H. Roßnagel, C. H. Schunck, S. Mödersheim, D. Hühnlein (Hrsg): Open Identity Summit 2020,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2020 163

Privacy by Design Architecture Composed of Identity Agents Decentralizing Control over Digital Identity

Kalman C. Toth[1] and Ann Cavoukian[2] and Alan Anderson-Priddy[3]

Abstract: Proposed is an identity architecture that satisfies the principles of privacy by design, decentralizes control over digital identity from providers to users, mitigates breach and impersonation risks, and reduces dependency on remote access passwords. The architecture is composed of interoperating identity agents that work on behalf of their owners and deploy digital identities that are virtualized to look and behave like identities found in one's wallet and contacts list. Encapsulating authentication data, identity agents strongly bind owners to their digital identities and private keys enabling them to prove who they are, protect their private data, secure transactions, conduct identity proofing, and reliably delegate consent. Identity agents also off-load application services from identity-related and privacy-related tasks. A gestalt privacy by design process has been used to discover the architecture's privacy requirements and design elements and systematically reason about how the design elements satisfy the privacy requirements. Identity-related functionality has been intentionally compartmentalized within identity agents to focus development on creating trustworthy software. A reference model for development derived from the described identity architecture is proposed.

Keywords: privacy, privacy by design, digital identity, authentication, verification, security.

1 Introduction

The identity architecture proposed herein has been motivated by the alarming growth in identity theft, impersonation, fraud and lost privacy due to private data collection by service providers, remote access password vulnerabilities, and the web's patchwork of identity schemes. Large-scale breaches (e.g. Facebook, Google, Capital One, Marriott, Sony, Target, JP Morgan, Home Depot, Equifax) have disclosed social security numbers, personal information, medical records, credit reports, credit card records, bank accounts, voter data, and other such sensitive information. Authorities are deeply concerned about threats to our critical infrastructure including power, transportation and voting systems.

The identity architecture satisfies the principles of privacy by design, decentralizing digital identities to owners enabling them to prove who they are, protect their private

[1] NexGenID, Portland, Oregon 97205, USA, kalmanctoth@gmail.com
[2] Global Privacy & Security by Design Centre, Toronto, M4S 2X6, Canada, ann.cavoukian@gpsbydesign.com
[3] Portland State University, OIT, Portland, Oregon 97207, USA, andersonpriddy@gmail.com

data, secure transactions, elevate identity assurances, and reliably delegate consent.

2 Privacy by Design Dependent on Capable Digital Identity System

Explained by Ann Cavoukian in [Ca17], the principles of privacy by design include minimizing private data disclosure and collection; safeguarding private data, securing transactions end-to-end; delegating consent to access private resources; and establishing privacy as the system default setting to ensure acceptable levels of privacy protection.

Consider that privacy can be lost when underlying identity schemes are broken. Users can be tricked by rogue data collection sites to disclose passwords, second factor access codes, and social security numbers; weak passwords can be broken and used by imposters to access private data in the cloud; transactions can be intercepted; and delegated consent can be defeated when stolen identities are used by imposters. In other words, enhanced privacy is highly dependent on the efficacy of the identity system used.

3 **Decentralizing I**dentity Reduces Risks for Providers and Users

Many web services have become honeypots for identity theft partly because of the enormous volume of private and identifying data they collect. To address this problem, writers including [Al16], [So18] and [WW19b] have proposed deploying *self-sovereign identities* and *decentralized identifiers* (DIDs) to shift control over identity to users.

The architecture described in this paper decentralizes digital identity by providing users trustworthy agents that help them make safe identity and privacy-related decisions. Installed on the user's personal device, an identity agent protects and tightly binds the owner to her digital identities. Her identity agent helps her decide which digital identities to create and use, what private data to protect, which consent permissions to reliably delegate, how to elevate identity assurances, and what private data to disclose. Minimizing disclosure limits how much data service providers need to collect which reduces breach risk while dispersing the attack surface.

4 System Concept: Decentralized and Privacy Enhanced Identity

Depicted in Fig. 1, the system concept for the proposed identity architecture [TA18], [TA19a], [TA19b], [TCA20][4] is composed of *identity agents* and *digital identities* installed on the devices of users and providers (e.g. smart phones, servers, laptops).

[4] "Electronic Identity and Credentialing System", US Patent 9,646,150 B2, May 9, 2017.

