

GI-Edition

Lecture Notes in Informatics

Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)

Open Identity Summit 2022

07.–08. July 2022
Copenhagen

Proceedings

325

H. Roßnagel, C. H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2022

GESELLSCHAFT
FÜR INFORMATIK



Heiko Roßnagel, Christian H. Schunck,
Sebastian Mödersheim (Eds.)

Open Identity Summit 2022

07. - 08.07.2022
Copenhagen, Denmark

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-325

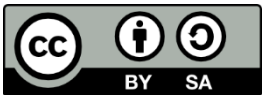
ISBN 978-3-88579-719-7
ISSN 1617-5468

Volume Editors
Heiko Roßnagel | Christian Schunck
Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering
Nobelstr. 12, D-70569 Stuttgart, Germany
heiko.rossnagel|christian.schunck@iao.fraunhofer.de
Sebastian Mödersheim
Sebastian Mödersheim
Technical University of Denmark Compute
Richard Petersens Plads, Building 324, Room 180
DK-2800 Kgs. Lyngby, Denmark
samo@dtu.dk

Series Editorial Board
Andreas Oberweis, KIT Karlsruhe,
(Chairman, andreas.oberweis@kit.edu)
Torsten Brinda, Universität Duisburg-Essen, Germany
Dieter Fellner, Technische Universität Darmstadt, Germany
Ulrich Flegel, Infineon, Germany
Ulrich Frank, Universität Duisburg-Essen, Germany
Michael Goedicke, Universität Duisburg-Essen, Germany
Ralf Hofestädt, Universität Bielefeld, Germany
Wolfgang Karl, KIT Karlsruhe, Germany
Michael Koch, Universität der Bundeswehr München, Germany
Peter Sanders, Karlsruher Institut für Technologie (KIT), Germany
Andreas Thor, HFT Leipzig, Germany
Ingo Timm, Universität Trier, Germany
Karin Vosseberg, Hochschule Bremerhaven, Germany
Maria Wimmer, Universität Koblenz-Landau, Germany

Dissertations
Steffen Hölldobler, Technische Universität Dresden, Germany
Thematics
Agnes Koschmider, Universität Kiel, Germany
Seminars
Judith Michael, RWTH Aachen, Germany

© Gesellschaft für Informatik, Bonn 2022
printed by Köllen Druck+Verlag GmbH, Bonn



This book is licensed under a Creative Commons BY-SA 4.0 licence.

Preface

Welcome to the “Open Identity Summit 2022”, which has been jointly organized by the Special Interest Groups BIOSIG within the German Computer Science Society (Gesellschaft für Informatik) and the Technical University of Denmark.

The international program committee performed a strong review process according to the LNI guidelines with at least three reviews per paper and accepted 50 % of the 16 submitted papers as full scientific papers.

Furthermore, the program committee has created a program including selected contributions of strong interest (further conference contributions) for the outlined scope of this conference.

We would like to thank all authors for their contributions and the numerous reviewers for their work in the program committee.

Copenhagen, 10th of May, 2022

Heiko Roßnagel
Fraunhofer IAO

Christian H. Schunck
Fraunhofer IAO

Sebastian Mödersheim
DTU Compute

Conference Chairs

Heiko Roßnagel, Fraunhofer Institute for Industrial Engineering IAO
Christian H. Schunck, Fraunhofer Institute for Industrial Engineering IAO
Sebastian Mödersheim, Technical University of Denmark Compute

Programme Committee

Franco Arcieri, Italy
Tamas Bisztray, Norway
Arslan Broemme, Germany
Christoph Busch, Germany
Victor-Philipp Busch, Germany
Jos Dumortier, Belgium
Daniel Fett, Germany
Lothar Fritsch, Norway
Walter Fumy, Germany
Igor Furgel, Germany
Marit Hansen, Germany
Olaf Herden, Germany
Gerrit Hornung, Germany
Detlef Houdeau, Germany
Detlef Hühnlein, Germany
Tina Hühnlein, Germany
Aws Jaber, Norway
Luigi Lo Iacono, Germany
Ulrike Korte, Germany
Michael Kubach, Germany
Andreas Kühne, Germany
Sebastian Kurowski, Germany
Herbert Leitold, Austria
Milan Markovic, Serbia
Tarvi Martens, Estonia
Gisela Meister, Germany
Sebastian Mödersheim, Denmark
Alexander Nouak, Germany
Sebastian Pape, Germany
René Peinl, Germany
Henrich Pöhls, Germany
Kai Rannenberg, Germany
Alexander Roßnagel, Germany
Heiko Roßnagel, Germany
Christian H. Schunck, Germany
Rachelle Sellung, Germany
Jon Shamah, United Kingdom
Maurizio Talamo, Italy
Don Thibeau, United States
Karsten Treiber, Germany
Samuel Wairimu, Sweden
Tobias Wich, Germany
Thomas Wieland, Germany
Alex Wiesmaier, Germany
Jan Zibuschka, Germany
Jan Ziesing, Germany
Frank Zimmermann, Switzerland

Hosts and Partners

BIOSIG – Biometrics and Electronic Signatures (www.biosig.org)
The special interest group “Biometrics and Electronic Signatures” (BIOSIG) within GI e.V. is dedicated to the fundamentals, methods, techniques, processes and implementations used to guarantee the authenticity and integrity of entities.

mGov4EU – Mobile Cross-Border Government Services for Europe (https://www.mgov4.eu/)
mGov4EU pushes forward the practical use of inclusive mobile Government services in Europe, bringing such services in line with EU citizens’ expectations for safe, resilient and sustainable mobile communication. Innovating electronic identity management, storage of data and the exchange of electronic documents are key elements.

Table of Contents

Open Identity Summit 2022 – Regular Research Papers

Lothar Fritsch, Marie Mecaliff, Kathinka Wik Opdal, Mathias Rundgreen and Toril Sachse
Towards robustness of keyboard-entered authentication factors with thermal wiping against thermographic attacks ……..……15

Isaac Henderson Johnson Jeyakumar, David W Chadwick and Michael Kubach
A novel approach to establish trust in verifiable credential issuers in Self-Sovereign Identity ecosystems using TRAIN ……....27

Nikos Fotiou, Evgenia Faltaka, Vasilis Kalos, Anna Kefala, Iakovos Pittaras, Vasilios A. Siris and George C. Polyzos
Continuous authorization over HTTP using Verifiable Credentials and OAuth 2.0…….....39

Michael Kuperberg and Robin Klemens
Integration of Self-Sovereign Identity into Conventional Software using Established IAM Protocols: A Survey ……..….....51

Steffen Schwalm, Daria Albrecht and Ignacio Alamillo
eIDAS 2.0: Challenges, perspectives and proposals to avoid contradictions between eIDAS 2.0 and SSI…….....63

K. Valerie Carl, Timothy Markus Christian Zilcher and Oliver Hinz
Corporate Digital Responsibility and the current Corporate Social Responsibility standard: An analysis of applicability ……..……..75

Takashi Norimatsu, Yuichi Nakamura and Toshihiro Yamauchi
Flexible Method for Supporting OAuth 2.0 Based Security Profiles in Keycloak…….....87

Sebastian Kurowski and Christian H. Schunck
Risk variance: Towards a definition of varying outcomes of IT security risk assessment……99

Open Identity Summit 2022 – Further Conference Contributions

Nicolas Fähnrich, Matthias Winterstetter and Michael Kubach
A user-centric approach to IT-security risk analysis for an identity management solution ……113

Andrea Horch, Christian H. Schunck and Christopher Ruff
Adversary Tactics and Techniques specific to Cryptocurrency Scams….....119

Hermann Strack, Sebastian Karius, Marlies Gollnick, Meiko Lips, Sandro Wefel, and Robert Altschaffel
Preservation of (higher) Trustworthiness in IAM for distributed workflows and systems based on eIDAS……..…..125

Nicolas Fähnrich and Heiko Roßnagel
Online tool for matching company demands with IT-security offerings.....131

Paul Bastian, Carsten Stöcker and Steffen Schwalm
Combination of x509 and DID/VC for inheritance properties of trust in digital identities……..137

Open Identity Summit 2022

Regular Research Papers

Towards robustness of keyboard-entered authentication factors with thermal wiping against thermographic attacks

Lothar Fritsch[1], Marie Mecaliff[2], Kathinka Wik Opdal[3],
Mathias Rundgreen[4], Toril Sachse[5]

Abstract: Many authentication methods use keyboard entry for one of their authentication factors. Keyboards factors have been compromised exploiting physical fingerprints, substances from fingers visible on keys, with acoustic recordings through mobile phones, and through video reflections captured by high-resolution cameras used for video conferencing. Heat transfer from human fingers to keypads is an additional attack channel that has been demonstrated. There are few mitigation measures published against this type of attack. This article summarizes the feasibility of thermographic attacks against computer keyboards and against door pin pads, as well as the efficiency of the scrubbing technique deployed in order to counter thermographic attacks. For this purpose, a series of experiments with small, mobile thermal cameras were carried out. We report findings such as time intervals and other constraints for successful laboratory observation of authentication factors, describe scrubbing methods and report the performance of those methods.

Keywords: password hijacking, infrared camera, thermographic attack, thermal imaging, authentication factor, identity management, information security, scrubbing, thermal lock picking.

1 Introduction

PIN codes and passwords as well as touchscreen-entered patterns are widely used authentication factors. Their compromise can lead to the collapse of individual digital identities as well as to the degradation of whole identity ecosystems [Fr20]. The common feature of most solutions is the transfer of an authentication secret to a computer input device through physically pressing or moving fingers of the human hand over the device. This physical contact transfers body heat from the finger to the touched device. Such heat is measurable with thermography cameras which measure the infrared head emissions from object surfaces. Sensors are recently integrated in small devices such as the FLIR C5 pocket camera[6] and the hardened Android phone CAT S62 Pro[7],

[1] Oslo Metropolitan University (OsloMET), Department of Computer Science, Postboks 4, St.Olavs plass, 0130 Oslo, Norway, lotharfr@oslomet.no, https://orcid.org/0000-0002-0418-4121

[2] Institut National des Sciences Appliquées de Toulouse, France, mariemecaliff@gmail.com
[3] Oslo Metropolitan University (OsloMET), Department of Computer Science, s187533@oslomet.no
[4] Oslo Metropolitan University (OsloMET), Department of Computer Science, rundgreen@me.com
[5] Oslo Metropolitan University (OsloMET), Department of Computer Science, s341837@oslomet.no
[6] FLIR C5 thermal camera, https://www.flir.com/products/c5/ as of 18.1.2022
[7] CAT S62 Pro mobile phone with thermal camera, https://www.catphones.com/en-gb/cat-s62-pro-smartphone/ , as of 18.1.2022

which are available at prices below 1000 EUR. Both the price range as well as the deployability of the cameras outside laboratory settings increase the feasibility and likelihood of thermal attacks. We therefore investigated the feasibility of attacks and their prevention.

The remainder of the article is structured as follows: First, we summarize the background of the project by a summary of thermal attacks and their mitigation in academic literature. Next, we define our research questions and describe the experimental setup and the results of experimentation. Finally, we discuss our results and summarize open issues.

1.1 Background thermal hacking

Attacks against PIN pads: Point-of-sales attacks against PIN code security have been investigated by Singh. et al [SBS19] with the goal to investigate the influence of camera distance, time passed between entry and capture, angle of photography and ambient temperature. Li et al [Li19] built a demonstrator that extracts the sequence of typed keys from a numerical keypad in laboratory experiments. They found influence factors such as typing speed, ambient temperature and typing speed as well as the number of repetitive keys used. In a second publication, Li et al [Li18] present experimental results for three countermeasures that reduce attack success from 30% to 10% based on these observations (see Tab. 1). Mowery et al [MMS11] demonstrate automated extraction of keypad key patterns from 10.000 to 24 possible 4-digit PIN codes. They note influence factors such as keypad heat absorption, material reflectivity, ambient temperature and lightness of finger pressure during typing. However, they do not recommend countermeasures.

Touch screens and pads: Abdrabou et al [Ab20] experimented with the thermal capture of security patterns and gestures on touch screens and touch pads of mobile computers and device. They achieved experimental success rates ranging from 14.81% (touch pad taps) up to 60% (gestures). They note that tap patterns transfer less heat than gestures painted with sliding fingers. Temperature differences between different test participants' body temperatures were complicating analysis. No countermeasures were suggested. Complementary experimentation with mobile phone touch screens performed by Abdelrahman et al. [Ab17] investigated the automated extraction of touch PIN codes and authentication patterns from video sequences. They achieved success rates between 80% and 100 in lab settings, however noted that automated analysis suffers from overlapping lines in authentication patterns. The article proposes several countermeasures, summarized in Tab. 1.

Computer keyboards: Kaczmarek et al [KOT19] studied password entry on computer keyboards. They notice differences in heat traces left by different typing styles. Notably, typing with fingers sliding over keys that are not pressed as part of the passwords complicates password extraction. Strong reduction in password search complexity is found. The authors speculate about countermeasures, however, do not experiment with

them (see Tab. 1). Wodo et al [WH16] investigate a wide range of key pads and keyboards in an exploratory study. They note that materials, surfaces, ambient and keyboard temperatures as well as timing constraints influence thermal print visibility. They reference defense countermeasures, which are classified below.

Known countermeasures. Below, countermeasures against thermal attacks mentioned in the surveyed literature are listed and classified into types and maturity.

Countermeasure	Reference	Maturity
Block line of sight with shield	[SBS19], [AKSA17], [WH16]	Proposal
Reflective surface	[SBS19]	Proposal
Curved shape of keypad for diffusion	[SBS19]	Proposal
Delay entry transaction until cooled off	[SBS19]	Proposal
Minimum distance camera to keypad	[SBS19]	Proposal
Temperature control of keypad	[SBS19], [WH16]	Proposal
Materials with low heat conductivity	[SBS19], [KOT19]	Proposal
Blinding with illuminated keypad	[SBS19], [AKSA17]	Proposal
Wiping with CPU-generated heat	[AKSA17]	Proposal
Heating of surface	[AKSA17]	Proposal
Touch-to-heat wiping of thermal print	[SBS19], [Li19], [KOT19], [AKSA17], **[Li18]**	Proposal, **Experiment**
Blowing of warm air over keypad	[Li19], **[Li18]**	Proposal, **Experiment**
Randomized virtual pad on touchscreen	[SBS19], [AKSA17]	Proposal
Ambient light against key mapping	[Li19]	Proposal
Passwords with repetitive keys	[Li19], [AKSA17], **[Li18]**	Proposal, **Experiment**
Very long passphrases	[AKSA17]	Proposal
Add authentication factor or medium	[AKSA17], [WH16]	Proposal
Use temperate finger substitute for entry	[KOT19]	Proposal

Tab. 1: Countermeasures against thermal attacks

Our background summary in Tab. 1 clearly shows the lack of empirical validation of thermal attack countermeasures in published literature. Only for three of the measures, experiments of their effect have been published.

Targets of attacks: All attacks found in literature were deployed in laboratory settings. Most effort is put into attacking PIN pads, followed by touch-based patterns and then computer keyboards. Below, the mapping of literature on attack target is listed.

PIN pads: [SBS19], [Li19], [Li18], [MMS11], [WH16]

Touch screens: [Ab20], [AKSA17]

Touch pads (laptop): [Ab20]

Keyboards: [KOT19], [WH16]

Digital door lock: [WH16]

The specification of countermeasures in the above publication falls short of detailed descriptions of how the countermeasures must be applied in order to succeed. Neither material specifications, light intensities, temperature intervals or descriptions of wiping movements are described in a level of detail that would allow for the reproduction of the experiments.

1.2 Research questions

In this article, we summarize the findings of experimentation that targeted two research questions:

1. Are IR fingerprints exploitable in a practical attack scenario with small thermal cameras? This research question investigates the practicability of attacks against a door PIN pad system.
2. How can IR attacks get mitigated with simple means for everyday application? Which methods that do not need technical modifications, and which work with minimal effort, can mitigate thermal attacks?

Digital door locks with PIN pads combine three experimental advantages: They are available indoors in controlled temperatures and lighting conditions. They do not move, but rest in locked position. And, most important, their users pass the door after PIN entry while leaving the PIN pad with thermal fingerprints behind. Thus, our feasibility study examines door locks with PIN pads as attack targets [RS21].

The empirical foundation of the proposed countermeasures against thermal attacks is weak. Therefore aims our second research question at systematic trials of easily applicable mitigation techniques in order to qualify how they work, and under which circumstances they are effective [Me21],[Wi21].

2 Experimentation and results

We carried out three studies with experiments in order to investigate the research questions: a feasibility test of thermal attacks against digital door locks with PIN pad [RS21], and experiments deploying countermeasures against thermal attacks against computer keyboard password entry [Wi21],[Me21]. In this section, we describe the experiments.

2.1 Practicability of attacks against digital door locks with PIN pads

First observations made by targeting operative door locks at OsloMET's buildings. Fig. 1 shows various types of indoor and outdoor PIN pads in thermal images. Note how shape, surface materials and heat emissions from internal electronics influence the visibility of the heat prints on key "5" in the middle of the key pads.

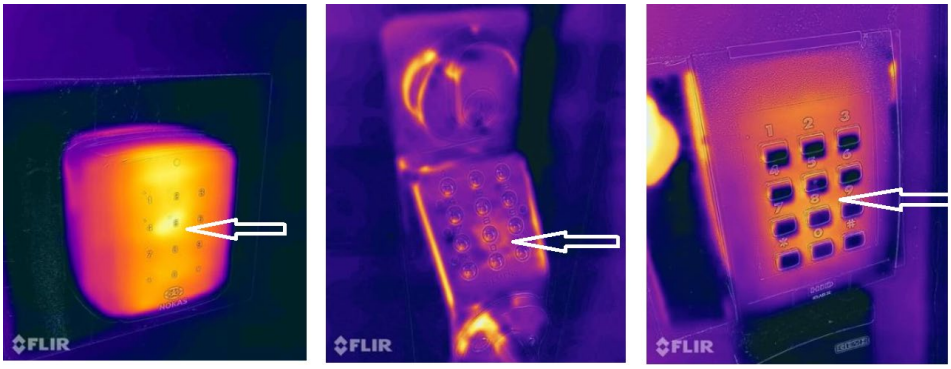


Fig. 1: Thermal image of various locks at OsloMET after pressing "5". Left: Internal heat obfuscates print. Center: round metal keys diffuse heat radiation on keys. Right: internal heat and plastic caps hide print [RS21].

As a consequence of these observations, a door lock was borrowed from a locksmithing shop in Oslo. It was set up in a lab where lighting conditions could be controlled. The lock's temperature was measured at room temperature and refrigerated in order to simulate outdoor measurement. Fig. 2 shows the lab setup.

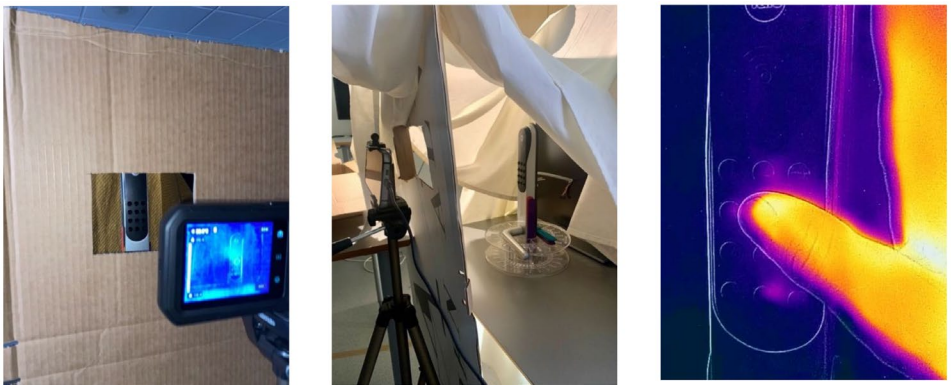


Fig. 2: Laboratory setup. Left: Camera positioning with mask. Center: shielding cloth against ambient light. Right: Thermal recording of PIN entry [RS21].

Experimentation was carried out by recording videos of PIN entries of a 4-digit and a 6-digit PIN for each temperature. Experiments were repeated for 10 rounds. Recordings

were done for 40 seconds after entry. Visibility of the heat prints was generally degraded after this time period. Visibility of PIN keys was measured by analyzing amplified color contrast values in the video still frames with the help of video editing software (5KPlayer for MacOS, Digiarty Software). Results show that the refrigerated lock shows heat prints longer than the lock at room temperature. Visibility of the PIN keys ranged from 3 seconds to 15 seconds. The average independent of temperature was 6.91 seconds (see Fig. 3). The longest visible print was detectable for 30,4 seconds.

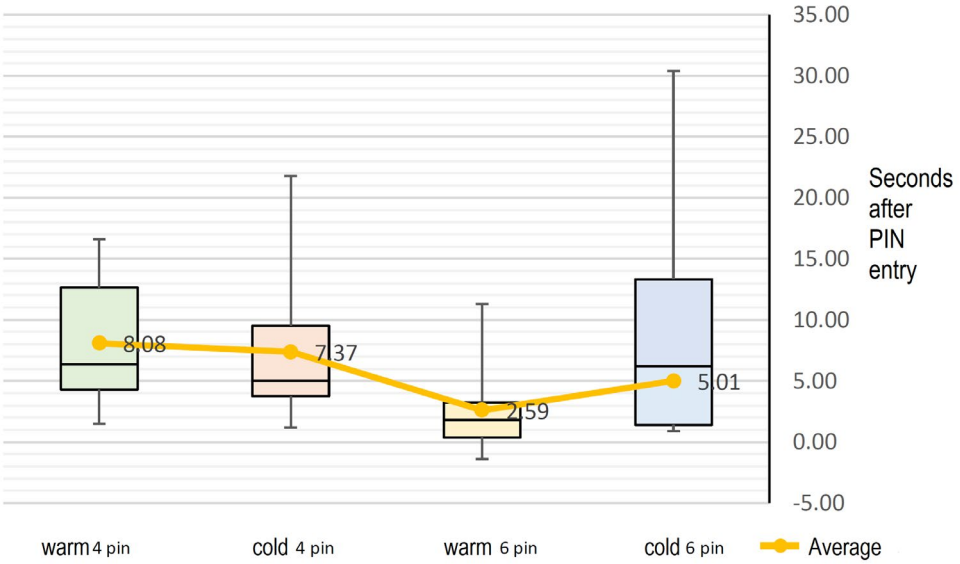


Fig. 3: Visibility of 4- and 6-digit PIN on lock at room temperature and refrigerated. Results obtained from contrast-enhanced video recordings. Average visibility: 6.91 seconds [RS21].

In summary, thermal lock-picking attacks on door lock PIN codes are feasible, however in specific contexts. The type of lock its surface materials, and ambient temperatures as well as ambient light will influence success. In warm environments, the time window between closing the door and the fading of the heat prints will be short. Some locks generate internal heat, which outshines the thermal prints.

2.2 Examination of wiping techniques on computer keyboards

The second research question investigated the applicability and the effectiveness of mitigation methods against thermal attacks. Reviewing the methods from literature (see Tab. 1), four methods were chosen for their ease of use, and their availability without modifications to the target keyboard:

a) Application of flat hand to keyboard: Heat transfer from pressing a flat hand to the keyboard after password entry was used to camouflage pressed keys in a larger heart print.

b) Hot gel pack: Applying a medical gel pack warmed up to body temperature, heat was transferred to the keyboard in order to camouflage keys pressed.

c) Cold gel pack: Cooling the thermal fingerprints using a chilled medical gel pack applied to a keyboard with the intention to remove the thermal prints.

d) Scrubbing: Moving hand randomly over the keyboard (once and repeatedly) with the intention to camouflage the pressed keys in additional thermal prints.

e) Blowing: Applying warm air blown over the keyboard in order to conceal the thermal prints from password entry.

The experimental setup was built in an air-conditioned lab room at OsloMET. A PC keyboard was placed on a table. A camera tripod with the thermal camera was mounted next to the keyboard and calibrated. A refrigerator as well as hot and cold water and a water cooker/coffee maker were available for heating and cooling gel packs. Experiments were run several times, assessing the effect of the wiping method as well as the timing constraints of the visible artifacts. The lab setup is shown in Fig. 4.



Fig. 4: Experimental setup for password wiping studies [Me21].

In a first round of pre-experiments, camera calibration was done. Through small series of tests, cool-off time intervals for the keyboard were found. Issues arose when the team found out that different persons emit different amounts of heat. Two experimenters had relatively low surface temperatures of their fingertips, such that they after experimentation decided to standardize their hand temperatures with the help of a bottle with warm water heated to a controlled target temperature. For experimentation, two passwords of 8 and 16 characters length were chosen: ILOVEYOU and SMITTEVERNTILTAK. The latter password contained double and tripe use of letters, which was found to have an impact on detectability of pressed keys. Experimentation was done in 10 rounds of measurements with timed typing, time measurements and interval photography using the thermal cameras.

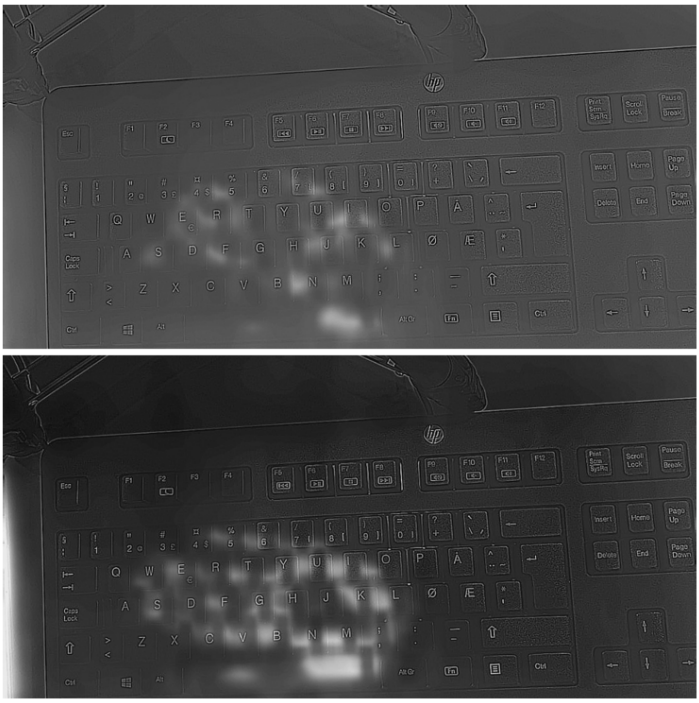


Fig. 5: Hand pressing technique. Top: 2.5s, bottom: 5s application [Wi21].

Not surprisingly, application time was a major determinant of wiping heat transfer. Fig. 5 shows the difference in transferred heat from 2.5 and 5 seconds of pressing the whole hand against a keyboard after typing a password. Individual keys were still identifiable after 2.5 seconds. In the experiments with moving hands or blowing air, speed or exposure time was equally relevant. As shown in Fig. 6 b), application of warm air from own longs has a considerably higher wiping effect when blown for 5 seconds with higher pressure. The application techniques and their effect are summarized in Tab 2 below.

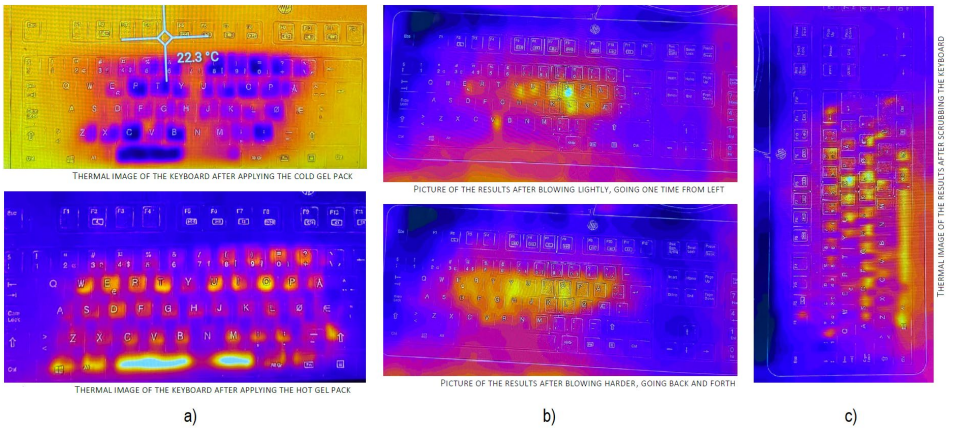


Fig. 6: Wiping: a) cold and hot gel pack; b) blowing warm air; c) wiping with hand [Me21].

In the course of experimentation it turned out that the wiping methods had to get applied much longer than the original password typing in order to securely transfer enough heat. Applications under 2.5 seconds revealed pressed keys, while application times of 5 seconds or more had a sufficient mitigation effect. Results will be discussed more in the next section of the article.

2.3 Summary of results

Results from experimentation show that any of methods a) to e) can mitigate the thermal attack. An application time of approx. 5 seconds will sufficiently wipe the thermal print. The different wiping techniques show different effects, and they require different preparations. Results are summarized in Tab 2.

Method	Application time	% keys identified	Preparation overhead
Cold gel pack	5s	31	Gel pack chilled to 8 C.
Hot gel pack	5s	31	Gel pack pre-heated to. 42 C
Moving hand slow	2s	44	Warming up hand if needed
Blowing	5s	59	Inhale and blow on keyboard
Pressing hand	5s	0	Warming up hand if needed
Moving hand once	5s	0	Warming up hand if needed
Moving hand once	2.5s	15	Warming up hand if needed
Moving hand multiple	5s	0	Warming up hand if needed
Moving hand multiple	10s	0	Warming up hand if needed
Control experiment (no wiping)		93	-

Tab 2: Effect and timing of wiping methods.

Various issues require care when wiping methods are applied. Moving hands is variable in speed and pressure applied. Self-discipline for slow movements is required. The same holds for the blowing method. A 5-second blow from the lungs is a practice many use once per year on the occasion of blowing the candles on birthday cakes, not when opening doors on a daily basis. Pressing the flat hand is a well-controllable movement, however it may suffer from the hand not being warm, due to either physiological conditions or recent exposure to cold environments. Finally, warm gel packs help standardizing temperature and application surface, are however items that need extra attention and preparation.

3 Discussion & conclusion

We showed that thermal attacks are a practical vulnerability that can get exploited in the area of PIN-protected door locks. This finding complements earlier research about practical point-of-sales-terminals as attack targets for payment card PIN theft. Further theoretical attacks target keyboard passwords and screen-lock patterns on mobile devices. However, the staging of attacks in these scenarios will require the attacker to have an opportunity to photograph devices shortly after passcode entry. We found that thermal prints are visible up to nearly one minute in thermal cameras.

Simple and practical countermeasures are available for individuals. They complement physical protection measures such as shielding, distance, heating and reflective surfaces. Hand pressing, hand movements, blowing hot air or the application of heat packs can effectively render thermal prints invisible, and thereby mitigate attacks. User-deployed wiping techniques should be part of security briefings and of user policies wherever keyboard-entered authentication factors are used. We suggest the inclusion of thermal wiping techniques into security awareness materials. One promising way of promotion will be the production of themed awareness gifts in the form of hand-warming heat packs with a thermal attack reminder as part of security awareness work.

Future research. Future research will investigate the social context for successful thermal attacks. Experimentation with secured campus access areas and their users will reveal the rates of succeeding with thermal photography versus being spotted and compromised in practical settings Further attention should be used on authentication factors for digital identities, since the potential damage of compromise is large. PIN-pad enabled tokens can get compromised in similar ways as keyboards.

Acknowledgement. We thank OsloMET's research group for Universal Design, especially Professor Weiqin Chen, for providing laboratory space for experimentation.

Bibliography

[Ab20] ABDRABOU, YASMEEN ; ABDELRAHMAN, YOMNA ; AYMAN, AHMED ; ELMOUGY, AMR ; KHAMIS, MOHAMED: Are Thermal Attacks Ubiquitous? When Non-Expert Attackers Use Off the shelf Thermal Cameras. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. New York, NY, USA : Association for Computing Machinery, 2020 — ISBN 978-1-4503-7535-1, S. 1–5

[Ab17] ABDELRAHMAN, YOMNA ; KHAMIS, MOHAMED ; SCHNEEGASS, STEFAN ; ALT, FLORIAN: Stay Cool! Understanding Thermal Attacks on Mobile-based User Authentication. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, *CHI '17*. New York, NY, USA : Association for Computing Machinery, 2017 — ISBN 978-1-4503-4655-9, S. 3751–3763

[Fr20] FRITSCH, LOTHAR: Identification collapse - contingency in Identity Management. In: *Open Identity Summit 2020*. Bd. P305. Bonn : Gesellschaft für Informatik e.V., 2020. — Accepted: 2020-05-27T12:09:21Z — ISBN 978-3-88579-699-2

[KOT19] KACZMAREK, TYLER ; OZTURK, ERCAN ; TSUDIK, GENE: Thermanator: Thermal Residue-Based Post Factum Attacks on Keyboard Data Entry. In: *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, *Asia CCS '19*. New York, NY, USA : Association for Computing Machinery, 2019 — ISBN 978-1-4503-6752-3, S. 586–593

[Li19] LI, DUO ; ZHANG, XIAO-PING ; HU, MENGHAN ; ZHAI, GUANGTAO ; YANG, XIAOKANG: Physical Password Breaking via Thermal Sequence Analysis. In: *IEEE Transactions on Information Forensics and Security* Bd. 14 (2019), Nr. 5, S. 1142–1154

[Li18] LI, DUO ; ZHANG, XIAO-PING ; ZHAI, GUANGTAO ; YANG, XIAOKANG ; ZHU, WENHAN ; GU, XIAO: Modeling Thermal Sequence Signal Decreasing for Dual Modal Password Breaking. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, S. 1703–1707

[Me21] MECALIFF, MARIE: *How to secure passwords against infrared camera attacks*, *Project report DATA3710* (Student report). Oslo : Department of Computer Science, Oslo Metropolitan University, 2021

[MMS11] MOWERY, KEATON ; MEIKLEJOHN, SARAH ; SAVAGE, STEFAN: Heat of the moment: characterizing the efficacy of thermal camera-based attacks. In: *Proceedings of the 5th USENIX conference on Offensive technologies*, *WOOT'11*. USA : USENIX Association, 2011, S. 6

[RS21] RUNDGREEN, MATHIAS ; SACHSE, TORIL: *Thermohacking*, *Project report DATA3710* (Student report). Oslo : Department of Computer Science, Oslo Metropolitan University, 2021

[SBS19] SINGH, GURVINDER ; BUTAKOV, SERGEY ; SWAR, BOBBY: Thermal Print Scanning Attacks in Theretail Environments. In: *2019 International Siberian Conference on Control and Communications (SIBCON)*, 2019, S. 1–6

[Wi21] WIK OPDAL, KATHINKA: *Hvor sikre er PIN-koder mot angrep basert på termografi?*, *Project report DATA3710* (Student report). Oslo : Department of Computer Science, Oslo Metropolitan University, 2021

[WH16] WODO, WOJCIECH ; HANZLIK, LUCJAN: Thermal Imaging Attacks on Keypad Security Systems: In: *Proceedings of the 13th International Joint Conference on e-Business and Telecommunications*. Lisbon, Portugal : SCITEPRESS - Science and Technology Publications, 2016 — ISBN 978-989-758-196-0, S. 458–464

A novel approach to establish trust in verifiable credential issuers in Self-sovereign identity ecosystems using TRAIN

Isaac Henderson Johnson Jeyakumar[1], David W Chadwick[2] and Michael Kubach[3]

Abstract: Self-sovereign identity (SSI) promises to bring decentralized privacy friendly identity management (IdM) ecosystems to everyone. Yet, trust management in SSI remains challenging. In particular, it lacks a holistic approach that combines trust and governance frameworks. A practical and scalable mechanism is needed for verifiers to externally verify their trust in credential issuers. This paper illustrates how TRAIN (Trust mAnagement INfrastructure), an approach based on established components like ETSI trust lists and the Domain Name System (DNS), can be used as a trust registry component to provide a holistic approach for trust management in SSI ecosystems. TRAIN facilitates individual trust decisions through the discovery of trust lists in SSI ecosystems, along with published credential schemas, so that verifiers can perform informed trust decisions about issued credentials.

Keywords: Self-sovereign identity, SSI, digital identity, decentralized identity, identity management, trust registry, trusted issuers, trust lists, IdM.

1. Introduction

The concept of Self-sovereign identity (SSI) [Al16] is currently widely debated in the digital identity community, among practitioners, politicians, as well as academics. It promises to put the end user (citizen) in total control of revealing their identity. The end user's credentials are managed by themselves and directly presented to verifiers (service providers) without the involvement of third parties. The issuers of the credentials (i.e., identity providers) are not involved in the process of presentation. While SSI architectures often use blockchains or other distributed ledger technologies (DLT), this is not a necessity [F22].

Some pursue SSI hoping to achieve their vision of an independent citizen identity. Others

[1] University of Stuttgart, Institute of Human Factors and Technology Management IAT, Allmandring 35, Stuttgart, 70569, Isaac-Henderson.Johnson-Jeyakumar@iat.uni-stuttgart.de.
[2] Crossword Cybersecurity, Capital Tower, 91 Waterloo Rd, South Bank, London SE1 8RT, david.chadwick@crosswordcybersecurity.com
[3] Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO, Nobelstr. 12, 70569 Stuttgart, Germany, michael.kubach@iao.fraunhofer.de

see SSI in the context of the political project of data sovereignty that the European Union (EU) is pushing forward with the revision of eIDAS [EU21]. SSI has been called the next evolutionary step in the world of digital identity, the future of digital identity, and more [IN20], [Si18].

However, despite the high hopes that are connected to SSI, the technology still has to overcome certain significant challenges before it can become widely adopted in the market and live up to the expectations of its proponents [Ku20]. One fundamental hurdle that has already been widely discussed but has not yet been fully solved in practice is trust management in these decentralized identity ecosystems that SSI is creating [KR21]. This is where our paper contributes. In the following sections, we first elaborate on the trust and governance challenges that current SSI systems are facing. Subsequently, in section 3 we present related work and initiatives, as this topic has of course been studied by other researchers and developers. In section four we present the general TRAIN approach to trust management for SSI that we have developed over the last year. Section 5 then gives more details on how TRAIN enables verifier-centric trust decisions to be made in order to establish trust in issuers. In section 6 we indicate where future work is still needed, before we conclude the paper in section 7.

2. Trust and governance challenges in Self-sovereign identity

SSI claims to solve the trust issues in identity management systems by focusing on Decentralized identifiers (DIDs), distributed ledgers, crypto key rotations, and Zero-knowledge proofs. However, these architectural elements only address "technical" trust, which is only one part of the overall trust management in IdM system. On the other side is "institutional" trust which addresses and defines criteria for relying parties to be accepted as legitimate actors in the ecosystem. Unless institutional trust is ensured, every issuer can claim to be legitimate which can lead to different security attacks in the ecosystem. For example: a framework for institutional trust may define that only certain governmental institutions are entitled to issue Identity Documents (IDs) to its citizens. If everyone was eligible to issue citizen IDs then there might be a lot of fake IDs circulating.

