DCC DIGITAL CREDENTIALS CONSORTIUM

Building the digital credential infrastructure for the future

A White Paper by the Digital Credentials Consortium

Editors :

- Kim Hamilton Duffy (Chair, W3C Credentials Community Group)
- Hans Pongratz (CIO, TUM)
- J. Philipp Schmidt (Director, Learning Initiative, MIT Media Lab)

Contributing Authors:

- James Chartrand (McMaster University)
- Stuart Freeman (Applications Developer, C21U, Georgia Tech)
- Ulrich Gallersdörfer (TUM)
- Matt Lisle (Director of Digital Learning Technologies, C21U, Georgia Tech)
- Alexander Mühle (Hasso Plattner Institute, University of Potsdam)
- Sélinde van Engelenburg (Delft University of Technology)

Lead Reviewing Author:

- Krishna Rajagopal (William A. M. Burden Professor of Physics, Dean for Digital Learning, MIT)

Reviewing Authors:

- Nimisha Asthagiri (Chief Architect, edX)
- Matteo Bertazzo (Product Manager, CINECA)
- Michael Burke (Registrar, Harvard)
- Brian Canavan (Senior Associate Registrar, MIT)
- Paolo Cherubini (former Deputy Rector, University of Milano-Bicocca; Steering Committee for Learning & Teaching, EUA, University of Milano-Bicocca)
- Ike Chuang (Senior Associate Dean of Digital Learning, MIT)
- Camille Crittenden (Executive Director, CITRIS and the Banatao Institute, UC Berkeley)
- Darek DeFreece (Managing Director, New Academic Ventures, UC Berkeley)
- Julien DePauw (Tecnológico de Monterrey)
- Jeff Dieffenbach (Associate Director, Integrated Learning Initiative, MIT)
- Dick H.J. Epema (Professor of Computer Science, Delft University of Technology)
- Jose Escamilla (TecLab Director, Tecnológico de Monterrey)
- Matthias Gottlieb (Senior Researcher, TUM)
- Steven Harmon (Director of Innovation, C21U, Georgia Tech)
- Oliver Heyer (Research, Teaching and Learning Services: Director of Projects, Development and Operations, UC Berkeley)
- Irving Hidrogo (Educational Technology Project manager, Tecnológico de Monterrey)
- Michael Kan (Executive Director of HarvardX, Harvard)
- Timo Kos (Executive Director, Extension School, Delft University of Technology)
- Henry Leitner (Chief Innovation Officer and Associate Dean, Senior Lecturer on Computer Science, Harvard)
- Nida van Leersum (Policy Advisor, Delft University of Technology)
- Gary Matkin (Dean of Continuing Education, UC Irvine)
- Mihnea Moldoveanu (Vice Dean of Learning and Innovation, University of Toronto)
- Melissa Pool (University Registrar, McMaster University)
- Ishwar K. Puri (Professor and Dean of Engineering, McMaster University)
- Jan Renz (Research Scientist, Hasso Plattner Institute, University of Potsdam)
- Bertha Saldivar (Technologies for Education, Tecnológico de Monterrey)
- Sanjay Sarma (Fred Fort Flowers and Daniel Fort Flowers Professor of Mechanical Engineering, Vice-President for Open Learning, MIT)
- Brian Subirana (Director, Auto-ID Lab, MIT)
- Marinke Sussenbach (Manager Education and Student Affairs, Delft University of Technology)

- Tracy Tan (Director, MicroMasters Program, MIT)
- Dustin Tingley (Deputy Vice Provost for Advances in Learning, Professor of Government, Harvard)
- Willem van Valkenburg (Manager Teaching & Learning Services, Delft University of Technology)
- Diana Wu (Dean, UC Berkeley Extension and New Academic Ventures, UC Berkeley)
- Maria White (Assistant Dean of Engineering, McMaster University)
- Alan Wolf (Managing Director of Academic Technology Services, Harvard)
- Walter Wong (University Registrar, UC Berkeley)

External Reviewers:

We are grateful for the comments, questions, and other input received from the following individuals. Inclusion in the list below is not meant to constitute endorsement. Any errors in the final white paper are the responsibility of the authors.

- Alessandro Aglietti (Co-founder, Growbit)
- Roman Beck (Professor in Information Systems and Blockchain Economist, IT University of Copenhagen)
- Hennie Bulstra (Convenor User Group Diplomas and Credentials, European Blockchain Partnership)
- William Claxton (Founder and CEO, NextID)
- Perrine de Coëtlogon (Chargée de mission Blockchain & Education, University of Lille)
- Brian Fleming (Executive Director, Sandbox Collaborative, Southern New Hampshire University)
- Lorenzo Gentile (Research Assistant, University of Copenhagen)
- Andrew Law (Director Business Innovation, The Open University, UK)
- Herman de Leeuw (Executive Director, Groningen Declaration Network)
- Kerri Lemoie (Tech Strategist & Researcher, OpenWorks Group)
- Mark Leuba (VP Product Management, IMS Global)
- Phil Long (Special Advisor, UTO, Arizona State University; Senior Scholar, CNDLS, Georgetown University)
- Snorre Lothar von Gohren Edwin (CTO, Diwala)
- Lluís Alfons Ariño Martín (CIO, Universitat Rovira i Virgili)
- Greg Nadeau (Manager, Public Consulting Group and Chair, IMS Global Comprehensive Learner Record CLR Workgroup)
- Nate Otto (Director, Badgr Platform, Concentric Sky)
- Simone Ravaioli (Director Strategic Partnerships, Digitary)

About the Digital Credentials Consortium

The Digital Credentials Consortium was founded by leading universities with expertise in the design of verifiable digital credentials. Together, we are designing an infrastructure for digital credentials of academic achievement.

Founding Members

Delft University of Technology (The Netherlands)

Georgia Institute of Technology (USA)

Harvard University (USA)

Hasso Plattner Institute, University of Potsdam (Germany)

Massachusetts Institute of Technology (USA)

McMaster University (Canada)

Tecnologico De Monterrey (Mexico)

TU Munich (Germany)

UC Berkeley (USA)

UC Irvine (USA)

University of Milano-Bicocca (Italy)

University of Toronto (Canada)

Table of Contents

Context

The Case for Digital Academic Credentials

Technology is profoundly changing higher education, but the way we issue and manage academic credentials, which represent learning outcomes and achievements, has not yet taken advantage of the possibilities of digital technology.

What would an academic degree look like if it was designed today? Or a professional certificate? Or a certificate for an online course? As the question of trusted verification and authentication of learning and credentials poses itself with increased urgency we need to redesign the way we issue, recognize and transact with academic credentials.

Adding digital technologies enables three broad areas of benefits:

- It increases the efficiency of exchanging and evaluating credentials,
- It provides more reliable ways to protect and verify the credentials, thereby reducing the opportunity for fraud,
- It expands learners' control over their credentials, enabling a verifiable history of lifelong learning.

The Digital Credentials Consortium proposes to modernize the concept of credentials, bringing benefits to learners and to relying parties (such as employers) by improving how skills and competencies are conveyed and recognized. Our hope is that this will strengthen trust and enable additional value in how academic credentials are considered within nations and globally.

What is a Digital Credential? (Documents vs Envelopes)

A digital credential can be imagined as a combination of two components: a document and an envelope into which that document is placed. The document is like the piece of paper a university issues to a graduate, which might contain the name of the recipient as well as a description of the credential they received. The envelope protects the content of the document so it cannot be changed and it reliably communicates the authenticity of its contents.

Our efforts focus on the envelope and the system that provides safe delivery and storage of multiple envelopes—similar to the postal service for mail. The envelope contains information about who issued the credential and to whom it was issued. It creates robust links to the identity of an issuer (e.g., a specific university) and the learner (e.g., a particular learner).

Also, both the identities and the integrity of its content can be verified to detect fraud or tampering.

Other efforts are underway that focus primarily on the document; examples include the European Qualifications Framework and the IMS Global Comprehensive Learner Record. They present different ways in which skills and competencies can be described. Given the global nature of our effort, we will support different approaches as much as possible.

A University-led Effort - Digital Credentials Consortium

Our mission is to create a trusted, distributed, and shared infrastructure that becomes the standard for issuing, storing, displaying, and verifying digital academic credentials.

The Digital Credentials Consortium (DCC) is primarily concerned with use-cases in higher education, but we see our work as part of a broader effort to bridge post-secondary and lifelong learning, connecting traditional institutions of higher education, non-formal education providers, as well as the workplace, through interoperable standards.

Our goal is to contribute to an education landscape that increases learner agency and promotes more equitable learning and career pathways.

What makes this different from other initiatives? Our effort is entirely driven by institutions of higher education. We are committed to open source and open standards and are actively working with standards groups to complement existing efforts [see Relationship to Other Credential Standards and Initiatives]. Finally, we expand on previous efforts in a number of important ways:

- More flexible ways to express the identities of issuers and learners that tie into existing university services.Stronger privacy-by-design and privacy-by-default with attention to regional legal frameworks such as the GDPR.
- More reliable revocation mechanisms and credential lifecycle management.
- More direct learner agency over one's lifelong learning record.
- Higher level of consistency between the machine-readable data of the credential, the human-readable visual representation, and the necessary output formats—paper or digital.

The initial group of international universities leading this effort has deep expertise and experience in designing digital credentials. Our focus is the design of the standard and development of a transparent governance model that keeps the learner's rights at the center.

We are not developing commercial products and services; instead, we are working with technology companies, online learning platforms and IT vendors to create a vital ecosystem of options to choose from. We will also work with employers to integrate verification services into their hiring workflows. By working together, we intend to put into practice a new standard for learner-controlled, privacy-preserving credentials, in a manner that ensures interoperability and avoids vendor lock-in. To this end, the consortium will incubate standards openly within the framework of a W3C community group, with draft specifications and reference implementations released under the W3C Software and Document License, and will consider collaborating with other standards bodies as appropriate.

Guiding Principles

Our work is informed foremost by learners' welfare, rights, and agency.

Learners

Learners retain primary control over their credentials. Learners' consent is required for issuance of digital credentials. Learners decide to whom they grant access. Barriers to receiving and managing credentials are minimal to enable broad participation.

