

1. 개요


한글의 문서 파일은 사용자가 따로 지정하지 않는 한 .HWP를 기본 확장자로 가진다. 문서 파일에 저장되는 내용은, 실제 사용자가 입력한 문서의 내용과 문자 장식 정보뿐만 아니라 문서를 편집할 당시의 글꼴에 대한 정보, 조판에 영향을 주는 설정 사항(용지 종류, 여백 정보 등)도 포함된다.

한글 문서 파일 형식 5.0은 2000년 10월 이후에 출시된 한글 제품군(한글 워디안, 한글 2002, 한글 2005, 한글 2007, 한글 2010, 한글 2014, 한글 2018 등)에서 생성되며, 문서 버전에 따라 큰 골격은 유지되나, 추가적인 정보들에 의해 약간의 차이가 있다.

한글 문서 파일 형식 5.0은 파일의 크기를 최소화하기 위하여 압축 기능을 이용한다. 압축된 문서 파일도 기본적인 정보를 저장하는 부분은 압축되지 않으며, 실제 압축되는 부분은 사용자가 입력한 본문과 그림 관련 데이터 부분이다.

한글은 문서 파일의 압축에 zlib.org의 zlib을 사용했다. zlib은 웹상에 소스가 공개되어 있는 공개 소프트웨어이다. zlib은 zlib License를 따르며, 이는 소스의 자유로운 사용이 가능하며 해당 소스를 사용한 2차 산출물에 대한 소스 공개 의무가 없다. 자세한 사항은 zlib에 포함되어 있는 라이선스 문서 파일을 참조하기 바란다.

한글 문서 파일 형식 5.0의 구조는 윈도우즈의 복합 파일(Compound File)에 기초를 두며, 문자 코드는 ISO-10646 표준을 기반으로 한다. 대부분의 문자 정보는 유니코드(UTF-16LE) 형식으로 전달되고, 저장된다.

 Compound File에 대한 접근 방법은 OLE관련 자료 또는 MSDN을 참고
StgOpenStorage(), IStorage::Open(), ...