



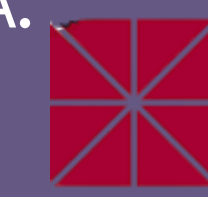
Contact us:



Gaurab Subedi¹, Jorge Fonseca¹, Amanda Leisgang Osse^{2,3}, Andrew Hooyman⁴, Jefferson Kinney^{2,3}, Jeffrey Cummings, MD, ScD^{2,3}.

¹Department of Computer Science, University of Nevada, Las Vegas, Las Vegas, USA, ²Kirk Kerkorian School of Medicine, University of Nevada, Las Vegas, Las Vegas, NV, USA, ³Chambers-Grundy Center for Transformative Neuroscience, Las Vegas, NV, USA, ⁴Physical Therapy, Crean College of Health and Behavioral Sciences, Chapman University, Orange, CA, USA.

Funded by the Alzheimer's Drug Discovery Foundation (ADDF).



CHAPMAN UNIVERSITY



Alzheimer's Drug Discovery Foundation

1 Background & Objective

Why this matters

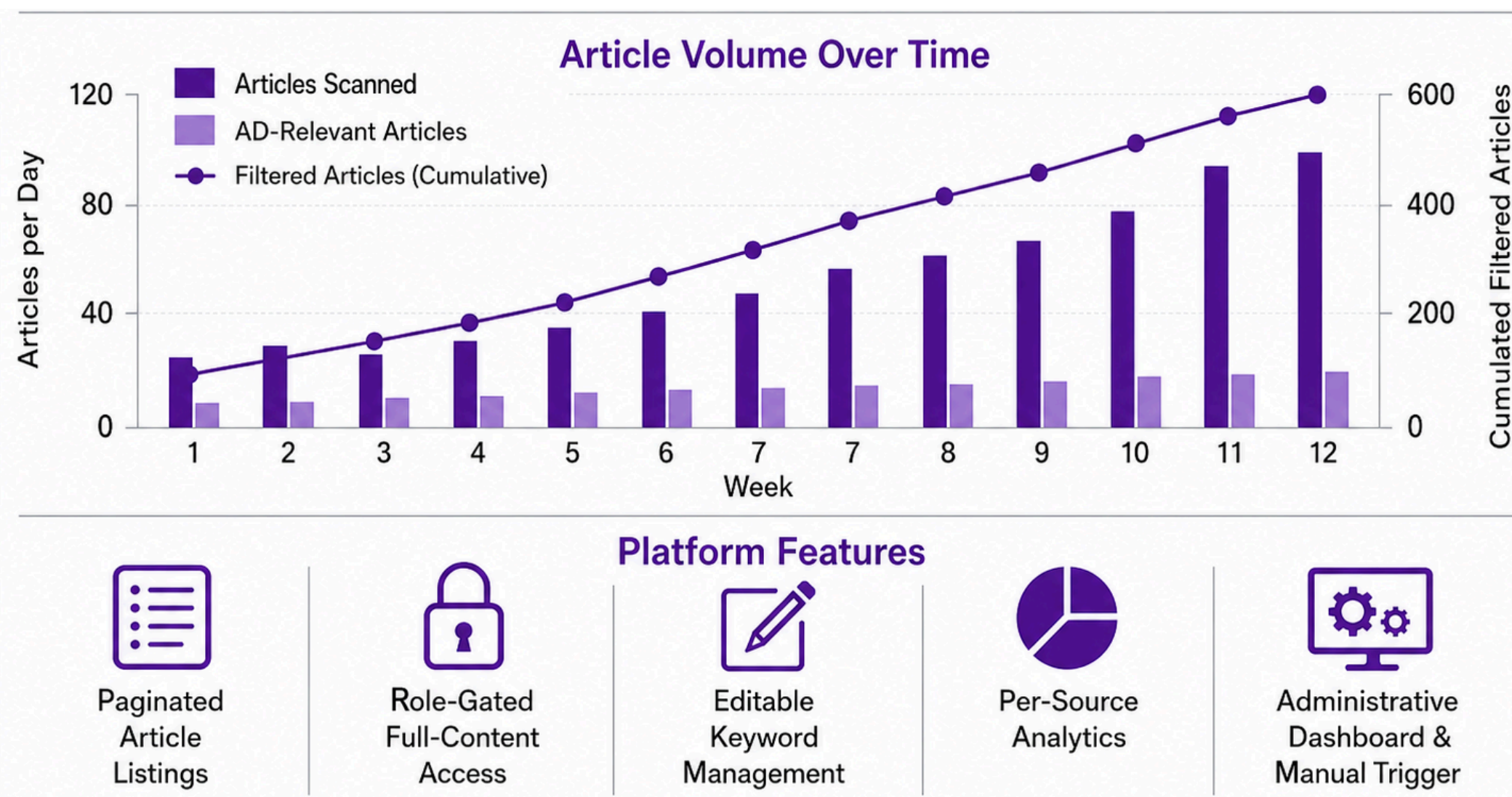
- Alzheimer's disease research updates are distributed across government agencies, pharmaceutical companies, academic medical centers, clinical trial sources, and biomedical news platforms.
- These updates are important for tracking emerging therapies, biomarker advances, regulatory decisions, and clinical trial activity.
- Manual monitoring is slow, fragmented, and prone to missed or irrelevant information.

