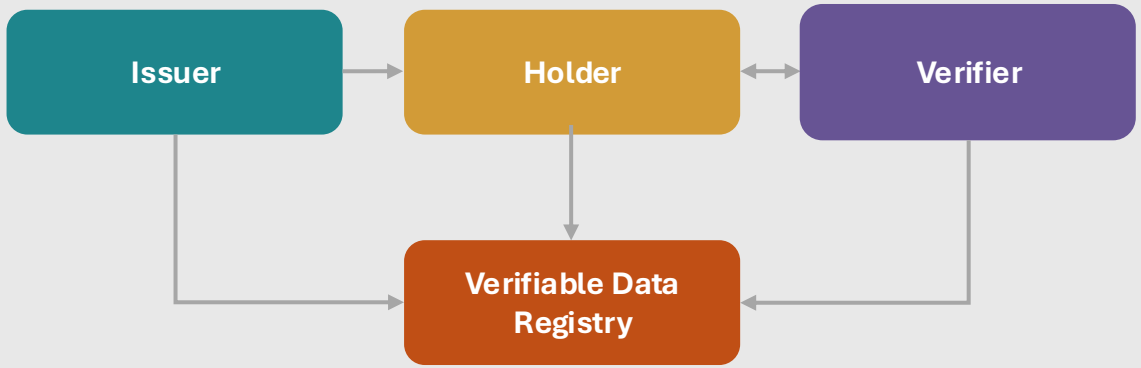
SoK: Data Minimization in Verifiable Credentials

ESORICS 2026, 21st DPM Workshop
17 September 2026

Norbert Henseler[1,2], Robin Doss[1], Anh Dinh[1], Marthie Grobler[2], Tina Wu[2]

[1]Deakin University, Australia

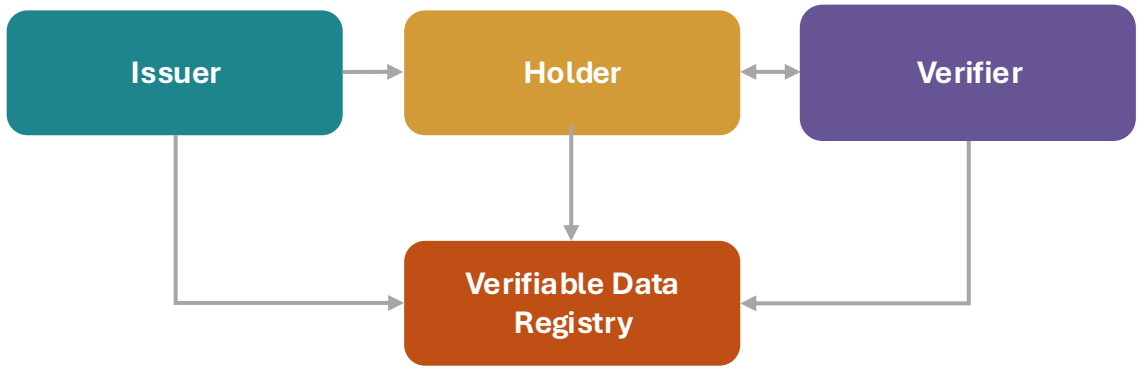
[2]CSIRO, Australia



Data Minimization is an Ecosystem Property

What are Verifiable Credentials (VCs)?

A VCs are tamper-evident digital credentials whose issuer and integrity can be cryptographically verified. They are Standardized by the W3C.



VC Ecosystem

VCs can support Selective Disclosure (SD), allowing holders to present only required claims.

Why Data Minimization in VCs Matters

The verifier often needs a fact, not the underlying data

Standard Practice	Minimized Disclosure
Date of birth	Proof age ≥ 18
Full address	Residency claim
Employment details	Evidence of employment status
Financial History	Eligibility, threshold proof
Copies retained for later use	Retention for a purpose

Image source: Pillar Catholic

The bank does not necessarily need the complete VCs. It needs enough information to make its decision.

The Problem Goes Beyond Selective Disclosure

Selective Disclosure (SD) can still begin with an over-sized VC or an over-reaching request.