Issuers

Issuers control to whom they issue credentials, the particular achievement that the credential represents, and which credential options are available to the learner. Issuers can revoke credentials according to their institution's policies. Barriers to issuing credentials are minimal to enable broad and diverse participation.

Trust

Trust in the integrity of the infrastructure is embedded in its design and flows from transparency. Everyone is able to review how the infrastructure and processes work. Trust in the integrity of the credentials is established cryptographically. Credentials can be verified without consulting the original issuer.

More information about this effort can be found on the project website at https://digitalcredentials.mit.edu.

About this white paper

This white paper sets out the design considerations of the system architecture. It serves as the foundation for the development of reference implementations, software libraries, and deployment prototypes by the participating universities. It describes technology choices we are making, the tradeoffs they come with, and the state of our current thinking.

The paper is intended for a general audience, but contains sufficient technical detail to invite review by technology system designers and digital credential developers.

We also intend for the white paper to act as a call-to-action that will help clarify to others (not just partners mentioned in this paper, but other individuals, organizations and corporations) how they can work with us.

I Scope, Requirements, Terminology

Scope

The system we are designing can be used to issue and verify many different types of academic credentials, ranging from university degrees and diplomas, to individual course credits, to alternative credentials (including microcredentials) for online courses or face-to-face workshops. Issuers decide the information they must include in the credential. Our system will not change the way universities provide instruction, assess learning, or make decisions about awarding credentials. It simply offers a more powerful and convenient way to share, manage, and verify the credentials.

Requirements

This section describes core requirements that the system is designed to fulfill.

Prioritize learner agency and privacy

Require learner consent for issuing credentials: The learner is at the center of transactions related to their credentials. Both learner and issuer consent is required to issue a credential.

Issue credentials that optimize for learner flexibility and privacy: Learners must be able to use credentials flexibly, avoiding lock-in to a specific system. At the same time, the credentials issued by the system will use privacy-enhancing measures to ensure that only the learner has that freedom, while limiting other parties who may want to exploit data about the learner. This includes longer term considerations as well, such as the learner’s desire for handling of their data after they are deceased.

Enable seamless verification without involvement of the issuer: Learners can present their credentials for frictionless verification without requiring the issuer to be involved. This is particularly important during situations where the issuing institution may not be reachable, which could happen if a university has closed, or if broader political or infrastructure problems make contact infeasible.

The usability of the verification process is a primary consideration to avoid pitfalls such as social engineering exploits that occur when relying parties do not understand verification status messages or other poor visual cues. The verification process must be robust and trustworthy, but implementable in a way that supports seamless integration into a variety of tools.

Minimize the need for disclosure: Sharing credentials requires the minimum necessary amount of disclosure, in particular for any personally identifying information (PII). For example, learners need not send a transcript if all that's requested is the equivalent of a diploma, even though the transcript contains a superset of the data. Traditional credentials do not handle this well; to prove you are over the age of 21, in most jurisdictions you must show a driver's license or other government-issued ID that reveals far more information, such as your exact date of birth and address.

Prevent tracking: Minimize the ability of the issuer or other parties to track activities of the learner or correlate information about them. For example, the learner may share credentials without involving or even informing the issuer. In addition, this approach minimizes opportunities for the scraping of credential information without consent of the learner, e.g. to create a learner profile by correlating transactions that are recorded on a public blockchain.

Enable recovery: Replacements for lost credentials should be easily requested and securely received from the issuer or a trusted third-party, such as standard-compliant vendors.

Offer multiple options for credential storage: The learner can choose where to store and manage their credentials.

Provide trusted learner identity: The learner can cryptographically prove that a credential is about themself. The system enables learners to use their credentials during digital interactions with other systems that require authentication.

Enable Trust

Prevent tampering and fraud: The system should minimize credential forgeability, making credentials tamper-evident in content and presentation, and offering reliable means of establishing authenticity.

Allow only necessary auditability: Issuers must be able to audit their own issuance and revocation events to determine proper system behavior and detect fraudulent activity. However, audit logs should minimize PII and only be accessible to users authorized by the issuer.

Provide display integrity: The visible credential and underlying data must be easily verifiable as consistent. Humans rely on a range of cues (e.g., watermarks, signatures) to make a decision about a credential's integrity, but an under-emphasized aspect of digital credentials is the integrity of how credentials may be displayed on different screens or devices (“display integrity”). Designing for easily verifiable human-readable displays is essential.

Support Diverse Use-Cases and Technology Best-Practices

Support different issuers and types of credentials. The system must provide a standard way to verify credentials from many different sources and support different types of credentials (and credential data standards). Content of the credential can conform to a variety of schemas and vocabularies chosen by the issuer. Some well-known examples used in academic credentialing include PESC, EQF, CTDL, CASE, CLR, ELMO, Open Badges, and Schema.org.

Remain efficient, scalable, fault-tolerant, and highly available: High-certainty verification of credentials is possible with minimum time and cost overhead and scales to the demands of the global higher education system. All aspects that are relied on for learner usage are required to be highly available with appropriate consideration of points of failure.

Ensure longevity: The system must ensure that credentials can be used by learners at a minimum throughout their lifetime.

Design for sustainability: The system should avoid excessive resource needs and ensure that the technical design and governance structures can evolve over time to support additional and new use-cases.

Prevent lock-in: No part of the standards or implementing systems requires use of a proprietary solution or specific vendor, though vendors are encouraged to build standards-compliant solutions. It’s especially critical that learners have control over where their data resides and are locked neither into a specific provider nor solution.

Enable integration with existing infrastructure: Issuing functionality is designed to be easy to integrate into existing university student information systems, and offer the features demanded by registrars and other university groups involved in issuing credentials, such as ease of issuance, revocation, recordkeeping, etc.

Build on open standards: The open standards used in this approach may be used and adapted for a variety of governance models. Initially, issuer identity verification support will roll out in a trust-building, conservative manner that only takes responsibility for maintaining the identity of members of the initiative.

Provide accessibility: In addition to adhering to accessibility guidelines and standards (such as the W3C Web Content Accessibility Guidelines), we will seek partnerships to ensure that we are promoting accessibility best practices. Credentials and systems based on this standard should be broadly accessible, including to those using assistive technologies.

Support international use: Our approach will ensure usability in different languages, jurisdictions, and conventions.

Terminology

We use the term **learner** to reflect that the individual in our use cases has lifelong learning experiences in a variety of contexts, such as the traditional classroom, online learning, formal/informal training, and other experiences. The learner is the central actor. Subsequently, the pronouns “they/them” will be used as a singular for “learner” throughout. See additional information related to the alignment of this term to other data models in the “Terminology Alignment” appendix.

The **issuer** is an entity issuing credentials to the learner, in our use-case this tends to be an academic institution. This term was chosen to convey that this is an open standard usable by any credential issuer. While some components of the system are specific to the Digital Credentials Consortium, all follow open standards.

A credential is a set of claims (attributes about a learner) made by an issuer. A **verifiable credential** is a tamper-evident credential where the authorship can be cryptographically verified.

Our use of the terms **credential, verifiable credential,** and **verification** (as well as related forms) comes from the Verifiable Credentials Data Model. Appendix - Terminology Alignment provides additional details about this terminology alignment, but an essential definition (and concept) for this paper is **verification**, which is the evaluation of whether a verifiable credential is an authentic and timely statement of the issuer.

The **relying party** is any organization or person the learner chooses to share a credential with. This can be a potential employer, a bank, an educational institution, or any other party asking for credentials that the learner consents to.

An **identity registry** is a specific type of verifiable data registry used to mediate the verification of identifiers, public keys, and other relevant data. This includes **issuer registries**, used for the verification of issuers, and potentially **learner registries** on an opt-in basis.

A **revocation registry** is a specific type of verifiable data registry used to store, and enable retrieval of, revocation status of a credential.

A **wallet** is a term we use to refer to software installed on a learner’s mobile device, or accessible through a web browser, to manage their credentials and profiles. We realize that this term may not be the best conceptual match for these functions, and is subject to change if a better expression emerges in the credential ecosystem.

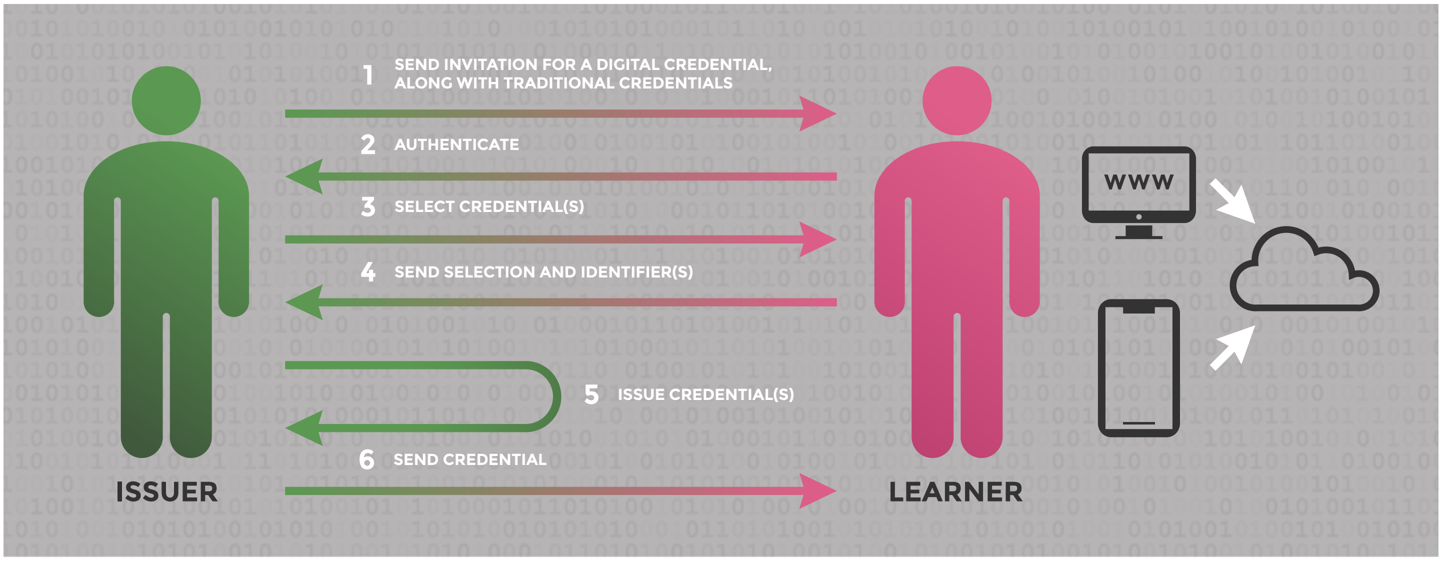
II Features of the System

The credential system supports the following core features and interactions. The consortium will develop an open source, standards-based reference implementation to support these interactions, as described in Components.

