AffixIO Technical Paper · WP-006
June 2026
affix-io.com

AFFIXIO WHITE PAPER · WP-006

A PII-Free KYC Schema by Design: Structural Data Minimisation via Zero-Knowledge Identity Circuits

Pass KYC checks while your database stays empty of names and IDs.

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

ABSTRACT

KYC vaults become breach targets. AffixIO verifies eligibility attributes with ZK identity circuits so issuers see yes/no outcomes, not passport scans. GDPR Article 25 minimisation is enforced by the wire format.

CONTENTS

SECTION 1

Introduction

KYC is a regulatory requirement in financial services, gambling, age-restricted goods, and a growing range of professional services. It requires organisations to verify that customers are who they claim to be and that they satisfy eligibility conditions (age, residency, sanctions status, credit risk) before providing regulated services. The conventional implementation collects and stores the PII used to perform this verification, creating a database of sensitive personal data that is a primary target for data breaches and a source of significant regulatory compliance cost.

GDPR requires that personal data be collected only to the extent necessary for a specified purpose and retained only as long as necessary. For KYC, the specified purpose is eligibility verification. Once eligibility has been established, the personal data used to establish it is no longer necessary for the original purpose. However, because conventional KYC systems store PII in the same record as the eligibility verdict, they create a structural coupling: you cannot retain the eligibility evidence without retaining the PII. The result is that organisations retain KYC PII far longer than the verification purpose requires, because the PII is embedded in the compliance record.

AffixIO's ZK KYC approach breaks this coupling. The PII is processed by an identity verification service; the outputs of that verification (binary flags indicating that specific conditions are met) are passed to the ZK circuit as private witnesses; the circuit produces an eligibility proof. The eligibility proof and its Merkle anchor are the compliance record. The PII is not in the compliance record. The compliance record can be retained indefinitely without retaining PII.

SECTION 2

The KYC PII Problem

Conventional KYC systems collect and store the following categories of PII: full legal name, date of birth, residential address history, government identity document type and number, facial image or biometric data, and in some jurisdictions, tax identification number. This data is collected at onboarding and stored in the organisation's customer database for the duration of the customer relationship and for a mandatory retention period (typically five to seven years under AML legislation) after the relationship ends.

The risks this creates are well-documented. A data breach exposing KYC records can expose government document numbers, facial images, and address histories for the entire customer base, causing harm to customers that is extremely difficult to remediate (you cannot change your passport number easily). The GDPR fines for breaches of sensitive personal data at scale are substantial. The ICO's guidance on data minimisation explicitly identifies KYC data retention as a common area of non-compliance.

Two partial solutions are widely deployed but insufficient. First, data vault segregation stores KYC PII in a separate, more-restricted database, reducing but not eliminating breach exposure. The PII is still stored; it is just stored in a separate place. Second, pseudonymisation replaces PII with references to a separate lookup table. The pseudonymisation is reversible by the operator, and the combined system still contains all the PII.

The ZK approach is structural rather than procedural. By design, the compliance record contains only the eligibility proof. There is no field in the record for PII. The system cannot inadvertently retain PII in the compliance record because the schema does not support it.

SECTION 3

Identity vs. Eligibility Verification

Effective ZK KYC depends on clearly distinguishing between two questions that conventional KYC conflates. Identity verification asks: "Is this person the individual they claim to be?" This requires checking their biometric or documentary identity against a trusted source. It inherently involves PII processing. Eligibility verification asks: "Does this individual satisfy the conditions required to receive this service?" This requires knowing certain facts about the individual (their age exceeds a threshold, their country of residence is not sanctions-listed, their credit risk is within acceptable bounds) but does not require knowing the specific PII values that establish those facts.

The ZK insight is that eligibility verification requires only binary outputs from identity verification. "This person's date of birth indicates they are over 18" is an eligibility fact that can be derived from a date of birth without transmitting the date of birth itself. The circuit receives a single bit: verified_over_18 = 1 or 0. It does not receive the date of birth. The ZK proof proves that the bit was computed correctly (i.e., from a real date of birth that actually satisfies the threshold) without revealing the date of birth.

This separation is the core architectural insight. Identity verification remains within the identity verification service, which processes PII in accordance with its own data protection obligations and is deleted or minimised as quickly as the verification purpose allows. The eligibility outputs (binary flags) cross the service boundary and enter the ZK circuit. PII does not cross this boundary.

SECTION 4

The KYC Circuit Design

The AffixIO identity eligibility circuit receives four private binary witnesses corresponding to four eligibility conditions, and produces a single public output: the eligibility verdict.

Circuit implementation omitted from public documentation.

