Open GR --WM

Your Local Graph RAG

All brains. No cloud bill panic.

- Turns PDF chaos into a structured knowledge graph you can chat with.
- Runs completely on your Mac with specialized local models.
- Built for real docs: annual reports, charts, and table-heavy narratives.

Privacy

100% Local

No external API calls

Model Roles

3 Specialists

Right model, right job

Feature 1: Smart Ingestion

Upload PDFs, route pages, skip wasted compute.

- Per-page routing: every page gets text extraction.
- Vision model only runs on pages with images or very low text density.
- Build modes let you trade speed vs depth.
- Normie translation: your laptop sweats only where it matters.

Routing

Page-level

No blanket vision burn

Control

Fast/Balanced/Thorough

One-click tuning

Feature 2: Three-Model Architecture

Speed, sight, and reasoning in one pipeline.

- Scanner: llama3.2 extracts entities and relations from text.
- Eyes: llama3.2-vision reads charts/tables from selected pages.
- Brain: deepseek-r1:14b handles final multi-hop reasoning in chat.
- Judge: fast local check catches unsupported claims before final output.

Scanner

llama3.2

Fast extraction

Brain

deepseek-r1:14b

Deep reasoning

Feature 3: Knowledge Graph Build

Entities + predicates become a queryable map.

- NetworkX MultiDiGraph stores rich directed relationships.
- Triples from text and vision are composed into one graph.
- Progress logs and stop controls keep long builds manageable.
- If parsing gets weird, robust JSON fallback logic prevents crashes.

Structure

MultiDiGraph

Direct + reverse neighbor lookup

Resilience

Robust Parser

Handles model filler

Feature 4: Persistent Graph Library

Build once. Reuse anytime.

- Auto-save graph, triples, chunks, and embeddings to disk.
- Load/unload from RAM on demand.
- Delete from disk to reclaim space.
- Disk size and source-file count are visible per graph.

Lifecycle

Load / Unload / Delete

Storage in your control

Metadata

Source-aware

Tracks augmented files

Open GR --WM

Feature 5: Graph Augmentation

Add more PDFs to the same graph over time.

- Choose 'Augment loaded graph' to merge new sources into existing knowledge.
- Embeddings and triples are appended; graph identity remains stable.
- Perfect for evolving research sets and periodic financial updates.
- Normie translation: one brain, many documents, less chaos.

Mode

Incremental

No full rebuild required

Continuity

Stable Graph ID

Same source of truth

Feature 6: Chat Studio

Grounded answers, citations, and visual callouts.

- Retrieval combines keyword graph neighbors + embedding top chunks.
- Answers cite [Graph] and [Chunk N] to show where facts came from.
- Hero figures and optional charts render in the response.
- Live thinking panel shows progress while the model reasons.

Grounding

Dual Retrieval

Graph + vector context

UX

Live Thinking

No dead-air waiting

Feature 7: Quality Guardrails

LLM judge checks before answer delivery.

- Judge model classifies answer as pass, retry, or not_found.
- Retry path asks for correction when grounding looks weak.
- If data is absent in context, app says so clearly.
- Normie translation: less confident nonsense, more useful honesty.

Judge

Fast Local

Low-latency validation

Fallback

Not Found

No fabricated answers

Open GR --WM

Feature 8: Aesthetic UX

Dark theme, playful glow, focus-first interactions.

- Two workspaces: Ingest + Graph Studio and Chat Studio.
- Progress bars, status bubbles, and dynamic states guide the user.
- Passive elements stay muted, active context lights up.
- Because serious tools can still look like they had coffee and style.

Design

Pastel on Dark

Readable and modern

Flow

Context-aware

Less tab confusion

How to Use It in 30 Seconds

Short path from PDF to answer.

1) Ingest tab: upload PDF and click Build (or Augment).
2) Load your graph from library when needed.
3) Chat tab: ask question, inspect citations and hero figures.
4) Repeat with new PDFs to keep your graph current.
5) Delete old graphs when your SSD starts side-eyeing you.

Time to Value

Minutes

Not quarters

Reality

Local compute

Expect warm laptop vibes