Local-Only Resume & Document Redactor (Anonymizer)

Product Specification Document

---

1. Concept Overview

A browser-based, privacy-first tool that detects and removes personally identifiable information (PII) from resumes and simple documents entirely on the client side, with no uploads to a server.

Core User Flow

1. Users drop in a resume (PDF, DOCX, or text)
2. The app highlights likely PII (names, emails, phone numbers, addresses, LinkedIn URLs, company names, locations)
3. Users confirm/review replacements (e.g., "John Doe" → "Candidate A", "Google" → "Company X")
4. The app generates an anonymized version ready for:
    - Blind hiring processes
    - Sharing portfolio/case studies without exposing client or candidate data
    - Publishing examples online safely

Privacy-First Approach

All processing happens in the user's browser, following the "zero-upload" privacy trend seen in modern PII redaction and privacy tools where sensitive data is removed before it ever leaves the endpoint. This is positioned as a lightweight alternative to heavy server-side PII APIs, and explicitly focused on HR/recruiting use-cases instead of call centers or big data warehouses.

---

2. Target Users and Use Cases

Primary Users

- **Recruiters and HR teams** running anonymized/"blind" hiring pipelines
- **Hiring agencies** screening candidates while hiding client or candidate identity in early stages
- **Individual job seekers** who want to share resumes online or in portfolios without exposing personal details
- **Consultants, designers, and freelancers** who need to redact client/company names in case studies and proposals

Industry guides show that automated PII redaction is increasingly used to remove sensitive elements while preserving enough structure for review and analytics, especially in regulated or privacy-sensitive workflows. This tool applies the same concept to documents like resumes and simple business docs, but runs locally in the browser.

Example Workflows

Recruiter Workflow: Recruiter uploads a stack of resumes (Pro feature: batch). Each is anonymized to hide name/contact details, but keeps skills, experience, and dates.

Job Seeker Workflow: Candidate drops their resume, anonymizes PII, and uses that version in public portfolios or as attachments in communities.

Consultant Workflow: Consultant pastes a case study, quickly redacts client name and identifiers to share on LinkedIn or a blog.

---

3. Business Model

3.1 Free Tier (Lead + SEO + Virality)

Features:

- Single-document anonymization (one resume/doc at a time)
- Basic PII detection: emails, phones, URLs, typical name patterns, obvious addresses
- Support for 1–2 formats (e.g., PDF upload + raw text input)

Monetization: Display ads on landing pages, usage guides, and lightly on the tool page. This ties into a broader trend where privacy/compliance and HR tooling can attract higher-value B2B advertisers, raising CPM/RPM versus generic utility sites.

3.2 Pro Tier (Main Revenue)

Pricing (global, USD):

- **Individual:** $19–$29 one-time lifetime or $5/month
- **Teams** (small agencies/HR shops): $15–$29/month for 3–5 seats

Micro SaaS case studies in the resume/recruitment tooling space show that $8–$49/month price bands are common and accepted for niche tools that reduce manual effort.

Pro Features:

- **Batch anonymization:** process multiple resumes/docs in a single run
- **Custom redaction rules:** user-defined keywords or patterns to always redact (internal project names, internal codes, etc.)

- **Advanced entity types:** company names, locations, university names, project names
- **Export mapping file (local):** a JSON/CSV mapping original→placeholder so agencies can reference back without exposing to external systems
- **No ads, priority feature access**

3.3 Future Upsell Opportunities

Template Packs: Anonymized case study templates, anonymized resume templates for portfolios, etc.

White-label Mode (later): Agencies pay more to use their logo/colors (still client-side).

Integrations (later): Simple Chrome extension to anonymize text in web textareas (LinkedIn posts, job posts, etc.).

