Deflating Mass-Surveillance Attempts in the Post-Snowden Era: Publicly-Traceable Conditional Decryptions

Francesco Bruschi[1], Marco Esposito[2], Andrea Rizzini[3] and Ivan Visconti[4]

[1]*Politecnico di Milano, professor and researcher at DEIB department*
[2]*Politecnico di Milano, PhD student at DEIB department*
[3]*Politecnico di Milano, PhD student at DEIB department*
[4]*Sapienza Università di Roma, professor and researcher at DIAG department*

Abstract

This paper presents the notion of *conditional backdoor*, a cryptographic paradigm that enables transparent, verifiable access to encrypted data based on predefined conditions. It replaces traditional backdoors with secure, auditable mechanisms leveraging witness encryption and blockchain-based enforcement. By treating lawful access as a computable predicate, the model aligns privacy preservation with regulatory compliance and accountability.

Keywords

Conditional Backdoor, Strong Privacy, Lawful Access, Blockchain, Witness Encryption, Auditability

1. Introduction

The concepts of privacy and security are intimately related. Throughout history, this relationship has permeated philosophical, sociological, technological, and political discourse, evolving across multiple domains of inquiry. Ancient societies viewed privacy primarily as personal autonomy and solitude, whereas security emphasized protection from external threats to the public and the individual alike [22, 28]. As technological capabilities for surveillance and mass communication evolved, privacy came to denote the state in which personal information remains concealed. Security, in turn, has come to refer to the tools and mechanisms that safeguard this state. From this perspective, security becomes the enabler of privacy: to keep something private is to secure it from unauthorized access. At its most extreme, security implies complete opacity—shielding not just information, but the very fact of its existence. Treating personal data as a secret implies that it might be disclosed to trusted others—while simultaneously raising the specter of disproportionate control and surveillance by those entrusted with maintaining that security [20]. As the second quarter of the 21st century approaches, the growing prevalence of cyber warfare, state-backed terrorism, and transnational criminal networks has compelled the European Union and its partners to pursue coordinated multilateral efforts to harden digital infrastructure across jurisdictional boundaries [8]. These efforts unfold in the aftermath of the post-9/11 U.S. policy era [24], during which the "nothing to hide" narrative framed privacy and security as inherently conflicting objectives. More recently, this binary trade-off model has been critically re-evaluated to avert unnecessary compromises of fundamental rights [27]—rights that remain at the normative core of the EU's legal order [16]. This reconsideration found partial legal articulation also in the USA, which curtailed aspects of bulk metadata collection and signaled recognition that unchecked surveillance erodes civil liberties [23]. Yet this recalibration was not a universally shared legal or normative shift. Around the same period, the UK enacted the IPA [25]—building upon the RIPA [26]—which expanded state powers through mandated bulk data collection and introduced the controversial "double-lock" authorization mechanism, combining ministerial approval with judicial

DLT2025: 7th Distributed Ledger Technology Workshop, June, 12-14 2025 - Pizzo, Italy

francesco.bruschi@polimi.it (F. Bruschi); marco.esposito@polimi.it (M. Esposito); andrea.rizzini@polimi.it (A. Rizzini); ivan.visconti@uniroma1.it (I. Visconti)
0000-0001-5669-2788 (F. Bruschi); 0000-0001-6984-8569 (M. Esposito); 0009-0002-9652-6703 (A. Rizzini); 0000-0003-2381-5846 (I. Visconti)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2025-11-24

oversight for surveillance warrants. In parallel with these regulatory developments, attention has turned to the means by which fundamental rights—particularly the right to privacy—can truly be preserved. Privacy-preserving technologies represent a second axis of this debate, revealing how security and privacy can be orthogonal rather than oppositional [21]. Encryption, in particular, has exemplified this relationship for decades. It enables the construction of secure systems that do not rely on centralized trust or discretionary access controls, but rather on asymmetric capabilities grounded in formal adversarial models [6]. In other words, modern cryptography rearranges power by putting ”the knife in users' hands”: even individuals with limited computational resources can encrypt messages that remain secure against any realistic adversary, including those with nation-state capabilities [17].

