osstextextractor

doc parser