

Client-Side PII Masking & Governance

Architectural Defense Against Frontier AI Exfiltration, Regulatory Fines (\$50M OAIC), and Sovereign Cloud Liability

01. EXECUTIVE SUMMARY & THREAT VECTORS

SECTION 01 / 04

The Paradigm Failure of Traditional Cloud & Proxy Redaction

Modern enterprise adoption of Large Language Models (LLMs) creates an unprecedented data leakage vector. When employees upload commercial briefs, legal agreements, medical files, or loan files into external generative AI platforms, unredacted corporate records transit overseas data centers. Traditional cloud-based redaction proxies fail because documents must be transmitted to a third party before inspection. Redactorr enforces a **zero-trust, client-side isolation model**: 100% of pattern evaluation, optical recognition, and tokenisation executes inside local memory before payload egress.

Legacy Cloud Proxy Risk

- Document uploaded to intermediary cloud proxy.
- Plaintext PII/PHI stored in server RAM and temp disks.
- External vendor holds cryptographic keys.
- Exposure under US CLOUD Act subpoenas.
- High latency (2–5 seconds per multi-page document).

Redactorr Client-Side WASM Model

- Zero network payload sent to Redactorr servers.
- 15-Layer AEGIS regex/NER engine runs in browser/VPC.
- Enterprise retains 100% local key custody.
- Complete compliance with Australian Privacy Principle 8.
- Sub-100ms local execution speed.

02. STATUTORY MANDATES & MONETARY EXPOSURE

AUSTRALIAN PRIVACY FRAMEWORK

STATUTORY MANDATE	REGULATORY REQUIREMENT	NON-COMPLIANCE RISK	REDACTORR PROTECTIVE ARCHITECTURE
Privacy Act 1988 (Cth) / APP 11	Mandates entities take active, reasonable steps to protect personal info from misuse, interference, and loss.	Statutory fines up to \$50,000,000 or 30% of adjusted turnover.	PROTECTED Pre-flight redaction ensures unredacted records never enter untrusted AI memory.
APP 8 (Cross-Border Data Disclosure)	Accountability for personal data transferred outside Australia to overseas cloud LLM endpoints.	Direct statutory liability for overseas cloud breaches or training scrapers.	COMPLIANT Only tokenised labels (e.g. [PERSON_1], [BSB_1]) cross borders.
Notifiable Data Breaches (NDB)	Mandatory OAIC and public notification within 30 days of serious data compromise.	Class action litigation, executive reputation destruction, shareholder loss.	ZERO EGRESS Zero document persistence removes the statutory condition for breach notification.

03. AUSTRALIAN BREACH FORENSICS & MITIGATION STUDIES

EMPIRICAL INCIDENT ANALYSIS

Empirical analysis of major Australian cyber catastrophes demonstrates how presentation-layer redaction and zero-egress sanitisation prevent catastrophic exfiltration events.

INCIDENT & VICTIM	BREACH VECTOR & STOLEN DATA	ROOT CAUSE FAILURE	REDACTORR ARCHITECTURAL PREVENTION PROOF
Medibank Private (2022) 9.7M Australians	Exfiltration of diagnostic records, Medicare IDs, psychiatric summaries, and clinical billing files.	Plaintext clinical notes held across shared database pools with broad internal read access.	Clinical Tokenisation (Layer 07): Medicare IDs, practitioner numbers, and diagnoses masked into structural tokens ([MEDICARE_1]). Raw clinical text is unreadable if exfiltrated.
Optus (2022) 9.8M Customers	Identity payloads containing Passports, Driver Licences, DOBs, and addresses scraped via unauthenticated API.	Presentation layer directly returned raw statutory identity credentials instead of masked tokens.	Identity Defense (Layer 02 & 04): Pre-flight regex and OCR identify Australian driver licences and passport checksums instantly, rendering API responses benign.
HWL Ebsworth (2023) 65+ Federal Agencies	4TB of confidential legal briefs, court pleadings, and government litigation documents published on dark web.	Unrestricted legal drafting with external AI tools and file sharing across unencrypted matter files.	Legal Matter Isolation (Layer 06 TITAN): Client names, privileged settlement terms, and court citations are scrubbed in browser memory before legal drafting.
Latitude Financial (2023) 14M Credit Files	Massive historic credit applications, TFNs, bank statements, and salary records exfiltrated via vendor login.	Tabular data in spreadsheets and scanned PDFs retained without cryptographic token mapping.	Tabular Layout Guard (Layer 13 NEXUS): Cells containing TFNs, BSBs, and financial data are tokenised with spatial layout preserved, making datasets unexploitable.

