Privacy-by-Design Enterprise Computing Architecture for Secure Data Integration and Saudi PDPL Compliance.

Abstract

Saudi enterprises integrate customer, employee, patient and partner records across enterprise applications, data warehouses, analytics platforms and external partners, and every integration flow now has to be justified under the Kingdom's Personal Data Protection Law (PDPL). Published evidence suggests that compliance is still immature: an inspection of 723 popular Saudi websites shortly before enforcement found widespread gaps in cookie notices and privacy policies, and organisations face difficulty translating legal provisions into technical controls. This paper designs a privacy-by-design enterprise computing architecture for secure data integration and PDPL compliance and evaluates it analytically. A design-science method was used, with requirements derived from an integrative review of 30 peer-reviewed journal articles from the years 2020 to 2025. Two anchor studies framed the review: a legal analysis of PDPL objectives and their balance with insurers' use of personal data, and the inspection of Saudi website privacy practices. The resulting architecture links a machine-readable regulatory knowledge layer, a consent and purpose registry, a privacy-aware integration layer, protected storage and analytics, data subject rights services, and an accountability layer for monitoring, audit and breach response, with a controlled gateway for transfers outside the Kingdom. The evaluation maps PDPL themes reported in the literature to architectural components and rates the published evidence for nine criteria. No new experiments were conducted. Evidence is strongest for machine-readable compliance checking, consent-driven data provisioning and data minimisation, mixed for privacy-enhancing technologies and performance overhead, and limited for Saudi-specific validation and for the treatment of cross-border transfers, on which published characterisations of the law differ. The paper closes with an implementation sequence and a research agenda.

UNDER PEER REVIEW IN IJAR

1. Introduction

Enterprise data integration has always been a technical problem of schemas, interfaces and data quality. In Saudi Arabia it has also become a legal one. The Personal Data Protection Law was enacted in 2021 and has quickly drawn academic and practitioner attention (Alkhedhairy, 2025). Commentators describe it as a major step towards aligning the Kingdom's regulatory environment with international data protection norms, drawing on the European General Data Protection Regulation (GDPR) while retaining provisions that reflect national interests (Suliman, 2025). For enterprises, this means that the routine work of moving personal data between operational systems, warehouses, analytics services and partners must now be designed, documented and defended against legal obligations.

The available evidence indicates that this shift is not yet reflected in practice. When 723 popular websites in Saudi Arabia were examined three months before the law's official enforcement, roughly 85% of the sites that used cookies showed no cookie banner or policy, and 39% did not provide the privacy policy that the PDPL requires (Alhazmi & Daghistani, 2025). In a survey of 357 clinicians in Saudi public and private healthcare, only 7% reported high familiarity with the legal implications of artificial intelligence and confidence that AI tools complied with data laws was low (Alanazi, 2025). Organisations that process personal data also find it hard to translate legal provisions into a technological context, and non-compliance exposes them to financial, legal and reputational risk (Alkhamsi & Alqahtani, 2024).

Privacy by design offers a response: privacy protections should be built into the architecture of systems rather than added after deployment. The idea is well established in principle, yet the engineering literature shows that developers often prioritise functional over privacy requirements and lack tools and institutional support (Alhazmi & Arachchilage, 2021), and that adoption of privacy-enhancing technologies (PETs) remains low in practice (Klymenko et al., 2025). Most technical work on compliance has also been framed around the GDPR, so Saudi enterprises have few designs that connect privacy engineering with their own regulatory setting.

This paper addresses that gap through a design-science study. It derives requirements from what peer-reviewed sources report about the PDPL and its context, combines them with evidence from privacy and compliance engineering, proposes a privacy-by-design enterprise computing architecture for secure data integration, and evaluates the architecture analytically against

UNDER PEER REVIEW IN IJAR

explicit criteria. Two limits are stated at the outset. First, the PDPL is described only as far as the cited sources describe it; the paper does not interpret the statute directly and is not legal advice. Second, the evaluation uses published evidence rather than new experiments or benchmarks.

The next two sections state what the study sets out to do and how it was done. Section 4 reviews what the literature reports about the PDPL, Section 5 reviews privacy engineering mechanisms, Section 6 presents the architecture and Section 7 evaluates it. Section 8 discusses implications and limitations, and Section 9 concludes.

2. Aim, Objectives and Guiding Questions

The aim is to design a privacy-by-design enterprise computing architecture that supports secure integration of personal data across enterprise systems in Saudi Arabia and helps organisations demonstrate compliance with the PDPL, and to evaluate that architecture transparently against criteria and published evidence. The objectives are: (1) to identify, from peer-reviewed sources, the PDPL themes and contextual conditions that bear on enterprise data integration; (2) to identify engineering mechanisms that translate data protection obligations into system behaviour; (3) to design an architecture that embeds these mechanisms in the integration life cycle; and (4) to evaluate the architecture against explicit criteria, recording where evidence is strong, mixed or missing.

Three research questions guide the study. RQ1: Which PDPL-related obligations and contextual conditions, as reported in the peer-reviewed literature, should shape the design of enterprise data integration in Saudi Arabia? RQ2: Which architectural components and mechanisms can embed these obligations by design across the integration life cycle? RQ3: How well is the proposed architecture supported by published evidence, and which gaps and trade-offs remain?

