RAG and Ai-ready news data rebuilt, verified and constantly updated

RAGdoll   2025/08/05

RAGdoll

CONTENTS

RAGdoll

RAGdoll: The First Live News Feed for RAG

01

World's First Live News Feed

RAGdoll pioneers the first live news feed tailored for retrieval-augmented generation systems, ensuring seamless integration with AI technologies.

02

Structured and Copyright-Clean

Our platform delivers reconstructed news from verified open data, ensuring copyright-clean content that is safe for both private and commercial use.

03

Versatile Formats

RAGdoll provides news in structured formats such as JSON, CSV, and text, catering to diverse needs from analysis to LLM training.

RAGdoll

"....

Who We Are:

Part of Noah Analytics

RAGdoll is a specialist data reconstruction company supported by Noah Analytics, a trusted name in verified news reconstruction.

Cambridge Origins

Founded in Cambridge, Noah Analytics has developed large-scale verified news reconstruction systems used by national and B2B publication brands.

RAGdoll

Mission: Clean Data for AI Teams

· · ·

High-Quality Data

RAGdoll provides structured, verified, and copyright-clean news to eliminate reliance on paywalled or legally ambiguous sources.
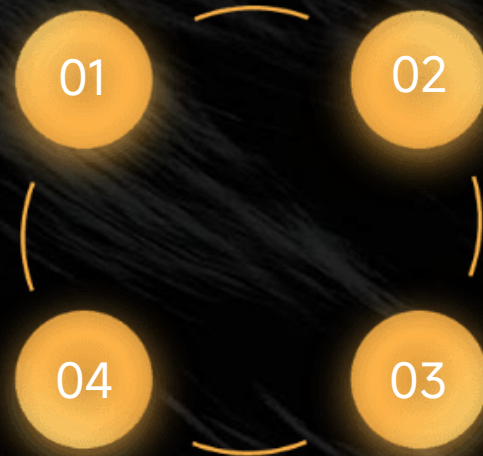
Improved Explainability

Our mission includes enhancing explainability, ensuring transparent and accurate RAG performance.

01   02

04   03

Support for AI Systems

Our data is designed to support data engineers, analysts, AI developers, and corporate intelligence teams.

Safe for Enterprise Use

RAGdoll ensures legal safety, providing data that is defensible and audit-ready for enterprise use.
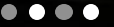
RAGdoll

Market Gap: Why Scraping Fails AI

Market Gap

Legacy news scraping provides unstructured, paywalled, and inconsistent data, failing to meet the needs of AI systems for reliable, legally safe, live data.

RAGdoll

What Makes RAGdoll Different

01

Original Content

RAGdoll reconstructs news using verified open facts, ensuring original content without paywalled wording.

02

Structured Outputs

Our platform delivers structured metadata, timestamps, and source links in various formats, ensuring enterprise-grade explainability and auditability.

RAGdoll

Reconstruction Pipeline Overview

Client Definition
01 Clients define verticals via schemas or keywords, and RAGdoll monitors global open sources to extract verifiable fact patterns.

Verification and Reconstruction
02 We cross-check timestamps, discard unverified items, and rebuild stories in original language with structured taxonomy.

Delivery Options
03 RAGdoll delivers machine-readable bundles to RAG stores, dashboards, or APIs in any language, ensuring seamless integration.

RAGdoll

Output Formats & Delivery Modes

Diverse Formats

RAGdoll delivers JSON for RAG ingestion, CSV for analysts, text for editors, and vectors for semantic retrieval.

Flexible Delivery

We offer webhook delivery, REST API access, S3 bucket delivery, and direct database push, ensuring compatibility with any technical environment.

RAGdoll

RAG-First Engineering Benefits

01

Structured Fields

RAGdoll's structured fields ensure compatibility with retrieval-augmented generation systems.

02

Clean Factual Grounding

Our clean factual grounding removes paywall noise, improving search relevance and traceability.

03

Timestamp Accuracy

Accurate timestamps enhance the reliability of RAGdoll's data, supporting time-aware models.

04

Language Neutrality

Language neutrality ensures RAGdoll's data can be used across different languages without compromising quality.

