

PolicyLoop

MVP Product Requirements Document

Version 1.0 | Confidential | April 2026

Product	PolicyLoop	Author	Product Lead
Version	1.0 — MVP	Status	Draft for team review
Target customer	Sri Lanka Data Protection Authority	Launch target	GPA December 2026

1. Problem Statement

Data protection authorities in emerging and mid-capacity markets cannot effectively process public consultation responses on draft regulations. They receive either too few submissions (low participation) or too many unstructured ones (impossible to synthesise manually). The result is policy that does not reflect the best available input — and no public record of how submissions influenced the final regulation. Drafters default to organised, well-resourced stakeholders whose submissions are legible, entrenching elite capture.

PolicyLoop solves the drafter's problem first. If the drafter cannot use the tool, stakeholder participation has no channel. If the channel exists and is trustworthy, participation improves. The response matrix — a structural, non-optional feature — closes the loop and makes future participation worthwhile.

2. Goals & Success Metrics

MVP Goal

Ship a working product before the first contract is signed. Use the working product to close Sri Lanka's data protection authority as first customer, DIFC as second, and launch at the Global Privacy Assembly (GPA) in December 2026 with two live case studies.

Acceptance Criteria — Sri Lanka Consultation

Metric	Target	How measured
Synthesis quality	≥85% precision and recall vs human expert benchmark	Benchmark test on 3–5 historical DPA consultations before go-live
Drafter decision time	≤3 min per clause cluster, no training	Observed usability test with Sri Lanka drafter
Submission friction	Complete response in ≤30 min	Timed test with 3 representative stakeholder types
Response matrix published	Within 5 business days of consultation close	System timestamp on matrix publication

Stakeholder diversity	≥3 distinct stakeholder types in submission corpus	Submission type breakdown in consultation report
-----------------------	--	--

3. User Personas

Persona 1 — The Regulator Drafter

Who	A policy analyst or legal officer at a data protection authority. Responsible for managing a public consultation, reviewing submissions, producing a revised draft, and publishing a response.
Wants	Process 50–500 submissions on a draft regulation within a 4–6 week window, identify the key positions, produce a defensible revised draft, and publish a structured response showing how each point was handled.
Fails now	Drowning in unstructured PDFs and email submissions. No systematic way to cluster themes. Defaults to submissions from organised industry groups because those are the only ones legible. Spends weeks on manual synthesis with no guarantee of completeness.
Success	Can review all clause-level positions in a structured view, record decisions in the platform, and generate a response matrix at the click of a button. Time from consultation close to published response drops from weeks to days.

Persona 2 — The Organised Stakeholder

Who	A lawyer, compliance officer, policy analyst, or executive at a regulated entity, law firm, industry association, or civil society organisation. Engages regularly with regulatory consultation processes.
Wants	Submit a structured, persuasive response that actually influences the final regulation. Track whether the submission was considered. Build a record of policy engagement.
Fails now	Consultation portals are generic, unstructured, and give no feedback. Submissions feel like they go into a void. No confirmation that the response was received meaningfully or had any effect.
Success	Submits a clause-level response in a structured format. After consultation closes, receives a notification showing exactly which of their positions were accepted, partially accepted, or rejected — with the regulator's reasoning.

Persona 3 — The Public Submitter

Who	An individual citizen or small organisation with a stake in the regulation but no specialised legal or policy capacity. English-literate, web-capable.
Wants	Submit a view on a regulation that affects them. Understand what the regulation says in plain language.
Fails now	Consultation documents are impenetrable. Submission formats assume legal expertise. No feedback after submission.
Success	Reads a plain-language summary of each clause. Submits a freeform comment attached to the relevant clause. Receives outcome notification. Does not need legal training to participate meaningfully.

4. Core User Flows

Flow A — Regulator Drafter

- 1. Upload draft** — Uploads draft regulation (PDF or Word). System parses into clause tree and assigns stable clause IDs. Drafter reviews and corrects parsed structure.
- 2. Configure consultation** — Sets title, description, deadline. Attaches targeted questions to specific clauses. Sets eligible submitter types. Publishes consultation.
- 3. Monitor submissions** — Real-time dashboard showing submission volume, stakeholder type breakdown, and per-clause engagement heatmap. Individual submissions readable in full.
- 4. Review AI synthesis** — For each clause, views clustered positions with supporting submissions, stakeholder attribution, and evidence quality. Corrects clusters where AI has misfired.
- 5. Record decisions** — For each cluster, records decision (Accepted / Partially Accepted / Rejected / Noted) with free-text reasoning. Exports revised draft as Word document with track changes.
- 6. Publish response matrix** — System generates structured response matrix. Drafter reviews and publishes. Every submitter is notified automatically.

