue certificates or academic information and accreditations.
2. It must be possible to verify the validity of the certificate content and associated accreditations even after the dissolution of the institution that originally issued the certificate and all its information systems.
3. It must be usable in all types of education: formal, non-formal, and informal.
4. It must be flexible enough to accommodate different types of academic information, accreditations, etc., regardless of the underlying information models.
5. It must be scalable and capable of supporting transactions and operations related to the registration, verification, and query of information by the model's end users.
6. It must comply with personal data protection regulations, and particularly with the General Data Protection Regulation (GDPR) in cases where it affects individuals located in the European Union, which is the subject of validation in this report.

The model, as already explained, is based on the scientific publication entitled "Application of Blockchain in Education: GDPR-Compliant and Scalable Certification and Verification of Academic Information" [6], which presents the system developed by GAVIN.

This section details the technical architecture designed in the GAVIN project to create a blockchain-based initiative that complies with the GDPR. The design follows a privacy-by-design approach, incorporating both off-chain and on-chain components to manage credential data securely and transparently. The architecture is modular and scalable, aiming to support institutional collaboration through a consortium-based governance model. The following subsections describe the core components and mechanisms that ensure both compliance and operational efficiency.

4.1. Off-Chain Data Storage

To comply with GDPR requirements, particularly the rights to rectification and erasure, the considered model [6] stores academic information on institutional off-chain servers and on blockchain anonymized personal data to verify this academic information. GAVIN, the practical implementation of [6] stores identifiable data off-chain on secure institutional servers and the on-chain proposal is implemented by storing only hashed representations (HMAC-SHA3) [36] of credentials, which are signed by the issuing entity, are stored on the blockchain.

4.2. Blockchain Layers,Access Control and Use Escenarios

The conceptual model is presented in the following figure 1:

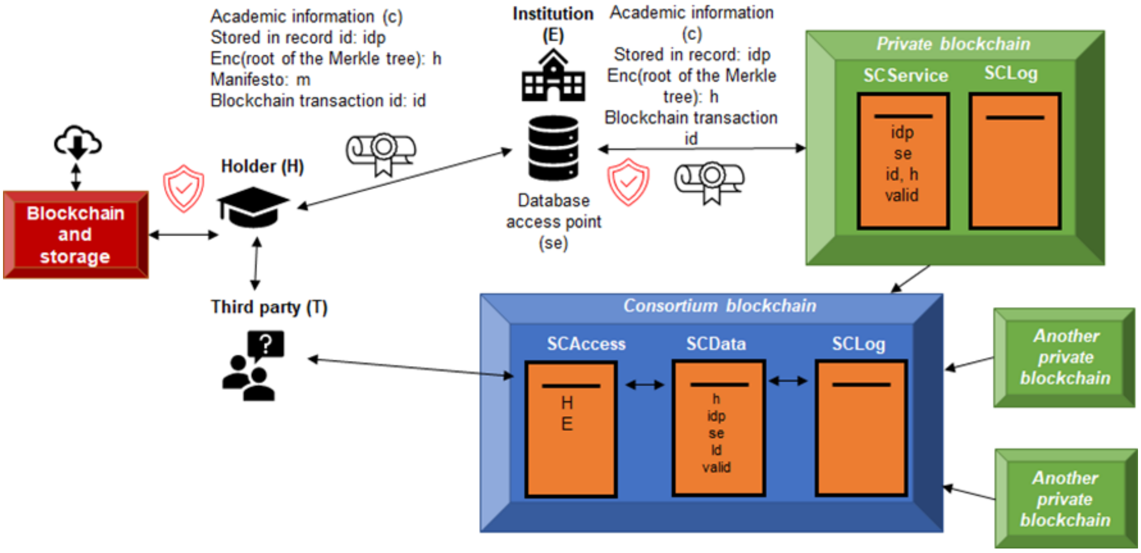

Figure 1. Conceptual model applied in the GAVIN project.

In the proposed model, three main actors are identified: H, the holder of the academic data; E, the entity issuing the certificates; and T, the authorized third-party verifier. Unlike other approaches that transfer all data to the blockchain, in this architecture the original academic information (c) remains stored in the issuing institution's internal systems, thus preserving institutional control over the records. What is generated to verify authenticity is a structured manifest (m), represented as a Merkle tree [37]. This scheme allows the different elements of the certificate to be organized hierarchically and flexibly, adapting to heterogeneous academic data structures.

Each branch of the tree represents an information node in accordance with the issuing entity's data model. In addition to the academic content, additional metadata such as certificate validity, expiration dates, internal storage references, or unique identifiers can be integrated. For each element in the tree, an HMAC (Hash-based Message Authentication Code) value is generated with SHA3, using an individual secret key to reinforce privacy and protect against preimage or length extension attacks. This key, different for each piece of data, is shared between the issuing entity and the data subject. This guarantees both authenticity and the possibility of sharing information in a fragmented manner, without having to expose the entire set. The value resulting from applying HMAC to each element in the tree is encrypted using the issuing entity's private key, thus generating an anonymized version of the data (h), which is stored in a smart contract called SCService, deployed on a private blockchain. This network is accessible through an API, allowing its integration with existing IT systems without the need for structural redesigns. Along with this encrypted information, other non-personal data can be stored, such as internal identifiers, pointers to the original database (se), or validity indicators, giving the model great versatility without compromising privacy. The private blockchains used in this scheme are managed by consortia of institutions, organized according to geographic or functional criteria. The data recorded on these networks can be periodically synchronized with a federated consortium blockchain, which consolidates non-personal information from the different private blockchains. Each recorded transaction is immutably reflected in the chain, preventing retroactive manipulation by members of the participating institutions. Any attempt to alter or delete an academic record without authorization will be traced thanks to the use of the SCLog smart contract, which keeps an unalterable record of all interactions and access.

