AffixIO Technical Paper · WP-008
June 2026
affix-io.com

AFFIXIO WHITE PAPER · WP-008

Zero-Knowledge Proofs as GDPR Article 25 Infrastructure: Data Minimisation by Design

When the schema has nowhere to put PII, minimisation is real.

AffixIO | United Kingdom | affix-io.com | June 2026

ABSTRACT

Article 25 asks for privacy by design, not checkbox compliance. Zero-knowledge proofs make minimisation structural: witnesses enter the prover, only digests hit storage. We explain how AffixIO's record service schema enforces this by construction.

CONTENTS

SECTION 1

Introduction

GDPR Article 25 is the data protection by design and by default requirement. It has been in force since May 2018 and is widely cited in data protection literature, but its practical implementation is poorly understood. Most organisations satisfy Article 25 through procedural measures: privacy reviews in the software development lifecycle, data minimisation assessments when launching new data processing activities, and access control configurations that limit who can access personal data. These are legitimate and necessary measures. They are also fundamentally policy-dependent: they work as long as the policies are followed and the configurations are maintained correctly.

The ICO's guidance on Article 25 recognises that "by design" means that privacy considerations must be integrated into the design of systems and business processes, not bolted on after the fact. But "integrated into the design" is often interpreted as "addressed during the design process" rather than "enforced by the design of the system." A system that addresses privacy during design but does not enforce privacy through its structure can have its privacy properties changed by configuration without triggering the same design review process.

Zero-knowledge proofs offer a qualitatively different implementation of Article 25. In a ZK proof system, the personal data that is the input to a computation is a private witness. The output is a proof that the computation was performed correctly. The proof is public; the input is not. The system cannot be configured to reveal the input without replacing the proof system with a different one, because the ZK property is intrinsic to the proof system's mathematical structure. Data minimisation is not a configuration option; it is a mathematical property of the system.

SECTION 2

Article 25 Requirements

Article 25 has two distinct paragraphs with distinct requirements. Article 25(1) requires that, at the time of determining the means for processing and at the time of the processing itself, the controller shall implement appropriate technical and organisational measures designed to implement data minimisation principles effectively. Article 25(2) requires that the controller implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.

The EDPB guidelines on Article 25 identify seven data protection principles that the by-design and by-default requirements must implement: data minimisation, purpose limitation, storage limitation, accuracy, confidentiality and integrity, accountability, and transparency. For each principle, the guidelines require both "by design" measures (built into the system) and "by default" measures (the default settings protect privacy without requiring action).

Article 25(3) provides a certification mechanism: controllers may demonstrate compliance by adherence to approved codes of conduct or certification mechanisms. No certification mechanism specifically for ZK proof systems has been approved, but the analysis in this paper suggests that ZK proof systems would satisfy the technical criteria for Article 25 certification better than most conventional approaches.

SECTION 3

Procedural vs. Structural Data Minimisation

Procedural data minimisation relies on policies, processes, and configurations to limit data collection and retention. A privacy impact assessment identifies excessive data collection; a remediation project removes the identified excess; a review cycle checks periodically that the remediation is maintained. This approach can be highly effective when consistently applied. Its weakness

is that consistency requires ongoing effort: configurations drift, policies are not followed in edge cases, and new features introduce new data collection without triggering the review process.

Structural data minimisation, by contrast, makes excessive data collection technically impossible through the design of the system. A form that has no field for date of birth cannot collect date of birth. A database schema that has no column for IP address cannot store IP addresses. A ZK proof system that accepts personal data only as private circuit witnesses cannot expose that personal data in its outputs, regardless of configuration.

The distinction matters for Article 25 compliance because "by design" implies that the privacy property is built into the system, not applied to it. A system with structural data minimisation satisfies the "by design" requirement more robustly than a system with procedural data minimisation, because the structural property cannot be undone by configuration change alone.

SECTION 4

ZK Proofs as Structural Data Minimisation

A ZK proof system enforces data minimisation through three structural properties. First, private witnesses are inputs to the circuit that are never included in the proof output. The mathematical structure of the ZK proof guarantees that knowing the proof does not enable a verifier to determine the witness values (beyond what is implied by knowing the circuit is satisfied). This is not a configuration option: it is inherent to the proof system's soundness properties.