Standardization and interoperability are arguably the most relevant governance challenges for SSI [St21]. Currently, an increasing number of initiatives are trying to implement SSI or SSI-like solutions and contribute to standardization efforts in order to achieve interoperability between the different ecosystems under development. SSI today is in a stage similar to the early days of the transmission and network protocols before the internet, i.e., TCP/IP protocols, were firmly established. Isolated SSI islands with proofs of concepts in different technologies prevail. It will be necessary to achieve a widespread adoption of standards in order to achieve global interoperability.

At this point, for example, the international W3C consortium has launched efforts for standards such as DIDs and Verifiable Credentials (VCs), aiming to standardize SSI data models [W321a], [W319]. But so far, protocols are out of their scope. These efforts are

being undertaken by the Open ID Foundation, which is enhancing the OpenID Connect (OIDC) protocol to support W3C VCs and DIDs [Op22]. In addition, the Trust over IP Foundation (ToIP) has committed itself to building a holistic architecture for digital trust on the Internet [Tr22]. In particular, its desire to support the ability to port VCs between different networks will be crucial for the widespread adoption of person-based SSI. Likewise, the number of providers will initially be decisive to reach a broad mass of users.

The European Union Agency for Cyber Security (ENISA) in its recent report regarding leveraging the self-sovereign identity concept to build trust [EN22] mentioned the requirement of Governance frameworks in the Certification of Wallets, Audit, and Oversight of DID Controllers, VC Issuers, DIDs, and VC registries.

3. Related work and initiatives

The requirement of a trust registry and trust anchor has been recognized as a challenge by researchers and institutions. The Trust Over IP Foundation has recognized the need to address the trustworthiness of the various parties involved in the SSI ecosystem. Hence, a trust registry specification working group [To22] has been established, which is addressing the challenges pertaining to trust and governance in the SSI ecosystem and aims to develop a trust registry framework specification.

Moreover, the GAIA-X Federation Services (GXFS) in the GAIA-X initiative, which aims to build their IdM on open SSI components, has acknowledged this challenge as well. The project, which is developing a federation of data infrastructures and service providers for Europe, has presented the requirement for trust anchors as one part of the trust services in one of their recent publications [Ga21].

Finally, in the current pandemic situation, many governments have begun designing and implementing Covid certificate systems. With the EU Digital Covid Certificate, the lack of a global trust architecture and ready-to-deploy tools to build compatible systems in other countries could not be clearer. Consequently, the Linux Foundation for Public Health (LFPH) has launched the Global Covid Credential Network (GCCN) [Gl22] to address this gap by adapting and operationalizing the Interoperability Blueprint of the Good Health Pass Collaborative, an industry coalition that has defined principles and standards for Covid certificates. GCCN has also established a special working group called the Trust Registry Network, which also addresses the requirement of the trust registry component.

4. The role of TRAIN in the SSI ecosystem

TRAIN stands for "**TR**ust m**A**nagement **IN**frastructure". It was a subproject run by Fraunhofer-Gesellschaft in the EU NGI eSSIF-Lab project [ES22]. The basic conceptual approach of TRAIN as a lightweight trust infrastructure was first published in [KR21].

TRAIN makes use of the global, well-established, and trusted infrastructure of the Internet Domain Name System DNS as its root of trust. The DNS is already used and trusted by everyone. However, it is susceptible to cache poisoning and MITM attacks, which can lead to false results being returned. DNSSEC [Ar05] has been specified to ensure that the results are authentic and have not been tampered with. Consequently, TRAIN uses DNSSEC whenever this is available.

The basic technology used by TRAIN has already been developed and validated in several pilots of the EU LIGHTest project (which developed the general context for trust in digital transactions). For the reference architecture of this approach please refer to [Wa17].

TRAIN addresses the issue of establishing trust in certain institutions in the SSI ecosystem beyond that achieved through mainly technical means, i.e., cryptography. An example would be the verification of the credibility of credential issuers, e.g., to find out whether the credential issuers really are who they claim to be. At the time of writing this paper (February 2022), there are around 113 different registered DID methods (and probably many more unregistered ones) [W321b]. Each method might have a different technical backend implementation. But irrespective of the technical infrastructure behind a certain SSI infrastructure, verifying the institutional trust between different components still remains an open issue. This is where TRAIN steps in. Using TRAIN, the verifier has the possibility of subscribing to one or several trust schemes that can be defined by trustworthy institutions (the roots of trust), thereby giving the verifier the opportunity to verify the credibility of issuers, regardless of their technical infrastructure.

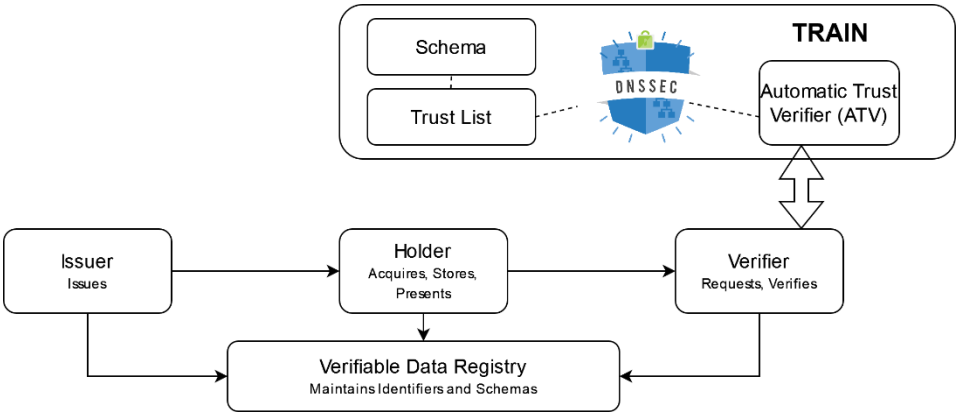


Figure 1 Integration of TRAIN into W3C VC SSI ecosystem

The architecture and integration of TRAIN into the W3C VC SSI ecosystem is shown in Figure 1. Although the TRAIN infrastructure uses the DNS for lookups, the trust schemes, JSON schemas and trust lists are distributed on the web and are not stored in the DNS. There can be different instances of trust lists and trust schemes hosted by different trust scheme operators (institutions providing trust schemes), and the verifier alone can decide which existing trust schemes and trust lists (for example: eIDAS) to trust. The TRAIN

API acts as a bridge between the SSI ecosystem and the TRAIN infrastructure. The TRAIN ATV component can be deployed as a cloud service but can be operated by the verifiers themselves in their own infrastructure as well.

5. Enabling individual and distributed trust decisions based on TRAIN

Currently TRAIN assists verifiers in making trust decision about VC issuers (although it could be extended in future to allow users to make trust decisions about verifiers). In essence, a VC issuer can make any statement it wishes (true or false) about the trust schemes it is a member of, whilst the verifier uses the TRAIN infrastructure to determine whether any statement is true or not.

5.1. Creation and publication of trust schemes

Any DNS owner can create their own trust scheme and become a trust scheme operator e.g., tso.com (see Figure 2). Similarly, every VC verifier decides which trust scheme operators to trust. Trust scheme operators decide which VC issuers are members of its trust scheme and therefore are trusted to issue VCs of a certain type with a certain schema. The trust scheme operator publishes the members of its trust scheme in a trust list that is based on the ETSI standard TS 119 612. A trust scheme operator, e.g., tso.com, with scheme "example" may also trust the trust scheme e.g., "ssi", of another trust scheme operator, e.g., company.uk. Consequently, it may wish to include the members of another trust scheme as being equivalent to its own members. The trust scheme operator would therefore add pointer resource records (PTR RRs) to its DNS trust scheme entry (as described below) to point to these other equivalent trust schemes. The use of PTR RRs forms mappings between Trust Schemes and Trust Lists. TRAIN offers the flexibility to create different trust schemes mapped to different trust lists according to the requirements of a certain trust framework.

5.2. DNS structure

The DNS controller creates a DNS entry with the name of its trust scheme e.g., example.tso.com., or ssi.company.uk. Then below this, two further DNS entries named *_trust* and *_scheme* respectively are created as shown in Figure 2. The names of these two entries were specified by the EU Lightest project [Wa17], and TRAIN is following those guidelines. Here, *example* is the scheme name, *tso.com* is the authority responsible for the scheme, and *_scheme._trust are* standardized constant terms used across the TRAIN trust infrastructure. The bottom entry, e.g., _scheme._trust.example.tso.com, contains one or more PTR RRs. Each PTR RR points to a DNS entry where the location of an ETSI trust list can be found, in a URI RR. This use of PTR RRs allows one trust scheme to point to several ETSI trust lists, for example, one EU country could point to the equivalent trust

schemes each located in a different country of the EU. It also allows an ETSI trust list to be incorporated into multiple trust schemes.

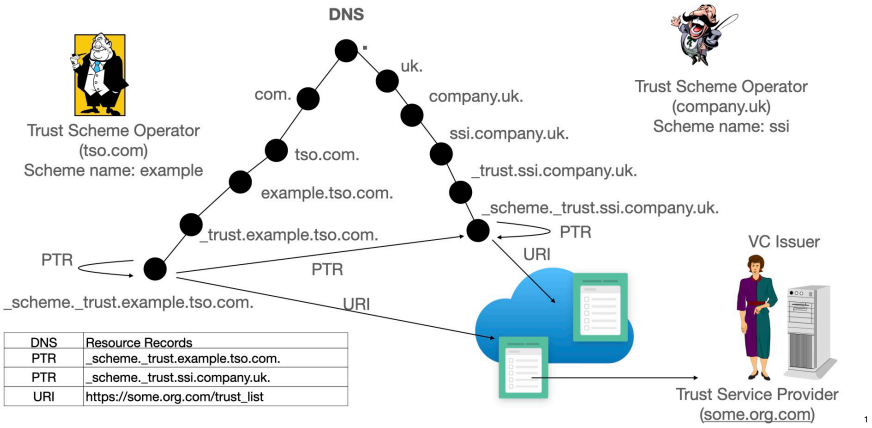


Figure 2. TRAIN use of DNS Records

5.3. Trust lists and JSON schema formats

Trust Lists used by TRAIN follow the ETSI TS 119 612 standard [ET16] and list all the enrolled entities (Issuers) in a specific data file/format certified by the issuing authority. An exemplary trust list is given in the following:

```
<TrustServiceProvider>
 <TSPInformation>
  <TSPName>
   <Name xml:lang="en">BGE</Name>
  </TSPName>
  <IssuerName>
   <Name xml:lang="en">https://vc.bge.verifiablecredentials.net</Name>
  </IssuerName>
  <TSPTradeName>
   <Name xml:lang="en">VATES-11111111</Name>
  </TSPTradeName>
  <TSPAddress>
   <PostalAddresses>
   <ElectronicAddress>
  </TSPAddress>
  <TSPInformationURI>
   <URI xml:lang="en">https://www.inclusion.gob.es/en </URI>
  </TSPInformationURI>
 </TSPInformation>
 <TSPServices>
  <TSPService>
   <ServiceInformation>
     <ServiceTypeIdentifier>
      https://train.trust-scheme.de/schema/gasBill-
schema.json</ServiceTypeIdentifier>
```

```
    <ServiceName>
      <Name xml:lang="en">Gas Bill</Name>
    </ServiceName>
    <ServiceDigitalIdentity>
      <DigitalId>
             <X509Certificate>...</X509Certificate>
      </DigitalId>
    </ServiceDigitalIdentity>
      <ServiceStatus>
  http://ehic.essif.trust-scheme.de/ServiceTypes/Servicestatus/granted
</ServiceStatus>
        <StatusStartingTime>2021-05-11T00:00:00Z</StatusStartingTime>
    </ServiceInformation>
  </TSPService>
 </TSPServices>
</TrustServiceProvider>
```

Every trusted VC issuer's details are described under the attribute *<TrustServiceProvider>*. The ID of the issuer is under the attribute *<IssuerName>*. Each VC issuer in the trust list has a Service Type Identifier under the attribute *<ServiceTypeIdentifier>*. This is a URL, and the web page that it points to should contain the JSON schema (including the @context property) for the VCs that are issued for this Service Type. In this way the verifier can find out which attributes the issuer is trusted to issue[4]. This trust list also offers the flexibility to the service provider to add different services with different schemas. An example of such a JSON Schema follows:

```
{
    "$schema": "http://example.com/gasBill",
    "issuer": "https://vc.bge.verifiablecredentials.net",
    "@context": [
        "https://www.w3.org/2018/credentials/v1",
        "https://bge.co.uk/VCcontext/v1",
        "https://schema.org/"],
    "credential_type": "GasBill",
    "type": "object",
    "properties": {
        "name": {
            "name": "credentialSubject.name",
            "type": "string",
            "maxLength": 64 },
        "address": {
          "name": "credentialSubject.address",
          "type": "object",
          "properties": {
            "streetAddress": {
                "name": "credentialSubject.address.streetAddress",
                "type": "string",
                "maxLength": 64 },
```

[4] The W3C VC Data Model specifies the credentialSchema property which allows issuers to publish the location of their VC schemas inside the VCs they issue. By including the same URL in the Trust List allows verifiers to trust that the issuer's schema location is correct.

```
                "postalCode": {
                    "name": "credentialSubject.address.postalCode",
                    "type": "string",
                    "maxLength": 64 },
                "addressLocality": {
                    "name": "credentialSubject.address.addressLocality",
                    "type": "string",
                    "maxLength": 64 },
                "addressCountry": {
                    "name": "credentialSubject.address.addressCountry",
                    "type": "string",
                    "maxLength": 64 } },
              "required": [ "streetAddress", "postalCode",
"addressLocality", "addressCountry" ]},
            "previousRead": {
                "name": "credentialSubject.gasBill.previousRead",
                "type": "number",
                "exclusiveMinimum": 0 },
            "presentRead": {
                "name": "credentialSubject.gasBill.presentRead",
                "type": "number",
                "exclusiveMinimum": 0 },
            "units": {
                "name": "credentialSubject.gasBill.units",
                "type": "number",
                "exclusiveMinimum": 0 },
            "kwh": {
                "name": "credentialSubject.gasBill.kwh",
                "type": "number",
                "exclusiveMinimum": 0 },
            "price/kwh": {
                "name": "credentialSubject.gasBill.price/kwh",
                "type": "number",
                "exclusiveMinimum": 0 },
            "amountDue": {
                "name": "credentialSubject.gasBill.amountDue",
                "type": "number",
                "exclusiveMinimum": 0 } },
        "required": [ "name", "address", "previousRead", "presentRead",
"units", "kwh", "price/kwh", "amountDue" ]
}
```

5.4. TRAIN ATV

The TRAIN Automatic Trust Verifier (ATV) is designed to verify the trustworthiness of a VC issuer given minimal information. It only requires two inputs one is the trust scheme name, that is embedded as a DNS name in the VC (see section 5.4), and the other one is the URI of the VC issuer, obtained from the VC. The URI of the issuer is flexible and may depend on the backend technology being used by the VC ecosystem. For example: the URI can be a DID that could be anchored in a blockchain/distributed ledger, but it could also be a https URL from a PKI. The TRAIN ATV is not restricted by the backend technology behind the VC in the SSI ecosystem.

The TRAIN ATV will first attempt to connect to the DNS name server that holds the entries of the Trust Scheme Operator using DNSSEC. This provides an unbroken chain of trust from the root DNSKEY RR set to the Trust Scheme Operator's DNS entries. However, if DNSSEC is not available, it will use standard DNS. The reason for this is that support for DNSSEC is not within the control of the Trust Scheme Operator, and so we prefer to allow TRAIN to be used by those Trust Scheme Operators that are willing to accept the risks now, rather than forcing them to wait until DNSSEC is available to them. We recognize that this leaves the trust scheme open to certain attacks, such as DNS MITM and cache poisoning, but Trust Scheme Operators can perform this risk assessment before deciding to use TRAIN without DNSSEC.

The TRAIN ATV will read the PTR RRs, dereference the URI RRs, and expect to find an ETSI trust list published at this https URL. It will then check if the VC Issuer is listed in this DNS named trust list, and if so, will tell the verifier that the issuer is a trusted member of this trust scheme operated by this "DNS name". Likewise, it will tell the verifier the URL of the issuer's VC schema. In this way it does not matter whether the issuer was telling the truth or not in its issued VC. The TRAIN API and the DNS controller/trust scheme operator establish the root of trust.

The source code of the ATV is freely available under Apache 2.0 [ES22]. ATVs can be run by anyone, so there can be multiple distributed copies of this service running in clouds or locally, and verifiers only need to keep pointers to one or more of them to provide them with backup services or completely under their own control.

5.5. Configuration at the issuer side

Every VC that is issued by any issuer that supports the TRAIN trust scheme must contain a standard Terms of Use property containing the DNS names of the trust scheme(s) that the issuer is a member of. It must also contain a standard credentialSchema property listing the URL where the schema can be found, along with the syntax of the schema. These of course could be true or false statements. In any case as the Verifier will check them using the TRAIN API – what counts in the end is the actual inclusion of the details into the trust list of the Trust Scheme Operator. This enrolment is another process that is beyond this paper. The exact format of the TRAIN Terms of Use property is given below:

```
"termsOfUse": [{
    "type": "https://train.trust-scheme.de/info",
    "trustScheme": ["example.tso.com", "ssi.company.uk"]
}]

"credentialSchema": {
    "id": "https://train.trust-scheme.de/schema/gasBill-
schema.json",
    "type": "JsonSchemaValidator2018"

  }
```

According to the W3C VC recommendation, each Terms of Use must have a globally unique type, and we have reserved the value "*https://train.trust-scheme.de/info*" to refer to the TRAIN Terms of Use type.

5.6. Configuration at the verifier side

All the verifier has to do is configure the DNS names of the trust schemes that it trusts, and the URL(s) of the TRAIN ATV API(s) to call to verify their membership lists. When it receives a VC, it extracts the asserted trust schemes made by the issuer in the ToU property, and if it trusts any of the listed trust schemes, it calls the TRAIN API, passing it the URI of the issuer (taken from the VC) and the DNS name of the trusted trust scheme that the VC Issuer purports to be a member of.

```
{
   "Issuer": "https://vc.bge.verifiablecredentials.net",
   "Trust_Scheme_Pointer": "ssi.company.uk"
}
```

The TRAIN API will then check if the VC issuer is a member of any of the trust lists pointed to by this trust scheme, and if so, return the Service Type URL to the Verifier. The verifier can check that this URL is identical to the one in the credentialSchema property, and if it is, use the schema contained at this URL to validate that the attributes in the received VC match the schema for this Service Type.

6. Limitations and future work

The integration of TRAIN with VC issuers and verifiers has been described in this paper, but the integration of TRAIN with other components like the VC holder still remains to be done. This work is currently being specified by the OpenID Foundation and the Decentralized Identity Foundation (DIF). DIF is defining presentation definitions, as part of the presentation exchange specification [DI22], which allows a verifier to indicate to a holder, which VCs it should return in a verifiable presentation. OpenID Connect is being enhanced so that it can transfer presentation definitions from the verifier to the holder, and verifiable presentations from holders to verifiers, using the OIDC SIOPv2 protocol. The latest draft of OpenID Connect for Verifiable Presentations [Op22] contains informative implementation guidelines describing how issuers, holders and verifiers can utilise the TRAIN trust scheme approach.

Currently ETSI Trust Lists only support X.509 PKI public keys. Other credential infrastructures such as WHO and EU COVID-19 certificates also use X.509 PKIs. Clearly X.509 PKIs are globally accepted and operational, which is why we chose them for our initial TRAIN trust infrastructure. Several SSI infrastructures that work with X.509 PKIs have already successfully integrated TRAIN into their infrastructures. Further work describing how to incorporate DID public keys in ETSI Trust lists is planned, thereby

offering the possibility of integrated trust lists with both X.509 PKI and DID public keys.

Moreover, TRAIN is being developed further towards an architecture that could accommodate DIDs as trust-scheme pointers besides DNS HostNames. This would enable a TRAIN trust infrastructure that does not rely on the DNS System, but could rely on alternative trust anchors, such as blockchains or other distributed ledgers.

Finally, TRAIN can be used to hold trust lists of verifiers, and holders could use this to determine which Relying Parties can be trusted to receive their identity attributes. Specifying this in detail is also further work that is planned.

7. Conclusion

In order to accomplish the promise of a bright future for identity management, SSI solutions need to address some fundamental trust and governance challenges. Although the current SSI offers modern cryptographic trust to enhance the privacy of users, the challenge of institutional trust still needs to be addressed. Without institutional trust, it is impossible for a relying part to verify the credibility of a VC issuer.

The TRAIN approach leverages the DNS, an already proven and universally accepted trust anchor, to provide a trust management infrastructure for SSI in a distributed manner. This is an important first step in providing the necessary credibility to make SSI also attractive for relying parties.

Bibliography

[Al16] Allen, C.: The Path to Self-Sovereign Identity, GitHub, 26/4/2016., https://github.com/ChristopherA/self-sovereign-identity, accessed: 5/2/2020.

[Ar05] Arends, R. et.al.: DNS Security Introduction and Requirements, Internet Engineering Task Force, Request for Comments RFC 4033, doi: 10.17487/RFC4033, 3/2005.

[DI22] DIF Presentation Exchange. https://identity.foundation/presentation-exchange/, accessed: 11/2/2022).

[EN22] ENISA, “Digital Identity: Leveraging the SSI Concept to Build Trust,” European Union Agency for Cybersecurity., Athens, Heraklion, 1/2022.

[ES22] ESSIF-LAB, eSSIF-TRAIN by Fraunhofer-Gesellschaft | eSSIF-Lab., https://essif-lab.eu/essif-train-by-fraunhofer-gesellschaft/, accessed: 11/2/2022.

[ET16] ETSI: Electronic Signatures and Infrastructures (ESI), General Policy Requirements for Trust Service Providers. ETSI, Sophia Antipois Cedex, France, European Standard ETSI EN 319 401, 2016.

[EU21] EU Commission, A cybersecure digital transformation in a complex threat environment — Brochure, Shaping Europe’s digital future, 28/1/2021.

[F22] Federal Office for Information Security, A brief guideline on self-sovereign identities. 22/02/2022, Bonn.

[Ga21] Gaia-X Federation Services (GXFS) White Paper - Gaia-X Ecosystem Kickstarter, GAIA-X., https://gaia-x.eu/sites/default/files/2021-12/GXFS_1.pdf, 12/2/2021.

[Gl22] Global COVID Certificate Network (GCCN), Linux Foundation Public Health., https://www.lfph.io/global-covid-certificate-network/, accessed: 12/1/2022.

[IN20] INATBA, Decentralized Identity: What is at Stake?, INATBA Identity Working Group, Nov. 2020., https://inatba.org/news/inatba-identity-working-group-publishes-position-paper-on-decentralised-identity/, accessed: 8/2/2022.

[KR21] Kubach, M.; Roßnagel, H.: A lightweight trust management infrastructure for self-sovereign identity, in Open Identity Summit 2021 - Lecture Notes in Informatics (LNI) - Proceedings, Roßnagel, H., Schunck, C.H., and Mödersheim, S., Eds. Bonn: Köllen Druck + Verlag GmbH, pp. 155–166, 2021.

[Ku20] Kubach, M.; Schunck, C.H.; Sellung, R.; Roßnagel, H.: Self-sovereign and Decentralized identity as the future of identity management?, in Open Identity Summit 2020 - Lecture Notes in Informatics (LNI) - Proceedings, Bonn: Köllen Druck + Verlag GmbH, pp. 35–47, 2020.

[Op22] OpenID Connect for Verifiable Presentations, 24/4/2022., https://openid.net/specs/openid-connect-4-verifiable-presentations-1_0.html, accessed: 24/4/2022.

[Si18] Simons, A.: Decentralized digital identities and blockchain: The future as we see it, Microsoft Blog, https://www.microsoft.com/en-us/microsoft-365/blog/2018/02/12/decentralized-digital-identities-and-blockchain-the-future-as-we-see-it/, accessed: 5/2/2020.

[St21] Strüker, J. et.al.: Self-Sovereign Identity - Foundations, applications, and potentials of portable digital identities. Project Group Business & Information Systems Engineering of the Fraunhofer Institute for Applied Information Technology FIT, Bayreuth, 2021.

[To22] ToIP TSS0001>: Trust Registry Specification. Trust over IP Foundation, 2021., https://github.com/trustoverip/tswg-trust-registry-tf, accessed: 12/1/2022.

[Tr22] Trust Over IP - Defining a complete architecture for Internet-scale digital trust, Trust Over IP., https://trustoverip.org/, accessed: 18/2/2022.

[W319] W3C, Verifiable Credentials Data Model 1.0, W3C Recommendation 19/11/2019., https://www.w3.org/TR/vc-data-model/, accessed: 6/2/2020.

[W321a] W3C, Decentralized Identifiers (DIDs) v1.0, W3C Working Draft 8/2/2021, 2/8/2021., https://www.w3.org/TR/did-core/, accessed: 9/2/2021.

[W321b] W3C, DID Specification Registries," W3C Working Group Note 2/11/2021, 2021., https://www.w3.org/TR/did-spec-registries/#did-methods, accessed: 18/2/2022.

[Wa17] Wagner, S.; Kurowski, S.; Laufs, U.; Roßnagel, H.: A mechanism for discovery and verification of trust scheme memberships: the LIGHTest Reference Architecture, in Open Identity Summit 2017 - Proceedings, Lecture Notes in Informatics (LNI), Bonn: Köllen Druck + Verlag GmbH, pp. 81–92, 2017.

H. Roßnagel, C.H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 39

Continuous authorization over HTTP using Verifiable Credentials and OAuth 2.0

Nikos Fotiou, Evgenia Faltaka, Vasilis Kalos, Anna Kefala, Iakovos Pittaras, Vasilios A. Siris, George C. Polyzos[1]

Abstract:

We design, implement, and evaluate a solution for achieving continuous authorization of HTTP requests exploiting Verifiable Credentials (VCs) and OAuth 2.0. Specifically, we develop a VC issuer that acts as an OAuth 2.0 authorization server, a VC verifier that transparently protects HTTP-based resources, and a VC wallet implemented as a browser extension capable of injecting the necessary authentication data in HTTP requests without needing user intervention. Our approach is motivated by recent security paradigms, such as the Zero Trust architecture, that require authentication and authorization of every request and it is tailored for HTTP-based services, accessed using a web browser. Our solution leverages JSON Web Tokens and JSON Web Signatures for encoding VCs and protecting their integrity, achieving this way interoperability and security. VCs in our system are bound to a user-controlled public key or a Decentralized Identifier, and mechanisms for proving possession are provided. Finally, VCs can be easily revoked.

Keywords: Access control; Authentication; Zero Trust

1 Introduction

In the recent years, the global pandemic made remote working a necessity rather than an option. Nevertheless, this came at a cost: according to a recent research 74% of organizations attribute business-impacting cyber attacks to vulnerabilities in technology put in place during the pandemic.[2] For this reason, more and more enterprises embrace security approaches such as the *Zero Trust* paradigm. The main concept of Zero Trust is "never trust, always verify", which means, among other things, that every request should be authenticated and authorized. This architecture begs for new, secure, lightweight access control solutions, with increased interoperability and without adding privacy threats. In this paper, we propose a security solution that can be used for providing authorization for every HTTP request. Our solution leverages *Verifiable Credentials* (VC) [Ma19] and provides efficient VC management, improves interoperability, and enhances user security and privacy.

[1] Athens University of Economics and Business, Mobile Multimedia Laboratory {fotiou,eugeniafaltaka,kalos20, kefala,pittaras,vsiris,polyzos}@aueb.gr

[2] https://www.tenable.com/press-releases/seventy-four-percent-of-organizations-attribute-damaging-cyberattacks-to

doi:18.18420/OID2022-03

Our solution considers users of an enterprise wishing to access an HTTP-based protected resource using a web browser. Both users and the protected resource may be located in networks outside the administrative realm of the enterprise. From a high-level perspective our solution operates as follows: Users interact with an authorization server, owned or controlled by the enterprise, using a *wallet*, implemented as a browser extension. The authorization server responds with a Verifiable Credential (VC) that contains the capabilities of the user, which is stored in the user's wallet. Then, the user interacts with a protected resource (through the web browser) and includes in the corresponding requests: (i) the received VC and (ii) a Proof of VC Possession. A VC *verifier*, acting as a transparent HTTP proxy, intercepts the communication between the web browser and the protected resource, validates the VC based on pre-configured rules and confirms the Proof of Possession. If all checks succeed the verifier forwards the HTTP request to the resource.

Our solution also enables authorization servers to provide an efficient revocation mechanism. This revocation mechanism includes a compact list of revoked VCs encoded in a self-verifiable data structure. The revocation list can be received directly from the issuer, or it can be provided indirectly; it can be even included in a resource access request.

Compared to legacy OAuth 2.0 solutions, our system provides the following advantages:

- The generated VCs are bound to a user-controlled identity, therefore they can be stored for a longer interval and they cannot be used by entities that have intercepted them (unlike, for example, mere "bearer tokens"). Hence, client applications do not have to interact often with an authorization server.
- Our system uses VCs as "access tokens". VCs support richer semantics, can be used for evaluating complex access control policies, and facilitate interoperability.
- Our system provides an efficient mechanism for checking the revocation status of an access token/VC.

Compared to related VC-based solutions, our design provides the following advantages:

- Our system builds on the widely used and well supported OAuth 2.0 flows for managing the lifecycle of a VC. These flows are implemented in a browser based wallet achieving continuous and secure authorization over HTTP without user intervention.
- Our system leverages JSON Web Tokens (JWT) [JBS15b] and JSON Web Signatures (JWS) [JBS15a] for representing and protecting VCs. These are widely used and standardized solutions (as opposed to, for example, linked-data proofs).
- Our system supports user authentication using both public public keys, as well as "Decentralized Identifiers" (DIDs) [W321]. DIDs allow users to rotate their secret keys without having to receive a new VC.

The remainder of this paper is organized as follows. In Section 2 we detail the design of our solution. In Section 3 we present the implementation and evaluation of our solution. We discuss related work in Section 4 and we conclude our paper in Section 5.

2 Design

In this section we detail the components of our system and their interactions (also illustrated in Fig. 1).

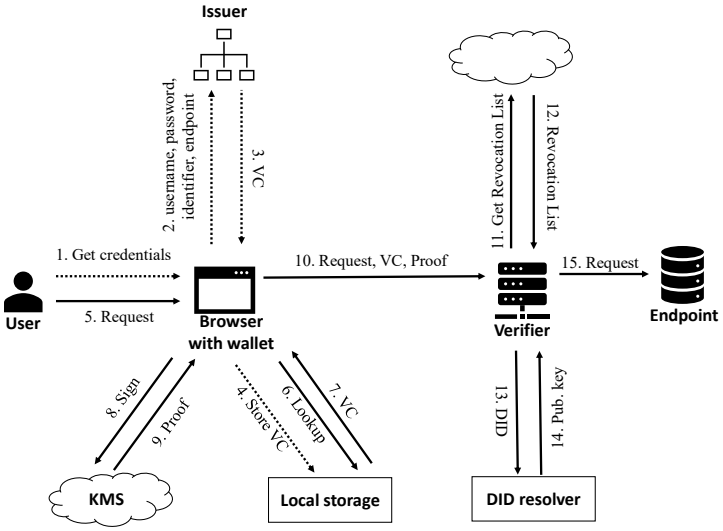


Fig. 1: An instance of our system

2.1 Components

Our system is composed of a VC *issuer*, a VC *verifier*, and a *wallet*. The goal of our system is to allow authorized users to invoke *operations* on *resources* stored in an particular *endpoint*, for example, perform a "list" operation, on a "folder", stored in a Cloud storage system. An endpoint is identified by a URL, denoted by $URL_{Endpoint}$, and it can be oblivious to our system. Our system supports the use of *Decentralized Identifiers* (DID) as a mean for identifying various entities. DIDs are URIs which resolve to a DID document that contains information related to the DID, such as ways to cryptographically authenticate the DID owner. The structure of a DID document and the DID resolution mechanism are specific to each DID *method*. We provide more details about how DIDs are implemented in our system n section 3.1

The VC issuer is an OAuth 2.0 authorization server extended with VC issuing capabilities. Each VC issuer is identified by an ID_{issuer}, which can be a URL or a DID. Furthermore,

each issuer owns a public-private key pair and we assume a secure method for resolving an ID_{issuer} to the corresponding public key (e.g., through static configuration, using the web PKI, or by performing a DID resolution). Issued VCs are encoded as JWTs, and signed using a JSON Web Signature (JWS) and the private key of the issuer. VCs in our system "describe" the capabilities of a VC *subject* over a protected endpoint. Additionally, a VC issuer maintains a VC revocation list.

The VC verifier is an HTTP proxy that intercepts HTTP requests towards a protected endpoint. The VC verifier is able to verify the validity, the status, and the ownership of VCs included in the intercepted requests. Additionally, the VC verifier acts as a *policy enforcement point* by validating whether or not a VC can be used for executing the requested operation over a resource.

The wallet is a web browser extension that interacts with the VC issuer and verifier using standard OAuth 2.0 flows. The wallet is responsible for storing the received VCs, and for including them in the corresponding HTTP requests. Additionally, the wallet generates and manages user owned identifiers, which can be public keys or DIDs. Such an identifier is included in a VC and the associated secret key is used by the wallet for generating a VC *proof of possession*. Users may have multiple identifiers, as well as multiple wallets.

2.2 Interactions

Our system involves a configuration step, after which the system components can interact with each other using OAuth 2.0 flows.

2.2.1 System configuration

This is a step usually executed during a set-up phase. With this step an issuer is configured with policies that specify the capabilities that correspond to a user. Additionally, users *register* with the issuer (at least) a wallet. With this registration process, users create a username and password for their wallet and assigns to it a subset of their capabilities. The generated username and password will be later used by the wallet in order to retrieve VCs from the issuer. Finally, verifiers are configured with a list of trusted ID_{issuer} identifiers and (if required) with their corresponding public keys.

2.2.2 VC request and issuance

With this step, a user's wallet requests from the issuer a VC that can be used for accessing a particular endpoint. A VC request is in essence an OAuth 2.0 access token request using the client credentials grant (section 4.4 of [He12]); in our system the corresponding "client

credentials grant" is the wallet's username and the password registered to the issuer during the configuration phase. The wallet includes in this request an identifier (which can be either a public key, or a DID). The wallet may re-use an existing identifier or it may generate a new one, specific to that particular VC request. The issuer verifies the provided username and password, and retrieves the capabilities associated with them. Then it creates a VC that includes these capabilities and the provided identifier, encodes it as a JWT, and signs it. An example of a VC as used in our system follows. As it can be seen, the standard *iss* and *aud* JWT claims are used for denoting the issuer and the target endpoint of the VC. If the provided identifier is a DID it is included in a *sub* claim; if it is a public key the *cnf* claim as defined in RFC7800 is used. The VC may include additional JWT claims that control its validity period. Finally, the *vc* claim includes information that can be used for determine the VC's revocation status, as well as a list of "resources" and allowed "operations".

```
{
    "iss": ID_issuer,
    "aud": URL_Endpoint,
    "sub": User owned DID,
    "vc":{
      "@context": [...],
      "id": "credential 1",
      "credentialStatus": {...},
      "credentialSubject": {
        "type": ["CapabilitiesCredential"],
        "Resource1": [ "Operation 1", "Operation 2" ]
      }
    }
  }
```

2.2.3 VC revocation

Our revocation mechanism is based on the system described in [SDS20]; a similar approach for VC revocation is followed by a recent W3C draft [Gr20]. In order to support revocation, an issuer maintains a revocation list that covers all *not expired* VCs it has issued. This list is a simple bitstring and each VC is associated with a position in the list. In particular, each revocable VC includes a property named *revocationListIndex* that specifies the position of the VC in the revocation list. Revoking a VC means setting the bit corresponding to the VC to 1. Since the list includes only non-expired VCs, its size is tolerable for most use cases. For example, an issuer that issues on average 100 VCs per day with lifetime equal to one month, would only need 30×100 bits to store its revocation list. Any entity can verify the status of a non-expired VC that supports this revocation mechanism, by examining the value of the bit of the corresponding revocation list.

A revocation list is included in a JWT, signed and timestamped by the issuer. This JWT can be retrieved directly by the issuer, or indirectly, e.g., the issuer can store it in an online location, or even in a blockchain. Each revocable VC includes a property named *statusListCredential* which is a "pointer" (e.g., a URL) to the revocation list location. A verifier can retrieve the revocation list by itself, or require from users to include it in their requests. In all cases, the verifier has to validate the signature of the JWT and determine its "freshness".

2.2.4 Endpoint access

A user can request from an endpoint to perform an operation over a resource. This request is transmitted over HTTP using the user's web browser. If the $URL_{Endpoint}$ is included in the *aud* claim of a stored VC, the wallet retrieves this VC from its local storage and prepares a *proof of possession*. This proof is generated according to "Demonstration of Proof-of-Possession at the Application Layer" (DPoP) [D.20] OAuth 2.0 extension. DPoP has been designed for HTTP communication and achieves PoP in a single message. In particular, with DPoP the wallet creates JWS signed using the key the corresponds to the user identifier included in the VC. The DPoP payload includes at least a unique, sufficiently large random number, the HTTP *method* of the request, the HTTP *URI* of the request, and the *time* when the proof was created. Then, the wallet includes the VC in the *Authorization* HTTP header of the request and the generated proof in a *DPoP* HTTP header. The request is received by the verifier that acts as an HTTP proxy. The verifier initially validates the included VC. In particular, it examines if the VC is signed by a trusted issuer and if the value of *aud* claim equals to the the $URL_{Endpoint}$. Additionally, if the VC includes claims that control its validity period, it examines if the VC is valid. Then, it extracts the user identifier included in the VC. If that identifier is a DID, the verifier performs a DID document resolution and retrieves the corresponding public key. Then, it verifies the signature of the provided DPoP using the public key associated with the user identifier, and it examines if the DPoP is "sufficiently fresh", if it includes the correct HTTP method and URI, as well as if the included random number has not been "recently" used. If the VC is revocable, it examines the status of the VC by retrieving the revocation list from the VC issuer. Finally, it verifies if provided VC includes the capabilities that are necessary for invoking the requested operation. If all checks succeed, the verifier forwards the request to the endpoint.

3 Implementation and evaluation

We have implemented[3] our issuer as a .net core web application, and our verifier using Python3 and the jwcrypto library[4]. Moreover, we integrated DIDs in our system using DIF's

[3] Pointers to GitHub repositories of our implementations can be found in `https://mm.aueb.gr/projects/zerotrustvc`

[4] `https://jwcrypto.readthedocs.io`

Universal Resolver [DI21], and we have implemented our wallet as a Firefox extension. We provide more details about the use of the Universal Resolver and about our browser-based wallet in the following subsections. Then we present our evaluation scenario and we discuss the performance and the security properties of our solution.

3.1 Support for DIDs

Our system supports DIDs, which is common practice in most VC systems. The DID standard allows each DID method to define its own way for resolving a DID to the corresponding document. To avoid any further complexity and to contribute to the interoperability of our project, we rely on DIF's Universal Resolver for DID document resolution. The Universal Resolver performs DID resolution across many different DID methods by providing a universal API. Internally, this is accomplished through an architecture consisting of drivers for each supported method.

Our current implementation supports the did:web method [Mi21]. This is a DID method that bootstraps trust by leveraging an existing web domain's reputation. A did:web DID is constructed based on a URL which when resolved results in the corresponding DID document. Our verifier implementation uses a local instance of the Universal Resolver to resolve did:web DIDs. It interacts with it through an API that receives as input a did:web DID and responds with the corresponding public key.