RAGdoll

Training & Fine-Tuning with Clean News

01

Safe Corpora

RAGdoll provides safe corpora for LLM fine-tuning, ensuring data is free from toxic or copyrighted text.

02

Domain Adaptation

Our consistently styled, timestamped entries support domain adaptation and classifier training.

03

Benchmarking

RAGdoll's verified entries enable benchmarking and construction of high-precision vertical corpora.

RAGdoll

Sector Use Cases Overview

Sector Use Cases

RAGdoll's reconstructed feeds outperform raw RSS across sectors like pharma, finance, energy, defence, insurance, and market due-diligence by delivering filtered, structured, legally safe updates.

RAGdoll

Historical Depth & Instant Backfills

01

Historical Reconstruction

RAGdoll can retrospectively rebuild years of news for any vertical within hours, leveraging billions of verified open-source records.
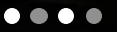
02

Immediate Analysis

This enables immediate time-series, sentiment, gap, and proxy analyses without waiting for slow manual archives.

RAGdoll

Value for Analysts & Publishers

Noise-Free Data

Analysts gain noise-free, defensible data with explainable lineage, accelerating research cycles and enriching dashboards.

Improved Visibility

Desk analysts and strategists gain improved visibility with RAGdoll's high-quality data.

Reduced Costs

Publishers receive ready-to-print original copy and structured data that reduce newsroom costs while launching niche verticals at scale.

Enhanced Modeling

RAGdoll's data supports enhanced modeling and better dashboards for analysts and publishers.

RAGdoll

Corporate & Consultancy Advantages

01

02

03

Live Feeds

Consultancies and risk units access live, traceable, sector-precise feeds for dashboards and due-diligence reports.

On-Demand Intelligence

RAGdoll replaces slow manual gathering with on-demand reconstructed intelligence that withstands audit scrutiny.

Corporate Use

M&A and strategy teams benefit from RAGdoll's live data for dashboards and internal tools.
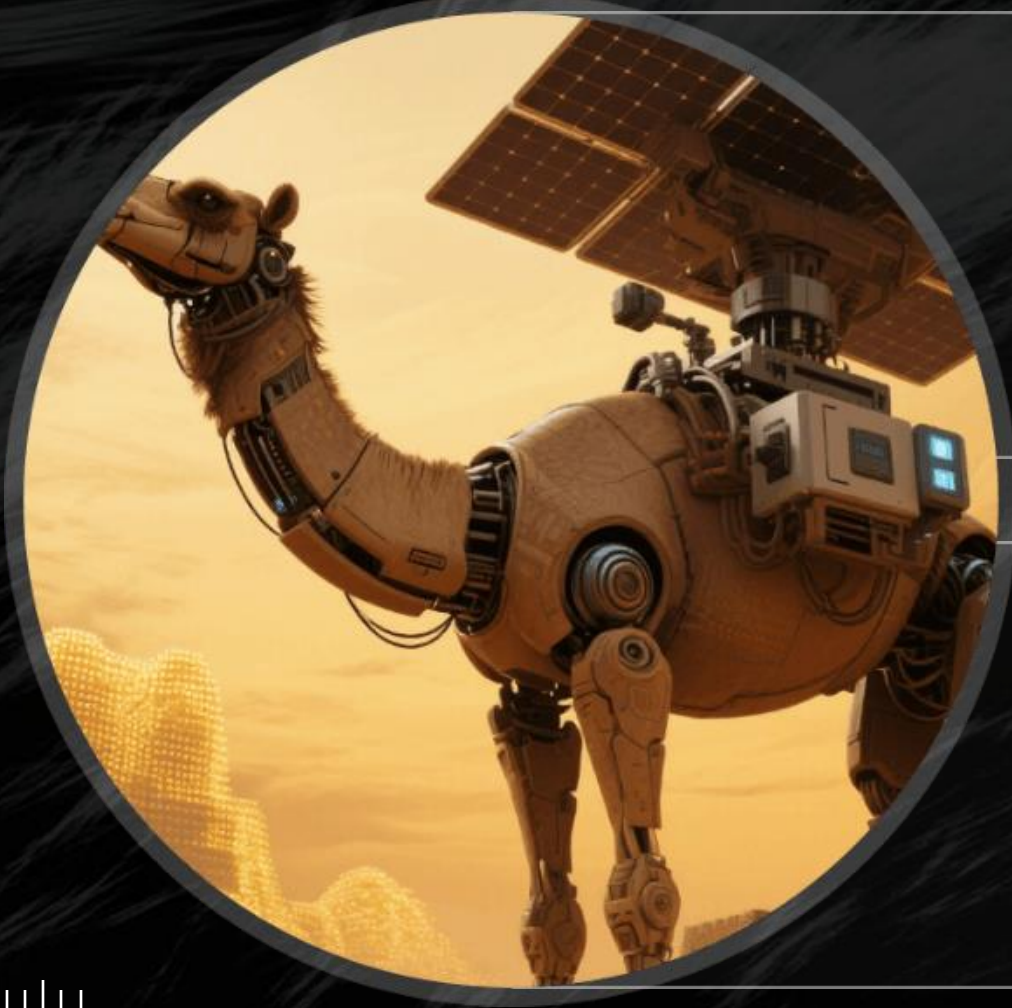
RAGdoll

