PARADOCS

The World's Largest Paranormal Database

COMPREHENSIVE STATUS ASSESSMENT

Product Readiness & Market Positioning Review

February 22, 2026

EXECUTIVE SUMMARY: ParaDocs has achieved ~75% of the vision outlined in the dev-handoff and user-journey documents. The platform has a mature database (~258K reports), sophisticated analytics, 8 data source adapters, and a polished UI. However, critical gaps remain in data quality assurance at scale, content moderation tooling, mobile optimization of key flows, and SEO infrastructure needed before public launch. The competitive landscape shows a $31B+ market with no modern all-in-one paranormal database, representing a significant first-mover opportunity. This document details what's working, what needs fixing, and how to get market-ready.

1. WHAT'S WORKING WELL

These are the features and systems that are aligned with the vision documents, implemented correctly, and ready for production or near-production use.

1.1 Core Database & Schema

Status: STRONG. The PostgreSQL/Supabase/PostGIS schema is comprehensive and well-normalized. 23 migrations track clean evolution. 258,000+ reports across all phenomenon categories. Key tables (reports, phenomena, report_phenomena junction, detected_patterns, profiles, subscriptions) are all properly structured with UUIDs, soft deletes, audit trails, and JSONB for flexible metadata. The 11-category taxonomy with 40+ named phenomena entries provides excellent classification depth.

1.2 Multi-Source Ingestion Engine

Status: STRONG. 8 source adapters (NUFORC, BFRO, Wikipedia, Reddit, IANDS, NDERF, Ghosts of America, Shadowlands) with a well-designed adapter pattern. Quality filtering, AI-powered title improvement via Claude, automatic phenomenon identification, and bulk processing capabilities. The pipeline architecture (Source > Adapter > Quality Filter > Title Improvement > Phenomenon ID > DB Insert) is sound and extensible.

1.3 User Experience & Design

Status: STRONG. The dark-themed, space-inspired UI is visually distinctive and on-brand. Tailwind CSS provides responsive design across breakpoints. Framer Motion animations add polish. The 4-step submission wizard, constellation research map, and gamification system (streaks, ranks, challenges) align well with the user-journey document's engagement strategy. The Explore page's dual-mode (personalized feed + advanced browse) is exactly what the dev-handoff specified.

1.4 Pattern Detection & Analytics

Status: STRONG. PostGIS DBSCAN clustering for geographic patterns, temporal anomaly detection, seasonal pattern analysis, and flap/wave identification. The insights dashboard with AI-generated narratives, skeptic mode for alternative explanations, methodology transparency panels, and uncertainty displays shows real intellectual rigor. This is a genuine differentiator that no competitor offers.

1.5 Constellation / Research Mapping

Status: GOOD. D3/Canvas-based network visualization lets users log investigations, draw connections between cases, build theories, and track research verdicts. The 5-level explorer rank system (Stargazer to Master Archivist) and research streak gamification drive engagement. Public profile sharing enables community discovery. This maps directly to the user-journey document's 'Investigator' persona.

1.6 Subscription & Monetization Infrastructure

Status: GOOD. Stripe integration with checkout, billing portal, webhook handling, and tier-based feature gating. The FeatureGate component, PaywallGate, UsageMeter, and UpgradeCard components create a complete subscription UX. Cancellation flow with retention offers is implemented. This aligns with the dev-handoff's monetization goals.

1.7 OG Images & Social Sharing

Status: COMPLETE (just fixed). Dynamic OG image generation via Vercel Edge functions for both branded fallback images and per-report social cards. Twitter Card meta tags included. The root cause of the extended

outage was 'export var config' vs 'export const config' for Edge runtime detection, which is now resolved. Both /api/og/fallback and /api/og/report endpoints are confirmed working with proper meta tags in _app.tsx and index.tsx.