Our contribution. We consider the notion of *conditional backdoor* and explore its role in regulating access to encrypted data. We treat *identifiability* as a special case where decryption reveals the subject's identity, and we frame this within a cryptographic threat model involving authorized and adversarial entities. We further formalize the requirement that decryption must leave a verifiable trace, and investigate whether current or emerging technologies can enforce access conditions while preventing undetectable decryption. Finally, we discuss the potential of identifiability as a compliant-by-design approach to lawful access, like in the context of KYC frameworks.

1.1. The “good” backdoor problem

Criminal and terrorist networks have adapted quickly to digital tools, now coordinating through encrypted messaging apps rather than physical meetings [5]. This shift has prompted many governments to pursue new regulations for lawful access. Yet encryption schemes like AES produce ciphertext indistinguishable from random noise. This makes it technically and legally questionable to assume the presence of encrypted content based solely on its appearance—yet some legislative frameworks implicitly rely on this assumption when mandating key disclosure under penalty. How can one be compelled to decrypt something that cannot even be proven to be ciphertext? A possible answer lies in *anamorphic encryption* [15], which allows users to embed a hidden additional plaintext to a regular plaintext into a ciphertext, disclosing only the regular one under coercion. This technique essentially involves hiding a plaintext—an approach that is generally steganographic in nature and can be trivially implemented whenever sufficiently long random strings can be exchanged. In fact, encryption schemes like AES have pseudorandom ciphertexts that can therefore replace random strings wherever they are used. Even the popular TLS protocol includes random strings exchanged by client and server. Moreover, invoking such techniques may backfire: in adversarial jurisdictions, the mere availability of alternative decryption paths could be treated as intent to obstruct justice or grounds for escalated penalties. To frame the broader problem, we denote by D a hypothetical universal decryptor enabling lawful access to protected data. D is not a technical artifact per se, but a policy ambition that could be instantiated in multiple ways—for example:

- Imposing key length limits, making ciphertext breakable with sufficient computational resources (though not necessarily only by authorities);
- Mandating key escrow, e.g., publishing encryption keys encrypted under the authority's public key.

In either case, implementation depends on user cooperation, which might be legally enforced but cannot be cryptographically guaranteed. Without such collaboration, D remains aspirational. Focusing on the escrow model, these systems grant authorities two problematic powers:

1. *Arbitrary decryption*: Once in possession of the private key, authorities can decrypt data unconditionally, regardless of legal conditions or time constraints.
2. *Invisible access*: Decryption leaves no forensic trace—no cryptographic signal that access has occurred, nor any evidence enabling attribution. The process is opaque and unaccountable by design.

Even assuming trustworthy governance, these issues persist. A system cannot be secure *only* because its current administrators are benevolent [19]. Insider abuse, data leaks, and political regime change all undermine the viability of a "good" backdoor. Historically proposed mechanisms such as **key escrow** or **mandated backdoors** have consistently introduced structural vulnerabilities [1, 3]. They either centralize risk or embed opaque override channels that become attack surfaces in their own right.

A more robust direction may involve designing access mechanisms that are *verifiable*, *conditional*, and *publicly auditable*. Traditional architectures are fundamentally ill-suited for this task. By contrast, public blockchains offer immutable ledgers, decentralized enforcement, and consensus-based logic that could, in principle, condition decryption on observable public events. While this does not yet resolve the problem, it meaningfully shifts it—from silent institutional override to accountable cryptographic access.

2. Framework analysis

Privacy is often framed as a tradeoff between control and access, but in technical systems, we can formalize it more rigorously. We start by defining *strong privacy* as a set of enforceable and composable guarantees that resist unauthorized access not through trust or policy, but through cryptographic hardness and public verifiability. A central motivation for strong privacy is *compliance*—not merely as conformity to regulation, but as the capacity of a system to provably uphold constraints on data access. This is especially salient when identification or attribution is involved. It entails lawful, auditable, and minimally identifying access, consistent with the principle of contextual integrity, where privacy violations arise when information flows deviate from role- and context-specific norms [14, 2]. We say a system offers strong privacy if it satisfies the following high-level properties:

1. **Semantic security:** Observing outputs—whether ciphertexts, sanitized statistics, or intermediate states—should not allow an adversary to infer protected information significantly better than without such access.
2. **Quantifiable risk:** The system defines a measurable and enforceable risk model. Depending on the context, this may take the form of statistical guarantees, computational indistinguishability, or incentive-compatible deterrents (e.g., economic penalties or game-theoretic disincentives).
3. **Auditability:** Every authorized access to protected data—whether through decryption, query execution, or statistical release—must be observable and attributable within the system's trust and threat model.
4. **Temporal resilience:** Privacy holds not just at time t_0 but remains robust as auxiliary information accumulates. Strong privacy anticipates $t_1 > t_0$ scenarios, including AI inference, metadata leakage, or post-quantum cryptanalytic advances.

