Towards Academic and Skills Credentialing Standards and Distributed Ledger Technologies

Morné Pretorius[a], Nelisiwe Dlamini[b] and Sthembile Mthethwa[c]
Council for Scientific and Industrial Research (CSIR),
Information and Cyber Security Centre (ICSC),
Pretoria, South-Africa

Keywords: Distributed Ledger Technology, Blockchain, Education, Standardisation, Standards, Verifiable Credentials, Skills Tracking.

Abstract: Today's internet-connected world is moving towards evermore digitisation. Consequently, the education system globally is experiencing various problems whilst trying to keep up with this disruptive and ongoing change that is introduced. One way to alleviate the problem is standardising how skills and academic achievement are quantified, digitised, authenticated and persisted to achieve a means of automated verification and matching of the current need with what skill-sets are available. This research aims to provide a starting point towards a standardised future solution which considers existing emerging standards and technologies to provide skills tracking capability. The existing standards, data schema, technologies and techniques are discussed and an existing real-world prototype architecture is identified. This prototype's terminology is then mapped to the emerging World Wide Web Consortium (W3C) standards which will serve as a baseline design.

1 INTRODUCTION

During hunter-gatherer times, there was no separation between learning and playing and thus education used to be an exploratory pass time, which later changed with the need for children to do forced agricultural labour (Gray, 2008). As agricultural automation increased, the need for childhood labour began to decrease, which freed up children's time upon which the concept of compulsory childhood education was gradually introduced from the 16th to 19th century. Even today education is predominantly seen as a separation of work and play, or rather: work, then play (Gray, 2008).

Traditional 19th and 20th-century education systems have served the world for decades by developing students' skills, availing systems that are set up to provide training and awarding their learning with recognition and qualifications. Consequently contributing to the preparedness and readiness of the workforce.

Digital advancements have reshaped the world and led to innovations in education and training. This changes traditional classroom-based learning in various ways by introducing new learning methodologies or, is perhaps leading us closer to the original hunter-gatherer way of education. This learning is backed by experience accumulated throughout a person's lifetime, skills developed by completing projects, lessons learned via online teaching systems, and self-guided information gathering (The Mozilla Foundation et al., 2012). These help the person to keep abreast of current skills and obtain knowledge, especially in a time where 21st-century skills (navigating and detecting truth in vast collections of online websites and information sets) are required by most if not all employers. This self-directed interest-driven learning aims to (The Mozilla Foundation et al., 2012):

- establish a broader accreditation and recognition ecosystem;
- granularly track traditional curriculums along with new skills and literacies;
- inspire and assist students to seek new learning avenues;
- capture obtained skills and achievements across different contexts;
- expose or communicate captured skills and achievements to relevant stakeholders and
- facilitate cross-context learning.

[a] https://orcid.org/0000-0002-7665-4778
[b] https://orcid.org/0000-0002-8635-8948
[c] https://orcid.org/0000-0001-7961-5240

Pretorius, M., Dlamini, N. and Mthethwa, S.
Towards Academic and Skills Credentialing Standards and Distributed Ledger Technologies.
DOI: 10.5220/0010340602490257
In *Proceedings of the 7th International Conference on Information Systems Security and Privacy (ICISSP 2021)*, pages 249-257
ISBN: 978-989-758-491-6

One problem with the existing education system is the continuous proliferation of information and disruptive technologies, where these education systems often find themselves lagging and do not recognise post-curricular or informal skills and knowledge gained as the learner or person progresses. Instead, educational institutions mostly require a learner to be enrolled to study on a formally defined learning path. The formal achievements of the learner are awarded, leaving out potentially important skills, knowledge and competencies obtained throughout self-directed or lifelong learning.

There is thus an urgent need to establish an internationally recognised method in which all skills and knowledge can be consolidated, digitised, authenticated and persisted in a manner where the learner has control over their information and then gets to amend or update it (Chakroun and Keevy, 2018).

Distinctions should be made to avoid unintended debates around terminologies, most prominently known as macro-credentials and micro-credentials (Chakroun and Keevy, 2018). Macro-credentials refer to the more traditional and formal super-sets that represent university degrees and micro-credentials refer to subsets that could make-up (or stack) an equivalent macro-credential. For example, a macro-credential could be an engineering degree whereas a micro-credential could be the subjects that make-up the overall degree or, it could be a series or selection of short courses that form the equivalent of the subjects needed to represent an engineering degree.