Each witness represents a condition evaluated by the identity verification service. `identity_verified` is 1 if the identity check passed (document valid, biometric match above threshold). `over_age_threshold` is 1 if the verified age meets or exceeds the configured threshold (18, 21, or 25 depending on the service). `not_sanctioned` is 1 if the individual's name, nationality, and document number were checked against applicable sanctions lists and produced no matches. `residency_permitted` is 1 if the verified country of residence is on the permitted list for the service.

The circuit produces YES (1) only if all four conditions are satisfied. If any condition fails, the circuit produces NO (0). The ZK proof certifies that the computation was performed correctly for the specific set of witnesses provided. The verifier knows the output (YES or NO) but not the individual witness values. A regulatory auditor examining the governance record knows that the KYC check was performed and whether it passed, but cannot determine from the governance record alone which specific condition caused a NO outcome.

SECTION 5

Interface with Identity Verification Services

The ZK KYC circuit does not perform identity verification. It consumes the outputs of identity verification. AffixIO's implementation is designed to integrate with standard commercial identity verification APIs (Onfido, Jumio, Yoti, Veriff) via a thin adapter layer that converts the API's verification response into the four binary witnesses.

WITNESS	SOURCE FIELD IN IDV API	CONDITION
identity_verified	check.result = "clear"	Overall check outcome is clear
over_age_threshold	report.date_of_birth	Today minus DOB exceeds threshold in years
not_sanctioned	watchlist.result = "clear"	Sanctions check returned no matches
residency_permitted	report.nationality or address.country	Country in permitted list for this service

The adapter layer processes the IDV API response in memory, extracts the four binary values, and immediately discards the full API response. The four binary witnesses are passed to the ZK circuit. At no point are PII values (name, date of birth, document number, facial image) written to persistent storage by the AffixIO platform. The IDV provider's own data retention and deletion procedures govern how long the PII is retained within their infrastructure.

SECTION 6

The PII-Free Record Schema

The KYC governance record stored in AffixIO's compliance record service has the following schema. It contains the proof digest (SHA-256 hash of the proof bytes), the circuit identifier (`kyc-v1`), the eligibility outcome (YES or NO), the Merkle leaf hash, the signed Merkle root, the timestamp, and a hashed session identifier. There is no name field, no date of birth field, no document number field, no address field, and no nationality field.

This is not achieved by configuration or access control. The record format has no fields for these values. A developer building a new integration who attempts to add a name field to the KYC record would need to modify the schema definition and the circuit adapter, both of which are under version control and subject to code review. The absence of PII fields is enforced by the schema structure, not by a policy that could be overridden by configuration.

> **Schema enforcement:** The PII-free property is a consequence of the record format, not a policy choice. No configuration change can cause PII to be stored in the KYC governance record without a schema change that would be visible in version control.

The hashed session identifier in the record is a SHA-256 hash of a session token that the user receives at onboarding. The session token is ephemeral and does not contain PII. The hash allows the organisation to link multiple governance records to the same onboarding session without storing any identifying information in the governance record.

SECTION 7

Merkle Anchoring and Audit Trail

KYC governance records are anchored in the same SHA-256 Merkle tree used for AI governance records. Each KYC eligibility proof is inserted as a leaf in the tree. The resulting Merkle root is published to the append-only roots log and signed with the ML-DSA-65 key. The complete audit trail for KYC operations is the set of signed roots from the KYC operations log, together with the individual proof digests.

This provides an important compliance property: the completeness of the KYC audit trail is provable without examining the content of any individual record. A regulator who wants to verify that all onboarding sessions generated a KYC eligibility proof can count the leaves in the Merkle tree and compare to the count of onboarding sessions. If the counts match, every session has a corresponding proof. If they do not match, the discrepancy is evidence of an anomaly in the KYC process. The count comparison does not require reading any record content, so no PII is exposed in the completeness verification step.

SECTION 8

GDPR Article 25 Analysis

GDPR Article 25 requires that data protection be implemented by design and by default. The ICO's guidance on Article 25 identifies three key requirements: technical and organisational measures must be implemented at design time (not retrofitted), those measures must effectively implement the data minimisation principle, and the default settings must implement maximum privacy without requiring action by the data subject.

The PII-free KYC schema satisfies all three requirements. Data minimisation is implemented at design time: the schema was designed without PII fields. The schema effectively implements data minimisation: no PII is stored in the KYC governance record as a structural consequence of the schema design. The default settings are maximally privacy-protective: there is no configuration that stores more data by default; additional data storage requires schema modification.

The Article 25 analysis for conventional KYC systems typically involves a data minimisation assessment that identifies which PII fields can be reduced or anonymised after verification. The PII-free ZK approach bypasses this assessment entirely: the system never processes PII in the governance layer, so there is nothing to minimise. The GDPR data protection impact assessment (DPIA) for this system is substantially simpler than for conventional KYC, because the KYC governance system is not a PII processor.