Objective:

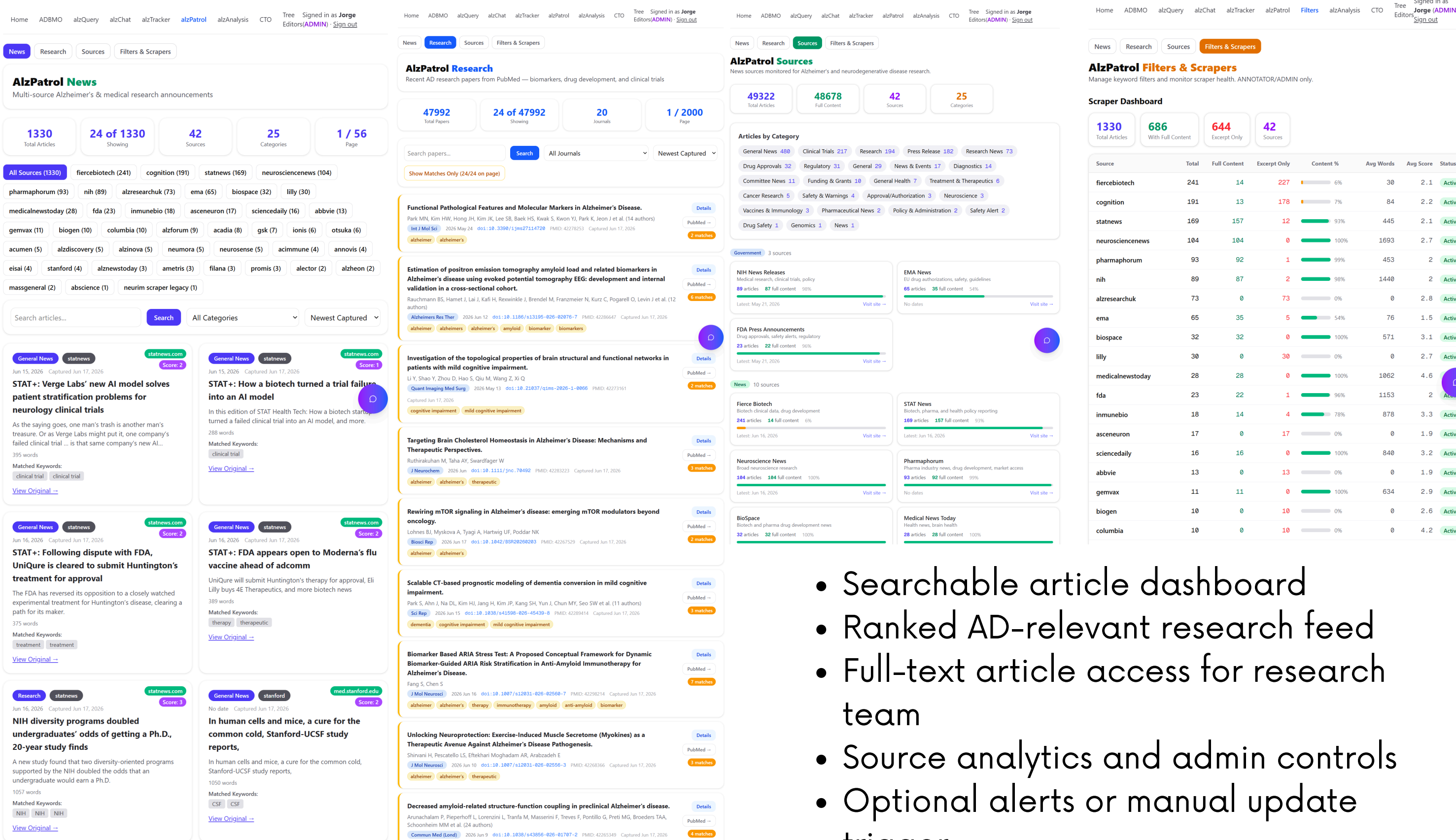
- Develop an automated information surveillance system that continuously aggregates, filters, classifies, and organizes Alzheimer's disease research updates into a centralized searchable platform.

3 Results

- 51 connectors across 4 AD research source types.
- 80–100 articles scanned daily from 35 active sources.
- 562 AD-relevant articles collected, including 189 full-text articles.
- <20 ms API response with search, analytics, and admin tools.