04. 15-LAYER AEGIS ENGINE ARCHITECTURAL CONTROLS

ZERO-DATA RETENTION PIPELINE

Deterministic Pattern Shield (Layers 01–05) High-performance WASM regex arrays matching TFNs (with Modulus-11 checksums), Australian Medicare numbers, BSB/Account pairings, Passport formats, and phone numbers with 99.8% precision.	Spatial & OCR Engine (Layers 06–10) In-browser layout parsing for PDFs, Word documents, and images. Preserves tabular alignment and markdown structures while replacing sensitive entities with semantic tokens.
Local Contextual NER (Layers 11–13) Small-footprint, on-device Named Entity Recognition resolving ambiguous executive names, addresses, and organizations without transmitting text over the wire.	Cryptographic Verification (Layers 14–15) Generates a tamper-proof SHA-256 hash of every reviewed document session, creating an immutable audit trail for internal security reviews and OAIC compliance reporting.

05. CFO FINANCIAL LIABILITY & COST-BENEFIT MODEL

ENTERPRISE RISK MITIGATION ROI

Metric / Cost Driver	Unprotected Enterprise Workflow	Redactorr Client-Side Protected	Net Financial Advantage
Average Australian Breach Cost	\$4,030,000 per incident (IBM Security Cost of a Data Breach Report).	\$0 (Zero data stored on cloud servers; no exfiltratable plaintext).	100% Elimination of centralized server liability.
OAIC Maximum Civil Penalties	Up to \$50,000,000 per serious interference with privacy.	\$0 (Full compliance with APP 11 & cross-border APP 8 safeguards).	\$50M Risk Shield for Board of Directors and C-Suite.
Manual Redaction Labor (100 Staff)	18 min per complex brief = 7,500 hrs/yr (\$450,000/yr at \$60/hr).	2 min per brief with AI assist = 833 hrs/yr (\$50,000/yr).	\$400,000 / Year direct operational productivity saving.
Air-Gapped & Sovereign Deployment	Prohibitive custom proxy engineering costs (\$250k+ setup).	Zero external infrastructure footprint; deploys via simple web or VPC binary.	Instant Deployment with zero infrastructure overhead.

06. CISO ARCHITECTURAL CONTROLS MATRIX

VERIFIABLE SYSTEM GUARANTEES

Control Area	Implementation Specification	Audit Proof Artifact
Zero Data Persistence	Documents processed in Ephemeral WASM memory. Zero disk writes or server-side logging.	Client-side memory profiler & Network tab inspector verification.
Encryption at Rest & In Flight	AES-256 client-side session storage with zero plaintext egress. TLS 1.3 for license checks.	Cryptographic verification ledger & local key isolation.
Deterministic Integrity	SHA-256 proof hash calculated per document redaction event.	Exportable audit ledger with cryptographic verification receipts.
Key Custody	Client generates and retains all encryption keys in browser local storage / KMS.	Zero vendor key escrow; complete sovereign control.

Enterprise Sign-Off & Governance Authority

Redactorr Pty Ltd operates under Australian corporate jurisdiction. For air-gapped VPC deployments, on-premise container images, and custom industry rule packs, contact our Enterprise Security Architecture team at security@redactorr.com.

REGISTERED ENTITY: REDACTORR PTY LTD ABN: 91 487 406 432 LOCATION: MELBOURNE, AUSTRALIA