3.2 Browser-based wallet

User's wallet is implemented as a browser extension. The extension is tasked with requesting the credentials from the issuer and presenting them to the verifier. Internally, each credential is stored to the browser's supplied storage and is indexed based on the URL included in the *aud* claim. This not only allows for fast lookups, but also for syncing those credentials between browsers in different devices. Furthermore, users will also have the ability to back up their credentials to some other source (i.e., the cloud, the file system etc.).

Users provide to the extension an ID_{issuer} and the appropriate username and password. Then, the extension either creates a new cryptographic key pair, or re-uses an existing did:web DID and communicates with the issuer. If successful, the extension will read the credential from the issuer's response, parse it and update its internal state. This is the only process that involves user intervention. To present a credential, the extension in the background listens for any HTTP request made to a URL for which there is a saved credential that includes that URL in the *aud* claim. When such a request is made, the extension will retrieve that credential and create a fresh DPoP value using the appropriate cryptographic key material. Then, the extension injects the DPoP and the VC as new HTTP headers, in the original HTTP request.

The extension is implemented for the Firefox browser. To manage the user's cryptographic key material we have adopted 2 different strategies; in-browser key management and using external key management systems (KMS). In the former case, the extension saves private keys encrypted in the browser's storage, while public keys are saved in clear. To encrypt a private key we use AES with a key derived from a user supplied password using PBKDF2. In the latter case, we rely on a Cloud-based KMS (but any external KMS can be trivially supported). In that case, private keys are stored in the Cloud and when the wallet needs to sign something (i.e., to create a DPoP), the cryptographic hash of the raw data is sent to the KMS. This method, although it adds an additional round trip, it alleviates the need for the extension to manage private keys itself.

3.3 Cloud Storage scenario

A Cloud Storage access scenario is implemented in order to evaluate the proposed architecture. In this use case, the protected resource is a Google Cloud Storage Bucket. Buckets are the basic containers in Google Cloud Storage and are used in order to organize data and control access to them.

According to our use case an employee of the enterprise wishes to access some data from a bucket via a web interface. The user must first request a VC through the browser extension acting as her VC wallet. The issued VC specifies certain capabilities such as *read*, *write*, *upload*, or *list*. The web interface is provided by a Python3 script, which implements the Google Storage API: this script is the protected endpoint. This script interacts with the Cloud storage using a "service" account, hence both the script, as well as the Cloud storage provider are oblivious to the used VCs, as well as to the user management system and the access control policies of the enterprise.

3.4 Overhead

We have measured the VC issuing processes, DPoP generation, and access request verification in a desktop PC equipped with an Intel i5 5540 CPU and 8GB RAM, running Windows 10 using EdDSA and ES256 signature algorithms for JWS. All operations require less than $0.1ms$.

Since VCs and DPoPs are transmitted in HTTP headers they are encoded using base64. The base64 encoding of a VC that includes two resources and two operations is 656 bytes. Similarly, the base64 encoding of a DPoP is 440 bytes.

A revocation list is stored in a signed JWT. This JWT also includes the *iss* claim, which defines the issuer, and the *iat* claim, which defines the date and time at which the token was issued. Such a signed token, which is encoded using base64 encoding, generated using

ES256 JWS algorithm, including the verification key in the JWS header and a revocation list with 4000 entries, is 1431 bytes long.

3.5 Security properties

Our solution leverages OAuth 2.0, whose security properties have been formally verified [FKS16], for managing the lifecycle of VCs and it provides proof of possession, preventing this way many security attacks. Additionally, our system achieves the following security properties:

Increased availability. Non-revocable VCs can be verified without needing the issuer to be online. When it comes to revocable VCs various optimizations can be considered for decreasing the dependence on the availability of the issuer, e.g., cache revocation lists for some time, store revocation lists in alternative locations. Similarly, verifiers do not have to maintain any user specific state since all the information required to make an access control decision is included in each request; verifiers are only required to maintain for a limited time the nonce included in a DPoP in order to prevent replay attacks.

Efficient access control management. User and access control policy management is implemented independently of the protected endpoint, since granting or revoking an access right does not involve any communication with the endpoint (or the verifier). Furthermore, by implementing the VC verifier as an HTTP proxy, we allow transparent protection of any HTTP-based resource.

Attack surface reduction. In our solution the amount of verifications a verifier needs to perform is less compared to a system that relies on Access Control Lists (ACLs), which are inflexible, do not scale well, and are difficult to use and upgrade [Ka06]. In our system, a verifier has only to verify the validity and the possession of the VC included in an access request. Furthermore, verifiers are not required to store any additional secret information to implement our protocols, neither do they have to maintain user accounts. Moreover, a user is allowed to use multiple wallets and assign to each wallet different capabilities. For example, a wallet used in a "travel laptop" may have less capabilities compare to a wallet used in a secured, well-administrated PC. Finally, our wallet selects the appropriate VC by itself, by matching the requested URL with the URL included in the "aud" claim of each VC: this approach is less prone to security attacks compared to most of the existing approaches that require user intervention, e.g., they require from users to scan a QRcode.

Resilience to attacks. Our system is resilient to many types of attacks. Since the VCs are bound to a user owned identifier, our system is not affected by attackers-in-the-middle that intercept the communication towards a protected endpoint. These attackers, can neither modify the transmitted VCs without being detected, nor re-use the captured VCs to their own purposes. Similarly, our system allows different user identifiers per VC, hence, even if the private key that corresponds to an identifier is breached, the captured VCs can only

be used for accessing a specific endpoint. Moreover, by including a DID such as did:web as a user identifier in a VC, a user can rotate the private key that corresponds to that DID without having to receive a new VC. Therefore, users can even proactively rotate their keys.

4 Related work

The problem of designing efficient authorization and access control solutions for Zero Trust Architectures (ZTA) is well known and there are many efforts that try to address it [Ya20, LHK20]. Lukaseder et al. [LHK20] discuss how the Zero Trust Model can be applied to open networks, such as in a network of a university. Furthermore, they implement and present a Zero Trust network framework, called Alekto that authenticates and authorizes users in order to take access control decisions and compute trust scores for an eLearning system. Yao et al. [Ya20] propose a dynamic and fine-grained access control and authorization solution for ZTA, which is composed of an access control agent, a user identity authentication module, an access control engine, and a trust evaluation engine. The main differences between our work and these works are that our solution uses VCs as access tokens, which include the client capabilities. The VCs can have longer lifetimes and they can be stored in secure wallets as opposed to mere bearer tokens.

Similarly to this work, Lagutin et al. [La19] try to integrate VCs and DIDs into the OAuth 2.0 protocol. In their solution, which is designed for constrained devices, they use VCs and DIDs as authentication grants. Clients use these grants to obtain access tokens from the authorization server. Our solution follows a reverse approach: clients use a username and password as a grant to obtain VCs. This has the advantage that authorization is enforced when requesting access to a resource. The solution in [FSP21] also combines VCs with OAuth 2.0 in order to provide capabilities-based access control. Our approach, improves the solution presented in [FSP21] by allowing users to use multiple identifiers and even use different identifier per VC, by adding support for Decentralized Identifiers, and by considering a wallet. In our system users can have multiple wallets and each wallet can be assigned different access rights: this property has many security advantages.

Our system leverages VCs for expressing capabilities because VCs are well understood techniques being standardized. Additionally, supporting a specific VC type is straightforward, hence interoperability can be supported with low effort. Related approaches that can be used instead of VCs in a system similar to ours are Macaroons [Bi14], and Authorization Capabilities for Linked Data (ZCAP-LD) [C.20]. Similarly, our system builds on OAuth 2.0, which is a widely used standard, hence our solution can be easily integrated in existing systems. Additionally, OAuth 2.0 has been designed specifically for HTTP services. Other related works propose new protocols such as the Credential Handler API [Ls21], and the Presentation Exchange protocol [BZR22].

5 Conclusion

In this paper we proposed a security solution that allows continuous authorization over HTTP. Our solution uses Verifiable Credentials (VCs) to store user capabilities. Additionally, it leverages OAuth 2.0 and a browser-based wallet to provide fast VC lifecycle management, without requiring any user intervention. Our solution implements proofs of possession preventing this way VC sharing. Additionally, our solution supports Decentralized Identifiers, allowing users to rotate their private keys without having to receive a new VC.

Our solution provides additional advantages, which are not highlighted by the use case considered in our paper. For example, our solution allows multiple issuers, it supports re-using the same VC for accessing different endpoints of the same type, and the considered revocation mechanism does not reveal to an issuer information about the user that tries to access an endpoint. Similarly, our solution allows wallets to include in user request a "fresh" copy of the revocation list, enabling this way offline verifiers, which can be of particular importance in cases such as IoT systems. Future work in this area involves the replacement of DPoP with Webauthn assertions, the application of our solution in the IoT by adding support for more efficient VC encodings (e.g., using CBOR), as well as for IoT specific protocols (e.g., CoAP), and the integration of VC verifier into the endpoints themselves.

Acknowledgments

The work reported in this paper has been funded in part by European Union's Horizon 2020 research and innovation programme through subgrant *Enabling Zero Trust Architectures using OAuth2.0 and Verifiable Credentials (ZeroTrustVC)* of project *eSSIF-Lab*, under grant agreement No 871932 and by the *Research Center of the Athens University of Economics and Business*.

Bibliography

[Bi14] Birgisson, Arnar; Politz, Joe Gibbs; Úlfar Erlingsson; Taly, Ankur; Vrable, Michael; Lentczner, Mark: Macaroons: Cookies with Contextual Caveats for Decentralized Authorization in the Cloud. In: Network and Distributed System Security Symposium. 2014.

[BZR22] Buchder, D.; Zundel, B.; Reidel, M.: Presentation Exchange v2.0.0. Working Group Draft, DIF, 2022. https://identity.foundation/presentation-exchange/.

[C.20] C. L. Webber, M. Sporny Eds: Authorization Capabilities for Linked Data. Draft Community Group Report, W3C, 2020. https://w3c-ccg.github.io/zcap-ld/.

[D.20] D. Fett et al.: OAuth 2.0 Demonstration of Proof-of-Possession at the Application Layer (DPoP). RFC draft, 2020.

[DI21] DIF Identifiers and Discovery Working Group: , DID Universal Resolver, 2021. https://github.com/decentralized-identity/universal-resolver.

[FKS16] Fett, Daniel; Küsters, Ralf; Schmitz, Guido: A Comprehensive Formal Security Analysis of OAuth 2.0. CCS '16, Association for Computing Machinery, New York, NY, USA, p. 1204–1215, 2016.

[FSP21] Fotiou, Nikos; Siris, Vasilios A.; Polyzos, George C.: Capability-based access control for multi-tenant systems using OAuth 2.0 and Verifiable Credentials. In: 2021 International Conference on Computer Communications and Networks (ICCCN). pp. 1–9, 2021.

[Gr20] Group, W3C Credentials Community: Revocation List 2020. Draft community group report, W3C, 2020. https://w3c-ccg.github.io/vc-status-rl-2020/.

[He12] Hardt (ed.), D: The OAuth 2.0 Authorization Framework. RFC 6749, IETF, 2012.

[JBS15a] Jones, M.; Bradley, J.; Sakimura, N.: JSON Web Signature (JWS). RFC 7515, IETF, May 2015.

[JBS15b] Jones, M.; Bradley, J.; Sakimura, N.: JSON Web Token (JWT). RFC 7519, IETF, 2015.

[Ka06] Karp, Alan H.: Authorization-Based Access Control for the Services Oriented Architecture. In: Fourth International Conference on Creating, Connecting and Collaborating through Computing (C5'06). pp. 160–167, 2006.

[La19] Lagutin, Dmitrij; Kortesniemi, Yki; Fotiou, Nikos; Siris, Vasilios A.: Enabling Decentralised Identifiers and Verifiable Credentials for Constrained IoT Devices using OAuth-based Delegation. Proceedings of DISS 2019 Workshop on Decentralized IoT Systems and Security, p. 6, 2019.

[LHK20] Lukaseder, Thomas; Halter, Maya; Kargl, Frank: Context-based Access Control and Trust Scores in Zero Trust Campus Networks. In: SICHERHEIT 2020. Gesellschaft für Informatik e.V., Bonn, pp. 53–66, 2020.

[Ls21] Longley, D.; sporny, M.: Credential Handler API 1.0. Draft Community Group Report, W3C, 2021. https://w3c-ccg.github.io/credential-handler-api/.

[Ma19] Manu Sporny et al.: Verifiable Credentials Data Model 1.0. W3C Recommendation, W3C, 2019. https://www.w3.org/TR/verifiable-claims-data-model/.

[Mi21] Michael Prorock et al.: did:web Method Specification. Editor's draft, W3C, December 2021. https://w3c-ccg.github.io/did-method-web/.

[SDS20] Smith, Trevor; Dickinson, Luke; Seamons, Kent: Let's Revoke: Scalable Global Certificate Revocation. In: Network and Distributed System Security Symposium. 2020.

[W321] W3C Credentials Community Group: Decentralized Identifiers (DIDs) v1.0. W3C Proposed Recommendation, W3C, 2021. https://www.w3.org/TR/did-core/.

[Ya20] Yao, Qigui; Wang, Qi; Zhang, Xiaojian; Fei, Jiaxuan: Dynamic Access Control and Authorization System Based on Zero-Trust Architecture. In: 2020 International Conference on Control, Robotics and Intelligent System. Association for Computing Machinery, New York, NY, USA, pp. 123–127, 2020.

H. Roßnagel, C.H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 51

Integration of Self-Sovereign Identity into Conventional Software using Established IAM Protocols: A Survey

Michael Kuperberg[1], Robin Klemens [2]

Abstract: Self-Sovereign Identity (SSI) is an approach based on asymmetric cryptography and on decentralized, user-controlled exchange of signed assertions. Most SSI implementations are not based on hierarchic certification schemas, but rather on the peer-to-peer and distributed "web of trust" without root or intermediate CAs. As SSI is a nascent technology, the adoption of vendor-independent SSI standards into existing software landscapes is at an early stage. Conventional enterprise-grade IAM implementations and cloud-based Identity Providers rely on widely established pre-SSI standards, and both will not be replaced by SSI offerings in the next few years. The contribution of this paper is an analysis of patterns and products to bridge unmodified pre-SSI applications and conventional IAM with SSI implementations. Our analysis covers 40+ SSI implementations and major authentication protocols such as OpenID Connect and LDAP.

Keywords: SSI; Self-Sovereign Identity; DID; Decentralized Identifiers; VC; Verifiable Credentials; IAM; Integration; Interoperability; Protocol; OIDC; OpenID Connect; OAuth; SAML; LDAP; X.509 Client Certificates; Kerberos; Active Directory; ADFS

1 Introduction and Problem Statement

Traditional implementations of Identity and Access Management (IAM) in enterprises include products such as Microsoft Active Directory or RedHat KeyCloak, and protocols such as OpenID Connect (OIDC), SAML 2.0, and LDAP. More recently, hosted IAM implementations from cloud-based vendors such as Auth0, Azure or AWS have gained popularity. Still, many companies opt for a hybrid landscape, combining on-premise IAM core deployments with cloud-based applications. Security-wise, PKI (Public Key Infrastructure) standards such as X.509 (incl. certificates for authentication and other tasks) ensure both interoperability and centralized governance, using Certificate Authorities (CAs) and Certificate Revocation Lists (CRLs).

At the same time, a new paradigm has gained momentum *outside* enterprise-internal setups: Self-Sovereign Identity (SSI) [PR21] is a term describing user-centered, user-administered decentralized approach and role model. SSI goes beyond authentication by establishing formats and tools for Verifiable Credentials (VCs). The prevailing implementation of SSI is based on W3C-issued standards for VCs [W3b] and Decentralized Identifiers (DIDs) [W3a].

[1] DB Systel GmbH, Jürgen-Ponto-Platz, 60329 Frankfurt, Germany michael.kuperberg@deutschebahn.com
[2] Institute for Internet Security, Westfälische Hochschule, Germany, and Service-centric Networking, Technische Universität Berlin, Germany, klemens@internet-sicherheit.de

doi:18.18420/OID2022-04

By design, the W3C standards for SSI are substantially different from conventional enterprise-focused IAM and PKI. Consequently, enabling SSI without rewriting existing applications requires additional integration efforts to integrate SSI into the enterprise world and conventional IAM protocols. Furthermore, lack of SSI support at the level of operating systems and web browsers means that additional software has been built for administering DIDs and VCs on devices, resulting in software-based *SSI wallets* such as Lissi-Wallet [Li]. OS-level support or even direct HW support for DIDs and VCs may arrive in the future.

The contribution of this paper is an analysis of solutions which enable the integration of SSI into IAM infrastructures for human users, both for cloud/internet applications and conventional/legacy software. To structure the analysis, we define SSI integration patterns, visualize them and illustrate their impact on the conventional IAM architectures. The analysis of specific products is performed using publicly available information, i.e. the identified software is not subjected to deployment, pilots, assessments, or security analysis.

The paper is structured as follows: Sec. 2 contains the foundations and Sec. 3 presents related work. We define the criteria and the methodology for the evaluation (Sec. 4.1), categorize the 40+ SSI solutions to filter out those which we found to not offer any integration with conventional IAM protocols (Sec. 4.2), and describe the remaining seven products in more detail (Sec. 4.3). Sec. 4.4 presents the comparison results, and Sec. 5 concludes.

2 Foundations

Different architectures and protocols are used for IAM, and new ones are introduced steadily. Yet there is a dominant pattern found in modern web-based applications over the Internet: end users often have the opportunity to login using an account hosted by a separate *Identity Provider* (IdP). The IdP is often owned by a different company, e.g. Google or Facebook, which has many users and mines their data. The resulting "social login" is the public internet equivalent of intra-company Single Sign-On as one IdP can be used across several *Service Providers* (SPs). A the same time, one SP can support a choice of several IdPs.

The SP-IdP pattern is illustrated in Fig. 1 on the left. Authorization is not shown in Fig. 1 because it is implemented very differently depending on the use case: intra-company SSO often includes centralized authorization data, turning an IdP into an IAM system; but certain SPs may keep their authorization data internally as well. In public web applications delivered over the Internet, even when using social login, authorization is mostly kept within SPs. Still, authorization may rely on identity data (such as location, age, gender, etc.) that is stored by the IdP and can be passed to the SP over standardized formats, such as OAuth.

The central precondition for this traditional setup is trust. For web-based applications, trust is established through PKI, specifically through the "trust anchor" set of root CAs. Root CA certificates come pre-bundled with browser downloads or with operating systems. Root CAs issue certificates to intermediate CA, which in turn issue certificates to servers, websites, end users, executable code, etc. CRLs provide additional security.

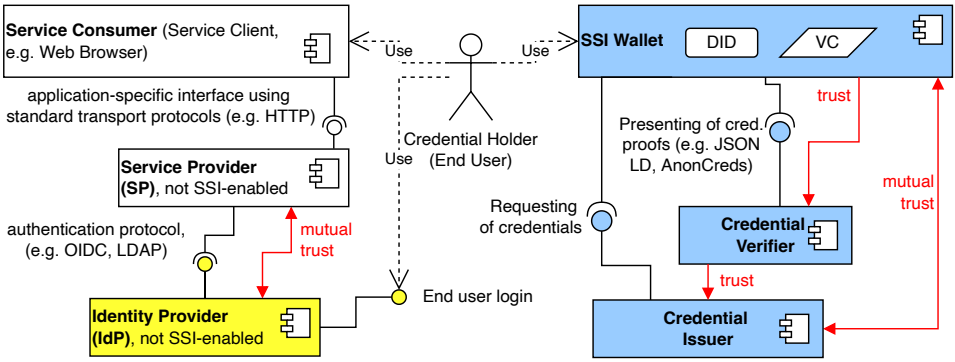


Fig. 1: Non-SSI authentication components (on the left) and SSI components (on the right). Authorization details, PKI infrastructure and SSI Verifiable Data Registry (Identifiers & Schemas) are not shown.

SSI addresses a well-known issue of social logins: while the use of centralized IdPs increases end users' comfort, it also makes end users (and SPs) more dependent on the centralized IdPs (SSI also claims additional benefits, such as machine-readable digitalization of real-life assertions, e.g. possession of driving licenses, etc.). SSI is shown in Fig. 1 on the right. It is also based on asymmetric cryptography but differs by focusing on a *decentralized*, user-controlled exchange of signed assertions. For the explanation of the Issuer, Verifier, Holder and the individual protocols, [PR21] provides a very good in-depth reference while also explaining how SSI fulfills *passwordless login* when implemented natively. Note that the trust relationships for SSI differ from those in the non-SSI case (see red arrows in Fig. 1).

To ensure SSI protocol-level interoperability and to speed up adoption, standards for defining and interchanging DIDs and VCs have been created. In particular, most SSI implementations are not based on hierarchic certification schemas but rather on the peer-to-peer and distributed "web of trust" without root or intermediate CAs. The decentralized approach is also where SSI needs to solve a scaling problem: instead of *limited-scale* "trust anchor" of a few root CAs, all three trust relationships on the right of Fig. 1 must be established for each new wallet holder and/or each new DID - but also for each Credential Verifier and each Credential Issuer. There is no universal solution for this issue, yet this issue is outside our scope.

While the adoption of the existing vendor-independent SSI standards is trying to gain a foothold in the enterprise world (and on the public internet), work on integrating SSI into enterprise environments and landscapes is also far from being completed or standardized. Within enterprise settings (where some services and applications are internal to a company), delegated authentication is often implemented with company-own deployments of identity providers (e.g. Active Directory or RedHat Keycloak). Company-own SSO often extends to both web applications and "traditional" rich clients; it uses such protocols as SAML or Kerberos, but also OAuth and OIDC. The resulting implementation of IAM often exposes

the LDAP protocol interfaces to connect third-party applications to IAM - especially when generic authorization support is needed additionally.

Often, the pre-existing software SP cannot be modified or replaced to support SSI protocols and standards - and even if it could, a co-existence of SSI and non-SSI-based-IAM may require a complex synchronization (e.g. user lockout in IAM must be mirrored in the SSI terms). Also, it is desirable that any changes to the pre-existing IAM solution must be as backward compatible as possible. Consequently, we identified the two following solution patterns that logically focus on IAM functionality:

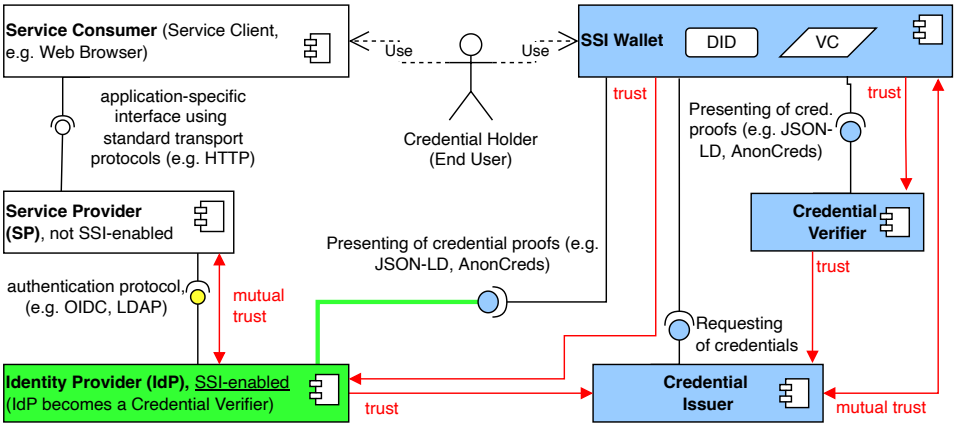


Fig. 2: Pattern A: IdP modified to support SSI protocol(s) for direct interaction with SSI Credential Holders. IdP assumes the role of Credential Verifier; the original Verifier may remain or be removed. Other components remain unchanged. Again, PKI and Verifiable Data Registry are not shown.

- Pattern A: use an SSI-enabled IdP (we illustrate this in Fig. 2) which offers conventional, non-SSI interfaces to the SP but functions as Credential Verifier towards the SSI roles (Holder and Issuer)
- Pattern B: augment the Credential Verifier (we illustrate this in Fig. 3) which provides non-SSI authentication interfaces to the IdP (for "authentication delegation") while leaving the SP *and the IdP* unchanged

Pattern B in Fig. 3 can also be varied into a third pattern, Pattern C, by introducing an *additional* component ("identity broker" or "bridge") between IdP and Verifier. The difference from Pattern C to Fig. 3 (with Pattern B) is that the Verifier remains unchanged while the additional bridge translates between the SSI and non-SSI protocols and data.

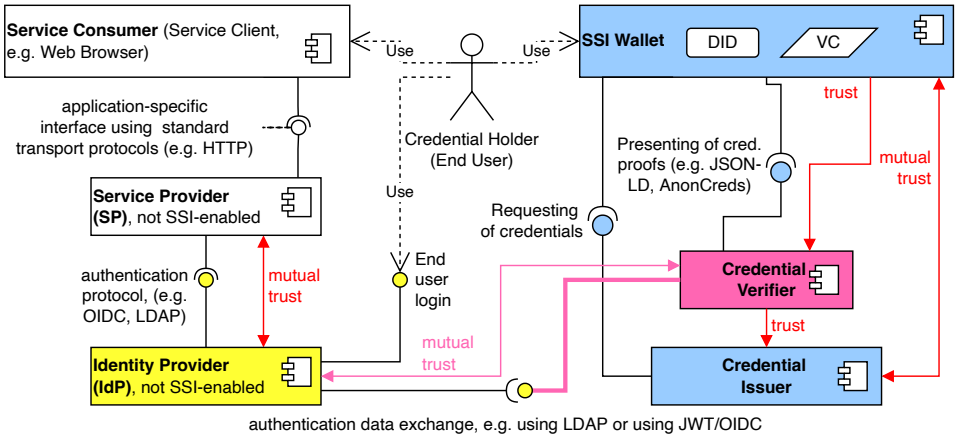


Fig. 3: Pattern B: Credential Verifier modified to support not-SSI protocol(s) for authentication delegation from IdP to Verifier. Pre-existing IdP interfaces are used, and the IdP implementation is not modified. Other components remain unchanged; PKI and Verifiable Data Registry are not shown.

3 Related Work

Several approaches that integrate SSI in/with established non-SSI IAM protocols have already been described. While we analyze individual implementations later in Sec. 4, this section also includes publications describing *comparisons* and surveys of solutions, as well as on theoretical proposals. Among overview-type survey publications such as [ČT21; FCA19; KP; Se21], we haven't found a criteria-driven comparison of SSI products/solutions covering support of traditional authentication/authorization protocols.

Published concepts that have no *publicly available* implementation include [HK20] and [Yi]. Yildiz et al. [Yi] design and implement a prototypical bridge between SSI authentication and SAML-based IdP-to-SP integration. They develop a hybrid solution that switches from username/password to login via VC with minimal authentication flow changes. VaultPoint, the system developed by Hong and Kim [HK20], complies with OAuth2 and combines SSI with smart contracts deployed on Ethereum. The smart contracts allow users to perform authentication and authorization using their own devices. However, the smart contracts store personally identifiable data which cannot be deleted - an approach that is not compliant with the EU GDPR. Thus, we will not consider VaultPoint in the evaluation described below.

Specifications without implementations include [Sa] and [Te]. Sabadello et al. [Sa] describe in their DID Auth document an approach of authentication with a focus on DIDs. In total, the authors present 10 different architectures to complete DID-based authentication by enabling the identity owner to prove control of a DID to a relying party. The OIDC specification "OpenID Connect for Verifiable Presentations" published by Terbu et al. [Te] extends OIDC to support presentation claims over VCs. This allows (1) existing OIDC Relying Parties to

accept VCs as claim sources and (2) new applications built with VCs to use OIDC as an integration layer for credential holders. In addition, the specification enables VC interchange in conjunction with Self-Issued OpenID[3] providers and traditional OICD providers.

Grüner et al. [GMM21] conduct a comparative evaluation of interoperability and portability of schemes for SSI Identity Management Systems (IdMS). As part of their research, the authors analyze the interaction of the user and SP with the IdP. They list OIDC, OAuth2, and SAML 2.0 as traditional IdM compared to DIDAuth and DID as SSI samples. However, their research doesn't evaluate the integration of SSI into traditional IAM protocols.

In constrast to our paper, these eight publications [ČT21; FCA19; GMM21; HK20; KP; Sa; Se21; Te; Yi] do not contain a systematic analysis or comparison of existing offerings.

4 Comparison

4.1 Comparison Methodology Criteria

We evaluate the products based on declared support for six conventional IAM protocols. Comparisons of auth* protocols are frequently made in Internet discussion groups, but we did not identify a *peer-reviewed* publication that would analyze usage frequency of auth* protocols in products, or even compare/rank them. Therefore, we chose the protocols based on our industry experience. The first three (OIDC, SAML 2.0, LDAP) are critically important for a product's relevance and adoption in enterprise environment with preexisting and legacy software, although the specific needs vary in each setting. The remaining three (X.509, Kerberos, AD-native protocols) are less frequent and thus are rather optional. Still, a product supporting one or several of those will be more useful in enterprise settings, in particular where device management is in place (rolling client X.509 for authentication without passwords) or where non-web applications ("fat clients") are in wide use. It should be noted that the protocols we have chosen are neither mutually replaceable nor universal (e.g. LDAP supports the querying of group membership whereas OIDC does not, being restricted to *basic profile infos* [Op]). The main aspects of the individual protocols are as follows:

1. OIDC (an abbreviation for OpenID Connect; it runs on top of OAuth 2.0, resulting in a combination of authentication and authorization), since OIDC is a major integration protocol for web applications, especially on the public internet
2. SAML 2.0 (authentication and authorization), since this is a major SSO protocol for applications in enterprise environments
3. LDAP (incl. LDAPS) (authentication and authorization), since this is the protocol commonly used for legacy centralized IAM in enterprise environments

[3] https://openid.bitbucket.io/connect/openid-connect-self-issued-v2-1_0.html

4. X.509 client certificates (authentication and authorization)
5. Kerberos (only authentication)
6. Active Directory native procotols[4] (authentication and authorization)

As for the comparison itself, we do not perform any tests to verify that a product's advertised features are indeed implemented, and adhere to standards. In other words, we rely on the vendor-provided public information (yet we invested considerable time to clarify incomplete and conflicting statements, and to have the results published). We do not execute black-box or white-box compatibility/functionality tests or even a proper implementation audit.

Note that the interoperation between SSI and established auth* protocols does not cover other essential operational concerns, especially those common to enterprise settings: lifecycle management, compliance, security management, reporting, etc. Also, note that in the comparison below, we have only included SSI solutions which adhere to the W3C standard for DIDs and VCs. There exist further SSI solutions which employ custom protocols (incl. non-disclosed protocols), but we have decided not to include them into the paper's scope because IAM is all about interoperability, exchangeability, and proven standards.

4.2 Filtering out SSI Solutions which do not meet any of the Comparison Criteria

We have studied publicly available information and documentation of 40+ active solutions. Of these, seven solutions claim out-of-the-box support for at lease one of the procotols in Sec. 4.1: (1) SSI Preview in Azure ID, (2) MATTR OIDC Bridge (3) OpenID-SSI_Login, (4) esatus SOWL, (5) Spherity, (6) SSI4A, and (7) VC-OAuthN OIDC.

For the remaining 30+ solutions, we did not identify support for *any* of the conventional IAM protocols that form our six evaluation criteria (cf. Sec. 4.1): (8) Aloaha, (9) Bitnation, (10) Blockchain Helix, (11) cheqd, (12) DigiME, (13) DockIO, (14) Eddits, (15) Element [Mi20b], (16) ESSIF, (17) evan.network, (18) Evernym, (19) Gemalto's SSI effort [Th], (20) "IBM Verify Credentials" (21) ID2020, (22) Identity.com, (23) Idento.one (24) Jolocom, (25) Namecoin, (26) Peaq, (27) selfDID, (28) SelfKey, (29) Seraph ID, (30) Shocard, (31) Sovrin, (32) Serto.id and (33) veramo (two projects created when uPort was split in 2021), (34) SSIBAC, (35) Trinsic (which offers integration with Zapier, but no conventional IAM protocols out-of-the-box), (36) TrustCerts, (37) Veres.one.

Additionally, we found that >10 offerings that used to be active (and have/had websites) appear frozen/abandoned as of Jan. 2022, e.g. Abacus, Block.id, Ethense, FinID, KYC.legal, MHMD, PeerMountain, Persona, Proof/Tierion, Protea, SpidChain, Tenz-ID, etc.

[4] The AD product supports not just LDAP, but also the protocols from `https://docs.microsoft.com/en-us/openspecs/windows_protocols/ms-adod/5ff67bf4-c145-48cb-89cd-4f5482d94664`, such as SAMR/SAMS; the ADFS (Active Directory Federation Services) is an optional add-on that is needed to bring OIDC/OAuth support to AD. Azure Active Directory is a cloud-only offering related to but distinct from AD.

4.3 Compared Products, Implementations, Standards and Initiatives

In this section, we discuss the remaining seven candidate offerings in alphabetical order and summarize our findings in Sec. 4.4 incl. Table 1.

1. *Azure AD verifiable credentials* [Mi20a; Mi21b] is Microsoft's SSI implementation preview that builds on the cloud version of Active Directory (AD); their SSI strategy is found in [Mi21a]. The preview corresponds to Pattern A (cf. Fig. 2). It also partners with multiple identity verification providers to connect the virtual world to the physical world. Concerning adoption, the only mention (as of January 2022) is "the National Health Service (NHS) in the UK". There is no information on when the implementation will leave the preview status and become generally available ("GA").

2. MATTR offers commercial solutions for OIDC-enabled Credential Issuers and Credential Verifiers in JSON-LD format. MATTR OIDC Bridge [MA] is a closed-source extension to MATTR Core with OIDC. The Bridge defines how an OIDC IdP can be extended to support SSI-based authentication leveraging DIDs and VCs and corresponds to our Pattern C (see Sec. 2). The primary function of the Bridge is mapping the presented Credentials claims in JSON-LD format to the OIDC format. At the time of writing this paper, MATTR and the OIDC Bridge only support schema.org for publishing custom data vocabulary; its marked adoption is unclear. However, MATTR stands out by providing pricing information on its public homepage.

3. OpenID-SSI_Login is a prototype described in [Lu20a] by Lux et al., and the source code is open [TU20]. The authors integrate SSI into OIDC: they extend a preexisting IdP and replace the required attributes (within the OICD standard set of claims) with SSI VCs. Thus, OpenID-SSI_Login follows our Pattern A (cf. Fig. 2). [Lu20a] reports that the prototype has been tested with Sovrin and with Hyperledger Indy. No information about the adoption or next releases of OpenID-SSI_Login is available.

4. SOWL [AG] functions as an *Identity Provider* by exposing OIDC, OAuth, SAML 2.0, LDAP, and similar protocol interfaces while maintaining SSI credentials internally. Thus, SOWL follows our Pattern A (cf. Fig. 2) and provides authorization support as well (e.g. over LDAP). SOWL is a closed-source commercial solution and the license fees for the server-side components are fixed by negotiation (the website does not provide any pricing). No information about SOWL's market adoption is available.

5. Spherity Digital Identity Management Toolkit [Sp; St] is a closed-source, commercial SSI implementation targeting IDs for both humans and things (IoT). The Toolkit corresponds to our Patterns A and B (cf. Fig. 2). The toolkit is accompanied by a server-based "Cloud Identity Wallet" (also offered as a SaaS), with an API to connect applications to it and an SDK for integration into mobile apps. In [St], the integration of the Cloud Wallet into IAM landscapes (using LDAP, SAML, OIDC etc.) is announced. The Spherity offerings are marked as General Availability (GA) and pricing is subject to negotiations; no information about market adoption is available.

6. SSI4A [Me] ("SSI for All") is a research project completed in 2019 with no further development since then. The architecture of SSI4A is described in a scientific paper [GMM19] and matches our Pattern A (cf. Fig. 2). The prototype supports uPort and Jolocom (see Sec. 4.2) as SSI solutions. The website says that the users "can obtain attestations about their email address" and names a university portal as the single "integrated application to provide SSI authentication via SSI4A". The source code is not open-source but still publicly available, i.e under a "view only" license.

7. VC-OAuthN OIDC [BC] is another open-source, research-grade project concerned with achieving VC-based authentication using OpenID Connect. As of January 2022, it sees active development and reports tests compatibility with the VON network implementation, using the standardized DIDComm protocol [De] for the messaging between the OpenID Provider and the Identity Holder. The implementation corresponds to our Pattern A (cf. Fig. 2 in Sec. 2). The documentation explains the rationale, architecture and implementation very well. As in [Lu20b], the attributes for the ID token are extracted from VCs provided by the Identity Holder.

4.4 Comparison Results

	Availability; License	OIDC and/or OAuth	SAML 2.0	LDAP	*X.509 client certif.*	*Kerberos*	*AD native*
Azure AD SSI Preview	Preview, cloud-only; commercial	**both**	*no*	*no*	*no*	*no*	**yes**
MATTR OIDC Bridge	GA, cloud-only; commercial	**both**	*no*	*no*	*no*	*no*	*no*
OpenID-SSI_Login	Prototype; open source (ASL 2.0)	**both**	*no*	*no*	*no*	*no*	*no*
SOWL	GA; commercial	**both**	**yes**	**yes**	*no*	*no*	**yes**
Spherity	GA, commercial	**OAuth**	*no*	*no*	*no*	*no*	*no*
SSI4A	Prototype; "read-only" code license	**both**	*no*	*no*	*no*	*no*	*no*
VC-OAuthN OIDC	Prototype; open source (ASL 2.0)	**both**	*no*	*no*	*no*	*no*	*no*

Tab. 1: Native out-of-the-box support of six conventional IAM protocols to connect *applications* to identity providers (see criteria in Sec. 4.1), for the SSI implementations described in Sec. 4.3. Note that Spherity additionally envisions connecting its *Cloud Wallet* to IdPs using OIDC, SAML 2.0 etc.

Table 1 summarizes our findings and we can conclude that *as of January 2022*, support for conventional IAM protocols varies significantly, with OIDC being the most widely supported one, and no Kerberos support. As we did not perform any tests (performance, compatibility, security) and further factors (costs, support, stability/SLAs, etc.) are not considered at this stage, the results do not allow to rank individual offerings or to compare them to each other. It is noteworthy that none of the four commercial offerings is open-source.

5 Conclusions and Future Work

In this paper, we have addressed a key aspect of SSI adoption: which tools and frameworks can support the integration of pre-existing, unmodified applications with SSI concepts and protocols? To start with, we described architectural patterns that can be used for such an integration, by augmenting conventional IAM architectures with additional capabilities and components. Then, we defined a set of protocols (OIDC, SAML 2.0, LDAP, and three others) as criteria for comparing offerings, based on publicly available information.

Of the analyzed 40+ offerings, only seven provide the necessary capabilities by implementing at least one of the necessary protocols. Of these seven, three are research-grade projects, and only one of these three is seeing further development. The remaining four include one preview-status implementation from a major cloud vendor and three GA offerings.

Our research shows that the offerings work in significantly different ways and that there are rather few standardization attempts or best patterns for these aspects, i.e. beyond the SSI protocol level. While we do not recommend a specific product and do not rank the surveyed offerings, we have observed that it is hard to derive information from open documentation and that code examples or integration tutorials are relatively infrequent.