Issuance	Storage, Presentation	Request
Issuer Which claims to put into a VC?	**Holder** What sensitive data remains in the wallet? Which data to share with whom?	**Verifier** What is requested, retained?

Prior work is mostly transaction based and SD.

Our SoK asks where Data Minimization (DM) occurs, who controls it and which techniques/mechanisms can support DM?

Two Questions, three Contributions

RQ1: Which DM techniques have been proposed for VCs, and how can they be organized?

RQ2: Where have these techniques been applied, and how transferable are they across domains?

C1

Taxonomy of Techniques

Organizes DM techniques and mechanisms

C2

Lifecycle Framework

Maps responsibilities and data flow to issuer, holder, verifier.

C3

Domain Mapping

Shows coverage and transferability across application domains.

The goal is to systematically map actors, techniques , lifecycle stages, application domains and propose a taxonomy + framework.

57 Studies after Systematic Screening

1,169 database hits → **892** after deduplication → **352** full text → **54** included → **3** 2026 update → **57** synthesis

Sources

- IEEE Xplore + Scopus
- Forward/backward citation search

Scope

- Peer-reviewed work
- Jan 2020 - Aug 2026

Synthesis

- Technique, mechanism
- Domain/use case
- Comparative evaluation

Deliberately broad search: VC-specific studies + transferrable DM techniques

Taxonomy of DM Techniques and Mechanisms

Selective Disclosure (SD)

VC specific

- Atomic credentials
- Commitments / hash-based
- SD signatures
- Zero-knowledge proofs

Focus: holder presentation

Rule-Based DM

Transferable

- Horizontal DM (hDM)
- Vertical DM (vDM)
- Transformative DM (tDM)

Useful for holder, verifier, issuer

ML-Based DM

Transferable, least mature

- Audit methods
- PII filtering
- Federated learning

No VC-specific minimizer found

SD dominates VC research. Rul-based and ML-based techniques broaden where DM can be used.

DM is an Ecosystem Property

Issuance	Storage/Presentation	Request
Issuer	**Holder**	**Verifier**
• Avoid unnecessary claims • Choose proof/signature design that enables SD	• Protect data at rest • Minimized stored data • Use SD to disclose only required claims	• Request only necessary claims • Do not correlate presentations • Minimize retention after presentation

Holder-side SD is insufficient. Issuers determine what can be disclosed; holders control what is disclosed, and verifiers are responsible for what is requested.

DM as Privacy-by-design principle depends on all actors.

Evidence across Domains

DM Technique	Education	Retail	Finance	Health	Prof. Credentials	Legal Identity	IoT	Domain independent
Selective Disclosure	●	●	●	–	●	●	●	●
Rule-Based DM	–	–	–	●	–	●	●	●
ML-Based DM	●	–	–	●	–	–	–	●

Dot = at least one reviewed study in that domain; exact counts are in the paper.

Selective Disclosure

Broadest domain coverage; strongest evidence around presentation-time minimization.

Rule-Based DM

Re-useable regulatory, policy or domain constraints.

ML-Based DM

Mostly transferable evidence; not yet mature as a VC workflow mechanism.

The strongest evidence is for minimizing what holders present, not what issuers create or verifiers ask for.

Five Gaps across the Research Area

1. Issuer Schemas

Avoid over-collection to cover a wide rage of the use-cases.

2. Verifier Requests

Generate minimized, purpose-bound requests.

3. Post-disclosure Control

Revocation, retention, reuse and unlinkability remain unresolved.

4. End-to-End Evaluation

Benchmark complete VC workflows, not isolated mechanisms.