Flow B — Organised Stakeholder

- 1. Discover consultation** — Lands on the consultation page via direct link or future alert feature. Reads consultation description and deadline.
- 2. Read draft** — Views draft rendered clause by clause with plain-language summaries. Downloads full draft if needed.
- 3. Submit response** — Answers regulator's targeted questions. Adds clause-level commentary against specific provisions. Optionally selects verified-anonymous mode.
- 4. Verify identity (if anonymous)** — Platform verifies company registration or individual identity. Regulator sees only a stakeholder profile tag. Identity is not disclosed.
- 5. Receive outcome** — After consultation closes, receives email notification with link to response matrix showing how their positions were treated and the regulator's reasoning.

Flow C — Public Submitter

- 1. Arrive at consultation** — Lands on public consultation page. Reads plain-language description.
- 2. Browse draft** — Reads clause summaries. Identifies the clause or clauses relevant to them.
- 3. Submit comment** — Types freeform comment against a clause. Provides name and email. Submits.
- 4. Receive outcome** — Email notification when consultation closes with link to how the relevant clause was addressed.

5. Feature Scope — In / Out Table

Every feature discussed in planning appears in one column or the other. Nothing sits in a grey zone. Changes to this table require explicit decision at product-lead + CTO level with date recorded.

IN SCOPE — MVP	OUT OF SCOPE — MVP
✓ Draft regulation upload (PDF / Word)	X Pre-policy call for input (open-ended)
✓ AI clause tree parsing with stable IDs	X AI drafting assistant
✓ Plain-language clause summaries (2 reading levels)	X Sandbox (any mode)
✓ Targeted question builder per clause	X Inter-agency coordination tool
✓ Consultation configuration and publishing	X Complaints and enforcement analytics
✓ Structured stakeholder submission interface	X Stakeholder subscription product
✓ Clause-by-clause commentary fields	X Peer policy intelligence dashboard

✓ Verified-anonymous submission layer	X Cross-consultation analytics
✓ Stakeholder type classification	X Public deliberation layer (agree/disagree on positions)
✓ Real-time drafter dashboard	X WhatsApp / SMS / USSD / voice input
✓ AI synthesis: position extraction and clustering	X Languages other than English
✓ Stakeholder type and evidence attribution per cluster	X Mobile app
✓ Human-in-the-loop cluster correction	X White-labelling per regulator
✓ Drafter decision recording with reasoning	X Third-party integrations (Slack, CRM, etc.)
✓ Revised draft export (Word with track changes)	X API access for third parties
✓ Response matrix generation and publication	X Advanced stakeholder analytics
✓ Submitter outcome notification (email)	X Participation history tracking
✓ Role-based access control	X Consultation alert / feed features
✓ Full audit logging	X In-platform drafting collaboration
✓ Data residency controls (UAE, Sri Lanka)	X Multi-jurisdiction dashboard
✓ Encryption at rest and in transit	X Blockchain / verifiability features
✓ SSO for government users	X Impact simulation of any kind
✓ Demo instance with populated historical consultation	

6. Data Model Specification

The clause-anchored data model is the foundational architecture decision. Every other feature is built on top of it. It must be correct at MVP — retrofitting it later is prohibitively expensive. This section defines entities and relationships in plain language. Full database schema is owned by the engineering lead in the technical spec.

Consultation

A single regulatory consultation event.

- `id` – UUID, globally unique
- `title`, `description`, `jurisdiction`
- `status` – `draft` | `open` | `closed` | `published`
- `open_at`, `close_at` – timestamps
- `created_by` – regulator user ID
- `version_history` – array of draft document references

Clause

An atomic unit of the draft regulation with a stable identifier.

- `id` – UUID, stable across all revisions of the document
- `consultation_id` – parent consultation
- `path` – human-readable address (e.g. Part 2 > Section 4 > Clause 3)
- `text` – current version of clause text
- `plain_summary` – AI-generated plain-language summary
- `targeted_questions` – array of questions set by the regulator
- `revision_history` – array of {`text`, `timestamp`, `author`}

Submission

A complete response from one stakeholder to one consultation.

- `id` – UUID
- `consultation_id`
- `submitter_type` – `entity` | `association` | `civil_society` | `academic` | `individual` | `government`
- `identity_mode` – `identified` | `verified_anonymous`
- `identity_tag` – shown to regulator if anonymous (e.g. 'Financial services firm, 100-500 employees, Sri Lanka')
- `submitted_at` – timestamp
- `targeted_question_responses` – keyed by question ID
- `clause_comments` – array of `ClauseComment` (see below)
- `attachments` – array of file references

ClauseComment

A single stakeholder comment attached to a specific clause.