The proposed system includes three main components:

- **SCData:** A smart contract stored on a consortium blockchain that stores the anonymized values of credentials, implemented using HMACs in the developed prototype.

- **SCAccess:** A smart contract stored on the consortium blockchain that verifies access permissions granted by data subjects. In the current implementation of the project a Blockchain Application Firewall (BAF) [38] controls data queries and enforces user permissions, functioning as a decentralized access gateway, but this element could not be necessarily used in other implementations.
- **SCLog:** A smart contract in charge of registering all the accesses to the information.

Once the verification data (h) is recorded in the consortium blockchain, access is governed by the SCAccess contract, which maintains control over permissions and authorizations for consultation, both by the holder (H) and, where applicable, by the issuing entity (E). The data stored in the SCData contract is always anonymous; it does not allow direct identification of the holder, but it does allow verification of the validity of the linked academic information, provided that T has the corresponding authorization.

The system contemplates various validation scenarios depending on the level of access desired by the holder:

- Scenario 1: Full verification: H sends T the full certificate along with the manifest (m), including all the keys necessary to reconstruct the Merkle tree. T requests the encrypted value (h) from the SCAccess contract, verifies it against SCData, and checks its integrity.
- Scenario 2: Partial verification: H decides to share only part of the academic content. To do so, he transmits to T only the relevant elements and verification paths of the Merkle tree. After receiving the encrypted tree root from SCData (after authorization validated by SCAccess), T can verify that the partial information corresponds to a valid, unrevoked certificate.
- Scenario 3: Direct query to the issuing authority: H delegates the delivery of the information to T to E. In this case, T accesses the database pointer (stored in SCData), and after its authorization is validated in SCAccess, it can obtain the certificate (c) directly from E. This interaction is logged in SCLog.

It should be noted that in scenarios 1 and 2, the system can continue to function even if the issuing entity disappears, since its direct intervention is not required for verification. However, if both E disappears and H loses its keys, data recovery would be unfeasible. To mitigate this risk, the model provides that, in the event of an institution's cessation of activity, academic information will be transferred to a distributed storage system, controlled by smart contracts and encrypted with keys previously shared between E and the authorized data subjects. In this context, if the data subject intentionally destroys their key, or if the issuing institution's key ceases to exist, the stored data would become inaccessible. This loss, being irreversible, could be interpreted as an effective way to comply with the right to be forgotten under the GDPR, given that the information would be permanently encrypted beyond recovery. The procedure for granting access is for the data subject to expressly authorize a third party, identified on the platform, to access the data necessary to verify the academic information. This verification data is stored on the blockchain so that it cannot be accessed without the corresponding permission. Furthermore, the data is hashed with an SHA3 HMAC, and the result is subsequently encrypted.

From a technical perspective, GAVIN implements the model considered in the implementation of a consortium blockchain based on Ethereum technology with several private blockchains, following the model presented in the previous point, and offering the following functionalities:

- Store any type of academic information in a free format.
- Implement smart contracts on the platform.
- Allow the interconnection of blockchains (consortium).
- Implement storage with a blockchain-based access control system, where data is stored encrypted and access is restricted to the data subject.

The theoretical model presented is technology-agnostic, and there are some issues that current technology does not easily address. Therefore, the technical reference model considered meets the requirements of the proposed model, including GDPR compliance, but must also address some challenges.

To complete the system architecture, there will be another blockchain or equivalent with a distributed storage system and an access control mechanism that would serve as a backup in the event that an institution ceases to exist or disappears, and the user does not have their academic certifications. They could be recovered from this system as long as the institution issuing the academic data records them there before disappearing. Access to this information will be controlled, and it will be the data subject's decision to allow the records to be sent to this backup infrastructure in extreme cases in which the institution issuing the academic information disappears. Note that this does not affect the verification of academic data, which can continue to be performed normally with the rest of the designed system. Its usefulness is limited to scenarios in which the issuing institution disappears and, at the same time, the data subject loses the academic information that can be shared with a potential verifier. In the case of the prototype developed at GAVIN, a distributed storage based on MariaDB [39] was used, in which the academic data of certain users will be stored encrypted, with a sufficiently robust encryption system and using a different secret key known to the data holder and generated by the academic institution.

In addition to being encrypted, the data in this distributed database is not directly exposed to the Internet. Instead, a private Ethereum-based blockchain is used to perform effective access control. It is responsible for storing a record identifier mapped to the database table (regid) and the address of the account authorized to access said data (address).

Thus, in the technical implementation, when a given user H wants to retrieve their data, they make a request to the system, which is responsible for checking whether or not they have permission, verifying through a signature from the user themselves that they own a certain blockchain address (address). If so, user H is returned the regid associated with the address (address). Furthermore, during the execution of this access control process, an event is generated that records access to the blockchain for the purpose of retrieving the data. This event is saved in a system access log for retrieving academic certificates. Once the user has obtained the RegID corresponding to their academic data, they can connect to the MariaDB database and request the stored content associated with that RegID. The database returns the stored data in encrypted form, and the user can decrypt it using the key k, which only they know. This approach ensures that, even if the RegID of the encrypted data is known, only the authorized user who possesses the corresponding key k can decrypt and access their academic data.

4.3. Consortium Governance, Interconnection and Scalability

In order to provide guarantee the governance and scalability of the platform based on blockchain, the GAVIN project employs a multi-layered architecture:

- Multiple private blockchains, managed by academic institutions.
- A central consortium blockchain aggregating anonymized information to verify academic information and verifying access rights.

