DPA Guard — Product Requirements Document (PRD)

1. Document Control

- **Product Name:** DPA Guard
- **Internal Codename:** GDPR DPA Checker
- **Document Type:** Product Requirements Document (PRD)
- **Version:** 1.0 (Draft)
- **Status:** In Review
- **Owner:** Product / Engineering
- **Audience:** Product, Engineering, Security, Legal, Pilot Customers
- **Target Market:** EU / Germany-based SMEs (20–200 employees)
- **Last Updated:** 2026-01-10

---

2. Purpose

Define the requirements for an MVP that helps SMEs **review Data Processing Agreements (DPAs) fast and safely**. The product is **decision support, not legal advice**.

The MVP must: - Extract key DPA clauses - Assess risk using a structured **Playbook** (Green/Yellow/Red) - Provide **evidence-first** outputs (quotes + section references) - Suggest negotiation asks / fallback wording - Export a shareable report (PDF + copyable table)

---

3. Problem Statement

DPAs are required for vendor onboarding but are often reviewed by non-lawyers under time pressure. Common failure modes: - Missing/weak GDPR Article 28 processor obligations - Unclear subprocessors, transfers, deletion, audits, breach timelines - Poorly structured documents making manual review slow

Teams need a **reliable, repeatable** way to surface issues and produce a negotiation-ready change request pack.

---

4. Goals and Success Metrics

4.1 Goals (MVP)

1. Reduce time-to-first-review of a DPA.
2. Produce consistent, playbook-based risk flags with explicit evidence.

3. Enable vendor negotiation with a “change request pack” (table + email draft).
4. Provide an “Escalate to Legal” path when high risk remains.

4.2 Non-Goals

- Legal advice or “approve/deny” authority
- Full CLM workflows, e-signature
- Broad contract types beyond DPA/GDPR Addendum
- Agentic auto-negotiation, auto-apply redlines (Phase 2)
- OCR for scanned PDFs (optional later)

4.3 Success Metrics (Definition of Done)

- Works reliably on **≥20 sample DPAs** (text-based PDF/DOCX)
- **Every flagged issue** includes evidence (exact quote + section references)
- Produces:
- 1-page exec summary
- Clause-by-clause risk table
- Negotiation pack (commentary table + email draft)
- PDF export
- Pilot users report meaningful time savings and confidence in negotiation outputs

---

5. Users, Personas, and Primary Use Cases

5.1 Primary Personas

1. **Ops / Procurement Lead (SME)**
2. Needs quick “go/no-go with changes” decision
3. Wants a vendor email + change requests
4. **Security / DPO / Privacy Champion**
5. Focuses on transfers, TOMs, audits, breach, subprocessors
6. Needs evidence and escalation triggers

5.2 Core Use Cases

- Review a new SaaS vendor DPA before onboarding
- Generate a negotiation-ready list of requested changes
- Provide an internal approval packet for leadership/legal

---

6. Scope

6.1 In-Scope (MVP)

- **Contract type:** DPA / GDPR Addendum only
- **Language:** English-first

- **Inputs:** PDF (text-based) and DOCX
- **Outputs:**
- Executive summary (1 page)
- Clause risk table (copyable)
- Negotiation pack (commentary table + email draft)
- PDF export

6.2 Out of Scope (MVP)

- Scanned PDFs requiring OCR
- Word tracked changes output
- German-language clause evaluation
- Full vendor onboarding workflows

---

7. User Journey and Key Screens

7.1 Reference Flow

1. **Create New Review**
2. Select document type: DPA
3. Enter minimal context
4. **Upload Document** (PDF/DOCX)
5. **Processing**
6. Text extraction → segmentation → classification → evaluation
7. **Results**
8. Decision recommendation + top risks
9. Clause table with risk labels and evidence
10. **Generate Negotiation Pack**
11. Commentary table + email draft
12. **Export**
13. PDF report

7.2 MVP Screens (Web)

- Upload & context form
- Processing status page
- Results page (exec summary + risk table)
- Export/download page (PDF + copyable table)

---

8. Required Context Inputs (Minimal)

Collected at review creation to tailor Playbook logic: - **Customer role:** Controller / Processor (or unknown) - **Processing region:** EU-only / includes non-EU / unknown - **Data types:** customer data / employee data / special categories (optional) - **Vendor type:** SaaS / hosting / support / other (optional)

If the user skips context, the system uses conservative defaults and flags uncertainties.