Identity agents decentralize control over identity from service providers to users while satisfying the principles of privacy by design for their joint benefit - safely managing disclosure, data collection, privacy protection, transaction security and delegated consent.

Identity agents empower users and administrators by virtualizing digital identities such that they look and behave like physical credentials in their digital wallets and contact lists. They tightly bind owners to their digital identities, consent tokens, PINs, keys, and other artifacts by encapsulating the authentication data of the owner. Identity agents also off-load application services from identity and privacy management tasks and interoperate with other identity agents, digital identity exchange services, and proof-of-existence identity registries. Digital identities specify an identifier plus selected attributes/images characterizing the owner, or little or no identifying information for pseudonymous and anonymous uses. Depending on perceived identity correlation risks, an identifier can be unique within a given context, globally unique, or pair-wise unique [WW19b].

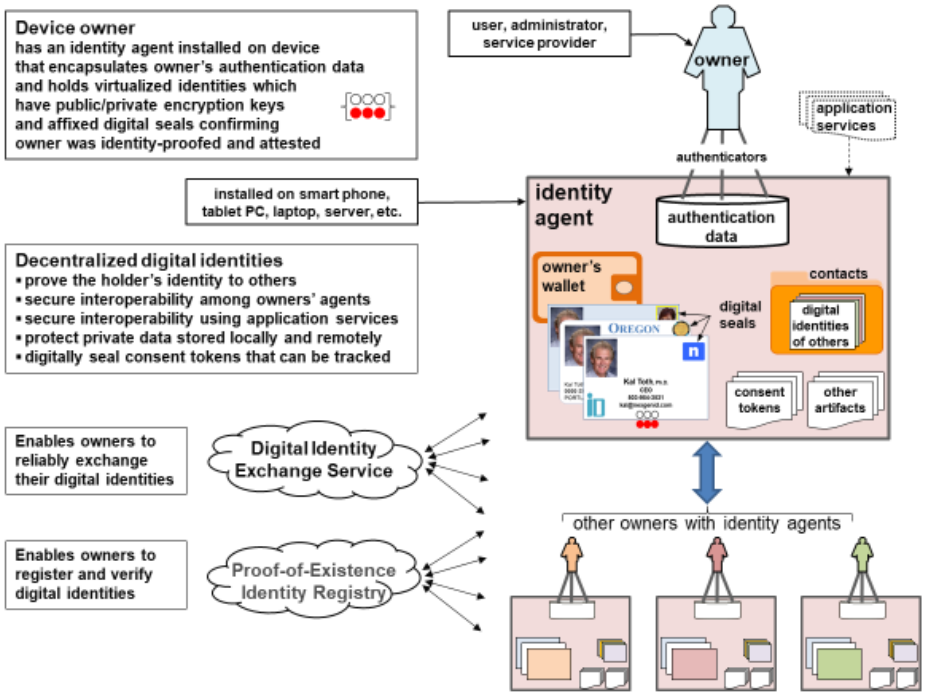

Fig. 1: System concept diagram

Each digital identity created by an owner is allocated multiple public/private encryption key-pairs, each pair used for designated purposes when selected to prove identity, secure

transactions, and delegate consent to access private data. The private embossing key of a digital identity can be used to digitally seal attestations to digital artifacts [TCA20][5]. Such attestations cannot be repudiated because the owner controls her device, identity agent, selected digital identity, and embossing key used to bind her identity and attestation to the digital artifact (e.g. to a digital identity, a consent token, or a legal document).

For example, a requesting owner can present his digital identity and identifying data to an issuing owner who proofs his identity. If successfully proofed, the issuer can use one of her digital identities to issue a digital seal affixing her attestation plus her identity to the requester's digital identity which the issuer cannot repudiate and can be verified.

5 Privacy by Design Process and Validation

Early in development, system engineers routinely use a gestalt process to define requirements and evaluate designs, iterating until they converge on an acceptable design satisfying the requirements. Fig. 2 depicts the privacy by design process used to discover and validate the architecture's privacy requirements (R) and design elements (D). Upon thoroughly iterating over all the privacy requirements and design elements, the listed requirements and elements were validated. Each iteration contributed to the goal of showing that the principles of privacy by design were satisfied.