Issue/Receive Credential

The "Issue Credential" task is initiated by the issuer sending an invitation to the learner to receive a digital credential. If the learner chooses to receive the credential, they authenticate with a site or service maintained by the issuer. The learner is presented with the credentials they are eligible to receive (which is determined by the learner's achievement record). The learner makes their selections and provides the identifier(s) with which they want the credential to be associated. The issuer generates the credential(s) and sends it/them to the learner.

As discussed in Flexible Identity Model, this process allows the learner to associate their credentials with real-world identities, such as eID, university ID, bank ID, or even emerging self-sovereign ID (an approach enabling learners to control when, to whom, and how they share data about themselves). Some of these identity models further enable the learner to cryptographically prove control over the identifier and associated documents.


1 SEND INVITATION FOR A DIGITAL CREDENTIAL, ALONG WITH TRADITIONAL CREDENTIALS
2 AUTHENTICATE
3 SELECT CREDENTIAL(S)
4 SEND SELECTION AND IDENTIFIER(S)
5 ISSUE CREDENTIAL(S)
6 SEND CREDENTIAL
ISSUER
LEARNER
WWW

At least during the prototyping phase we recommend that learners continue to receive traditional credentials (the same way they are being issued today) and that the new digital credentials be phased in as optional.

A specific focus will be made to develop and offer tutorials and training for issuing and managing credentials, including how to inform students of the many benefits of accepting and learning to manage their credentials.

Store/Retrieve/Manage Credential

Learners store and manage their credentials in a variety of different ways:

- A wallet that is installed on a device, such as a smartphone.
- A website with storage and management features that is either hosted by universities, service providers, or technically proficient users themselves.

Credential portability across systems is a requirement that is enabled through standards described in Credential Ecosystem Standards and Protocols.

Share Credential

The learner's consent is central to credential exchange and verification. Since the learner controls their credentials, the learner may choose to share them with any relying party.

The system allows learners to share credentials in different contexts, in which the learner wants choice over which credentials to reveal (due to privacy or relevance concerns). Initially, this is enabled at the time of issuance by providing multiple credentials of different granularity, as described in Credential Data Model. We are also exploring emerging techniques to enable more advanced, fine-grained options for additional data minimization and selective disclosure (through zero knowledge proofs and redaction) for future iterations.

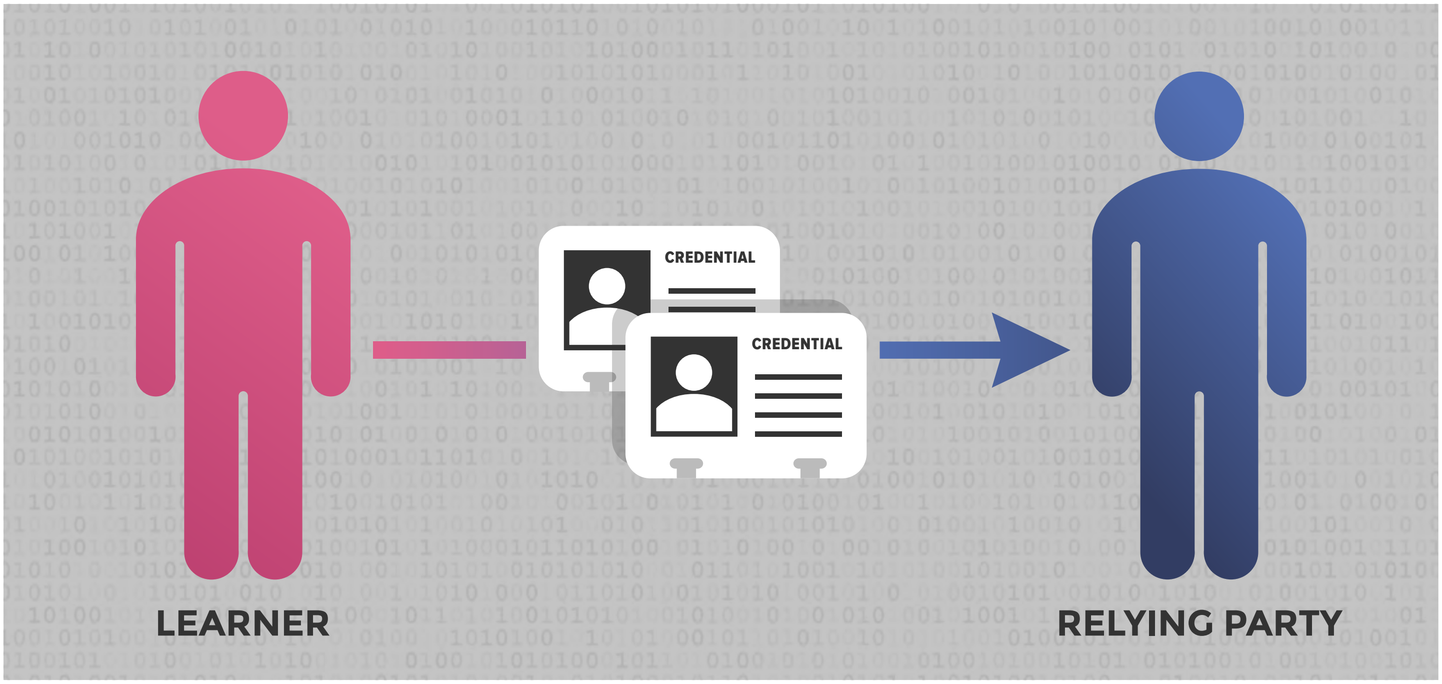
Different "Share Credential" contexts include:

- Share on social networking site (minimal sensitive personal data).
- Share with potential employer or educational institution.

The first context corresponds to credentials that the learner is prepared to share and make available for verification, such as proof of graduation or certification. In this case, the social networking site may expose instant verification through a widget (i.e., an embeddable code) in the site. If a relying party does not wish to rely exclusively on the site's verification widget, they could independently verify the content of the credential through a verification tool of their choice that conforms to our system's verification standards (here, we envision a marketplace of different products and services offering verification).

In the second case, which may apply to a credential with more sensitive details (such as a transcript) the learner may not wish to share the credential broadly. Instead, they would choose to share the credential with parties they approve, who would then use a standard-compliant means of verification. They could share it in two ways:

1. Informal method of sharing of content (e.g. file upload, email).
2. Credential exchange protocols (based on emerging standards, described in Credential Ecosystem Standards and Protocols).



The standards-based tools enabling these learner use cases are described in Components.

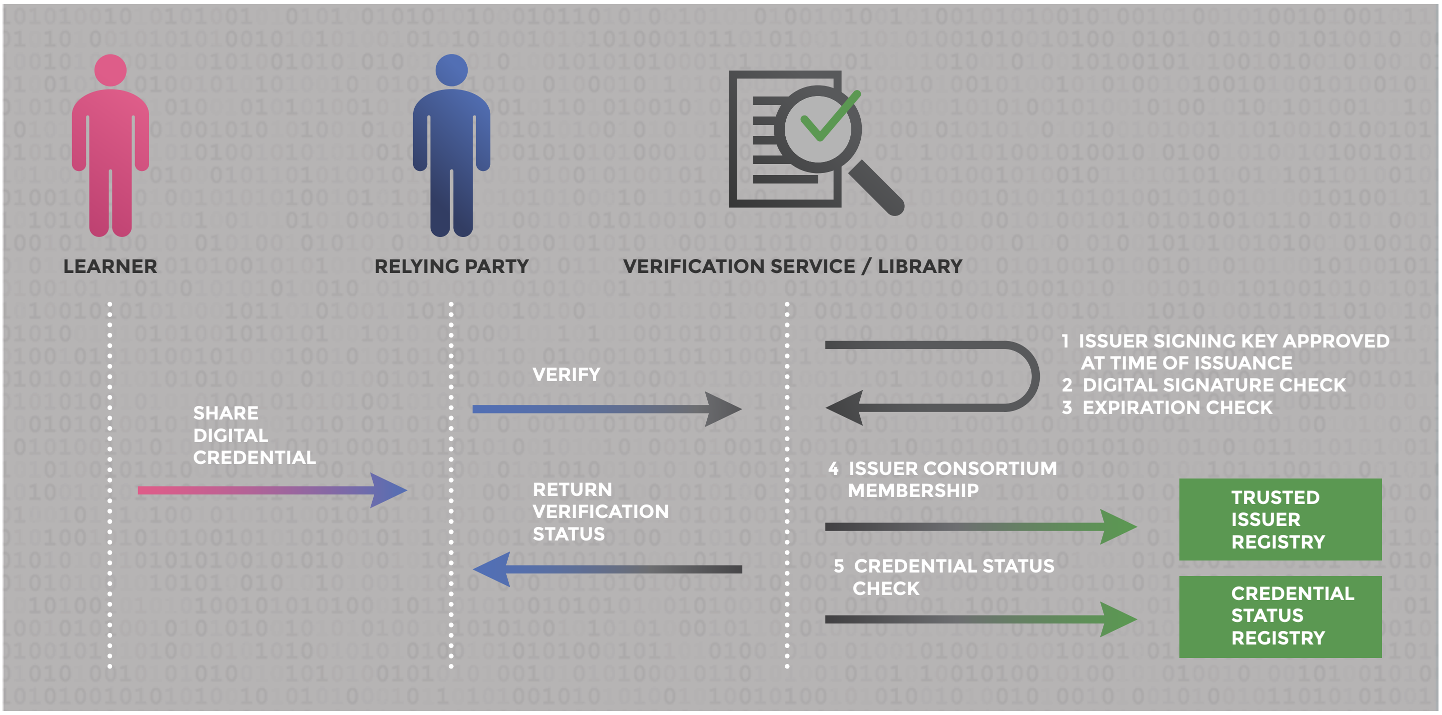
Verify Credential

The verification process is initiated after a learner shares their credential(s) with a relying party that then uses a standard-compliant means of verification. Establishing the identity of the issuer is critical to establishing trust in the credential. A credential issued by the Digital Credentials Consortium will contain pointers to sources of trust approved by the consortium to be used during verification. This does not imply that the verification protocol outlined here is restricted to consortium-issued credentials; this is an open standard and so a relying party may choose to honor credentials complying with the standard anchored to different sources of trust.