In our future work, we plan to create a reference architecture for integrating pre-SSI architectures with SSI concepts, including authorization aspects. We also plan to perform hands-on tests of the described products to investigate the performance and scalability of hybrid SSI-IAM solutions. Furthermore, we intend to research the intersection of the specifactions of W3C [W3a; W3b] and OIDC [Te] more closely.

Acknowledgements

Andreas Grüner, Axel Küpper, Mirko Mollik, Artur Philipp, and Sebastian Weidenbach supplied very helpful review findings and pointed us to several new initiatives and tools.

References

[AG] esatus AG: esatus SOWL, URL: https://esatus.com/solutions/self-sovereign-identity/sowl/?lang=en, visited on: 01/31/2022.

[BC] BCgov: Verifiable Credential Authentication with OpenID Connect (VC-AuthN OIDC), URL: https://github.com/bcgov/vc-authn-oidc, visited on: 01/31/2022.

[ČT21] Čučko, Š.; Turkanović, M.: Decentralized and Self-Sovereign Identity: Systematic Mapping Study. IEEE Access 9/1, pp. 139009–139027, 2021.

[De] Decentralized Identity Foundation, URL: https://identity.foundation/didcomm-messaging/spec/, visited on: 01/31/2022.

[FCA19] Ferdous, M. S.; Chowdhury, F.; Alassafi, M. O.: In Search of Self-Sovereign Identity Leveraging Blockchain Technology. IEEE Access 7/1, pp. 103059–103079, 2019.

[GMM19] Grüner, A.; Mühle, A.; Meinel, C.: An Integration Architecture to Enable Service Providers for Self-sovereign Identity. In: 2019 IEEE 18th Intl. Symposium on Network Computing and Applications (NCA). Pp. 1–5, Sept. 2019.

[GMM21] Grüner, A.; Mühle, A.; Meinel, C.: Analyzing Interoperability and Portability Concepts for Self-Sovereign Identity./, p. 11, 2021.

[HK20] Hong, S.; Kim, H.: VaultPoint: A Blockchain-Based SSI Model that Complies with OAuth 2.0. Electronics 9/8, 2020, ISSN: 2079-9292, URL: https://www.mdpi.com/2079-9292/9/8/1231.

[KP] Kaneriya, J.; Patel, H.: A Comparative Survey on Blockchain Based Self Sovereign Identity System. In: 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, pp. 1150–1155, URL: https://ieeexplore.ieee.org/document/9315899/.

[Li] LiSSI: The LiSSI wallet - The new solution for Identities. Digital, decentralised and self-sovereign. URL: https://lissi.id/mobile, visited on: 01/31/2022.

[Lu20a] Lux, Z. A.; Thatmann, D.; Zickau, S.; Beierle, F.: Distributed-Ledger-based Authentication with Decentralized Identifiers and Verifiable Credentials. In: BRAINS2020. Pp. 71–78, 2020.

[Lu20b] Lux, Z. A.; Thatmann, D.; Zickau, S.; Beierle, F.: Distributed-Ledger-based Authentication with Decentralized Identifiers and Verifiable Credentials, 2020, arXiv: 2006.04754 [cs.DC].

[MA] MATTR, URL: https://learn.mattr.global/docs/platform/extensions/oidc-bridge/overview, visited on: 01/31/2022.

[Me] Meinel, C.: SSI4A is a gateway to enable the easy use of Self-sovereign Identity (SSI) solutions and a seamless integration into applications, GitHub Repository: https://github.com/agruener2000/ssixa-core, URL: https://ssixa.de, visited on: 01/31/2022.

[Mi20a] Microsoft: Verify once, use everywhere - Join our list and we'll let you know when our Public Preview is ready. 2020, URL: https://didproject.azurewebsites.net, visited on: 12/30/2021.

[Mi20b] Misc.: Element - DID Method implementation using the Sidetree protocol on top of Ethereum and IPFS, https://github.com/decentralized-identity/element, 2020, URL: https://element-did.com, visited on: 01/31/2022.

[Mi21a] Microsoft: Own your digital identity - Discover decentralized identity, a new way to provide ownership of personal data. 2021, URL: https://www.microsoft.com/en-ww/security/business/identity-access-management/decentralized-identity-blockchain, visited on: 01/31/2022.

[Mi21b] Microsoft: Verify once, use everywhere - Use a line of code to verify any data about anyone, while protecting privacy. 2021, URL: https://www.microsoft.com/en-ww/security/business/identity-access-management/verifiable-credentials, visited on: 01/31/2022.

[Op] OpenID Foundation: OpenID Connect Specification, URL: https://openid.net/developers/specs/, visited on: 01/31/2022.

[PR21] Preukschat, A.; Reed, D.: Self-Sovereign Identity. Manning Publications, 2021.

[Sa] Sabadello, M.; Hartog, K. D.; Lundkvist, C.; Franz, C.; Elias, A.; Hughes, A.; Jordan, J.; Zagidulin, D.; Rusu, E.; Powers, A.; Callahan, J.; Andrieu, J., URL: https://github.com/WebOfTrustInfo/rwot6-santabarbara/blob/master/final-documents/did-auth.md, visited on: 01/31/2022.

[Se21] Sedlmeir, J.; Smethurst, R.; Rieger, A.; Fridgen, G.: Digital Identities and Verifiable Credentials. Business and Information Systems Engineering 63/5, pp. 603–613, 2021.

[Sp] Spherity: Spherity, URL: https://spherity.com, visited on: 01/31/2022.

[St] Stöcker, C., URL: https://medium.com/spherity/on-ssi-enabled-idp-solutions-d382abc4b433, visited on: 02/15/2022.

[Te] Terbu, O.; Lodderstedt, T.; Yasuda, K.; Lemmon, A.; Looker, T., URL: https://openid.net/specs/openid-connect-4-verifiable-presentations-1_0-07.html, visited on: 01/31/2022.

[Th] Thales: Self-sovereign identities at work - Digital identity 2.0, URL: https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/digital-identity, visited on: 12/30/2021.

[TU20] TU Berlin SNET Research Group: OpenID-SSI Login - A bridge between the OpenID-Connect and Self-Sovereign Identity / DIDComm World, 2020, URL: https://github.com/TU-Berlin-SNET/DIMS-openid-ssi_login.

[W3a] W3C: DID (Decentralized Identifier) Data Model and Generic Syntax 1.0, URL: https://github.com/WebOfTrustInfo/rebooting-the-web-of-trust-fall2016/blob/master/draft-documents/DID-Spec-Implementers-Draft-01.pdf, visited on: 01/31/2022.

[W3b] W3C: Verifiable Claims Data Model and Representations, URL: https://www.w3.org/TR/verifiable-claims-data-model/, visited on: 01/31/2022.

[Yi] Yildiz, H.; Ritter, C.; Nguyen, L. T.; Frech, B.; Martinez, M. M.; Kupper, A.: Connecting Self-Sovereign Identity with Federated and User-centric Identities via SAML Integration. In: 2021 IEEE Symposium on Computers and Communications (ISCC). Athens, Greece, pp. 1–7, ISBN: 978-1-66542-744-9.

eIDAS 2.0: Challenges, perspectives and proposals to avoid contradictions between eIDAS 2.0 and SSI

Steffen Schwalm[1], Daria Albrecht[2], Ignacio Alamillo[3]

Abstract: The proposal for review of the eIDAS Regulation from 2021 has opened strong expectations for a deep change in traditional identity models. The user-centric identity model proposed starts with the creation of European Digital Identity Wallets that will enable citizens' control over their data in identification and authentication processes without control by entities providing the identification services. Likewise, with the proposed legal rules for giving legal certainty to electronic ledgers and blockchains, [eIDAS2]opens possibilities to decentralization, especially for the provision and management of user's attributes. The implementation of qualified trust services for attestations or electronic ledgers limits decentralization by requirement of a trusted 3rd party. Standardization will be key in assuring interoperability at the EU level. What are the challenges and opportunities of eIDAS 2.0? And what are the main focuses and needs of (European) standardization? These and other questions will be analysed and discussed in the paper.

Keywords: eIDAS, SSI, self-sovereign identity, identity model, digital wallet, eID

1 Introduction

Unique identification of legal or natural entities as well as their objects – the basement for a digital identity – allows the verification of companies (Do they really exist?), the person acting for the company (Do they really exist?) and their authorization (Is Alice authorized to act for company A?).

Digital identities are currently typically issued by a centralized authority. Despite the widely used but privacy critical social identities, the main electronic identification means of natural entities are government eID issued by member states. While Italian, Danish or Estonian eID are widely used, although notified on different Level of Assurance, the utilization of German eID is still low. Especially in those countries where little us of the government eID is made, many other identification procedures such as BankID (identification by bank and typically one time bank transfer), video identification or fully automated identification always based on a government (mostly notified) eID became popular in the different industries e.g. Finance, Insurance, Health Care or Public Sector. Current government eID and private identification procedures are mainly focused on natural or legal entities. But digital identities contain much more, such as attributes and

[1] msg group, Robert-Bürkle-Str. 1, 85737 Ismaning, Germany
[2] msg group, Robert-Bürkle-Str. 1, 85737 Ismaning, Germany
[3] Universidad de Murcia, C. Campus Universitario, 11, 30100, Murcia, Spain

cbe doi:18.18420/oid2022-05

evidence related to natural or legal entities like vaccination passports, authorization (power of attorney) or diplomas. Those proofs are currently mostly represented by digital documents in pdf or equivalent typically presented via mail, portals etc.

In parallel decentralized digital ecosystems occurred in the context of emergence of distributed ledger technologies. DLT by its distributed design makes it easy to establish decentralized digital business models cross-industry and cross-country between. The technology gains it`s biggest added value in transactions between > 3 parties which don`t trust each other and so trust in a distributed network which is immutable by design [Wer18], [Ko21], [Tr20]. In the context of DLT and decentralized ecosystems also the new paradigm of self-sovereign-identities has to be mtioned. SSI promise identity owner full control over its identity and attributes [Allen]. All identity information is stored decentralized and only the holder should decide whom he`ll give access or transmit identification information. One main postulate is that in DLT based on SSI a trusted 3rd party is not necessary anymore since DLT is used as decentralized PKI and immutable by design – so SSI may be trustworthy by itself [Wer18], [Ko20]. ENISA mentioned in one of it`s last reports that some main initiatives e.g. the strategic Show Case projects in Germany[4], funded by Federal Ministry of Econonmy and Climate Protection use DLT as decentralized PKI and emphasized the privacy advantages due to selective disclosure and Zero Knowledge Proof-Mechanism [ENISA22]. According to ENISA the utilization of DLT may be a step to created trust in SSI.

Currently SSI lacks the legal trust because current [eIDAS1] mainly focused on government eID not integrating the new SSI-paradigm. With the eIDAS Bridge the EU just developed possible legal and technical solution to bridge centralized approach of [eIDAS1] referenced to government eID and (qualified) trust services with decentralized manner of DLT and possibly SSI [Al20]. Accelerated by success of DLT and developments like [EBSI] in Europe but also the limited utilization of existing (centralized) eID, the EU-Commission just revised eIDAS and proposed a re-engineered regulation in June 2021 – recognizing decentralization on one hand and requirement of legal trust on the other one. This paper specifically focuses on whether the [eIDAS2] is complementary or contradictory to the Self-Sovereign-Identity (SSI) concept [Allen], how it may solve the challenge of legal trust in DLT and/or SSI and which challenges and chances the new version of eIDAS offers in respect to the digital identity models in Europe. In first step the main changes of eIDAS 2.0 will be described. Based on the paper discusses possible issue and contradictions between eIDAS 2.0 and SSI. The discussion focus on comparison of EU digital Wallet and the SSI-Principles, the chances and limits of decentralization in eIDAS 2.0 and last but not least the role of DLT in context eIDAS 2.0. The paper finalizes with a perspective on how eIDAS 2.0 and foreseeable underpinning standards should focus on to establish trustworthy self-sovereign identity including legal compliance and trust.

[4] Schaufenster sichere digitale Identitäten

2 Main legal changes in proposed new eIDAS 2.0

2.1 Overview

In June 2021, the European Commission published the proposal on regulation amending [eIDAS1] from 2014 with the aim to establish a framework for a European Digital Identity or, in other words, [eIDAS2]. The main goal of the proposed update is not a replacement but further development of [eIDAS1] in the context of decentralization and the upcoming SSI-paradigm, on one hand, but also the critical assessment and identified areas for improvement in [eIDAS1], on the other hand. The main changes in [eIDAS2] refer to electronic identification. Concerning trust services, only some additional services related to electronic identification were added and some logical gaps where closed.

2.2 Main changes on electronic identification and European Digital Identity Wallet

The main changes in eIDAS [eIDAS2] on electronic identification cover following topics:

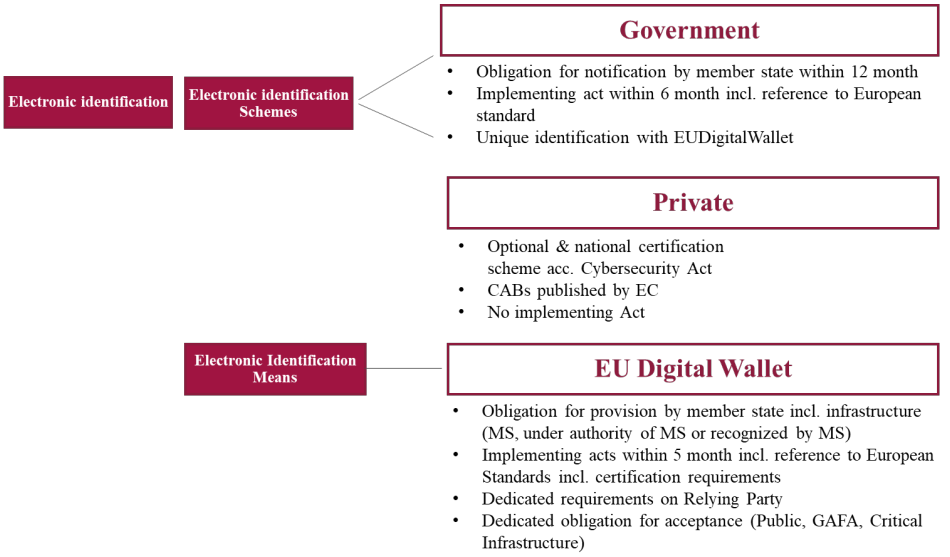


Figure 1: Proposal on [eIDAS2]: Main changes on electronic identification and European Digital Identity Wallet

[eIDAS2] proposal defines in Art. 6a the obligation for every member state to notify one identification within 12 months after the regulation will become applicable. Mandatory implementing acts referencing to European technical standardization shall be published

by European Commission within 6 months after new regulation is published. So, in comparison to [eIDAS1], the new regulation requires that at least one identity scheme from each member states shall be notified (Art. 10 and following). Considering that notification is one pre-condition for mutual recognition of identity, the obligation for notification can be mentioned as step forward in the wider utilization of eID in Europe. The presumable biggest change in [eIDAS2] is the requirements for every member state to provide a European Digital Identity Wallet to its natural entities. The Wallet could be published:

- By the member state
- Under authority of the member state
- Recognized by the member state

This makes also private wallet possible under the recognition of the member state. The European Digital Identity Wallet will contain the core identity currently covered by government eID as well as additional attributes or verifiable credentials acc. to W3C-standards so driver license, diplomas or the vaccine passport of its holder. This means that [eIDAS2] strictly follows the identity triangular of SSI. Every citizen will become a holder of a European Digital Identity Wallet and should become able to decide on his/her own, to whom he/she releases the identity information. The wallet consolidates core identity and attributes all together, but it must be taken into account that, due to cybersecurity reasons, the government eID will typically be stored on secure hardware components, normally a secure element or an e-sim, and only attributes will be stored in the wallet as a software component [Anke21], [TR03159]. In addition to that, the creation of (qualified) electronic signatures should be possible with the European Digital Identity Wallet. Technical details as well as security requirements for European Digital Identity Wallet will be defined in the ongoing European Standardization at ETSI and CEN. On the other hand, directly corresponding with the European Digital Identity Wallet, the new qualified attestation services acc. Art. 45a-e [eIDAS2] must be taken into account. Only qualified trust services providers offering such qualified attestation services are allowed to access European Digital Identity Wallet. Recognizing this close relationship between qualified attestation services and the wallet, [eIDAS2] contains the same requirements for mandatory implementing acts referring on European Standards for both – wallet and attestation service. Therefore, only the issuer into the European Digital Identity Wallet must be qualified attestation services. Consequently, [eIDAS2] crosses digital identity means and (qualified) trust services – they determine each other. To issue (qualified) attestation the trust service needs access to trust sources provided by member states e.g. public registries which requires their digital availability.

This means, in summary, that the new European Digital Identity Wallet will especially contain interface to qualified attestation service and relying party and shall fulfil LoA high acc. Art. 8 [eIDAS2]. The obligations on acceptance have to be emphasized: Not only public services, also any member of critical infrastructure entities (which means financial sector, utilities, health care etc.) as well as big internet companies such as

Google, Apple, Facebook or Amazon are forced to accept the European Digital Identity Wallet (Art. 12b). Similar to [eIDAS1], the member state is fully liable for providing the European Digital Identity Wallet as well as the eID-Scheme. A qualified attestation service takes the full liability risk like all QTSP, acc. Art. 13. This means that eIDAS limits the risk for users significantly in [eIDAS2] as well. The following picture, oriented on the Architecture Reference Framework [ARF 22] gives an overview on how the different parties may fit together:

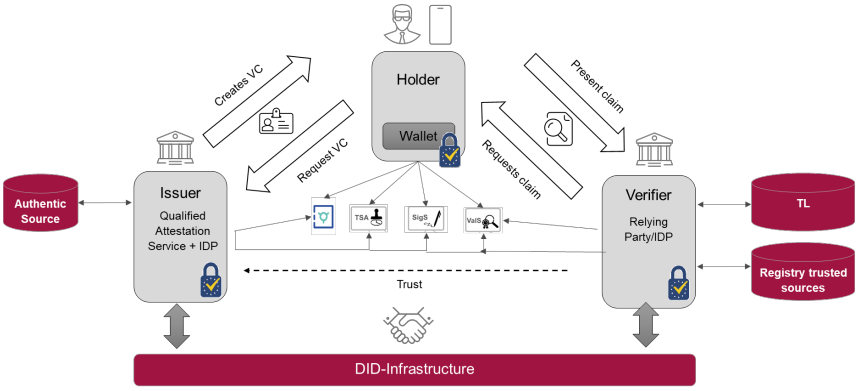


Figure 2: Possible interaction different parties in eIDAS 2.0

2.3 Main changes regarding (qualified) trust services and trust service providers

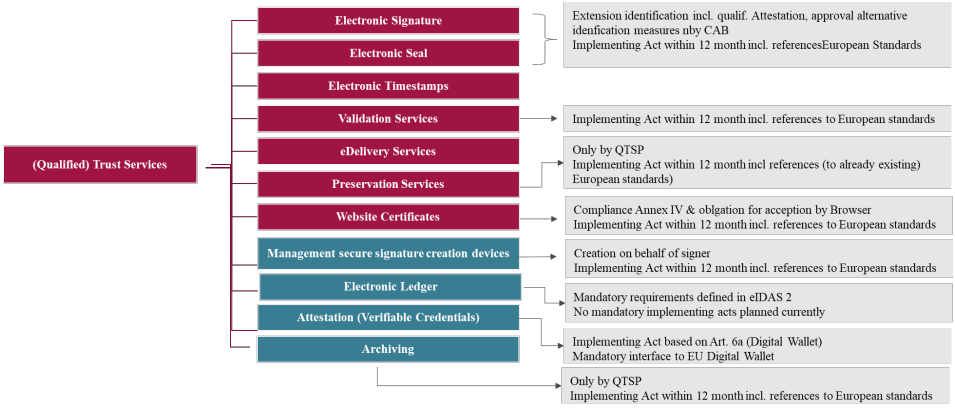


Figure 3: Proposal on eIDAS 2.0: Main changes regarding (qualified) trust services and trust service providers

In addition to the new qualified attestation services, [eIDAS2] also introduces the following new trust services for Electronic Ledger, so trust services for DLT (Art. 45g) This means that [eIDAS2] ensures trust in distributed ledger by (qualified) trust service

providers ensuring at least a minimum level of proven security and interoperability. Interestingly, [eIDAS2] does not contain the requirement of mandatory implementing acts referring to European standards only for the electronic ledger. Similar to [eIDAS1] all QTSP take the full liability risks (Art. 13) including the onus at their side – the trust chain is still the same [Ko21], [Zac20].

3 Possible issues and contradictions between [eIDAS2]and SSI

3.1 European Digital Identity Wallet and SSI-Principles

Since [eIDAS2] requires creation of (qualified) electronic signatures in combination with a wallet under the requirement of acceptance, the wallet might become the key tool for trustworthy digital transactions in regulated environments. Regarding the less success of only government issued eID in eIDAS 1.0 one main requirement for the success of EU-digital wallet is the distribution of providers. All possibilities given by eIDAS 2.0 so issued by member state, under authority of member state or recognized by member state should be used by all member states because the foreseeable competition of different public and private providers will ensure diversity according to different users' needs. The fact that EU Digital Wallet has to be issued to every legal or private entity in Europe eIDAS 2.0 achieves principle of representation and equity, the need for certification against European standards ensures it`s interoperability. With their wallet the user decides about relying party he wants to interact – the control of wallet it always on user`s side. eIDAS 2.0 contains obligation for acceptance but not utilization of wallet for the user and at same time opens the ecosystem for all interested parties as long as they fulfil the security requirements on e.g. trust services or relying parties [Al22]. With clear identification and authentication, the new regulation avoids a security findings and vulnerabilities like in German IDWallet where core identity information could be delivered to any unproven relying party without wither any identification nor authentication [FragSt21], [BSI19], [Ko20], [DINTS31648].

The table below gives an example how SSI-principles and [eIDAS2] may fit together:

SSI Principle	Fulfilment by eIDAS 2.0
Representation	Notified eID Scheme and European Digital Identity Wallet
Interoperability	Certified European Digital Identity Wallet, conformity assessed QTSP and notified eID as well as eIDAS nodes; Common European standards referenced by implementing acts

SSI Principle	Fulfilment by eIDAS 2.0
Decentralization	European Digital Identity Wallet and proven issuer as well as relying parties
Control and Agency	European Digital Identity Wallet, proven issuer and relying party
Participation	Only obligations for acceptance - no obligation to use the wallet nor the identities
Equity and Inclusion	Equal regulation for whole EU and EFTA
Useability, Accessibility and Consistency	Certified European Digital Identity Wallet and qualified trust service providers based on common European standards proved by accredited CAB
Portability	Any identities or attestation from European Digital Identity Wallet can be moved. Details should be defined in European standards
Security	State of the art security requirements defined in common European standards mentioned by implementing acts. Proved by CAB during certification of wallet, relying party or conformity assessment of QTSP. Trust provable via TrustList
Verifiability and authenticity	Verifiability and authenticity of attestations, signatures, seal, timestamps provable via (qualified) validation services, attestation services etc.
Privacy and minimal disclosure	Ensured by European Digital Identity Wallet and the fact that only holder decides which information he`ll provide but due to fact that relying parties are approved, the holder can really be sure to whom he/she will provide which information. Selective Disclosure and ZeroKnowledgeProof included
Transparency	European-wide regulation with common acts and mandatory European standards

SSI Principle	Fulfilment by eIDAS 2.0
	which are the basement for notification of eID-schemes, certification of European Digital Identity Wallet, relying parties, QTSP and all information published

Table 1: Possible match eIDAS 2.0 and SSI-Principles

3.2 Decentralization and its limits in eIDAS 2.0

[eIDAS2] defines the main legal framework for trustworthy digital transactions with centralized and decentralized digital identities and in the consequence a valid records management in Europe. The regulations take into account that SSI is not implemented on a green field but in an existing environment where centralized digital identities are established, widely used and, in regulated industries, fulfil the legal requirements [Ko18], [Ko20] [Anke21]. If SSI should be a sustainable alternative instead of centralized digital identities, legal compliance and trust are main pre-condition and trust given by notified eID-Scheme, certified EU-DigitalWallet and verifiable credentials by certified and supervised qualified attestation services which are fully liable. This means [eIDAS2] ensures a trustworthy decentralization with the entanglement of legal requirements in the law and its implementing act with mandatory European standardization. Clear and proven liability, security and interoperability of trust services and identity enable legal certainty of SSI with the disadvantage that a full decentralization with self-created credentials independent from any trusted 3rd party is not possible. In parallel, [eIDAS2] ensures with its mandatory implementing acts the achievement of SSI-principles on interoperability, security and so participation, equity and inclusion. The reason is that the implementing acts will reference common European standards for all member states and ensure same technical framework for each European Digital Identity Wallet and SSI in Europe in accordance with the SSI-Principle of representation [Ku20].

The fact that [eIDAS2] requires notification of government issued eID (or recognized/under authority of/by member state) as well as certification of private identification scheme by CAB – same with European Digital Identity Wallet the new regulation limits the decentralization of SSI because a trustworthy 3rd party is always necessary under eIDAS, but also to fulfil burden of proof in any regulated industry [We18], [DINTS31648]. However, this apparent disadvantage is one main added value of eIDAS 2.0, because for the first-time self-sovereign-identities gain legal trust and become usable in regulated environments with its needs for burden of proof and documentation requirements which must be made evident in non-repudiated manner against trusted 3rd parties. [eIDAS2] ensures a legally compliant verifiability and proven security and makes execution of SSI principles on security, authenticity and verifiability possible. Without legal compliance SSI would remain academic [Al20], [Sedl21],

[Ko20]. By ensuring trust in SSI, [eIDAS2] also limits its decentralization and therefore creates the boundaries of decentralization and SSI principle of participation evident. If there should be reliability that the legal or natural entity is really what it seems to be, a verified and secure identification is essential. This procedure, however, would set an entry requirement for the participation in the ecosystem.

3.3 DLT in the context of eIDAS 2.0

Basically [eIDAS2] is technology neutral. Neither for the (qualified) attestations, nor the identification scheme nor the identification means a concrete infrastructure is required. No DLT is mandatorily needed to implement Self-Sovereign-Identity. SSI is much more an identity and access management concept where on one hand the identity holder decides to whom he will give which part of his identity information and on the other hand does not have to give the full identity information in all cases but only the needed parts. Technically no DLT is mandatorily needed for SSI – the attestations may also be created in a centralized PKI which would recognize the fact that a centralized authority – the qualified attestation service issues the attestation based on (typically centralized trusted sources provided by member states) [Co20]. Nevertheless, some SSI proposals make use of functions supported by DLTs, such as DID-anchoring (of information of the qualified attribute attestations) or revocation information propagation [Sedl21], [Ku20].

DLT currently lacks a clear and legally compliant identification of parties taking part in the network, as well as unique evidence for authenticity and integrity of its transactions. Regarding the fact that DLT is immutable by design this main property is in contradiction to privacy law e.g. GDPR and its rights of the affected person (e.g. right for erasure, right for correction). Same with lack of standards for interoperable data exchange of on-chain data what limits the right for data portability according to GDPR [Ko20], [DINSPEC4997]. Similar vulnerabilities are the less long-term crypto stability, preservation of evidence and Proof of Existence which is critical for utilization in regulated environments with their often-complex documentation requirements, burden of proof until the end of the common decade long retention period [We18], [Sa17], [Ko21]. Without fulfilling basic criteria for trusted transactions and records management DLT is not feasible to be used in regulated environments [DINTS31648]. With QTSP for DLT the eIDAS 2 ensures legal trust in DLT because the QTSP will foreseeably act as de facto gatekeeper. The other advantage is that [eIDAS2] just solve the liability problem in DLT. According to Art. 13 eIDAS every QTSP is fully liable for its business. Since Art. 13 was not changed, this also applies to QTSP for Electronic Ledger and implies a Public or Private Permissioned DLT to ensure that there is always a provider operating and providing the DLT-network. With this approach [eIDAS2] ensures proven security in DLT. Because DLT might be used as decentralized PKI for SSI especially the EBSI it`s difficult to understand why the [eIDAS2] proposal does not contain the requirements for mandatory implementing acts referencing European Standards for QTSP for Electronic Ledger.

4 Perspectives of eIDAS 2.0 and necessary standardization

The proposal on new eIDAS-regulation proposes the first regulation on trustworthy self-sovereign-identities gaining legal trust and compliance. With the obligation for member states to provide one notified eID-Scheme but also European Digital Identity Wallet for their member states, the new eIDAS ensures a secure digital identity for each citizen. The close combination of wallet and (qualified) attestation services ensure legal trust not only in self-sovereign-identities and verifiable credentials but also actual data sovereignty and proven security for the user due the notification of eID-Scheme, certification of the wallet as well as certification of the qualified trust service provider. The risk for the user of a European digital identity is limited because member states and QTSP take the full risk for their schemes, European Digital Identity Wallet and attestation. It´s positive that [eIDAS2]is technology neutral and does not require DLT as infrastructure for SSI but also mentions QTSP for Electronic Ledger and, in this way, achieves proven security and trust for DLT. The extensive requirements on mandatory implementing acts linked to European standards enable the technical harmonization and limit national specifics. The creation of coherent and comprehensible European standardization framework gains as more importance as the standards will be referenced by the mainly mandatory implementing acts acc. to eIDAS 2.0 proposal. Against this background the standardization should especially focus on eID-schemes, EU-DigitalWallet and Attestation services first. Delegated authentication protocols like OIDC and OAuth2 are established and so interoperability is not a challenge currently [Hue19]. In W3C the work concerning DID-resolver is ongoing [Resolv] – a collaboration would be meaningful to identify relevant subjects for Europe and ensuring international feasibility of European SSI-standardization [Bast22]. The standardization may also focus on interoperability between centralized and decentralized digital identities to ensure comprehensive digital transactions notwithstanding if the natural or legal entity owns wallet or stored their identities at a centralized identity provider and only shares them with a relying party. Standardization supporting eIDAS 2.0 shall avoid reinventing the wheel. There are established and feasible standards e.g. for creation or preservation of signature, seal, timestamps; thus, only the gaps should be closed [ESI].

Currently, [eIDAS2] and related standardization mainly focus to store core identity information based on notified identity scheme on hardware of mobile devices and only the attestation in the wallet software itself [TR03159]. This means that core identity information of European citizens will be stored in non-European hardware whose specification are not disclosed or completely open source. Necessary European standards should focus on appropriate security measures for a fully hardened but also interoperable wallet which technical specifications and implementations are open source and therefore completely provable for 3rd parties [BSI19], [Al20], [Ko20], [Ko21]. The ongoing work on eIDAS Toolbox should consider this. It is also worth mentioning that some critical issues should be considered in the final version. For instance, a clearer statement for the certification and acceptance of wallets provided by private companies against the requirements of European Digital Identity Wallet to avoid restrictions on competition

should be provided. Since DLT may be used as infrastructure for SSI, there also should be mandatory implementing acts in eIDAS with references to European standards to ensure technical harmonization. Regarding the SSI-principles, it can be stated that there is no fundamental contradiction with the [eIDAS2]to be seen. The [eIDAS2] makes it possible for SSI-principles to become reality recognizing that decentralization has to be restrained to an acceptable level for achieving legal trust and data sovereignty. If a holder can`t trust an identity, issuer or verifier, he cannot act self-sovereign.

Bibliography

[Al20] Alamillo Dr. I-: SSI eIDAS Legal Report. How eIDAS can legally support digital identity and trustworthy DLT-based transactions in the Digital Single Market. Brussels 2020.

[Allen] Allen, C.: https://github.com/WebOfTrustInfo/self-sovereign-identity/blob/master/self-sovereign-identity-principles.md

[Anke21] Anke J. et al: Self-Sovereign Identity as the Basis for Universally Applicable Digital Identities. HMD 58, 247–270 (2021)

[ARF22] European Digital Identity Architecture and Reference Framework. Outline. 2022

[BSI19] Federal Office for Information Security (BSI): Towards Secure Blockchains. Concepts, Requirements, Assessments. 2019

[Co21] Corici A. et. al: Towards Interoperable Vaccination Certificate Services. 17th International Conference on Availability, Reliability and Security (ARES 2021) mGov4EU - Mobile Cross-Border Government Services for Europe 08 2021

[eIDAS1] Regulation (EU) No 910/2014 of the European Parliament and of the Council - of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. eIDAS, 2014.

[eIDAS2] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity {SEC(2021) 228 final} - {SWD(2021) 124 final} - {SWD(2021) 125 final}

[ENISA22] DIGITAL IDENTITY. Leveraging the Self-Sovereign Identity (SSI) Concept to Build Trust. European Union Agency for Cybersecurity.2022

[ET20b] ETSI Group Report 003. Permissioned Distributed Ledger (PDL). Application Scenarios

[ET21b] ETSI TS 103 732 V1.1.1. CYBER; Consumer Mobile Device Protection Profile.2021

[GDPR] Regulation (EU) 2016/ 679 of the European Parliament and of the Council - of 27 April 2016 - on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/ 46/ EC (General Data Protection Regulation). GDPR, 2016.

[Hue18] Huehnlein D.: Towards Universal Login. In: Roßnagel, H., Schunck, C. H.,

Mödersheim, S. & Hühnlein, D. (Hrsg.), Open Identity Summit 2020. Bonn: Gesellschaft für Informatik e.V.. (193-200). DOI: 10.18420/ois2020_18

[IS20a] ISO 22739:2020: Blockchain and distributed ledger technologies - Terminology, 2020.

[Ko20] Korte, U. et. al.: Criteria for trustworthy digital transactions – Blockchain/ DLT between eIDAS, GDPR, Data and Evidence Preservation. OpenIdentity Summit 2020. Lecture Notes in Informatics (LNI). Proceedings. Bonn 2020 S. 49-60

[Ko21] Korte, U. et. Al.: Records Management and Long-Term Preservation of Evidence in DLT. In: Roßnagel, H., Schunck, C. H. & Mödersheim, S. (Hrsg.), Open Identity Summit 2021. Bonn: Gesellschaft für Informatik e.V.. (131-142)

[Ku20] Kubach M. et. al.: Self-sovereign and Decentralized identity as the future of identity Management?. In: Roßnagel, H., Schunck, C. H., Mödersheim, S. & Hühnlein, D. (Hrsg.), Open Identity Summit 2020. Bonn: Gesellschaft für Informatik e.V.. (S. 35-47). DOI: 10.18420/ois2020_03

[Me80] Merkle, R. C.: Protocols for Public Key Cryptosystems. In: 1980 IEEE Symposium on Security and Privacy. IEEE, Oakland, CA, 1980. S. 122-134.

[OE17] OECD Digital Economy Outlook 2017. Organisation for Economic Co-operation and Development OECD, Paris, 2017.

[Resolv] https://github.com/decentralized-identity/universal-resolver

[Sedl21] Sedlmaier J., Smethurst R., Rieger A.: Digital Identities and Verifiable Credentials. Business & Information Systems Engineering 5/202

[TR03159-1] Technical Guideline TR-03159. Mobile Identities Part 1: Security Requirements for eIDAS LoA “substantial” Version 1.0 Draft 2 26. August 2019, Federal Office for Information Security. Bonn 2019

[Sa17] Sato, M.; Matsuo, S.'i.: Long-Term Public Blockchain: Resilience against Compromise of Underlying Cryptography. In ICCCN: 26th International Conference on Computer Communications and Networks (ICCCN) July 31-August 3, 2017, Vancouver, Canada. IEEE, Piscataway, NJ; 2017, S. 1–8

[Tr20] Treiber, K. Die Blockchain als zentrale Schnittstelle führt zur Verschmelzung unterschiedlicher Branchen. In: Die Zukunft ist dezentral. Frankfurt School of Finance & Management. Frankfurt 2020

[UN17] UN United Nations Commission on International Trade: UNCITRAL model law on electronic transferable records. United Nations, New York, 2017

[We18a] Weber, M. et al.: Records Management nach ISO 15489. Einführung und Anleitung. Beuth Verlag, Berlin, 2018.

[Wer] Werbach, K.: The Blockchain and the New architecture of Trust. Massachusetts Institute of Technology. 2018

[Zac20] Zaccharia et. al.: EU eIDAS-Regulation: Article-by-Article Commentary. Brussels 2020

Corporate Digital Responsibility and the current Corporate Social Responsibility standard: An analysis of applicability

K. Valerie Carl[1], Timothy Markus Christian Zilcher[1] and Oliver Hinz[1]

Abstract: Corporate Digital Responsibility (CDR) takes a key role in developing, deploying, and managing digital technologies, products, and services responsibly and ethically. New technologies offer new chances but also expose new threats, especially related to privacy and data security that managers need to cope with. CDR puts privacy and data security attempts in a broader context to provide a more holistic approach to Corporate Responsibilities and to strengthen consumer trust in corporate activities. However, managers still face a lack of CDR guidelines that support the implementation of CDR activities. Existing guidelines related to Corporate Responsibilities, like the ISO standard 26000, provide guidance on Corporate Social Responsibility (CSR) addressing socially responsible and sustainable behaviour. However, current standards do not cover CDR directly. As such, the purpose of this contribution is to evaluate the applicability of the existing CSR standard to CDR to pave the way for CDR standardization in the future.

Keywords: Corporate Digital Responsibility, Digitalization, Ethical Guidelines, Standardization.

1 Introduction

Advancements in digital technologies and an omnipresent digitalization of personal and professional lives allow for networks of devices that communicate via the Internet and perform fully automated tasks without any human interaction. The IoT emerged as an essential building block for many applications and systems. Despite the value creation and innovative technologies, consumers are especially concerned regarding the risks related to privacy and data security. These concerns even deepened due to data breaches and cyber-attacks [Vi19] and lead to a lack of trust. The possible hazard of privacy and security related issues can cause economic, ethical, or legal issues for consumers and firms alike [Ba19]. Prior research suggests that consumers' perception of their data security is critical for Internet or e-commerce technology adoption [Lu02]. To address these uncertaintys of digital technologies properly in a more comprehensive way and to support and to promote trust in corporate activities, a guiding framework that supports the ethical and responsible behavior in a digital world is necessary. In this context, the concept of Corporate Digital Responsibility (CDR) is gaining importance. CDR is closely related to the concept of Corporate Social Responsibility (CSR), both subsumized under the umbrella of Corporate Responsibilities. While corporate responsibilities for a company's impact on social and

[1] Goethe University Frankfurt/Main, Chair of Information Systems and Information Management, Theodor-W.-Adorno-Platz 4, D-60323 Frankfurt am Main, Germany, {kcarl, ohinz}@wiwi.uni-frankfurt.de, https://orcid.org/0000-0003-4655-1046 [Carl] / https://orcid.org/0000-0003-4757-0599 [Hinz]

cbe doi:18.18420/OID2022-06

economic aspects as well as their consequences is directed in the concept of CSR [MR02], CDR is a derivative of it with regard to digital issuses. CDR should give guidance to organizations how to handle potential negative consequences and how to use the opportunities of digitalization. CDR puts, inter alia, privacy and data security attempts in a broader context to provide a more holistic approach to corporate responsibilities and to strengthen consumer trust in corporate activities in a digitized world.