This model ensures scalability, interoperability and resilience in the event of institutional dissolution.

The technical reference model used in GAVIN project is presented in the following figure 2:

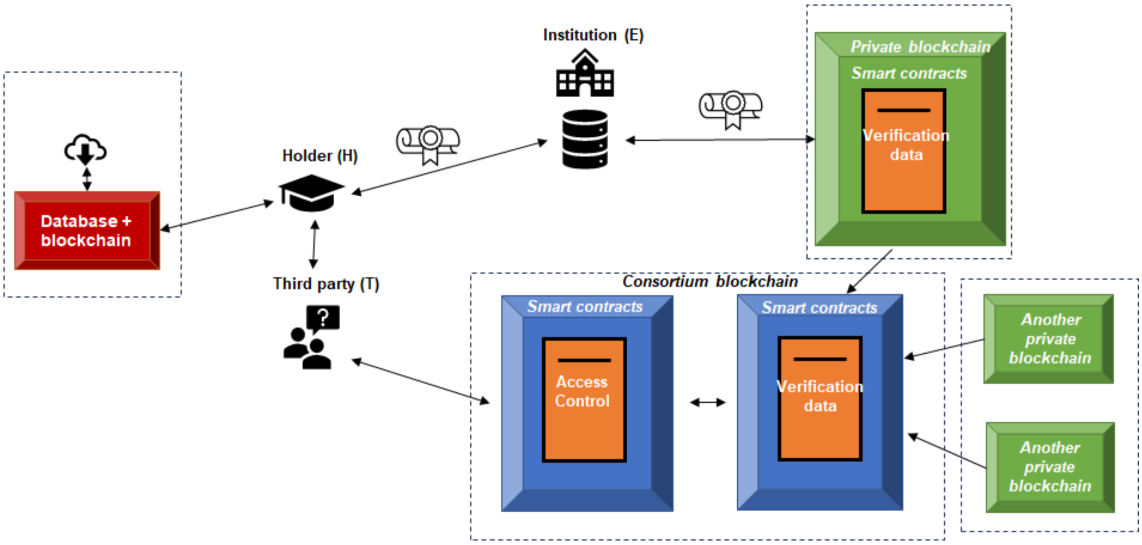

Figure 2. Technical reference model used in the GAVIN project.

The main difference observable between Figure 1, which represents the initially proposed design [6], and Figure 2, which shows the component architecture considered for implementing a prototype using real-world technology, is that the consortium blockchain has been split into two separate blockchains. This change was made because it is not technically feasible to restrict access to data stored on a blockchain (specifically, the data contained in the Access smart contract) if access to that blockchain is granted. Therefore, in the GAVIN project, the consortium blockchain has been divided into two distinct blockchains: one is responsible, through the SCAccess contract, for verifying, granting, and revoking access permissions to academic verification data, while the second blockchain is where the actual academic verification data resides. However, it is important to point out that the current technology for interconnecting these Consortium blockchains has a relevant limitation: when an origin blockchain makes a query to an external blockchain, the result of this interaction is also recorded in A. This behavior represents a challenge in terms of privacy, since it could allow an unauthorized user (U), with read access to the initial blockchain, to view evidence of queries previously made by legitimately authorized third parties (T), potentially exposing protected academic information without U having the right to access it. Since this behavior does not derive from a defect in the proposed model, but from technical restrictions inherent to current blockchain platforms, an intermediate security solution was implemented in the GAVIN prototype: the Blockchain Application Firewall (BAF) [38]. This additional layer is located between the verifying party (T) and the blockchain where the SCAccess is located. In networks such as Ethereum, third parties access the blockchain through the JSON-RPC interface, connecting directly to specific nodes of the system to issue transactions or make queries. This method exposes a critical risk: any entity capable of executing RPC calls could extract information publicly stored on the blockchain, including data previously verified by other users, even without specific authorization. This problem stems from the fact that, by design, many current blockchain platforms restrict modification of stored data, but not its viewing. Some solutions allow for some level of access control, but these are typically limited to per-account permissions rather than at the node level, which is insufficient for the proposed model, which requires more granular access management. Given this situation, GAVIN incorporated the BAF as an intermediary mechanism. This middleware acts as an intelligent filter, evaluating incoming requests before allowing their transmission to the blockchain node. Its operation is based on predefined rules that determine whether a specific request meets the criteria necessary to be accepted. For example, if a third party (T) attempts to access information contained in a specific smart contract within the blockchain where SCAccess is located, the BAF will validate whether T has the necessary permissions and, only if so, will allow the request to pass; otherwise, it will deny it. In this way, the BAF enables flexible, efficient, and adaptable access control, avoiding unnecessary exposure of sensitive data and complying with regulatory requirements such as those established in the GDPR. It also prevents the abusive use of data stored in the blockchain for unauthorized purposes, such as the mass analysis of academic patterns. Furthermore, the model defines that all access requests must be logged. Ideally, this logging would be carried out in a specific smart contract (SCLog) deployed on blockchain A. However, achieving full on-chain traceability would entail considerable overhead in terms of storage and performance, compromising the system's efficiency. As a temporary solution while blockchain technologies evolve in capacity, it has been decided to delegate detailed access traceability to the BAF, using a centralized database such as MariaDB. Access events are stored there, along with all relevant metadata. To maintain the integrity of this information, a periodic notarization mechanism has been incorporated: every 24 hours (or at a specified frequency), a hash digest of the generated logs is calculated, and this value is stored in blockchain A via the SCLog smart contract. This hybrid technique provides an additional layer of trust. In the event of malicious manipulation of the log stored in the BAF database, any alteration would be detectable thanks to the discrepancy between the current hash and the one previously recorded in the blockchain. This guarantees a balance between operational efficiency and robustness in access traceability.