The protocol cryptographically enforces that credentials may not falsely be attributed to a source of trust that it is not a part of. Supposing that the source of trust is the Digital Credentials

Consortium, a relying party will receive a credential and check the following:

1. The issuer's signing key was approved for use at the time of issuance, because the time stamp on the issuance precedes any expiration or revocation of the key.
    - The timestamp may be provided by a blockchain transaction, timestamp authority, or other means.
    - The signing key is approved by the issuer (according to the source of trust the issuer designates in the credential).
2. The digital signature indicates the credential has not been tampered with and that the issuer's signing key was in fact used to sign.
3. The credential has not expired, according to the expiration terms provided by the issuer in the credential.
4. The issuer’s identifier can be found in the Digital Credentials Consortium Registry (see “Trusted Issuer Registry” in Components for details).
5. The credential’s status has not changed to an unusable state (for example, it hasn’t been revoked by the issuer—see “Credential Status Registry” in Components for details).


LEARNER
RELYING PARTY
VERIFICATION SERVICE / LIBRARY
SHARE DIGITAL CREDENTIAL
VERIFY
RETURN VERIFICATION STATUS
1 ISSUER SIGNING KEY APPROVED AT TIME OF ISSUANCE
2 DIGITAL SIGNATURE CHECK
3 EXPIRATION CHECK
4 ISSUER CONSORTIUM MEMBERSHIP
5 CREDENTIAL STATUS CHECK
TRUSTED ISSUER REGISTRY
CREDENTIAL STATUS REGISTRY

Verification returns a success/failure status along with reasons for failure; such as expired status.

Not shown here is proof of control of a credential; i.e. that the learner is indeed the subject of the credential. That is covered in “Prove Control of Credential”.

Revoke Credential

The issuer is able to revoke the credential(s) they issued.

The learner does not need the ability to revoke a credential because only the learner can choose to share a credential. The learner cannot be forced to disclose any credentials they disavow. However, as discussed in Planning for Compliance, the system must provide ways to support a learner's "right to erasure".

Reissue Credential

Credential reissuance may be required for a variety of reasons, such as a name change associated with a life event, a typo in or loss of the original credential, or, in some cases, lost or compromised cryptographic keys. The tasks associated with reissuance are similar to those for the original issuance, except there is an additional step of revoking the original credential.

Onboard Issuer

The onboarding process establishes the issuer's credential signing keys and synchronizes the issuer's identifier and key information to a registry for use during credential authenticity checks. As discussed later in Open, Trust Promoting, the Digital Credentials Consortium will maintain a registry of verified (member) issuers.

We plan to extend the onboarding process and allow various options for self-registration or use of other registries in the future. In the meantime, other organizations can already use the system and create credentials, but these will not be fully verified against the consortium registry (since those issuers are not listed in the registry).

Prove Control of Credential

Digital credentials enable the learner to prove they are the intended recipient of the credential, avoiding some forms of fraud that exist with paper credentials. The strength of proof can vary according to the type of credential.

An example of a weaker proof that other systems have used for credentials that involve lower stakes is by embedding the learner's email address (or hash of it) into the credential itself.

A somewhat stronger approach could embed real-world identities (or again, a hash of the associated identifier), such as eID, university ID, bank ID, and so on, at the discretion of the learner (as described in the Issue/Receive Credential section above). The learner could then corroborate their identity with external documentation or proof (student card, etc.)

A cryptographic option, which may be combined with other approaches described here, would be to rely on a public key that is embedded in the credential, for which the credential holder can prove ownership. However, this approach suffers from other limitations, such as when the learner loses control over the corresponding private key.

We are particularly interested in proof of control enabled via the Verifiable Credentials Data Model, where a learner identifier can be expressed flexibly (as a URI), enabling stronger forms of control and key lifecycle management via Decentralized Identifiers or Solid Profiles. This approach enables integration into strong authentication ceremonies, such as those enabled through the FIDO2 open standard.

III Components of the System

This section describes the foundational components of a learner-centric credentialing ecosystem. As we progress through our prototype phases, we will enable more component options, thereby enabling more learner options for control, flexibility, and privacy.

Foundations

Open, Trust-Promoting

The Digital Credentials Consortium optimizes the infrastructure for openness whenever possible. The relying party can choose their preferred verification service provider and/or apply their own extended criteria, such as the set of issuers they deem trustworthy.

Activities on the decentralized infrastructure can be transparently observed and, in combination with internal records of the centralized components, a complete audit trail can be created. Administrators can therefore monitor credential issuance and revocation.
While audit trails and records of the issuance system itself are determined by issuer requirements and not considered part of the standard, we will suggest best practices for detecting fraud and avoiding misuse.

Trust remains the highest goal for the infrastructure and it is essential for establishing issuer authenticity. To align the early stages of development with the high requirements for issuer authenticity, we decided for the initial implementation to define a closed registry for issuers, while aiming for an open, self-governed registry. As the protocol is open, any other party is allowed to make its own decisions about sources of truth. Details of these implementation choices and components follow.

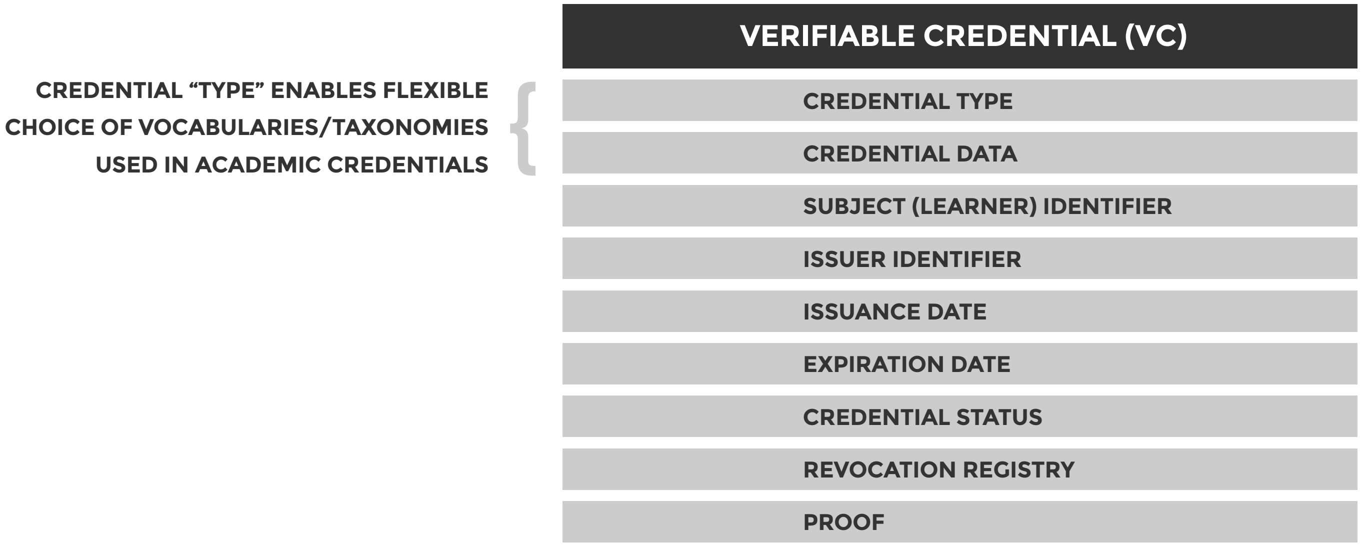
Credential Data Model

The credential “envelope” format is the primary driver of interoperability. The emerging W3C Verifiable Credentials (VC) Data Model defines a means of expressing digital credentials in a tamper-evident way that puts the learner at the center. Its structure is lightweight, satisfying the “envelope” properties, and allowing flexible definition of schemas and interoperability across different kinds of credentials.

We have selected Verifiable Credentials for our design because this method fulfills the requirements listed in Section I above and offers the following properties:

- Digital, tamper-evident, and machine-readable.
- Learner is at the center of exchanges of their data; their consent is required to exchange.
- Learner may cryptographically prove the credential is about them.
- Verification does not require consulting the issuer.
- Content of the credential can conform to a variety of schemas and vocabularies chosen by the issuer. Some well-known examples used in academic credentialing include PESC, EQF, CTDL, ELMO, Open Badges, and Schema.org.
- Robust consideration of privacy, security, accessibility, and internationalization recommendations, making it suitable for high-stakes credentials such as transcripts.

The figure demonstrates the elements of a Verifiable Credential. The credential data (the “document” or “payload”) can vary according to the issuer’s credentialing requirements, while allowing a standard/consistent means of validating credentials.


VERIFIABLE CREDENTIAL (VC)
CREDENTIAL “TYPE” ENABLES FLEXIBLE CHOICE OF VOCABULARIES/TAXONOMIES USED IN ACADEMIC CREDENTIALS
CREDENTIAL TYPE
CREDENTIAL DATA
SUBJECT (LEARNER) IDENTIFIER
ISSUER IDENTIFIER
ISSUANCE DATE
EXPIRATION DATE
CREDENTIAL STATUS
REVOCATION REGISTRY
PROOF

While the issuer is responsible for the credential content (including schema and vocabularies), it is recommended that issuers provide credentials at a level of granularity that is most useful to learners and relying parties. One approach is for the issuer to supply multiple credentials with a variety of granularity levels for the learner to share, for example issuing an entire transcript with a list of courses as one credential, or issuing individual credentials for each course, which can then be aggregated into a transcript by the relying party. We favor solutions that offer a higher degree of flexibility, but maintain the issuer's control over the granularity of their credentials (for example, if the issuer chooses to disallow cherry-picking only high grade courses when presenting an official transcript).

