AI resources governance with OpenDID: Strategy & Roadmap

Lenin Chacón
Tecnológico de Costa Rica
San Jose, Costa Rica
lenchaber@outlook.com

Kevin Moraga
Tecnológico de Costa Rica
Heredia, Costa Rica
kmoraga@tec.ac.cr

ABSTRACT

The accelerated adoption of artificial intelligence (AI) in academia has created new demands on computing infrastructure, pushing institutions to rethink how resources are allocated. This paper proposes a governance model for fair and efficient access to university-owned GPU clusters, grounded in decentralized identity (DID) and verifiable credentials (VC). Drawing from scheduling theory in operating systems, we outline how eligibility proofs—such as GPA, coursework, or research needs—can be used to assign priority in resource queues. We present a roadmap for implementation within a Latin American university context, with an initial prototype already under development. Our approach aims to reduce friction in resource access while reinforcing institutional policies on equity and transparency. By aligning DID-based authentication with cryptographic proofs and scheduling logic, we offer a scalable strategy that could serve as a reference for other educational environments seeking responsible AI infrastructure governance.

CCS CONCEPTS

• **Security and privacy** → **Authentication**; **Access control**; • **Computer systems organization** → *Distributed architectures*; • **Social and professional topics** → *Computing and business management*.

KEYWORDS

Artificial Intelligence (AI), Decentralized Identifiers (DID), Self-Sovereign Identity (SSI), Verifiable Credentials (VC), Zero-Knowledge Proofs (ZKP), Batch Scheduling, Equitable Access

1 INTRODUCTION

Access to artificial intelligence (AI) resources for scientific research and the development of proprietary models has become one of the main challenges for universities around the world. Conventional processors lack the necessary resources to train these models, so specialized processors (GPU/TPU) must be used, in addition to a technical support infrastructure that can be costly in terms of initial investment, energy consumption, and concurrent use. Therefore, for Latin American institutions, which often have limited budgets, ensuring access to these tools for their students and researchers is a significant challenge that must be addressed strategically.

The current situation evokes a long period in the history of Costa Rican computing. It was in 1968 [5] when the first computer arrived on Costa Rican soil thanks to IBM. At that time, costs were high and equipment was scarce, so, given this inconvenience, equipment use was centrally assigned in batches. Users handed in their punch cards and waited for them to execute sequentially. This dynamic was appropriate for the constraints of the time, but it's not far removed from the current environment with AI resources: a scarce and expensive commodity that must be managed with safety and efficiency criteria.

More than five decades later, the challenge seems to reappear like a ghost: how to strategically distribute AI computing resources fairly and efficiently to drive academic research and innovation.

2 AI COMPUTING COSTS

The training of AI models involves analyzing costs from two perspectives: local infrastructure and cloud services, each with distinct financial implications.

In a local environment, the cost per CPU or GPU clock cycle is composed of three factors: energy consumption, hardware depreciation, and operational expenses (technical staff, cooling, physical space, etc.). For example, a high-end GPU card with a consumption of 250 W, running intensively for 10 hours, may require more than 3.5 kWh, in addition to the average electricity cost, which results in a significant expense. Therefore, when considering the initial investment in equipment, under continuous operation, the cost should be amortized over a period of three to five years. In the cloud, the cost is calculated per hour of use. Currently, reserving a next-generation GPU unit (such as the NVIDIA H100) can range between 3–4 USD per GPU hour in reserved mode and exceed 30 USD per node hour in on-demand mode. Considering that training a medium-sized language model can require thousands of compute hours, final costs can exceed tens of thousands of dollars [1].