Yet, we can observe lively discussions in practice [e.g., He21], governance [e.g., Th17], and research [e.g., Lo21] that address the necessity and conceptualization of CDR. Thus, the theoretical debate on CDR evolves addressing the understanding of CDR and its scope [e.g., He21, Lo21, Mi21]. Nevertheless, practitioners still lack concrete guidance for the implementation of CDR activities compared to activities dedicated to the related concept of CSR (i.e., ISO 26000). While the ISO 26000 is a well-known guidance for the implementation of CSR in corporate practice there is still no equivalent standard covering CDR despite the already advanced digitalization. Considering the already omnipresent risks and challenges caused by the ongoing digitalization, CDR guidance as a standard is neededd. To the best of our knowledge, no previous research evaluated the applicability of the related CSR standard to the context of CDR. Thus, goal of this publication is to evaluate the need for adjustments to the current CSR standard, consequently the transformation into a superordinate Corporate Responsibility standard. Otherwise, there might be a need for the development of a separate standard to adequately address the topic of CDR.

CDR implements guidelines for the company's interaction with several stakeholder groups including, e.g., shareholders, employees, consumers, or the society itself [Lo21]. As various stakeholder groups do not always share the same interests, the guidelines CDR provides might not fit all stakeholder groups equally well. Consequently, the applicability evaluation of the ISO 26000 to the context of CDR focuses on business-to-consumer companies and their activities aiming at consumers. Hence, we pave the way for the standardization of CDR and provide guidance on the implementation of CDR in corporate practice. Following, the next section provides a definition of CSR and CDR. Section 3 then assesses the applicability of the CSR standard to CDR. Finally, we discuss the necessity of a specific CDR standard, this study's implications, and future research paths.

2 Corporate Social and Digital Responsibility

CSR and CDR are correlated and both part of Corporate Responsibilities, however research and practice should focus separately on CDR as is addresses the specific risk and challenges of the currently unfolding digitalization [e.g., Lo21]. CSR describes the responsibility of companies to align themselves with the expectations, goals, and values the society and stakeholders have. According to CSR, companies should take the economic, social, and ecological consequences of their actions into account [Ag11] and provide improvement to the quality of life by taking social responsibility. While organizations must follow legal obligations (i.e., regulations, laws) when offering products or services, CSR

intends them to align their behavior with ethically responsible conduct according to "what is right, just and fair, even when they are not obliged to by the legal framework" [MM07, p.337]. Consequently, voluntariness to improve the social well-being of stakeholders affected by the company's economic activities is at the core of CSR [Fr18]. CSR activities can also support corporate interests [Wi21] and are applicable to all sizes, industries and types of companies [Fr18]. The degree and type of CSR implementation varies and depends on the influence of, e.g., stakeholders, regulations, or applicable standards.

In 2010, the International Organization for Standardization published the ISO standard 26000 "Guidance on social responsibility". Despite its non-certifiability, this standard should serve as a guideline for organizations to act within the purpose of social responsibility and contribute to their sustainable development. These guidelines are applicable to organizations of all types. According to the ISO 26000, social responsibility should be an integral part of a company's core strategy. The central attribute of social responsibility is an organization's initiative to integrate social and environmental considerations into its decision-making process and to be accountable for the impact of its decisions and activities on society and the environment. Therefore, organizations have to identify stakeholders, and take their interest and expectation into account. The CSR standard aims to encourage organizations to go beyond compliance with the law, making it a fundamental duty of any organization and an essential part of its social responsibilities. Hence, ISO 26000 demands both transparent and ethical behavior that contributes to a sustainable development. In summary, firms should integrate CSR throughout the organization, their relationships, and regarding stakeholders' interests.

CDR is an independent concept that complements the principles of CSR by addressing the challenges and peculiarities of a digitized world [Lo21]. To this end, CDR puts associated risks of digital technologies, e.g., privacy and data security issues, in a broader context to provide a more holistic approach to Corporate Responsibilities and to strengthen consumer trust in corporate activities in a digitized world. Despite growing research efforts on CDR and its conceptualization [e.g., He21, Lo21], to the best of our knowledge, no previous research evaluated the applicability of the established CSR guidance to the context of CDR. Hence, this work lays the foundation for future research on CDR and a potential standardization of the concept by evaluating the status-quo of research concerning CDR and the applicability of a current standard related to Corporate Responsibilities (i.e., ISO 26000). In the past, the CDR debate brought up an approach consisting of eight dimensions to describe CDR and the concept's scope [Th17]: (i) access, (ii) dispute resolution and awareness, (iii) economic interests, (iv) education and awareness, (v) governance and participation, (vi) information and transparency, (vii) privacy and data security, and (viii) product safety and liability. Some (national) regulations, like the GDPR, already cover distinct sub-fields of CDR. Nevertheless, CDR activities exceed the legally binding (national) minimum requirements and rather describes the voluntary acceptance of additional responsibilities. In countries that already require compliance with high standards, e.g., with respect to privacy and data security, activities related to CDR require higher levels of voluntary responsibility than in countries with lower legal standards. Thus, the activities relatable to CDR vary between different countries as the legal requirements always specify

the minimum level. However, since the concept of CDR applies worldwide, the concept sets country-independent minimum standards, which may be tightened by national laws.

3 Applicability of the CSR standard to the context of CDR

Aim of this study is to evaluate whether the existing CSR standard, ISO 26000, is applicable to the context of CDR to form a superordinate Corporate Responsibility standard. Thus, we assess the current coverage of peculiarities of CDR, possible adjustments for adequate coverage, and the need for extensive additions to the CSR standard. This paves the way for the potential standardization of CDR and thus easy guidance for companies on how to implement CDR in practice. The evaluation grounds on eight dimensions of CDR [Th17] and associated sub-dimensions derived from theory and practice.

3.1 Access

Companies can support consumers' *access* to (basic) digital technologies, products, and services. Especially in a digitized world, *access* gains tremendous importance [DT21]. The CDR dimension *access* covers *physical* and *mental access*.

Physical access refers to the ability of individuals to physically access technologies. Hence, organizations can facilitate and enable safe access to digital technologies, products, and services. The CSR standard includes the demand for the dissemination of technologies, reasonably priced technologies, and preserving access in the event of a non-payment. However, issues not covered are specifics such as access to hardware, software, and Internet connection, which represent important parts of this sub-dimension.

Mental access includes corporate practices that increase consumers' prior knowledge and facilitate usage. The CSR standard requests firms to strengthen consumer knowledge generation. However, further explanations on consumers' prior knowledge and mental usage requirements are missing and would need a detailed representation.

Summing up, the CSR standard covers important areas of the dimension *access* but it lacks more extensive issues related to both sub-dimensions. For example, Internet access and ease of use should complement the existing standard. However, there is a possibility to widen the focus of the CSR standard to cover this dimension appropriately.

3.2 Dispute resolution and awareness

Dispute resolution and awareness presents another dimension of Corporate Responsibilities in the digital context. Companies can implement adequate mechanisms for resolving consumer complaints and potential redress for harm endured from transactions [CV16]. Correspondingly, CDR proposes an adequate way of *contact* regarding dispute resolution and redress for consumers, as well as a fair handling *process*.

Regarding the first *contact*, the CSR standard formulates the obligation to provide information to consumers in order to ensure a transparent and accessible process. Hence, the CSR standard contains an obligation for companies to enable easy accessibility of mechanisms, e.g., when complaints occur. Besides, the CSR standard specifies that dispute resolution should involve no or minimal costs for consumers and should proceed without waiving their rights. When consumers file for a complaint, the procedure should be simple and easily accessible (e.g., in terms of language, education, distance, physical and mental limitations). Consequently, this CDR sub-dimension is widely covered.

The second sub-dimension concerns the *process* of dispute resolution and dealing with complaints. The CSR standard requires the handling of a complaint according to a specified system and within a predictable period. Further should this *process* deviate from court procedures but the standard prohibits the circumvention of legal regulations. As such, it incorporates a fundamental principle of Alternative Dispute Resolution methods, which are particularly relevant in the digital environment. Consequently, the CSR standard widely covers this sub-dimension of CDR. However, consumer orientation lacks, which means to exhaust all options for solving the problems with consumers. Further, a focus on Online Dispute Resolution mechanisms is missing which seems particularly appropriate for dealing with complaints in the course of digital transactions.

Summing up, most of the requirements of *dispute resolution and awareness* are included in the CSR standard. However, peculiarities concerning the dispute handling *process* in the digital context are missing. Hence, there is a need to strongly develop and add to the existing CSR standard to cover this CDR dimension extensively.

3.3 Economic interests

The digital context can reinforce the mismatch between the interests of consumers and companies. Following the principles of CDR, firms protecting consumers' *economic interest* also protect their own future profits. Hence, this dimension covers, e.g., fair *competition* policies [e.g., Ra16], *pricing* [e.g., HHS11], or *interoperability* [e.g., Le13].

A functioning *competition* represents the idea of a competition, which enables the market mechanism to function optimally. The CSR standard notes the importance of functioning *competition* for innovation, cost efficiency, equity, economic growth, and standard of living. Therefore, companies should not engage in anti-competitive behavior to achieve an unfair competitive advantage and rather obey competitive law. In contrast, the CSR standard lacks mentioning monopolistic structures that are fundamentally opposed to functioning competition. With regard to digital markets, the reference to problems of market definition and the determination of market shares is missing. Besides, there is no legal consideration of network effects. However, network effects describe the changed value of a market or platform due to an additional market user and represent an important factor for companies, especially in the digital context [HOS20].

The second sub-dimension refers to *pricing*. In particular, price discrimination occurs

when firms charge different prices for the same product or service of the same quality [HHS11]. The CSR standard predicates that any distinction between people that results in an impairment of equal treatment should be avoided. This includes differentiating prices over time, between consumers, and between circumstances. However, there is a lack of more concrete evidence on the problem of price discrimination.

The sub-dimension *interoperability* refers to the ability of different systems, techniques, or organizations to work together, using a common technical standard. The current CSR standard does not covers this sub-dimension even though it contributes to avoid lock-in effects and therefore protect consumers' *economic interests*.

Concluding, even though the CSR standard widely covers the sub-dimension referring to *competition* except for monopolistic structures and network effects, the sub-dimensions of *pricing* and *interoperability* need broader coverage in the CSR standard. The implementation is partially lacking and without the integration of these sub-dimensions, an application of the current CSR standard to the concept of CDR is inconceivable.

3.4 Education and awareness

Education and awareness covers a broad range of application fields, e.g., consumer awareness regarding social, economic, and ecological consumption consequences [Th17]. The enhancement of more sophisticated digital technologies (e.g., blockchain technology) amplifies consumers' need for education. Hence, CDR encourages companies to *educate consumers* also raising their *awareness* for consumption consequences.

The first sub-dimensions refers to *consumer education* and should help consumers make informed consumption decisions. The CSR standard already requires firms to provide information to consumers to enable informed, responsible consumption decisions with knowledge of their rights and obligations. The CSR standard states that companies should foster *consumer education*, paying attention to the increased needs of disadvantaged (e.g., economic) consumer groups. *Consumer education* topics include, e.g., product safety, price and quality of products, and sustainability. Consequently, the CSR standard covers far-reaching parts of this sub-dimension. However, there is a lack of specified *consumer education* measures in different consumption stages.

Besides conventional *consumer education*, the second sub-dimension *awareness* aims to create consumer awareness of environmental, social, and economic consequences of consumption. The current CSR standard states that *awareness* is about paying attention to the impact of consumption decisions on other market participants, as well as on the common good, as opposed to simply pursuing individual interests.

These CDR sub-dimensions coincide almost completely with principles already covered by the CSR standard. However, there is need for alignment regarding the specification of the timing of *consumer education* measures (i.e., before, after, or during service).

3.5 Governance and participation

The dimension *governance and participation* entails adequate corporate participation mechanisms [Th17]. This CDR dimension consists of three sub-dimensions: *consumer feedback*, *consumer organization involvement*, and *product development*.

The sub-dimension *consumer feedback* covers requirements for companies to respond to the concerns expressed by consumers. These requirements include facilitating consumer-focused employee behavior, providing social skills training, and creating a pleasant corporate culture to encourage consumer feedback. However, a detailed guidance on the development of these capabilities and the design of feedback mechanisms lacks.

The CSR standard currently does not cover the sub-dimension related to *consumer organization involvement*. Consumer organizations advocate for the interests of private consumers and provide information on matters of private consumption. Thus, companies should survey representative groups of the community on business issues and participate in local forums, however a specific reference to consumer organizations lacks. Hence, there is a need for amendments advising firms to incorporate *consumer organizations*.

In almost the same manner, the CSR standard lacks references to the sub-dimension *product development*. *Product development* describes the process of creating a product, starting with the analysis of future trends and the incorporation of consumer needs up to the market launch. Although the CSR standard advices to include stakeholder groups' opinions, there is no mention of concrete participation in terms of *product development*. Further, the CSR standard could distinguish between opportunities for participation at different stages of the *product development* process. Besides, relevant methods related to crowdsourcing and similar digital possibilities lack.

In summary, the topic of *consumer feedback* is already part of the current CSR standard. Nevertheless, the CSR standard should exceed coverage related to employee behavior and a concretization of feedback mechanisms. The other two sub-dimensions are almost completely absent in the CSR standard. Accordingly, extensive additions would be necessary here so that the existing CSR standard also covers the CDR concept.

3.6 Information and transparency

Information and transparency are prerequisites for informed decision-making, therefore anticipated by consumers. This dimension addresses several application scenarios, e.g., the product scope, ecological footprint, or pricing [GGK10].

The provision of information to consumers about products, services, and measures taken by the company forms the sub-dimension *information*. The CSR standard already covers the required disclosure of truthful and unbiased information about products, services, terms and conditions, impacts on society, the economy, and the environment. Provided

information should be complete and understandable to enable informed consumption decisions. Yet, the CSR standard already covers an extensive part of the *information*-related CDR issues. Nevertheless, a concrete request for adequate information on data protection agreements and the link to specific issues in the digital context are missing.

The sub-dimension *transparency* fosters transparency of information. The CSR standard specifies that companies should disclose information about their decisions, behaviors, and potential social impacts. Hence, the CSR standard meets the fundamental requirement of *transparency*. However, it does not contain any more specific provisions, e.g., related to transparency of revenue generation—peculiarities of the digital context.

Concluding, the current CSR standard already addresses large parts of the *information and transparency* sub-dimensions, but the obligation to provide *information* specific to the digital context (e.g., digital business models, data protection agreements) lacks. To cover this CDR dimension a more far-reaching focus of the CSR standard is necessary.

3.7 Privacy and data security

Data privacy covers consumers' ability to control their data, whereas *data security* implies the protection of data against possible risks [BC11]. Hence, the concept of CDR fosters the protection of *privacy and data security* exceeding regulations voluntarily.

The responsible handling of data in terms of collecting and using data forms the sub-dimension of *privacy*. The CSR standard acknowledges the increased importance of personal data as a resource for digital products and services in the context of larger databases and digital communication technologies. It demands the consent of consumers at the time of data collection. Besides, the CSR standard covers the responsible and restricted use. To provide a more detailed evaluation of this CDR dimension, we employ an established framework, the six privacy protection goals, articulated by Hansen et al. [HJR15] to systematically assess the current coverage of the CSR standard regarding this CDR sub-dimension. While having slight overlap with the CSR standard, the six privacy protection goals address two further fields: *unlinkability* states, inter alia, that data protection relevant data cannot be linked across domains and *intervenability* describes the possibility of intervening in ongoing or planned data processing operations relevant to data protection. Both represent important privacy goals, which are worth considering including in the CDR standard. Hence, the CSR standard partly covers *data privacy.* Nevertheless, extensive amendments are needed to cover this topic in the digital context.

Besides, the CSR standard partly covers the second sub-dimension *data security* and requires appropriate security mechanisms, ensuring the protection of personal data. Yet, the CSR standard only implements basic aspects of *data security*. More far-reaching regulations such as potential physical security risks due to unauthorized access to personal data, security risks of data mining with regard to personal information, or recommendations for actions to avoid cyber-attacks remain unmentioned. Besides, a reference to other standards

such as the ISO 27000 series dealing more concretely with the subject of information security to cover the topic of CDR more appropriately lacks.

Hence, the current CSR standard covers *data privacy* for digital technologies superficially. However, while there are basic references to appropriate security mechanisms and privacy, concrete guidance for more *privacy and data security* lacks. Accordingly, far-reaching additions are necessary to represent this CDR dimension appropriately.

3.8 Product safety and liability

The dimension of *product safety and liability* addresses safe operations and the firm's liability in case of potential injuries (i.e., physical and mental harm). The digital context also makes it difficult to trace the damage back to its source [Sm17]. Consequently, the concept of CDR requires firms to protect consumers' *safety* from mental and physical risks also providing adequate *liability* and accountability in case of harm.

The first sub-dimension *product safety* deals with regulations on the safety of consumer products. The CSR standard states that companies should foster safe product operation and provide information on the safe use of products and services, both for proper and the expected improper use. Besides, the CSR standard requires firms to anticipate and remedy further potential risks and hazards. Nevertheless, the CSR standard has substantial gaps with regard to the *safety of digital products*. In particular, the CSR standard does not meet the challenges of *product safety* specific to the digital context like social media.

The second sub-dimension deals with *product liability*. According to the CSR standard, *product liability* refers to liability for compensations against the manufacturer for damage caused to the end user because of a defective product. Further, it also concerns the accountability for potential (human) rights violations. However, calls for accepting additional responsibilities in the sense of liability and specific requirements for internal company liability regulations are missing. In terms of digital products and services the consideration of *liability*-specific challenges such as intelligent algorithms, are not addressed. In addition to digital products and services, the CSR standard lacks *liability* regulations covering physical products sold via digital intermediary platforms.

Summing up, the CSR standard covers the requirement of *product safety* in general. Not covered is the *product safety* and *liability* in the specific digital context. Thus, there is a need for an extensive addition to the CSR standard or the establishment of an own standard to adequately address this CDR dimension and the peculiarities of the digital context.

4 Conclusion

Aim of this study is to evaluate whether the current standards and norms addressing Corporate Responsibilities are applicable to the context of CDR. Hence, we examined the

coverage of CDR dimensions by the ISO standard 26000, a standard providing guidance on CSR. ISO 26000 already addresses some of the CDR dimensions. However, ISO 26000 only focuses on corporate activities and consumer interaction in general. Still, the digital context poses peculiarities that exceed previous responsibilities [e.g., Lo21, Mi21] and the current CSR standard by far. Hence, Corporate Responsibilities within the digital context should receive an extended connotation that exceeds the understanding of CSR. Consequently, taking into account the detailed insights on the coverage of each of the eight CDR dimensions and the overall evaluation of applicability, this study suggests developing a CDR standard comparable to ISO 26000 that addresses the peculiarities and unique challenges of a digitized world. Alternatively, the standard 26000 would need to be extended extensively to include specific instructions to cover the digital context. However, since previous research recommends considering CDR and CSR as separate concepts [e.g., Lo21], one could better account for the specifics of the digital context by also developing two related, partially overlapping, but separate standards.

Consequently, this publication makes several theoretical contributions. Firstly, this study presents the concept of privacy and data security in the broader context of Corporate Responsibilities. Hence, this research adopts a broader approach to privacy and data security than numerous other research endeavors, motivating research on both as distinct topics. Secondly, this publication advances the current research base related to CDR by providing an in depth understanding of the scope of each CDR dimension. Research on CDR is still in its infancy [e.g., Lo21]. Therefore, it is of tremendous importance to develop consensus on the scope of CDR and its associated dimensions to pave the way for the standardization of the concept, thus providing guidance on the implementation of CDR. Hence, this publication intends to enhance the discourse on the understanding of CDR to support future standardization. Thirdly, this research contributes to the understanding of potential overlaps and divergences between the concepts of CSR and CDR.

From a practical point of view, this research translates the theoretically derived concept of CDR [e.g., Lo21] to corporate practice. This paper offers practitioners guidance for the implementation of CDR in practice, and hence how to address consumer trust issues related, e.g., to privacy and data security. Thus, the scope and applicability analyses serve as a first orientation for practitioners aiming at the implementation of CDR in their companies. Besides, this research should ignite the discussion on how to develop a standard that addresses CDR and its dimensions adequately. To this end, we provide a first assessment on the applicability of the current CSR standard, possible additions so that CDR can be covered, and the evaluation of a potential standard of its own. Based on the derived assessment, we suggest establishing a specific standard that addresses CDR and the peculiarities of the digital context. Hence, this research contributes to the solidification of CDR in corporate practice and a future standardization. Such a standardization can provide additional guidance for firms on how to implement CDR and corroborates a common view of the definition and conceptualization of CDR.

Despite our best efforts, this study is not without limitations. Firstly, this study focuses on one CSR standard, ISO 26000. To assess the applicability of a commonly used standard

in detail, this focus was necessary. Besides, there is no consensus in research nor practice on one framework describing the scope of CDR albeit sharing core values and a common understanding. However, we encourage future research to assess the applicability of other standards and norms related to Corporate Responsibility, also incorporating further CDR frameworks. Secondly, a focus on the interaction of firms with one specific stakeholder group was necessary. Nevertheless, CDR addresses several stakeholder groups like employees or society in general. Consequently, we motivate future research to address this gap and to assess the applicability of current standards on these aspects of CDR. The limitations again highlight the need for consensus on the nomenclature and scope of CDR. Despite its shortcomings, this research made a first step towards the standardization of CDR and thus supported the establishment of the concept in practice.

Acknowledgement

This work has been supported by the Hessian State Chancellery – Hessian Minister of Digital Strategy and Development under the promotional reference 6/493/71574093 (CDR-CAT).

Bibliography

[Ag11] Aguinis, H.: Organizational Responsibility: Doing Good and Doing Well. APA Handbook of Industrial and Organizational Psychology, Vol 3: Maintaining, Expanding, and Contracting the Organization. American Psychological Association, Washington, DC, US, pp. 855-879, 2011.

[Ba19] Baumann, A.; Haupt, J.; Gebert, F.; Lessmann, S.: The Price of Privacy: An Evaluation of the Economic Value of Collecting Clickstream Data. Business & Information Systems Engineering 61/4, pp. 413-431, 2019.

[BC11] Bélanger, F.; Crossler, R. E.: Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems. MIS Quarterly 35/4, pp. 1017-1041, 2011.

[CV16] Clifford, D.; Van Der Sype, Y. S.: Online Dispute Resolution: Settling Data Protection Disputes in a Digital World of Customers. Computer Law & Security Review 32/2, pp. 272-285, 2016.

[DT21] Díaz Andrade, A.; Techatassanasoontorn, A. A.: Digital Enforcement: Rethinking the Pursuit of a Digitally-Enabled Society. Information Systems Journal 31/1, pp. 184-197, 2021.

[Fr18] Frederick, W. C.: Corporate Social Responsibility: From Founders to Millennials. Corporate Social Responsibility, Business and Society 360. Vol. 2. Emerald Publishing Limited, Bingley, pp. 3-38, 2018.

[GGK10] Granados, N.; Gupta, A.; Kauffman, R. J.: Research Commentary—Information Transparency in Business-to-Consumer Markets: Concepts, Framework, and Research Agenda. Information Systems Research 21/2, pp. 207-226, 2010.

[He21] Herden, C.; Alliu, E.; Cakici, A.; Cormier, T.; Deguelle, C.; Gambhir, S.; Griffiths, C.; Gupta, S.; et al.: "Corporate Digital Responsibility": New Corporate Responsibilities in the Digital Age. Sustainability Management Forum 29, pp. 13-29, 2021.

[HHS11] Hinz, O.; Hann, I.-H.; Spann, M.: Price Discrimination in E-Commerce? An Examination of Dynamic Pricing in Name-Your-Own Price Markets. MIS Quarterly 35/1, pp. 81-98, 2011.

[HJR15] Hansen, M.; Jensen, M.; Rost, M.: Protection Goals for Privacy Engineering. 2015 IEEE Security and Privacy Workshops. IEEE Computer Society, pp. 159-166.

[HOS20] Hinz, O.; Otter, T.; Skiera, B.: Estimating Network Effects in Two-Sided Markets. Journal of Management Information Systems 37/1, pp. 12-38, 2020.

[Le13] Lewis, G. A.: Role of Standards in Cloud-Computing Interoperability. Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS). IEEE Computer Society, Wailea, Maui, HI, U.S., pp. 1652-1661, 2013.

[Lo21] Lobschat, L.; Mueller, B.; Eggers, F.; Brandimarte, L.; Diefenbach, S.; Kroschke, M.; Wirtz, J.: Corporate Digital Responsibility. Journal of Business Research 122, pp. 875-888, 2021.

[Lu02] Luo, X.: Trust Production and Privacy Concerns on the Internet: A Framework Based on Relationship Marketing and Social Exchange Theory. Industrial Marketing Management 31/2, pp. 111-118, 2002.

[Mi21] Mihale-Wilson, A. C.; Zibuschka, J.; Carl, K. V.; Hinz, O.: Corporate Digital Responsibility – Extended Conceptualization and a Guide to Implementation. European Conference on Information Systems (ECIS) 2021, Marrakech, Morocco, 2021.

[MM07] Matten, D.; Moon, J.: Pan-European Approach: A Conceptual Framework for Understanding CSR. In (Zimmerli, W. C.; Richter, K.; Holzinger, M. eds.): Corporate Ethics and Corporate Governance. Springer, Berlin, pp. 404-424, 2007.

[MR02] Maignan, I.; Ralston, D. A.: Corporate Social Responsibility in Europe and the U.S.: Insights from Businesses' Self-Presentations. Journal of International Business Studies 33/3, pp. 497-514, 2002.

[Ra16] Ransbotham, S.; Fichman, R. G.; Gopal, R.; Gupta, A.: Special Section Introduction—Ubiquitous IT and Digital Vulnerabilities. Information Systems Research 27/4, pp. 834-847, 2016.

[Sm17] Smith, B. W.: Automated Driving and Product Liability. Michigan State Law Review 2017/1, pp. 1-74, 2017.

[Th17] Thorun, C.; Vetter, M.; Reisch, L.; Zimmer, A. K., https://www.bmjv.de/G20/DE/ConsumerSummit/_documents/Downloads/Studie.pdf?__blob=publicationFile&v=1, accessed: 09/07/2019.

[Vi19] Vial, G.: Understanding Digital Transformation: A Review and a Research Agenda. The Journal of Strategic Information Systems 28/2, pp. 118-144, 2019.

[Wi21] Wickert, C.: Corporate Social Responsibility Research in the Journal of Management Studies: A Shift from a Business-Centric to a Society-Centric Focus. Journal of Management Studies 58/8, pp. E1-E17, 2021.

Flexible Method for Supporting OAuth 2.0 Based Security Profiles in Keycloak

Takashi Norimatsu[1,2], Yuichi Nakamura[1] and Toshihiro Yamauchi[3]

Abstract: Keycloak is identity and access control open-source software. When used for open banking, where many OAuth 2.0 clients need to be managed and a different OAuth 2.0-based security profile needs to be applied to each type of API, the problem of increasing managerial costs by the Keycloak administrator occurs because Keycloak's security profile logic depends on the client settings, and the logic cannot be changed for each client's request. This paper proposes its solution by separating the security profile logic from the client settings, and by changing the security profile for each client's request based on the content of the request, and actual security profiles Financial-grade API (FAPI) are implemented to Keycloak. The paper calculates managerial costs in both the existing and proposed methods in scenarios managing FAPI, and compares the results. The comparison shows that using the proposed method reduces costs. Our implementations are contributed to Keycloak.

Keywords: OAuth 2.0, Security Profile, FAPI, Open Source, Keycloak, Open Banking

1 Introduction

OAuth 2.0 [Ha21] is a widely used web-based authorization protocol. It is defined as a framework, so it can be used flexibly in a wide range of use cases. This flexibility, however, might introduce security holes if it is used incorrectly or inappropriately. To prevent such problems, detailed methods for using OAuth 2.0 securely have been developed. These are called *security profiles*.

Some organizations have standardized and published security profiles. Examples of such security profiles are Financial-grade API Security Profile 1.0 Baseline (FAPI1-baseline) [Fi21a] and Advanced (FAPI1-advanced) [Fi21b] by the OpenID Foundation (OID-F).

The security profiles of several in-service open banking systems are based on FAPI1-advanced. Examples of such security profiles are Open Banking Security Profiles [Op21a] in the UK, Consumer Data Right (CDR) security profile [Co21] in Australia, and Open Banking Brasil Financial-grade API Security Profile 1.0 [Op21b] in Brazil.

[1] Hitachi, Ltd., 6-6, Marunouchi 1-chome, Chiyoda-ku, Tokyo, 100-8280 Japan
[2] Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530 Japan
[3] Faculty of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530 Japan, https://orcid.org/0000-0001-6226-5715

cbe doi:18.18420/OID2022-07

In systems supporting security profiles based on OAuth 2.0, the authorization server is a key component because it plays a central role in OAuth 2.0. A variety of proprietary and open-source identity and access management (IAM) software support the functionality of authorization server. Keycloak[4] is one example of such IAM open-source software. It is written in Java and is used widely for authentication and authorization purposes [NK20]. Keycloak can be used free-of-charge but there are some charged services[5].

Keycloak manages several kinds of entities. Figure 1 shows some of the relationships among them. The figure only depicts the entities and relationships that relate to the topic of this paper.

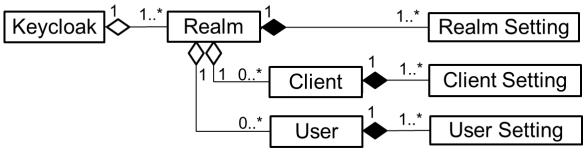


Fig. 1: Some of the relationships among entities managed by Keycloak

Keycloak creates realms that are separated and not accessible to each other. Keycloak manages clients, users and other entities within a realm. Each realm has its own settings.

Keycloak manages a client application ("Client" in Figure 1 and generally referred to as just "client"), which provides some services to end users, and treats this client as an OAuth 2.0 client. A client has several kinds of settings. Some are defined by OAuth 2.0 and called Client Metadata [Jo21] while others are defined by Keycloak. A client can send several kinds of requests to Keycloak, and these requests (for example, authorization requests and token requests) are defined by OAuth 2.0. When a client sends a request to Keycloak, Keycloak processes the request based on the client settings in Keycloak.

Keycloak manages each end user ("User" in Figure 1) of a client. For example, in an authorization code flow of OAuth 2.0, Keycloak needs to authenticate a user and get consent from the user to allow the client to access the user's own resources.

Applying a security profile by Keycloak means that Keycloak processes a request from a client and judges whether the request satisfies the requirements of the security profile. If the requirements are satisfied, Keycloak returns a normal response. If they are not satisfied, Keycloak returns an error response.

Existing open banking systems need to manage many clients because the systems are applied nation-wide. Open banking systems also need to support multiple types of APIs that require different security levels because the systems need to apply a different security profile for each type of API.

[4] https://www.keycloak.org, accessed: 06/07/2021.
[5] https://access.redhat.com/products/red-hat-single-sign-on, accessed: 16/12/2021.

When Keycloak is used as an authorization server in open banking, two problems increase the managerial costs of a Keycloak administrator. The first problem is that the Keycloak administrator needs to manage many client settings to apply security profiles. The second problem is that the Keycloak administrator needs to create a realm for each security profile and needs to manage the realm and entities included in the realm. The cause of these problems is that Keycloak's security profile logic depends on the client settings and cannot be changed for each client's request.

To resolve these problems, we designed a flexible method for supporting several security profiles and implemented it with Keycloak. To resolve the first problem, the method can separate the logic related to security profiles from the client settings in Keycloak, so that the client settings of all clients do not need to be set up in Keycloak to apply a security profile. This separation reduces the costs for managing clients. To resolve the second problem, the method can change the security profile dynamically for each client request based on the content of the client request, so that a realm does not need to be created for each security profile. Making such changes dynamically reduces the costs for managing realms and entities included in the realms.

To show that the proposed method can resolve the problems, we calculated the managerial costs needed to apply several security profiles for many clients, and compared the resulting costs with costs from before implementing the proposed method. The comparison shows that the proposed method reduces managerial costs.

The implementations were contributed to Keycloak, reviewed by Keycloak maintainers, and successfully merged into Keycloak’s main branch.

The rest of this paper is structured as follows: Section 2 describes, in detail, the problems that occur when Keycloak is used for open banking. Section 3 gives details on the proposed method for resolving the problems. Section 4 shows the evaluation of the proposed method to show that it can resolve the problems. Section 5 describes the results of submitting the implementation to the upstream Keycloak repository. Section 6 gives works related to the proposed method. Finally, Section 7 gives a conclusion.

2 Problems related to applying security profiles

In open banking, an appropriate security profile should be applied to protect APIs. As Figure 2 shows, the straightforward method (the *client settings-based method*) of applying a security profile with Keycloak is to find the values of the client settings (in Keycloak) relevant to the security profile, set such values to the client settings, and process the client’s request by following the values. If existing client settings cannot cover a security profile, adding new client settings and logic becomes necessary.

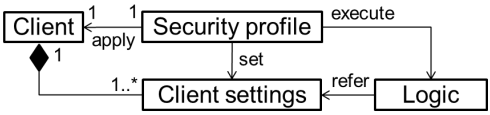


Fig. 2: Applying a security profile to a client by the client settings (in Keycloak)

Requirements for deploying an authorization server in open banking are the following:

- Requirement 1: An authorization server needs to manage many clients and end users, because open banking is applied nation-wide. For example, 319 clients have been registered for Open Banking in the UK[6].
- Requirement 2: In open banking, an authorization server needs to manage several security profiles to protect several types of APIs that require different security levels. For example, one type of API (a read API) is used to retrieve an end user's bank account's balance and transaction history. Another type of API (a write API) is used to initiate a payment service on behalf of the end user. Protecting each type of API requires different security levels. In these examples, FAPI1-baseline is intended to be used to protect a read API while FAPI1-advanced is for a write API.

However, applying security profiles by the client settings-based method for open banking is difficult due to the following two problems.

- Problem 1: The client settings of every client need to be set up appropriately for a security profile. The amount of managerial operation for managing clients increases when there are many clients.
- Problem 2: Only one security profile can be applied to one client in a realm. If multiple security profiles need to be managed, one realm needs to be created for each security profile. The amount of managerial operations for managing realms and entities included in the realms also increases when there are multiple security profiles.

These problems cause difficulties. Increasing the required operations increases costs. In addition, the increase in operations increases the risk of operational mistakes, which often cause security incidents.

3 Policy-based method for applying security profiles

To resolve the problems described in section 2, a flexible method (the *policy-based method*) for applying a security profile was designed. Its design principles are as follows.

[6] https://www.openbanking.org.uk/fintechs/, accessed: 09/01/2022.

- Design principle 1: To resolve problem 1, the logic related to security profiles is separated from client settings so that the client settings of every client do not need to be configured to be appropriate for a security profile.
- Design principle 2: To resolve problem 2, a security profile applied to a request from the same client in a realm can be changed dynamically so that another realm does not need to be created for each security profile.

Figure 3 shows a schematic diagram of the policy-based method. “Request Context” indicates a request’s content and context, such as the HTTP header’s value. “Security policy” determines which security profile is applied to a client’s request based on the request context and client setting. To apply a security profile, the policy-based method executes its logic to judge whether the request satisfies a security profile’s requirements regardless of the client settings.

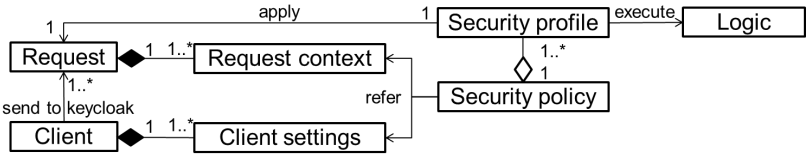


Fig. 3: Applying a security profile to a request from a client regardless of the client settings

By following the above design principles, *client policies*, the framework for processing a request from a client, are implemented as an implementation of the policy-based method.

Figure 4 shows logical components of the client policies. As shown below, four types of components are defined for client policies: executor, profile, condition. and policy.

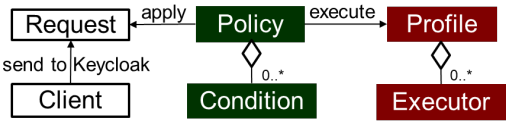


Fig. 4: Logical components of client policies

An *executor* is a component that includes the logic needed to apply part of a security profile. By following design principle 1, the logic does not depend on client settings.

A profile is a component that includes all the logic needed to apply a security profile itself. A profile consists of several executors. FAPI1-baseline, FAPI1-advanced, and FAPI-CIBA security profile (FAPI-CIBA) are implemented as executors and profiles.

A *condition* is a component that includes the logic determining whether a profile is to be applied to a client’s request. By following design principle 2, the condition can use the context data of a client’s request (for example, parameters including a request and the HTTP context) to determine whether a profile is to be applied to a client’s request. Therefore, based on the content of the request, Keycloak can change the security profile.

A *policy* is a component that includes all the logic needed to determine whether a profile is to be applied to a client's request. A policy consists of several conditions. If all evaluation results of the conditions are positive, the policy applies a security profile.

Figure 5 shows how client policies work. The example has two policies, the FAPI1-baseline policy and the FAPI1-advanced policy. There are also two profiles, the FAPI1-baseline profile and the FAPI1-advanced profile. The FAPI1-baseline policy decides to apply the FAPI1-baseline profile if a client's request includes the "read_account" scope value of OAuth 2.0. The FAPI1-advanced policy decides to apply the FAPI1-advanced profile if the client's request includes the "bank_transfer" scope value. Then:

- When a client sends a token request with a scope including "read_account", the FAPI1-baseline policy decides to apply the FAPI1-baseline profile to the request, but the FAPI1-advanced policy decides not to apply the FAPI1-advanced profile.
- When a client sends a token request with a scope including "bank_transfer", the FAPI1-baseline policy decides not to apply the FAPI1-baseline profile to the request, but the FAPI1-advanced policy decides to apply the FAPI1-advanced profile to the request.

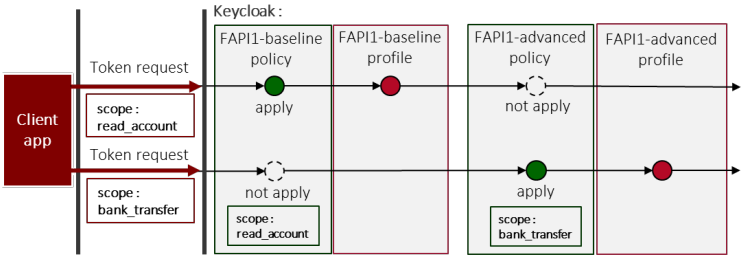


Fig. 5: An example of how client policies work

4 Evaluation

The policy-based method was evaluated to confirm that it can resolve the problems described in section 2. Using Keycloak 15.0.2, we derived the theoretical costs for managing security profiles by a Keycloak administrator in both the client settings-based method and the policy-based method. We calculated the managerial costs in scenarios managing FAPI for both the client settings-based method and the policy-based method and compared the costs.

4.1 Assumptions

To focus the discussion on managing security profiles, only entities related to security profiles (namely, realms, clients, and client policies) are considered. To simplify the

discussion, it is assumed that the managerial cost for manually specifying any setting for one entity is the same and is considered to be 1.

