

데이터 문해력과 디지털 글쓰기

오픈소스 소프트웨어: 빛에듀(Bit-Eud)

이광춘 / 한국 R 사용자회

11/4/22

목차

서문	3
I 글쓰기	4
1 디지털 글쓰기	5
1.1 디지털 문서 제작 패러다임	5
1.2 문서로 보는 민주주의	6
1.3 디지털 문서 저작	6
1.4 과학문서 저작	7
1.5 사례: 한국 R 컨퍼런스(2021)	7
2 문자와 종이 역사	8
2.1 언어와 문자	8
2.2 매체 역사	8
2.3 문자	10
2.3.1 알파벳	10
2.3.2 문자 시스템	11
2.3.3 글쓰기 방향	11
2.3.4 표의 문자	11
2.4 저작	12
II 디지털 글쓰기	14
서문	15
III 빛에듀(Bit-Edu)	16
3 Summary	17
참고문헌	18

서문

Part I

글쓰기

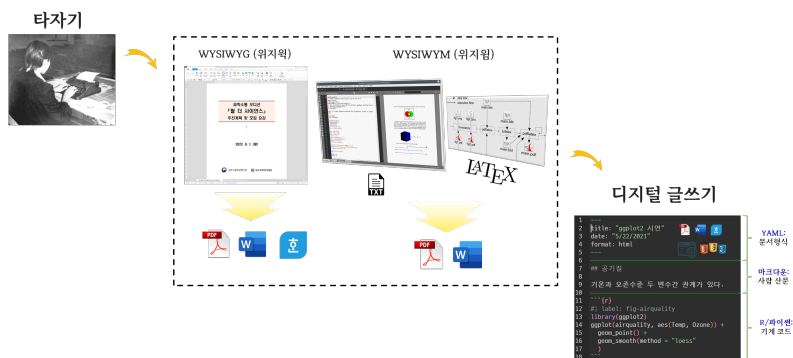
1 디지털 글쓰기

1.1 디지털 문서 제작 패러다임

신속하고 빠르게 누구나 짧은 학습을 통해서 문서를 저작하고 출판할 수 있는 방식으로 인기를 얻은 아래한글 혹은 MS워드 워드프로세서를 사용하는데 이는 위지위그(WYSIWYG: What You See Is What You Get, “보는 대로 얻는다”)에 기초한 것으로 문서 편집 과정에서 화면에 포맷된 낱말, 문장이 출력물과 동일하게 나오는 방식이다. 이전 타자기 대비 경쟁력이 있었고 이를 바탕으로 문서편집기 소프트웨어가 대세로 떠올랐다.

위지위그의 대척점에 있는 것이 위지위임(WYSIWYM, What You See Is What You Mean)으로 대표적인 것인 *LaTeX* 으로 구조화된 방식으로 문서를 작성하면 컴파일을 통해서 최종 문서가 미려한 출판가능한 PDF, PS, DVI 등 확장자를 갖는 출판결과물을 얻을 수 있다.

*LaTeX*이 갖는 장점은 이제 프로그래밍 코드(R / 파이썬 / SQL / 셸스크립트 / 자바스크립트 등)을 문서저작에 포함되며 **literate programming** 패러다임이 대세로 떠올랐고 현존하는 다양한 문서저작 문제에 대한 응답으로 자리를 매김하고 있다.¹



¹[출처: <http://www.cbinews.co.kr/news/articleView.html?idxno=36859>]

1.2 문서로 보는 민주주의

대한민국은 권력을 입법권, 행정권, 사법권으로 나눠 권력을 나누었다. 세부적으로 보면 다소 차이는 있겠지만 법을 만드는 것은 결국 문서를 저작하는 것이고 대한민국 정부는 법에 담긴 문서를 실행하는 기관이고 사법부는 법 문서를 해석하는 기관으로 볼 수 있다. 결국 민주주의는 문서를 기반으로 존립하는 사회체제라고 볼 수 있다. 최근 들어 인공지능이 부상하면서 문서저작을 보조하던 수단에서 이제는 사람보다 더 나은 통찰력을 보이는 분야도 속속 등장하고 있다. 이런 측면에서 인공지능 기계를 현 민주주의 체계 내에 어떻게 녹여낼 것인지 관심이 높아지고 있다.



1.3 디지털 문서 저작

과거 둘로 쪼개져서 과학기술 문서저작 즉 논문 등은 *LaTeX* 기반 위지윅(WYSIWYM) 패러다임이 확고한 자리를 잡았지만 위지위그(WYSIWYG) 패러다임 기반 아래한글 워드와 같은 문서저작편집기 성능이 높아지면서 사실상 둘 사이 경계도 허물어졌다.