- `id` – UUID
- `submission_id` – parent submission
- `clause_id` – clause this comment addresses
- `text` – free-form comment
- `stance` – supports | opposes | seeks_amendment | neutral | not_stated
- `evidence_refs` – array of attachment references cited

Cluster

A group of submissions sharing a semantically distinct position on a clause.

- `id` – UUID
- `clause_id` – clause this cluster addresses
- `label` – short human-readable label (AI-generated, drafter-editable)
- `summary` – 2-4 sentence summary of the position
- `submission_ids` – array of submissions in this cluster
- `stakeholder_type_breakdown` – count by type
- `evidence_quality` – AI-assessed: strong | moderate | weak | none
- `proposed_redline` – AI-suggested amendment language (nullable)
- `correction_log` – array of {action, before, after, corrected_by, timestamp}

Decision

A drafter's recorded response to a specific cluster.

- `id` – UUID
- `cluster_id`
- `outcome` – accepted | partially_accepted | rejected | noted
- `reasoning` – free-text, required before consultation can close
- `decided_by` – drafter user ID
- `decided_at` – timestamp

ResponseMatrix

The public document generated at consultation close.

- `id` – UUID
- `consultation_id`
- `published_at` – timestamp
- `entries` – array of {clause_id, cluster_id, cluster_label, outcome, reasoning}
- `notification_sent_at` – timestamp when submitters were emailed

VerifiedIdentity

Stored separately from main data model with stricter access controls.

- `id` – UUID
- `submission_id` – foreign key to Submission
- `real_identity` – encrypted; company registration or individual ID
- `verification_method` – `company_register` | `id_document` | `email_domain`
- `verified_at` – timestamp

NOTE: This table is stored in a separate database instance with independent access controls.

Key constraint: Clause IDs persist across all revisions of the draft. When a drafter renumbers provisions, the IDs do not change. Path labels update; IDs do not. This is the single most important invariant in the system.

7. AI Synthesis Specification

The synthesis engine is the product's core AI capability and its primary technical risk. This section specifies behaviour precisely so that engineers and the ML team are building to the same standard and evaluating against the same criteria.

What goes in

- All ClauseComments for a given `clause_id`, with submitter type, stance, and evidence references
- The full clause text for context
- The regulator's targeted questions for that clause (to understand what the regulator was asking about)

What comes out (per clause)

- **Clusters** — semantically distinct positions, not sentiment buckets. If 17 submissions oppose Clause 4 but for three different reasons, that is three clusters not one.
- **Cluster label** — 5–10 word description of the position. Drafter-editable.
- **Cluster summary** — 2–4 sentences capturing the argument and its basis.
- **Stakeholder attribution** — count and type breakdown of submissions in each cluster.
- **Evidence quality assessment** — strong (empirical data cited), moderate (reasoned argument), weak (assertion only), none.
- **Proposed redline** — where the cluster implies a specific textual change, a candidate amendment. Nullable. Drafter can accept, modify, or ignore.
- **Conflict flags** — where two clusters hold irreconcilable positions, flagged for drafter attention.

Human-in-the-loop intervention points

- Drafter can split one cluster into two (if the AI over-grouped distinct positions)
- Drafter can merge two clusters into one (if the AI over-split similar positions)
- Drafter can reclassify a submission from one cluster to another
- Drafter can edit the cluster label and summary
- All corrections are logged in the cluster's `correction_log` with timestamp and user ID
- Corrections are used to improve the model for subsequent consultations

Uncertainty handling

- If the model cannot determine a submission's stance on a clause with >70% confidence, it flags the submission as 'needs review' rather than silently misclustering it
- A clause with fewer than 3 submissions does not trigger clustering — submissions are shown directly
- If clustering produces only one cluster (all submissions agree), this is surfaced explicitly as 'consensus' rather than a single-cluster synthesis

Evaluation criteria (pre-launch gate)

Metric	Definition	Target
Cluster recall	Did the system identify every distinct position a human expert would identify?	≥85%
Cluster precision	When the system groups submissions under one cluster, do human experts agree?	≥85%
Evidence attribution	When the system says a position is supported by evidence, is that evidence in the submission?	≥90%
Drafter correction rate	Fraction of clusters the drafter accepts without modification	≥75% (monitored, not gating)

Benchmark dataset: 3–5 completed public consultations from comparable DPAs (Singapore PDPC, Philippines NPC, Ghana DPC) where human expert clustering is available or can be produced. Must hit all gating targets before Sri Lanka consultation goes live.