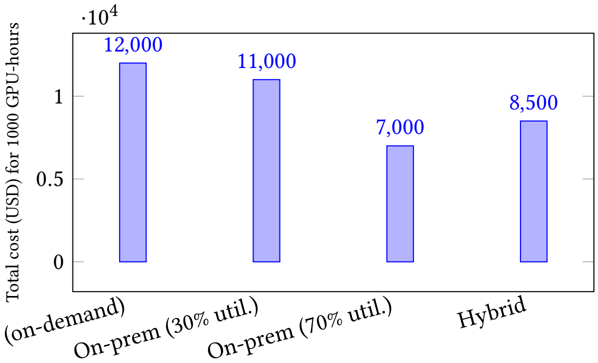

Figure 1: Illustrative cost comparison for a 1000 GPU-hour across different infrastructures. The values are representative and may vary depending on the use case.

From the detailed cost analysis, it becomes evident that neither cloud nor local infrastructure offers an ideal scenario. On one hand, the cloud may be convenient for experimental and short-duration workloads, while a solid in-house infrastructure is justified when usage is continuous and projects high levels of utilization. Thus, while infrastructure is an important consideration, effective governance of access remains the central axis for optimizing AI model training costs.

3 OPENDID, SSI Y VERIFIABLE CREDENTIALS

To implement solid governance, reliable identification and verification mechanisms are essential. This is where the cornerstone comes into play: decentralized identity (DID) and verifiable credentials (VC), technologies promoted by the W3C [13][14].

The proposal is to issue a digital student ID card that acts as a verifiable credential. This card, which will be published by the Instituto Tecnológico of Costa Rica (TEC), contains relevant information such as grade history, certified courses, enrollment status, and more. Additionally, these mechanisms have other advantages such as selective disclosure and zero-knowledge proofs (ZKPs), where students could be more cautious with the information they choose to share, choosing to share only what is necessary, for example:

- ”I have passed the Fundamentals of AI course.”
- ”My GPA is higher than 85.”

The following figure represents a typical case of a trust triangle and how it would look in the current scenario.

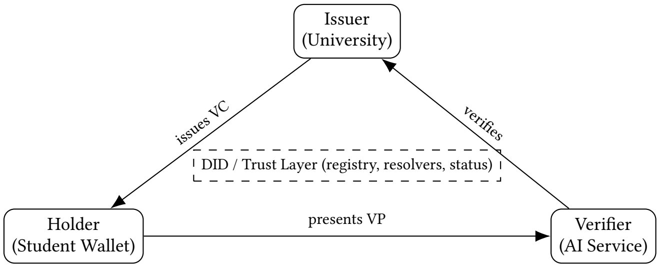

Figure 2: SSI trust triangle with a DID/Trust layer anchoring identifiers, keys, and status.

Several recent studies have explored the use of decentralized identity for authentication and access control beyond financial contexts. Park et al. proposed *OmniOne*, a blockchain-based SSI platform integrated with OpenID Connect for cloud authentication [9], while Turkanović et al. introduced *EduCTX*, a blockchain-based platform for issuing and verifying academic credits [12]. These approaches demonstrate the feasibility of adopting DID/VC frameworks for secure and auditable resource governance within academic environments.

3.1 Actors in decentralized digital identity

Under this proposal, the main actors in decentralized digital identity are three: the issuing institution (**Issuer**), in this case the TEC, which grants verifiable credentials [14] to students; the student now acts as the **Holder**, storing the credentials in their digital wallet; and finally, the **Verifier**, which in this case would be the verification process for accessing AI resources or specialized computing processing.

This model is an alternative that seeks to avoid relying on centralized authentication repositories, since validation is performed through verifiable evidence against credentials issued and signed by the institution. This demonstrates that trust does not depend on a closed database like a traditional system, but rather on a decrypted digital identity ecosystem.

3.2 OpenDID and the blockchain-based trust layer

The component that would act as a ”trust layer” is known as OpenDID [9][13][14], a framework that enables the interoperable management of decentralized identifiers (DIDs). OpenDID establishes mechanisms for registering, resolving, and verifying identifiers, enabling different applications to efficiently trust the origin and veracity of credentials. This, following the example, could be the resource allocation scheduling system.