As described in section 2, it is assumed that a security profile can be applied by setting valid values for the client settings (in Keycloak) of a client managed by Keycloak. However, to apply FAPI1-baseline, FAPI1-advanced, and FAPI-CIBA used afterwards in the examples, some logic in Keycloak's body code becomes necessary. Such additional coding is ignored in the evaluation.

4.2 Theoretical Costs

The calculations consider the following three task patterns for managing security profiles. In each pattern, the theoretical cost is derived from the number of settings that need to be managed.

1. Initializing the Keycloak environment that supports security profiles.
2. Adding new security profiles to existing the Keycloak environment.
3. Modifying security profiles of existing the Keycloak environment.

The following variables are defined.

N_{SREL} ≡ The number of realm settings

N_{CLI} ≡ The number of clients

N_{SCLI} ≡ The number of client settings

N_{SPF} ≡ The number of security profiles applied to the same client

N_{SSPF} ≡ The number of settings for a security profile

N ≡ The number of settings for managing security profiles

Figure 6 shows which entities the variable corresponds to in both the client settings-based method and the policy-based method.

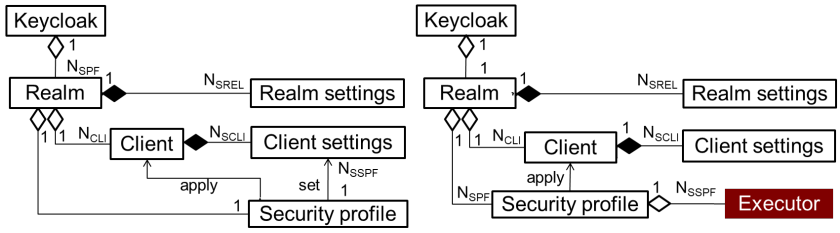


Fig. 6: Numerical relationships among security profile related entities

In the client settings-based method, N_{SSPF} is the number of client settings relating to a security profile; however, in the policy-based method, N_{SSPF} is the number of executors for a security profile.

Table 1 shows the managerial cost for each task pattern.

Pattern	N (client settings-based)	N (policy-based)
Initializing Keycloak for security profiles	$N_{SPF} \times (N_{SREL} + N_{CLI} \times (N_{SCLI} + N_{SSPF}))$ $\propto N_{SPF} \times N_{CLI}$	$N_{SREL} + N_{CLI} \times N_{SCLI} + N_{SPF} \times N_{SSPF}$ $\square N_{SPF} + N_{CLI}$
Adding new security profiles	$N_{SPF} \times (N_{SREL} + N_{CLI} \times (N_{SCLI} + N_{SSPF}))$ $\square N_{SPF} \times N_{CLI}$	$N_{SPF} \times N_{SSPF}$ $\square N_{SPF}$
Modifying security profiles	$N_{SPF} \times N_{CLI} \times N_{SSPF}$ $\square N_{SPF} \times N_{CLI}$	$N_{SPF} \times N_{SSPF}$ $\square N_{SPF}$

Tab. 1: Theoretically derived managerial costs for security profiles

In pattern 1, if the client settings-based method is used, the Keycloak administrator needs to create a realm for each security profile, and create clients and set up their client settings for the security profile for each realm. If the policy-based method is used, the Keycloak administrator needs to create only one realm, create clients in the realm and set up executors for each security profile.

In pattern 2, if the client settings-based method is used, the Keycloak administrator needs to perform the same tasks as in pattern 1. If the policy-based method is used, the Keycloak administrator needs to only set up executors for each added security profile.

In pattern 3, if the client settings-based method is used, the Keycloak administrator needs to set up client settings for all clients for a modified security profile. If the policy-based method is used, the Keycloak administrator needs to only set up executors for each modified security profile.

4.3 Calculating the Costs in the Examples

Three examples for the above task patterns are suggested: initializing the Keycloak supporting FAPI1-baseline, adding FAPI1-advanced, and modifying FAPI1-advanced to FAPI-CIBA. Keycloak has 117 realm settings ($N_{SREL} = 117$) and 110 client settings[7] ($N_{SCLI} = 110$), and it is assumed that this security profile is applied to requests from 1000 clients ($N_{CLI} = 1000$).

[7] https://www.keycloak.org/docs/15.0/server_admin/index.html#_oidc_clients, accessed: 09/01/2022.

Using the client settings-based method, 13 client settings (Proof Key for Code Exchange Code Challenge Method, Client Authenticator, Root URL, Admin URL, Base URL, Web Origins, Backchannel Logout URL, Valid Redirect URIs, JWKS URL, Valid Request URIs, CIBA Client Notification Endpoint URL, Consent Required, and Full Scope Allowed) need to be set up properly ($N_{SSPF} = 13$) to support FAPI1-baseline.

Using the policy-based method, 6 executors[8] (Secure Session Enforce, PKCE Enforcer, Secure Client Authenticator, Secure Client URIs, Consent Required, and Full Scope Disabled) need to be set up ($N_{SSPF} = 6$) to support FAPI1-baseline.

Using the client settings-based method, 21 client settings (Access Type: bearer, Access Type: public, Client Authenticator, Root URL, Admin URL, Base URL, Web Origins, Backchannel Logout URL, Valid Redirect URIs, JWKS URL, Valid Request URIs, CIBA Client Notification Endpoint URL, Use ID Token as a Detached Signature, User Info Signed Response Algorithm, ID Token Signature Algorithm, Access Token Signature Algorithm, Request Object Signature Algorithm, Signature Algorithm, Consent Required, Full Scope Allowed, OAuth 2.0 Mutual TLS Certificate Bound Access Tokens Enabled) need to be set up ($N_{SSPF} = 21$) to support FAPI1-advanced.

Using the policy-based method, 11 executors (Secure Session Enforce, Confidential Client Accept, Secure Client Authenticator, Secure Client URIs, Secure Request Object, Secure Response Type, Secure Signing Algorithm, Secure Signing Algorithm for Signed JWT, Consent Required, Full Scope Disabled, and Holder of Key Enforcer) need to be set up ($N_{SSPF} = 11$) to support FAPI1-advanced.

Using the client settings-based method, 22 client settings (21 of them are the same for FAPI1-advanced, and CIBA Backchannel Authentication Request Signature Algorithm) need to be set up properly ($N_{SSPF} = 22$) to support FAPI-CIBA.

Using the policy-based method, 14 executors (11 are the same for FAPI1-advanced, Secure CIBA Authentication Request Signing Algorithm, Secure CIBA Session Enforce, and Secure CIBA Signed Authentication Request) need to be set up ($N_{SSPF} = 14$) to support FAPI-CIBA.

Table 2 shows managerial costs for managing security profiles. In the example, “client” means the client settings-based method and “policy” means the policy-based method.

Example	N_{SSPF} (client)	N_{SSPF} (policy)	N (client)	N (policy)
Initializing Keycloak for FAPI1-baseline	13	6	123117	110123
Adding FAPI1-advanced	21	11	131117	11
Modifying FAPI1-advanced to FAPI-CIBA	22	14	22000	14

Tab. 2: Calculated managerial costs for security profiles based on the examples

[8] https://www.keycloak.org/docs/15.0/server_admin/index.html#_client_policies, accessed: 09/01/2022.

4.4 Discussion

Considering the theoretically derived managerial costs and their examples shown in Table 1 and Table 2, we can conclude that using the policy-based method is a good option from the perspective of managerial costs for the following usage situations while the client settings-based method is an acceptable option in other cases.

- When different security profiles need to be applied to requests from the same client
- When many clients need to be supported
- When security profiles need to be added or modified frequently

5 Contribution to the Keycloak Upstream Repository

To contribute the implementation of client policies and security profiles (FAPI1-baseline, FAPI1-advanced and FAPI-CIBA) to the Keycloak upstream repository, we needed to confirm that the implementation for security profiles complies with the FAPI specifications. To validate the implementation, we used OpenID Conformance Suite provided by OID-F[9] and confirmed that Keycloak code including the implementation could pass the conformance tests for both FAPI1-advanced and FAPI-CIBA. After this validation, the implementation was contributed to the Keycloak project, reviewed by Keycloak maintainers, and merged into Keycloak's main branch in Keycloak 15[10].

6 Related work

The policy-based method introduces the idea of Attribute Based Access Control (ABAC) to change the security profile applied to a client's request based on the content of the request. [EY05] established ABAC's formal model and proved that this model can represent the Mandatory Access Control (MAC) and Discretionary Access Control (DAC) model. Generally, ABAC's access control part can be achieved by the policy-based method. [PRR16] categorized the access control part into a logic-based policy and enumerated policy. [PRR16] proposed Label Based Access Control for enumerated policies and established its formal model. [XRR12] derived the minimum requirements of ABAC for representing MAC, DAC, and Role Based Access Control (RBAC) and formalized the minimum ABAC model satisfying these requirements.

OAuth 2.0 itself is so flexible that it has been used in a wide range of use cases for web-based authorization. [Fe17] applied OAuth 2.0 to accessing data collected by IoT devices. They modified the OAuth 2.0 protocol by considering a situation where IoT-device

[9] https://openid.net/certification/about-conformance-suite/, accessed: 06/07/2021.
[10] https://www.keycloak.org/docs/latest/release_notes/index.html#financial-grade-api-fapi-improvements-fapi-ciba-and-open-banking-brasil, accessed: 09/12/2021.

resources are limited and constrained. [Fe17] used RBAC to access such data. [ZRS19] applied OAuth 2.0 to access services provided by microservices in a container-orchestrated platform. [AAM19] applied and modified OAuth 2.0 for user centric identity management. All these studies use OAuth 2.0 to control access to resources provided by resource servers via APIs in OAuth 2.0's context by using ABAC. Unlike these studies, the proposed policy-based method in this paper utilizes the idea of ABAC for access control at OAuth 2.0's endpoints to determine whether an authorization server like Keycloak accepts a request from a client at these endpoints.

7 Conclusion

For Keycloak to support security profiles based on OAuth 2.0 and meet large-scale use cases such as open banking, the problem of increasing Keycloak managerial costs must be resolved. To resolve the problem, a policy-based method was proposed and implemented as client policies to enable flexible management of security profiles. Actual security profiles like FAPI1-baseline, FAPI1-advanced, and FAPI-CIBA have been implemented using client policies.

To confirm that the proposal resolves the problem, three scenarios of initializing, adding, and modifying security profiles were considered. Costs were calculated for these scenarios, and the results show that managerial costs of the Keycloak administrator decrease compared with the existing client settings-based method.

To validate the implementation of FAPI security profiles, we proved that it complies with the FAPI security-profile specifications by passing FAPI conformance tests. All implementations were contributed to Keycloak and merged into Keycloak's main branch. In the future, other security profiles like FAPI 2.0 [Fe21], the next major version of FAPI 1.0, will be implemented and contributed to Keycloak.

Acknowledgement

We would like to express our thanks to Stian Thorgersen (Keycloak project lead), Marek Posolda (Keycloak development team), FAPI-SIG's members, and the Keycloak community. They encouraged us to hold further discussions and provided us with advice about our implementation of client policies.

Bibliography

[AAM19] Abubakar-Sadiq, Shehu; António, Pinto; Manuel, E. Correia: Privacy Preservation and Mandate Representation in Identity Management Systems. In (Álvaro Rocha, et.al. Eds.): Proc. the 2019 14th Iberian Conference on Information Systems and

Technologies (CISTI), 2019.

[Co21] Consumer Data Right Security Profile, https://consumerdatastandardsaustralia.github.io/standards/#security-profile, accessed: 06/07/2021.

[EY05] Eric, Yuan; Jin, Tong: Attributed based access control (ABAC) for web services. In (Randall Bilof, ed.): Proceedings of the IEEE International Conference on Web Services (ICWS), vol. 856, pp. 561-569, 2005.

[Fe17] Federico, Fernández et.al.: A model to enable application-scoped access control as a service for IoT using OAuth 2.0. In (Noel, Crespi et.al. Eds.): Proc. the 2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN), pp.322-324, 2017.

[Fe21] Daniel, Fett: FAPI 2.0: A High-Security Profile for OAuth and OpenID Connect. In (Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim Eds.): Proc. Open Identity Summit 2021, pp. 71-81, 2021.

[Fi21a] Financial-grade API Security Profile 1.0 - Part 1: Baseline, https://openid.net/specs/openid-financial-api-part-1-1_0-final.html, accessed: 06/07/2021.

[Fi21b] Financial-grade API Security Profile 1.0 - Part 2: Advanced, https://openid.net/specs/openid-financial-api-part-2-1_0.html, accessed: 06/07/2021.

[Ha21] Dick, Hardt et.al.: The OAuth 2.0 Authorization Framework, https://datatracker.ietf.org/doc/html/ rfc6749, accessed: 06/07/2021.

[Jo21] Michael, Jones et.al.: OAuth 2.0 Dynamic Client Registration Protocol, https://datatracker.ietf.org/doc/html/rfc7591, accessed: 06/12/2021.

[NK20] Yuichi, Nakamura; Kazufumi, Enomoto: Authentication and Authorization Based on OSS for Secure System Interoperation, https://www.hitachi.com/rev/archive/2020/r2020_05/05a04/index.html, accessed: 16/12/2021.

[Op21a] Open Banking Security Profiles, https://standards.openbanking.org.uk/security-profiles, accessed: 06/07/2021.

[Op21b] Open Banking Brasil Financial-grade API Security Profile 1.0 Implementers Draft 3, https://openbanking-brasil.github.io/specs-seguranca/open-banking-brasil-financial-api-1_ID3.html, accessed: 04/12/2021.

[PRR16] Prosunjit, Biswas; Ravi, Sandhu; Ram, Krishnan: Label-based access control: An ABAC model with enumerated authorization policy. In (Elisa Bertino, et.al. Eds.): Proc. the 2016 ACM International Workshop on Attribute Based Access Control, pp. 1-12, 2016.

[XRR12] Xin, Jin; Ram, Krishnan; Ravi, Sandhu: A unified attribute-based access control model covering DAC, MAC and RBAC. In (David Sadek, et.al. Eds.): DBSec'12: Proceedings of the 26th Annual IFIP WG 11.3 conference on Data and Applications Security and Privacy, pp. 41-55, 2012.

[ZRS19] Zehan, Triartono; Ridha, Muldina Negara; Sussi: Implementation of Role-Based Access Control on OAuth 2.0 as Authentication and Authorization System. In (Hendri Irawan , et.al. Eds.): Proc. the 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), pp. 259-263, 2019.

Risk variance: Towards a definition of varying outcomes of IT security risk assessment

Sebastian Kurowski[1], Christian H. Schunck[2]

Abstract: Assessing IT-security risks in order to achieve adequate and efficient protection measures has become the core idea of various industry practices and regulatory frameworks in the last five years. Some research however suggests that the practice of assessing IT security risks may be subject to varying outcomes depending on personal, situational and contextual factors. In this contribution we first provide a definition of risk variance as the variation of risk assessment outcomes due to individual traits, the processual environment, the domain of the assessor, and possibly the target of the assessed risk. We then present the outcome of an interview series with 9 decision makers from different companies that aimed at discussing whether risk variance is an issue in their risk assessment procedures. Finally, we elaborate on the generalizability of the concept of risk variance, despite the low sample size in light of varying risk assessment procedures discussed in the interviews. We find that risk variance could be a general problem of current risk assessment procedures.

Keywords: Risk Analysis, Risk Assessment, Risk Management, IT-Security, Information Security

1 Introduction

Risk analysis has become an important cornerstone of information security management. For instance, the EU General Data Protection Regulation (GDPR) requires security measures to be adequate in light of the risk for the data subjects rights and liberties (Article 32, paragraph 1 and article 24, paragraph 1, GDPR). Industrial Frameworks such as the VDA Information Security Assessment (ISA) [VD15] require a security level and thus risk associated characterization of security measures. These are just two example of frameworks that have shifted towards a risk-based approach, putting the justifiability of security measures at their core. This development seems reasonable, since managing information security around assessed information security risks allows organizations not just to choose the right security measures, but to align their budgets accordingly, and to

[1] Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO, Team identity management, Nobelstr. 12, Stuttgart, 70569, sebastian.kurowski@iao.fraunhofer.de
[2] Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO, Team identity management, Nobelstr. 12, Stuttgart, 70569, christian.schunck@iao.fraunhofer.de https://orcid.org/0000-0002-7917-8180

The research presented in this paper was partially funded by the federal ministry for economic affairs and climate action under grant no. 01MT19006B and partially funded by the Zero Outage Industry Standard Ltd.

have reasonable justification if incidents happen despite taken efforts. However, these advantages can only materialize, if the assessment of risks is reliable and factual. Reasonable justifications can only stand if the risk has been regarded beyond possible doubt. The optimal budget can only be determined if risks have been assessed without any biases. In the largely positivistic research area of IT- and information security the factuality of risk assessments is often assumed implicitly. Yet Baskerville drew an argument for the interpretivistic nature of risk assessments [Ba91], indicating that these may be biased by the person interpreting the risk. Additionally, Luhmann [Lu90] provides an argument for the subjectivity of risks by arguing that a risk is an anticipation of observed threats. This can also incidcate that the assessment of risks may not only be subject to individual, subjective traits but also to factors surrounding the anticipation of a risk and the observation of a threat. Finally, the dissertation by Mersinas [Me17] shows that decision making and attitude of security deciders can be influenced by risk aversion and affinity.

This raises the question: Can risk assessments vary based on non-risk related traits?

In this contribution we coin the term risk variance as varying outcomes of risk assessments. We provide the results of semi-structed interviews with nine decision makers from the IT and information security domain in different organizations on the existence of variance in risk assessments. We then discuss the findings along the existing body of knowledge on influencing factors of decision making and arrive at a definition of risk variance. We also discuss how general the problem of risk variance, and the identified factors could be. The following section provides a first characterization of what could possibly characterize risk variance, followed by a brief discussion of the current state of the art of risk assessments.

2 State of the Art

2.1 Variation and biases in security decision making

Some publications discovered biases in security decision making. Hyeun-Suk et al. [Hy12] showed that security decision makers would tend to assess other companies as more vulnerable than their own company. Mersinas [Me17] showed that decision makers indicate subjective affinity or aversion towards certain risk scenarios. Still, the subjectivity of organizational analyses, organizational decision-making, and thus also risk assessment is a rare research subject in information and IT security research.

However, extensive research exists from the field of psychology, sociology and economics. The research of Kahneman and Tversky [KT79] shows that individuals can indicate affinity or aversion towards specific risk scenarios, which matches the findings by Mersinas [Me17]. Nosofsky [No83] and later Benjamin et al. [Be09] showed that criterion selections can change based on the presentation of those criterions (criterion

noise). E.g. the number of criterions can increase criterion noise, while providing overviews can decrease it. Gilboa and Schmeidler [GS89] showed that subjects tend to regard known scenarios as more significant than unknown scenarios. And finally, Hermand et al. [He03] find that the target that a risk applies to (risk target) influences the significance of that risk for the assessing individual. They showed that risks that apply to strangers are perceived more likely, than those that apply to the assessors. This shows that while there is few existing evidence for varying outcomes of decision-making processes in IT- or information security, there is a large body of knowledge on possible individual, situational, presentational (i.e. criterion noise), and contextual influences, that may as well apply to IT- or information security.

2.2 Variation in risk assessment approaches

These factors, however, do not play any role in current risk assessment approaches. Good practices and norms such as ISO/IEC 27005:2018 [Is18], OCTAVE, OCTAVE FORTE [AD02], ITU X.1208 , NIST SP 800-122 , BSI-Standard 200-3, factor analysis of information risk (FAIR), or the French “expression des besoins et identification des objectivs de sécurité” (EBIOS) do not take assessor traits, situational traits, or any other influencing factors into account. The only existing norm that considers its organization surrounding is NIST SP 800-30, which requires risk assessments to be structured along the organization’s hierarchy. Peer reviewed literature on risk assessments on the other hand largely considers automation approaches, over variation minimization. Zhang and Rao use neural networks [Zr20] for risk assessments, Shakibazad and Rashidi [Sh20] build upon pre-assessed vulnerability scores which are assumed to be objective, Riesco and Villagra build assessments on large semantic networks [RV19], Rios et al. use attack trees [Ri20], and James [Ja19] derives risk assessments based on deterministic finite automation. None of these approaches take the variability of inputs or the variable interpretation of outputs into account. But even in non- or semi-automated approaches, assessment procedures reducing or avoiding possible variances do not play a role. Teng et al. [Te20] employ an analytical hierarchy process (AHP) [Sa88] in order to weigh different risks. While this could potentially decrease criterion noise, it does not weigh on the other possible influencing factors. Sektas-Bilusisch et al. [Se20] combine focus groups with a formal model in order to assess risks. However, they as well do not take any possible variations in account. This shows that the variation of risk outcomes based on individual, situational, presentational, and contextual cues is not yet considered within industry practices, norms, or research.

3 On the relevance of risk variance

Since no direct evidence of risk variance could be obtained from existing literature, yet the existence of this problem seemed to be plausible in light of the body of knowledge of other research domains, we conducted an interview series with nine different decision

makers from nine different organizations.

3.1 Sampling and Data Capturing

The interviews aimed at verifying or falsifying the existence of risk variance in the IT- and information security risk assessment processes of these companies. Additionally, details on how risks are conducted, which norms are used, what role security plays within the organization, and if risk variance was observed, how the organization mitigates this variance were sought. These interview aims provide both exploratory and confirmatory research questions. Therefore, a semi-structured interview methodology was used [My09] as it allowed the interviewers to deviate from the question script in order to further explore the responses of interviewees. The interviews were conducted as part of a funded project by an association for IT availability. This allowed for the acquisition of interview partners from the members of this association. While the thematic frame of the association (along with a small sample size) may hinder the generalizability of these findings, we were still able to acquire interviewees from different functions including information security, quality management, sales and executive roles. Interviews were conducted by two interviewers. Given that interviewees consented to recording, all interviews were recorded for later analysis and deleted after the analysis was finished. No interviewee objected to the interview being recorded. A third researcher transcribed the interviews, which were then used for analysis.

3.2 Data analysis

Due to the explorative properties of semi-structured interviews, one of the main tasks of the analysis methodology was to reduce the possible variety of statements without losing too much information. Qualitative content analysis (QCA) was chosen for this purpose [EK08][Sc19]. QCA provides for interview transcripts to be analysed with a thematic framework of main themes and sub-themes. The use of code systems for analysis within the thematic framework is not obligatory. Therefore, code systems were not used in the analysis of the interviews. Although these represent a considerable reduction of the data [GL13], an ex-ante elaboration of code systems would get in the way of the explorative character of the data. An elaboration of codes during the analysis, as used for example in grounded theory based analysis approaches [GS71][HJ03] also did not seem profitable, as an elaboration of explanatory substantive and general theories [Ur09] would go beyond the scope of this publication. Furthermore, due to the number of interview partners (n=9), no value was seen in quantitative analysis, which ultimately led to the decision not to use code systems. The thematic framework was used by two researchers working independently to interpret the transcribed responses. These interpretations were then checked for agreement by both researchers. Discrepancies were resolved in a meta-interpretation. This meta-interpretation was finally used for a narrative summary, similar to a narrative review of literature [Ja16]. The thematic framework used for the analysis is presented in the Appendix. This represents the respective main topics on which the

analysis is based. For example, risk analysis questions should be considered in terms of their degree of systematisation (standards used, use of standards, risk factors considered, weighting/measurement of risk, abstract description of approach). Analysis-initiating factors should be distinguished in terms of regular and irregular factors. The regularity of the analysis was considered exclusively in terms of the period after which an analysis is repeated. Influences on monetary planning and reserves were considered in terms of their existence, the nature of the influence and the monetary aspects influenced. In contrast, purposes of the risk analysis that go beyond this were not to be explored in greater depth. The influence of risk analysis on the company was additionally regarded by the thematic framework. The monetary influence was in the foreground, since this could play a supporting role for the concept of efficiency under the assumption that entrepreneurial action can be reduced to the exchange of monetarily measurable resources. However, additional purposes of risk analysis can indicate its value for the company's success. The occurrence of risk variance was analysed with regard to its existence in principle and possible reasons for it. If risk variances occur in the company, possible limiting countermeasures were recorded. If none occurred, possible preventive countermeasures were considered. However, this case did not occur with any of the interview partners. Finally, the analysis of the demographic questions aimed to analyse the current perspective on the company, the professional proximity to risk analyses, relevant previous experience and the relevance of the topic of information security for the organisation itself, both in absolute terms and in relation to other important (open) topics such as customer satisfaction, or shareholder value. In the course of the analysis, it became apparent that interviewee 8 could not give any organisation-specific answers due to his role as a security consultant. Since the statements therefore referred to his general view, but not to a specific company, the answers were excluded from the development of the meta-interpretation. This results in an effective sample of (n=8).

3.3 Findings

Tab. 2 shows that all interviewees claimed that they have observed risk variance in their risk assessment outcome. This is especially interesting, as the standardization of the risk assessment process varies from standardized according to international norms, standardized according to company specific processes, semi-standardized with checklists and templates to ad-hoc improvised assessment processes. Obviously the systematicity of assessments does not mitigate risk variance sufficiently.

R.V.*	Impact as..	Probability as..	Risk aspects	Standardization
✓	Business Impact	Quantitative	B, S, O	ISO 27k process
✓	Financial Impact	No information	S, P, Fi, Pr	No information

R.V.*	Impact as..	Probability as..	Risk aspects	Standardization
✓	Financial Impact if possible	Semi-Quantitative / Quantitative if possible	S	Standardized company specific process
✓	Expert Opinion / Data if possible	Quantitative based on expert opinion / Data if possible	S, B, O, Pr, Prod, Ma, Qu, IT	Standardized company specific process
✓	Qualitative	Qualitative	IT, B	Standardized company specific process
✓	Customer-depending	Customer-depending	App	Customer depending
✓	Liability	No information	IT, Fi	Improvised
✓	Financial	Qualitative based on expert opinion	BC	Semi-standardized

* Risk Variance, ✓ Risk variance observed, B = Business Risk, S = Security Risk, O = Organizational Integration, P = Privacy Risk, Fi = Financial Risk, Pr = Price Risk, Se = Service Risk, Pr = Provisioning Risk, Prod = Production Risk, Ma = Marketing Risk, Qu = Quality Risk, IT = IT Risk, App = Application Downtime, BC = Business Continuity

Tab. 1 Observed risk variance and risk assessment characteristics mentioned by the interviewees

The same holds for the role of quantification. Some researchers, e.g. [Zu20] sometimes confuse quantification with objectiveness of results. However, our results clearly show that no matter, whether percentage point expert values, ordered non-numerical risk classes, or actual data is used, risk variance always exists within the processes. Finally, there does not seem to be an influence between the broadness of considered risk aspects. Whether risk assessments include the identification of consequences to, or influences from application downtime only, or multiple different aspects within the company, risk variance is always observed.

The factors which interviewees saw as reasons for the varying risk assessments however included risk affinity or aversion, knowledge of the domain, understanding of psychology, empathy, professional background, domain of work, contextual understanding, personality, and the situation of decision-making. Surprisingly, the professional domain was mentioned as a reason for risk variance by three different interviewees. One interviewee mentioned that IT security people might have a focus on exploits but not on topics like emergency crisis management or business continuity. Other interviewees stressed the different views between Chief Financial Officers (CFO) and Chief Information Security Officers (CISO) stressing that the CFO "…didn't see the

importance of the security as the [CISO] did." Instead, the "…CFO was more interested in reducing […] expenses related to what the [CISO] office was demanding.".

The contextual understanding of a risk scenario was mentioned by two interviewees. One mentioned that a risk assessor can have a different understanding on the services that are provided to customers, how valued the customers are, etc. Another interviewee even mentioned that "sometimes business and sales tell you this is a must win. So, then risk is looked at differently.".
Knowledge of the domain which is affected by the risk scenario, and an understanding of psychology and empathy in order to "…ask the right questions in the right way and the right times" was mentioned by one interviewee. Interestingly domain knowledge was not mentioned by any other interviewee. However, being able to ask the right questions seems to be related with the situation of decision-making, that has been mentioned by another interviewee. This interviewee claimed that the risks vary based on who it is and also how the decision is made, e.g. after detailed discussions or as an ad-hoc decision. We therefore noted the understanding of psychology and the situation of decision making both as the situation of the risk assessment in Table 3. Finally, individual differences and personality was mentioned by three different interviewees without further details on the specific traits. For instance, one interviewee mentioned that "…managers have very different personalities…", another one told us that the assessment itself is an "…individual decision.".

4 Towards a definition of risk variance

The previous section showed that risk variance is an issue with the interviewee's companies. The observation of risk variance also aligns well with findings on decision-making biases from the fields of psychology, sociology, and economics. However, the reasons given by interviewees for risk variance seem to vary.

Risk affinity or risk aversion is being mentioned by most interviewees. However, it is only mentioned with high, very high, and in one case an unclear assessment of the importance of security in the organization. It is also independent from the IT security focus of these interviewees' professional experiences. It seems hardly surprising that risk affinity or risk aversion seems to play a role when observing risk variances in organizations with high and very high importance of security. The breadth of possible discussed risk scenarios could be much larger in these companies, unveiling risk affinity or aversion towards certain scenarios more easily. This confirms hypothesis 1.

Interestingly, knowledge of the domain of a risk scenario was only mentioned by one interviewee from a security framework implementation perspective in a company with high importance of security. But if considered together with the contextual understanding of a risk scenario, it spans beyond IT scenarios and is observed with organizations that emphasize security both highly and very highly. It could be that

domain-unknowing risk assessors that assess risk scenarios under naïve or overly pessimistic scenarios are more observed with organizations that put more emphasis of risk assessments in more parts of the company, due to the high or very high importance of security. This would also explain why the contextual understanding is also observed by the interviewee with customer representative and management of outsourcing experience. The processual situation in which the risk assessment (situation of risk assessment) is conducted in, is also mentioned together with a high and very high importance of security and by interviewees with management and quality assurance experience. The professionally influenced focus on processes of these interviewees may lead to this observation. The domain of the assessors on the other hand also played a role with medium, high, and customer-focused high importance of security in the companies. It is observed by security management, enterprise architecture, and executive level management professionals. Such professions usually cooperate with various individuals from different domains. The different thought approaches, e.g. of law, psychology, sociology, business management and computer science could yield different conclusions. This however can only be observed by individuals that have worked with different domains as for instance security managers, executive managers, or enterprise architecture managers. All mentioned reasons for risk variance so far seem to be attributable to the interviewees capability of observing them. The variance between the different professional experiences and the importance of security all aligns well with the mentioned reasons. The claim of generalizability thus is almost of esoteric nature. Since we can conclude that: Risk affinity or aversion towards risk scenarios, knowledge of the domain that a risk scenario affects, contextual understanding of the risk scenario, the processual environment of the risk assessment, and the domain of the assessor seem to be general reasons for risk variance. If they are not observable, it currently seems plausible that the reason for this lack of observation may be the lense of the observer and not the non-existence of the reason. Risk variance is therefore to be defined as a variation of outcomes of IT- and information security risk assessments based on individual traits (risk affinity / aversion [KT79][Me17], knowledge of the domain [GS89], contextual understanding of the scenario), the processual environment (presentational cues [Be09][No83], social cues [Lu90]), and the domain of the assessor. Risk target [He03] was the only possible aspect of risk variance that was not mentioned by the interviewees. However, this could also be due to the lack of observability beyond experimental setups and targeted questioning of individuals.

5 Conclusion

This contribution provides insights from an interview series with 9 interviewees on the issue of risk variance. It uses the existing body of knowledge along with the insights from the interviews in order to arrive at a definition of risk variance. It also discussed the possible generalizability of these findings, beyond conceptual or sampling-based generalizability. We found that risk variance is an issue with all interviewees. The reasons for risk variance however vary slightly between the interviewees. Yet, this can

be explained by the different lenses of the interview partners. Additionally, the mentioned reasons align well with the body of knowledge on possible influencing factors of decision making under uncertainty. We therefore assume that risk variance is a generalizable issue. The definition provided in this contribution however is not possibly conceptually saturated. Hereby the severity of risk variance is not necessarily depending on the variation between two assessments by the same person. It can be severe however, if compliance goals are not met, due to variations between the assessments conducted by the organization and the assessments conducted by auditing parties, or their subcontractors. Additionally risk variance can implicate that budget decisions made on an educated argument are suddenly biased by individual, situational, and contextual traits. The question on possible mechanisms that contribute to the identified reasons for risk variance therefore is relevant. The identified reasons in this contribution seem to be due to a lack of understanding of scenarios, a lack of contextual understanding, different social cues, presentation, knowledge, and risk affinity / aversion. Except for the social cues, all of these reasons could potentially be founded in the conceptual richness of the term security risk. I.e. the FAIR ontology involves attacker models, economic models, along with IT specific terminology. This richness of concepts could increase the room for interpretation, failing to frame the decision-making biases, and thus arriving at variation of the resulting assessment. If this is true, then an epistemologically founded re-definition of the concepts of risk with the goal to minimize the conceptual richness of the term could indeed help to minimize risk variance. This however is subject to further research.

6 Annex

#	Question Type	Classification	
		Name	**Description**
Risk assessment procedures in the organization			
1	*Risk assessment procedures in the organizations*	Used Standards	Name of the standards that are used as part of the risk assessment
		Use of Standards	Role that these standards play in the risk assessment (e.g. as a baseline)
		Risk aspects at play	Parts that are considered as related to the IT security risk
		Weight / size of risk	Quantified or Qualified risk values
		Process	Participants, tasks and their execution order
2	*Triggers for risk assessments*	Irregular triggers	Irregular events that result in a (re-)assessment of risks
		Regular triggers	Regularly occuring events that result in a (re-)assessment of risks

#	Question Type	Classification	
		Name	**Description**
3	*Regularity of risk assessments*	Timespan	Regularity of reassessments
4	*Impact of risk assessments on financial aspects in the organization*	Influencing relationship	Is there an influence on any money matters?
		Type of influence on money matters	How are money matters influenced?
		Influenced aspects of money matters	What kind of money matters are influenced?
5	*Impact on other purposes*	Name of Purpose	Name of the purpose for which a risk assessment is used
Risk variances in the risk assessment procedures			
6	*Existence of risk variance*	Existence of Risk Variance	Existence or in the past observed variances of risk assessments
		Reason for Risk Variance	Assumed reasons for varying risk assessments
6a)	*Mitigating risk variance (if risk variance exists)*	Name of Measure	Measure to limit the outcome of varying risk assessments
6b)	*Preventing risk variance (if risk variance does not exist)*	Name of Measure	Measure to avoid the outcome of varying risk assessments
Demographic questions			
9	*Current role and responsibility*	Name of Role	Name of the role
		Information security related tasks	Tasks with relation to security if not implied by the role
10	*Professional experience*	Information security or IT Experience	Experience in years on security/IT or security/IT related topics
11	*Importance of information security*	Relevance of information security	Order of relevance of information security in the organization
		Relativization of information security	Relativization of information security relevance order in light of other topics or personal opinion of the interviewee

7 Bibliography

[AD02] Alberts, C.J., Dorofee, A.: Managing information security risks: the OCTAVE approach. Addison-Wesley Longman Publishing Co., Inc. (2002).

[Ba91] Baskerville, R.: Risk analysis as a source of professional knowledge. Comput. Secur. 10, 8, 749–764 (1991).

[Be09] Benjamin, A.S. et al.: Signal detection with criterion noise: applications to recognition memory. Psychol. Rev. 116, 1, 84 (2009).

[EK08] Elo, S., Kyngäs, H.: The qualitative content analysis process. J. Adv. Nurs. 62, 1, 107–115 (2008).

[GS89] Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. 18, 2, 141–153 (1989). https://doi.org/10.1016/0304-4068(89)90018-9.

[GS71] Glaser, B.S., Strauss, A.: A.(1967). The discovery of grounded theory. N. Y. 581–629 (1971).

[GL13] Glaser, J., Laudel, G.: Life with and without coding: Two methods for early-stage data analysis in qualitative research aiming at causal explanations. Forum: Social Qualitative Research, 14 (2). ISSN 1438-5627. (2013).

[He03] Hermand, D. et al.: Risk target: An interactive context factor in risk perception. Risk Anal. 23, 4, 821–828 (2003).

[HJ03] Hughes, J., Jones, S.: Reflections on the use of Grounded Theory in Interpretive Information Systems Research. In: Proceedings of the ECIS 2003 Conference. pp. 1–10 , Naples, Italy (2003).

[Is18] ISO/IEC: Information technology - Security techniques - Information security risk management. ISO/IEC, Geneva, CH (2018).

[Ja16] Jahan, N. et al.: How to conduct a systematic review: a narrative literature review. Cureus. 8, 11, (2016).

[Ja19] James, F.: A Risk Management Framework and A Generalized Attack Automata for IoT based Smart Home Environment. In: 2019 3rd Cyber Security in Networking Conference (CSNet). pp. 86–90 IEEE, Quito, Ecuador (2019). https://doi.org/10.1109/CSNet47905.2019.9108941.

[KT79] Kahneman, D., Tversky, A.: Prospect Theory: An Analysis of Decision under Risk. Econometrica. 47, 2, 263 (1979). https://doi.org/10.2307/1914185.

[Lu90] Luhmann, N.: Technology, environment and social risk: a systems perspective. Organ. Environ. 4, 3, 223–231 (1990).

[Me17] Mersinas, K.: Risk Perception and Attitude in Information Security Decision-making. Royal Holloway, University of London (2017).

[My09] Myers, M.: Qualitative research in business and management. Sage Publications Ltd, London (2009).

[No83] Nosofsky, R.M.: Information integration and the identification of stimulus noise and criterial noise in absolute judgment. J. Exp. Psychol. Hum. Percept. Perform. 9, 2, 299 (1983).

[RV19] Riesco, R., Villagrá, V.A.: Leveraging cyber threat intelligence for a dynamic risk framework. Int. J. Inf. Secur. 18, 6, 715–739 (2019).

[Ri20] Rios, E. et al.: Continuous Quantitative Risk Management in Smart Grids Using Attack Defense Trees. Sensors. 20, 16, 4404 (2020). https://doi.org/10.3390/s20164404.

[Sa88] Saaty, T.L.: What is the analytic hierarchy process? In: Mathematical models for decision support. pp. 109–121 Springer (1988).

[Sc19] Schreier, M. et al.: Qualitative Content Analysis: Conceptualizations and Challenges in Research Practice—Introduction to the FQS Special Issue" Qualitative Content Analysis I". In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research. (2019).

[Se20] Sektas-Bilusich, D. et al.: A Risk-Based Framework to Inform Prioritisation of Security Investment for Insider Threats. Int. J. Saf. Secur. Eng. 10, 1, 49–57 (2020). https://doi.org/10.18280/ijsse.100107.

[Sh20] Shakibazad, M., Rashidi, A.J.: New method for assets sensitivity calculation and technical risks assessment in the information systems. IET Inf. Secur. 14, 1, 133–145 (2020). https://doi.org/10.1049/iet-ifs.2018.5390.

[Hy12] hyeun-Suk, R. et al.: Unrealistic optimism on information security management. Comput. Secur. 31, 221–232 (2012).

[Te20] Teng, Y. et al.: Algorithm for quickly improving quantitative analysis of risk assessment of large-scale enterprise information systems. In: 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). pp. 2512–2515 IEEE, Chongqing, China (2020). https://doi.org/10.1109/ITNEC48623.2020.9085010.