8. Non-Functional Requirements

Performance

- Page load ≤2s on standard broadband for all primary drafter and stakeholder views
- Synthesis job for 500 submissions completes in ≤15 minutes
- Response matrix generation ≤30 seconds

Availability

- 99.5% uptime during active consultation windows
- Planned maintenance windows communicated 72 hours in advance

Security

- Encryption at rest (AES-256) and in transit (TLS 1.2+)
- Role-based access control: regulator admin, regulator drafter, stakeholder, public (read-only)
- Full audit log of all read and write operations on consultation and submission data
- SSO support for government identity providers (SAML 2.0 / OIDC)
- Verified anonymous identity layer in a physically separate database instance with independent access controls and separate encryption keys
- SOC 2 Type I readiness by month 10 of build
- Annual penetration testing from month 12

Data residency

- Sri Lanka consultation data: stored in Asia-Pacific region (Singapore or Mumbai preferred)

- DIFC consultation data: stored in UAE region
- Configurable per-consultation residency for subsequent customers
- No cross-jurisdiction data replication without explicit regulator consent

Privacy

- GDPR-compliant data processing (applicable to EU-based submitters)
- Sri Lanka PDPA compliant
- UAE PDPL compliant (for DIFC)
- Data processing agreements in place before any customer deployment
- Submitter personal data deleted on request per applicable law
- Retention schedule: verified identity records purged 6 months after consultation close; submission content retained per regulator's record-keeping obligations

Accessibility

- WCAG 2.1 AA for public submission interface
- Tested with screen readers before Sri Lanka go-live

9. Timeline

Phase	Months	Key deliverables	Gate to proceed
0 — Foundation	0–2	Seed close, team formed, architecture spec signed off, Sri Lanka and DIFC relationships formalised in writing	Technical co-founder in place; seed term sheet signed
1 — MVP build	2–7	All in-scope features shipped, synthesis benchmark passed, demo instance live with populated historical consultation	Synthesis hits ≥85% on benchmark; drafter usability test passed
2 — First contract	7	Sri Lanka DPA signs contract after seeing working product	Signed contract in hand before consultation opens
3 — Live consultation	7–11	Sri Lanka runs full consultation on PolicyLoop. Response matrix published. Case study written.	Response matrix published within 5 days of close
4 — DIFC + GPA	11–14	DIFC contract signed (month 11–12). GPA December 2026 launch with two case studies.	Two complete case studies before GPA date

Hard constraint: If MVP slips past month 8, the Sri Lanka consultation cannot complete before GPA 2026. GPA then becomes a vision pitch, not a launch. Month 7 is the deadline that must not move.

10. Open Questions Log

Decisions not yet made. Each must be explicitly closed before it blocks a build decision. Owner is responsible for driving resolution by the date shown.

#	Question	Owner	Resolve by	Status
1	Which specific Sri Lanka consultation (which regulation, which timeline) will be the first live run?	CEO	Month 2	Open

2	What is the commercial structure for Sri Lanka pilot — donor-funded, paid, or free with LOI?	CEO	Month 2	Open
3	Which historical DPA consultations will form the synthesis benchmark dataset?	CTO	Month 3	Open
4	White-label per regulator or shared marketplace? Affects branding, URL structure, and deployment architecture.	Product Lead	Month 3	Open
5	What verification methods are acceptable for verified-anonymous submissions in Sri Lanka context?	Product Lead	Month 3	Open
6	Which LLM provider(s) for synthesis? API vs self-hosted? Data residency implications?	CTO	Month 3	Open
7	Multilingual support timeline — Sinhala and Tamil for Sri Lanka public submissions?	Product Lead	Month 4	Open
8	DIFC relationship structure — design partner, paid pilot, unpaid pilot, or investor?	CEO	Month 4	Open

11. Non-Negotiables

Clause-anchored architecture	The clause ID system is the foundational decision. Every other feature depends on it. Retrofitting it after launch is prohibitively expensive. It must be correct at MVP, even though users do not see it directly.
Response matrix as structural property	The system does not permit a consultation to close without a published response matrix. This is not a configurable feature. It is not a toggle. Customers who cannot accept this self-select out. It is the product's moral spine and its primary differentiator from every existing civic engagement platform.

This PRD is the operating document for the MVP build. Changes to feature scope, the data model, or the non-negotiables require explicit sign-off from the product lead and CTO. The open questions log is reviewed at the start of each sprint and updated to reflect actual decisions. A PRD that does not update as the product develops is fiction.