One example of micro-credentials is that of Massive Open Online Courses (MOOC) which started surfacing as early as 2007 (Table 1) and has gained traction with an increase from 16–18 million in 2014 to about 35 million registered users in 2015 (Chakroun and Keevy, 2018). One challenge that arises with introducing MOOCs is to find a means to match the equivalent of what MOOC represents to that of what macro-credentials represent. If this can be done, then learners can be notified of the amount of work they need to do to achieve an equivalent university degree and how far along a particular career path they are which could foster motivation towards formal, online or practical achievements. Another challenge is to ensure the integrity and fairness of MOOC when matching the equivalent competency needed to achieve a formal or traditional degree and to make this equivalent legible to an employer when the learner reaches a point of employability (Chakroun and Keevy, 2018). An advantage that could arise from a digitised skills and achievements track record is to match the current labour market need to what skills and knowledge are available. However, doing this implies analysis of the credential data which implies access control, anonymisation and encrypted storage if privacy is of any concern to the public to comply to the Privacy by Design (PbD) requirements of the General Data Protection Regulation (GDPR) (Raul et al., 2017; DLA Piper, 2017). In the South African context, there are also requirements for mandatory and voluntary disclosure of personal information and data safeguarding under the Protection of Personal Information Act (POPIA) (DLA Piper, 2017).

Table 1: List of online micro-credential issuer initiatives.

Year Founded	Country	Initiative
1997	USA, Australia	IMS Digital Credentialing
2007	Ireland	ALISON
2008	Germany	iVersity
2011	USA	Udacity
2012	USA	Coursera
2012	USA	EdX
2012	United Kingdom	FutureLearn
2012	Japan	Schoo
2012	Spain	UniMOOC
2012	Brazil	Veduca
2013	China	Ewant
2013	China	XuetangX
2013	Australia	Open2Study
2013	Saudi Arabia	Rwaq
2013	Russia	Universarium

Table 1 lists some current MOOC initiatives, which illustrate the early stages of credential digitisation. The problem with these issuers is that the credentials they issue are hosted by themselves on central data repositories which makes these credentials and claims (e.g. macro- and micro-credentials) about an individual less discoverable and interoperable with the larger education ecosystem than it could be. It is also difficult to link these obtained credentials to a persistent identity, as they usually use e-mail-based authentication methods to service their clients or learners. Furthermore, from the perspective of more traditional certificate authorities and education institutions that issue macro-credentials, they mostly manage their learner data, courses and results using inefficient and centralised systems such as excel sheets (Gräther et al., 2018). This makes it tedious to search for information and present an overview which also induces manual verification.

This research is not about any particular MOOC or their associated strengths and weaknesses but rather aims to illustrate the role that accompanying

technologies, such as Verifiable Credentials (VC) and Distributed Ledger Technology (DLT), could play in macro-credential or micro-credential services' advancement. The goal is to serve the research community towards either re-using existing efforts or to assess the efforts that will be needed to implement a digitally verifiable credentialling system in future. Ideally, this system should comply with data privacy legislation and existing standards. Regarding interoperability, most DLTs lack standards adoption as most are siloed with no architecture, software design and interoperability unification approach (International Telecommunication Union (ITU), 2019). Associated standards and interoperability requirements that are recognised and accepted internationally are still emerging and can serve as a baseline for user protection and performance of these technologies (Faridi, 2020). The contribution of this study can therefore be summarised as follows: a) A technology prototype architecture which incorporates novelty by complying with emergent standards using emergent technologies and with user centric data control as required by recent data regulations. b) Establish a baseline design through mapping existing prototype terminology and components to standards which will allow other sectors or groups to learn from and re-use ideas. The required system then presents multiple novelties because traditional systems differentiate themselves to this one in various ways:

- All issued credentials and identities are hosted on a central data repository, which is susceptible to a single point of failure.
- Traditional systems make it difficult to link credentials to a persistent identity as they usually use email-based authentication methods to interact with users, i.e. learners.
- Identities are less discoverable and interoperable and are susceptible to censorship.

