

# TikaInHadoop

## Running Tika in Hadoop

On very rare occasions, Tika can fail catastrophically: infinite hang or out of memory errors. There may be other features of Tika that make it useful for developers to share notes on how to run Tika at scale. This page is intended to gather lessons learned and offer pointers for running Tika in the Hadoop framework.

## Useful Parameters

## Lessons Learned

## Links

- William Palmer's blog post on running Tika in Hadoop – [Tika to Ride](#)

## Frameworks

- Julien Nioche's [Behemoth](#)
- William Palmer's [Nanite](#)

## Hadoop-Adjacent Frameworks (might or might not use Hadoop, but the entry point is not pure Hadoop)

- [Apache OODT Crawler](#)