2. WHAT NEEDS IMPROVEMENT

2.1 Data Quality at Scale (CRITICAL)

Status: NEEDS WORK. This is the single biggest blocker to public launch. With 258K reports, quality variance is enormous. The dev-handoff document emphasizes the mass ingestion plan needs 'extensive filtering requirements to ensure quality.' Current gaps:

- Quality scoring is basic - the quality-filter.ts does obviousness checks, length validation, and spam detection, but doesn't handle duplicate detection across sources, hallucinated location data, or semantic quality assessment
- No deduplication pipeline - reports from NUFORC/Reddit/other sources describing the same event aren't merged or flagged
- Credibility is manually assigned - no algorithmic scoring based on evidence, corroboration, witness count, or source reliability
- No content moderation queue - user submissions go straight to the database without human review workflow
- Missing quality metrics dashboard - no way to visualize data quality distribution across the 258K records
- Geocoding accuracy unknown - bulk-geocoded records may have imprecise or incorrect coordinates

Recommendation: Build an automated quality scoring pipeline that evaluates each report on 8-10 dimensions (evidence strength, witness count, description detail, location specificity, temporal precision, source reliability, corroboration potential, narrative coherence). Surface a quality score alongside credibility. Implement fuzzy deduplication using title/location/date similarity matching.

2.2 Filtering & Search Depth (HIGH PRIORITY)

Status: PARTIAL. The explore page has 10+ filter dimensions (category, phenomenon, country, state, credibility, date range, evidence type, content type, featured, sort). This is good. However, several gaps remain versus the dev-handoff spec:

- No radius-based geographic search - 'show me reports within 50 miles of my location' isn't supported
- No full-text search operators - users can't do phrase matching, boolean operators, or field-specific searches
- No saved searches/alerts - users can't save a filter configuration and get notified of new matches
- Subcategory filter UX is cluttered - with 40+ phenomena, the multi-select needs better grouping/search
- No cross-referencing filters - can't find reports that match multiple phenomena simultaneously
- Map filtering is limited - the map page has basic category filters but not the full filter suite

Recommendation: Implement PostGIS ST_DWithin for radius search. Add Supabase full-text search with ts_vector/ts_query. Build saved search functionality with email/push alerts. Bring the full filter suite to the map page.

2.3 SEO & Discoverability (HIGH PRIORITY)

Status: NEEDS WORK. Currently blocked by the beta 'noindex, nofollow' meta tag (intentional). But beyond removing that tag, the SEO infrastructure is thin:

- No JSON-LD structured data - report pages, phenomena pages, and the map should have Schema.org markup
- No sitemap.xml - with 258K report pages and 40+ phenomena pages, a dynamic sitemap is essential
- No robots.txt optimization - need to guide crawlers to high-value pages
- Missing canonical URLs - duplicate content risk from filter URLs

- No breadcrumb structured data - important for Google SERP rich results
- Page titles need optimization - should include location, category, and year for long-tail SEO

Recommendation: This should be the #1 pre-launch priority. Implement next-sitemap for dynamic sitemap generation. Add JSON-LD for Report, Place, Article, and BreadcrumbList schemas. Optimize title/description templates for each page type. Implement canonical URLs for all filterable pages.

2.4 Mobile Experience (MEDIUM PRIORITY)

Status: PARTIAL. Tailwind responsive classes are used throughout, and the Quick Facts mobile reorder was recently fixed. However:

- Map page is clunky on mobile - Leaflet on small screens needs touch-optimized controls and bottom-sheet report cards
- Submission wizard could be smoother - the 4-step flow needs mobile-specific UX (larger touch targets, swipe navigation)
- No PWA offline support - manifest.json exists but no service worker for offline reading
- Dashboard sidebar is desktop-oriented - needs a proper mobile navigation pattern
- Constellation map is desktop-only - D3/Canvas visualization doesn't scale to mobile screens

2.5 Admin & Moderation Tooling (MEDIUM PRIORITY)

Status: NEEDS WORK. The admin dashboard exists with stats and activity feeds, but lacks the operational tools needed for launch:

