

5412 Peer Review

Introduction:

An intro would be helpful to understand the goal of this system. I think it's for fake news detection(?)

Data:

Data types look good and well defined. Everything looks like it has a reason to be there and I can't see anything that should be present that isn't.

Architecture:

I think you have well thought out tiers. They seem to make a lot of sense and I like the idea of caching to reduce load on the lower level tiers. One non-trivial part of this that I don't see an explanation for is how you're collecting news data. What news sources are you going to be working with? Ones with established APIs? Will you have a fleet of web scrapers that are crawling through these sites? A lot of unanswered questions on how you actually get the important data.

Evaluation:

I really like these evaluation metrics. They seem well thought out and exactly the types of things that a service like this would need to be good at to be effective in practice. Explanations of how you're going to simulate partitions would be nice but that's not necessary given the rubric.

Milestones:

Seem reasonable and doable. Good pacing.