Digital Transformation of Education Credential Processes and Life Cycles – A Structured Overview on Main Challenges and Research Questions

Ingo R. Keck
Scientific Data Management
TIB Leibniz Information Centre
for Science and Technology
Hannover, Germany
Email: Ingo.Keck@tib.eu

Maria-Esther Vidal
Scientific Data Management
TIB Leibniz Information Centre
for Science and Technology
Hannover, Germany
Email: Maria.Vidal@tib.eu

Lambert Heller
Open Science Lab
TIB Leibniz Information Centre
for Science and Technology
Hannover, Germany
Email: Lambert.Heller@tib.eu

Abstract—**In this article, we look at the challenges that arise in the use and management of education credentials, and from the switch from analogue, paper-based education credentials to digital education credentials. We propose a general methodology to capture qualitative descriptions and measurable quantitative results that allow to estimate the effectiveness of a digital credential management system in solving these challenges. This methodology is applied to the EU H2020 project QualiChain use case, where five pilots have been selected to study a broad field of digital credential workflows and credential management.**

Keywords*–*Credentials; Education credentials; Digitisation; Challenges in digitisation.

I. Introduction

Education credentials are an important part of our modern life. Pupils exit schools with a set of marks certified on their final school report, then, based on these results, they are able to apply for acceptance at higher education institutes or for apprenticeship. Students and employees continue to collect credentials at university, at work or via other ways of education. Even today, when digitisation has entered into almost every part of our lives, these education credentials often still are printed and written on paper. These paper-based credentials present several problems in practice. For example, managing of these credentials applying for a job position is tiresome for the applicant and even more so for the company that offers the position. Indeed, most companies nowadays require scans of the paper credentials and will only check the validity of the originals once the candidate for the position has been selected, to avoid the manual labour involved. Additionally, surveys show that lying about education and employment credentials is a common problem. According to a survey by CareerBuilder [1], 58% of employers have caught a lie on a resume. Similar findings arise from another recent survey by StatisticBrain [2], which reports that over half of resumes and job applications (53%) contain falsifications and over three quarters (78%) are misleading. Digitisation of education credentials has the potential to make credential handling both easier and more secure. Nevertheless, it is important to ask the correct questions to be able to investigate how well a solution performs in the implementation and management of digital education credentials.

The main contribution in this work in progress article is to present the main challenges encountered in education credential management and usage, and in the changes from analogue to digital credential workflows. We propose specific questions that will allow an qualitative and quantitative assessment of the performance of a credential management system and infrastructure in regard of these challenges (given in Table I). Finally, we introduce the use case of the EU Horizon 2020 project QualiChain [3], where these research questions will be evaluated with the help of the participants in the project's pilots.

The article is organised as follows: In Section II, we elaborate the different challenges we encountered while analysing the reports and questionnaires provided by the QualiChain pilots. In Section III, we propose a set of questions for every challenge presented in the previous section. In Section IV, we present the use case of QualiChain. The article closes with Section V where our conclusions and future work are outlined.

II. Challenges in Education Credential Management

How can the performance of a solution offering the issuing, management and verification of digital education credentials be evaluated? Based on the results acquired in [4], we propose to segment the questions of interest into three subtopics, that follow the process of changing from an analogue to a digital setting:

A. *Challenges of paper-based credentials*;

B. *Challenges of transition to digital credentials*; and

C. *Challenges of digital credentials*.

In the following sections, we present these experienced difficulties and propose ways how to measure the performance of a presented solution for the implementation and management of digital education credentials.

A. Challenges of Paper-Based Credentials

Paper-based credentials are the state of the art and have a history dating back to medieval times. Their use over centuries makes it obvious that, before digitisation, they were widely seen as the best solution. However, the developments in the last decades and the move to digital workflows increased the pressure on analogue, paper-based credentials and lead to increasing problems, especially in the field of fraud prevention.