[Ur09] Urquhart, C. et al.: Putting the 'theory' back into grounded theory: guidelines for grounded theory studies in information systems: Guidelines for grounded theory studies in information systems. Inf. Syst. J. 20, 4, 357–381 (2009). https://doi.org/10.1111/j.1365-2575.2009.00328.x.

[VD15] VDA: Information Security Assessment. Verband der Automobilindustrie (VDA), Berlin, Deutschland (2015).

[ZR20] Zhang, Y., Rao, Z.: Research on Information Security Evaluation Based on Artificial Neural Network. In: 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE). pp. 424–428 IEEE, Shenzhen, China (2020). https://doi.org/10.1109/AEMCSE50948.2020.00098.

[Zu20] Zuo, J. et al.: Comprehensive Information Security Evaluation Model Based on Multi-Level Decomposition Feedback for IoT. Comput. Mater. Contin. 65, 1, 683–704 (2020). https://doi.org/10.32604/cmc.2020.010793.

Open Identity Summit 2022

Further Conference Contributions

H. Roßnagel, C.H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 113

A user-centric approach to IT-security risk analysis for an identity management solution

Nicolas Fähnrich[1], Matthias Winterstetter[2], Michael Kubach[1]

Abstract: In order to build identity management (IdM) solutions that are secure in the practical application context, a holistic approach their IT-security risk analysis is required. This complements the indispensable technical, and crypto-focused analysis of risks and vulnerabilities with an approach that puts another important vector for security in the center: the users and their usage of the technology over the whole lifecycle. In our short paper we focus exclusively on the user-centric approach and present an IT-security risk analysis that is structured around the IdM lifecycle.

Keywords: identity lifecycle; user-centric; risk analysis; IT-security; cybersecurity; social engineering; identity management; IdM

1 Introduction

For identity management (IdM), the call to put the user into the center of the development efforts for new solutions is not new. Actually, it has already been one key aspect of Cameron's 7-laws of identity [Ca05]. Lately, the widely popular term Self-sovereign identity (SSI), going back to 10 postulated principles by Allen [Al17], puts user control into the center as well. Many see the future for digital identity in this concept and it is in mentioned in many high-profile initiatives, e.g. by the EU or the German government [Bu21d], [Bu21c]. Now it is not the topic of this paper to discuss the whether the implicit or explicit assumption by proponents of these claims that it is the failure to put the user into the center which is the reason why privacy friendly IdM solutions have failed on the market so far. However, if we pursue the path towards self-sovereign identity further to build systems that allow users to fully own and manage their identity without having to rely on a third party [Mü18] this user also should be put into the center of the IT-security risk analysis of such systems. Traditionally, IT-security risk analysis focuses on vulnerabilities of software, hardware, or network systems. The exploitation of humans as attack vectors via so called social engineering attacks is often neglected [BP16]. Generally, the research on social engineering is still in an early stage, when it comes formal definitions, attack frameworks and attack templates [MLV16]. If we look at common procedures in IT-security risk analyses, that follow standards like "ISO/IEC 27001", the NIST Cybersecurity framework [Na18] or the German IT-Grundschutz [Bu21b], we find a similar approach that starts off with a detailed

[1] Fraunhofer IAO, Nobelstraße 12, 70569 Stuttgart, firstname.lastname@iao.fraunhofer.de
[2] Universität Stuttgart, Institut für Arbeitswissenschaft und Technologiemanagement (IAT), Allmandring 35, 70569 Stuttgart, matthias.winterstetter@iat.uni-stuttgart.de

doi:18.18420/OID2022-09

documentation of the IT-infrastructure of a given application. In addition, all data processing operations are documented, including all data categories. Based on these data categories, person-related data is further investigated to derive the individual protection needs. In a following step, possible IT-security threats are identified by a catalogue comparison and potential threats are rated and documented. A risk is then derived based on the probability of occurrence of the identified threats and the protections needs of the processed person-related data. The underdeveloped field of social engineering in IT-security risk analysis might be one reason why IdM-projects tend to rely on the traditional approach to security risk analysis and make do with considering technical aspects, considering the human factor at most in an unstructured and superficial manner. Still, while a thorough analysis of vulnerabilities of software, hardware and network systems is certainly indispensable, is nevertheless incomplete. For an approach that puts users at the center of the control of such valuable data as identity information, we also need to put them at the center of our security risk analysis. In a research project to build an ecosystem of secure and trustworthy digital identities [ON22] we are currently putting our call to action. With this short paper we hope to enrich the discussion on user-centricity in IT-security research, in particular with focus on (self-sovereign) IdM, as a currently underdeveloped field. Through an early publication of the first results of our work we aim to collect valuable feedback than can guide further development efforts. Hence, the remainder of this short paper after this brief introduction presents the user centric approach to IT-security risk analysis we developed for the IdM research project. Then we conclude the paper with a discussion of the preliminary results, limitations, and next steps.

2 Proposed user centric approach

The technical side of IT-security risk analysis, such as the infrastructure design, the choice of security mechanisms, crypto protocols, and authentication methods is without a question of great importance for the overall security of applications. However, the relative disregard of the end user side can lead to significant security problems in practice use, especially in the case of IdM solutions where highly sensitive personal data are being handled, for example, to conclude contracts. Even more so, if the user is the only point of control without reliance on a third party as if in SSI-approaches. In fact, studies analysing the current state of IT-security and attacks show that users are heavily involved in the majority of cyberattacks, with multiple social engineering techniques being used [Bu21a], [G 21]. In the absence of an existing structured procedure for IdM solutions that focuses on the end user side in addition to the technical security aspects, we developed a new approach that builds on the identity lifecycle of Meints and Royer [MR08] (Figure 1).

This enables us to thoroughly analyse every process step within every IdM-lifecycle phase e.g., the interaction between identity providers and end users within the “Revision/Auditing” phase that can potentially be exploited by attackers to commit an attack on the IdM system. To identify relevant attack vectors we made use of Mitre’s Attack Pattern Enumeration and

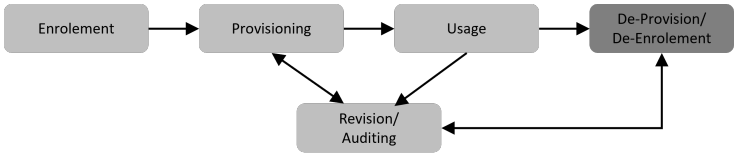


Fig. 1: Identity Lifecycle based on Meints and Royer [MR08]

Classification (CAPEC) [MI22] as it is actively managed, regularly updated and commonly regarded as a global knowledge base for IT-security threats.

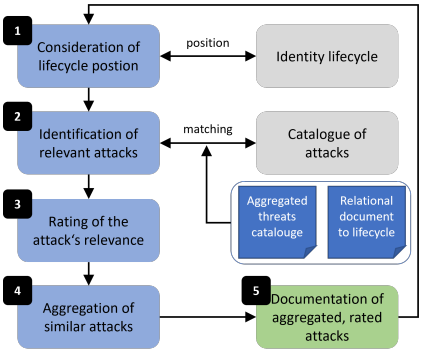


Fig. 2: Procedure for user-centric identification, rating, and documentation of relevant attacks

The developed procedure consists of the following steps and is illustrated in Figure 2:

1. Iterative consideration of one identity lifecycle phase at a time.
2. Identification of relevant attack vectors given in Mitre's Attack Pattern Enumeration and Classification (CAPEC). To assist with this process we created a catalogue of non-technical attacks and a relational document to link the threats to the identity lifecycle.
3. Rating of the attack's relevance within the respective phase of the identity lifecycle considering the statistical frequency of occurrence and ease of carrying out the attack.
4. Aggregation of similar attacks and their respective rating in relevance.
5. Documentation of aggregated and rated attacks for the respective identity lifecycle phase.

In step 1, we set the phase in the identity lifecycle, e.g., "Enrolement", that is to be considered in the following steps of the analysis. A research in the CAPEC catalogue follows in step 2. This is the central phase of the process and consists of matching relevant cyberattacks with the respective identity lifecycle phase. More detail on how this step is performed and supported will be given below. As a result, we obtain a list of relevant cyberattacks which

are then rated in step 3. In this rating we consider different factors such as the statistical frequency of the attack type and the ease for an attacker to carry out the attack. Since this procedure leads to a potentially high number of possible attacks, we aggregate these into corresponding categories in step 4. In step 5 we document the aggregated and rated attacks for the respective identity lifecycle phase. Then the next iteration of the procedure starts again with step 1 and the subsequent phase of the identity lifecycle. This is repeated until all phases of the identity lifecycle have been completed. As our end result, we receive a completed table of categorized and rated cyberattacks that are assigned to the respective phases of the identity lifecycle that should be considered in the architecture and the development of the underlying application. This procedure ensures a high degree of thoroughness since possible attacks are not only matched against the technological aspects of the IdM solution but the end user as potential attack vector is considered as well and this is done along the complete identity lifecycle in a structured form.

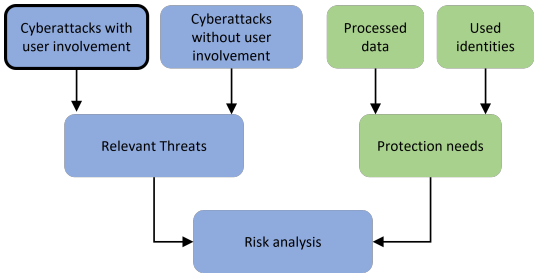


Fig. 3: Attack-based approach for the identification of cyberattacks with user involvement in the context of a risk analysis

This attack-based approach forms the basis for the identification of cyberattacks that require user interaction. It should be noted that this is only a subset of the total quantity of possible cyberattacks, since also technical attacks without user interaction have to be considered. When combining the identified attacks and complementing technical attacks with the protection requirements of the respective application, a comprehensive risk analysis is to be performed (Figure 3). To assist with the identification process of relevant attacks in step 2 of the described procedure, we require a concise overview of possible threats that could affect the system in consideration. To cover this aspect, we aggregated the most relevant non-technical threats, listed under the CAPEC Domains of Attack [MI22]. We focused on non-technical threats since these are usually relevant for user-centric attacks on a system. The threats were documented in an excel-table for easy review. In addition to aggregating the non-technical threats presented by CAPEC, we extended the provided information with a threat evaluation, a generic example of an attack and potential counter measures. The threat evaluations consist of the “kill chain” (description of the phases of an attack) phases in which the threats are most likely to occur, the required technical and social skills for execution, a rough evaluation of the required effort and the likelihood of occurrence. The generic example of the threat provided in the document is separated into flow one through three, describing the three stages of the threat in detail. The actual execution of an attack

regarding the threat may consist of more or less steps than the provided three, as they are only meant to give the reader a general idea of how an attack could occur. Regarding the potential counter measures we provide a list of viable options. We assigned the generic counter measures “user training” and “guidelines” to all threats, as these measures are among the most effective means for dealing with non-technical threats. In total, the document contains 16 non-technical threats (excluding the aggregated threats) for IdM systems. As user identities move through the different phases of the identity lifecycle they become more susceptible to some of the listed threats and less susceptible to others. While all threats are at least partially applicable during the “Usage” phase of the lifecycle, threats like identity theft become more relevant during the “Provisioning”, “Usage” and “Revision/Auditing” phases. Opposed to that, threats like spoofing and phishing can already be relevant during the “Enrolment” phase. To enable an easier identification of relevant threats in each phase of the lifecycle we additionally created a second, relational document to link threats to the lifecycle phases to show if they are applicable in the respective phase. A generic example for each applicable threat is provided for additional explanation thus allowing for an easier identification of threats. This secondary document is meant to serve as a bridge between the overview of the aggregated threats we provided and the lifecycle to ease the identification of relevant attacks in the second step shown in Figure 3.

3 Conclusion

IT-security is key to achieve trustable IdM solutions. Current development approaches focus primarily on technical security aspects although it has shown that end users are an important attack vector on IT-systems through social engineering attacks. Now that developments such as the trend towards SSI put the user in an even more responsible position, a failure to methodologically integrate the human factor into security considerations for the development of novel IdM systems creates an even bigger problem for their overall security. In this paper we therefore present a user-centric approach that leverages the identity lifecycle by on Meints and Royer. Using an iterative approach, each phase of the identity lifecycle is considered to match relevant attacks exploiting the human vector using Mitre’s CAPEC. As a result, we get a documentation of relevant and rated cyberattacks including adequate mitigation measures that can be used in the further development of the IdM solution. This short paper elaborates on our approach covering the non-technical attacks. The analysis of this subset of attacks certainly has complemented as shown in Fig. 3. Furthermore, other standards and guidelines cover non-technical aspects as well, however we think that our approach provides a structured procedure to evaluate the user’s role in attacks over the whole IdM-lifecycle into the risk analysis. This can potentially lead to a higher security level especially for end users. This short paper presents work in progress. We will test our approach as part of the risk analysis of the IdM solution developed in the ONCE research project [ON22] and optimize it further based on the lessons learned.

References

[Al17] Allen, C.: The Path to Self-Sovereign Identity, https://github.com/ChristopherA/self-sovereign-identity, 2017, visited on: 02/05/2022.

[BP16] Beckers, K.; Pape, S.: A serious game for eliciting social engineering security requirements. In: 2016 IEEE 24th International Requirements Engineering Conference (RE). IEEE, pp. 16–25, 2016.

[Bu21a] Bundesamt für Sicherheit in der Informationstechnik: Die Lage der IT-Sicherheit in Deutschland 2021, Bonn, 2021.

[Bu21b] Bundesamt für Sicherheit in der Informationstechnik: IT-Grundschutz: Informationssicherheit mit System, https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/it-grundschutz_node.html, Mar. 2021, visited on: 02/15/2022.

[Bu21c] Bundesministerium für Wirtschaft und Energie: Digitale Identität, https://api.bundesregierung.de/resource/blob/992814/1898280/d9819a40553a9543b9e8f3acb620b0c2/digitale-identitaet-neu-download-bundeskanzleramt-data.pdf?download=1, 2021, visited on: 02/15/2022.

[Bu21d] Bundesministerium für Wirtschaft und Energie: High-level scope (ESSIF) - EBSI Documentation - CEF Digital, https://ec.europa.eu/digital-building-blocks/wikis/pages/viewpage.action?pageId=379913698, 2021, visited on: 02/15/2022.

[Ca05] Cameron, K.: The Laws of Identity, https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf, 2005.

[G 21] G DATA CyberDefense AG: Cybersicherheit in Zahlen, Hamburg, 2021.

[MI22] MITRE: Common Attack Pattern Enumeration and Classification (CAPEC): A Community Resource for Identifying and Understanding Attacks, https://capec.mitre.org/, 2022, visited on: 02/10/2022.

[MLV16] Mouton, F.; Leenen, L.; Venter, H. S.: Social engineering attack examples, templates and scenarios. Computers & Security 59/, pp. 186–209, 2016.

[MR08] Meints, M.; Royer, D.: Der Lebenszyklus von Identitäten. Datenschutz und Datensicherheit 32/3, p. 201, 2008.

[Mü18] Mühle, A.; Grüner, A.; Gayvoronskaya, T.; Meinel, C.: A survey on essential components of a self-sovereign identity. Computer Science Review 30/, pp. 80–86, 2018.

[Na18] National Institute of Standards and Technology: Cybersecurity Framework Version 1.1. https://www.nist.gov/cyberframework/framework, 2018, visited on: 02/15/2022.

[ON22] ONCE: Online einfach anmelden, https://once-identity.de/, 2022.

Adversary Tactics and Techniques specific to Cryptocurrency Scams

Andrea Horch [1], Christian H. Schunck [2] and Christopher Ruff [3]

Abstract: At the end of the year 2020, there was a steep uptrend of the cryptocurrency market. The global market capitalization of cryptocurrencies climbed from 350 billion US$ in October 2020 to almost 2.5 trillion US$ in May 2021 and reached 3 trillion US$ in November 2021. Currently, there are more than 17,600 cryptocurrencies listed on CoinMarketCap. The ample amount of money within the market attracts investors as well as scammers and hackers. Recent incidents like the BadgerDAO hack have shown how easy it is to steal cryptocurrencies. While all the standard scamming and hacking techniques such as identity theft, social engineering and web application hacking are successfully employed by attackers, new scams very specific to cryptocurrencies emerged, which are the focus of this paper.

Keywords: cryptocurrency, scam, distributed ledger technology, blockchain, digital wallet, digital identities

1 Introduction

The charts of Bitcoin and altcoins (all other coins, which offer an alternative to Bitcoin) on CoinMarketCap show two huge continous uptrends (bullruns). The first very steep uptrend took place in 2017/2018, the other one started at the end of 2020 and is still ongoing. Even though scammers have always been active in cryptocurrency the amount of scams and hacks have increased sinificantly with the global marketcap. According to [Sta20] the value of cryptocurrencies lost to security threats increased nine-fold between 2020 and 2021. This paper gives an overview novel kind of scams, which are very specific to the cryptocurrency ecosystem. In the following we give a brief overview of the literature, introduce cryptocurrency specific terminology, present factors that make attacks cryptocurrency specific and give examples of such attacks and then discuss the results in the conclusion.

Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO, Nobelstrasse 12, 70569 Stuttgart, Germany, firstname.lastname@iao.fraunhofer.de

[1] https://orcid.org/0000-0001-9384-316X

[2] https://orcid.org/0000-0002-7917-8180

[3] https://orcid.org/0000-0003-0484-4131

cbe doi:18.18420/OID2022-10

2 Literature Review

Security problems, hacks and scams related to cryptocurrencies and relevant tools have been reported in several papers. The earliest classification of scams in the context of bitcoin was carried out in [VM15] where scam have been classified into four main categories: high yield investment programs (HYIPs) or Ponzi schemes, mining scams, scam wallets and exchange scams. The authors of [Ba21] provide a more recent overview of scams and present a fairly broad taxonomy, which does not fulfil rigorous requirements such as proposed in [NVM13]. Some of the identified scams have only a loose link to cryptocurrencies. For example, typical ransomware attacks are included in the taxonomy, where the role of cryptocurrency is limited to being a means for ransom payment. Other work closely investigate smart contracts which are very deeply linked to the technologies underlying cryptocurrencies: attackers both exploit vulnerabilities in existing smart contracts [ABC17] and engineer apparently vulnerable smart contracts with hidden traps as honeypots [TSS19]. The study presented in [Tr22] reviews the current state of knowledge on kinds of existing cryptocurrency fraud. It provides a raw classification scheme and definitions of the frauds identified, but does not give any hints on the cryptocurrency-specific attributes of the presented frauds.

3 Terminology

Coins, Tokens, NFTs: According to [Le22] a cryptocurrency is a digital measure of value that can be tracked and transferred without the need of an intermediate authority (e.g. bank or government). Cryptocurrencies are built on and exist on a network called blockchain. A blockchain is an unbounded and immutable (append-only) digital ledger, which stores the information on every transaction made on a network in the form of linked blocks [Fo22]. The native cryptocurrency of a blockchain is called “coin”. Other currencies, which are not the native currency of a blockchain, but built on it, are called “tokens”. A non-fungible token (NFT) is a digital asset on a blockchain. The use of a blockchain allows to prove the authenticity and ownership of the NFT [Ri21]. An NFT is not a currency as currencies are not unique, e.g. all Bitcoins have the same value. NFTs are unique and e.g. used for artwork, where every artwork is unique and has a different value [Ri21].

Digital Wallets: Digital wallets are software applications used to interact with a cryptocurrency or blockchain. Wallets allow viewing balances, making transactions and other interactions with the underlying blockchain (i.e. staking, using smart contracts, etc.) [SSB20]. Digital wallets are a concept to more easily interact with the public key cryptography (PKC) functionality that is the basis of most blockchains and digital ledgers and store the key pair required to access and transfer the funds on a blockchain. The public key serves as an address whereas the private key is used as a password and should never be shared with anyone. Wallets facilitate the pairing of private and public keys and allow users to sign transactions using their private keys. The cryptographic function for building the key pairs allows to generate a public key from a private key, but not vice versa.

Cryptocurrencies use 256bit numbers as private keys, which are up to 77 figures long and cannot be easily remembered by humans. A wallet software often provides 12 to 24 word key phrases called "seed phrases" generated out of 2,048 words of a dictionary [Me22] which can be used to reconstruct public keys. However, once a private key is inaccessible it cannot be recovered and the corresponding funds become inaccessible as well. Since digital wallets can additionally hold digital attributes and certificates, the novel scam techniques are also relevant in the context of securing self-sovereign and decentralized identities based identity management schemes.

Smart Contracts: A smart contract is machine readable code stored on a blockchain network or distributed ledger. The contracts are self-verifying, self-executing and tamper resistant. Storage, execution, computation and documentation are handled by the underlying network, removing the need for a single trusted third party. Thus, smart contracts allow one or more parties to enter agreements or agree on certain actions defined by the contract in a transparent and "trustless" way. There are numerous use cases for smart contracts, such as transparent autonomous supply chain documentation, financial services or real estate transactions and documentation [MPJ18].

Decentralized Exchanges: In contrast to centralized exchanges (CEX), Decentralized Exchanges (DEX) don't rely on a central authority that has custody of the transacted funds, tokens or coins but instead allows users to (mostly anonymously) transact peer-to-peer using smart contracts, while still having control of their private keys. The benefit of having full control over the funds often comes with trade-offs like scalability, low liquidity, high transaction fees, price slippage, front running and missing regulatory compliance [Ts20].

Decentralized Finance: Decentralized Finance (DeFi) is a generic term describing financial technologies and services based on distributed ledgers (e.g. blockchains) and smart contracts. DeFi cuts out the middleman (i.e. financial institutions) and instead provides said services based on smart contracts stored on an immutable distributed ledger. Depending on the services, this can eliminate fees, shorten processing times and does not require approval from third parties such as banks.

Airdrops: Airdrops involve the distribution of tokens or coins to wallets or addresses for free. This is often used for marketing and promotion purposes, to increase visibility and usage of a coin, or an underlying platform. To be eligible for an airdrop, users often have to complete certain tasks, such as following or sharing a project on social media or using a certain platform while some airdrops do not require any user interaction.

4 Novel scams using cryptocurrency-specific approaches

We analysed hundreds of recent cryptocurrency scams and attacks and identified important factors, which make an attack or scam cryptocurrency-specific: these scams use functionalities of a blockchain (1) to distribute coins/tokens (e.g. airdrop scams), (2) to move coins/token on the blockchain (e.g. honeypot to drain wallets) or (3) to manipulate

trades on the blockchain (e.g. Sandwich Attack). There are attacks and scams using blockchain technologies such as smart contracts (e.g. exit scams). We still regard these scams to be cryptocurrency specific but note that the blockchain-specific part could be substituted with a not blockchain-specific technology and thus the scam could also work outside of the cryptocurrency ecosystem.

Airdrop Scams: Airdrop scams are phishing tokens airdropped to random wallets of a blockchain in order to lure the wallet owners to phishing websites of fake exchanges by showing a high conversion rate for the airdropped token [Tu22]. In August 2021 the scammers airdropped $SHIB tokens on Binance Smart Chain (BSC) to random users showing up in their wallets to be worth around 1,000 USD. Wallet owners who visited the scam website were asked to approve their wallets to swap the tokens. Approving the smart contract gave the scammers access to drain the wallets and the funds were stolen [Bs21].

Scam Tokens: Decentralized exchange platforms like Uniswap (https://uniswap.org/) allow an open and free listing of new tokens, which benefits new projects to launch fast, and at low costs. These advantages for new projects also help scammers to run fake coins and scam projects with low efforts [Ma21]. Scammers use different approaches to get the cryptocurrencies of victims. A very popular way is the creation of fake token imitations where the scammers search for new legitimate tokens on decentralized exchanges and create a similar token listing, e.g. $SHIB and $SH1B.

Smart Contract-based Scams: An example for a smart contract based scam is a smart contract-based honeypot where scammers post private keys or seed phrases in chatrooms (e.g. on telegram). The post looks like a mistake by an unexperienced user, but it was posted on purpose by a scammer. Honeypot wallets hold a significant number of tokens, which can only be moved by paying a fee using a corresponding "gas" token (e.g. $ETH on Ethereum). The victims who decide to exploit the ostensible user's mistake are thus lured into spending gas tokens in order to move the user's tokens to their own wallet. But an underlying smart contract foresees that the gas tokens sent will be instantly moved to a wallet owned by the attacker who created the smart contract behind the honeypot [Mc18].

Sandwich Attacks: Bots of malignant traders search for pending large trading transaction of other traders on the blockchain. A bot sniffs out a transaction and front-runs the victim trader by purchasing the same asset as the victim. The front-run is possible by paying a higher gas fee, which gives a higher priority in the transaction queue. By placing the front-run trade the attacker manipulates the price of the asset and the victim suffers a higher slippage (price difference between the point in time a transaction was submitted and the time the transaction is confirmed) for its transaction and pays a higher price for the purchase. The attacker now back-runs the victim's transaction and then gets a higher price for selling the asset. The attack is called sandwich attack because the attacker front-runs and back-runs the original pending transaction, which is sandwiched in between [Da21].

5 Discussion and Conclusion

In this paper we presented a number of recent attacks schemes specific to cryptocurrencies. The field develops very quickly, and new scams emerge almost daily. Therefore this analysis is only preliminary and cannot be regarded as comprehensive and complete. Furthermore cryptocurrencies facilitate many conventional fraud schemes and scams due to the difficulty of tracing cryptocurrencies. Our medium-term goal is to use an approach similar to ATT&CK matrices to comprehensively present and analyse all the tactics used in attacks on cryptocurrencies in a web-based format so that it can be extended and updated as new techniques emerge and tactics become more elaborate.

6 Acknowledgments

This work was partially supported by the project ONCE (01MN21003F, once-identity.de), funded by the German Federal Ministry for Economic Affairs and Climate Action.

Bibliography

[ABC17] Atzei, N.; Bartoletti, M; Cimoli, T.: A Survey of Attacks on Ethereum Smart Contracts SoK, Berlin, Heidelberg, vol. 10204, pp. 164–186, 2017.

[Ba21] Bartoletti, M. et.al.: Cryptocurrency Scams: Analysis and Perspectives. In: IEEE Access, vol. 9, pp. 148353–148373, 2021.

[Bs21] BSCScan.com: BSCScan SHIB Scam, https://bscscan.com/token/0xab57aef3601cad382aa499a6ae2018a69aad9cf0#comments, accessed: 21/02/2022.

[Da21] Das, A.: DEFI Sandwich Attack Explain. https://medium.com/coinmonks/defi-sandwich-attack-explain-776f6f43b2fd, 2021, accessed: 18/02/2022.

[Fo22] Fool.com: “What Are Crypto Tokens? https://www.fool.com/investing/stock-market/market-sectors/financials/cryptocurrency-stocks/crypto-tokens/, accessed: 21/02/2022.

[Le22] Ledger.com: WHAT IS CRYPTOCURRENCY? https://www.ledger.com/academy/basic-basics/about-crypto/what-is-cryptocurrency, accessed: 21/02/2022.

[Ma21] Maksimenka, I.: How to Identify and Avoid Uniswap Scams. https://coinmarketcap.com/alexandria/article/how-to-identify-and-avoid-uniswap-scams, 2021, accessed: 18/02/2022.

[Mc18] McIntosh, R.: Hack the Hackers: ‘Honeypot’ Crypto Scam Targets Would-Be Coin Thieves. https://www.financemagnates.com/cryptocurrency/news/hack-

hackers-honeypot-crypto-scam-targets-coin-thieves/, 2018, accessed: 18/02/2022.

[Me22] Medium.com: The ultimate guide to private keys and recovery seed phrases. https://medium.com/coinmonks/the-ultimate-guide-to-crypto-private-keys-and-recovery-seed-phrases-556ae60e59e7, 2022.

[MPJ18] Mohanta, B. K.; Panda, S. S.; Jena, D.: An Overview of Smart Contract and Use Cases in Blockchain Technology. In: 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–4, 2018.

[NVM13] Nickerson, R.C.; Varshney, U.; Muntermann, J.: A method for taxonomy development and its application in information systems. In: European Journal of Information Systems, vol. 22, no. 3, pp. 336–359, 2013.

[Ri21] Rizvi, S.: A Complete Beginner's Guide to NFTs, https://trustwallet.com/blog/a-complete-beginners-guide-to-nfts, 2021, accessed: 21/02/2022.

[SSB20] Suratkar, S.; Shirole, M.; Bhirud, S.: Cryptocurrency Wallet: A Review. In: 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), pp. 1–7, 2020.

[Sta20] Statista.com: Total value of cryptocurrency lost to and recovered from theft and other attacks between March 2020 and February 2022. https://www.statista.com/statistics/1285057/crypto-theft-size/, accessed: 21/02/2022.

[Tr22] Trozze, A. et.al.: Cryptocurrencies and future financial crime. In: Crime Science 11, BioMed Central Ltd, article no. 1, 2022.

[Ts20] Tsai, W. -T. et.al.: Decentralized Digital-Asset Exchanges: Issues and Evaluation. In: Proceedings of the 2020 3rd International Conference on Smart BlockChain (SmartBlock), pp. 1–6, 2020.

[TSS19] Torres, C. F.; Steichen, M.; State, R.: The Art of the Scam: Demystifying Honeypots in Ethereum Smart Contracts. In: Proceedings of the 28th USENIX Conference on Security Symposium, USA, pp. 1591–1607, 2019.

[Tu22] Tunny, J.: What are Scam Airdrop Tokens on Binance Smart Chain and Why Are They So Prevalent? https://www.bsc.news/post/what-are-scam-airdrop-tokens-on-binance-smart-chain-and-why-are-they-so-prevalent, accessed: 18/02/2022.

[VM15] Vasek, M; Moore T.: There's No Free Lunch, Even Using Bitcoin: Tracking the Popularity and Profits of Virtual Currency Scams. In (Böhme, R., Okamoto, T. ed.): Financial Cryptography and Data Security, FC 2015, Lecture Notes in Computer Science, Springer, Berlin, , vol. 8975, pp. 44-61, 2015.

Preservation of (higher) Trustworthiness in IAM for distributed workflows and systems based on eIDAS

Hermann Strack[1], Sebastian Karius[2], Marlies Gollnick[3], Meiko Lips[4], Sandro Wefel[5], Robert Altschaffel[6]

Abstract: The secure digitalisation of distributed workflows with different stakeholders (and trust relationships) using systems from different stakeholder domains is of increasing interest. Just one example is the workflow/policy area of student mobility. Others are from public administration and from economic sectors. According to the eIDAS regulation, eID and trust services (TS) are available across EU - upcoming also EUid & wallets (eIDAS 2.0) - to improve security aspects (providing interoperability or standards). We present some security enhancements to maintain higher trustworthiness in Identity and Access Management (IAM) services for different policy areas with mandatory, owner-based and self-sovereign control aspects - based on eIDAS and different standards and the integration of views/results from deployed or ongoing projects (EMREX/ELMO, Europass/ EDCI, eIDAS, EUid, Verifiable Credentials, NBP initiative, OZG implementation, Self-Sovereign Identities SSI, RBAC, ABAC, DAC/MAC, IPv6) and a trustsistor.

Keywords: eIDAS eID & TS (2.0), EUid, IAM, LoA, authentication, access control, notarisation, NBP initiative, OZG, Self-Sovereign Identities SSI, RBAC, ABAC, DAC/MAC, IPv6, trustsistor

1 Introduction

Digitization of workflows in different fields like Education, Public Administration, Health Services and Business needs for compliance realizations, checks and balances according to their policies. This includes the implementation and integration of security and trust services, as well as trusted entities/roles, using methods of security by design and management. Obviously, strong authentication and access control would improve the security against different threats and vulnerabilities from outside or inside the domains or interest groups involved. This includes, for example, exploiting vulnerabilities to obtain identities, roles or other data, or abusing user roles and administrator rights.
Important intermediate as well as final results at workflow level are documents, certificates and diplomas, with security requirements for integrity, authenticity and privacy, which also meet the requirements for reliable archiving. The integration of PKI

[1] Hochschule Harz, FB Automatisierung und Informatik, 38855 Wernigerode, hstrack@hs-harz.de
[2] Hochschule Harz, FB Automatisierung und Informatik, 38855 Wernigerode, skarius@hs-harz.de
[3] Hochschule Harz, FB Automatisierung und Informatik, 38855 Wernigerode, mgollnick@hs-harz.de
[4] Hochschule Harz, FB Automatisierung und Informatik, 38855 Wernigerode, mlips@hs-harz.de
[5] Martin-Luther-Universität Halle-Wittenberg, Institut für Informatik, 06120 Halle (Saale), sandro.wefel@informatik.uni-halle.de
[6] Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, 39106 Magdeburg, Robert.Altschaffel@iti.cs.uni-magdeburg.de

cbe doi:18.18420/OID2022-11

based eIDAS, eID and TS would support securing the workflows and their policies and roles accordingly. This applies in particular in the IAM field, with access control policies and architecture elements in an EU wide interoperable resp. standardized manner. See ETSI[7] / TR BSI[8] for further new developments, e.g. eIDAS 2.0. see [KSSR20; KSSK20]. In chapter 2 we present the current status of the KOLIBRI NBP project, in which the authors' institutions are involved (BMBF funded). In chapter 3, we provide an outlook for security improvements in various policy areas. We will apply our experiences from implementing portions of the National Educational Platform (NBP) with Level of Assurance/LoA "high" to additional policy and IAM protection areas, including network segmentation, workflow/access controls based on trust and separation of duties (SOD).

2 National Educational Platform Initiative (NBP)

The project "KOLIBRI" has implemented a prototype for the National Educational Platform (NBP) in Germany based on open source and standards. In addition, important eIDAS components got successful security evaluations (e.g. Common Criteria ISO 15408). All types of educational institutions are enabled to connect to the platform in a secure and privacy-preserving manner, also via standards (ongoing) on metadata level.

The research prototype of the project "KOLIBRI" implements the following features: Security & Privacy (regarding eIDAS/eID & TS standards, GDPR, OZG[9]), an identity broker and authorisation system with Single Sign-On (SSO), central collaboration services, connectors/metadata for decentralised Identity Management Systems and Identity Providers (IDP), connection to user wallets (with SSI/eIDAS 2.0 functions), and connections to EU services and standards: EMREX/ELMO, Europass/EDCI/VC [Min17]. In particular, the integration of SSO (Single Sign-On) by „KOLIBRI" takes into account the different levels of assurance (LoA) for the strength of authentication security according to the EU eIDAS regulation (LoA: low, substantial, high).This is important for cross-domain user integration and SSO, also at LoA "High" using eID. More additional attributes such as group membership can be transmitted.

A central Identity Broker enables the connection of the identity providers (IDP) of the satellite systems of the education providers. In order to enable citizens without special educational membership to have secure identities with full legally binding at the document transmission level, the login was also connected via a governmental eID service provider with eID card enabled login (OZG-Nutzerkonto). This can be used for legally binding

[7] https://www.etsi.org/newsroom/news/1111-2016-07-etsi-publishes-european-standards-to-support-eidas-regulation
[8] https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03130/TR-03130_TR-eID-Server_Part3.html
[9] OZG – Online-Zugangsgesetz/Online Access Law, OZG-Nutzerkonten: https://www.onlinezugangsgesetz.de

processes between state educational institutions, authorities and the platform, without extra qualified electronic signature (QeS).

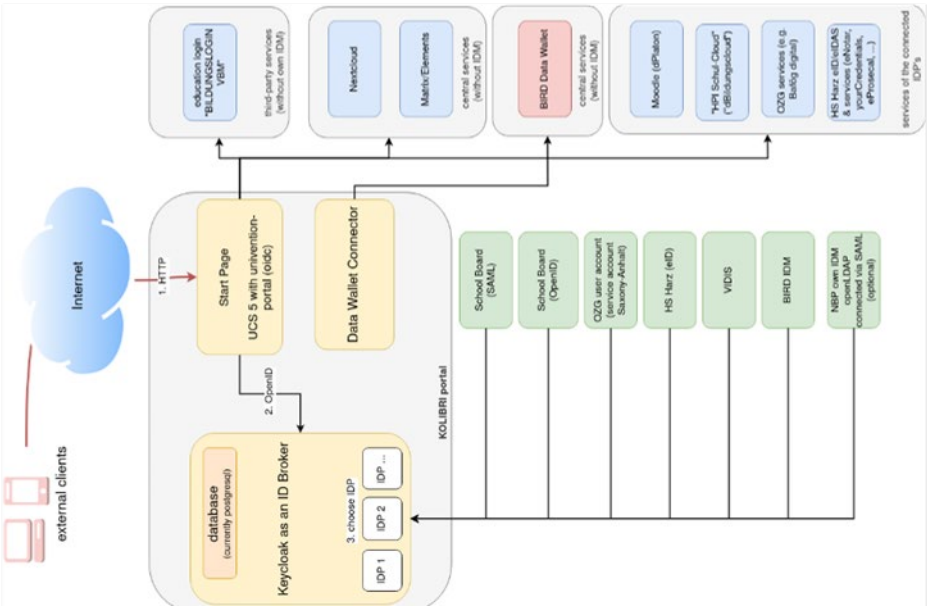


Fig. 1: KOLIBRI NBP: Keycloak as an ID Broker and eID/OZG services as IDP

The applications of the Harz University of Applied Sciences connected in the KOLIBRI platform use the online features of the German ID card to provide various services. These services are: eProsecal - provides highly secure authenticated access (LoA "high") for various university services/actors through an ID card-based logon and account process, including fully legally binding. Processes with multiple multi-user/role references (n:m) can also be mapped. Users have access to the data of all released processes via their eProsecal basic account and can also securely share them with other users and even with authorities in a legally binding manner (eID-based sharing). This results in a SSI (Self-Sovereign Identity) wallet function, without blockchain integration, based on the properties of the eID system in Germany. eInternship – internship management/contracts between university and company; eTor/eTestate - registration/attendance management for exams and lab practicals; eColloquium - signed colloquium exam forms/certificates (ELMO/EDCI/VC); eNotar - offers the possibility to provide documents with a qualified signature in a legally binding secure manner (public service laws in DE VwVfG §3a/§33, [SBKO19]). These services can be combined with further eIDAS-based services such as time-stamping and long-term storage services with oversignature (BSI Standard TR ESOR[10], eIDAS Preservation Standard). In addition, EU-wide time-dependent

[10] https://bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeRichtlinien/TR03125/BSI_TR-ESOR-LEIT.pdf

cybersecurity crypto requirements are covered, with the accompanying cryptoalgorithm management (EU SOGIS) providing a further basis for trust.; YourCredentials – extend the eNotar principle to the authentication and notarisation of derived identities and attributes (e.g. by RBAC/ABAC, [AHAZ19]) from their trust domains, e.g. for wallets for eIDAS 2.0 / EUid[11]. This enables the handling of multiple identities of a person arising from different phases/providers in that person's life. The notarisation of identity assignment by the trust service is extensible to relationships of related identities such as parents and children.