At this point, a permissioned blockchain like Hyperledger Fabric or Besu, plays a leading role: it does not store personal information or academic records, but provides guaranteed immutability and traceability of multiple DID records, associated keys, and credential status (valid, suspended, or revoked). In this way, the blockchain acts as a trusted public registry managed by smart contracts, while the student would remain in control of their credentials in their digital wallet [10].

This permissioned architecture aligns with consortium governance models typical in academic or governmental networks, where participating institutions jointly maintain validator nodes under defined trust and access policies. Such a setup provides a balance between transparency and operational control, ensuring that identity registries and credential states remain verifiable without exposing private student data.

3.3 Latin American context and digital sovereignty

This discussion of decentralized digital identity is not unique to the Latina population.

Locally, citizens is highlighted [3] [4] the need for public institutions to adopt and integrate identity models that add technological sovereignty to citizens, models that increase their control over their data.

This approach is directly and clearly linked to sovereign identity models [10], where individuals are the primary custodians of their personal information, where validation occurs without intermediaries.

Applied to the context of the TEC, this approach means that students would not only be entitled to a digital ID, but would be the primary administrators of their own identity attributes, deciding how and with whom that information is shared. It’s an ideal scenario, where students are at a solid age to recognize the virtues of this new approach, which would provide almost immediate feedback in an increasingly digitalized era.

3.4 Access governance and scheduling theory

In operating systems, resource allocation to processes is addressed by **scheduling algorithms** [11]. This algorithm describes multiple approaches, among which batch algorithms and priority algorithms stand out. The priority scheduling algorithm determines that tasks with the highest assigned value are executed before those with the lowest value, ensuring that resources are used according to the criteria defined by the system policy.

Applied in a university context, this perspective allows us to imagine a scenario where the use and access to GPUs is managed as if they were processes in a queue. Students who meet certain criteria, such as maintaining a high GPA or passing specific AI-related courses, receive higher priority in the execution queue. In this way, instead of depending on an administrative process, access is determined automatically, fairly, and verifiably, replicating the logic of operating systems.

In practical deployments, this governance layer can be integrated with existing HPC schedulers such as Slurm [7] or PBS [6]. In such cases, the DID/VC-based verification service would issue authenticated job tokens or priority weights that are consumed by the scheduler's native queueing mechanism (e.g., FairShare or PriorityQueue). This design enables interoperability with existing cluster infrastructures while introducing verifiable and privacy-preserving access policies.

3.5 AI resource planning based on verifiable credentials

Now, following the university scenario, this approach allows us to imagine an ecosystem where access to GPUs is managed as if students were processes in a queue. Each student could present their CV, demonstrating their eligibility level—for example, a high GPA, having taken an advanced AI course, or a justified research need — which the scheduler would take into account when assigning a priority level in the queue.

In this way, an administrative process would be automated, making the process more fair, auditable, and decentralized. The logic of operating systems is quite similar: an "identity" with certain attributes is presented, and the system assigns the task according to the defined policies.

The parallel is clear: Just as in the 1960s [5] computing time had to be carefully allocated to optimize a scarce resource, today university policy is required to support the efficient use of AI resources under a transparent and equitable framework.

3.6 Priority Function Definition

To guarantee equitable access, the scheduling mechanism must translate the student's verifiable attributes and the estimated resource usage time into a measurable priority score. Inspired by classical scheduling algorithms from operating systems [11], each GPU request is associated with a value P, calculated as a weighted combination of verifiable credentials and estimated duration W:

$$P = w_1 \times G + w_2 \times C + w_3 \times R \tag{1}$$

where:

- G = normalized GPA value (0–1 scale), proven through a Verifiable Credential (VC),
- C = indicator of completed AI courses (1 if key courses were completed),
- R = research necessity factor, issued by a supervisor or instructor,

The weights (w_1, w_2, w_3) are defined equitably with a value of 0.33 each, allowing a balance between academic merit, curricular progress, research necessity, and operational efficiency. The estimated duration W, is used in the ordering rule which penalizes excessive workloads, prioritizing requests that require less computing time to optimize overall resource utilization.