SECTION 9

AML and FATF Implications

Anti-money laundering regulation requires that organisations retain KYC records for five years after the end of a customer relationship. This retention requirement is one of the primary reasons KYC PII is retained so widely and for so long. Under conventional KYC, compliance with the AML retention requirement means retaining PII for five-plus years.

The ZK KYC approach changes the analysis. The governance record retained for five years is the eligibility proof, not the PII. The eligibility proof demonstrates that a KYC check was performed and passed. It does not reveal the specific PII that was checked. Whether this satisfies the AML retention requirement depends on the specific jurisdiction and the applicable national legislation implementing the FATF recommendations.

In the UK, the Money Laundering, Terrorist Financing and Transfer of Funds (Information on the Payer) Regulations 2017 require retention of "a copy of, or the references to, the evidence" obtained during due diligence. An eligibility proof that cryptographically demonstrates a successful verification was performed could satisfy the "references to" requirement. Organisations adopting ZK KYC should obtain legal advice on this question before deploying in a regulated financial services context. We think ZK KYC reduces but does not eliminate the need for conventional PII retention in AML contexts, and that the appropriate deployment is as a complement to reduced-retention conventional KYC rather than as a complete replacement.

SECTION 10

Regulatory Acceptance

ZK-based KYC is an emerging area with no established regulatory consensus. The European Banking Authority has published guidance acknowledging cryptographic identity verification methods, but has not specifically addressed ZK proof-based eligibility verification. The UK FCA's guidance on digital identity acknowledges that "novel methods of identity verification" may be considered, but requires that they satisfy the same standards as conventional KYC in terms of the reliability of the verification outcome.

The technical case for regulatory acceptance of ZK KYC is strong. The eligibility proof provides better evidence of a successful verification than a log record, because it is mathematically verifiable and tamper-resistant. The PII-free property reduces data breach risk substantially. The post-quantum signatures provide long-lived integrity for records that must be retained for

five years. Regulatory acceptance is expected to follow technical acceptance as ZK proof systems become more widely understood in financial regulatory contexts.

SECTION 11

Known Limitations

The ZK KYC approach has several limitations that should be considered before adoption. The binary witness architecture means that if a witness is incorrectly computed by the adapter (for example, if a sanctions list check returns an incorrect result due to a system error), the ZK proof certifies an incorrect computation. The proof proves the computation was performed; it does not prove the computation used correct data. The identity verification service's quality is the upstream dependency on which the correctness of the witnesses depends.

The approach relies on a trusted identity verification service to process the PII. This does not eliminate PII processing; it moves it to a third party who operates under their own data protection obligations. Organisations must assess the IDV provider's compliance posture and data retention practices as part of their overall DPIA, even though AffixIO's own systems do not process PII.

The KYC circuit currently supports four binary conditions. More complex eligibility rules (income thresholds, multi-factor risk scoring, dynamic sanctions list lookups) require additional circuit development. Each new condition requires a new circuit version with its own proving and verification keys. The circuit library is expanding, but complex eligibility rules require bespoke development that is not available off the shelf.

SECTION 12

Conclusion

PII-free KYC is achievable today using zero-knowledge proof technology. The architectural key is the separation between identity verification (which requires PII and is performed by a specialist IDV service) and eligibility verification (which requires only binary outputs of identity verification and can be performed by a ZK circuit). The KYC governance record contains only the eligibility proof, anchored in a Merkle tree and signed with post-quantum keys. No PII enters the governance layer. Data minimisation is enforced by the schema structure, not by policy or configuration.

The regulatory advantages are concrete: GDPR Article 25 compliance is achieved by design, data breach exposure in the compliance system is eliminated, and the long-lived integrity of KYC records is guaranteed by ML-DSA-65 signatures that resist future quantum attacks. The regulatory landscape for ZK-based KYC is evolving, and organisations deploying this approach should maintain dialogue with their regulatory supervisors. AffixIO's ZK KYC circuit is available as part of the open ZK circuit library.

Related reading

- WP-017: ZK Selective Disclosure for eIDAS 2.0 and the EUDI Wallet
- WP-008: Zero-Knowledge Proofs as GDPR Article 25 Infrastructure
- WP-014: Double-Spend Prevention for Zero-Knowledge Proofs

Frequently asked questions

What does PII-free KYC mean?

Proving a customer meets policy without persisting names, document numbers, or images in your systems.

Does this satisfy FATF requirements?

Travel rule and AML checks can consume eligibility proofs from regulated issuers rather than raw identity payloads.

Who holds the source identity?

Trusted issuers or wallets; verifiers only see proof digests and outcomes.