3. Method: Design Science Grounded in an Integrative Review

Design science develops artefacts, in this case an architecture, that address a defined problem and evaluates them against stated criteria. The method suits the topic because the knowledge needed to design for compliance is distributed across law, requirements engineering, information security and data management, and because the expected output is a design that practitioners can adapt. The study followed four activities: problem explication and requirements derivation, design, analytical evaluation, and communication of the artefact together with its limitations.

UNDER PEER REVIEW IN IJAP

Requirements were derived from an integrative review. A scholarly metadata index was searched in November 2026 with combinations of legal terms ("Personal Data Protection Law", "PDPL", "Saudi Arabia", "GDPR", "data protection"), design terms ("privacy by design", "data protection by design", "privacy requirements", "compliance checking") and data management terms ("data integration", "data sharing", "record linkage", "data warehouse", "consent management"). Publisher-level searches and searches of an open bibliographic index were added because the PDPL literature is small. A source was retained if it was a peer-reviewed journal article in English, its final publication year was between 2020 and 2025, and it either discussed the PDPL or its regional context or reported a design, mechanism or empirical finding on privacy and compliance engineering. Items without verifiable metadata, conference papers and opinion pieces were excluded.

Two anchor studies shaped the search. The first is a critical legal analysis of PDPL objectives, its comparison with other regimes such as the GDPR and its balance between data subjects' interests and the sharing of personal data with insurers (Alkhedhairy, 2025). The second is an inspection framework applied to popular Saudi websites that reports their alignment with PDPL requirements on cookies, consent and privacy policies (Alhazmi & Daghistani, 2025). The final corpus contains 30 articles: eight on the PDPL and the Saudi or Gulf context, eight on compliance modelling and automated checking, four on privacy requirements and developer practice, and ten on privacy-enhancing technologies and secure data integration.

Each statement in Sections 4 and 5 was checked against the source it cites. Statements about the PDPL are confined to what the cited articles report, without article numbers or dates beyond those given in the sources. For the evaluation, every criterion was rated against the corpus as supportive (several studies report a working mechanism with evidence of effect), mixed (mechanisms exist but with material trade-offs or narrow validation), or limited (little or no published evidence). No performance tests, user studies or deployments were carried out for this paper.

4. The PDPL and Its Context as Reported in the Literature

Legal commentators agree on the basic orientation of the law. The PDPL was enacted in 2021, and a critical analysis judged its measures to appear sound while cautioning that three years of operation are too short for a definitive assessment of whether it balances data subjects' interests

UNDER PEER REVIEW IN IJAR

against the risks of sharing personal data with insurers (Alkhedhairy, 2025). Another evaluation describes the law as inspired by the GDPR but distinctive in certain respects, pointing in particular to strict requirements for keeping data in the Kingdom and to stronger regulatory control, and asks whether its two enforcement bodies, the National Data Management Office (NDMO) and SDAIA, the Saudi authority for data and artificial intelligence, have the powers and resources to secure compliance (Suliman, 2025). The same analysis asks whether the law builds trust in the digital economy or creates impediments that could discourage international cooperation.

The literature is less unanimous on international transfers. Alongside the characterisation of localisation requirements as stringent (Suliman, 2025), a comparative study of health data law in Saudi Arabia, Oman and Qatar reports that, unlike states with strict localisation rules, the three countries adopted free international transfer as the starting principle, subject to legal conditions and exceptions (El-Khoury & Albarashdi, 2025). A wider study of Gulf Cooperation Council laws discusses transfers based on adequacy decisions or suitable safeguards and concludes that the region's legislation still has room to develop (Alhababi, 2024). For architecture, the prudent inference is not to choose between these readings but to treat every outbound flow as a controlled, documented and reviewable event.

Sector context adds further obligations. Healthcare data receive heightened protection in the surveyed Gulf laws, with additional requirements for lawful processing, and strong compliance burdens may encourage investment in data-sharing methods that pursue anonymity (El-Khoury & Albarashdi, 2025). The use of personal data by Saudi insurers has been examined specifically, with a tentative conclusion that the law's treatment of such data is sound (Alkhedhairy, 2025).

Empirical studies describe a compliance gap. Besides the website findings already cited, which point to unmet rights such as being informed about data collection (Alhazmi & Daghistani, 2025), clinicians identified gaps in regulatory alignment with national laws including the PDPL as one of six legal barriers to AI adoption, and legal training was significantly associated with confidence in AI compliance (Alanazi, 2025). Among Saudi mobile users, privacy concerns and trust were positively associated with perceived usability of privacy and security settings, leading to recommendations for user-centric controls, transparent data handling policies and periodic previews of data management status (Saeed, 2024).

UNDER PEER REVIEW IN IJAR

One design response has already been proposed locally. A compliance framework translates PDPL legal terms into measurable data management standards covering all integral parts of data management, and a case study using two Arabic language models to check privacy policy adherence on Saudi websites achieved micro-averaged F1-scores of 93.32% and 92.46% (Alkhamsi & Alqahtani, 2024). This shows that PDPL requirements can be operationalised as checkable standards, although the case covers policy text rather than integration flows.

