

NiFi

NiFi - это ETL - ^(струминговый) инструмент, написанный на Java

XML - более жесткий, JSON - более гибкий;
у XML есть XML-схемы, а у JSON нет;
авро-файлы - JSON с JSON-схемами
(как в XML);

NiFi оперирует в основном с авро-файлами
(JSON и XML тоже);

Flow - файлы - промежуточные файлы
с данными

Процессор NiFi получает на вход Flow-файл;

Flow-файлы состоят из описания данных
(атрибутов) и самих данных (payload);

Единица работы - функция;

NiFi registry - GIT для ETL-процессов;