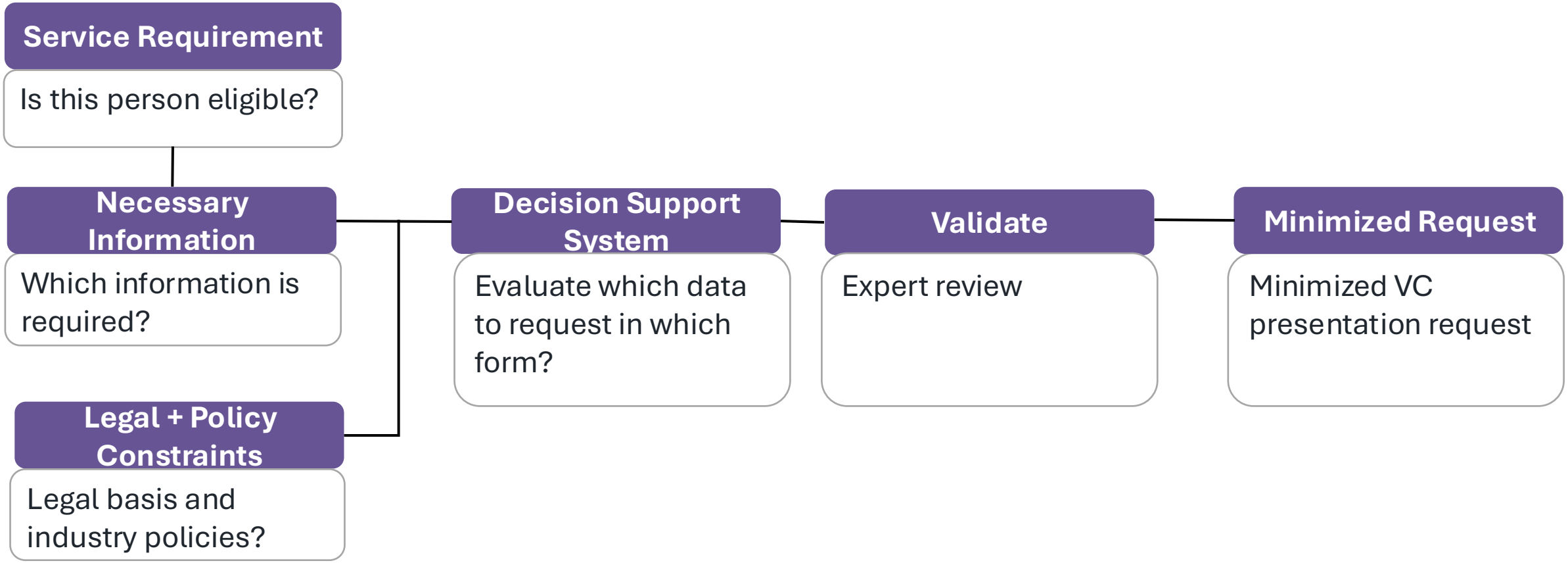
5. Post-Quantum Readiness

Compact, efficient quantum - secure SD.

The Gap is DM across the end-to-end VC lifecycle.

Decision Support for Minimal Verifier Requests

This is motivated by the SoK gap analysis; it is not a contribution of the paper.



The research aim is to make verifier request minimization measurable.

Three Takeaways

1. DM ≠ SD

SD is required, but DM depends on VC design, request formulation, retention

2. All Actors are responsible for DM

Issuers constrain what can be used; holders constrain what is disclosed; verifiers constrain what is requested and retained.

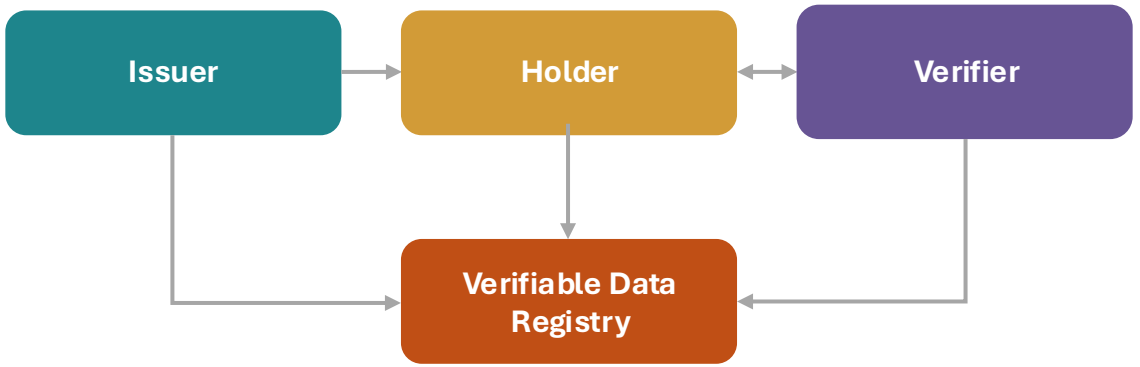
3. Verifier-, Issuer-Side DM is underdeveloped

Minimal, purpose-bound request generation and issuance of VCs with minimal number of claims are gaps for future research.