Section 2 briefly outlines our methodology. Section 3.1 summarises the concepts that are a starting point towards using standard terminology alongside technologies that would improve future education ecosystem interoperability. Section 3.2 discusses what relevant data schemas are available to be used within the skills and academic achievement context. Section 3.3 discusses paths towards utilising these technologies to improve existing micro-credentialling services and how they relate to blockchain or DLT. Section 3.4 briefly touches on what methods are available to selectively disclose subsets of credentials, which is a requirement by more recent regulations such as the GDPR and POPIA. Section 4 provides a mapping to the W3C standard's concepts as well as recommendations to further decentralise credential issuer authorisation and section 5 concludes with future work and ongoing challenges that are to be resolved if such a system is to be instantiated successfully.

2 METHOD

Information Systems (IS) find themselves instantiated in the natural world because of previous design choices that aim to satisfy human requirements. This has led to the research method know as Design Science Research (DSR) (Hevner and Chatterjee, 2010) and is the core method that we are required to use as we aim to either bring new systems into being or improve on existing systems. Where natural- and behavioural science aim to hypothesise, collect data, prove a hypothesis and develop theory, the DSR method aims to understand and improve or produce a new artifact which leads to its evaluation after instantiation (Hevner and Chatterjee, 2010). This work is the first phase, or rather conceptual and architectural assessment towards what is required to satisfy a need established by (Chakroun and Keevy, 2018). It is a push to move from what has and is being abstractly defined in the W3C emergent standards towards a concrete instantiation of one particular sector use case (Otto et al., 2019). We therefore aim to answer various questions:

1. What technologies and specifications exist that are needed to create an international credentialling and skills tracking system?
2. What technologies exist that enable the solution to minimise data disclosure (PbD) for regulatory compliance whilst still providing discoverability of the system actor identities?
3. What standards exist that could provide future international interoperability and also aid in automated digital credential verification?

Various searches were conducted to identify architectural components that are required to realise the required credentialling and verification system. Sources presenting trends in education digitisation efforts, academic and online course achievements standards along with DLT-based digitisation efforts in the education context were collected. The sources were then filtered based on their ability to align with the architectural requirements defined by leaders or pioneers in the educational digitisation field and standards advancement entities.

3 REQUIRED SYSTEM TECHNOLOGIES

3.1 Verifiable Credentials

VCs are an emerging W3C technology and standard that enables online entities to present claims or attestations about themselves or other entities in a cryptographically verifiable manner. This makes it possible for micro-credential services to prove that a particular qualification or competency claim is associated with an identity, or in the standard's context, a Decentralised Identifier (DID) (Reed et al., 2020). A DID aims to solve the single point of failure problem associated with centralised solutions where traditional identities are maintained by some custodian or central authority, for example, the Department of Home Affairs (DHA) or a certificate authority in the current internet ecosystem. Each DID can refer to one or several DID documents that contain meta-information about an entity or subject such as public keys (verification methods or contact information).

In the micro- and macro-credential context these DID documents and VCs can store academic or life long learning achievements or subsets of achievements. The VC and DID standards are agnostic of the physical data storage location of the DID and its associated document(s). This location is referred to as a Verifiable Data Registry (VDR) which is controlled by a controller which can be a person, organisation, autonomous software, logical thing or multiple of these, each with their own DID. This DID controller can also be the person (also referred to as the DID subject) which provides the person or subject the ability to be in control of their own verifiable claims or attestations. The standard is also agnostic of many implementation details, such as the cryptographic primitives used to generate a DID, how to persist credential- or meta-data related to a DID, and how to resolve or lookup a DID. Instead, generic syntax and requirements are defined to execute the basic Create, Read, Update, Delete (CRUD) operations on DID associated meta-data, which in the skills tracking context is the information that makes up a macro- or micro-credential.

It is important to realise the abstract nature of the four basic roles that are present within the VC standards as depicted in Figure 1. There is the *issuer* who generates a claim about a *subject* and binds it to that particular *subject*. Then there is the *verifier* who computes the cryptographic proof to verify claims about a particular *subject*, and finally, there is the *holder* who could be the *subject* or another entity

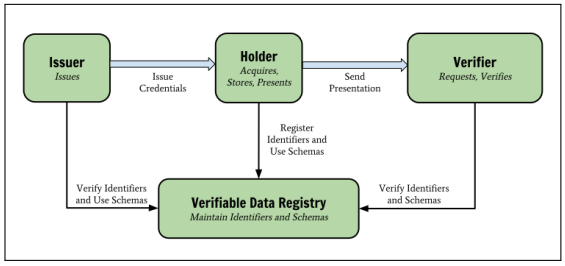

Figure 1: W3C VC data model ecosystem overview (Sporny et al., 2019).