In answer to RQ1, the literature indicates five themes that should shape integration design: transparency and notice towards data subjects; lawful processing, including consent where it is required; heightened safeguards for sensitive categories such as health data; controlled and documented transfers outside the Kingdom, given differing characterisations of the transfer rules; and demonstrable accountability towards SDAIA and the NDMO as enforcement bodies. A sixth, contextual requirement is capability: legal awareness and training remain weak in at least some sectors. Table 1 summarises the sources behind these themes.

Table 1. PDPL-related themes reported in the peer-reviewed literature and their design implications

Source	Focus	Reported finding	Design implication
Alkhedhairy (2025)	PDPL objectives; insurers	Law enacted in 2021; measures appear sound, but too early for a definitive assessment	Treat rules as revisable; support sector-specific purposes
Suliman (2025)	PDPL effectiveness	GDPR-inspired; stringent localisation and regulatory control; SDAIA and NDMO enforcement	Document flows for regulators; control data location
El-Khoury and Albarashdi (2025)	Health data in three Gulf states	Heightened protection of health data; free transfer principle with conditions and exceptions	Extra safeguards for sensitive data; reviewable transfer rules
Alhababi (2024)	Gulf transfer rules vs GDPR	Transfers based on adequacy or suitable safeguards; room for legislative development	Transfer gateway with recorded basis and safeguards
Alhazmi and Daghistani (2025)	723 Saudi websites	About 85% of cookie-using sites lacked banners or policies; 39% lacked privacy policies	Consistent notice and consent services
Alkhamsi and Alqahtani (2024)	PDPL compliance framework	Legal terms translated into measurable data management standards	Machine-readable PDPL rule catalogue
Alanazi (2025)	357 Saudi clinicians	Low legal familiarity; gaps in alignment with PDPL among barriers	Training and governance alongside technology

UNDER PEER REVIEW IN IJAR

Source	Focus	Reported finding	Design implication
Saeed (2024)	Saudi mobile users	Concerns and trust linked to perceived privacy usability	User-centric controls and data management previews

5. Privacy Engineering Mechanisms for Integration

The engineering literature offers mechanisms for turning such themes into system behaviour, although nearly all were developed for the GDPR. The first family makes regulation machine-readable. A complete conceptual model of the GDPR in UML was produced with traceability to the regulation, 35 compliance rules encoded in OCL and 20 variation points for specialising the generic model to national law (Torre et al., 2021). This two-tier idea, generic and specialised, transfers naturally to a PDPL specialisation. A policy language developed in a European research project expresses consent, business policies and regulatory obligations, and supports automated checks that processing complies with given consent and that business processes comply with obligations (Bonatti et al., 2020). A data protection by design tool applied a regulation-to-code process that translates regulatory provisions into technical and organisational measures and then into software code, using a knowledge graph of informed consent, and was demonstrated in insurance and smart cities (Chhetri et al., 2022).

The second family governs data provisioning. Consent ontologies can annotate database schemas so that declarative mappings generate datasets that comply with given consent at the moment they are needed, while provenance is captured for each step (Debruyne et al., 2020). Where consent may not be enough, as in some business-to-business and business-to-consumer sharing, contracts can supply the legal basis, and a knowledge graph-based tool binds legal bases to data sharing contracts and verifies contract compliance with minimal overhead in smart city and insurance scenarios (Tauqeer et al., 2022). Dynamic consent, in which data subjects change their permissions over time, requires precisely defined security and privacy properties and can draw on differential privacy, blockchain, zero-knowledge proofs and cryptographic procedures (Khalid et al., 2023).

The third family minimises and protects data within processes and pipelines. Intra-process data degradation supports continuous minimisation during business processes, process mining can check whether process execution respects data subject rights, and differential privacy can enable compliant process discovery (Zaman & Hassani, 2020). Business process models combined with

UNDER PEER REVIEW IN IJAR

simulation allow controllers to estimate the human resources, time and cost needed for compliance activities such as breach notification (Varela-Vaca et al., 2025). For analytical uses, input features of machine learning models can be removed or generalised using knowledge distillation, sometimes without any loss of accuracy (Goldsteen et al., 2022). Relaxed functional dependencies can reveal attributes from which sensitive values could be inferred, increasing dataset confidentiality while masking fewer values (Caruccio et al., 2020).

The fourth family secures integration across organisational and system boundaries. A conceptual framework for business-to-business data sharing combines encryption, access control, anonymisation and audit trails, and considers homomorphic encryption and blockchain (Li et al., 2024). For small firms that analyse data jointly, de-identification, data synthesis and federated learning were found to comply with the regulations reviewed, which require reasonable rather than absolute privacy efforts, although a highly knowledgeable adversary must still be considered where stakes are high (Li et al., 2021). Privacy-preserving record linkage frameworks emphasise that building protections into the architecture, through role separation and controlled data flows, matters as much as the cryptographic technique (Nie et al., 2025). Healthcare data warehouse research proposes security and privacy measures integrated by design at every stage from data input to presentation (Thantilage et al., 2023), and legal scholarship shows how anonymisation and pseudonymisation can reconcile protection with the free flow of health data (He, 2023). Two architecture-level metrics, the privacy risk expansion factor and the privacy exposure index, compare how design choices spread sensitive data across an ecosystem (Chereja et al., 2025). In e-commerce, a framework combining legal, technological and procedural measures has been proposed to strengthen data security and consumer trust, alongside models designed to minimise personal data sharing (Morić et al., 2024). Decentralised data stores such as Solid shift control to individuals but still require actors, roles and processes to be aligned with regulatory concepts (Pandit, 2023).