By contrast, we argue that *weak privacy* systems rely on institutional trust, discretionary enforcement, or static policy assumptions that cannot be computationally or statistically enforced. Such systems include traditional access control regimes, opaque statistical disclosure frameworks, and "good" backdoor proposals (see Section 1.1) that assume trustworthy governance. Privacy thus becomes a question of *who can decrypt*, *under what conditions*, and *with what accountability*. Next, we explore technical mechanisms that offer fine-grained, programmable control to realize the above guarantees.

2.1. Properties of Conditional Backdoors

We consider a class of access mechanisms—*conditional backdoors*—relevant to systems where strong privacy is realized through encryption-based enforcement. The goal is to enable access to encrypted content under precisely defined and publicly verifiable conditions. These mechanisms operate within our broader framework of strong privacy (Section 2), but introduce additional structural constraints centered on conditionality and auditability.

Definition 1 (Conditional Backdoor)**.** A *conditional backdoor* consists of algorithms and protocols enabling a designated authority $\mathscr{A}$ to decrypt a ciphertext *ct* **if and only if** it knows a witness w such that $P(s, w) = 1$, where s encodes the access conditions specific to the user who encrypted the data (e.g., derived from a blockchain state root representing a judicial authorization for that particular user), and P is an efficient relation. The intent of decrypting must leave a publicly accessible trace certifying the attempt to access *ct*. This attempt, if legitimate, will participate in the generation of w.
Since in this paper we are only presenting some initial results of our ongoing research, we will present informal definitions only. We list now desiderata:

1. **Conditionality**: $\mathscr{A}$ can decrypt a ciphertext *ct if and only if* $P(s, w) = 1$ and w is known to $\mathscr{A}$. The relation P must encode lawful access conditions. The condition must be:
   - **Expressive**, allowing for composable logic (e.g., time locks, multi-party authorizations).
   - **Publicly decidable**, enabling anyone to check whether the decryption condition has been met.
2. **No Arbitrary Access** (Admissibility): There must exist no efficient algorithm $\mathscr{A}^*$, even colluding with infrastructure operators (e.g., cloud storage providers) and the authority that enables decryption without knowing a witness w satisfying $P(s, w)$ where s is the condition considered when the ciphertext was computed.
3. **No Silent Access** (Auditability[1]): To enable the decryption of *ct* originally encrypted for a condition s, the authority $\mathscr{A}$ must publicly leave a trace and this will lead to the generation of w such that $P(s, w) = 1$.
   The goal is to ensures the request of the authority be:
   - **Detectable**: Any concrete decryption capability must be publicly known.
   - **Timestamped**: The trace associated to the request includes a verifiable timing of the request.
   - **Attributable**: The trace of the request identifies the requester.
4. **Robustness (resilience to state manipulation)**: An efficient malicious $\mathscr{A}$ must not be able without a publicly observable trace to compute w such that $P(s, w) = 1$ and s is the condition that a user used to compute the encryption.

2.2. Conditional backdoors via witness encryption

We now discuss a direction that can lead to the construction of a conditional backdoor framework using some cryptographic tools. Given the properties of strong privacy in Section 2, a canonical approach is to enforce access through public predicates over verifiable states. The seemingly natural cryptographic tool for this purpose is witness encryption (WE) [9]. In this setting, the access policy is encoded as a predicate $P(s, w)$, where s is a public statement describing a verifiable condition (e.g., a court permission event has been recorded or the time-lock delay has elapsed), and w is a witness proving that statement s holds. WE enables the ciphertext to remain undecipherable until such a statement becomes true (i.e., until a witness exists proving the statement holds).

Witness Encryption Preliminaries. We briefly recall the notion (from Definition 3.1 in [9]) of witness encryption for an NP language L with corresponding witness relation R. A WE scheme consists of two algorithms:

- $\mathsf{WE.Enc}(1^\lambda, s, m)$: Takes a security parameter 1^λ, a statement s, and a message m, and outputs a ciphertext *ct*.
- $\mathsf{WE.Dec}(ct, w)$: Takes a ciphertext *ct* and a witness w, and outputs a message m or $\perp$.