(organisation, autonomous software vault) that acts as custodian on behalf of the *subject* (Otto et al., 2019). All the entities that fulfil these roles are referenced or referred to via their own DID. The abstract and variable nature of these roles are intended to extend the VC standard well beyond the skills tracking context and into healthcare, finance, retail, legal identity and devices and is beyond the scope of this article. Here we focus only on the developments that are in progress regarding the digitisation of the education sector.

3.2 Education Credentialing Schemas

From Table 1 it is evident that the pioneers in education technology and interoperability standards advancement are the Instructional Management Systems (IMS) global learning consortium, which has been active since 1997. They have initiated the development of the digital education credentials data schema specification known as Open Badges (OB) (Consortium, 2018b) which has been in active practical use for close to a decade (Otto et al., 2018) and is an open specification that is maintained and updated across many organisations such as Campus Labs, Credly, Mozilla, Digitalme, D2L Corporation and Pearson (Consortium, 2018b). A worldwide view regarding OB usage can be accessed at badgetheworld.org and a list of badge *issuers* are presented at openbadges.org.

The OB specification serves as lexicon or vocabulary that can structure the data that make up skills or educational achievements, much like how website developers can use schema.org to structure data on their pages to allow search engines to more easily discover or crawl and understand published data. This discoverability is also one of the ten design goals of the W3C VCs specification (Reed et al., 2020). These schemas define the make-up of the information to establish a form of re-use and interoperability. There are other significant similarities in how OBs function when compared to W3C VCs which lays good groundwork for future

standardisation and interoperability. This is due to an underlying data format that is used by both OBs and VCs known as JavaScript Object Notation-Linked Data (JSON-LD) and in addition, JSON-LD signatures which are used to authenticate claims or credentials cryptographically. JSON-LD provides web-developers the ability to structure their data according to a pre-defined data schema which is maintained by affiliates from Google, Microsoft, Yahoo and Yandex in the case of schema.org.

From an identity perspective, OBs use a string-type identifier (predominantly e-mail addresses) to associate an assertion or credential with an *issuer* and/or *subject*, although there are efforts underway to support the alternative DID-based identity integration within the "Open Badges Validator" (Otto et al., 2018). This would de-couple the OBs from central points of failure and provide badge holders more freedom as to where their badges are kept and provide decentralised verification in the future. Currently, OBs utilise centralised verification methods (`HostedBadge` and `SignedBadge`) which are only associated with an e-mail or centrally hosted identity which is more prone to attack and compromise. Two methods or options are listed by (Otto et al., 2018) that could align the OB schema or data model with that of the W3C DID and VC standards which would advance the efforts of the IMS digital credentials initiative in terms of security although a majority of the credential information remains only authenticated and not encrypted. It is thus important to consider the privacy guidelines defined in the W3C VC documentation to prevent domain cross-correlation and to provide sufficient anonymisation, particularly when references and proofs of credentials are stored on an immutable DLT network infrastructure which persists the data. This could stand in contrast to what is required by the GDPR which grants individuals the right to be forgotten (Raul et al., 2017) and requires credential mutability or revocation.

OBs also feature a badge "baking" specification (Consortium, 2018a) where micro- or macro-credential data is merged or embedded into an image file format such as a Portable Network Graphic (PNG) or Scalable Vector Graphic (SVG). This functionality is also consolidated with the W3C VC standards since it is currently represented as a JSON-LD document but should ideally also be representable as a DID document. Another analogous concept within the OB specification is the badge backpack which relates to the *holder* defined in the W3C VC standards. The OBs ecosystem also cannot currently issue the same `BadgeClass` or badge type from a different *issuer* (Otto et al., 2018) which is analogous to, for example, all chemical engineering degrees being issued by one university only. If these similarities can be aligned with the W3C DID and VC standards specifications, then the OB schema could become the de facto vocabulary to digitise education and skills or competencies and would be on the right track towards interoperability across the different *issuers*, *subjects*, *holders* and *verifiers*.