- No moderation queue - no way to review, approve, reject, or edit submitted reports in bulk
- No user management - no ability to ban, warn, or manage user accounts
- No content flagging workflow - users can't flag reports, and there's no triage system
- No automated spam detection - the quality filter works at ingestion but not on user submissions
- Missing audit logs - no trail of admin actions for accountability

2.6 Testing & Reliability (MEDIUM PRIORITY)

Status: NOT STARTED. No evidence of any testing infrastructure:

- No unit tests - quality filter, credibility scoring, pattern detection are untested
- No integration tests - API routes have no automated testing
- No E2E tests - critical flows (submit report, search, constellation) aren't tested
- No error monitoring - no Sentry or similar for production error tracking
- No performance monitoring - no Core Web Vitals tracking or lighthouse CI

3. FEATURE STATUS MATRIX

Comprehensive mapping of dev-handoff requirements against current implementation.

FEATURE	STATUS	NOTES
Report Database (258K+)	COMPLETE	PostgreSQL + PostGIS, well-normalized
Multi-source Ingestion	STRONG	8 adapters, quality filter, AI titles
Explore Page + Filters	GOOD	10+ dimensions, needs radius search & saved searches
Interactive Map	GOOD	Leaflet, 500 markers, needs full filter suite
Report Submission	GOOD	4-step wizard, evidence upload, anonymous option
Phenomena Encyclopedia	STRONG	40+ entries, AI content, report linking
Pattern Detection	STRONG	Geographic, temporal, seasonal, flap/wave
Credibility Scoring	PARTIAL	5-level manual system; needs algorithmic scoring
User Auth & Profiles	GOOD	Supabase auth, roles, reputation
Constellation Research Map	GOOD	D3/Canvas viz, theories, connections, ranks
Subscription/Stripe	GOOD	Checkout, portal, tiers, feature gates
Dashboard & Analytics	GOOD	Personal, public, admin variants
OG Images / Social Cards	COMPLETE	Dynamic Edge functions, branded fallback
Email Campaigns	GOOD	Resend integration, drip campaigns, digests
A/B Testing	GOOD	Basic framework with tracking
Data Quality Pipeline	NEEDS WORK	Basic filter exists; needs dedup, scoring, moderation
SEO / Structured Data	NEEDS WORK	noindex active; no sitemap, JSON-LD, canonical URLs
Content Moderation	MISSING	No review queue, flagging, or bulk admin tools
Testing Infrastructure	NOT STARTED	No unit, integration, or E2E tests
Error Monitoring	NOT STARTED	No Sentry, no performance tracking
API Rate Limiting	MISSING	Public APIs unprotected
API Documentation	MISSING	No OpenAPI/Swagger spec
Advanced Search	MISSING	No boolean operators, phrase matching
Saved Searches / Alerts	MISSING	Smart alerts cron exists but not user-facing
Mobile Native App	NOT STARTED	PWA foundation only

4. MASS INGESTION PLAN & QUALITY ASSURANCE

4.1 Current Ingestion Status

The ingestion engine currently supports 8 sources with varying quality levels:

SOURCE	TYPE	VOLUME	QUALITY	NOTES
NUFORC	UFO	~100K	Medium	Structured but self-reported, no verification
BFRO	Cryptid	~5K	High	Classified A/B/C, investigated by field researchers
Reddit	Mixed	~50K+	Variable	Batch importer, needs heavy quality filtering
Ghosts of America	Ghost	~30K+	Low-Med	User-submitted, minimal verification
Shadowlands	Ghost	~10K+	Low	Folklore-heavy, historical
Wikipedia	Mixed	~500	High	Encyclopedic, well-sourced
IANDS	NDE	~1K+	High	Academic research org
NDERF	NDE	~5K+	Medium	Self-reported but structured

4.2 Quality Assurance Requirements

To reach market readiness, the following quality layers must be implemented:

Layer 1: Automated Pre-Processing

- Duplicate detection - fuzzy matching on title + location + date within 30-day windows
- Language quality check - minimum character count, coherence scoring, gibberish detection
- Location validation - verify geocoded coordinates fall within claimed country/state boundaries
- Date plausibility - reject future dates, flag dates before 1800
- Source cross-referencing - flag when NUFORC + Reddit describe same event within 48hrs / 50mi