The final family concerns the people who build systems. A systematic review of privacy requirements elicitation catalogued 78 studies, of which 38 had not verified their methods in a practical case and only 26 had been tried in industry, and a survey of 198 practitioners found the Privacy Safeguard method best known (Canedo et al., 2023). A mapping study of 90 requirements engineering papers on the GDPR found that consent and overall compliance

UNDER PEER REVIEW IN IJAR

dominated while topics such as cookies and children's data received little attention (Negri-Ribalta et al., 2024). A process diagram for integrating PETs into software development, evaluated with five experts and 100 developers, highlighted the need for practitioner education (Klymenko et al., 2025).

In answer to the first part of RQ2, five mechanism families are available: machine-readable regulation and policy checking; consent- and contract-driven data provisioning; minimisation and inference control; secure interorganisational integration with PETs; and engineering process support. Their common weakness for the present purpose is that none has been validated against the PDPL in a published integration setting.

6. The Proposed Privacy-by-Design Architecture

Figure 1 shows the architecture. It places personal data flows at the centre and surrounds them with six components and a transfer gateway. The design follows three principles. Obligations are represented once, as machine-readable rules, and reused by every component. Every flow is purpose-bound: data move only when a registered purpose and lawful basis allow it. Evidence of compliance is produced as a by-product of operation rather than reconstructed for audits.

UNDER PEER REVIEW IN IJAP

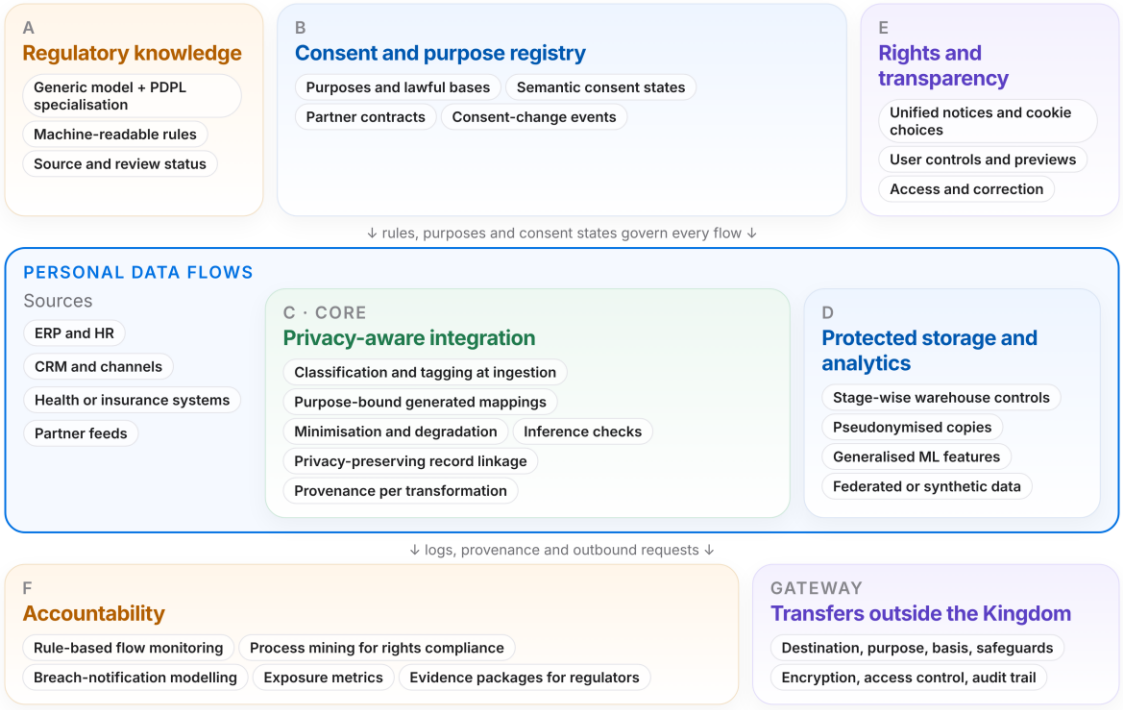


Figure 1. Privacy-by-design enterprise computing architecture for secure data integration and PDPL compliance.

Component A is a regulatory knowledge layer. Following the two-tier approach of Torre et al. (2021), it holds a generic data protection model and a PDPL specialisation maintained jointly by legal and engineering staff. Obligations are expressed in a policy language of the kind described by Bonatti et al. (2020) and refined through a regulation-to-code process (Chhetri et al., 2022). The compliance standards proposed for the PDPL by Alkhamsi and Alqahtani (2024) provide a starting catalogue. Because the published characterisations of some provisions, notably transfers, differ, each rule records its source and review status.