1) Fraud and Verification: Advances in digital printing make it continuously more difficult to protect paper-based credentials against fraud. As already mentioned, a survey by CareerBuilder [1] reports that 58% of employers have caught a lie on a resume and 33% of them have seen an increase in resume embellishments and fabrications like embellished skill sets (57%), embellished responsibilities (55%), dates of employment (42%), job titles (34%), academic degrees (33%), companies worked for (26%) and awards (18%). A different survey [2] states that over half of resumes and job applications (53%) contain falsifications and over three quarters (78%) are misleading. Most issuers do not have the capabilities to use advanced falsification protection in their paper credentials, compared to what is done, for example, for paper-based money. Without a general standard, it would also be impossible for a non-expert to decide if the credential in front of him/her has the correct characteristics, as there are over 3000 higher education establishments in the European Union alone [5]. Instead, institutions and states commonly register important credentials and allow interested individuals to inquire on the validity of a presented credential. The UK, for example, offers the Higher Education Datacheck service [6]. The use of this service is chargeable, and the process can take up to seven days [7]. The process is also highly manual and time consuming.

2) Dependence on Issuer: The problems with fraud make it difficult for other than official education establishments to issue education credentials. This leads to the problem that learners will be unable to furnish sufficient and incontestable proof over several types of qualifications gained outside this established system. In the job market, written recommendation statements (also easy to falsify) or contact persons of reference are used to compensate for this. These methods are also manual and time costing for the people involved. The challenge to correctly identify the issuer of such as statements is related to this problem. Additionally, this can be the reason why direct access to reference persons often is preferred, as in this case the authenticity of the reference person can be checked by other means, like contact over official phone numbers or email addresses.

3) Handling: Paper-based credentials are easy to handle and store for the bearer, but in situations where many credentials have to be collected, screened and analysed, the high manual handling costs make their use expensive. This leads to a time consuming and costly recruitment process. For staffing private and especially public sector organisations it can be challenging to efficiently handle competency management in large organisational structures, as was reported in our questionnaire collection at the QualiChain pilots.

4) Data Security: Using high-quality acid free paper and storage in low humidity and at room temperature in pest free environments, paper has successfully been archived over many decades. Additionally, data protection can be enforced by physical access restrictions that are commonly available. However, most users of paper-based credentials outside of official archives and libraries lack the means of long-term storage, which makes paper-based credentials vulnerable to loss and damage. This is made more severe by the impossibility to create identical copies of paper-based credentials.

B. Challenges of Transition to Digital Credentials

Any solution that asks users to move from a well-established analogue paper-based workflow to a digital workflow, will face challenges in this transition. In the following points we present the issues we encountered in our data collection.

1) Digitisation of Existing Credentials: Analogue credentials are put into existence using written text, images, drawings and security characteristics in various forms. To retain all this information in digital form is difficult, and to efficiently work with the content of the credential, it is necessary to convert the unstructured text, for example gained by a scan of the document, into structured data, that has been semantically enriched.

2) Interaction Between Analogue and Digital Workflows: While workflows for both digital and analogue paper-based credentials exist, it is desirable to cater for both types, if technically feasible and sensible. Often this will mean making manual adjustments possible in a digital workflow or to temporarily create digital twins of paper-based credentials to incorporate them into pure digital workflows. This can also mean that digital credentials are printed out, to be included in paper-based credential workflows.

C. Challenges of Digital Credentials

Digital representations of credentials have their own challenges, that may be quite different from the paper-based ones.

1) Private Data Protection: Digital data can easily be copied, and creating identical copies of digital data is part of the normal workflow in IT. If, for example, a digital credential is sent from the issuer over a secure channel to the credential holder, its actual data is copied multiple times in the process: The credential is copied from the data storage at the issuer to the network stack of the issuers system, then copied into a transport format, copied over various relays in the communication system till it is copied once more into the network stack of the receiver, unpacked and finally copied into the receiving application's memory. However, this characteristic of digital data makes it also easy to leak private data in the process. Where in paper-based credentials simple physical access control often is enough, for digital credentials, access control has also to be secured digitally.

2) Data Security: Digital data is stored in physical storage and this storage will degenerate over time. It is, therefore, important to be able to copy the digital credential to new physical storage and to continuously monitor the quality of the storage before the degradation leads to damaged data. In libraries the "lots of copies keep stuff safe" (LOCKS) model has been successfully implemented for electronic publications, based on the idea that independent copies of the same data in physical and geographical independent data stores ensure high data security and availability [8].