DM in VCs should be designed as a lifecycle property.

Questions & Feedback

How can the VC Ecosystem support only necessary and purpose bound data is disclosed?



Norbert Henseler

Deakin University, CSIRO

norbert.henseler@deakin.edu.au

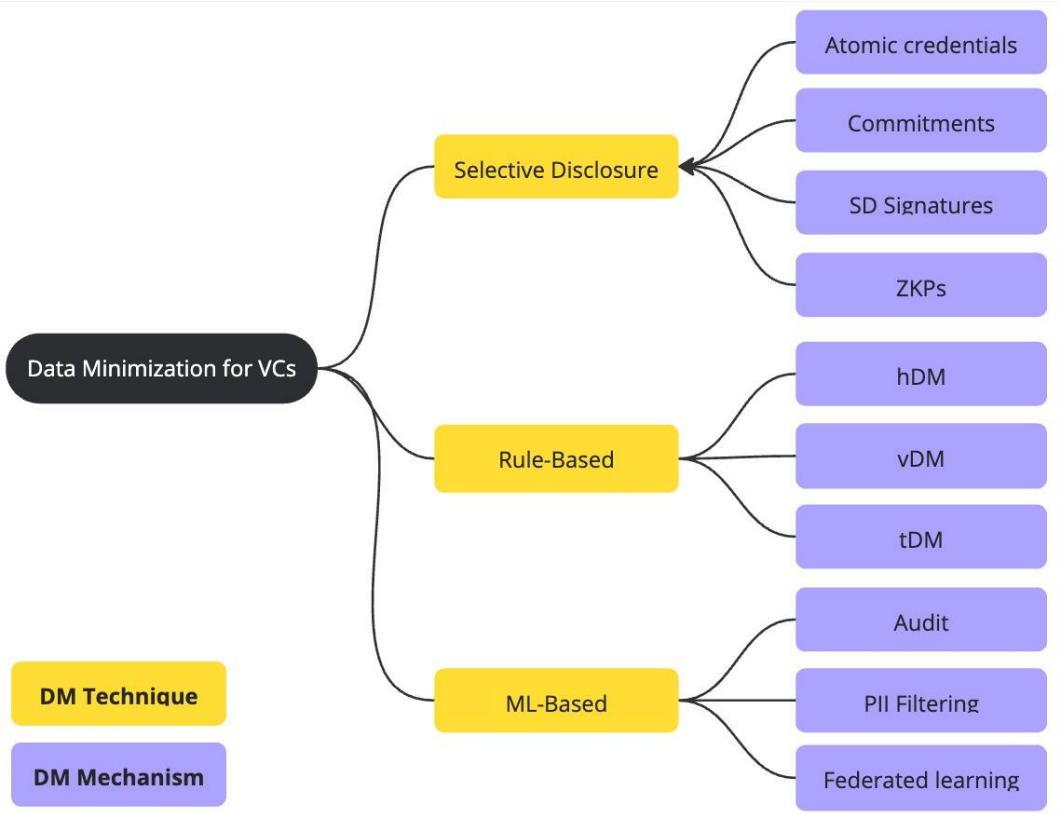
linkedin.com/in/norberthenseler

LinkedIn

Deakin University

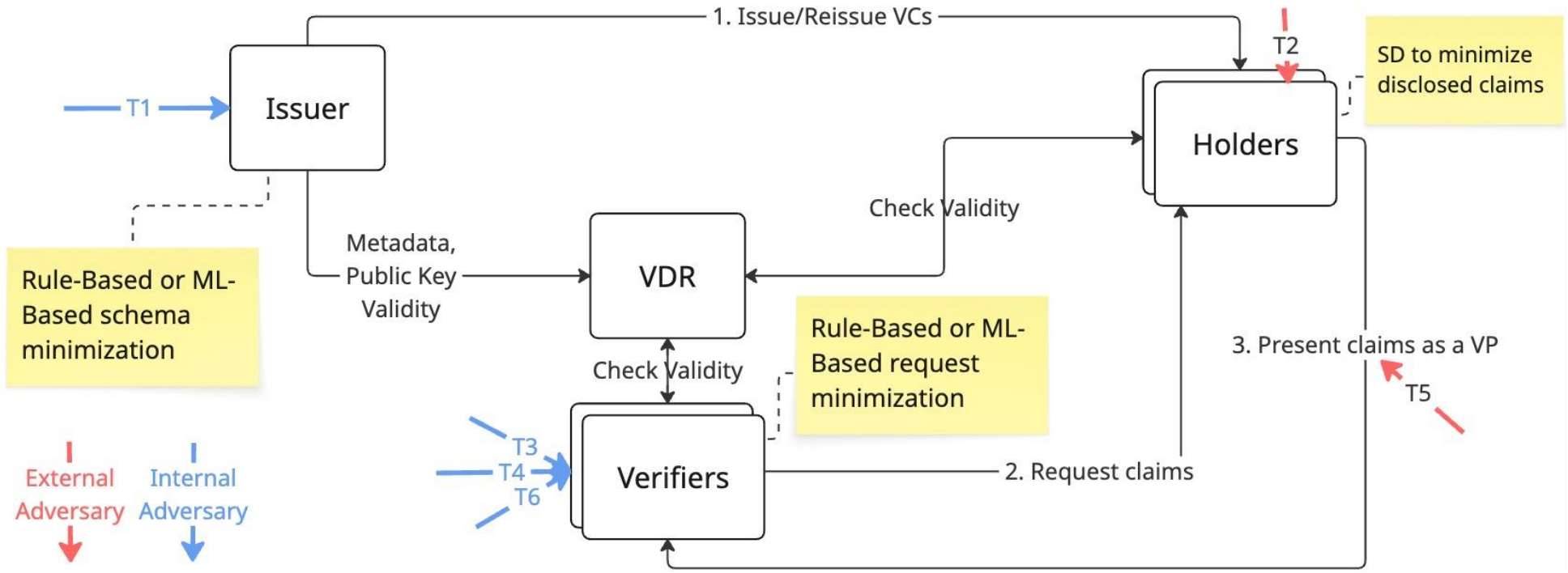
CSIRO

Appendix B: Taxonomy of DM Techniques and Mechanism



DM techniques: VC-specific SD, transferable Rule-Based DM, and ML-Based DM.

Appendix C: Role- and Lifecycle-Aware Framework



Holder-side selective disclosure alone is insufficient. Issuers determine what can be disclosed; holders control what is disclosed, and verifiers are responsible for what is requested.

Appendix D: Comparative Evaluation of DM Mechanisms

				Criteria					
	Technique	**Mechanism**	**n**	**Privacy Impact**	**Efficiency**	**Scalability**	**Domain Transferability**	**Regulatory Alignment**	**Quantum Resilience**
Taxonomy	**SD**	Atomic credentials	5	Medium	Medium	Low	Low	Medium	Implementation-dependent
		Commitments/hash-based	7	Medium	High	Medium	Low	Medium	Medium
		SD Signatures	9	High	Medium	Medium	Low	Medium	Low
		ZKP	10	High	Low	Medium	High	High	Medium
	Rule-Based	hDM	3	Low	High	High	Medium	Medium	Not inherent
		vDM	6	High	High	Medium	Medium	High	Not inherent
		tDM	4	High	Medium	Medium	Medium	High	Implementation-dependent
	ML-Based	Audit	5	Low	High	High	Low	Low	Not inherent
		PII Filtering	2	Medium	Medium	Low	Low	Medium	Not inherent
		Federated learning	4	Low	Low	Medium	Low	Low	Implementation-dependent

ZKPs offer strong transferrable privacy, but may lack efficiency and scalability

Rule-base mechanisms are better suited for issuer and verifier-side DM.