This results in a set of tamper-evident credentials that can be used by the learner in a variety of contexts. For example, if a relying party only needs to see that the student has a 4-year diploma, then the learner (or technically, the learner's wallet or agent) would share their diploma as opposed to a full transcript. At the same time, the learner isn't able to tamper with the content of a credential—attempting to do so will cause it to fail verification.

The Verifiable Credentials Data Model allows the learner to disclose any credentials they choose, but as the learner is at the center of the exchange, they can also choose not to disclose some credentials. The learner determines which irreducible credential they want to share, and the relying party decides what level of detail they are willing to accept.

Enabling learner control over which credential information is shared can also be achieved with data minimization / selective disclosure techniques, consistent with the Verifiable Credentials Data Model (see section 5.8 Zero Knowledge Proofs for one such technique).

The W3C Verifiable Credentials (VC) Data Model is still a relatively new standard and we will continue to assess and address any possible issues with our implementation by working with the relevant standards bodies, including the W3C Verifiable Credentials Working Group, W3C Credentials Community Group, W3C Decentralized Identifier Working Group, and Decentralized Identity Foundation.

Credential Ecosystem Standards and Protocols

The Verifiable Credentials data model provides the basis for an interoperable credentialing ecosystem. Emerging standards and protocols enabling flexible learner-focused credential storage and exchange are also essential to learner control and interoperability. These standards will be critical in the development of related tools over time.

Entities in the Credential Ecosystem

An example deployment of a credential ecosystem—including the entities that enable management and exchange of credentials and associated identities—are shown below. In this example, the learner interacts with a "wallet", which is software installed on a learner's device enabling secure on-device storage of cryptographic keys or other identification material.

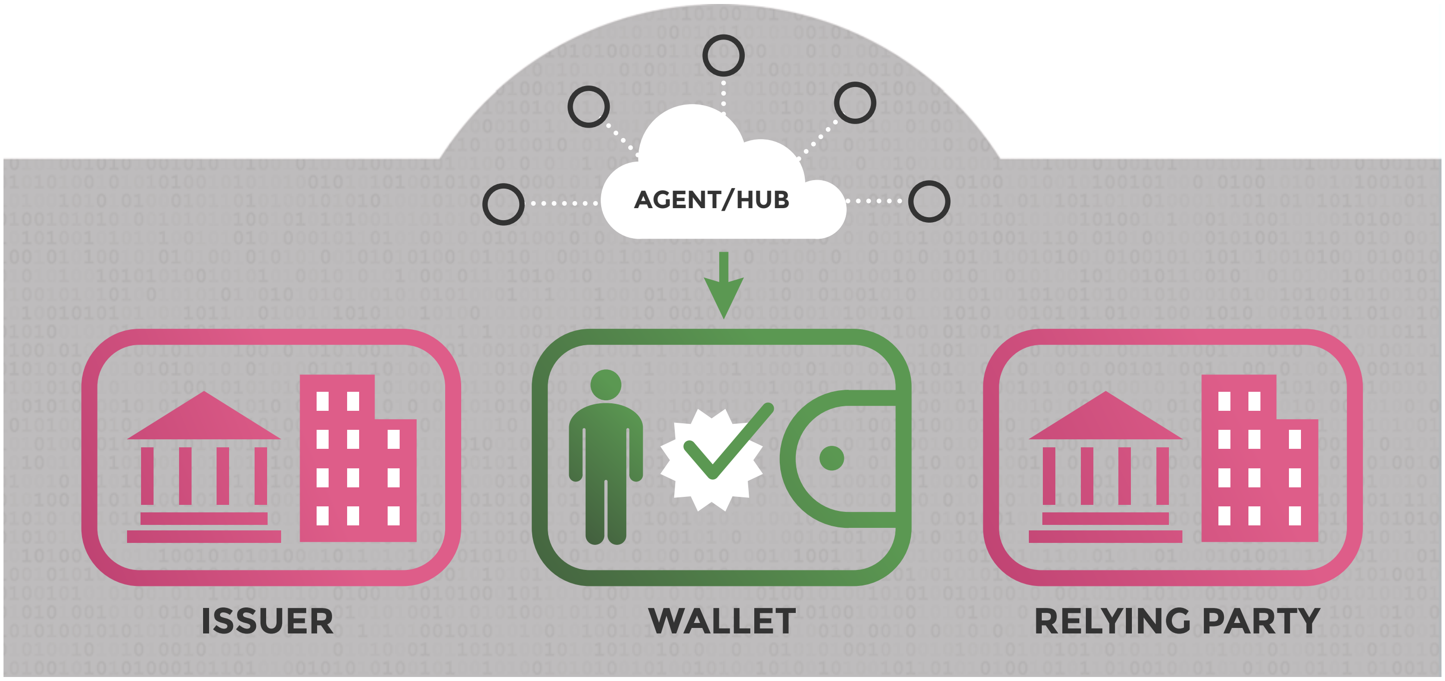


Image based on “A Comprehensive Guide to Self Sovereign Identity”, by Heather Vescent and Kaliya Young

Alternatively, a learner could also interact with a web-based interface, while still using strong authentication via FIDO2 specifications.

Credentials may reside on the same device, or may exist in learner-chosen secure encrypted storage locations that are currently referred to “hubs” or “vaults.” Some reference architectures include Identity Hubs and Encrypted Data Vaults.

An “agent” is the term used to describe software assisting the learner in the management, storage, and exchange of their credentials. The agent understands the protocols, but asks the learner for consent before sharing sensitive data.

Credential Exchange Protocols

Emerging protocols enable these operations and exchanges to happen interoperably. For example, the emerging credential exchange protocols involve an interaction by which:

1. The relying party’s agent communicates what credential schemas (or even issuers) they accept,
2. The learner’s agent parses this information and searches the learner’s data storage for matching credentials, and
3. The learner’s agent packages and sends the credential to the relying party.

Such exchanges are enabled through the draft specifications and protocols Credential Exchange Manifest and Credential Handler API.

Verification Procedure

The process by which a credential is verified is another foundational piece of our approach, and is described in Verify Credential earlier in the white paper.

Flexible Identity Model

Background: paper-based credentials
With paper-based credentials, the question of learner identity is resolved by comparison with other identity documents. For example, the student's full name, birthdate, and birthplace are printed on the document, allowing the requesting party to first look at the certificate and then at a provided official ID. If the identity on the official document matches with the identity on the certificate, the credential is valid.

Establishing issuer identity in paper-based credentials relies on a combination of conventions (watermarks, sealed envelopes) and existing trust mechanisms (direct send of the credential from the university). This approach could be replicated digitally, but there are new, emerging approaches that will provide the same level of assurance with regard to identity attributes.

Flexible Identity Model, supporting the growth of Self-Sovereign Methods
Our approach to identity is influenced by two primary factors:

1. Emerging self-sovereign identity and learner-controlled storage standards promise to pro mote learner (and issuer) control of identity and data.
2. Legislation or other regional requirements may add the need to link a credential with a specific form of identity, such as a national ID. Supporting such cases is necessary for the credential to be *usable* by the learner.

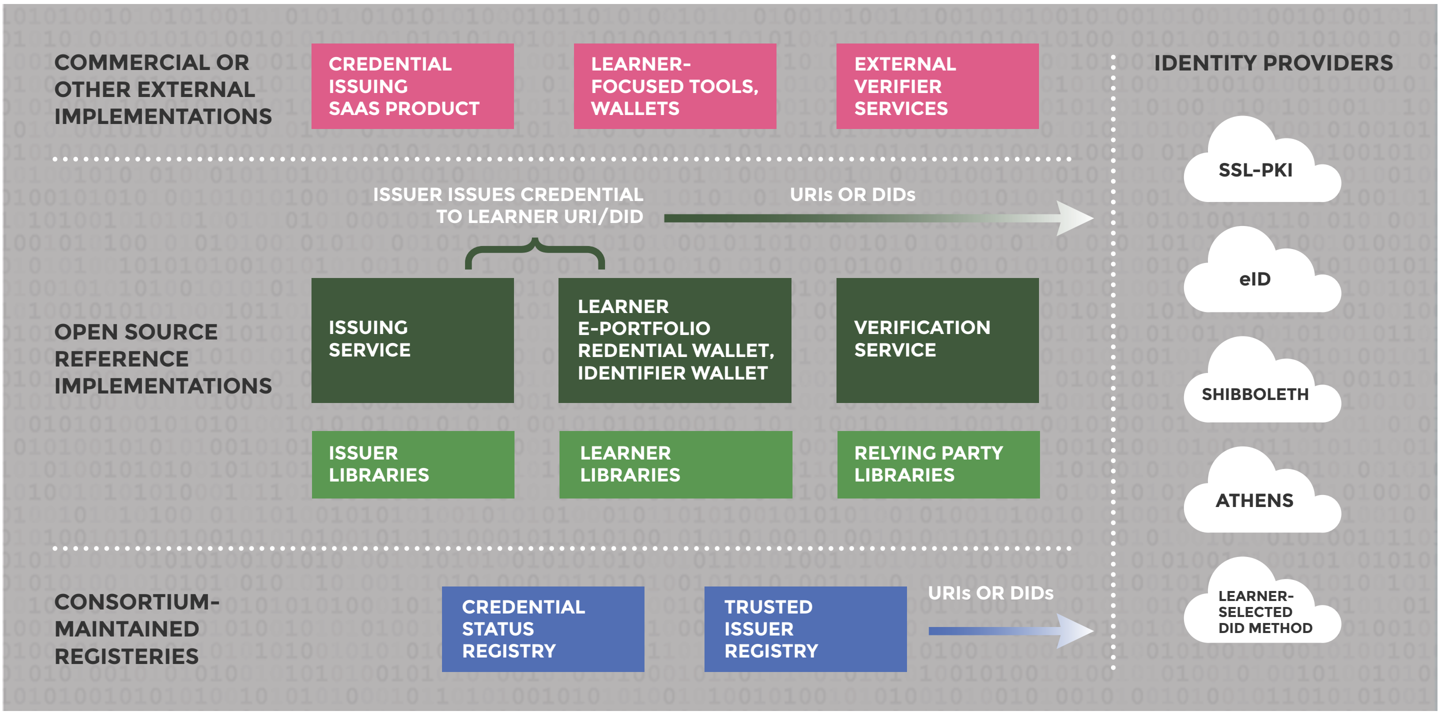
Our solution to this rests on the identifier approach used by the Verifiable Credentials Data Model. The VC Data Model provides a flexible framework for learner identification via URIs, enabling the use of traditional identity schemes as well as new emerging decentralized or self-sovereign identity schemes, such as those compliant with the Decentralized Identifier Data Model.