Component B is a consent and purpose registry. It records purposes, lawful bases, consent states and contractual bases for every category of personal data. Consent is represented semantically so that downstream mappings can query it (Debruyne et al., 2020), contracts with partners are linked to legal bases (Tauqeer et al., 2022), and changes in consent propagate as events, as dynamic consent requires (Khalid et al., 2023).

UNDER PEER REVIEW

Component C is a privacy-aware integration layer and the core of the design. At ingestion, it classifies personal and sensitive attributes and tags each record with its purpose and origin. Mappings between source systems and targets are generated from schema annotations, so that only attributes permitted for a purpose are delivered and provenance is written for each transformation (Debruyne et al., 2020). Minimisation and degradation rules remove or coarsen attributes as processes progress (Zaman & Hassani, 2020), and inference checks identify quasi-identifiers that could reveal sensitive values before data are released (Caruccio et al., 2020). Where records from different sources must be joined, privacy-preserving record linkage with separated roles is used instead of exchanging identifiers in clear text (Nie et al., 2025).

Component D provides protected storage and analytics. Warehouses and lakes apply security and privacy controls at each stage from input to presentation (Thantilage et al., 2023), with pseudonymisation as the default for analytical copies (He, 2023). Analytical and machine learning services receive generalised features where accuracy permits (Goldsteen et al., 2022). For joint analysis with partners, federated learning, synthetic data or de-identified extracts are preferred to raw transfers (Li et al., 2021).

Component E delivers data subject rights and transparency. It publishes privacy notices and cookie choices consistently across channels, addressing the gaps found on Saudi websites (Alhazmi & Daghistani, 2025), and offers user-centric controls with previews of how personal data are being managed (Saeed, 2024). Requests for access or correction are served from the provenance and registry records rather than from manual searches.

Component F is an accountability layer. It monitors flows against the rules in Component A, uses process mining to check that process executions respect data subject rights (Zaman & Hassani, 2020), models breach notification processes so that resources can be planned (Varela-Vaca et al., 2025), and produces evidence packages for regulators. Architecture-level exposure metrics are recalculated when new integrations are added (Chereja et al., 2025). Alongside these components, a transfer gateway mediates every flow to a destination outside the Kingdom, recording destination, purpose, basis and safeguards, so that the organisation can adapt as interpretation and guidance on transfers develop (Alhababi, 2024; El-Khoury & Albarashdi, 2025). Partner exchanges through the gateway combine encryption, access control, anonymisation and audit trails (Li et al., 2024).

UNDER PEER REVIEW IN IJAR

In answer to RQ2, the architecture embeds PDPL-related obligations through a shared regulatory knowledge layer, a consent and purpose registry, privacy-aware integration, protected storage and analytics, rights and transparency services, an accountability layer and a controlled transfer gateway. Figure 2 traces the five legal themes and the capability requirement from Section 4 to these components.

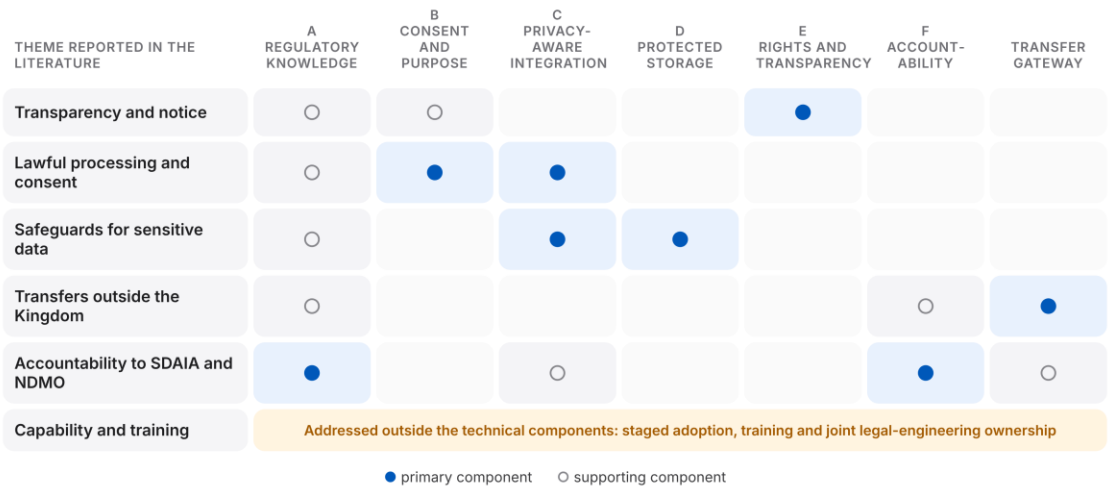


Figure 2. Traceability of PDPL-related themes reported in the literature to architectural components.

7. Analytical Evaluation

The architecture was evaluated against nine criteria drawn from the requirements in Sections 4 and 5: traceability of obligations, consent and purpose enforcement, minimisation and inference control, confidentiality in cross-boundary integration, data subject rights and transparency, accountability and auditability, control of transfers outside the Kingdom, operational overhead and usability, and organisational feasibility. Table 2 records the design response, the evidence rating and the residual risk for each. The question asked is whether published studies support the mechanisms the architecture relies on, not whether the architecture as a whole has been shown to work.