---

9. Clause Coverage (MVP)

The system must extract and evaluate at least the following clause types: 1) Parties & roles (Controller / Processor) 2) Subject matter & duration of processing 3) Nature & purpose of processing 4) Data categories + data subjects 5) Security measures / TOMs (Art. 32) 6) Subprocessors (list/notice/objection) 7) International transfers (SCC/transfer mechanism, locations) 8) Breach notification (timeline + procedure) 9) Assistance with data subject rights (DSAR) 10) Return/deletion after termination 11) Audit/inspection rights 12) Confidentiality of personnel 13) Liability / limitation (if present in DPA or referenced) 14) Governing law/ jurisdiction (if included) 15) Order of precedence with MSA (conflict handling)

---

10. Functional Requirements

10.1 Document Ingestion

FR-ING-01 Upload DPA as **PDF (text-based)** or **DOCX**. - Accept file size limits (TBD; MVP default: 25 MB) - Store original file securely for the review session

FR-ING-02 Extract text while preserving structure. - Preserve headings/numbered sections when possible - Store extracted text with section identifiers

FR-ING-03 Basic validation - Detect likely scanned PDFs (very low text density) and show: “OCR not supported in MVP.”

10.2 Segmentation (Clause/Section Splitting)

FR-SEG-01 Split extracted text into segments (sections/clauses). - Output: segment_id, heading/number (if available), text, page reference (if available)

FR-SEG-02 Allow merging/splitting segments in UI (Nice-to-have; may defer).

10.3 Clause Classification

FR-CLS-01 Assign each segment to one or more clause types. - Use deterministic rules + LLM classification as needed - Store confidence score per assignment

FR-CLS-02 Unmapped segments - If a segment is not mapped, label as “Other / Unclassified” and keep for traceability

10.4 Evaluation (Playbook-Based)

FR-EVAL-01 For each clause type in scope, evaluate against the SME Playbook. - Inputs: relevant segment(s) + user context + playbook snippet - Output per clause type: - risk_label: Green / Yellow / Red - short_reason

(2–3 sentences) - suggested_change (fallback wording or negotiation ask) - evidence: exact quote(s) from the segment text - references: section title/number (if available)

FR-EVAL-02 Evidence-first requirement - Every Yellow/Red must include at least one exact quote. - If evidence is missing/uncertain, label as Yellow/Red with reason “Not specified / unclear” and cite the surrounding context (e.g., “No breach timeline found”).

FR-EVAL-03 Consistency requirement - Evaluations must be based on: - (a) clause text - (b) user context - (c) playbook rules - The system must not invent clauses or cite text not present.

10.5 Executive Summary (Decision)

FR-SUM-01 Produce a 1-page summary including: - Recommended decision: **Sign / Sign with changes / Escalate to Legal** - Top 3–5 risks (with clause references) - Notes on key assumptions/unknowns (if context missing)

Decision logic (MVP): - Any **Red** → “Escalate to Legal” OR “Sign with changes” depending on playbook rule severity (configurable) - Many Yellows or critical missing clauses → “Sign with changes”

10.6 Clause Table Output

FR-TBL-01 Display a clause-by-clause table with: - Clause type - Risk label - Short reason - Suggested change - Evidence (quotes + section refs)

FR-TBL-02 Copyable table - Provide a “Copy” control that preserves column structure (e.g., clipboard TSV/ Markdown)

10.7 Negotiation Pack

FR-NEG-01 Generate a negotiation-ready pack: - Commentary table (focused on Yellow/Red) - Email draft addressed to vendor (customizable tone)

FR-NEG-02 Email draft must include: - Clear list of requested changes - References to clauses (section/ heading) - Non-legal-advice disclaimer language

10.8 Export

FR-EXP-01 Export to PDF: - Executive summary page - Clause risk table - Negotiation pack appendix (optional toggle)

FR-EXP-02 Ensure exported PDF contains evidence snippets and references.