A higher value of P represents a higher priority within the system, under which tasks would be ordered in descending order by P (from highest to lowest). This behavior corresponds to a **non-preemptive** priority scheduling policy, appropriate for batch processing environments (*batch scheduling*).

Batch scheduling and ordering rule. The system operates in discrete scheduling windows [11]. Within each batch, all requests are *frozen* at the start, and the queue is ordered only once under a non-preemptive scheme.

To combine merit (P_i) and estimated duration (W_i), the following ordering index is used:

$$\frac{W_i}{P_i + \varepsilon} \tag{2}$$

where $\varepsilon > 0$ avoids division by zero. Thus, shorter jobs with higher effective priority are served first, reducing the average turnaround time of the batch. When all P_i are equal, (2) reduces to the classic *Shortest Job First* (SJF) algorithm; with different P_i, a hybrid *Shortest & Weighted First* model is obtained, which minimizes the weighted sum of completion times [11].

Average turnaround time per batch. Given the non-preemptive order chosen in the batch, the completion times are:

$$C_k = \sum_{j=1}^{k} W_j, \tag{3}$$

and the average turnaround time is calculated as:

$$\bar{T}_R = \frac{1}{n}\sum_{k=1}^{n} C_k = \frac{1}{n}\sum_{j=1}^{n}(n - j + 1)\, W_j. \tag{4}$$

Minimizing $\bar{T}_R$ contributes to more efficient resource usage and improves the experience for users with short jobs without compromising system equity. This indicator will be reported per batch along with the distribution of times by priority deciles, in order to evaluate the model's effectiveness.

Practical analogy. The logic can be understood with a supermarket analogy: a person with few items in their cart (W_i small) and who also has preferential access (P_i high) goes first. In this way, the sum of average waiting times decreases, and the system achieves a balance between *global efficiency* and *merit recognition.*

Ties and robustness. In case of a tie in $\frac{W_i}{P_i}$, higher P_i is prioritized first, then lower W_i. If W_i presents uncertainty or variability, it can be discretized into categories {short, medium, long} and apply (2) within and between categories.

Altogether, this combined priority function allows optimization of AI resource access in university environments under a framework of verifiable equity and transparent governance.

3.7 Zero-Knowledge Proof Application

Zero-Knowledge Proofs (ZKPs) [2] enable the holder to demonstrate possession of an attribute without revealing its actual value. In this governance model, range proofs can be used to validate that a student's academic performance exceeds a defined threshold (e.g., "GPA > 85") while maintaining privacy. However, such proofs alone do not yield a numeric value that can be directly used in the priority function (Equation 1).

To preserve, both privacy and mathematical operability, a **threshold to symbolic value mapping** is proposed. The system defines GPA ranges that are independently verifiable through ZKPs, where each range corresponds to a normalized symbolic value G used in the priority calculation:

ZKP range	**Symbolic Value**	**Normalized G**
GPA > 90	"High Excellence"	1.00
$85 \leq$ GPA < 90	"Excellence"	0.90
$80 \leq$ GPA < 85	"Proficiency"	0.80
$70 \leq$ GPA < 80	"Satisfactory"	0.70
GPA < 70	"Below Threshold"	0.60

Table 1: Example of GPA verification ranges and corresponding normalized symbolic values.

This approach ensures that the verifier (e.g., the GPU scheduling system) can compute the priority score P without learning the student's exact GPA. The ZKP only certifies that the student belongs to a specific range, and the system translates this verified category into its corresponding symbolic value.

The proof process operates as follows:

(1) The issuer (university) defines a credential schema that includes the GPA attribute and publishes a set of threshold circuits (e.g., > 70, > 80, > 85, > 90).
(2) The holder (student) generates a ZKP asserting which threshold(s) they satisfy, without revealing the exact GPA.
(3) The verifier checks the proof against the issuer's public verification key and assigns the corresponding symbolic value of G to compute the final priority score P.