5. Legal Analysis

After analysing the technical model described in the previous section, this one evaluates the GAVIN architecture’s compliance with the GDPR. The analysis applies a doctrinal legal method, examining how each key GDPR principle and data subject right is operationalized within the system. Particular attention is given to the distribution of responsibilities among actors, the feasibility of data subject rights enforcement, and the management of international data flows.

Operations in private and consortium networks require clearly defining the obligations and limits of each participating entity, not only from a functional or technical responsibilities perspective, but also in the processing of personal data. In this regard, the GAVIN model provides for the formalization of joint responsibility agreements in accordance with Article 26 of the GDPR, where specific functions (primary controller, processors, etc.) are assigned and data flows are documented. This contractual framework serves as a reference for other institutional consortia and could be adapted according to standardized contractual templates available in organizations such as the AEPD or the EDPB.

The subsections that follow correspond to the core legal dimensions evaluated during the project’s legal-technical validation process.

5.1. Scope of the GDPR

As a starting point, we must consider the material and territorial scope of application.

Regarding the material scope, it is evident that it falls within the scope of application whenever the project will process personal data contained in or intended to be included in a file.

Regarding the territorial scope, the project also falls within the scope of application whenever the GDPR applies:

- On the one hand, to the processing of personal data when the controller or processor (RT/ET) has an establishment in the EU, regardless of whether the processing takes place in the Union or not.
- On the other hand, to the processing of personal data of data subjects located in the Union by a controller or processor not established in the Union, when the processing activities are related to:
    - the offering of goods or services to said data subjects in the Union, regardless of whether payment is required from them, or
    - the monitoring of their behavior, to the extent that this takes place in the Union.

It is therefore clear that the GAVIN project will also be subject to territorial scope.

5.2. Personal Data Processed by the GAVIN Project

The GAVIN project aims to use a blockchain to store academic certifications (in the broadest sense as explained above), which is considered data processing, as defined in Art. 4.1 GDPR as any operation or set of operations performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or modification, extraction, consultation, use, communication by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. It is evident that an academic certification will contain personal data, within the meaning of Art. 4.1 GDPR defines them as "any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is any person whose identity can be determined, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person." Therefore, it is clear that these certificates may contain, but are not limited to, the following data, including personal data:

- Registration Identifier (Code).
- University Registration Identifier (Code).
- National Registration Identifier (Code).
- Surname.
- First Name.
- Identity Document.
- Gender.
- Date of Birth.
- Country of Birth.
- City of Birth.
- Country where the student currently resides.
- Student's current city of residence.
- Postal Code.
- Student's email address.
- Student's phone number.
- Type of study.
- Degree.
- Area of specialization.
- Study ID.
- Issuing Entity ID.
- Start date of studies.
- End date of studies.
- Hours dedicated to completion.
- Number of academic credits (ECTS or equivalent).
- Language of study.
- Program subjects and content description.

Furthermore, as established by the Spanish Data Protection Agency, when a user of a blockchain-based system is a natural person, the public key or wallet address derived from the public key is also personal data, as is any unique identifier that can link that person's activity within the blockchain infrastructure. Furthermore, we understand that under no circumstances will it be necessary, and therefore, special categories of personal data will not be processed. These are defined in Article 9.1 of the GDPR, such as those that reveal ethnic or racial origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data intended to uniquely identify a natural person, data relating to health, or data relating to a natural person's sex life or sexual orientation. The processing of this type of data is prohibited, unless there is an exception to its processing as set forth in Article 9.2 of the GDPR. Likewise, we do not believe that the processing of data relating to criminal convictions and offenses or related security measures is necessary (Article 10 GDPR), as the processing of such data is also limited to being carried out under the supervision of public authorities or when authorized by Union or Member State law that provides appropriate safeguards for the rights and freedoms of data subjects. Furthermore, a complete record of criminal convictions may only be kept under the supervision of public authorities.

5.3. Implications for Compliance with the Principles of Article 5 of the GDPR

Article 5 of the GDPR sets forth the principles under which all personal data processing must be conducted, which constitute the core of this regulation. The following sections analyse the GAVIN project's compliance with each of these principles.

5.4. Data Flows in the Model and Characterization of the Participants as Data Controllers, Processors, or co-Controllers

The GDPR defines in Article 4.7 the "controller" or "data controller" as the natural or legal person, public authority, agency, or other body that, alone or jointly with others, determines the purposes and means of processing. If Union or Member State law determines the purposes and means of processing, the controller or the specific criteria for its appointment may be established by Union or Member State law.

It is, therefore, the key entity in a processing operation, as it tells us "who" processes personal data, and is therefore a sine qua non requirement for all data processing.

Thus, Article 5 establishes that the data controller (hereinafter the "TR") will be responsible for compliance with the principles established in this Article 5 and, furthermore, must be able to demonstrate this ("proactive accountability"). Therefore, compliance is not enough; we must also have evidence of such compliance. Given the configuration of the GAVIN project, already described as a consortium blockchain, we can consider the possibility of several data controllers acting jointly, determining the purposes and means.