[5] "Methods for Using Digital Seals for Non-Repudiation of Attestations", US Patent 9,990,309B2, 2-20-2018.

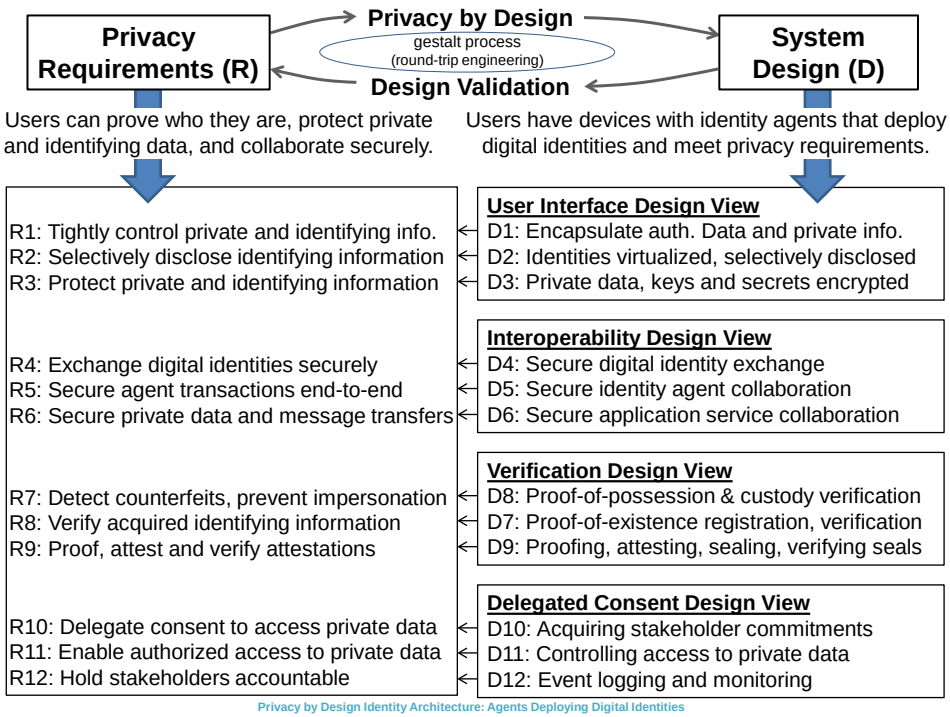

Fig. 2: Privacy by design process

The privacy by design validation process detailed in [TCA20] reasons about how the proposed identity architecture enables the following functions and features:

- Prevents compromise by encapsulating authentication data (e.g. biometric minutia, PIN hashes) used by device authenticators to verify owner presence and custody.
- Virtualizes the look and feel of physical identities rendering them as simple to use as passwords but with enhanced utility, usability and intuitive ease of use.
- Leverages an identity data model (e.g. [WW19a]) to specify civil digital identities for owners as well as so-called pseudonymous and anonymous identities.
- Allocates public/private encryption key-pairs to digital identities used to secure private data, transactions, messages, and consent tokens for the owner.
- Elevates identity assurances associated with digital identities by proofing [NI17] and affixing attestations to them with digital seals that cannot be repudiated.
- Registers digital identities that are hashed and digitally sealed in a proof-of-existence registry [Ro16], [TCA20][6] that other parties can verify.

[6] "Systems and Methods for Registering and Acquiring E-Credentials using Proof-of-Existence and Digital Seals", US Pat 10,127,378 B2, issued Nov. 13, 2018.

- Mitigates attack risks by leveraging an adaptation of the Diffie-Hellman key agreement method to exchange digital identities [Re99], [TCA20][7].
- Launches challenges to verify private key possession [ANL03] of presented digital identities and custody of owner devices to detect counterfeits and impersonation.
- Uses digital seals to affix requests, approvals and access permissions to consent tokens that can be tracked for accountability and cannot be repudiated.
- Establishes privacy by design default settings according to context and risks including using pseudonymous and anonymous identities; applying the adapted Diffie-Hellman exchange method; securing all transactions using digital identities; executing proof-of-possession and proof-of-custody challenges; and registering digital identities and consent tokens in a proof-of-existence registry.