Layer 2: AI-Powered Quality Scoring

- Evidence assessment - score 0-100 based on detail level, specificity, evidence types
- Narrative coherence - flag internally contradictory accounts
- Credibility indicators - multiple witnesses, physical evidence, official documentation
- Phenomenon confidence - how well does the report match its classified phenomenon?
- Source reliability weighting - BFRO Class A > Reddit anecdote > anonymous submission

Layer 3: Human Review Pipeline

- Moderation queue - user submissions require approval before publication
- Community flagging - users can flag reports for review (duplicate, spam, inaccurate)
- Expert verification - high-impact reports get specialist review

- Batch review tools - admins can approve/reject/edit in bulk

4.3 Scale-Up Ingestion Roadmap

Target: 500K+ reports by public launch. Additional sources to integrate:

- MUFON public case summaries - 135K+ cases, membership may be required for full access
- The Black Vault - 3M+ pages of declassified government documents
- Phantoms & Monsters archive - 45+ years of investigated cases
- Historical newspaper archives - pre-internet reports from newspaper digitization projects
- International databases - GEIPAN (France), MoD (UK), CEFAA (Chile)
- Academic papers - Journal of Scientific Exploration, Parapsychological Association

5. COMPETITIVE LANDSCAPE & MARKET POSITIONING

5.1 Market Opportunity

The paranormal/unexplained phenomena market represents a **$31.89 billion opportunity** (ghost tourism alone), with related horror entertainment at $112B and growing at 7.2% CAGR. **60% of Americans report paranormal experiences** and **75% believe in the paranormal**. The market is fragmented across dozens of outdated, single-phenomenon platforms with no modern all-in-one database.

5.2 Competitor Comparison

CAPABILITY	ParaDocs	MUFON	NUFORC	BFRO	Gaia	Reddit
All phenomena types	YES	UFO only	UFO only	Bigfoot	Mixed*	Scattered
Interactive map	YES	No	Basic	Limited	No	No
Credibility scoring	YES	No	No	A/B/C	No	Upvotes
Data analytics	YES	No	No	No	No	No
Modern UX	YES	No	No	Basic	Yes	Yes
AI-powered features	YES	No	No	No	Partial	No
Free tier	YES	$10.99/mo	Free	Free	$15.99/mo	Free
User community	Building	4K members	None	Small	806K subs	1.4M+

**Gaia is a streaming platform, not a database. Their 806K subscribers watch documentaries rather than research specific cases.*

5.3 Key Demographics

- Primary audience: 18-34 year olds - 39% most likely to believe in the paranormal; digital natives
- Gender skew: slight female majority - women more likely to believe in paranormal phenomena
- Platform preference: TikTok, YouTube, Reddit - short-form video dominates discovery for Gen Z
- Spending patterns: $50-1,500 on equipment - ghost hunting gear, streaming subscriptions, paranormal tourism
- Engagement triggers: news coverage drives 200%+ search spikes - government disclosures, viral sightings

5.4 ParaDocs Market Position

ParaDocs occupies a unique position as the **only modern, all-in-one paranormal database with credibility scoring and data analytics**. No competitor offers this combination. The closest analogy is 'Wikipedia meets IMDb for the paranormal' with research tools layered on top. The key positioning should be: **neutral, science-friendly but believer-inclusive**, distinguishing from both the credulity of ghost-hunting shows and the dismissiveness of mainstream science. This balance is critical for credibility with the primary 18-34 demographic who are interested but skeptical.

5.5 Revenue Model Alignment

The current Stripe subscription infrastructure supports a multi-stream approach:

- Freemium SaaS ($4.99-9.99/mo) - free browsing, premium analytics/alerts/API
- Affiliate revenue - ghost hunting equipment, books, paranormal tourism partnerships
- Enterprise licensing ($5K-50K+) - media companies (History Channel, Netflix), academic institutions
- Data partnerships - licensed API access for researchers/institutions

6. LAUNCH READINESS CHECKLIST

Prioritized list of what must be done before public launch, organized by urgency.

