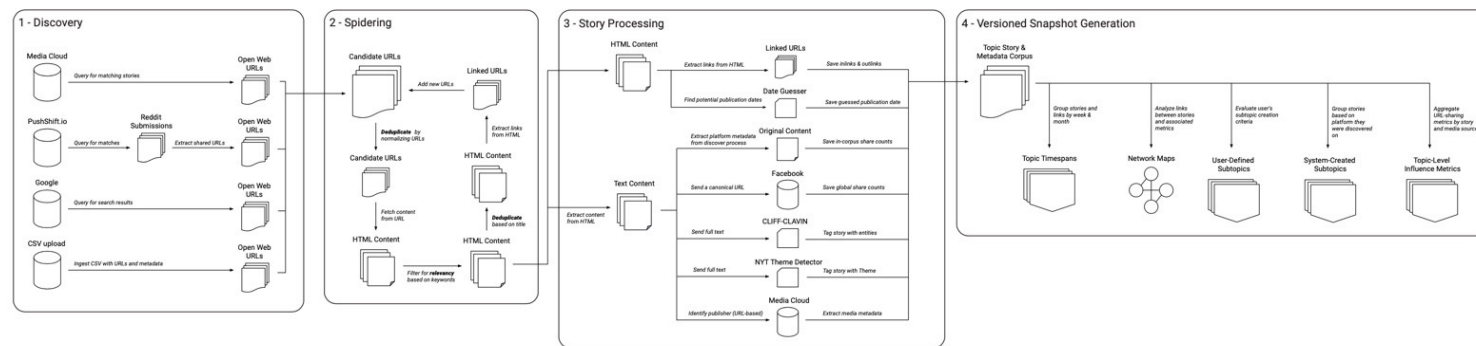


# Use the Media Cloud global news dataset to understand attention and influence online.

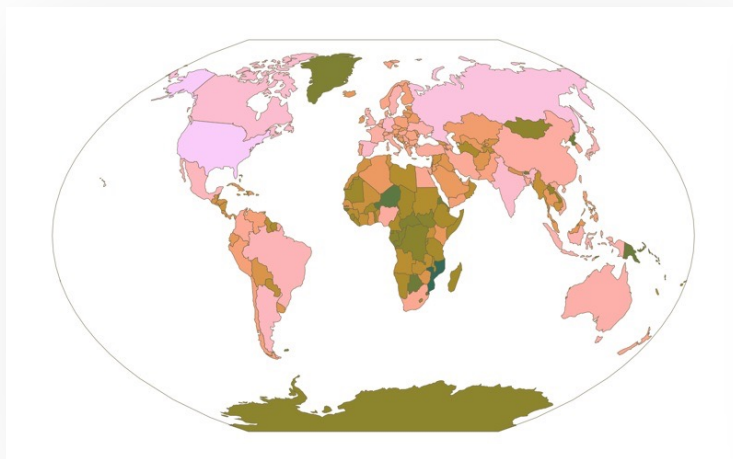
*Use the 1.8 billion stories in our system, or build your own corpus.*



Our ingest pipeline is built from **RSS feeds and spidering**. We **clean and deduplicate content** before adding **rich metadata** to support your media ecosystem research.



We have **global coverage in multiple languages**. Here's where the 231 million 2020 stories came from:



Use our API or online visual research support tools to investigate and acquire data. **All metadata is available for free.**



Roberts, Bhargava, Valiukas, Jen, Malik, Bishopm Ndulue, Dave, Clark, Etling, Faris, Shah, Rubinovitz, Hope, D'Ignazio, Bermejo, Benkler, Zuckerman