This model allows privacy-preserving prioritization: academic merit influences resource allocation without disclosing sensitive numerical data. It also enables the institution to update thresholds dynamically (e.g., per semester or department) while maintaining interoperability with Verifiable Credentials and DID-based authentication.

3.8 End-to-End Process Flow: Summary

Figure 3 presents the end-to-end workflow for AI resource governance built upon OpenDID and Verifiable Credentials. It integrates the previously described components—**University (Issuer)**, **Student (Holder)**, and **AI Resource System (Verifier + Scheduler)**—with the **Blockchain Registry** serving as a decentralized trust layer for credential integrity and status verification.

The process begins when the University issues a signed credential and registers its corresponding DID on the blockchain. The student stores this credential in their wallet and later submits a Verifiable Presentation that employs **Zero-Knowledge Proofs** to validate eligibility in a privacy-preserving manner. Upon receiving the request, the AI Resource System retrieves the issuer's verification data from the blockchain, confirms credential validity, and applies the **Shortest & Weighted First (SWF)** non-preemptive scheduling algorithm to organize queued jobs by priority. Execution results and usage logs are subsequently anchored to the blockchain to ensure transparency, traceability, and institutional compliance.

This representation consolidates the operational flow of the proposed model, illustrating how decentralized identity and cryptographic verification enable secure authentication, preserve privacy, and promote fair allocation of AI computing resources within academic environments.

4 ETHICAL AND SOCIAL ARGUMENT

Now, leaving aside the technical aspects, this model responds to an ethical principle: ensuring that students who demonstrate ingenuity and dedication can access tools they might not be able to access due to economic or administrative limitations. AI computing has become a driver of scientific innovation, so it is necessary to provide a multi-stakeholder access path to close a potential inequality gap.

Under this criterion, the administration and governance of these resources is presented as an efficiency mechanism, but above all, as an academic equity policy. By implementing a resource allocation strategy based on verifiable credentials with a priority scheduling algorithm, the TEC is putting an important card on the table: merit and academic need will be the parameters for allocating resources, not arbitrary privileges.

This principle offers the potential to be expanded in the future to interuniversity consortia in Costa Rica and Latin America, enabling the creation of shared infrastructures, promoting decentralized digital identity, and improving local cooperation.

5 IMPLEMENTATION ROADMAP

The proposed plan contemplates a 12-month horizon, with progressive phases that will provide feedback on technical feasibility and scale the solution to the institutional level.

5.1 Months 1–2: Authentication for access to AI resources

A pilot plan is being proposed for a digital ID with a DID for students who require access to GPUs/CPUs. They could access it by presenting a valid VC with the correct information. The initial objective would be to establish a basic authentication and usage traceability mechanism, without the need to manually manage traditional credentials.

5.2 Months 3–6: Institutional Authentication

In this phase, the goal is to establish an institutional framework. The digital ID card would be integrated with other TEC services (digital library, registration systems, virtual laboratories), consolidating student and faculty authentication under the SSI/VC model.