3.3 Distributed Ledger Technologies

A DLT is defined by the W3C as follows: *"A distributed database in which the various nodes use a consensus protocol to maintain a shared ledger in which each transaction is cryptographically signed and chained to the previous transaction"*(Reed et al., 2020). A DLT or blockchain is a form of public utility with two predominant operational settings, namely permissionless and permissioned. In the permissionless setting, anyone who wishes to take part in hosting the service is incentivised to join in, which provides the service by the people and for the people. In the permissioned setting, the number of participatory nodes are agreed upon beforehand and are trusted to provide the shared information service.

The intersection at which DLTs meet the W3C DID and VC standards begin with a DID method, which defines the implementation details necessary to make decentralised authentication possible instead of having central points of failure that host identities. Noteworthy is the ability of the W3C DID documents concept to also contain biometric service provider contact information to verify identity authenticity against a biometric vector as a multi-factor.

Many DID implementations have been defined by DLT development teams and are listed on the DID methods registry. If the OB data schema updates suggested by (Otto et al., 2018) are accepted by the greater IMS digital credentialling developer community and applied, it would make it possible to utilise any of the listed DLT networks to provide identification alongside the IMS OB skills and academic achievements data schema. By analysing the DID methods registry, as of June 2020, there are 58 different provisional DID method implementations in total. There are 3 implementations that run on the Bitcoin network, 4 for Hyperledger Fabric and 9 for Ethereum. Two of the 9 Ethereum implementations can also operate on additional networks from an interoperability point of view. The `did:signor` method can additionally operate on the Hedera Hashgraph, Quorum and Hyperledger Besu networks and the `did:gatc` method can function on the

Hyperledger Fabric, Hyperledger Besu and Alastria DLT networks.

The question that needs to be asked is what does the education credentialling context stand to gain when using a distributed system such as a DLT? The answer lies in the assessment around what properties are provided by these technologies? A DLT typically possesses 3 significant properties:

1. It provides a *temporal ordering* of events or claims without the need to synchronise the clocks of the network nodes that provide or host the network ledger service infrastructure and acts as a form of public utility.
2. It provides *immutability* that ensures historical integrity by including hashes (unique identifiers of a data frame) of the previous data frame or block in the next data frame and by non-deterministically gossiping this to all other nodes.
3. It provides *persistence* of data and resilience by redundantly storing multiple copies of the data in a distributed manner.

What is to be gained by the digitisation of any information is firstly a reduction in administrative overhead through automation, and this surely is a need for most traditional education institutions that still conduct paper-based verification. From the perspective of the MOOCs, who already possess digitisation, they stand to gain a form of globally recognisable *persistent* identity along with *persistent* credential information provided that these identity implementations comply with the aforementioned standards and that global consensus is reached around the adoption of said standards. The entire education ecosystem can gain historical information unification through the *temporal ordering* property, as they do not need to keep track of which claims were made in which order or at which time. This should also reduce fraud via automated auditability because of the *immutable* event history that resides on the ledger. Cryptographic verification also provides additional automation through the *immutability* property, along with built-in authentication through the cryptographic signatures that are linked to a DID.

Although there are many efforts underway (according to existing surveys) to provide education and skills digitisation and verification in conjunction with DLTs (Hameed et al., 2019; Yumna et al., 2019), very few of them aim towards the DID and VC standards and very few of them utilise the same standards-based roles terminology in their descriptions. One project that stood out in terms of alignment with the mentioned standards and the IMS OB data schema specification (Labs, 2019), was *Blockcerts* by Massachusetts Institute of Technology (MIT) Media Labs, which has submitted a proposal (Ronning and Chung, 2019) towards alignment and compliance.

Another project that stood out is the *Blockchain for Education* platform defined by (Gräther et al., 2018) because they incorporated the IMS OB data schema within their implementation. Also notable was the requirements they gathered through real-world surveys which identify role-players within the more traditional macro-credential setting as well as components for DLT inclusion.

3.4 Selective Disclosure

Recent data regulations require PbD and user centric data access control, which can be addressed by incorporating selective disclosure into the design. Selective disclosure is also related to the concept of generating a subset of the fields within a particular attestation or credential that claims something about a *subject*. In W3C terminology, this is referred to as a verifiable presentation. Because information can be easily copied from one hard-drive to another, it is more secure to generate a view or presentation of the data and to minimise the amount of information that gets disclosed.