> *We are evaluating which methods to include in consortium prototypes, and intend to explore methods with a range of characteristics: (1) methods that bridge to existing identification methods used by consortium issuers, (2) open-source, non-commercial methods, (3) interoperable methods provided by implementation partners.*

Components

Building on the foundation of this standards-based approach, we partition the system components as follows:

- Open source reference implementations: These are developed by the consortium, and include reusable libraries and learner-focused tools (such as wallets or e-portfolios) to promote learner control of their credentials and avoid vendor lock-in.
- Commercial or other external implementations: We envision a thriving ecosystem of service providers and technology companies who implement the standards and offer a variety of solutions to issuers, learners, and relying parties. They are free to use the open source reference libraries.
- Consortium-Maintained Registries: These registries must be maintained by the consortium. They are also standards-based and may be adapted by other issuers. These are used during the credential verification process to establish their integrity and authenticity.
- Identity Providers: As described in Flexible Identity Model, URIs or DIDs (as used in the Verifiable Credentials Data Model) enable flexible identification methods that can be chosen (independently) by the issuer and learner dependent on factors such as university convention, regional/local requirements, learner convenience, or even self-sovereign methods. The Identity Providers are external to the system, but the consortium may contribute to open source development of relevant identity methods supporting our use cases (e.g., bridge to eID or Shibboleth or even self-sovereign DID methods).


COMMERCIAL OR OTHER EXTERNAL IMPLEMENTATIONS
CREDENTIAL ISSUING SAAS PRODUCT
LEARNER-FOCUSED TOOLS, WALLETS
EXTERNAL VERIFIER SERVICES
IDENTITY PROVIDERS
ISSUER ISSUES CREDENTIAL TO LEARNER URI/DID
URIs OR DIDs
SSL-PKI
eID
SHIBBOLETH
ATHENS
LEARNER-SELECTED DID METHOD
OPEN SOURCE REFERENCE IMPLEMENTATIONS
ISSUING SERVICE
LEARNER E-PORTFOLIO REDENTIAL WALLET, IDENTIFIER WALLET
VERIFICATION SERVICE
ISSUER LIBRARIES
LEARNER LIBRARIES
RELYING PARTY LIBRARIES
CONSORTIUM-MAINTAINED REGISTERIES
CREDENTIAL STATUS REGISTRY
TRUSTED ISSUER REGISTRY
URIs OR DIDs

These components are described in the following sections.

Open Source Reference Implementations

A number of libraries will be provided as reference implementations by the Digital Credentials

Consortium to enable the use of our new system as either issuer, learner, or relying party. Due to the open-source nature of the system, other implementations adhering to the standard are permitted.

Issuer Libraries
Issuer libraries provide utilities supporting standard-compliant credential issuers. This includes tools for onboarding new issuers as well as for issuing and revoking credentials. Issuer libraries also define APIs supporting integrations with student information systems (SIS).

Learner Libraries
Learner libraries provide reusable functionality to support learner-focused tools and applications such as credential management applications or websites. These libraries support receiving, viewing, and sharing credentials as well as, and further management functions.

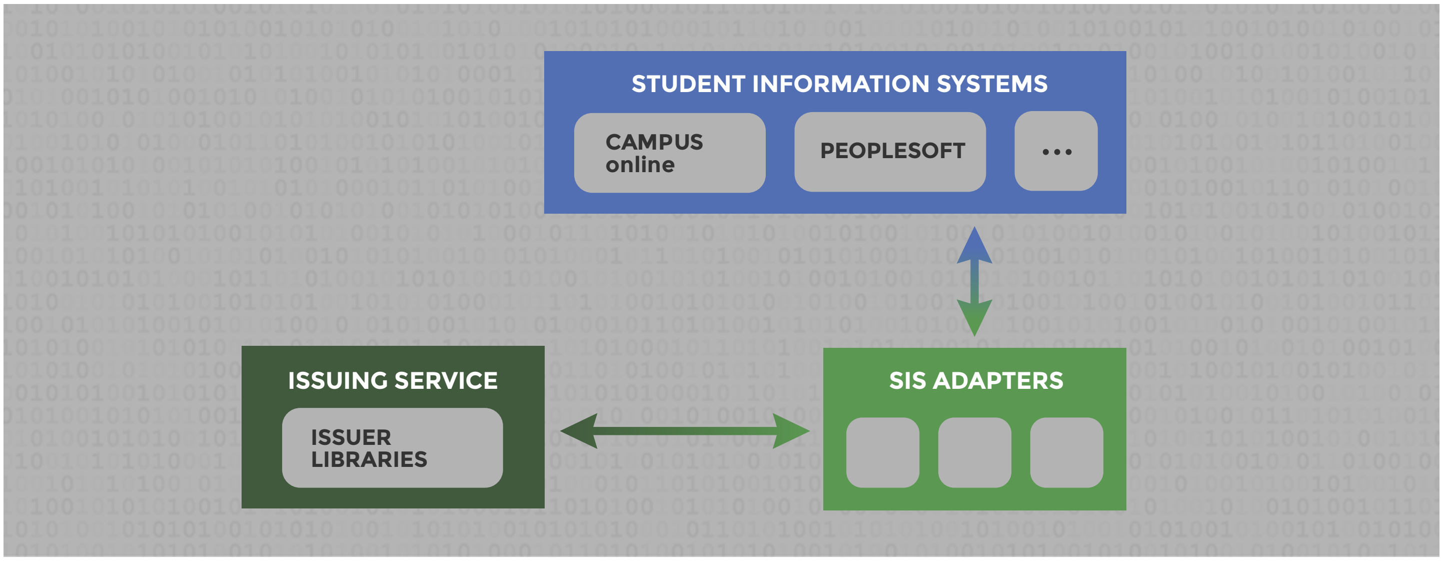
Relying Party Libraries
Relying parties or service providers wishing to verify any shared credentials can use this library to integrate into their systems or build custom verification workflows.

Services and Tools

The consortium will develop and provide services and tools based on the above libraries as reference implementations. The consortium will not offer commercial services or compete with service providers. We encourage and support other organizations to develop products/services based on these standards and libraries. Users may choose to use a standard-compliant alternative.

Issuing Services and Integration into Student Information Systems
The consortium will develop open source issuing tools and services (based on the issuer libraries) to be used by consortium issuers in early stages of the project. These may be adapted and deployed for use by other (non-consortium) issuers.

Central to this effort is ensuring that issuers can easily integrate the issuing tools and services into their existing Student Information Systems and workflows, thereby avoiding burdensome installation of new systems or data processing steps. To that end, the services will expose APIs enabling flexible integration with external data sources. These interfaces will be generic – not tied to any specific provider – so they can be adapted to a range of systems and workflows. Bearing in mind that our consortium members require specific SIS adapters, we will develop and deploy those SIS adapters for our own use and as reference implementations for others to use. Expected prototypes include CAMPUSonline, PeopleSoft, and UGOV adapters.



Verification Service
Verification services will be integrated into sites and registrar systems of consortium members. We will deploy and host verification services for the initial project partners for testing as the standard develops.

Learner-Focused Tools
The biggest gaps in current credentialing ecosystems are around learner scenarios. The consortium will encourage the creation of services and tools for learners in support of secure credential receipt, storage, and exchange.

These are expected to include:

- Standards-based Credential Wallet enabling mobile credential management.
- A web-based credential management interface, like an e-portfolio system, that is either hosted by universities, service providers, or technically proficient users themselves.
- Identity management tools, based on Flexible Identity Approach.

The foundations of these tools are described in Credential Ecosystem Standards and Protocols.

Consortium-Maintained Registries

Some of the functions needed to verify a credential and its issuer are provided by decentralized infrastructure. The main characteristic of these components is that they are publicly accessible and any stakeholder can use them. The write permissions for these components vary depending on the governance model. This allows any stakeholder such as a relying party to hold their own version of the components needed to verify a credential, thus reducing dependency on the infrastructure of third-parties. See Infrastructure for more detailed considerations.

Trusted Issuer Registry
A fundamental challenge with issuing digitally signed credentials is confirming that the public key used to sign a credential does in fact belong to the claimed institution.

To that end, the consortium will maintain a registry of the issuing identifiers for consortium members. The identifiers will be resolvable per the Flexible Identity Model. Our initial implementation of the registry will record DNS names and the associated members' SSL (Secure Socket Layer) certificates (referred to as SSL-PKI, Public Key Infrastructure, going forward), an approach that provides governance levers while also making it easier to later onboard new issuers.

Although the consortium will use SSL-PKI together with the consortium's registry, other organizations may use SSL-PKI with their own registry, or even SSL-PKI alone. Without a registry, however, the responsibility for verifying the authenticity of the signing key is pushed to the relying party—the relying party has to know which issuer identity to trust.

The registry additionally provides a mechanism for the consortium to maintain consistent and thoughtful development of the standard as we work through the early phases of deployment. It is critical to ensure the credentials produced by this effort have the same weight as the traditional (paper-based) credentials. As the system matures, we will work to further decentralize issuer verification.