Therefore, in this case, we must take into account the provisions of Article 26 of the GDPR, which states:

1. Where two or more controllers jointly determine the purposes and means of processing, they shall be considered joint controllers. The joint controllers shall determine, in a transparent manner and by mutual agreement, their respective responsibilities for compliance with the obligations imposed by this Regulation, in particular with regard to the exercise of data subject rights and their respective obligations to provide information referred to in Articles 13 and 14, unless and to the extent that their respective responsibilities are governed by Union or Member State law to which they apply. Such agreement may designate a contact point for data subjects.
2. The agreement referred to in paragraph 1 shall duly reflect the respective roles and relationships of the joint controllers in relation to the data subjects. The essential aspects of the agreement shall be made available to the data subject.
3. Regardless of the terms of the agreement referred to in paragraph 1, data subjects may exercise their rights under this Regulation against and against each of the controllers.

In this regard, we must emphasize that the joint controllers of the data processing and the blockchain platform must formalize a joint data processing agreement, without prejudice to the processing carried out by each institution on its blockchain, for which, in principle, they will be the sole controllers.

On the other hand, it must also be taken into account that the blockchain platform or platforms described in this project will require an information systems service provider, most likely in the cloud, which may follow a PAAS, IAAS, or SAAS model. In any case, we must remember that if the controller or joint controllers use a third party to provide these services (e.g., Amazon Web Services, Azure, etc.), personal data will be processed by that third party on behalf of the controller(s), a figure known as a data processor. This consideration is still under discussion and there is no definitive position on the matter from the Supervisory Authorities at the European level.

For these purposes, Art. Article 28 of the GDPR requires that the RT select a data processor that offers sufficient guarantees to implement appropriate technical and organizational measures to ensure that the processing complies with the requirements of the GDPR and guarantees the protection of the data subject's rights. Furthermore, a data processing agreement or contract must be formalized between the data controller and the data processor, regulating the purpose, duration, nature and purpose of the processing, the type of personal data and categories of data subjects, and the obligations and rights of the controller, as well as any subcontracting, in accordance with the minimum content established in Article 28.3 of the GDPR..

5.5. Application of GDPR Principles

The legal analysis of the GAVIN project validates compliance with the main GDPR principles:

- Lawfulness, Fairness, and Transparency (Art. 5.1.a): Clear privacy policies, user-informed consent, and transparent data processing.
- Purpose Limitation (Art. 5.1.b): Academic certification as the sole purpose, no other purposes being contemplated at present.
- Data Minimization (Art. 5.1.c): Only essential data is processed and stored off-chain. The data processed is only the HMAC (as recommended by AEPD [40]) and the final result is additionally encrypted before being saved on the blockchain.
- Accuracy (Art. 5.1.d): Regarding this right, the GAVIN project model can fulfill this right by deleting the HMAC of the previous academic certificate, rendering the data arbitrary, which would effectively amount to deletion. A new academic certificate could subsequently be generated with the correct data. In any case, it is worth mentioning that the issuing institution (I) should delete the data from its off-chain systems if permitted by law. It also marks the certificate as invalid, and access control mechanisms ensure that no one can access it, in addition to deleting all related information.
- Storage Limitation (Art. 5.1.e): Data remains available only while necessary or archived under public interest exceptions.
- Integrity and Confidentiality (Art. 5.1.f): According to this article of GDPR, personal data must be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction, or damage, by applying appropriate technical or organizational measures ("integrity and confidentiality"). The principle of data integrity is fulfilled by the blockchain technique. Furthermore, we must note that confidentiality is also guaranteed, since the academic transcript data is not visible on the blockchain since an HMAC function has been applied to it and it has been subsequently encrypted. The developed prototype implements a firewall, specifically the so-called "BAF" (Blockchain Application Firewall) [38], which prevents unauthorized access to verification data, although it is true that it is an element outside the consortium blockchain. The BAF performs filtering and access control, verifying which resources can be accessed within SCAccess. In this way, any user with access to said blockchain is prevented from viewing data requested by other entities, thus limiting access to information for which they do not have permissions. This is because with current blockchain technology it is not possible to perform this dynamic segmentation of permissions between accounts considered in the proposed model, so this perfectly valid technological solution is used for its real implementation in the reference technical model. In summary, encryption, access logs, and tamper-proof chains ensure data security.
- Accountability (Art. 5.2): The GAVIN project has defined and implemented appropriate technical and organizational measures to ensure and demonstrate that the processing complies with this Regulation, since, as explained, a robust pseudonymization or anonymization system is in place. Therefore, the impact of an attack or security breach would be low, given that the data stored in

the blockchain is the result of several HMAC generation processes with different keys and has been subjected to an encryption process. Furthermore, the BAF mechanism is in place, which provides additional security. It should be noted that the prototype records access notarization information (a hash summary of the log recorded in the BAF to potentially detect tampering) on a blockchain (equivalent to SLog, defined in the model) so that traceability of what has happened to their data can be maintained. Likewise, the holder can always check in SCAccess who has access permission to their academic information, and the BAF records the accesses and periodically notarizes them on a blockchain to detect potential tampering. In this way, the holder could check who has accessed specific academic information and when. This is done because current blockchain technology does not allow reading operations to be recorded, only those that modify data. This is how this article is fulfilled in the GAVIN Project.

5.5. Privacy by Default

Article 25 of the GDPR establishes that the data controller must adopt appropriate technical and organizational measures to effectively achieve compliance with the data protection principles and other requirements of the GDPR, with particular attention to the protection of the rights of data subjects.

These measures outlined in the GDPR are security measures, which may be technical, organizational, or other, but must be effective.

Logically, these security measures must be applied based on the state of the art, the cost of implementation, and the nature, scope, context, and purposes of the processing, as well as the risks of varying likelihood and severity that the processing entails for the rights and freedoms of natural persons. Consequently, security measures must be applied after a risk assessment.