There seem to exist two methodologies to achieve selective disclosure to the best of our knowledge. One is to use a zero-knowledge proof as described in (Sonnino et al., 2018) and the other is to pack the information into a Merkle tree data structure as described in (Patka, 2019). Zero-knowledge proof cryptography is iterative and asymptotically approaches the certainty that a claim is valid without revealing the actual knowledge or information itself, and care should be taken to ensure that computational efficiency is not compromised that the system's performance remains sufficient. Merkle tree proofs reveal and expose a subset of the VC to the verifier, but is more computationally efficient.

4 DISCUSSION

Standardisation is required to achieve the urgently required and international credentialling system described by (Chakroun and Keevy, 2018). What we found missing in existing DLT-based credentialling systems are the alignment with current internet identity and credentialling standards. These standards relate to the concept of Self Sovereign Identity (SSI) and provide a censorship resistant alternative to current e-mail or password-based identity. The

following lists are an attempt to relate the components and actors defined by (Gräther et al., 2018) and shown in Figure 2 to the concepts defined in the W3C DID and VC specifications and standards where actors can instead be identified by their own persistent and life-long identity or DID:

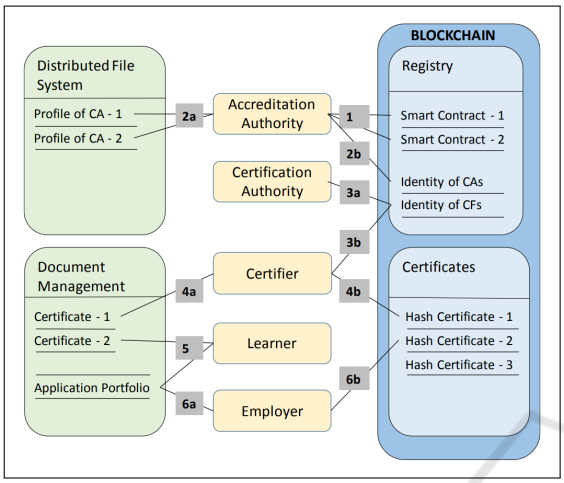

Figure 2: *Blockchain for Education* architecture (Gräther et al., 2018).

- *issuer* - Accreditation Authority (AA) which is identified by a quorum of authoritative entities (multiple DIDs), issues a credential to the authority at the next level down the hierarchy (see Figure 3). This authorisation credential could contain a multi-signature or threshold signature (Hwang and Chang, 2005). The signature's hash is calculated which the AA timestamps on the DLT so that anyone that looks up any k of n total signer DIDs (Gräther et al., 2018), and their associated public keys, can verify the authenticity and temporal or historic issuance.
- *issuer* - Certification Authority (CA) issues a certificate to the Certifier (MOOC or academic institution) and stores the hash of this certificate on the DLT. This certificate could also be verified by the learner in a similar way, except for using only a single DID and public key associated with the Certifier.
- *issuer* - Certifier which issues and signs the academic credentials achieved by the learner and timestamps its unique hash on the DLT.
- *subject* - Learners or certificate recipients that can present their DID to the employer or perhaps include a barcode or QR-code on their Curriculum Vitae (CV) that resolves to their DID and consequently their credentials under cryptographic conditions.
- *verifier(s)* - Employers that can resolve the learner's DID and *DID document* which contains the same hash that is verified against the hash that was stored on the DLT by the Certifier. This temporal integrity check includes the learner and issuing institutions DID associated asymmetric cryptographic signatures which could also be verified.

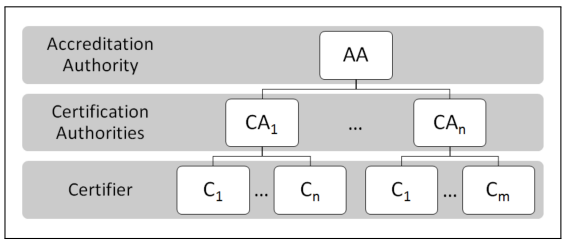

Figure 3: *Blockchain for Education* Identity Hierarchy (Gräther et al., 2018).

