Groundsource: Google Turns 25 Years of News into 2.6 Million Flood Records with Gemini

Kabui, Charles

2026-03-19

Google Research released Groundsource, a framework that uses Gemini to convert unstructured news articles into structured disaster data. The problem it solves is simple: accurate flood forecasting needs historical records, but existing databases like GDACS contain roughly 10,000 entries, nowhere near enough for global-scale AI. Groundsource analyzes news reports in 80 languages, translates them via Cloud Translation API, then runs Gemini through a strict

verification process. The model classifies whether an article describes an actual flood (not a warning or policy discussion), resolves relative time references like "last Tuesday" against publication dates, and maps locations down to neighborhoods using Google Maps Platform. The output: 2.6 million historical flood events across 150+ countries from 2000 to present. That is a 260x expansion over existing archives. Validation shows 82% of extracted events are accurate enough for real-world analysis, and the dataset captures 85-100% of severe floods that GDACS recorded between 2020 and 2026.

This data is already in production. Google's Flood Hub now uses it to deliver near-global urban flash flood forecasts up to 24 hours in advance. For communities in flood-prone regions, 24 hours of warning can mean the difference between evacuation and catastrophe. The full dataset is openly released, so any researcher can use it to build better hydrological models without needing proprietary infrastructure.

The deeper insight here is methodological. The world's news media collectively functions as a massive, multilingual sensor network for natural disasters, but the signal has always been locked in unstructured text. Google notes this same approach could extend to droughts, landslides, and avalanches. If it generalizes, the bottleneck for disaster forecasting shifts from data collection to model quality, and that is a much easier problem to solve.

Sources:

- Google Research Blog: Introducing Groundsource
- Groundsource Paper
- Flood Events Dataset (Zenodo)
- Google Flood Hub
- Flash Flood Forecasting Blog

Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. Read more: /terms-of-service