[1] In practical implementations, the trace might be initially visible only to designated auditors, with public disclosure mandated after a fixed delay Δ. However, this weakens the ”No Silent Access” guarantee during the interval $[t, t + \Delta]$.

Correctness. For any security parameter λ, for any $m \in \mathcal{M}$ (i.e. message space), and for any $s \in L$ such that $R(s, w)$ holds, we have:

$$\Pr\left[\mathsf{WE.Dec}\left(\mathsf{WE.Enc}(1^{\lambda}, s, m), w\right) = m\right] = 1 - \mathrm{negl}(\lambda)$$

For simplicity we will sometimes omit the security parameter. For technical reasons we will need a stronger form of witness encryption (i.e., extractable) and that for the sake of simplifying the notation this will remain implicit. In our construction, we instantiate WE where the witness relation R corresponds to our predicate P. Thus, $\mathsf{WE.Enc}(s, \cdot)$ encrypts under statement s, and decryption succeeds when provided a witness w such that $P(s, w) = 1$.

Construction 1 (Instantiation via WE). *We say that a ciphertext $ct = (ct_1, ct_2)$ from user U_i is conditionally backdoored if:*

- $ct_1 = \mathsf{WE.Enc}(s_i,\ \mathsf{Enc}(pk_{\mathsf{aid}}, k))$ *is a witness–encryption of* $\mathsf{Enc}(pk_{\mathsf{aid}}, k)$ *under the predicate $P(s_i, w)$, where s_i is the* user–specific statement *and* $\mathsf{Enc}(pk_{\mathsf{aid}}, k)$ *is a public-key encryption of a fresh symmetric key k under the authority's public key pk_{aid};*
- $ct_2 = \mathsf{Sym.Enc}(k, m)$ *is a symmetric encryption of message m under key k;*
- *Decryption of ct is computationally feasible if and only if the decryptor possesses both:*
    1. *A witness w such that $P(s_i, w) = 1$ (for this specific user U_i), and*
    2. *The secret key sk_{aid} corresponding to pk_{aid}.*

Notation.

- pk_{aid} / sk_{aid} — public/secret key of the (single) authority aid;
- s_i — public statement bound to user U_i (e.g. "there exists a finalised block from a checkpoint such that a storage slot contains an authorization for U_i");
- $P(s_i, w)$ — predicate expressing the decryption policy for U_i.

Remark 1 (Concrete instantiation of the predicate). *In practice, one can instantiate the pair (P, w) as follows.*

Statement. *Let $uid := \mathsf{ID}(U_i)$, we write*

$$s_i = (B,\ sl_{\min},\ \mathsf{aid},\ uid,\ h),$$

where B is the canonical blockchain prefix observed at encryption time, $sl_{\min}$ is the minimum slot index from which the access request may appear, aid *identifies the requesting authority, uid is the target user identifier, and h is a commitment to the access-request details.*

Witness. *Let*

$$w = (\sigma_{\mathsf{aid}},\ \pi,\ \tau),$$

where σ_{aid} is a valid signature by aid *on the request, π is a proof of legal authorisation (e.g. a signed court order or a Merkle inclusion proof within an authorisation registry), and τ is the on-chain transaction identifier logging the request.*

Predicate. $P(s_i, w) = 1$ *iff*

1. *an on-chain transaction τ exists in B at slot $sl \geq sl_{\min}$, signed by* aid*, requesting access to uid's data and embedding h;*
2. *σ_{aid} verifies under pk_{aid};*
3. *π attests that* aid *was legally authorised before slot sl (e.g. via a prior court-order transaction or a registry inclusion proof).*

All checks are polynomial-time, hence $L_P = \{\, s_i \mid \exists w\ :\ P(s_i, w) = 1\}$ lies in NP.

The explicit presence of pk_{aid} guarantees that, even when P is satisfied and anyone can open the witness encryption, *only* the authority holding sk_{aid} can ultimately recover m (see decryption below). Moreover, the user-specific component *uid* inside s_i enforces *selective* access: a witness w valid for U_i cannot satisfy $P(s_j, w)$ for any $j \neq i$.

Encryption algorithm.