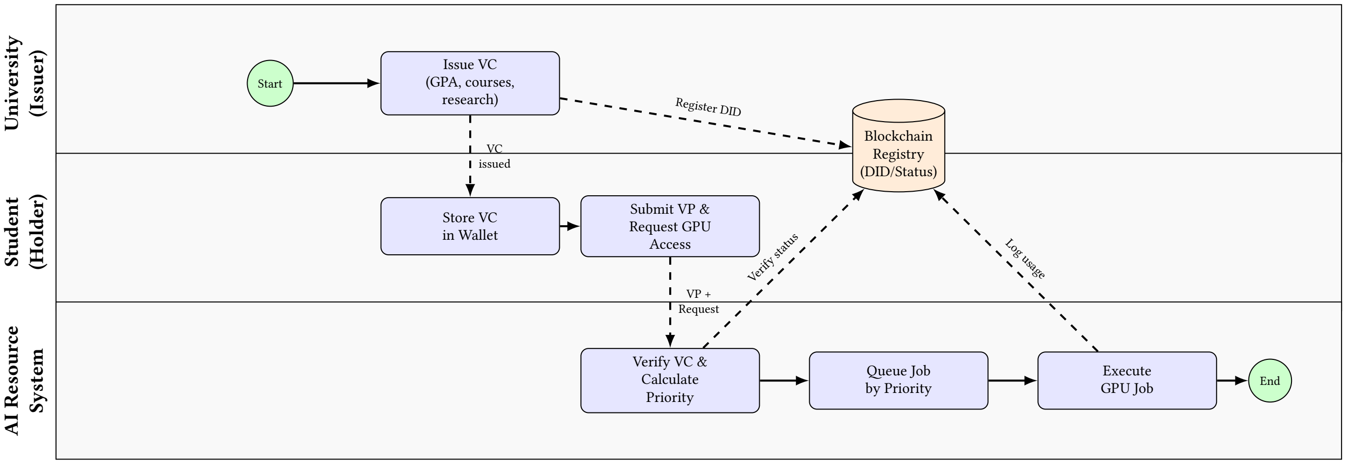

Figure 3: End-to-end process flow integrating OpenDID, verifiable credentials, and priority scheduling.

5.3 Months 6–12: Institutional scholarship system and digital voting

Incorporation of scholarship allocation mechanisms using verifiable credentials that incorporate mechanisms to recognize academic merit and special socioeconomic conditions, ensuring transparency in awarding scholarships. Additionally, a decentralized institutional voting module or academic consultations is proposed, resulting in a cohesive use of the identity ecosystem and zero-knowledge proofs.

5.4 Indicators to be evaluated

From this we derive the indicators that we must consider for its evaluation:

(1) Operational Efficiency: percentage of successful versus failed authentications.
(2) Scalability: number of students/users integrated per phase.
(3) Transparency and Equity: metrics for fair allocation of resources and scholarships.
(4) User Satisfaction: surveys and feedback from students and faculty.
(5) Institutional Compliance: integration with internal policies and alignment with state regulatory frameworks.

In the figure below, we can see the temporal structure and roadmap of the planned tasks.

6 CONCLUSIONS

Access to AI resources seems to be a new challenge that is already burdening higher education in Latin America. However, as Karl Marx said [8], ”history repeats itself”: scheduling theory and the search for efficient access to resources. However, these challenges may mean new ways to ”breaking out of the box” and confront these barriers with ingenuity, optimism, and innovation that integrates the most current technological resources. Under this scenario, the implementation of an emerging decentralized digital identity model can be justified, which also adds a fresh governance layer to TEC.

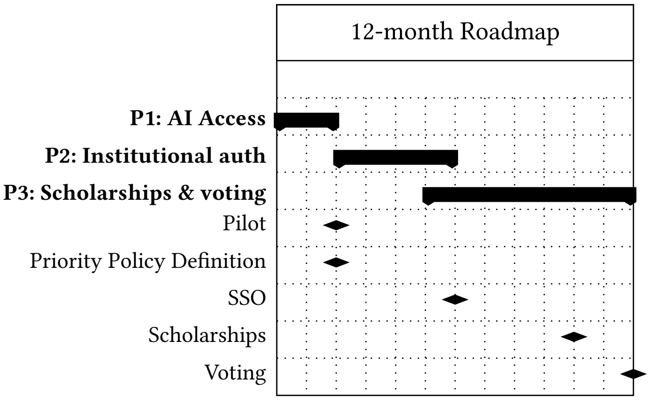

Figure 4: 12-month roadmap: P1 (Aug–Sep), P2 (Oct–Jan), P3 (Jan–Jul).