6 Discussion: Privacy by Design Reference Model for Identity

A reference model derived from the identity architecture is proposed to communicate essential methods, interfaces and protocols[8] to developers. A key objective will be to ensure that deployed identity agents are trustworthy, namely, that they reliably and correctly integrate authentication data, user interfaces, cryptographic mechanisms, identity proofing and attestation, programming interfaces, and collaboration protocols. These essential building blocks have been intentionally compartmentalized within identity agents to facilitate the development of software that is reliable and trusted.

Options for implementing such trustworthy software include open source and proprietary development, and possibly a combination of both. Software licensing, support, and maintenance arrangements can vary widely. Pundits argue that open source software is more, less, or just as secure as proprietary software. The debate revolves mainly around developers having or lacking visibility into the code. Arguments about the merits and shortcomings of these options address issues related to hacking vulnerability, security by obscurity, responsiveness to problems, development methods, skills and tools used, time-to-deliver, business failure, liability, and warranty.

Whichever option is adopted, effective software inspections, comprehensive testing, quality assurance, and configuration management are essential. Formal software engineering methods can be applied to identity agents to enhance trustworthiness.

[7] "Architecture and Methods for Self-Sovereign Digital Identity", US Patent (pending), provisional filed Oct. 8, 2018, utility application filed Nov. 12, 2018.

[8] Founding team intends to issue a license to developers similar to RedHat's patent promise to discourage patent aggression https://www.redhat.com/en/about/patent-promise.

7 Closing Remarks

The privacy by design process has progressively baked privacy into the identity architecture [TCA20]. Decentralizing identity from service providers to users decreases what providers need to collect while dispersing the attack surface. Identity agent owners can control and use their digital identities to reliably prove who they are, verify the identities of others, protect their private and identifying information, and reliably delegate consent. Because digital identities are intuitive, and owners can control what they disclose, they are less dependent on remote access passwords. Digital identities that have been proofed, attested and digitally sealed elevate identity assurances for all stakeholders.

Bibliography

[Al16] Christopher Allen: The Path to Self-Sovereign Identity, April 27, 2016, http://coindesk.com.

[ANL03] N. Asokan, Baltteri Niemi, Pekka Laitinen: On the Usefulness of Proof of Possession, 2nd Annual PKI Workshop, Apr. 28-29, 2003, pp.136-141.

[Ca17] Ann Cavoukian: Privacy by Design, The 7 Foundational Principles, https://ipc.on.ca/wpcontent/uploads/Resources/7foundationalprinciples.pdf, 2017.

[NI17] NIST Special Publication 800-63A: "Digital Identity Guidelines, Enrollment and Identity Proofing", Jan. 2017, https://doi.org/10.6028/NIST.SP.800-63a.

[Re99] E. Rescorla: Diffie-Hellman Key Agreement Method, RTFM Inc., June 1999.

[Ro16] Kiara Robles: Tool for Creating Verifiable IDs on the Blockchain, BlockchainMe, Dec. 2, 2016, https://github.com/kiarafrobles/blockchainMe.

[So18] Sovrin Foundation: Sovrin: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust, Version 1, January 2018, https://sovrin.org.

[TA18] Kalman C. Toth and Alan Anderson-Priddy: Architecture for Self-Sovereign Digital Identity, Computer Applications for Industry and Engineering (CAINE), New Orleans, LA, Oct. 8-10, 2018.

[TA19a] Kalman C. Toth and Alan Anderson-Priddy: Self-Sovereign Digital Identity: A Paradigm Shift for Identity, IEEE Security and Privacy, Vol. 17, No. 3, May/June 2019.

[TA19b] Kalman C. Toth and Alan Anderson-Priddy: Privacy by Design using Agents and Sovereign Identities, Information Security and Privacy Protection Conference (IFIP-SEC), Work in Process and Emerging Research, Lisbon, Portugal, June 25-27, 2019.

[TCA20] Kalman C. Toth, Ann Cavoukian and Alan Anderson-Priddy: Privacy by Design Identity Architecture Using Agents and Digital Identities, Annual Privacy Forum, Lisbon, Portugal, accepted for presentation and publication, June 4-5, 2020

(postponed).

[WW19a] World Wide Web Consortium (W3C): Verifiable credentials data model 1.0: Expressing verifiable information on the web, proposed recommendation, 9-5-2019.

[WW19b] World Wide Web Consortium (W3C): Decentralized Identifiers (DIDs) v1.0: Core Data Model and Syntaxes, WC3 Working Draft 09 December 2019.