From the aforementioned list in conjunction with Figure 3, it is shown that there are three layers of *issuers* in the identity hierarchy. The information formats associated with the DIDs are as follows:

- *DID Document* - CA Profile Information and associated public key(s) used to establish a secure communications channel with a system actor.
- *Verifiable Credential (VC)* - Learner certificate or macro-credential or micro-credential information with attached cryptographic proofs such as signatures.

Entities that hold and persist information related to the actors:

- *holder* - Document Management System in this instance is centralised and could be further decentralised by a DID document that resolves to one or more VC, stored in an encrypted distributed vault.
- *holder* - Distributed File System relates to a storage mechanism that holds any public information such as the DID document which holds actors' public keys that establish a secure communications channel between actors.

Both the *Blockcerts* and *Blockchain for Education* platforms store only the hash of the claims or credentials on a blockchain, which makes sense from a privacy and GDPR compliance perspective where the blockchain merely acts as a temporal (immutable event history) and authenticity verification mechanism. This also limits the overuse of blockchain storage capacity since it is a form of public utility in the permissionless DLT setting. Both implementations currently use public keys

as identifiers showing scope for DID inclusion as proposed in (Ronning and Chung, 2019), because their "strictly hierarchical" identity architecture is prone to a single point of compromise at the AA (Gräther et al., 2018).

Another notable system proposed by (Gresch et al., 2019) is focussed to suit specific needs of the University of Zurich (UZH). It does not introduce any form of standardisation across other educational entities to bring them in alignment. Particularly since various organisations are now exploring standardisation to enable interoperability of technologies and improvements in identity. The lack in alignment with standards is concerning, since standards could prove very helpful, i.e. harmonizing terminology and definitions used, minimising redundancy, etc. What we have found valuable was the overlap in requirements from (Gresch et al., 2019) and (Gräther et al., 2018) which will be used in the second DSR iteration to consolidate lower level requirements for further analysis and particular technology selection.

5 CONCLUSION AND FUTURE WORK

This research has focussed on the required technical components, standardisation efforts and data schema specifications that are currently underway, and how they relate to the use of DLTs in education and skills credentialling. DLTs provide a single source of truth that could bring international education authorities in alignment with each other and aid in standardisation and auditability across ecosystem players while automating credential verification processes. DLT could solve the data proliferation problem where information systems currently have duplicate data that operate in centralised and mostly isolated silos.

Technological components are discussed that are required to achieve the end goal of digitising, consolidating, authenticating and persisting all attested knowledge and achieved skills alongside a life-long persistent identity. Current prototypes that align with required technologies and that use DLT in education credentialling are identified and mapped to the W3C emerging standard terminologies and will serve as a starting architecture towards implementing a credential and skills tracking system that complies with current data privacy laws within our research and development group. The next phase in this DSR process is to consolidate specific requirements from literature and assess DID compliant DLTs in terms of their ability to satisfy these requirements.

Some problems cannot be solved by technologies such as DLTs and cryptography until a widespread agreement is reached around how skills and knowledge can be recognised and quantified through frameworks to enable skills matching to the market need as well as matching the make-up of traditional macro-credentials through standardisation. One example effort in this regard is the Competencies and Academic Standards Exchange® (CASE®) from the IMS global learning consortium and its associated competency framework tool: OpenSALT. Further analysis and comparison are required between competency measurement frameworks such as:

- World Reference Levels (WRL) from the United Nations Educational, Scientific and Cultural Organisation (UNESCO).
- European Qualifications Framework (EQF) from the European Quality Assurance Register (EQAR) and the European Association for Quality Assurance in Higher Education (ENQA).
- National Qualifications Framework (NQF) from the Department of Higher Education & Training (DHET) in South Africa.
- Connecting Credentials from the Lumina Foundation in the United States of America (USA).

Notably, the South African Qualifications Authority (SAQA) has begun an initiative to align the WRLs with the South African NQF but it is still unclear how far they have progressed (Jaftha and Samuels, 2017; South African Qualifications Authority (SAQA), 2018). The accreditation bodies that have implemented the listed frameworks could form the *issuer* hierarchy from Figure 3 or an accreditation group. If at least the *issuer*, *subject* and *verifier* entities from Figure 2 are each assigned a DID, the signature issuance hierarchy from Figure 2 could be replaced with a multi-signature or threshold signature strategy to form a flat or ring structure of signatures where anyone with a DID could verify the issued credentials.

