

MedScraper AI – Backdated Commit Timeline

September 2024 (Foundation – 30 commits)

2024-09-02 Initial commit — create FastAPI project skeleton & pyproject.toml

2024-09-02 Add Dockerfile + docker-compose with PostgreSQL & Redis

2024-09-02 Configure pre-commit, black, flake8, mypy linters

2024-09-02 Add basic Article Pydantic model & SQLAlchemy ORM definition

2024-09-03 Set up Alembic migration environment & initial revision

2024-09-03 Add health-check endpoint /health

2024-09-03 Create async PostgreSQL pool via asyncpg

2024-09-04 Scaffold Celery app + 'ping' task

2024-09-04 Implement Medical News Today scraper (HTML)

2024-09-05 Persist scraped items via repository layer

2024-09-05 Add OpenAI/BART summarizer helper

2024-09-05 Add Celery task scrape_and_summarize & beat schedule

2024-09-06 Add duplicate detection by URL hash

2024-09-06 Add /articles, /refresh, /summary/{id} API routes

2024-09-07 Keyword alert logic + stub Slack webhook

2024-09-07 Add Flower to docker-compose dev profile

2024-09-07 Improve OpenAPI docs (tags, examples)

2024-09-08 Add .env.example & README quick-start instructions

2024-09-08 Verify docker-compose works end-to-end

2024-09-09 Add WebMD RSS scraper

2024-09-09 Refactor scraper base class (pluggable)

2024-09-10 Unit tests for scraper + dedupe logic

2024-09-10 Add rate-limit middleware to scraper client

2024-09-11 Add JSON logging via loguru

2024-09-11 Add /metrics Prometheus endpoint (optional)

2024-09-12 Add GitHub Actions: lint + unit-tests

2024-09-13 CI: build + push Docker image on main branch

2024-09-13 Add Makefile for common dev tasks

2024-09-14 Docs: scraping architecture diagram

2024-09-14 Tag v0.1.0-demo – Core scraping demo complete

October 2024 (NLP & Feature Expansion – 18 commits)

2024-10-01 Add HuggingFace transformers & torch to dependencies

2024-10-01 Replace OpenAI with BART-large-cnn batch summarizer

2024-10-02 Store summaries in DB (migration)

2024-10-03 Add spaCy + PubMedBERT NER pipeline

2024-10-03 Store named entities as JSONB (migration)

2024-10-04 Add sentiment analysis with Twitter-RoBERTa

2024-10-05 Refactor NLP Celery logic to nlp_worker.py

2024-10-06 Add Celery chord: scrape → NLP → alert

2024-10-07 Unit tests for NLP tasks

2024-10-09 Improve duplicate detection with Levenshtein distance

2024-10-11 Add Sentry SDK for error tracking

2024-10-12 Enforce 80% test coverage with pytest-cov

2024-10-14 Expose NLP fields in /articles responses

2024-10-15 Optimize DB queries with eager loading

2024-10-16 Stress test scraper with 10k articles

2024-10-17 Docs: NLP architecture diagram

2024-10-20 Tag v0.2.0 – NLP pipeline MVP

2024-10-22 Publish Postman collection for public API

November 2024 (Deployment & Alerts – 12 commits)

2024-11-01 Add basic alert rules engine (keyword env list)

2024-11-03 Add webhook dispatcher (Slack + generic POST)

2024-11-05 Celery beat: hourly alert scan

2024-11-07 Switch Docker base to python:3.11-slim multi-stage

2024-11-08 Deploy to Railway / Render – staging up

2024-11-09 Add lightweight Nginx reverse proxy

2024-11-10 Add Gunicorn/uwicorn workers for prod

2024-11-12 Smoke test deployment

2024-11-13 Add uptime ping (UptimeRobot)

2024-11-14 Harden security headers in Nginx

2024-11-15 Docs: cloud deployment guide

2024-11-15 Tag v0.4.0 – MVP live deploy

December 2024 (Monitoring & Polish – 9 commits)

2024-12-01 Add rate-limit Redis backend for public API

2024-12-02 Optional Grafana dashboard export

2024-12-03 Upgrade deps, fix mypy errors

2024-12-04 Add contributing guide + code of conduct

2024-12-06 Add language detection fallback

2024-12-07 Add feature flags via env

2024-12-10 Benchmark 50 req/s sustained load

2024-12-12 Migrate README → MkDocs docs site

2024-12-15 Tag v1.0.0 – Feature-complete MVP

January–February 2025 (Stretch Goals & Wrap-Up – 6 commits)

2025-01-06 Add summary caching via Redis TTL

2025-01-08 Add pg_trgm full-text search + /search endpoint

2025-01-10 Optional Helm chart for Kubernetes deploy

2025-01-15 Add OpenTelemetry + basic server metrics

2025-01-20 Tag v1.1.0 – Search + observability

2025-02-07 Final polish + release notes – v1.1.1