

1. 파일 시그니처란?

•정의: 파일 고유의 포맷으로 특정 byte로 이루어져 있으며,

파일의 내용을 해석하기 위해 포함하고 있는 고유값.

•Header와 Footer

-Header:파일의 처음에 있는 시그니처

-Footer:파일의 마지막에 있는 시그니처

File type	Header signature	Footer signature
PDF	25504446 (%PDF) 25 50 44 46 2D 31 2E (%PDF-1. 3)	25 25 45 4F 46(ASCII-%%EOF)
PNG	89 50 4E 47 0D 0A 1A 0A	49 45 4E 44 AE 42 60 82
JPEG/JPG	FF D8 FF E0(JFIF 표준) FF D8 FF E8(특수용도)	FF D9
GIF	47 49 46 38 37 61(GIF97a) 47 49 46 38 39 61(GIF89a)	00 3B
ZIP	50 4B 03 04	50 4B 05 06
MP3	49 44 33 03	
BMP	42 4D(ASCII-BM)	

2. 문자 코드 [아스키 코드/엡시딕 코드/유니코드] + Base64란?

•아스키 코드

- 문자 데이터를 7비트로 표현하는 대표적인 문자 인코딩 방식으로, 128개의 고유한 숫자 값으로 다양한 문자들을 표현한다.
- 컴퓨터 통신 장치 데이터 처리, 제어 문자 처리 등 사용

•엡시딕 코드

- IBM 컴퓨터 운영 체제에서 사용되는 8비트 문자 인코딩으로, 숫자, 알파벳, 특수문자, 제어문자 등 다양한 문자를 표현한다.
- IBM 중심의 대형 컴퓨팅, 금융기관 시스템 등 사용

•유니코드

- 전 세계의 문자를 컴퓨터에서 일관되게 나타내고 처리하기 위한 국제적인 표준 문자. 처리법을 말하며, 기본적으로 2 바이트로 문자를 표현한다.
- 다국어 지원, 표준화 장점으로 웹, 프로그래밍, 소프트웨어, 데이터 포맷 등 사용

•Base64

- 8비트로 이루어진 6비트 이진 데이터를 문자 코드에 영향을 받지 않는 공통 ASCII 영역의 문자들로 변환하는 인코딩 방식

3. 빅엔디안과 리틀엔디안

•빅엔디안

- 메모리에 데이터 저장->상위 바이트를 먼저 저장하는 것
- 낮은 주소의 데이터에 높은 바이트부터 저장하는 방식으로, 선형방식과 유사하여 메모리에 저장된 데이터를 순서대로 읽기 쉬워 데이터 해석이 용이하다.

0x1000	0x12
0x1001	0x34
0x1002	0x56

•리틀엔디안

- 메모리에 데이터 저장->하위 바이트를 먼저 저장하는 것
- 낮은 주소의 데이터에 낮은 바이트부터 저장하는 방식으로, 선형방식의 반대로 데이터를 읽어야 하며, 계산, 현대 주요 CPU에서 용이하다.

0x1000	0x56
0x1001	0x34
0x1002	0x12