3) Data Management: Unlike their paper-based siblings, digital credentials can only be perceived by the user if their content or metadata is rendered in a perceivable form (usually visual). Management systems need to ensure that users know what is stored and what is transmitted if requested.

TABLE I. PROPOSED RESEARCH QUESTIONS TO EVALUATE THE PERFORMANCE OF A DIGITAL EDUCATION CREDENTIAL MANAGEMENT SYSTEM IN SOLVING THE CHALLENGES EXPERIENCED BY THE USER.

Challenge	Question	Units
Fraud protection and verification	How is the system protected against fraud?	qualitative
	What are the costs of a successful attack against the fraud protection?	time, money
Issuer dependence	What are the requirements for an issuer of digital credentials?	qualitative
	How much does issuing a credential cost?	time, money
Handling	Describe the workflow of a credential in the system.	qualitative
	How much does handling of a credential in the workflow cost?	time, money
Data security	How is the credential stored in the system?	quantitative
	Is the credential data format public and open?	yes/no
	How many independent copies of the credential are stored in the system at any time?	number
	How is the credential secured against accidental loss or data change?	quantitative
	How is the credential secured against unauthorised, but intentional, loss or change of data?	quantitative
Digitisation of existing credentials	How can existing analogue credentials be included into the digital workflow?	quantitative
	Is the content of the analogue credential converted to structured data to the same level of detail as digital credentials?	yes/no
Interaction between analogue and digital workflows	How can the system interact at the same time with digital and analogue credentials	quantitative
	How much increases the effort in the workflow, if digital and analogue credentials are mixed?	time, money
Private data protection	How is the private data stored in the system protected against unauthorised access?	quantitative
	What are the costs of a successful attack against the private data protection?	time, money
Data management	How is the data managed from the user perspective?	quantitative
	Can the user tell at any time of the workflow, what data exactly he/she is working with?	yes/no
	Can the user tell at any time of the workflow, who is able to access the data in question?	yes/no
Data sovereignty	How is data sovereignty enforced in the system?	quantitative
	Can the holder of the credential decide at any time of the workflow, who is able to access the data in question?	yes/no
	How much does it cost the user to store the data under his/her exclusive physical access?	time, money
	What are the costs of a successful attack against the access protection (access, denial of service, data change)?	time, money
	If there are other possibilities of storage, how convenient are they to the user?	time money
	What are the costs of a successful attack against these other storage possibilities (access, denial of service, data change)?	time, money

4) Data Sovereignty: The ease of copying of digital data allows for the storage of digital credentials physically far from the users, for example, on the cloud. However, this also means that the actual data then is outside the physical oversight of the user. The term ”data sovereignty” [9] has been coined in recent years to describe ”the idea that users, being citizens or companies, have control over their data” [10].

III. PROPOSED RESEARCH QUESTIONS

In this section, we collect the questions whose answers will be utilised to validate the effectiveness of a system devised to achieve the challenges presented in the previous Section II. Each presented topic translates into a set of questions. We start each topic with a question asking for a qualitative description of how the proposed solution approaches the relevant challenge and then, by adding quantitative questions that should enable us to measure the effect that the proposed solution has on each challenge in a given use case. Using this mixed qualitative and quantitative approach, it should be possible to compare a digital credential solution to the status quo of non-digital workflows.

In Table I, our research questions are presented; they are grouped according to the challenges presented in Section II. The challenge *data security* affects both digital and paper-based credentials in very similar ways, so we were able to combine all relevant questions into one field.