Just as in 1969 [5], with the arrival of the first computer, batch allocation was necessary for resource allocation and use, a similar scheme could be proposed with AI resources. However, this time, an automatic allocation mechanism is being introduced. With OpenDID and verifiable credentials, the TEC has a golden torch to lead an initiative that combines technical efficiency, social equity, and alignment with international best practices.

ACKNOWLEDGMENTS

To the staff of the TEC Computer School, for their guidance, support, and kindness throughout this process.

REFERENCES

[1] Amazon Web Services. 2025. AWS EC2 Pricing for NVIDIA H100 Instances. https://aws.amazon.com/ec2/pricing/.

[2] Benedikt Bünz, Jonathan Bootle, Dan Boneh, Andrew Poelstra, Pieter Wuille, and Greg Maxwell. 2018. Bulletproofs: Short Proofs for Confidential Transactions and More. In *Proceedings of the 2018 IEEE Symposium on Security and Privacy (SP)*. 315–334. https://doi.org/10.1109/SP.2018.00020

[3] CAF. 2020. Identidad Digital Soberana: Hacia la construcción de una infraestructura pública digital en América Latina y el Caribe. https://scioteca.caf.com/handle/123456789/1664. Banco de Desarrollo de América Latina (CAF).

[4] Alejandra Chomczyk, Jorge Madariaga, Eduardo Molina, and María Allende-López. 2020. Regulación de blockchain e identidad digital en América Latina: El futuro de la identidad digital. https://publications.iadb.org/es/regulacion-de-blockchain-e-identidad-digital-en-america-latina-el-futuro-de-la-identidad-digital. Banco Interamericano de Desarrollo (BID).

[5] Universidad de Costa Rica. 1968. Conozca a Matilde, la primera computadora del país. https://vinv.ucr.ac.cr/es/noticias/conozca-matilde-la-primera-computadora-del-pais. Vicerrectoría de Investigación, Universidad de Costa Rica. Accedido en 2025.

[6] Richard L. Henderson. 1995. Job Scheduling Under the Portable Batch System. In *Proceedings of the Workshop on Job Scheduling Strategies for Parallel Processing*. Springer, 279–294. https://doi.org/10.1007/3-540-60153-8_34

[7] Morris W. Jette, Andy B. Yoo, and Mark Grondona. 2002. SLURM: Simple Linux Utility for Resource Management. In *Proceedings of Job Scheduling Strategies for Parallel Processing (JSSPP 2003)*. 44–60. https://doi.org/10.1007/10968987_3

[8] Karl Marx. 1852. *El 18 Brumario de Luis Bonaparte.* Die Revolution. Frase célebre: ”La historia ocurre dos veces: la primera vez como tragedia, la segunda como farsa.”.

[9] Sang Joon Park, Myung Gil Kim, and Yong Woo Kim. 2020. OmniOne: Blockchain-based Self-Sovereign Identity Platform for OpenID Connect. In *Proceedings of the 2nd IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*. IEEE, 148–153. https://doi.org/10.1109/DAPPS49028.2020.00028

[10] Alex Preukschat and Drummond Reed. 2021. *Self-Sovereign Identity: Decentralized digital identity and verifiable credentials.* Manning Publications.

[11] Andrew S. Tanenbaum and Herbert Bos. 2015. *Modern Operating Systems* (4th ed.). Pearson.

[12] Muhamed Turkanović, Marko Hölbl, Kristjan Košič, Marjan Heričko, and Aida Kamišalić. 2018. EduCTX: A Blockchain-Based Higher Education Credit Platform. *IEEE Access* 6 (2018), 5112–5130. https://doi.org/10.1109/ACCESS.2018.2789929

[13] World Wide Web Consortium (W3C). 2022. Decentralized Identifiers (DIDs) v1.0. https://www.w3.org/TR/did-core/. W3C Recommendation.

[14] World Wide Web Consortium (W3C). 2023. Verifiable Credentials Data Model v2.0. https://www.w3.org/TR/vc-data-model-2.0/. W3C Recommendation.