Table 2. Analytical evaluation of the architecture against nine criteria

Criterion	Design response	Evidence rating	Residual risk

UNDER PEER R

Criterion	Design response	Evidence rating	Residual risk
Traceability of obligations	Generic model with PDPL specialisation; policy rules	Supportive	PDPL rules not yet validated
Consent and purpose enforcement	Registry; consent-driven mappings; contract bases	Supportive	Demonstrated on synthetic data
Minimisation and inference control	Degradation; feature generalisation; dependency checks	Supportive with caveats	Task-specific validation
Cross-boundary confidentiality	Privacy-preserving linkage; federated or synthetic data; encrypted exchange	Mixed	Strong adversaries; conceptual designs
Rights and transparency	Unified notices; user controls; provenance-based requests	Mixed	No enterprise-scale evaluations
Accountability and auditability	Monitoring; process mining; breach simulation; exposure metrics	Supportive in parts	Components untested in combination
Transfers outside the Kingdom	Gateway recording destination, basis and safeguards	Limited	Differing readings of the law
Overhead and usability	Selective use of PETs; low-overhead checks	Mixed	Costs mostly unreported
Organisational feasibility	Staged adoption; training; joint legal-engineering ownership	Limited	Skills and awareness gaps

Evidence is supportive for traceability of obligations. Complete conceptual models with traceable rules (Torre et al., 2021), machine-understandable policies with automated checking (Bonatti et al., 2020), a regulation-to-code tool demonstrated in two domains (Chhetri et al., 2022) and a PDPL-specific compliance framework (Alkhamsi & Alqahtani, 2024) show that obligations can be represented and checked. Evidence is also supportive for consent and purpose enforcement, since just-in-time dataset generation (Debruyne et al., 2020) and contract compliance verification with low overhead (Tauqeer et al., 2022) have been implemented and demonstrated, albeit on synthetic data or scenarios.

For minimisation and inference control, the evidence is supportive with caveats. Feature generalisation preserved model accuracy in some cases (Goldsteen et al., 2022), dependency-based analysis reduced the number of values that had to be masked (Caruccio et al., 2020), and intra-process degradation is supported by tools (Zaman & Hassani, 2020). Each mechanism, however, was validated on particular data and tasks.

Confidentiality in cross-boundary integration receives mixed support. Frameworks for business-to-business data sharing (Li et al., 2024) and record linkage (Nie et al., 2025) are largely

UNDER PEER REVIEW IN IJAP

conceptual or technology-agnostic. Privacy-preserving joint analysis is judged compliant with the regulations reviewed, but a highly knowledgeable adversary remains a concern where the stakes are high (Li et al., 2021). Data subject rights and transparency are also rated mixed: the need is well documented in Saudi Arabia (Alhazmi & Daghistani, 2025; Saeed, 2024), but published evaluations of rights-handling services in integrated enterprise environments were not found.

Accountability and auditability are rated supportive in mechanism but untested at enterprise scale. Process mining for rights compliance (Zaman & Hassani, 2020), simulation of breach notification processes (Varela-Vaca et al., 2025) and exposure metrics for architectures (Chereja et al., 2025) all exist, yet no study combined them in one accountability layer. Control of transfers outside the Kingdom is rated limited: the architecture provides a gateway and documentation, but published sources characterise the PDPL's transfer rules differently (El-Khoury & Albarashdi, 2025; Suliman, 2025), and the gateway's rules cannot be validated until interpretation is settled.

Operational overhead and usability are mixed. The contract verification tool reported minimal overhead (Tauqeer et al., 2022), but the reviewed studies of PETs and dynamic consent do not report operational costs in enterprise settings (Khalid et al., 2023; Li et al., 2024), and adoption of PETs remains low (Klymenko et al., 2025). Organisational feasibility is limited. Developers report unfamiliarity with privacy principles and missing support (Alhazmi & Arachchilage, 2021), most privacy requirements methods lack industrial validation (Canedo et al., 2023), research attention is uneven across topics (Negri-Ribalta et al., 2024), and Saudi clinicians show low legal readiness (Alanazi, 2025). These findings suggest that the architecture's success depends at least as much on skills and governance as on technology.

In answer to RQ3, the architecture rests on well-supported mechanisms for representing obligations, enforcing consent and purpose and minimising data, on mixed evidence for secure cross-boundary integration, transparency services and overhead, and on limited evidence for transfer control and organisational feasibility. The central trade-off is between the assurance gained from automated, fine-grained enforcement and the cost, complexity and skills it requires.

UNDER PEER REVIEW IN IJAR

8. Discussion, Implications and Limitations

The study's main contribution is to move PDPL compliance from the level of policy documents to the level of integration architecture. Most Saudi evidence describes the law or measures compliance gaps on public-facing websites. Most engineering evidence addresses the GDPR. The architecture connects the two by treating the PDPL as a specialisation of a generic data protection model and by binding every integration flow to a registered purpose, a lawful basis and a provenance trail. This matters because integration flows, rather than websites, are where much personal data is copied, joined and exposed inside enterprises.