1. **Key encapsulation:** Compute $ct_k = \mathsf{Enc}(pk_{\mathrm{aid}}, k)$ and then
$$ct_1 = \mathsf{WE.Enc}(s_i, ct_k).$$
2. **Data encryption:**
$$ct_2 = \mathsf{Sym.Enc}(k, m).$$

The resulting output is:
$$ct = (ct_1, ct_2)$$

Decryption algorithm. Given ct and a valid witness w for U_i:

1. Anyone computes $ct_k = \mathsf{WE.Dec}(ct_1, w)$;
2. Only the authority derives $k = \mathsf{Dec}(sk_{\mathrm{aid}}, ct_k)$;
3. Recover $m = \mathsf{Sym.Dec}(k, ct_2)$.

Illustrative predicate families.

- *Inclusion predicates:* $P(s, w) = 1$ if w proves inclusion of a value in a finalised public state (e.g. a court-signed order).
- *Zero-knowledge predicates:* $P(s, w) = 1$ if w attests, in zero knowledge, that procedural or jurisdictional conditions hold.
- *Temporal predicates:* $P(s, w) = 1$ if w proves that a delay Δ has elapsed since a timestamp committed earlier on-chain.

Quantifiable Risk. As introduced in Section 2, our definition of strong privacy requires that access violations carry measurable and enforceable consequences. In this ideal construction, the risk of unauthorized access is quantified as a deterrence function:

$$R(P, A) = (\Pr[\texttt{slash}(A)], \texttt{Cost}(A), \texttt{Trace}(A))$$

where A is an access attempt and P is the governing predicate. This tuple captures: (1) the probability that a violation is detected and punished (e.g., slashing or exclusion), (2) the economic cost imposed on violators, and (3) the degree to which the access action is observable.

A system satisfies the *Quantifiable Risk* property if, for all $A \not\models P$, we have $R(P, A) \geq \theta_{\min}$ for some deterrence threshold $\theta_{\min}$. This ensures that adversarial access becomes either computationally infeasible, economically irrational, or publicly accountable.

2.2.1. Adversary models and assumptions

We consider PPT adversaries attempting to recover m without proper authorization. Threats include:

- Secret forking to simulate satisfying states.
- Collusion with infrastructure (e.g., TEEs or validators).
- Exploiting improperly scoped predicates.

Our approach in the construction relies on the security of the underlying encryption schemes, and on the fact that the public state s is tied to a finalized, immutable source (e.g., on-chain finality). The above can intuitively guarantee that no efficient adversary can recover m from $ct = (\mathsf{WE.Enc}(s, \mathsf{Enc}(pk, k)), \mathsf{Sym.Enc}(k, m))$ unless it knows w such that $P(s, w) = 1$ (i.e. $s \in L_P$, the NP language induced by P) and possesses the corresponding secret key sk. Using extractable witness encryption, this guarantee is strengthened: any successful decryption implies the adversary possesses an efficiently extractable witness[2].

This framework establishes an idealized interface—decryption conditioned on arbitrary NP predicates—whose security is purely cryptographic. It serves as a design target for partial realizations that reinterpret the witness relation using time, consensus, or attestation mechanisms. We now turn to concrete systems that approximate this ideal through engineering compromises and domain-specific assumptions.

