I am an Apache tika user.