For practitioners, the evaluation suggests an implementation sequence. Organisations should first build an inventory of personal data flows and a consent and purpose registry, because every other component depends on them. They should then encode a small set of high-value rules, for example on notice, purpose limitation and transfers, and apply privacy-aware mappings and minimisation to the integrations that carry sensitive data such as health or insurance records. Accountability services and exposure metrics follow, and PETs such as federated learning or privacy-preserving linkage are added where joint analysis with partners justifies their cost. Training should run in parallel, since weak legal readiness and developer unfamiliarity are among the most consistent findings in the corpus.

For regulators and researchers, the review highlights where knowledge is thin. Published commentary differs on how the PDPL treats transfers outside the Kingdom, and enforcement capacity is still being assessed (Suliman, 2025). Peer-reviewed Saudi studies of PDPL compliance inside enterprise systems were not found. Useful next steps would be case studies that implement a PDPL rule set on real integration flows, measurement of the overhead of privacy-aware integration in Saudi organisations, and comparative work on how sector regulations for health and insurance interact with the general law.

Several limitations apply. The corpus is small because the peer-reviewed PDPL literature is small, and some PDPL sources are practitioner-oriented legal journals rather than computing venues. Descriptions of the law come from secondary sources, so the paper cannot confirm specific provisions and should not be read as legal guidance. Most engineering mechanisms were developed for the GDPR, and their transfer to the PDPL is argued rather than demonstrated. Several claims were checked only against abstracts when full texts were unavailable, and the

UNDER PEER REVIEW IN IJAR

review was conducted by a single researcher, which increases the risk of selection bias. Finally, the evaluation is analytical; the architecture has not been implemented or tested.

9. Conclusion

The study set out to show how obligations associated with the PDPL can be built into the way Saudi enterprises integrate personal data. Drawing on 30 peer-reviewed studies, it identified five PDPL-related themes and a capability requirement, five families of privacy engineering mechanisms, and an architecture of six components and a transfer gateway that embeds obligations into every integration flow. The evaluation found strong support for machine-readable obligations, consent-driven provisioning and minimisation, mixed support for cross-boundary confidentiality, transparency services and overhead, and limited evidence for transfer control and organisational feasibility in Saudi settings. The architecture is offered as a reasoned design and a basis for empirical validation, not as a proven solution. Its practical message is that PDPL compliance becomes demonstrable when purposes, legal bases and provenance travel with the data through the enterprise.

UNDER PEER REVIEW IN IJAR

References

1. Alanazi, A. (2025). Assessing clinicians' legal concerns and the need for a regulatory framework for AI in healthcare: A mixed-methods study. *Healthcare*, *13*(13), Article 1487. https://doi.org/10.3390/healthcare13131487

2. Alhababi, H. H. (2024). Cross-border data transfer between the GCC data protection laws and the GDPR. *Global Journal of Comparative Law*, *13*(2), 178–200. https://doi.org/10.1163/2211906x-13020003

3. Alhazmi, A., & Arachchilage, N. A. G. (2021). I'm all ears! Listening to software developers on putting GDPR principles into software development practice. *Personal and Ubiquitous Computing*, *25*(5), 879–892. https://doi.org/10.1007/s00779-021-01544-1

4. Alhazmi, A., & Daghistani, A. (2025). Privacy practices of popular websites in Saudi Arabia. *Journal of Umm Al-Qura University for Engineering and Architecture*, *16*(1), 19–29. https://doi.org/10.1007/s43995-024-00085-x

5. Alkhamsi, N. N., & Alqahtani, S. S. (2024). Compliance framework for personal data protection law standards. *International Journal of Advanced Computer Science and Applications*, *15*(7). https://doi.org/10.14569/ijacsa.2024.0150751

6. Alkhedhairy, M. A. (2025). Balancing privacy and risk: A critical analysis of personal data use as governed by Saudi insurance law. *Laws*, *14*(4), Article 47. https://doi.org/10.3390/laws14040047

7. Bonatti, P. A., Kirrane, S., Petrova, I. M., & Sauro, L. (2020). Machine understandable policies and GDPR compliance checking. *KI - Künstliche Intelligenz*, *34*(3), 303–315. https://doi.org/10.1007/s13218-020-00677-4

8. Canedo, E. D., Bandeira, I. N., Calazans, A. T. S., Costa, P. H. T., Cançado, E. C. R., & Bonifácio, R. (2023). Privacy requirements elicitation: A systematic literature review and perception analysis of IT practitioners. *Requirements Engineering*, *28*(2), 177–194. https://doi.org/10.1007/s00766-022-00382-8

9. Caruccio, L., Desiato, D., Polese, G., & Tortora, G. (2020). GDPR compliant information confidentiality preservation in big data processing. *IEEE Access*, *8*, 205034–205050. https://doi.org/10.1109/access.2020.3036916

UNDER PEER REVIEW IN IJAR

10. Chereja, I., Erdei, R., Delinschi, D., Pasca, E., Avram, A., & Matei, O. (2025). Privacy-conducive data ecosystem architecture: By-design vulnerability assessment using privacy risk expansion factor and privacy exposure index. *Sensors*, *25*(11), Article 3554. https://doi.org/10.3390/s25113554

11. Chhetri, T. R., Kurteva, A., DeLong, R. J., Hilscher, R., Korte, K., & Fensel, A. (2022). Data protection by design tool for automated GDPR compliance verification based on semantically modeled informed consent. *Sensors*, *22*(7), Article 2763. https://doi.org/10.3390/s22072763