3 Protection of workflows/roles/systems via IAM & Trustsistor

With current technology the use of TPM attestations/attributes at IAM would enable additional higher LoA contexts, including for mandatory control policies across domain/system boundaries. In context of our NBP prototype, it can be used for enhanced protection of important workflow roles like eNotar or system administrator roles, also in scenarios where eID is not available or necessary [We22, Kl09, JA09]. Additionally, it is important to protect workflows & trusted roles (e.g. eNotar roles as well as system administrator roles and RBAC/ABAC control schemes), entities, documents and systems against attacks on network or systems vulnerability levels, especially hacking from outside and inside, or misuse of separation of duties (SOD) [MZNO19]. Important measurements are information flow protections and network segmentations based on classifications of networks, entities and systems using firewalls and data diodes[12] [BJBR14], see Fig. 2 (inspired by privacy/BLP/MAC/MLS policies). But to protect additionally against IAM attacks (bypassing), it could be combined with different LoA levels for IAM. Therefore, also Mandatory LoA IAM attributes (cryptografically protected/binding, e.g. by MACs, derived/based e.g. on YourCredentials notarisations, could be securely added to protocol messages by using sub-header principles as well as on document level. This can be done in an analogous manner to IPv6[13] and would be worth exploring for enhanced and extended authentication and access control layers based on firewall, data diodes, and access control components, e.g. for improving ZeroTrust[14] schemes. Therefore, we introduce the concept of a „Trustsistor“ TSO component that is integrated, e.g. into firewall or proxy components, and reinforces trust relationships by adding trust attributes of a TSP to IP flows, e.g. between client and server as additional IAM (mandatory) access control information (MACI contexts). The notion is similar in some sense to the "transistor concept" in electrical flows. This means, we would differentiate between a service user SU, a service provider SP and service access controller SPC, as well as a trust service provider TSP notarising ACI[15] trust attributes TACI, e.g., by signatures/MACs.

[11] https://ec.europa.eu/commission/presscorner/detail/de/IP_21_2663
[12] https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6949883
[13] https://www.ietf.org/blog/ipv6-internet-standard/
[14] https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf
[15] ACI: Access Control Information

Further we propose the components “Trustsistor-Injector (TSOI)” for injecting trust identifiers/labels into the IP flow on the part of the service client using e.g. IPv6 sub-headers and the “Trustsistor-Controller (TSOC)” for checking the required TACI attributes on the part of the service provider/controller according to a TACI access policy. By the way: for better multilateral system integrity security the TSOI/TSOC components should be protected by TPM. Based on the TSO model, secure implementation of access control policies can be done with additional TACI attributes on IP flows. The research conducted here was partly funded by 3 EU funded projects under the umbrella of ”CyberSec LSA”[16] (EFRE).

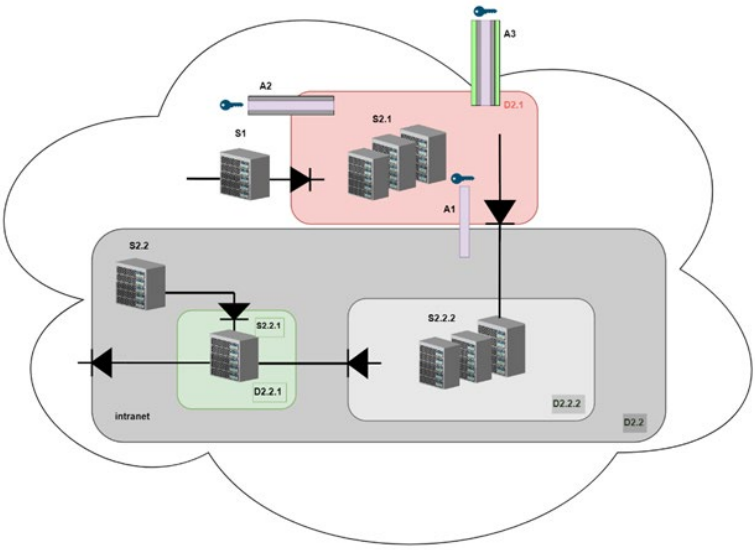


Fig. 2: Combined protections of privacy/flows/MAC and LoA Auth./IAM/trusts level at domains

4 Conclusion

The development of a prototype of the National Education Platform NBP revealed strengths and weaknesses of current Single Sign On (SSO) solutions. We showed that the use of eID-based authorization (LoA high) can be usefully employed in the area of SSO in the context of IAM. Using HW can significantly improve the security of platforms such as NBP and also the simplicity of authentication, since in best cases only a few strong IAM systems are needed. To further prevent security vulnerabilities such as access forgery, spoofing, leakage, etc. at the network transmission and security layer, we have outlined how the use of data diodes and network packets marked with Trust-ACI attributes

[16] https://cslsa.de

can preserve the security gained through strong authentication in conjunction with TACI notarisations at the network layer. This can be done by combining appropriate firewall rules and Trustsistor TSO injection and Trustsistor TSO controller components. Thus, the authorization defined at the IAM level is extendable by (mandatory) Trust Attributes (also LoA high), also at the network level.

Bibliography

[AHAZ19] Aftab, M.U.; Qin, Z.; Hundera, N.W.; Ariyo, O.; Zakria; Son, N.T.; Dinh, T.V. Permission-Based Separation of Duty in Dynamic Role-Based Access Control Model. *Symmetry* 2019, *11*, 669. DOI: 10.3390/sym11050669

[BJBR14] Bhatkalkar, B. J.; Ramegowda: A Unidirectional Data-flow Model for Cloud Data Security with User Involvement during Data Transit, 2014 International Conference on Communication and Signal Processing, IEEE Explore, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6949883

[EU14] EU: Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, 2014.

[HH20] Hühnlein, D.; Hühnlein, T.; Hornung, G.; Strack, H. (2020): Towards Universal Login. In: *LNI (Open Identity Summit 2020)*, 193–200. DOI: 10.18420/ois2020_18

[Kl09] Klenk, A.; Kinkelin, H.; Eunicke, C.; Carle, G.. 2009. Preventing identity theft with electronic identity cards and the trusted platform module. In *Proceedings of the Second European Workshop on System Security* (*EUROSEC '09*). ACM, NY, USA, 44–51. DOI: 10.1145/1519144.1519151

[KSSK20] Kusber T.; Schwalm, S.; Shamburger K.; Korte U.: Criteria for trustworthy digital transactions - blockchain/DLT between eIDAS GDPR, data and evidence preservation, In: *LNI (Open Identity Summit 2020)*, DOI: 10.18420/ois2020

[KSSR20] Kubach, M.; Schunck, C. H., Sellung, R. ; Roßnagel, H.: Self-sovereign and Decentralized identity as the future of identity management?. In: *LNI (Open Identity Summit 2020)*, 35-47. DOI: 10.18420/ois2020_03

[Min17] Mincer-Daszkiewicz, J.: EMREX and EWP offering complementary digital services in the higher education area, *Proceedings of EUNIS*, Münster, 2017.

[SBKO19] Strack, H.; Bacharach, G.; Klinner, S., Otto, O.; Schmidt, A.: eIDAS eID & eSignature for HEI/EDU Applications - eIDAS eID & eSignature based Service Accounts at University environments for crossboarder/domain access. In: *European Journal of Higher Education IT* 2019-1 https://www.eunis.org/erai/2019-1/

[We22] Web Authentication: An API for accessing Public Key Credentials, https://www.w3.org/TR/webauthn-2/#sctn-attestation, accessed: 21/02/2022.

H. Roßnagel, C.H. Schunck, S. Mödersheim (Hrsg.): Open Identity Summit 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 131

Online tool for matching company demands with IT-security offerings

Nicolas Fähnrich[1], Heiko Roßnagel[1]

Abstract: Small and medium sized companies (SMEs) are often insufficiently protected against cyberattacks although there is a wide range of cybersecurity guidelines, products and services available. In this paper, we present an online tool to support SMEs in improving their IT-security level by enabling them to identify critical business processes and to identify the most pressing protection needs by using a lightweight value chain-based approach. For using the online tool, no expert knowledge of the company's IT-infrastructure or implemented IT-security measures is required, since no assessment of cybersecurity threats but of the impact of potential damage scenarios on business processes is carried out. Based on a generated set of recommendations, companies are provided with suitable IT-security measures and corresponding offerings in a prioritized order. These offerings include services and products to implement the given recommendations.

Keywords: IT-security; expert system; value chain; bayesian network; SME; damage scenarios

1 Introduction

The ongoing trend towards digitalization enables companies to slim down processes, shorten response times and save costs. In contrast, there is an increasing threat from cyberattacks, which can cause significant damages to companies [BS20], [G 21]. Although there are numerous guidelines available to improve the IT-security level in companies like the German BSI IT-Grundschutz [BS21] or ISO/IEC 27001 [IS13], in practice the IT-security level is often insufficient, especially among SMEs [BS20], [Bs11], [Hi17]. This circumstance cannot be explained with a lack of IT-security offerings. We assume that the high complexity of existing guidelines and the heterogenous and wide range of IT-security products/services on the market combined with a low willingness to pay lead to a high entry barrier for companies that have so far invested little in IT-security. It's a challenge for companies to identify their protection needs and appropriate security measures, especially if they have not experienced any major damage from IT-security incidents so far. This paper, thus, presents the expert system "Smart Matching" to support SMEs in improving their IT-security level by giving prioritized recommendations and suitable technical/organizational measures for implementation including adequate offerings from a curated database. The expert system doesn't contain a security analysis based on implemented IT-security measures, but was rather designed in a way to be used by various company representatives without expert knowledge in IT-security and to offer a low-threshold entry to suitable IT-security solutions.

[1] Fraunhofer IAO, Nobelstraße 12, 70569 Stuttgart, firstname.lastname@iao.fraunhofer.de

doi:18.18420/OID2022-12

Companies gain insights into which areas of their business are particularly threatened and should be protected. The expert system was developed as part of the German national project TISiM [TI22b] which is financed by the Federal Ministry for Economic Affairs and Energy (BMWi) and is available free of charge as part of the webapp Sec-O-Mat [TI22a].

2 Related Work

There are numerous IT-security standards, guidelines and products/services available, which results in the challenge to identify suitable solutions that meet the respective company's requirements. Conventional approaches in IT-security consulting projects that follow standards like ISO/IEC 27001 [IS13] or BSI IT-Grundschutz [BS21] to support companies in this decision process start off with a documentation of the IT-infrastructure. Furthermore, all business processes with the involved IT-systems and data types are documented. In a following step, possible IT-security threats are identified and potential threats are rated and documented, whereupon a risk is derived. The current state of already implemented IT-security measures is documented and then compared to a target-state that is determined by the risk analysis and corresponding catalogues. These approaches are well established in practice and result in a detailed analysis of the specific IT-security demand including suitable recommendations for implementation. However, this procedure is resource-intensive and requires IT-security experts and the cooperation of company representatives who are familiar with the company's business processes and IT-infrastructure. Furthermore, a key element in various standards is the consideration of possible IT-security threats that lead to suitable mitigation measures, however this can potentially lead to misperceptions if this is not carried out by experts. For these reasons, the existing conventional approaches are not suitable for the given problem of an easy, low-threshold entry into IT-security and a consideration of possible damage scenarios of cyberattacks and their impact on business processes may be helpful to derive suitable measures. However, conventional approaches are indirectly used in the context of creating the knowledge base of our solution.

3 An expert system to assess the impact of damage scenarios and identify suitable IT-security measures

We have developed an expert system that enables SMEs to identify appropriate IT-security measures with little expense and without expert knowledge on part of the companies following a different approach by considering possible damage scenarios which can occur as a result of IT-security incidents. As part of our work in TISiM, it was our task to provide SMEs a low-threshold entry into IT-security for different company representatives without IT-security background, that can only assess the relevance of certain cybersecurity threats to a limited extent. They are, however, able to assess the impact of certain damage scenarios on their company. We therefore modelled the interrelationships between damage scenarios that affect the company's business processes and suitable recommendations for IT-security

to avoid them. We refrain from ascertaining the actual state of implemented IT-security measures as this cannot be determined reliably without expert knowledge. When developing a process model for supporting SMEs to comply with the EU-GDPR in earlier works, we already found many similarities regarding the business processes of companies in different industry sectors and with different sizes [FK19]. We therefore assume that most SMEs can be described based on their value chain activities with sufficient accuracy following Porter's value chain approach [Po85] and that other parameters such as the company size or the industry sector play a subordinate role. To achieve this, we identified major business processes, processed data categories and typical IT-systems and applications for every value chain activity. Based on the information regarding business processes and processed data, we have derived possible damage scenarios that directly affect the business processes, the processed data and the underlying IT-systems. In this procedure, the protection goals "Confidentiality", "Integrity" and "Availability" of the CIA-triad [Pe08] were used to derive damage scenarios for every business process within every value chain activity. The damage scenarios focus on the business impact and not on IT-security incidents that may cause them. This way, we ensure that various company representatives can adequately assess the damage scenarios without technical or IT-security expert knowledge. The procedure outlined above enables us to assess companies based on their value chain activities and the corresponding damage scenarios. To identify recommendations to increase the company's IT-security level, we chose an attack-based approach by using MITRE's "ATT&CK Matrix for Enterprise" [MI]. With this approach, we made assumptions for each value chain activity based on the identified business processes and IT-systems (e.g. "no portable media in production environment"). Considering these assumptions, we evaluated all attack techniques contained in the matrix according to whether the attack is relevant in the context of the considered value chain activity. Based on this, we identified associated mitigation measures which are given in the matrix. These mitigation measures are described in detail and address specific systems and are therefore not suitable in their given form. Therefore, we derived recommendations on a higher level based on them and aggregated similar recommendations. These recommendations are not new and can be found in various IT-security guidelines. However, in our approach we don't just want to identify recommendations relevant to a specific company but want to identify the most urgent ones where there is the greatest need for action. The recommendations are designed to be easy to understand and to be implemented through various suitable technical and/or organizational measures including adequate offerings (products/services) that are provided by the system. The expert system is based on a Bayesian network [En97], [Je01] that links the company's value chain activities to the recommendations. Every value chain activity forms a node in the Bayesian network that is connected to the nodes of the corresponding damage scenarios, which in turn are connected to the recommendations layer (Figure 1). In total 9 value chain activities with a total of 48 damage scenarios are mapped. The information which recommendations are suitable to address the respective damage scenarios is stored in conditional probability tables within a knowledge base that is built by using specially developed software tools for knowledge extraction and representation, which are described in more detail below. The knowledge base contains the recommendations that are individually rated by IT-security experts based

on their suitability/relevance to address specific damage scenarios for every value chain activity. The expert system is used as follows: Company representatives rate possible damage

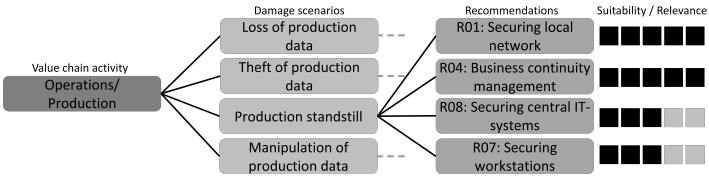


Fig. 1: Simplified, exemplary illustration of the bayesian network used in the expert system

scenarios based on the business impact for every value chain activity. The expert system matches these ratings with the recommendations stored in the knowledge base and outputs suitable ones in a prioritized order. The given recommendations and their priority are based on the information provided by the company representative on the impact of possible damage scenarios and the corresponding expert knowledge to address them. Companies can choose appropriate technical and/or organizational measures for implementing the recommendations that are output by the system. In addition, suitable services and products for each measure are output via a database query. These results can be filtered further, e.g. to find regional providers. By using the expert system, companies receive prioritized recommendations that match their individual requirements. In addition, appropriate technical and/or organizational measures including potential products, services and providers are suggested to implement the recommendations. We developed the expert system as a working prototype with the backend program including additional tools for building the knowledge base implemented in Python. The frontend of the underlying server/client-architecture was implemented in HTML and Javascript as functional mockup in order to be able to perform user tests at an early development stage. The production version of the frontend was implemented according to our specifications by Hochschule Mannheim and is available online [TI22a]. After the company representative has provided all the necessary information, these are transferred to the expert system running in the backend that returns a set of suitable recommendations in a prioritized order. In the backend application the company data is matched with the knowledge base. As a result, a set of recommendations sorted by priority is generated. This ensures a certain degree of transparency and the companies are made aware of the critical value chain activities. To provide additional guidance for the priority of individual recommendations, an indicator for the relative relevance is calculated for every recommendation. The knowledge base containing the values of the conditional probability tables of the Bayesian network is generated using the knowledge representation tool by calculating mean values of individual datasets generated by IT-security experts.

4 First insights from the company data

We evaluated the anonymously collected data that it transmitted to gain insights on the companies' value chain activities and particularly threatening damage scenarios. As shown in

Figure 2, the activities “infrastructure”, “marketing & sales”, “human resource management” and “customer service” were selected most frequently. The frequency distribution shows

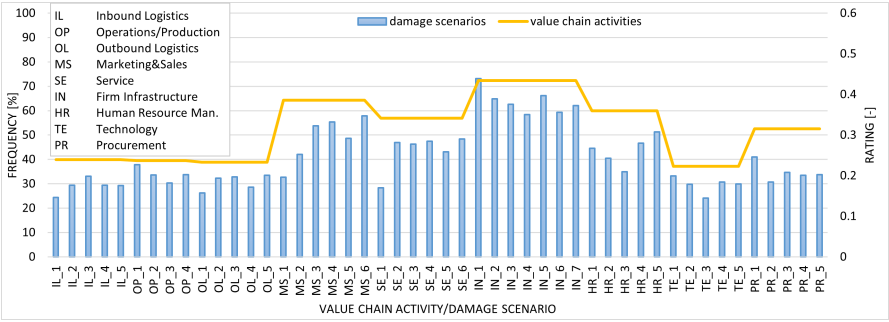


Fig. 2: Rating of damage scenarios and value chain activities (n=3511)

that most of the companies either don’t have production related value chain activities like “inbound logistics” or only use IT-systems within these activities to a limited extent. When considering the data regarding the rating of the damage scenarios within the respective value chain activities, the scenario “IN_1: no/limited availability of IT-systems” is rated the highest (0.61 in a range from 0 to 1), followed by the damage scenario “OP_1: production shutdown” (0.57), “IN_5: loss of company data” (0.55) and “MS_6: theft of customer/contract data” (0.54). The scenarios “SE_1: web presence not available” (0.30), “MS_1: online shop not available” (0.31) are rated the lowest. It’s an interesting result, that the highest rated damage scenarios reflect the results of recent studies, that identify ransomware attacks and the associated downtime and a standstill in production as the greatest threats to companies [Bi21]. These results are an indicator, that the developed solution is being used correctly.

5 Conclusion

The threat of cyberattacks is particularly challenging for SMEs, which are often not adequately protected from them and overwhelmed by complex guidelines and a wide range of IT-security products and services. We have therefore developed an online tool for matching company demands with IT-security offerings described in this paper, that supports SMEs in improving their IT-security level by providing a low-threshold entry into IT-security. The underlying model is based on Porter’s value chain to describe the companies regarding their business processes and possible damage scenarios from cyberattacks. Furthermore, we identified possible recommendations that the expert system can output by using an attack-based approach. For every recommendation a set of suitable technical/organizational measures is provided including suitable IT-security offerings. Companies can use the expert system by rating damage scenarios based on the business impact and receive suitable recommendations in a prioritized form with appropriate technical/organizational measures including adequate products/services. The developed expert system is a lightweight approach that can’t replace resource intensive IT-security consulting projects, especially

because there's no assessment of the company's current state of the IT-infrastructure and implemented IT-security measures. However, we performed several user tests with company representatives of different industry sectors and further optimized the expert system based on the results. Based on our experience from these user tests and an additional focus group discussion, we think that our tool fulfills its purpose to support companies to get started with improving their IT-security. However, only time can tell how the expert system performs in practice. Furthermore, the analyses shown in this paper are not a representative study, but rather the first results from the operation of the expert system that provide an initial insight into the company data. We will constantly optimize the expert system based on our lessons learned and extend the knowledge base with additional datasets created by IT-security experts to further improve the quality of the results.

References

[Bi21] Bitkom: Wirtschaftsschutz 2021, 2021.

[Bs11] BSI; secunet: Studie zur IT-Sicherheit in kleinen und mittleren Unternehmen, Bonn, 2011.

[BS20] BSI: Die Lage der IT-Sicherheit in Deutschland 2020, Bonn, 2020.

[BS21] BSI: IT-Grundschutz: Informationssicherheit mit System, Mar. 2021.

[En97] Enrique Castillo Jose Manuel Gutierrez, A. S. H.: Expert Systems and Probabilistic Network Models. Springer-Verlag, New York, 1997.

[FK19] Fähnrich, N.; Kubach, M.: Enabling SMEs to comply with the complex new EU data protection regulation. In: Open Identity Summit 2019. Gesellschaft für Informatik, 2019.

[G 21] G DATA CyberDefense AG: Cybersicherheit in Zahlen, Hamburg, 2021.

[Hi17] Hillebrand, A.; Niederprüm, A.; Schäfer, S.; Thiele, S.: Aktuelle Lage der IT-Sicherheit in KMU, tech. rep., Bad Honnef: WIK GmbH, 2017.

[IS13] ISO/IEC: Standard ISO/IEC 27001: Information technology - Security techniques - Information security management systems - Requirements, Geneva, 2013.

[Je01] Jensen, F. V.: Bayesian Networks and Decision Graphs. New York, 2001.

[MI] MITRE: ATT&CK Matrix for Enterprise, `https://attack.mitre.org/matrices/enterprise/`, visited on: 02/15/2022.

[Pe08] Perrin, C.: The CIA Triad, 2008.

[Po85] Porter, M. E.: Competitive advantage: creating and sustaining superior performance. New York, 1985.

[TI22a] TISiM: Sec-O-Mat, `https://secomat.de`, 2022.

[TI22b] TISiM: Transferstelle IT-Sicherheit im Mittelstand, `https://tisim.de`, 2022.

Combination of x509 and DID/VC for inheritance properties of trust in digital identities

Paul Bastian[1], Carsten Stöcker[2], Steffen Schwalm[3]

Abstract: The proposal for review of the eIDAS Regulation from 2021 has opened strong expectations for a deep change in traditional identity models. The new regulation starts with the creation of European Digital Identity Wallets that will enable citizens' control over their data in identification and authentication processes. Likewise digital identities and digital signatures are in place and interoperability between existing solutions mainly based on x509 certificates and decentralized PKI using DID/VC foreseeable. The paper provides various options in combining x509 and DID/VC approaches.

Keywords: eIDAS, SSI, self-sovereign identity, x509, DID, verifiable credentials, interoperability

1 Introduction

Unique identification of legal or natural entities as well as their objects – the basement for a digital identity – allows the verification of companies (Do they really exist?), the person acting for the company (Do they really exist?) and their authorization (Is Alice authorized to act for company A?). Digital identities and digital signatures are currently typically issued by a centralized authority using [x509] certificates as well as [OIDC] and [OAuth2]-protocols while e.g. DLT follow the DID/VC [W3C]. Both technical approaches are basically possible to execute the new SSI-paradigma but according to the comprehensible dissemination of x509/OIDC-approach the vice-versa interoperability is essential. This paper specifically discusses possible hybrid approaches on how to technically combine x509 and DID/VC. The paper is based on results of research projects from *GAIA-X Federation Services*[4] and *ID Union* where the authors take part in.

2 Hybrid-Approaches x509 and DID/VC

Conceivable hybrid approaches for combination of x509 and DID/VC are Embedding DID in x509 certificate, Derivation DID from x509 key pair, Encapsulated credential during onboarding in use case domain including issuance of identity credential, x509 based wallet and trusted verifier, Signed x509 in DID-Document, Using [eIDASBridge]. This

[1] Bundesdruckerei Group, Kommandantenstraße 18, 10969 Berlin, Germany
[2] Spherity GmbH, Emil-Figge-Straße 80, 44227 Dortmund, Germany
[3] msg group, Robert-Bürkle-Str. 1, 85737 Ismaning, Germany
[4] See https://gaia-x.eu/what-is-gaia-x/federation-services

cbe doi:18.18420/OID2022-13

list is not comprehensible so there are more approaches possible.

3 Discussion of hybrid approaches

3.1 Option 1: Embedding DID in x509 certificate

During issuance of a x509 certificate a signed DID will be embedded in the x509 certificate. It's necessary to ensure that a (qualified) trust service provider acc. [eIDAS] is needed to ensure that the DID is really linked to the identified natural or legal entity. This means basically that the onboarding process for x509 certificates must be changed such that the QTSP as an issuer of x509 certificates validates the signed DID of the identified natural or legal entities using a secure communication channel e.g., TLS acc. [TR02102]. Afterwards the DID will be integrated into the certificate as an x509 extension e.g., by a trusted resolver service so that the verifier gets the information how to resolve the DID from the DID document. This means that in the results the x509 includes a DID which can be resolved by a trusted 3rd party to ensure verifiability and useability of VC of wallet service endpoint of the given holder. Any other identity x509 attributes including the Root-CA can be used as usual without any change needed.

Option	Advantages	Disadvantages
Embedding DID in x509 certificate	Method for combination of x509 certificates with DID for inheritance of properties/credentials of verified entities	Change in x509 issuance process necessary

Table 1: Summary Option 1

3.2 Option 2: Derivation DID from x509 key pair

For the special use case that x509 certificates and DID use the same cryptographic primitives the key material of x509 may be used for evidence of control of the x509 certificate itself as well as the given DID document. Public and private key pair of the x509 certificate will be used for the creation of a new DID and DID document. This approach can be beneficial if dedicated crypto primitives are mandatorily required due to compliance, business or legal needs.

Option	Advantages	Disadvantages
Derivation DID from x509 key pair	Method for combination of x509 certificates with DID for inheritance of	Requires utilisation of same crypto primitives for x509 and DID as well as

	properties/credentials of verified entities	authoritative control for signatures in DID Document

Table 2: Summary Option 2

The picture below illustrates option 1 and 2:

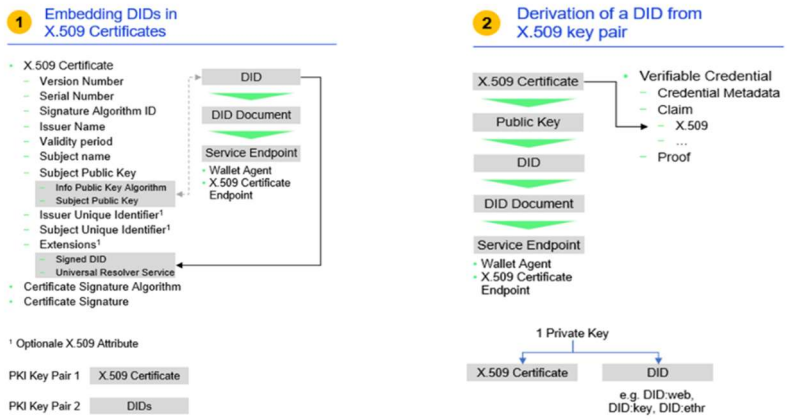


Figure 1: Option 1 and 2

3.3 Option 3: Encapsulated credential during onboarding

The idea of option 3 is that a verification service validates if the holder really owns two private keys so Private key for x509 certificate and Private key of self created (or created by TSP) DID. In this case the holder creates and signs the credential with his private keys to achieve an encapsulated data structure which contains both signatures. The aim is that the verifier is enabled to validate if the holder controls both private keys X509 identity proof, e.g. EV or QWAC, DID control of DID private key. Sequence of encapsulation does not matter and may be designed according to the communication protocols in use. If the verifier is an onboarding or verification service, the signature can be verified directly in the encapsulated credential itself. Additionally, the x509 verification service verifies the validity and trust chain of the x509 certificate. In the next step the can create verifiable credential for the holder where the trust is given by e.g., at the trusted issuer [BaseID], [OCI].

Option	**Advantages**	**Disadvantages**
Encapsulated credential	Method for combination of	Additional verification

Option	Advantages	Disadvantages
during onboarding	x509 certificates with DID for inheritance of properties/credentials of verified entities Encapsulated credentials are established approach No change in x509 specification	service (trusted third party) for issuance of VC

Table 3: Summary Option 3

3.4 Option 4: x509 based wallet and trusted verifier

X509 certificates may also be used to validate if the holder communicates with the infrastructure domain of a verified issuer or verifier. Under the assumption that those systems running in the same infrastructure domain the verifier may assume that the DID-based Wallet hosted in this domain is owned by the given holder. This implies that an SSI agent will create a channel to a service endpoint e.g., using an Aries DIDComm-Channel running over HTTPS. The communication is operated encapsulated with the assumption that if the user the x509 certificate of the outer channel is trustworthy that the endpoint of HTTPS is the mentioned DID subject and consequently the verifier is trustworthy too. The X.509 certificate can be an Extended Validation Certificate (EV).

Option	Advantages	Disadvantages
Wallet infrastructure with an X.509 Certificate (e.g. Extended Validation Certificate)	Easy to implement May solve issuer of trusted verifier	No solution for interoperability between x509 and DID/VC Only works for HTTPS-related communication while DIDComm also supports other channels like Bluetooth or NFC

Figure 5: Summary Option 4

3.5 Option 5: Signed x509 in DID-Document

Another option is to add a signed x509 certificate (e.g., signed by a qualified trust service

provider) in the DID-document of the holder and the certificate end point in the DID-document itself. The result is an encapsulated credential like option 2. During the addition of signed x509 in DID-document this must also be signed with its private key to update the DID-document including the x509 certificate.

Option	Advantages	Disadvantages
Signed x509 in DID-Document	Method for combination of x509 certificates with DID for inheritance of properties/credentials of verified entities	Addition of signed x509 in DID-document is not defined in W3C-DID-Specification, extension allowed Update of DID-document implies no secured link

Table 4: Summary Option 4

3.6 Option 6: eIDAS Bridge

Further option is the utilization of [eIDASBridge]. The [eIDASBridge]. was developed by the European Commission to establish a legal compliant link between SSI based on DID/VC and existing digital identities based on x509. It contains legal reports and technical specifications and ensures legal trust in SSI if [eIDAS2] is not fully applicable. The [eIDASBridge] implies that verifiable credentials are signed with an additional (qualified) electronic signature or seal of the issuer from a qualified trust service provider acc. to [eIDAS]. In result existing validation mechanism acc. [ETSIEN319102] can be used to make the authenticity and integrity of the VC evident against 3rd parties to fulfil the burden of proof and documentation requirements. [Ko20], [We18].

Option	Advantages	Disadvantages
eIDAS Bridge	Ensures legal trust of VC Verifiability of VC by any validation service acc. eIDAS	Less feasible for interoperable attribute exchange between x509- and DID/VC-based environments

Table 5: Summary Option 6

4 Outlook

The interoperability between x509 and DID/VC based digital identities as well as digital signatures can be mentioned as one of the most important success factors for SSI. The

paper discussed roughly different possiblities which will be analyzed in detail by the authors and may be part of further standardization..

Bibliography

[eIDAS1] Regulation (EU) No 910/2014 of the European Parliament and of the Council - of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. eIDAS, 2014.

[eIDAS2] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity {SEC(2021) 228 final} - {SWD(2021) 124 final} - {SWD(2021) 125 final}

[eIDASBridge] Burgos, O. et al: SSI eIDAS Bridge - Use cases and Technical Specifications. Brusssels 2020: https://joinup.ec.europa.eu/collection/ssi-eidas-bridge/document/ssi-eidas-bridge-use-cases-and-technical-specifications

[ETSIEN319102] ETSI EN 319 102-1 Electronic Signatures and Infrastructures (ESI). Procedures for Creation and Validation of AdES Digital Signatures; Part 1: Creation and Validation

[IS20] ISO/IEC 9594-8:2020 Information technology - Open systems interconnection - Part 8: The Directory: Public-key and attribute certificate frameworks

[Ko20] Korte, U. et. al.: Criteria for trustworthy digital transactions – Blockchain/ DLT between eIDAS, GDPR, Data and Evidence Preservation. OpenIdentity Summit 2020. Lecture Notes in Informatics (LNI). Proceedings. Bonn 2020 S. 49-60

[OIC] Hendrix, P et. Al.: Credential Issuer Conformance Criteria v2.0.0. W3C. 2020

[OIDC] OpenID Connect protocol: https://openid.net/connect/

[OAuth2] OAuth2 protocol: https://oauth.net/2/

[RFC5280] Cooper, D. et. Al.: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. 2008

[TR02102] Technical Guideline TR-020159. BSI TR-02102 Cryptographic Mechanisms. Federal Office for Information Security. https://www.bsi.bund.de/EN/Service-Navi/Publications/TechnicalGuidelines/tr02102/tr02102_node.html

[W320] W3C: Decentralized Identifiers (DIDs) v1.0. 2020.

[We18] Weber, M. et al.: Records Management nach ISO 15489. Einführung und Anleitung. Beuth Verlag, Berlin, 2018.

GI-Edition Lecture Notes in Informatics

P-299 M. Gandorfer, A. Meyer-Aurich, H. Bernhardt, F. X. Maidl, G. Fröhlich, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Digitalisierung für Mensch, Umwelt und Tier
Referate der 40. GIL-Jahrestagung
17.–18. Februar 2020,
Campus Weihenstephan

P-300 Michael Felderer, Wilhelm Hasselbring, Rick Rabiser, Reiner Jung (Hrsg.)
Software Engineering 2020
24.–28. Februar 2020
Innsbruck, Austria

P-301 Delphine Reinhardt, Hanno Langweg, Bernhard C. Witt, Mathias Fischer (Hrsg.)
Sicherheit 2020
Sicherheit, Schutz und Zuverlässigkeit
17.–20. März 2020, Göttingen

P-302 Dominik Bork, Dimitris Karagiannis, Heinrich C. Mayr (Hrsg.)
Modellierung 2020
19.–21. Februar 2020, Wien

P-303 Peter Heisig, Ronald Orth, Jakob Michael Schönborn, Stefan Thalmann (Hrsg.)
Wissensmanagement in digitalen Arbeitswelten: Aktuelle Ansätze und Perspektiven
18.–20.03.2019, Potsdam

P-304 Heinrich C. Mayr, Stefanie Rinderle-Ma, Stefan Strecker (Hrsg.)
40 Years EMISA
Digital Ecosystems of the Future: Methodology, Techniques and Applications
May 15.–17. 2019
Tutzing am Starnberger See

P-305 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim, Detlef Hühnlein (Hrsg.)
Open Identity Summit 2020
26.–27. May 2020, Copenhagen

P-306 Arslan Brömme, Christoph Busch, Antitza Dantcheva, Kiran Raja, Christian Rathgeb, Andreas Uhl (Eds.)
BIOSIG 2020
Proceedings of the 19th International Conference of the Biometrics Special Interest Group
16.–18. September 2020
International Digital Conference

P-307 Ralf H. Reussner, Anne Koziolek, Robert Heinrich (Hrsg.)
INFORMATIK 2020
Back to the Future
28. September – 2. Oktober 2020, Karlsruhe

P-308 Raphael Zender, Dirk Ifenthaler, Thiemo Leonhardt, Clara Schumacher (Hrsg.)
DELFI 2020 –
Die 18. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
14.–18. September 2020
Online

P-309 A. Meyer-Aurich, M. Gandorfer, C. Hoffmann, C. Weltzien, S. Bellingrath-Kimura, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Referate der 41. GIL-Jahrestagung
08.–09. März 2021, Leibniz-Institut für Agrartechnik und Bioökonomie e.V., Potsdam

P-310 Anne Koziolek, Ina Schaefer, Christoph Seidl (Hrsg.)
Software Engineering 2021
22.–26. Februar 2021,
Braunschweig/Virtuell

P-311 Kai-Uwe Sattler, Melanie Herschel, Wolfgang Lehner (Hrsg.)
Datenbanksysteme für Business, Technologie und Web (BTW 2021)
Tagungsband
13.–17. September 2021,
Dresden

P-312 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)
Open Identity Summit 2021
01.–02. Juni 2021, Copenhagen

P-313 Ludger Humbert (Hrsg.)
Informatik – Bildung von Lehrkräften in allen Phasen
19. GI-Fachtagung Informatik und Schule
8.–10. September 2021 Wuppertal

P-314 Gesellschaft für Informatik e.V. (GI) (Hrsg.)
INFORMATIK 2021Computer Science & Sustainability
27. September– 01. Oktober 2021, Berlin

P-315 Arslan Brömme, Christoph Busch, Naser Damer, Antitza Dantcheva, Marta Gomez-Barrero, Kiran Raja, Christian Rathgeb, Ana F. Sequeira, Andreas Uhl (Eds.)
BIOSIG 2021
Proceedings of the 20th International Conference of the Biometrics Special Interest Group
15.–17. September 2021
International Digital Conference

P-316 Andrea Kienle, Andreas Harrer, Jörg M. Haake, Andreas Lingnau (Hrsg.)
DELFI 2021
Die 19. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
13.–15. September 2021
Online 8.–10. September 2021

P-317 M. Gandorfer, C. Hoffmann, N. El Benni, M. Cockburn, T. Anken, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Künstliche Intelligenz in der Agrar- und Ernährungswirtschaft
Referate der 42. GIL-Jahrestagung
21. - 22. Februar 2022 Agroscope, Tänikon, Ettenhausen, Schweiz

P-318 Andreas Helferich, Robert Henzel, Georg Herzwurm, Martin Mikusz (Hrsg.)
FACHTAGUNG SOFTWARE MANAGEMENT 2021
Fachtagung des GI-Fachausschusses Management der Anwendungsentwicklung und -wartung im Fachbereich Wirtschafts-informatik (WI-MAW), Stuttgart, 2021

P-319 Zeynep Tuncer, Rüdiger Breitschwerdt, Helge Nuhn, Michael Fuchs, Vera Meister, Martin Wolf, Doris Weßels, Birte Malzahn (Hrsg.)
3. Wissenschaftsforum:
Digitale Transformation (WiFo21)
5. November 2021 Darmstadt, Germany

P-321 Veronika Thurner, Barne Kleinen, Juliane Siegeris, Debora Weber-Wulff (Hrsg.)
Software Engineering im Unterricht der Hochschulen SEUH 2022
24.–25. Februar 2022, Berlin

P-323 Christian Wressnegger, Delphine Reinhardt, Thomas Barber, Bernhard C. Witt, Daniel Arp, Zoltan Mann (Hrsg.)
Sicherheit 2022
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 11. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
5.–8. April 2022, Karlsruhe

P-325 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)
Open Identity Summit 2022
Fachtagung vom 07. - 08. July 2022, Copenhagen

All volumes of Lecture Notes in Informatics can be found at https://dl.gi.de/handle/20.500.12116/21.

The titles can be purchased at:

Köllen Druck + Verlag GmbH
Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn
Fax: +49 (0)228/9898222
E-Mail: druckverlag@koellen.de

Gesellschaft für Informatik e.V. (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and information systems), to document conferences that are organized in cooperation with GI and to publish the annual GI Award dissertation.

Broken down into
• seminars
• proceedings
• dissertations
• thematics
current topics are dealt with from the vantage point of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure high quality contributions.

The volumes are published in German or English.

Information: http://www.gi.de/service/publikationen/lni/

ISSN 1617-5468
ISBN 978-3-88579-719-7

Open standards and interfaces as well as open source technologies play a central role in the current identity management landscape as well as in emerging future scenarios in the area of electronic identification and trust services for electronic transactions according to the eIDAS regulation (2014/910/EU), innovative payment services according to the second payment services directive (PSD2) (2015/2366/EU), trustworthy and privacy enhancing solutions according to the general data protection regulation (2016/679/EU) and other innovative applications in the area of e-health, e-government, cloud computing and the internet of things for example.