Furthermore, these measures must ensure that they are applied in the very design of the procedure, that is, before its implementation. In this sense, we can affirm that the mere purpose of this report is already evidence of compliance with the principle of security by design. Furthermore, both the theoretical design and the technical reference model were also supervised by a legal team.

On the other hand, the data controller must also ensure that, by default, only personal data necessary for each of the specific processing purposes is processed. By "necessary data," we mean not only the amount of data, but also the duration of processing (retention period) and access to said data, which must be restricted.

Regarding the principle of privacy by default, we understand that it is fulfilled for the following reasons:

- As explained, both the conceptual theoretical model and the prototype developed at GAVIN comply with the principles of the GDPR, including minimization, retention limitation, and purpose limitation, as already explained.
- Regarding access, the model restricts access to only authorized users to consult the information, which is also implemented with the BAF in the reference technical model.
- Access to data by service providers is managed by users, who can choose when to provide such access and can restrict it at any time. Furthermore, all access to the blockchain is recorded.

On the other hand, the system developed to recover academic information to the owner in the event of its loss or if the issuing institution disappears could consist of distributed storage with a blockchain-based access control system, where data is stored encrypted and access is restricted to the owner. This would be applicable when the issuing institution disappears.

In particular, to solve this requirement in this project, a distributed storage system based on MariaDB will be used, in which the academic data of certain users will be stored encrypted, with a sufficiently robust encryption system and using a different secret key known by the data owner (holder) and generated by the academic institution.

In addition to being encrypted, the data in this distributed database is not directly exposed to the Internet. Instead, a private blockchain based on Ethereum technology is used to implement effective access control. It stores a mapping between a record identifier and the database table (regid) and the address of the account authorized to access said data (address).

Thus, in the technical implementation, when a given user H wants to retrieve their data, they make a request to the smart contract, which is responsible for checking whether or not they have permission, verifying through a signature from the user themselves that they own a specific blockchain address (address). If so, user H is returned the regid associated with the address (address). Furthermore, during the execution of this access control process, an event is generated that records access to the blockchain for the purpose of retrieving the data. This event is saved in a system access log for retrieving academic certificates. Once the user has obtained the RegID corresponding to their academic data, they can connect to the MariaDB database and request the stored content associated with that RegID. The database returns the stored data in encrypted form, and the user can decrypt it using the key k, which only they know.

This approach ensures that, even if the RegID of the encrypted data is known, only the authorized user who possesses the corresponding key k can decrypt and access their academic data. As with other parts of the system already described, it is not possible to dynamically restrict the data an external user who wants to obtain data from the private blockchain to implement the solution will have access to. Therefore, once again, a BAF is used to restrict which data on the private blockchain can be accessed externally (for example, only to a specific smart contract where the "address -> regid" mapping is stored), ensuring that the system will function without the possibility of additional data being leaked, such as transactions by academic institutions where new address-regid pairs are specified for the mapping.

The BAF itself, after verifying that the user is authorized, can retrieve the information from MariaDB and return it to the user, centralizing access control and data retrieval in this system. Furthermore, the BAF also notarizes each time the user accesses the data to download and periodically saves a hash summary of this information on the private blockchain. In this way, privacy and access control are guaranteed through the combined use of cryptography, a BAF, and smart contracts for access control and recording. In our opinion, and taking into account the current state of the art, this allows compliance with the GDPR principles of "secure by design" and "privacy by default." Additionally, the destruction of the user's k-key used to encrypt the original academic data converts the information into a non-recoverable sequence of information, which could also be deleted from MariaDB if requested and permitted by law.

5.8. Data Subject Rights

The project GAVIN fully supports the exercise of data subject rights:

- **Right to Access (Art. 15)**: The model's design allows for automatic access by the user (holder) at any time, without the need for third-party intervention. Furthermore, access to certification records that may be stored on the blockchain by third parties other than the data subject must be authorized by the data subject, who can also revoke this authorization at any time. In this sense, we can clearly infer that the GAVIN project is ready to fulfill this right.

- **Right to Rectification (Art. 16)**: This right can be fulfilled in two ways, as explained in the project:

    - One way to fulfill this right would be to remove the keys used to generate the HMACs for the different fields, meaning that the data stored in the blockchain would become arbitrary and dissociated, and therefore could not be associated with a holder. From this point on, the new information could be uploaded to the blockchain, meaning the academic certification would be "rectified" and the data would be accurate, thus complying with the principle of accuracy.

  - On the other hand, the project's blockchain could also fulfill this right by generating the new information and linking it to the previous information using the transaction identifier or other data.

- **Right to Erasure (Art. 17):** It is possible to exercise this right by deleting the keys and data used to generate the HMACs of the Merkle tree, whose signed HMAC was stored in the blockchain, such that the data stored in the blockchain now becomes arbitrary and dissociated data, without the possibility of recovering the original data from it (and access is withdrawn from everyone in SCAccess and, in addition, it is marked as invalid).
- **Right to Portability (Art. 20):** The Implementation in the reference technical model of the project GAVIN would also guarantee this right, as it allows the download of information in a structured, commonly used, and machine-readable data format.
- Right not be subject to a decision based solely on automated processing. (Art. 22): This right would not apply since the proposed model does not consider automated processes on or off-chain. The model proposes that, for the issuance of a new certificate, the issuing institution is responsible for executing this issuance. Furthermore, the interested party is the only one who can grant or revoke access permissions to their information. Finally, to verify data, the verifying entity is the one who actively interacts with the blockchain. Automated decisions are not made, nor are data subject profiles created. This is not applicable since there is no automated decision-making by either the promoters of the initiative or third parties, as mass access to user data is not possible, as the theoretical model and the technical implementation of the prototype have adequate access control mechanisms. Specifically, in the theoretical model, the record is made in the SCLog, while in the technical reference model, the record is achieved with the BAF, since read accesses are not recorded in the blockchain.