2.3. Workable Implementations

While the notion of WE posits a powerful ideal—encryption under the hardness of arbitrary NP problems—its realization under standard cryptographic assumptions remains elusive [9]. Most candidate constructions rely on indistinguishability obfuscation (iO) or multilinear maps, both of which face uninstantiated assumptions or impracticality for deployment [12]. Nonetheless, the rise of decentralized systems has inspired alternative approaches that approximate WE under social or cryptoeconomic assumptions, such as honest or rational majorities, or secure hardware. One system [18] uses smart contracts and a semi-trusted committee to emulate WE via verifiable secret sharing, enforcing correctness with zero-knowledge proofs and on-chain slashing. Because the security is economic rather than cryptographic, admissibility (CB.2) is only partial: a colluding threshold of committee members can decrypt the secret off-chain and leak it early, bypassing the on-chain predicate. Likewise, robustness against state manipulation (CB.4) is only partial, since share censorship or a fork prior to finality can reorder or omit the witness data and thereby satisfy—or delay—the predicate in an adversary-controlled branch. A similar honest-majority approach underlies DPSS-based systems [11], where secrets are stored and conditionally released via dynamic committees of blockchain miners; here, conditionality is enforced through on-chain predicates and zero-knowledge proofs, but admissibility, auditability, and robustness (CB.2–CB.4) all remain partial, as off-chain collusion or pre-finality reorgs can still undermine guarantees. The limitation specific to CB.4 does not apply to McFly [7]. McFly binds the ciphertext to the public key of the committee that will exist at height $h + \Delta$, a value that becomes immutable once block h is finalized; namely, an adversary cannot craft an alternative or premature chain state without reorganizing the chain past finality. Other related works such as[4] explores timed-release encryption as a special case of WE, where time itself acts as the witness, using anonymous committees and PRF chains to enable scalable and incentive-aligned disclosure. With this type of constructions, admissibility or auditability are not fully realized: early decryption remains possible for powerful adversaries (violating CB.2), and decryption can occur privately without leaving any public trace (violating CB.3). Unlike classical timelocks, recent work on timestamp-hiding commitments [13] further extends this line by using zero-knowledge proofs over incremental Merkle trees to prove time elapsed without disclosing absolute timestamps, adding a privacy-preserving axis to delay-based WE approximations. Here CB.4 would be largely guaranteed in practice, though small timing advantages remain possible due to allowed timestamp skew by block producers (e.g., ±13 seconds on Ethereum).

These and other constructions can be situated within a broader space of pragmatic WE approximations, characterized by their witness models, trust assumptions, timing guarantees, and resilience to early decryption (i.e., systems where a committee member or TEE could decrypt before the predicate is truly satisfied). Table 1 evaluates recent implementations not only by their witness mechanism but also by the extent to which they approximate the four cryptographic desiderata defined in Section 2.1.

[2] When no valid witness exists (i.e., $\forall w : P(s, w) = 0$), the witness encryption's soundness property ensures that the ciphertext reveals no information about m, even to computationally unbounded adversaries who do not possess sk. When a valid witness exists but the adversary does not know it, the adaptive witness indistinguishability of the WE scheme guarantees that, even if the adversary can choose statements and make decryption attempts adaptively, it cannot distinguish which of several valid witnesses (if any) underlies the ciphertext. For stronger guarantees, extractable witness encryption [10] ensures that any adversary capable of breaking the encryption must possess—in an extractable sense—a valid witness, preventing circumvention of access conditions through cryptanalytic means.

Implementation	Witness Type	CB.1	CB.2	CB.3	CB.4
		Conditional	*Admissible*	*No Silent Access*	*No State Manipulation*
Witness Encryption	NP instance	✓	✓	✓	✓
TEE-Based	Local attestation	✓	✗	Partial[3]	✗
Optimistic PVSS [18]	zk witness + committee	✓	Partial	✓	Partial
FaB-DPSS / eWEB [11]	DPSS + on-chain predicate	✓	Partial	Partial	Partial
McFly [7]	Committee-based future event	✓	✓	✗[4]	✓
Timed-Release [4]	VDF (time delay)	✓	Partial	✗	✓
Proof-of-Time [13]	Time-withheld zk proof	Partial[5]	✓	✓	Partial
Threshold Decryption[6]	Explicit submission	✓	Partial	✓	✗

Table 1
Pragmatic approximations to Witness Encryption, evaluated against the formal properties of conditional backdoors (CB.1–CB.4, see Section 2.1). A ✓ indicates full support; Partial denotes reliance on external trust or enforcement; ✗ indicates the property is not ensured in the threat model.

This reclassification reveals which systems offer enforceable access control with minimal trust and which rely on coordination, economic assumptions, or unverifiable enforcement. We argue that a promising direction would be to combine TEEs with optimistic, on-chain enforcement to approximate all four desiderata with minimal trust surface. In such a hybrid design, the TEE enforces conditional decryption locally and attests to witness satisfaction, while an on-chain smart contract accepts decryption outputs only if they are accompanied by attestations that can be challenged during a bounded dispute window. A fraud-proof mechanism—backed by reproducible witness evaluation or replayable transcripts—would ensure auditability (CB.3), mitigating the opacity of the enclave. By anchoring commitments and outcomes to an append-only ledger, and requiring slashing for misbehavior, such a system could also deter unauthorized access (CB.2) and offer partial resistance to state manipulation (CB.4), depending on the underlying chain's finality. Though still pragmatic, this construction would reduce reliance on any single trust assumption and leverages hardware only as a runtime enforcement layer, bounded by verifiable cryptoeconomic guarantees.