Second, the circuit defines precisely what computation is performed on the witnesses. The circuit is a fixed, auditable program. Changing the computation requires changing the circuit, which produces a different circuit identifier. Anyone can verify that the circuit corresponds to a specific computation by reading the circuit source code. This provides transparency about what processing is performed, satisfying Article 25's transparency dimension.

Third, the proof output is derived from the witnesses but reveals nothing about them beyond the fact that they satisfy the circuit's constraints. The output is a compact proof object and the circuit's public output value. The proof object cannot be reverse-engineered to obtain the witnesses. This is a consequence of the cryptographic hardness assumptions underlying the proof system (specifically, the discrete logarithm problem on the BN254 curve).

> **Structural guarantee:** In AffixIO's governance system, AI response content and user inputs are private witnesses to ZK circuits. The governance record contains only proof digests and binary circuit outputs. This property cannot be changed by configuration: changing it would require replacing the ZK circuit with a different kind of computation.

SECTION 5

AffixIO Case Study

AffixIO's AI governance system processes AI responses that may contain personal data (names, health conditions, financial situations mentioned by users in their queries). The governance system must evaluate whether each response satisfies the governance policy without retaining the response content. The ZK circuit approach achieves this.

The AI response text is evaluated by the policy enforcement layer, which extracts binary properties (whether the response is on an approved topic, whether cited sources are verified, whether the response is within the permitted scope). These binary properties become the private witnesses to the ZK circuit. The response text itself is never supplied to the circuit; it is evaluated by the policy enforcement layer in memory and immediately discarded. The binary properties are field elements with no personal data content.

The circuit produces a proof certifying that the binary properties were correctly evaluated and that they satisfy the policy conditions. The proof and the binary circuit output (YES or NO) are stored in the governance record. The

AI response text is not stored anywhere in the governance system. If the AI response contained personal data (for example, a user mentioned their medical history), that personal data is not retained in the governance layer.

The result is a governance system that processes AI responses without being a personal data processor in respect of the response content. The GDPR Article 25 analysis for AffixIO's governance layer does not need to address response content data, because the system's design prevents it from storing that data.

SECTION 6

The "By Design" Requirement

The "by design" requirement in Article 25(1) means that privacy-protective technical measures must be incorporated into the system from the outset, not added after the processing has been designed. For AffixIO's ZK governance system, the "by design" requirement is satisfied by the choice to use a ZK proof system at the architecture design stage.

The key evidence for Article 25(1) compliance is the schema definition for the governance record. The schema has no fields for response content, user input, names, or any other personal data. This is not a configuration choice: it is part of the schema definition that was established at design time. The absence of these fields in the schema is the technical measure that implements data minimisation by design. Any future development work that attempted to add personal data fields to the schema would require a schema migration, which is subject to version control and code review.

The EDPB guidelines on Article 25 specifically identify "structural" and "architectural" measures as the preferred forms of by-design implementation. ZK circuit-based data minimisation is precisely the kind of structural, architectural measure that the guidelines recommend.

SECTION 7

The "By Default" Requirement

The "by default" requirement in Article 25(2) means that the default settings of the system must be maximally privacy-protective, without requiring the data subject to take action to protect their privacy. For AffixIO's governance system, the "by default" requirement is satisfied by the structural impossibility of storing personal data in the governance record.

The most demanding interpretation of "by default" would require that the system, without any action by the data subject or the data controller, processes only the minimum necessary data. AffixIO's ZK system satisfies this interpretation: the schema enforces data minimisation regardless of how the system is configured, what data is passed to it, or what actions the data subject or data controller take. There is no configuration option that enables personal data storage in the governance record. The minimum necessary processing is also the only possible processing.

SECTION 8

DPIA Implications

Article 35 of GDPR requires a data protection impact assessment (DPIA) for processing likely to result in a high risk to individuals' rights and freedoms. AI systems that process personal data at scale are typically subject to DPIA requirements. The ZK governance architecture substantially reduces the scope and complexity of the DPIA for the governance layer.

A DPIA for a conventional AI audit system that retains response content must assess the risks of storing personal data at scale, identify measures to mitigate those risks, and evaluate the proportionality of the storage against the purpose. A DPIA for AffixIO's ZK governance layer does not need to assess the risks of storing response content, because the system does not store it. The DPIA scope is reduced to the proof digests, binary circuit outputs, and hashed identifiers that the system actually retains, none of which contains personal data in the ordinary sense.