5.6. Roles and Responsibilities

The model proposed in the GAVIN project identifies data controllers (institutions), data subjects (students), and processors (cloud providers). For consortium governance, joint controller agreements are necessary, as per Art. 26 GDPR.

5.7. Cross-Border Data Flows

It should also be noted that the GDPR assumes the free movement of personal data within the European Economic Area. However, any export of data to third countries will be considered an international data transfer.

In this regard, Article 44 of the GDPR establishes that transfers of personal data that are being processed or will be processed after their transfer to a third country or international organization will only take place if, subject to the other provisions of this Regulation, the controller and processor comply with the conditions set out in the GDPR, including those relating to onward transfers of personal data from the third country or international organization to another third country or international organization.

First, if data is to be transferred to third countries (e.g., the use of a cloud service hosted abroad), it is necessary that said country guarantee an adequate level of protection. GAVIN's analysis suggests that a consortium or private blockchain will be used. Therefore, these blockchains require validators and nodes to be physically located within the European Economic Area (EEA) or in one of the countries for which an adequacy decision has been issued, such as those mentioned above. Consequently, the fact that these are private and consortium blockchains means that the entities in charge oversee compliance with the network's conditions and policies.

In any case, if an international data transfer (IDT) were to be carried out within the framework of the GAVIN project or another project using this model, one of the aforementioned instruments would be necessary to comply with the GDPR.

6. Discussion

The implementation of the GAVIN model and its practical implementation offer valuable insights into the challenges and opportunities of reconciling blockchain systems with demanding data protection regulations. This section reflects on the project's key technical and legal contributions, identifies remaining obstacles, and considers broader implications for policy, standardization, and institutional adoption. By positioning GAVIN within ongoing regulatory and technological debates, we aim to assess its potential as a replicable framework beyond the educational domain

6.1. Innovations and Contributions

The GAVIN project represents a pioneering implementation where GDPR compliance is not an afterthought but a foundational criterion. Its use of off-chain storage, HMAC encryption, and BAF middleware to improve security and even adds logging features exemplifies privacy-by-design principles (Art. 25 of GDPR).

6.2. Technical and Legal Challenges

- **Inmutability vs. Erasure:** This challenge is achieved through deletion of off-chain data and hash revocation.
- **Shared Responsibility:** The final deployment of the model proposed by the GAVIN project requires robust legal frameworks among consortia.
- **International Hosting:** It is clearly an ongoing risk due to limited global GDPR harmonization.
- At the legal level, GAVIN also successfully addresses the complexities arising from co-responsibility in consortium networks. The proposal to formalize co-responsibility agreements between participating institutions and the correct identification of roles in data processing constitute a solid starting point that can be replicated in other institutional contexts.

However, certain limitations remain that deserve to be highlighted. These include the dependence on hybrid solutions (off-chain/on-chain) and centralized infrastructures such as the BAF for access traceability. Although these decisions are technically justified, they reveal that blockchain technology, in its current state, does not yet offer a completely autonomous and functionally self-sufficient solution to meet the necessary requirements. However, the overall result fulfills its objective of being compatible with data protection, as well as supporting all types of academic certification, scalability, and, through the use of APIs, not requiring structural changes to organizations' information systems to adapt and integrate with GAVIN.

6.3. Policy and Standardization Implications

Wider adoption of models like GAVIN requires standardization (e.g., ISO/IEC blockchain frameworks) and updated guidance from regulatory bodies tailored to decentralized contexts.

7. Conclusions

The GAVIN project affirms the viability of GDPR-compliant blockchain architectures for academic certification. By combining technological innovation with rigorous legal analysis, it offers a scalable and transferable model for trusted digital credentials that complies with the principles of the GDPR through a balanced combination of cryptographic techniques, distributed storage, smart contracts, and blockchain-based access control mechanisms, achieving a system that is not only compatible with the model, but also a functional, scalable, and legally robust solution for the issuance, validation, and even retrieval of academic information.

GAVIN not only responds to a real, latent need in today's digital society to verify any type of academic information, but does so based on a design that aligns with the data protection principles established in Article 25 of the GDPR. This "privacy by design" approach sets a relevant precedent for the development of decentralized systems in other sectors with high regulatory requirements, such as healthcare, justice, and finance, to name a few. Despite the progress made, the path toward standardizing blockchain architectures compatible with European data protection regulations still presents challenges. Legal interoperability between jurisdictions, the technological evolution of blockchain platforms, and the lack of specific guidelines from regulatory authorities remain open areas for research and innovation, along with, for example, automated compliance tools and AI-assisted credential verification.

Ultimately, GAVIN represents a necessary, real, and innovative proposal (in fact, both the model considered and the prototype created can be said to be academically pioneering) and can undoubtedly contribute significantly to the digital transformation of the educational ecosystem, offering a secure and GDPR-compliant alternative for the issuance, verification, and even retrieval of academic information in any format.