3. Call to action (conclusions)

This paper introduced the concept of conditional backdoor as a structured approach to reconciling privacy preservation with the need for regulated access to encrypted data. Rather than relying on traditional "good" backdoors often associated with unconditional and opaque access to encrypted data by the authorities, we explored a model where access is tied to verifiable, public conditions and enforced through cryptographic mechanisms. We have argued that embedding access logic into transparent and auditable protocols represents a meaningful shift: from trust-based assumptions to systems where legal and technical guarantees can coexist. The proposed framework is intended to reach a broad set of stakeholders in the research community, from policymakers and regulatory bodies to crypto-security engineers.

In this call to action we propose several priorities for further works:

1. **Decouple access from identification**, enabling data retrieval without default exposure of personal identities.
2. **Anchor access conditions to public blockchains**, ensuring traceability and independent verifiability.

[3] Auditability for TEE-based designs depends on external mechanisms such as host-side logging or remote attestation. No cryptographic audit signal is emitted from the enclave itself.

[4] Auditability is not cryptographically enforced: a dishonest committee could collude and decrypt off-chain without leaving a public trace. Full auditability would require protocol-level enforcement of on-chain decryption shares.

[5] A smart contract can enforce "access iff elapsed Δ," but nothing prevents the committer from leaking the secret off-chain before they submit a proof. Thus decryption would not cryptographically bound to the on-chain proof.

[6] Threshold-based schemes that are *not* anchored to a public blockchain state offer simplicity and clear conditional logic, but they rely purely on off-chain and honest-majority enforcement. This gives only partial admissibility (CB.2), and fails to protect against predicate forgery or state rewriting (CB.4).

3. **Deploy techno-legal pilots** under optimistic models with economic constraints (e.g., staking-based enforcement).
4. **Advance research and investment** in cryptographic primitives supporting fine-grained, programmable access control.
5. **Mandate traceable and contestable access**, ensuring that every authorized action is publicly observable and verifiable.

This shift can help rebuild trust, make oversight more transparent, and create a healthier balance between security and individual freedoms.

Acknowledgments. Ivan Visconti is member of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica.

Declaration on Generative AI In this work LLMs (specifically, ChatGPT) were used exclusively for grammar checks and the stylistic polishing of certain sentences. All outputs were reviewed by the authors to ensure that the original meaning was preserved and that no additional content was introduced by the LLM. In this paper, LLMs were NOT used for brainstorming or for generating original, non-human content.

References

[1] Hal Abelson et al. “The risks of key recovery, key escrow, and trusted third-party encryption”. In: *World Wide Web J.* (1997).

[2] Adam Barth et al. “Privacy and contextual integrity: Framework and applications”. In: *2006 IEEE symposium on security and privacy (S&P’06).* IEEE. 2006, 15–pp.

[3] Matt Blaze. “Protocol failure in the escrowed encryption standard”. In: *Proceedings of the 2nd ACM Conference on Computer and Communications Security.* 1994, pp. 59–67.

[4] Matteo Campanelli et al. “Encryption to the Future: A Paradigm for Sending Secret Messages to Future (Anonymous) Committees”. en. In: *Advances in Cryptology – ASIACRYPT 2022.* Ed. by Shweta Agrawal and Dongdai Lin. Vol. 13793. Series Title: Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2022, pp. 151–180. isbn: 978-3-031-22968-8 978-3-031-22969-5. doi: 10.1007/978-3-031-22969-5_6. url: https://link.springer.com/10.1007/978-3-031-22969-5_6.

[5] Matías Dewey and Andrés Buzzetti. “Easier, faster and safer: The social organization of drug dealing through encrypted messaging apps”. In: *Sociology Compass* 18.2 (2024), e13175.

[6] Whitfield Diffie and Martin E Hellman. “New directions in cryptography”. In: *Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman.* 2022, pp. 365–390.

[7] Nico Döttling et al. “McFly: verifiable encryption to the future made practical”. In: *International Conference on Financial Cryptography and Data Security.* Springer. 2023, pp. 252–269.