Credential Status Registry
The issuer may need to revoke a credential after issuance. For cases like this, credential status registries are part of our decentralized infrastructure. Issuers have control over their respective credential status registry where they can update the status of the validity of credentials. During the verification process this registry is consulted. To ensure this registry is privacy preserving, different data structures supporting data minimisation (such as Cryptographic Accumulators, Bloom Filters etc.) will be explored.

External Infrastructure

We recognize there is a need for trust anchors for issuer and learner identities that can’t be unilaterally dictated by our consortium. Further, learner-centric credentialing necessitates learner choice over credential data storage. Our design must therefore accommodate external infrastructures for the following components.

Credential Storage
The learner may choose their desired credential storage location. This may be a cloud storage account controlled by the learner or a trusted third party committed to storing learner credentials securely (for example, the university or third-party vendor may perform this service).

Learner Identity Providers
As described in "Flexible Identity Model," different digital credential projects will have different requirements for identification of the learner in the credential. In some cases, the learner may want to link their credentials to an official identity. This may include student identity systems (Shibboleth, Athens), national eID schemes (eIDAS), or other verified identity providers such as banks (BankID).

It is up to the individual credential project to decide what level of verification is needed to be accepted as an identity endpoint. While identity service providers are external to our system, the interfaces must be standards-based and interoperable. We are using Decentralized Identifiers (DID) to provide a standard interface to these different identity providers and services. DID service endpoints provide trust anchors through different external identity providers.

We will support the development of interfaces to a variety of identity services, as well as interfaces to self-sovereign alternatives where desired by the learner.

IV Other Considerations

Decentralized Infrastructure and Blockchains

As described in Components, a decentralized credentialing ecosystem is more robust, scalable, and flexible than a centralized system. Blockchain, a type of distributed ledger, has garnered excitement and hype–and often subsequent disillusionment–as a one-size-fits-all solution for decentralization. It has become clear over the last few years that a blockchain by itself does not solve all business use cases. However, when used strategically, blockchains, or some aspects of them, can benefit credentialing ecosystems by improving efficiency and trust around issuance and exchange of credentials.

Benefits

Blockchain can provide three main benefits for credentialing:

Redundancy
Decentralized networks (like a blockchain) provide redundancy, thereby reducing single points of failure common with centralized systems. Even though redundantly duplicated across a decentralized network, the data is nevertheless logically centralized. This effectively provides the best of both worlds–a central point of truth for inspecting credential status (such as whether it's revoked) but backed by the entire decentralized network, removing single points of failure present in centralized networks.

In contrast, consider two categories of centralization in existing digital credentialing solutions and their related risks:

- Solutions in which credentials must be retrieved and verified through an issuer (or a third party) website every time the learner shares a credential are vulnerable to service outages, or worse, disappearance (if, for example, the issuer decides not to host these files anymore, or if the issuer goes out of business).
- Some blockchain-based digital credential solutions attempt to remedy the above issue by anchoring a hash of the credential to a blockchain, but it still relies on the issuer (or third party) to host the issuer's public keys and credential revocation list. Verification of these credentials therefore is similarly vulnerable to service outages or failures due to missing file dependencies.

Immutability
Transactions in a blockchain are cryptographically chained to prior transactions, making them difficult to modify or remove after the fact. Any update to the blockchain state requires a new transaction that is visible to all parties in the public network. Independent parties can consequently collaborate and trust that all transactions posed to the ledger are legitimate.

Learners can reliably share their academic credentials and a relying party can verify the integrity of the credentials via the blockchain—without having to consult the issuing institution.

Issuers can reliably manage the status of credentials—say, revoking a credential if it was issued in error—through an immutable and trustworthy source of truth without introducing single points of failure.

Timestamping
To check credential validity, trusted timestamping is a conditio sine qua non. Anchoring information to a blockchain is a simple and reliable means to ensure timestamped credential integrity.

Cautions and Risks

The Gartner report "How to Position Blockchain Platforms to Increase Adoption" (summarized in this article) describes common pitfalls related to enterprise blockchain hype (enterprise blockchains are typically private/permissioned; see Appendix - Blockchain Archetypes for details). The gap between business requirements on the one hand, and the actual features provided by a blockchain solution on the other, can compromise the success of a project and can even lead to blockchain overuse—storing everything on a blockchain—when other solutions would be more scalable and less expensive. This problem is worsened by the absence or infancy of interoperable standards around data stored on a blockchain, risking vendor lock-in.

Worse still, in our use case, overuse of blockchain might compromise the learner's privacy. As discussed in Privacy by Design and Compliance with Frameworks, immutability of blockchains can interfere with the right to be forgotten, as provided by the European Union (EU) General Data Protection Regulation (GDPR).

Finally, the energy consumption and consequent carbon emissions due to current implementations of public permissionless blockchains (e.g., see recent study The Carbon Footprint of Bitcoin) lead to grave concerns about the environmental impact of such blockchain-based solutions.

To be clear, the cost of any individual transaction is not the concern. In fact, some of the authors of this paper have designed public blockchain-based systems that batch credential issuing such that only a single transaction per hundreds of thousands of credentials is needed. Our concern, rather, is holistic – we do not want to design a large-scale system based on an environmentally damaging foundation.

As a loose analogy, it is not the per gallon mileage of our hybrid vehicle we are concerned about, but the environmental cost of the highway (building it, maintaining it, etc). As a group, we have decided that we will not contribute to this problem. We do, however, believe strongly in open systems and protocols, and so would advocate for other consensus mechanisms, such as Proof of Stake, that still allow a public, permissionless foundation without associated adverse environmental effects. Systems incorporating such blockchains do not necessarily affect the underlying energy consumption: among other factors, the price of the underlying cryptoasset increases (or decreases) the incentive for participating as a miner, resulting in an altered energy consumption. Our approach will leverage the benefits of blockchains while avoiding pitfalls and risks described here.

Implementation Plans

Due to the environmental impact of Proof-of-Work consensus mechanisms, we will not deploy the initial system on a public permissionless blockchain.

We plan to take a phased approach:

- Develop and deploy the registries on a public permissioned blockchain.
- Monitor the evolution of consensus mechanisms in public permissionless blockchains and migrate when/if appropriate. We will manage the transition/migration in a way that creates minimal additional effort for issuers or learners. Existing credentials will remain valid and verifiable.

Our preference is to build on open standards and public access (see Appendix: Blockchain Archetypes).

The data stored on the blockchain must be reduced to only what's required for verification, and further adhere to best practices for ensuring learner privacy. This ongoing effort is described in Privacy by Design and Compliance with Frameworks.

Relationship to Other Credential Standards and Initiatives

As pointed out, our goal is to create a trusted, distributed, and shared infrastructure that becomes the standard for issuing, storing, displaying, and verifying digital academic credentials.

Our group has deep expertise and experience in the design of digital credentials and is committed to open source, open standards, and interoperability. We do not want to reinvent the wheel, but to use existing standards, combine them wisely and, if valuable, advance them. Another task is to develop a transparent learner-centric governance model. Members of our group are actively working with standardisation groups to complement existing efforts.

Credentials

- Groningen Declaration
- W3C Verifiable Credentials Data Model 1.0
- OpenBadges
- OpenCerts
- ELMO and EMREX
- Europass new
- IMS Global Comprehensive Learner Record, developed with guidance from the American Association of Collegiate Registrars and Admissions Officers (AACRAO)

Blockchain

- Ethereum & Hyperledger
- European Blockchain Partnership (EBP)

Identity management & Self-Sovereign Identity

- W3C Credentials Community Group
- W3C Decentralized Identifiers
- Decentralized Identity Foundation
- Eduroam & eduGain
- European Self-Sovereign Identity Framework (ESSIF)

Privacy by Design and Compliance with Frameworks

The GDPR as an Applicable Framework

For academic credentials to be useful they must be linkable to real-world identification data, which raises significant issues of data protection and privacy. Placing learner privacy at the core of our design is not just for compliance with legal frameworks, but also an ethical position we choose to take.

Protection of privacy is considered a human right (Universal Declaration of Human Rights, UN, art. 12). However, around the world there are various and inconsistent definitions of and protections for personal data. As a starting point, we follow the definition of personal data adopted for the General Data Protection Regulation (GDPR) by the European Commission, which states that "Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data."

Since we are developing a global solution, we chose to align most closely with the European Union (EU) GDPR, at least initially. At this point, the GDPR appears to prioritize individual rights and establish globally respected, broadly applicable, and widely used standards. For example, the GDPR regulates the protection and control over personal data in the European Union in part by regulating all parties that receive personal data from the EU, even if those parties are located outside the EU. Further restrictions apply for sharing personal data with parties outside the EU and in countries where data protection laws are not considered (by the EU) to be adequate. In any case, it means that the way in which credentialing data is protected outside the EU still can be affected by the GDPR.

This means that compliance with the GDPR should maximize our ability to design a system that is (a) legally relevant for many (if not all) potential users; (b) based on a useful point of reference for compliance with legal frameworks being developed in other parts of the world; and (c) that is broadly aligned with our core ethical principles.

In addition, we use a design process that can respond to the fact that laws and insights into their interpretation with respect to new technologies continue to be in development. We plan to incorporate various experts in reviewing our approach, collaborating with both legal and technology experts, to find solutions. We encourage others to review our approach with their legal counsel, and we welcome comments and feedback to improve compliance with emerging legal standards over time.

Design for Privacy

To plan for GDPR-compliance, it benefits system designers to use a privacy-by-design approach – that is, consider the privacy implications of every decision in the design process. This combines technology design decisions, decisions about the organizational context (how individuals and organizations interface with the technology), and ways to document levels of compliance with the guidelines.

Our privacy-by-design approach follows these three dimensions:

- **Technology Design -** Ensure learner privacy is core to the standards and systems we develop. Give as much prior notice about and control over the use of their data to the learners themselves through the design of the software and processes. For example, we require learner consent when digital credentials are issued.
- **Guidelines and Requirements -** Publish a set of requirements and guidelines for organizations (technology companies, universities, governments) interested in implementing the standard or interfacing with it.
- **Compliance -** Monitor the evolution of laws and regulations related to private data and public ledger technology to keep track of significant developments. Update our solution as needed to reasonably enable compliance with the GDPR and, to the extent reasonably feasible, other legal frameworks.