---

4. Architecture and Stack

4.1 Technical Stack

Frontend-only:

- React/Next.js (or similar) SPA/PWA
- Hosted on Vercel/Netlify (free tier)

PWA:

- Service Worker + Cache Storage for app shell
- Optional IndexedDB for storing user settings and Pro license tokens
- Best practices for PWAs emphasize secure HTTPS, careful caching, and avoiding storing sensitive data directly in cache; focus caching on static assets, not user content

4.2 PII Detection Logic

Start simple, iterate:

Phase 1 (Rule-based):

- Regex for email, phone numbers, URLs, postal codes
- Heuristics for names (capitalization, position in document, CV-specific patterns)
- Keyword lists for typical location words, common company suffixes (Inc, LLC, Ltd, GmbH, etc.)

This aligns with how many tools and libraries start with predefined categories and patterns for PII before layering AI.

Phase 2 (Smarter):

- Entity dictionaries for company names, locations, universities (pre-built lists or on-device assets)
- Optional lightweight NLP models in WASM/WebAssembly or via on-device JS models if performance allows, similar to how AI-based redaction tools extend rule-based detection for better coverage

Core Principle: All processing stays in memory within the browser; nothing is sent to a backend.

---

5. Things to Watch Out For While Building

5.1 Privacy and Security

Never upload user documents:

- Make it explicit in code and in your marketing that no file goes to your server
- Serve over HTTPS (needed for PWAs and recommended for any PII-related tool)

Avoid caching raw documents:

- Do not cache user files; only cache app assets
- Guides for PWAs suggest avoiding caching sensitive data and focusing on static resources instead

Clear privacy statement: A simple privacy page explaining: "All processing in your browser, no files stored on our servers" builds trust and aligns with privacy-by-design ideas used in PII tools.

5.2 UX and Accuracy

Show detections clearly:

- Highlight PII candidates and provide a side panel list
- Let users accept/reject replacements before finalizing

Be honest about limitations:

- Label results as "best-effort anonymization; user review required"
- This mirrors how serious PII-redaction tools talk about false positives/negatives and best-effort detection

Performance: Use Web Workers if needed for heavy processing on large documents so the UI stays responsive.

5.3 Legal / Compliance Positioning

You are a helper tool, not a compliance authority:

- Don't claim "GDPR-compliant by itself"; instead say it "reduces exposure of PII in documents" and that users are responsible for reviewing final output
- This is similar to how existing privacy/PII tools present themselves as aids in compliance, not legal guarantees

5.4 PWA & Offline Considerations

Focus on offline-friendly app shell:

- Allow the app UI to work offline for repeated usage after first load
- If user opens the app offline, show a custom offline page instead of a generic browser error, as recommended in PWA best-practices

Storage: Only store minimal configuration (user preferences, custom rules) in IndexedDB/localStorage; avoid auto-storing any original documents.

5.5 Monetization & Ads

Ad placement:

- Place ads on landing pages, guides, and result pages; keep the core editing area as clean as possible
- Follow AdSense/other networks' policies about ad density and layout (site quality and policy compliance are key for approval)

Pricing experiments: Start with a simple $19 lifetime plan and, if demand is strong, later introduce monthly/annual options, matching micro-SaaS patterns around job/resume tools.

---

6. Suggested v1 Scope

Must-have for v1:

- Upload or paste text from one resume/doc
- Detect and highlight: email, phone, URLs, common names, locations (basic heuristics)
- Let user confirm/adjust replacements and export anonymized text as:
    - Downloadable .txt or .docx (using a client-side library)
- Clear privacy statement and "all local" explanation
- Simple landing page describing use cases and benefits

Nice-to-have for v1.1:

- PDF input support with text extraction
- Basic PWA install support
- Custom rule editor (user can add words/phrases to always redact)

Summary

This specification provides a focused, global-fit micro-SaaS product: a browser-only resume/document anonymizer with a clear B2B/B2C privacy value proposition, simple frontend-only architecture, ad + Pro monetization model, and well-defined technical and UX considerations drawn from modern PII redaction and PWA best practices.