[8] European Commission. *ProtectEU: A European Internal Security Strategy.* Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2025) 148 final. Apr. 2025. url: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52025PC0148.

[9] Sanjam Garg et al. “Witness encryption and its applications”. en. In: *Proceedings of the forty-fifth annual ACM symposium on Theory of Computing*. Palo Alto California USA: ACM, June 2013, pp. 467–476. isbn: 978-1-4503-2029-0. doi: 10.1145/2488608.2488667. url: https://dl.acm.org/doi/10.1145/2488608.2488667.

[10] Shafi Goldwasser et al. “How to run turing machines on encrypted data”. In: *Advances in Cryptology–CRYPTO 2013: 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part II.* Springer. 2013, pp. 536–553.

[11] Vipul Goyal et al. “Storing and retrieving secrets on a blockchain”. In: *IACR International Conference on Public-Key Cryptography*. Springer. 2022, pp. 252–282.

[12] Aayush Jain, Huijia Lin, and Amit Sahai. “Indistinguishability obfuscation from well-founded assumptions”. In: *Proceedings of the 53rd annual ACM SIGACT symposium on theory of computing*. 2021, pp. 60–73.

[13] Alexander John Lee. “Proof of Time: A Method for Verifiable Temporal Commitments Without Timestamp Disclosure”. en. In: ().

[14] Helen Nissenbaum. “Privacy in context: Technology, policy, and the integrity of social life”. In: *Privacy in context*. Stanford University Press, 2009.

[15] Giuseppe Persiano, Duong Hieu Phan, and Moti Yung. “Anamorphic Encryption: Private Communication Against a Dictator”. en. In: *Advances in Cryptology – EUROCRYPT 2022*. Ed. by Orr Dunkelman and Stefan Dziembowski. Vol. 13276. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 34–63. ISBN: 978-3-031-07084-6 978-3-031-07085-3. DOI: 10.1007/978-3-031-07085-3_2. URL: https://link.springer.com/10.1007/978-3-031-07085-3_2.

[16] *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)*. Accessed: 16 April 2025. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

[17] Phillip Rogaway. “The moral character of cryptographic work”. In: *Cryptology ePrint Archive* (2015).

[18] Schwinn Saereesitthipitak and Dionysis Zindros. “Cassiopeia: Practical On-Chain Witness Encryption”. en. In: *Financial Cryptography and Data Security. FC 2023 International Workshops*. Ed. by Aleksander Essex et al. Vol. 13953. Series Title: Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2024, pp. 385–404. ISBN: 978-3-031-48805-4 978-3-031-48806-1. DOI: 10.1007/978-3-031-48806-1_25. URL: https://link.springer.com/10.1007/978-3-031-48806-1_25.

[19] Jerome H Saltzer and Michael D Schroeder. “The protection of information in computer systems”. In: *Proceedings of the IEEE* 63.9 (1975), pp. 1278–1308.

[20] Stefan Schuster et al. “Mass surveillance and technological policy options: Improving security of private communications”. In: *Computer Standards & Interfaces* 50 (2017), pp. 76–82.

[21] Yun Shen and Siani Pearson. “Privacy enhancing technologies: A review”. In: *HP Laboratories* 2739 (2011), pp. 1–30.

[22] Judith A Swanson. *The public and the private in Aristotle’s political philosophy*. Cornell University Press, 1994.

[23] U.S. Congress. *USA FREEDOM Act of 2015*. https://www.congress.gov/bill/114th-congress/house-bill/2048/text. Public Law No: 114-23. 2015.

[24] U.S. Congress. *USA PATRIOT Act of 2001*. Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act. 2001. URL: https://www.congress.gov/bill/107th-congress/house-bill/3162.

[25] UK Parliament. *Investigatory Powers Act 2016*. https://www.legislation.gov.uk/ukpga/2016/25/contents. Chapter 25. 2016.

[26] UK Parliament. *Regulation of Investigatory Powers Act 2000*. https://www.legislation.gov.uk/ukpga/2000/23/contents. Chapter 23. 2000.

[27] Govert Valkenburg. “Privacy versus security: Problems and possibilities for the trade-off model”. In: *Reforming European Data Protection Law* (2015), pp. 253–269.

[28] Samuel Warren and Louis Brandeis. “The right to privacy”. In: *Killing the Messenger: 100 Years of Media Criticism*. Columbia University Press, 1989, pp. 1–21.