P0 - Must Have Before Launch

#	TASK	EFFORT	IMPACT
1	SEO infrastructure (sitemap, JSON-LD, canonical URLs, remove noindex)	2-3 weeks	Critical for organic discovery
2	Data quality pipeline (dedup, automated scoring, validation)	3-4 weeks	Foundation of trust/credibility
3	Content moderation queue + admin review tools	2-3 weeks	Required for user submissions
4	API rate limiting and abuse prevention	1 week	Prevents scraping/abuse
5	Error monitoring (Sentry) + logging	3-5 days	Know when things break
6	Performance optimization & Core Web Vitals	1-2 weeks	Google ranking factor

P1 - Should Have for Launch

#	TASK	EFFORT	IMPACT
7	Radius-based geographic search (PostGIS ST_DWithin)	1 week	Key differentiator for map UX
8	Saved searches + email alerts for new matches	1-2 weeks	Retention & engagement driver
9	Mobile map UX overhaul (touch controls, bottom sheets)	1-2 weeks	60%+ traffic will be mobile
10	Algorithmic credibility scoring (evidence-based)	2-3 weeks	Core brand promise
11	Unit + integration tests for critical paths	2-3 weeks	Deployment confidence
12	Community flagging & report system	1-2 weeks	User trust & quality

P2 - Post-Launch Enhancements

#	TASK	EFFORT	IMPACT
13	Advanced full-text search (boolean, phrase, field-specific)	2 weeks	Power user feature
14	Public API with documentation (OpenAPI spec)	2-3 weeks	Developer ecosystem

15	Additional ingestion sources (MUFON, Black Vault, intl)	Ongoing	Data moat expansion
16	Mobile native app (React Native)	8-12 weeks	App store presence
17	TikTok/YouTube content strategy	Ongoing	Gen Z acquisition channel
18	Real-time WebSocket notifications	1-2 weeks	Live pattern alerts

7. OVERALL ASSESSMENT

ParaDocs is an **impressive 75%-complete product** that has already achieved technical parity or superiority to every competitor in the market on core database, analytics, and UX capabilities. The 258K-report database, 8-source ingestion engine, pattern detection, constellation research mapping, and subscription infrastructure represent months of sophisticated engineering work that aligns well with the dev-handoff and user-journey vision documents.

The gap to market-readiness is primarily operational, not technical. The platform needs data quality assurance at scale, content moderation tooling, SEO infrastructure, and monitoring/alerting rather than major new features. The core product vision is implemented and working.

Competitively, ParaDocs has a significant first-mover advantage. No competitor offers an all-in-one paranormal database with modern UX, interactive mapping, AI-powered analytics, credibility scoring, and community research tools. MUFON has credibility but outdated tech and only 4K members. NUFORC has data but terrible UX. Gaia has subscribers but is a streaming platform, not a database. Reddit has community but unstructured data. ParaDocs uniquely combines all of these strengths.

The $31B+ market is real and growing. Ghost tourism alone is projected to reach $40.82B by 2034. The 18-34 demographic is the primary growth segment with high digital engagement. Search volume spikes 200%+ around disclosure events. The timing is favorable with increasing government transparency on UAPs and mainstream acceptance of paranormal research.

With the P0 items addressed (estimated 8-12 weeks of focused development), ParaDocs will be positioned to launch as **the definitive paranormal research platform** with no direct competitor offering the same breadth and depth of features. The recommended launch strategy is to seed with Reddit communities (r/paranormal: 1.4M, r/UFOs, r/cryptids), partner with paranormal podcasters and YouTubers, and execute aggressive SEO targeting high-volume keywords like 'UFO database,' 'paranormal reports near me,' and 'cryptid sightings.'

Bottom line: The product vision is sound, the technical execution is strong, and the market opportunity is massive. The remaining work is achievable and well-defined. ParaDocs is closer to launch-ready than it might feel in the trenches of debugging Edge functions.