**ANALYSIS**

**License:**

The Apache 2.0 license is like the MIT license which is a permissive open-source license, which allows the use of software by individuals, academics, and for commercial purposes. It even allows us to share the modified and original code. The only thing that is required is to provide a copy of the license in our software and include the copyright details.   
  
Who uses Apache 2.0 license?

* Apache PDFBox
* Apache POI (Poor Obfuscation Implementation)

**Apache Tika:**

Apache Tika is an open-source content detection and analysis tool kit from the Apache Software Foundation.

**Usages**:

* File type detection (e.g., .docx, .pdf, .ppt, etc.)
* Text and Metadata content extraction

**Why Apache Tika?**

* Unified parser interface
* Low memory usage
* Language identification