12. Debruyne, C., Pandit, H. J., Lewis, D., & O'Sullivan, D. (2020). "Just-in-time" generation of datasets by considering structured representations of given consent for GDPR compliance. *Knowledge and Information Systems*, *62*(9), 3615–3640. https://doi.org/10.1007/s10115-020-01468-x

13. El-Khoury, M., & Albarashdi, S. (2025). Navigating the privacy landscape of healthcare-driven AI in the Middle East: Case studies from Oman, Qatar, and Saudi Arabia. *Social Sciences & Humanities Open*, *11*, Article 101492. https://doi.org/10.1016/j.ssaho.2025.101492

14. Goldsteen, A., Ezov, G., Shmelkin, R., Moffie, M., & Farkash, A. (2022). Data minimization for GDPR compliance in machine learning models. *AI and Ethics*, *2*(3), 477–491. https://doi.org/10.1007/s43681-021-00095-8

15. He, Z. (2023). From privacy-enhancing to health data utilisation: The traces of anonymisation and pseudonymisation in EU data protection law. *Digital Society*, *2*(2), Article 17. https://doi.org/10.1007/s44206-023-00043-5

16. Khalid, M. I., Ahmed, M., & Kim, J. (2023). Enhancing data protection in dynamic consent management systems: Formalizing privacy and security definitions with differential privacy, decentralization, and zero-knowledge proofs. *Sensors*, *23*(17), Article 7604. https://doi.org/10.3390/s23177604

17. Klymenko, A., Meisenbacher, S., Favaro, L., & Matthes, F. (2025). Supporting the integration of privacy-enhancing technologies into the software development life cycle. *SN Computer Science*, *6*(6), Article 592. https://doi.org/10.1007/s42979-025-04127-6

UNDER PEER REVIEW IN IJAR

18. Li, S.-C., Chen, Y.-W., & Huang, Y. (2021). Examining compliance with personal data protection regulations in interorganizational data analysis. *Sustainability*, *13*(20), Article 11459. https://doi.org/10.3390/su132011459

19. Li, W., Tse, W. K., & Chen, J. (2024). Privacy and security mechanisms for B2B data sharing: A conceptual framework. *Information*, *15*(6), Article 308. https://doi.org/10.3390/info15060308

20. Morić, Z., Dakic, V., Djekic, D., & Regvart, D. (2024). Protection of personal data in the context of e-commerce. *Journal of Cybersecurity and Privacy*, *4*(3), 731–761. https://doi.org/10.3390/jcp4030034

21. Negri-Ribalta, C., Lombard-Platet, M., & Salinesi, C. (2024). Understanding the GDPR from a requirements engineering perspective—a systematic mapping study on regulatory data protection requirements. *Requirements Engineering*, *29*(4), 523–549. https://doi.org/10.1007/s00766-024-00423-4

22. Nie, Z., Tyndall, B., Brannock, D., Gentles, E., Parish, E., & Banger, A. (2025). A framework for the design of privacy-preserving record linkage systems. *Journal of Cybersecurity and Privacy*, *5*(3), Article 44. https://doi.org/10.3390/jcp5030044

23. Pandit, H. J. (2023). Making sense of Solid for data governance and GDPR. *Information*, *14*(2), Article 114. https://doi.org/10.3390/info14020114

24. Saeed, S. (2024). Usable privacy and security in mobile applications: Perception of mobile end users in Saudi Arabia. *Big Data and Cognitive Computing*, *8*(11), Article 162. https://doi.org/10.3390/bdcc8110162

25. Suliman, H. O. H. (2025). Evaluating the effectiveness of Saudi Arabia's PDPL in the global digital economy. *Journal of Data Protection & Privacy*, *8*(1), 97. https://doi.org/10.69554/elux6976

26. Tauqeer, A., Kurteva, A., Chhetri, T. R., Ahmeti, A., & Fensel, A. (2022). Automated GDPR contract compliance verification using knowledge graphs. *Information*, *13*(10), Article 447. https://doi.org/10.3390/info13100447

UNDER PEER REVIEW IN IJAR

27. Thantilage, R. D., Le-Khac, N.-A., & Kechadi, M.-T. (2023). Healthcare data security and privacy in data warehouse architectures. *Informatics in Medicine Unlocked*, *39*, Article 101270. https://doi.org/10.1016/j.imu.2023.101270

28. Torre, D., Alferez, M., Soltana, G., Sabetzadeh, M., & Briand, L. (2021). Modeling data protection and privacy: Application and experience with GDPR. *Software and Systems Modeling*, *20*(6), 2071–2087. https://doi.org/10.1007/s10270-021-00935-5

29. Varela-Vaca, Á. J., Gómez-López, M. T., Morales Zamora, Y., & Gasca, R. M. (2025). Business process models and simulation to enable GDPR compliance. *International Journal of Information Security*, *24*(1), Article 41. https://doi.org/10.1007/s10207-024-00952-7

30. Zaman, R., & Hassani, M. (2020). On enabling GDPR compliance in business processes through data-driven solutions. *SN Computer Science*, *1*(4), Article 210. https://doi.org/10.1007/s42979-020-00215-x

UNDER PEER REVIEW IN IJAR