One concern from an implementation perspective is the amount of DID document resolutions that might be induced when implementing such a system and ensuring that the servers that host such digital documents are always available. Each DID and document lookup results in a communications call to a server, and care has to be taken during design time to keep DID document resolutions/lookups to a minimum.

REFERENCES

ITU (2019). Distributed ledger technology use cases. Technical Report FG DLT D2.1, International Telecommunication Union (ITU).

Chakroun, B. and Keevy, J. (2018). Digital credentialing: Implications for the recognition of learning across borders.

Consortium, I. G. L. (2018a). Open badges baking specification: Ims final release. Accessed: 2018-04-12.

Consortium, I. G. L. (2018b). Open badges v2.0: Ims final release. Accessed: 2018-04-12.

DLA Piper (2017). Data protection laws of the world, full handbook.

Faridi, O. (2020). Establishing proper standards is key to ensuring blockchain or dlt sector moves forward: Report.

Gräther, W., Kolvenbach, S., Ruland, R., Schütte, J., Torres, C., and Wendland, F. (2018). Blockchain for education: Lifelong learning passport. *European Society for Socially Embedded Technologies (EUSSET)*, 2(10).

Gray, P. (2008). A brief history of education: To understand schools, we must view them in historical perspective.

Gresch, J., Rodrigues, B., Scheid, E., Kanhere, S. S., and Stiller, B. (2019). The proposal of a blockchain-based architecture for transparent certificate handling. In *Business Information Systems Workshops*, pages 185–196. Springer International Publishing.

Hameed, B., Murad, M., Noman, A., Javed, M., Ramzan, M., Ashfaq, F., Usman, H., and Yousaf, M. (2019). A review of blockchain based educational projects. *International Journal of Advanced Computer Science and Applications*, 10(10).

Hevner, A. and Chatterjee, S. (2010). *Design Research in Information Systems*. Springer US.

Hwang, M.-S. and Chang, T.-Y. (2005). Threshold signatures: Current status and key issues. *International Journal of Network Security*, 1.

Jaftha, C. and Samuels, J. (2017). World Reference Levels: A global learning outcomes initiative to promote recognition of learning. Accessed: 2017-09-28.

Labs, M. M. (2019). Blockcerts:the Open Standard for Blockchain Credentials, Standards Alignment. Accessed: 2019-11-19.

Otto, N., Duffy, K. H., Singh, K., and Aoqui, L. G. F. (2018). Open badges are verifible credentials. Accessed: 2018-03-06.

Otto, N., Lee, S., Sletten, B., Burnett, D., Sporny, M., and Ebert, K. (2019). Verifiable credentials use cases: W3C working group note 24 september 2019. Accessed: 2019-09-24.

Patka, I. (2019). How share kit works: A guide to secure data sharing. Accessed: 2019-06-14.

Raul, A. C., Faircloth, F. E., Mohan, V. K., and Brown, S. (2017). The privacy, data protection and cybersecurity law review. Publication, Law Business Research Ltd.

Reed, D., Sporny, M., Longley, D., Allen, C., Grant, R., Sabadello, M., and Holt, J. (2020). Decentralized identifiers (DIDs) v1.0: Core architecture, data model, and representations. Accessed: 2020-05-29.

Ronning, A. and Chung, W. W. (2019). Blockcerts V3 Proposal: A white paper from Rebooting the Web of Trust IX. Accessed: 2019-11-19.

Sonnino, A., Al-Bassam, M., Bano, S., and Danezis, G. (2018). Coconut: Threshold issuance selective disclosure credentials with applications to distributed ledgers. *Computing Research Repository (CoRR)*, abs/1802.07344.

South African Qualifications Authority (SAQA) (2018). Annual performance plan for 2019/20. Accessed: 2019-10-25.

Sporny, M., Longley, D., and Chadwick, D. (2019). Verifiable credentials data model 1.0: Expressing verifiable information on the web. Accessed: 2019-11-09.

The Mozilla Foundation, Peer 2 Peer University, and The MacArthur Foundation (2012). Open badges for lifelong learning.

Yumna, H., Khan, M. M., Ikram, M., and Ilyas, S. (2019). Use of blockchain in education: A systematic literature review. In *Intelligent Information and Database Systems*, pages 191–202. Springer International Publishing.