Author Contributions: Conceptualization, A.G.V., C.D.-v.-E and D.E.G.; methodology, A.G.V., and D.E.G.; software, A.G.V. and C.D.-v.-E.; validation, A.G.V. and D.E.G. and Z.Z.; formal analysis, A.G.V. and D.E.G.; investigation, A.G.V. and D.E.G.; resources, A.G.V., C.D.-v.-E and D.E.G.; data curation, A.G.V., C.D.-v.-E and D.E.G.; writing—original draft preparation, A.G.V. and D.E.G.; writing—review and editing, A.G.V., C.D.-v.-E and D.E.G.; visualization, A.G.V., C.D.-v.-E and D.E.G.; supervision, A.G.V. and D.E.G.; project administration, A.G.V., C.D.-v.-E and D.E.G.; funding acquisition, C.D.-v.-E. All authors have read and agreed to the published version of the manuscript.

Funding: Grant TED2021-130828B-I00 funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

Acknowledgments: We would like to express our gratitude to the members of the GAVIN team at atlanTTic for their inspirational comments and insight.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. W. Gräther, S. Kolvenbach, R. Ruland, J. Schütte, C. F. Torres, and F. Wendland, “Blockchain for Education: Lifelong Learning Passport,” *1st ERCIM Blockchain Work. 2018, Reports Eur. Soc. Soc. Embed. Technol.*, no. 10, pp. 1–8, 2018, doi: 10.18420/BLOCKCHAIN2018_07.
2. O. S. Saleh, O. Ghazali, and M. E. Rana, “Blockchain based framework for educational certificates verification,” *J. Crit. Rev.*, vol. 7, no. 3, pp. 79–84, 2020, doi: 10.31838/jcr.07.03.13.
3. 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.
4. S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” 2008, doi: 10.2139/ssrn.3440802.
5. T. Lyons, L. Courcelas, and K. Timsit, *Blockchain and the GDPR*. The European Union Blockchain Observatory and Forum, 2018.
6. 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.
7. P. Voigt and A. Von dem Bussche, *The EU General Data Protection Regulation (GDPR)*. Springer Cham, 2017.
8. F. Casino, T. K. Dasaklis, and C. Patsakis, “A systematic literature review of blockchain-based applications: Current status, classification and open issues,” *Telematics and Informatics*, vol. 36. Elsevier Ltd, pp. 55–81, Mar. 01, 2019, doi: 10.1016/j.tele.2018.11.006.
9. S. Khezr, M. Moniruzzaman, A. Yassine, and R. Benlamri, “Blockchain Technology in Healthcare: A Comprehensive Review and Directions for Future Research,” *Appl. Sci.*, vol. 9, no. 9, 2019, doi: 10.3390/app9091736.
10. D. A. Zetzsche, D. W. Arner, and R. P. Buckley, “Decentralized Finance,” *J. Financ. Regul.*, vol. 6, no. 2, pp. 172–203, 2020, doi: 10.1093/jfr/fjaa010.

11. X. Wu, Z. Zou, and D. Song, "Ethereum Tools and Frameworks," in *Learn Ethereum: Build your own decentralized applications with Ethereum and Smart Contracts*, Packt, 2019, pp. 294–335.
12. HYPERLEDGER, "Whitepaper Introduction Hyperledger," *July 2018*, 2018. https://www.hyperledger.org/learn/white-papers.
13. N. Szabo, "Formalizing and Securing Relationships on Public Networks," *First Monday*, vol. 2, no. 9, 1997, doi: 10.5210/fm.v2i9.548.
14. A. Grech and A. F. Camilleri, "Blockchain in Education," Publications Office of the European Union, Luxembourg, 2017. doi: 10.2760/60649.
15. M. Sharples and J. Domingue, "The blockchain and kudos: A distributed system for educational record, reputation and reward," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9891 LNCS, pp. 490–496, doi: 10.1007/978-3-319-45153-4_48.
16. M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić, "EduCTX: A blockchain-based higher education credit platform," *IEEE Access*, vol. 6, pp. 5112–5127, 2018, doi: 10.1109/ACCESS.2018.2789929.
17. 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: 10.3390/app9122400.
18. 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.
19. 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.
20. 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.
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.
22. 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.
23. H. Yumna, M. M. Khan, M. Ikram, and S. Ilyas, "Use of Blockchain in Education: A Systematic Literature Review," vol. 11432, N. T. Nguyen, F. L. Gaol, T.-P. Hong, and B. Trawiński, Eds. Springer, Cham, 2019, pp. 191–202.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.

31. N. A. Malibari, "A Survey on Blockchain-based Applications in Education," in *2020 7th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2020, pp. 266–270, doi: 10.23919/INDIACom49435.2020.9083714.
32. F. Molina, G. Betarte, and C. Luna, "A Blockchain Based and GDPR-compliant Design of a System for Digital Education Certificates," *Clei Electron. J.*, vol. Vol. 26 No, 2023, doi: 10.19153/cleiej.26.1.3.
33. 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.
34. M. Finck, "Blockchain and the General Data Protection Regulation. Can distributed ledgers be squared with European data protection law?," 2019. doi: 10.2861/535.
35. Agencia española de protección de datos (AEPD), "Proof of concept Blockchain and the right to erasure," 2024. [Online]. Available: https://www.aepd.es/guias/Tech-note-blockchain.pdf.
36. H. Krawczyk, M. Bellare, and R. Canetti, "HMAC: Keyed-hashing for message authentication," 1997.
37. 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.
38. C. Delgado-von-Eitzen, M. J. Fernández Iglesias, L. E. Anido Rifón, and F. A. Mikic-Fonte, "Blockchain Beyond Immutability: Application Firewalls on Ethereum-Based Platforms," *Available SSRN 5094133*.
39. E. Kenler and F. Razzoli, *MariaDB Essentials*. Packt Publishing Ltd, 2015.
40. 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.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.