4 Platform Output



- Searchable article dashboard
- Ranked AD-relevant research feed
- Full-text article access for research team
- Source analytics and admin controls
- Optional alerts or manual update trigger

2 System Architecture

The platform uses a modular scraper architecture with 51 automated source connectors across four major source categories:

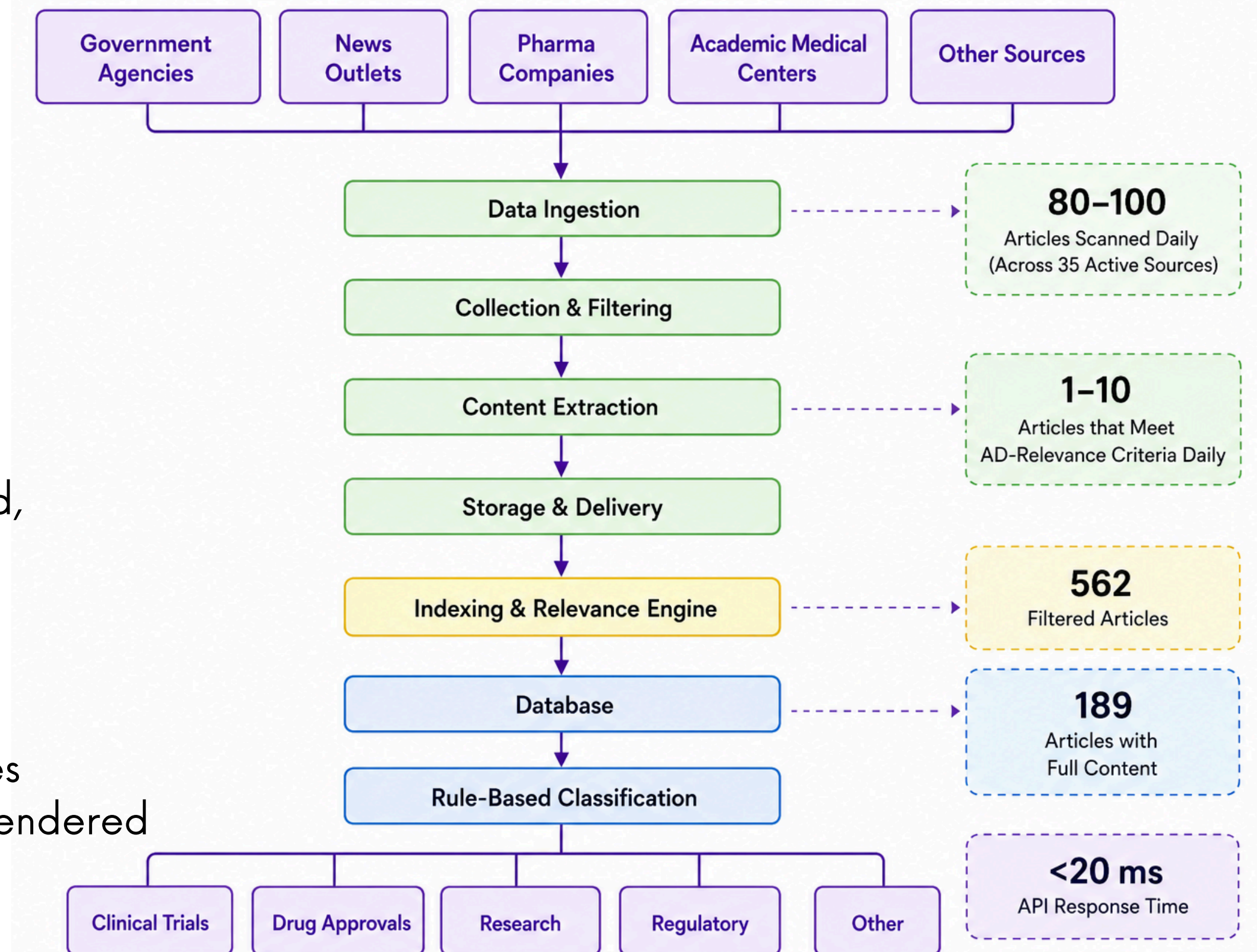
- Government agencies: NIH, FDA, EMA
- Biomedical news outlets: ScienceDaily, Fierce Biotech, Neuroscience News
- Pharmaceutical companies: Biogen, Eli Lilly, Eisai, Novartis
- Academic medical centers: Columbia, Stanford, Mass General, Brigham and Women's

Collection methods

- RSS feed parsing: 27 sources
- HTML extraction using HTTP requests: 18 sources
- Headless browser automation for JavaScript-rendered sites: 6 sources

Processing workflow

Data Sources → Collection → 58-Keyword AD Relevance Filter → Full-Text Extraction → SQLite Storage → Rule-Based Classification → Searchable Dashboard



Articles are categorized into clinically meaningful domains: clinical trials, drug approvals, research, regulatory, and other.

5 Conclusion

- Developed an automated AD research surveillance platform across regulatory, industry, academic, and news sources.
- The system reduces manual monitoring by collecting, filtering, classifying, and organizing AD-relevant updates.
- The platform supports faster awareness of emerging therapies, clinical trials, biomarkers, and regulatory decisions.

6 Future Work

- Expand source coverage by adding additional regulatory, industry, academic, and biomedical news sources.
- Extend the surveillance framework to Parkinson's disease and other neurodegenerative diseases.
- Add automated email alerts when new AD-relevant research updates, clinical trial news, biomarker findings, or regulatory information are detected.
- Improve relevance scoring and classification using semantic ranking and researcher feedback.