IV. USE CASE

The EU Horizon 2020 research and innovation action QualiChain ”targets the creation, piloting and evaluation of a decentralised platform for storing, sharing and verifying education and employment qualifications and focuses on the assessment of the potential of blockchain technology, algorithmic techniques and computational intelligence for disrupting the domain of public education, as well as its interfaces with private education, the labour market, public sector administrative procedures and the wider socio-economic developments.”[11] The fundamental idea of the project is to build an open source, distributed platform supporting the storage, sharing and verification of education credentials. This platform will allow for the implementation of additional services which will fulfil the needs of the participating actors, such as data analytics and decision support systems. QualiChain hosts five pilot projects distributed over Europe (for details please see [12]), where the system is tested in four real-world scenarios:

- Lifelong learning;
- Smart curriculum design;
- Staffing the public sector; and
- Providing HR consultancy and competency management services

We provided online questionnaires to support the participants in the pilots in the definition of the use cases, challenges and possible research questions, as well as to define key performance indicators. These questionnaires were filled in and discussed with the people involved in the pilots in early 2019. The process is discussed in detail in [4] and not repeated here for the sake of brevity.

V. Conclusion and Future Work

The intention of this article is to discuss the main challenges in education credential management and to present a methodology to both qualitative and quantitative measure a system's effectiveness in addressing them. Additionally, we aim at gathering feedback from the scientific community regarding these measurements and their adequacy. We apply this methodology to the use cases of the Horizon 2020 EU Project QualiChain, that cover a wide area of applications of education credentials. This will allow us an in deep evaluation of the project's performance. Based on the experience we will gather in this process, we plan to extend this work in the future to a full framework for the evaluation of the performance of education credential management solutions. This framework should be able to capture the whole life cycle of education credentials from creation and issuing over storage, management and access control, towards credential expiring or retraction.

Acknowledgement

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 822404 (QualiChain).

References

[1] *Fifty-eight percent of employers have caught a lie on a resume, according to a new careerbuilder survey*. [Online]. Available: https://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=8%2F7%2F2014&id=pr837&ed=12%2F31%2F2014 (visited on 2020-02-24).

[2] *Statisticbrain - resume falsification statistics*. [Online]. Available: https://www.statisticbrain.com/resume-falsification-statistics/ (visited on 2020-02-24).

[3] *QualiChain – decentralised qualifications' verification and management for learner empowerment, education reengineering and public sector transformation*. [Online]. Available: https://qualichain-project.eu (visited on 2020-02-24).

[4] I. Keck *et al.*, "D7.1 – qualichain pilots preparation handbook", Tech. Rep., 2019. [Online]. Available: https://alfresco.epu.ntua.gr/share/s/AQjQwquNRwOl-CtaHFDBLQ (visited on 2020-02-24).

[5] European Commission, *The role of universities in the europe of knowledge*, 2003. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:c11067 (visited on 2020-02-24).

[6] HECSU, *Prospects hedd verification + authentication*. [Online]. Available: https://hedd.ac.uk/ (visited on 2020-02-24).

[7] D. Matthews, "What blockchain technology could mean for universities", *Times Higher Education*, Aug. 2017. [Online]. Available: https://www.timeshighereducation.com/news/what-blockchain-technology-could-mean-for-universities (visited on 2020-02-24).

[8] V. A. Reich, "Lots of copies keep stuff safe as a cooperative archiving solution for e-journals", *Issues in Science and Technology Librarianship*, 2002. [Online]. Available: http://doi.org/10.5062/F47P8WCW (visited on 2020-02-06).

[9] R. Posch, "Digital sovereignty and it-security for a prosperous society", in *Informatics in the Future*, H. Werthner and F. van Harmelen, Eds., Cham: Springer International Publishing, 2017, pp. 77–86, ISBN: 978-3-319-55735-9.

[10] S. Amaro, *Europe's dream to claim its 'digital sovereignty' could be the next big challenge for us tech giants*, Nov. 2019. [Online]. Available: https://www.cnbc.com/2019/11/20/us-tech-could-face-new-hurdles-as-europe-considers-digital-sovereignty.html (visited on 2020-02-24).

[11] *QualiChain – decentralised qualifications' verification and management for learner empowerment, education reengineering and public sector transformation*, Nov. 2018. [Online]. Available: https://cordis.europa.eu/project/rcn/218758_en.html (visited on 2020-02-24).

[12] C. Agostinho *et al.*, "D2.2 – qualichain stakeholders' requirements and use cases", Tech. Rep., 2019. [Online]. Available: https://alfresco.epu.ntua.gr/share/s/EflUU9mbTESnrZo74WAZHg (visited on 2020-02-24).