Transparent Token Pricing Model

Token Pricing

RAGdoll sells content in £800 bundles containing 10 million tokens, yielding roughly 27,000 items. Clients may pause anytime, run unlimited users or verticals within allowance, and receive free image generation plus optional style guides in any language or length.

RAGdoll

Legal Safety & Copyright Integrity

Verified Open-Data

RAGdoll's verified open-data methodology avoids paywalled text, ensuring compliance and auditability.

Enterprise-Grade Safety

Our platform provides enterprise-grade legal defensibility for both commercial and internal AI deployments.

RAGdoll

Higher Value & Smarter Decisions

Faster Decisions

Our data empowers clients to build search systems, automated pipelines, risk reports, and sector timelines.

Enhanced Insight Engines

Our platform powers market insight engines, providing actionable intelligence for various sectors.

02

04

01

Accurate Intelligence

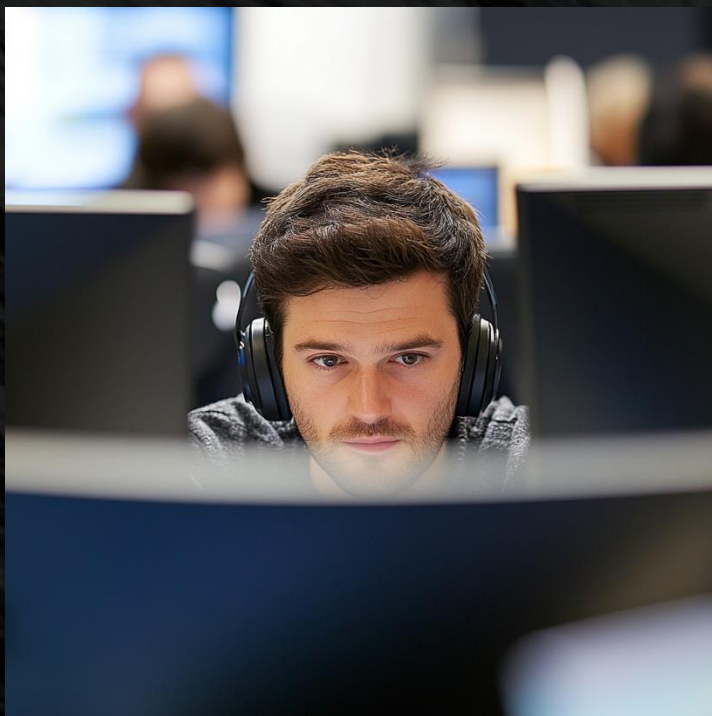RAGdoll delivers accurate, flexible, copyright-safe intelligence layered with deep historical data.

03

Evidence-Based Insights

RAGdoll supports faster, evidence-based decision-making with reliable data.

RAGdoll

Contact RAGdoll Today

Contact Information

RAGdoll, part of Worldwide AI Media Ltd,

Montford House, 26 Victoria Park Square, London E2 9PG.

Email: contact@noahwire.com

Phone: +44 7776 257777.

Reach out to start your copyright-clean news stream.

RAGdoll