10.9 Review History (MVP-level)

FR-HIS-01 Store reviews per organization/user: - Upload metadata, timestamp, vendor name (optional), results - Allow user to reopen and re-export

11. Playbook Requirements (Policy Engine)

11.1 Playbook Structure (Minimum)

Each rule must include: - rule_id (e.g., DPA-TR-01) - clause_type - requirement (what must be present/true) - severity (Low/Med/High) - red_flag patterns/keywords - preferred wording + fallback wording OR negotiation ask - rationale (1–2 lines)

11.2 Playbook Retrieval (Light RAG)

- Retrieve only the relevant playbook snippet per clause type.
- The evaluation step must internally record which rule(s) triggered, and present user-friendly reasons.

11.3 Default SME Posture

- Conservative EU-trust baseline
- Escalate on unclear transfers, weak deletion, missing audit rights, missing breach timeline, broad subprocessors without notice/objection, etc.

12. Non-Functional Requirements

12.1 Trust & Safety

NFR-TS-01 Evidence-first: No issue without evidence.

NFR-TS-02 Low hallucination: The system must not fabricate content, clause presence, or citations.

NFR-TS-03 Human-in-the-loop: Provide “Escalate to Legal” recommendation when high risk remains.

12.2 Privacy & Data Handling (EU baseline)

NFR-PR-01 Uploaded documents may contain sensitive business data. - Tenant isolation by organization - Encrypt in transit and at rest

NFR-PR-02 Data retention - Default retention policy (TBD) with delete-on-demand - Provide “Delete review” capability (MVP may be admin-only; but must exist in plan)

NFR-PR-03 Model usage transparency - Clear disclosure that documents are processed for review only - No training on customer documents (product promise)

12.3 Performance

NFR-PERF-01 Target time to results - MVP target: under 2 minutes for typical DPA (subject to doc size)

12.4 Reliability

NFR-REL-01 Deterministic pipeline steps where possible. - Log each step output for debugging

NFR-REL-02 Reproducibility - Same input + same playbook version should produce consistent results (allowing minor LLM variance; mitigate with structured prompting and temperature controls)

12.5 Security

NFR-SEC-01 Authentication and tenant isolation - Users belong to an org; data isolated by org

NFR-SEC-02 Audit logging - Track who uploaded and viewed exports

---

13. Compliance and Disclaimers

- The product provides **decision support** and is **not legal advice**.
- UI + exported reports must include a short disclaimer.

---

14. Reporting Requirements

14.1 Executive Summary Page

Must include: - Decision recommendation - Top risks - Key assumptions/unknowns - Vendor + review metadata (optional fields)

14.2 Clause Table

Must include: - Clause type - Risk - Reason - Suggested change - Evidence quotes + references

14.3 Negotiation Email

Must include: - Polite intro, request for updates - Bulleted change requests - Section references - Optional: attach negotiation table

---

15. MVP Acceptance Criteria (System-Level)

1. Uploads a PDF/DOCX DPA and produces results without manual intervention for at least 20 sample documents.
2. For each of the 15 clause types, the system either:
3. extracts + evaluates, or
4. marks as missing/unclear with a risk label and reason.
5. Any Yellow/Red row contains:

6. exact quote(s)
7. section/heading reference (if available)
8. Exported PDF includes executive summary + risk table and is readable and shareable.
9. Negotiation pack can be copied into an email and used directly.

---

16. Assumptions and Dependencies

- Input PDFs are text-based (not scanned)
- Playbook is available and versioned
- LLM provider available in EU-friendly deployment (deployment choice covered in technical design)

---

17. Risks and Mitigations (Product)

- **Risk:** Incorrect clause mapping → missed risks
- **Mitigation:** multi-signal classification + manual “Other” bucket + confidence scoring
- **Risk:** Hallucinated citations
- **Mitigation:** evidence extraction as a separate deterministic step; strict quoting rules
- **Risk:** Users hesitant to upload contracts
- **Mitigation:** security posture page + retention controls + EU trust messaging

---

18. Open Questions (To Resolve)

1. File size limits and retention defaults for MVP
2. Minimum context questions: which are mandatory vs optional?
3. Decision recommendation thresholds (how many Reds/Yellows trigger escalation)
4. Required export formats beyond PDF (CSV/Docx?)
5. Org features: SSO needed in MVP or later?

---

19. Next Document

After PRD review, we will create the **Technical Design Document (TDD)** covering architecture, data model, pipeline, prompting strategy, playbook retrieval, storage/security, and PDF generation.