The reduced DPIA scope translates to reduced compliance cost and a simpler ongoing compliance programme. DPIAs must be reviewed periodically and when the processing changes materially. A governance system that does not process personal data is less likely to trigger DPIA review obligations than one that does.

SECTION 9

ICO Guidance Alignment

The ICO's guidance on data protection by design and by default (updated in 2023) identifies specific technical measures that satisfy Article 25. Several of these are directly implemented by ZK proof systems.

The ICO guidance lists "anonymisation and pseudonymisation" as by-design technical measures. ZK proof outputs are neither anonymised nor pseudonymised personal data; they are cryptographic objects derived from personal data but from which the personal data cannot be recovered. This is a stronger property than anonymisation in the GDPR sense, which requires that re-identification is not reasonably possible. For ZK outputs, re-identification is cryptographically impossible (under standard assumptions), not merely unreasonably possible.

The guidance also lists "encryption" as a by-design measure. The ZK proof system uses cryptographic commitment schemes and elliptic curve pairings that are at least as strong as encryption in their data protection properties. The ICO's encryption measure is intended to prevent unauthorised access to personal data; ZK's zero-knowledge property prevents any party, authorised or not, from recovering the personal data from the proof.

SECTION 10

Application Categories

ZK-based Article 25 compliance is most valuable in four categories of AI and digital services applications.

Regulated AI content systems that must audit AI responses for policy compliance without retaining the content of those responses (the AffixIO case). The ZK governance record proves compliance without storing the content being governed.

Identity and eligibility verification systems that must demonstrate the result of a verification without retaining the personal data used to perform it (the KYC case, see WP-006). The ZK eligibility proof demonstrates the result without storing the identity documents.

Health and social care systems that must demonstrate that eligibility decisions were correctly made (based on age, diagnosis, or entitlement criteria) without retaining the health information used in those decisions. ZK health eligibility circuits prove the decision without retaining the health data.

Age and eligibility verification for regulated content where the service must prove that age verification was performed without retaining the age documents that were checked (the age verification case, see WP-009).

SECTION 11

Limitations

ZK proof-based Article 25 compliance has two significant limitations. First, it applies to the governance or audit layer of a system, not to the primary processing layer. The AI system that generates responses may itself process personal data; the ZK governance layer addresses only the governance record. Article 25 compliance for the primary AI processing requires separate measures.

Second, the ZK system's data minimisation property depends on the correctness of the circuit and the adapter layer. If the adapter layer incorrectly includes personal data in the circuit's witness representation, that data would be processed (though not stored) by the circuit. The formal claim is that personal data is not stored in the governance record; the weaker claim that it is not processed at all cannot be made without analysing the full adapter implementation.

SECTION 12

Conclusion

Zero-knowledge proof systems implement GDPR Article 25's data protection by design and by default requirements more robustly than conventional procedural approaches, because data minimisation is a structural property of the cryptographic system rather than a policy or configuration. AffixIO's governance architecture demonstrates this in production: the governance record schema has no fields for personal data, and this absence is enforced by the schema structure, not by policy.

The regulatory implications are significant. A ZK-based Article 25 implementation simplifies DPIA scope, reduces data breach exposure in the governance layer to near zero, and provides a structural "by design" argument that is more defensible than a procedural one in regulatory inquiries. For AI governance systems, identity verification, and health eligibility systems, ZK-based Article 25 implementation should be considered the engineering baseline.

Related reading

- WP-006: PII-Free KYC by Design with Zero-Knowledge Identity Circuits
- WP-009: Privacy-Preserving Age Verification with Zero-Knowledge Proofs
- WP-003: The Proof-Not-Log Paradigm for AI Audit Trails

Frequently asked questions

What is data minimisation by design?

Building systems whose data model cannot accumulate unnecessary personal data, rather than relying on retention policies alone.

How do ZK proofs help ICO audits?

You demonstrate that stored records contain only cryptographic commitments verifiable against published keys.

Is consent still required?

Yes where lawful basis demands it; ZK reduces what you hold after consent is given.