Technology Design

From a technical point of view, there are many technology design choices that relate to protection of personal data and compliance with the GDPR and other legal frameworks. Minimally, this includes long-established secure data handling practices, such as encrypting data in transit and at rest, access control mechanisms, and time limits after which data is deleted automatically. We believe the the GDPR encourages an even more comprehensive consideration of privacy implications, and therefore will use a privacy-by-design approach throughout the system design. Some examples include minimizing the footprint of personal data—even pseudonymous data (such as IP addresses and salted hashes of data)—that could be re-identifiable. Technical solutions such as zero-knowledge proofs can offer even more protections, reducing the amount of data the learner has to disclose during the validation process of a credential.

Guidelines and Requirements

Other personal data protection measures are more organizational and require clear guidelines describing implementation and use of the system. One example is to require, systematically, obtainment of consent from the data subject (i.e., the person that can be identified by the data) before data can be processed or stored for certain purposes (where consent is legally required), and also a commitment from the issuer and/or relying party of the personal data to only store or process the data for specific purposes.

The guidelines will clarify the baseline rules and regulations that our system and solution are aiming to satisfy. We intend that this transparency should simplify a user's ability to assess the legal sufficiency of the system for that user's own needs. Secondly, the guidelines will explain our perspective on which party, in various cases, we believe or expect would be responsible for protecting data and implementing the guidelines. Finally, the guidelines will discuss what measures may be taken to protect data in various specific cases.

Compliance

In order to document compliance, we attempt to be as clear as possible regarding (1) who is responsible for implementing the guidelines, and (2) what is their legal basis for processing the data. If this is vague, there is the risk that privacy is not protected adequately. Who is responsible will depend on a variety of factors, such as which party is the source of the data, what laws are applicable, and who is in control of the system where the data is stored. With respect to processing, the GDPR recognizes six possible reasons for processing personal data, including not only consent but also "legitimate interests" and when the data is necessary to perform a contract for the data subject. Several rights and obligations of the data subject, controllers, and processors depend on which of the foregoing reasons is the legal basis for processing.

Our work builds on cutting-edge technology, and the laws that should be complied with are often quite new as well. For example, the GDPR only became enforceable in May 2018. This means that in some cases, it may not be fully clear yet how this regulation should be complied with or how it may be enforced in a multilateral and multinational setting. Furthermore, there is sometimes a lack of clarity how the GDPR relates to other laws and requirements, for example the duty of the state to archive citizens' data. In France, education achievements and learning outcomes have to be registered for fifty years. In the Netherlands, "holding terms" apply to transcripts, grades, degrees, etc. There is an obligation to hold transcripts for 2 years, but after that, they should be "burnt."

A specific issue we are considering is how to comply with the 'right to erasure" that is provided to data subjects in the GDPR. According to the GDPR, a data subject can request that an entity controlling their personal data (under the GDPR, a "data controller") erase such data. To minimize the chance of a request for erasure not being honored (even if that outcome may be legal) and personal data being available on a blockchain forever, our approach minimizes the personal footprint on a blockchain from the outset. That said, there remain open questions around the implications of the right to erasure. The right of erasure is not absolute, and how a controller responds to a request will depend in part on why the controller obtained and used the data in the first place, i.e., the underlying legal basis for processing.

We are closely monitoring developments of the European Commission, which is concerned with implementing and enforcing the GDPR, and believe this consortium can help advance the discussion in areas of ambiguity and help evolve our understanding on how to implement the legal requirements of the GDPR and other similar frameworks.

Audit Records

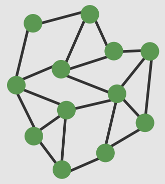
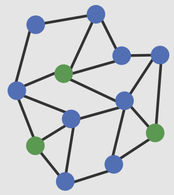
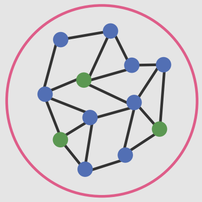
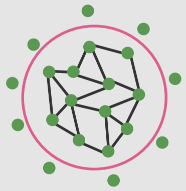
While digital credentialing approaches address some forms of fraud in existing systems, they may introduce others. Our design decisions have the goal of maximizing security and trust in the approach. When deploying a credentialing system, a different category of threats emerges (outside of the scope of the standard and reference libraries). Our reference implementations and best practices guidance will support the ability to perform audits such as subsequently comparing an unexpected batch (reported by a different system). This will be done in a way that doesn't expose additional recipient info—an issuer only needs to know that the issued set is larger than the expected set in order to initiate a batch revocation and reissuance.

Appendix - Blockchain archetypes

A blockchain is a specific type of distributed ledger technology. A blockchain is made up of a decentralized network of individual parties. These parties, called nodes, agree about data stored in the network and the rules under which this data can be changed. Each block in such a chain represents a single state; the newest block represents the current state. The state changes are represented by transactions issued by participants which are stored alongside a block. To fit different needs, a wide variety of blockchain implementations, consensus mechanisms, and architecture models is available.

In a nutshell, the main characteristics of blockchains are:

- Consensus: All transactions are validated by participating nodes; a consensus mechanism ensures the integrity of the ledger.
- Decentralized: Peer-to-Peer (P2P) network, no intermediary required.
- Cryptography: Hashes and digital signatures ensure data integrity and ownership.
- Immutability: Data once written cannot be altered or removed from the network.

BLOCKCHAIN TYPE	EXPLANATION	EXAMPLE	VISUALIZATION
PUBLIC PERMISSIONLESS BLOCKCHAINS	**In these blockchain systems, everybody can participate in the consensus mechanism of the blockchain. Also, everyone in the world with a connection to the internet is able to transact and see the full transaction log.**	**Bitcoin, LiteCoin, Ethereum**	
PUBLIC PERMISSIONED BLOCKCHAINS	**These blockchain systems allow everyone with a connection to the internet to transact and see the transaction log of the blockchain, but only a restricted amount of nodes can participate in the consensus mechanism.**	**Ripple, private versions of Ethereum**	
PRIVATE PERMISSIONED BLOCKCHAINS	**These blockchain systems restrict both the ability to transact and view the transaction log to only the participating nodes in the system, and the architect or owner of the blockchain system is able to determine who can participate in the blockchain system and which node can participate in the consensus mechanism.**	**Rubix, Hyperledger**	
PRIVATE PERMISSIONLESS BLOCKCHAINS	**These blockchain systems are restricted in who can transact and see the transaction log, but the consensus mechanism is open to anyone.**	**(Partially) Exonum**	

Modified from: Table 2, p. 16: ALLESSIE D, SOBOLEWSKI M, VACCARI L, PIGNATELLI F (Editor), Blockchain for digital government, EUR 29677 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76- 00581-0, doi:10.2760/942739, JRC115049 https://publications.jrc.ec.europa.eu/repository/bitstream/JRC115049/blockchain_for_digital_government_online.pdf

Appendix - Terminology Alignment

Verifiable Credentials Data Model

In this paper, the term learner is used interchangeably with a Verifiable Credential subject. Further, the subject is the same as the holder in the use cases described in this paper.

The following terms are equivalent to Verifiable Credentials Data Model terminology: credential, verifiable credential, issuer, relying party.

In this paper, we haven't called out the separate role of verifier; instead we chose to emphasize the verification process to avoid additional complexity by calling out this separate role. However, as with the Verifiable Credentials Data Model, the standard allows the roles of verifier and relying party to be separate.

Verifiable Credentials Use Cases

The tasks described in section II are based on Verifiable Credentials Use Cases User Sequences.

NIST Draft Whitepaper: Emerging Blockchain Identity Management Systems

Our terminology and architecture aligns with NIST's whitepaper (draft) "A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems" (https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.07092019-draft.pdf), as follows.

Our architecture is most closely described as "Off-chain Objects coupled with Global Identifiers Registry for Issuers"; "Off-chain Objects coupled with Global Credentials Registry".

The Trusted Issuer Registry corresponds to the "Global Identifiers Registry" for issuers. Credentials are stored as off-chain objects. The Credential Status Registry is used as in the Off-Chain Objects coupled with Global Credentials Registry. The Issuance Registry is an Anchors Registry.

References

- W3C Verifiable Credentials Data Model 1.0: https://w3c.github.io/vc-data-model
- W3C Verifiable Credentials Use Cases: https://w3c.github.io/vc-use-cases/
- W3C Decentralized Identifiers: https://w3c-ccg.github.io/did-spec/
- W3C Decentralized Identifier Method Registry: https://w3c-ccg.github.io/did-method-registry/
- NIST A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.07092019-draft.pdf
- How to Position Blockchain Platforms to Increase Adoption: https://www.gartner.com/document/3909078?ref=solrAll&refval=221177808&qid=b96eab586d1bd06d8898f
- Gartner reveals seven mistakes to avoid in blockchain: https://www.gartner.com/en/newsroom/press-releases/2019-06-12-gartner-reveals-seven-mistakes-to-avoid-in-blockchain
- “A Comprehensive Guide to Self Sovereign Identity”: https://www.amazon.com/Comprehensive-Guide-Self-Sovereign-Identity-ebook/dp/B07Q3TXLDP?SubscriptionId=AKIAILSHYYTFIVPWUY6Q&tag=duckduckgo-brave-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=B07Q3TXLDP
- European Qualifications Framework https://www.cedefop.europa.eu/en/events-and-projects/projects/european-qualifications-framework-eqf
- IMS Global Comprehensive Learner Record https://www.imsglobal.org/activity/comprehensive-learner-record
- W3C Software and Document License https://www.w3.org/Consortium/Legal/2015/copyright-software-and-document
- Decentralized Identity Foundation Identity Hubs https://github.com/decentralized-identity/identity-hub/blob/master/explainer.md
- Encrypted Data Vaults https://github.com/WebOfTrustInfo/rwot9-prague/blob/master/final-documents/encrypted-data-vaults.pdf
- FIDO/FIDO2 specifications https://fidoalliance.org/fido2/
- W3C Web Content Accessibility Guidelines https://www.w3.org/TR/WCAG21/