하지만, 과학기술 논문 저작에 진실성, 재현성 등이 강조되고 빅데이터 기반 과학연구가 큰 성과를 보이고 프로그래밍이 과학기술 논문 저작에 주요한 역할을 하면서 이런 추세를 반영한 과학기술 저작 패러다임이 떠오르고 있다.

이러한 추세를 반영한 도구의 발전도 함께 이뤄지면서 과학기술 저작은 물론 거의 모든 문서 저작분야에 큰 변화가 일어나고 있다.

1.4 과학문서 저작

- 심각한 현재상황
- 마크다운 기초
- 고급 마크다운
- 문서저작 실무
- 수학 수식

1.5 사례: 한국 R 컨퍼런스(2021)

- 컨퍼런스 웹사이트
- 컨퍼런스 전 기획
 - 컨퍼런스 포스터
 - 컨퍼런스 후원
- 컨퍼런스 본 행사
 - 발표 슬라이드
 - 발표 템플릿
- 컨퍼런스 후 설문
 - 설문조사 보고서
- 기타
 - 탄소중립 신재생 에너지 캠프

2 문자와 종이 역사

2.1 언어와 문자

전세계인구는 77.5억(2020년 기준)명이며 언어는 대략 7,100개 언어가 존재하고 있는 것으로 파악되고 있으며 영어, 중국어, 힌디어, 스페인어, 프랑스어, 아랍어, 벵골어, 러시아어, 포르투갈어, 인도네시아어 순으로 많은 언어인구를 갖고 있다.

한국어 언어는 한글을 문자로 갖고 있지만 모든 언어가 문자(Writing System)를 갖는 것은 아니다. 다르게 보면 동일한 뿌리를 갖는 문자를 공유하는 언어도 많다. 예를 들어,

- 알파벳: 라틴어(영어, 불어, 독어, 이탈리아어 등), 켈틱, 발틱, 슬로박, 등
- 한자: 중국어, 대만어
- 아랍문자: 아랍, 인도-이란 어족(페르시아, 펀잡 등)
- 가나: 일본
- ...

문자가 존재하기 위한 전제조건이 있는데 크게 다음 세가지를 들고 있다.

- 언어 : 한국어, 영어, 중국어, ...
- 아이디어: 생각, 콘텐츠, ...
- 기호: 숫자, 문자, 그림, ...
 - 언어
 - 숫자와 수식
 - 교통표지 등 표식
- 매체: 종이, 파피루스, 디스플레이

2.2 매체 역사

메시지를 전달하기 위해서 의미를 담을 수 있는 상형문자가 전세계 동시 다발적으로 개발되었으나 동시에 추상적인 개념

(예를 들어, 숫자)을 전달하기 위한 문자체계도 필요하여 개발됨과 동시에 문자를 담을 수 있는 매체에 대한 발전도 함께 이뤄졌다.

초기 동굴벽화처럼 벽면에 그림을 그려서 의사를 전달하는 체계가 출현한 것은 기원전 16,500년 라스코 동굴벽화를 비롯하여 반구암각화에도 그 흔적을 어렵지 않게 찾을 수 있다. 기원전 3,000년 수메르 문명이 출현하면서 수십만의 사람들이 도시에 모여살게 되면서 세금 징수, 계약 등 산적인 정치, 사회, 경제 문제를 해결하는데 필수적으로 문자가 필요하게 되었다. 동시에 세금과 계약관계를 기록하고 보존할 수 있는 방식으로 점토를 사용하였고, 이집트에서도 유사한 목적으로 점토판 대신에 파피루스를 종이처럼 사용하는 방식을 사용했다.



그림 2.1: 문자 매체의 진화

문자를 담을 수 있는 매체도 생산과 유지관리에 비용이 많이 들고, 상형문자가 담을 수 있는 정보에 한계가 보임에 따라 지중해에서 수많은 물품을 거래하던 페니키아인들은 기존 상형문자체계를 벗어나 소리에 의미를 담을 수 있는 방식으로 진화시켜 무한대에 가까운 의미를 담을 수 있는 문자를 만들어냈는데 기원전 1,200년 페니키아 문자를 창안하게 되었다. 이후, 그리스 헬레니즘 문화를 거치면서 줄리어스 시저 시대 라틴 알파벳(대문자)이 공표되면서 오늘에 이르게 되었다. 라틴 알파벳을 담을 수 있는 매체로 양피지(Parchment)가 사용되었고 필경사(Scribe)가 한땀 한땀 손으로 필사를 했다.

동굴벽, 점토판, 파피루스 종이, 양피지를 거쳐 드디어 106년 채륄(Cai Lun)이 종이를 발명하면서 문자문화에 큰 변화가 생기게 되었다. 이후 서기 800년경 목판인쇄술이 발명되고 1450년 구텐베르크 인쇄술을 통해 성경이 대량 인쇄되어 르네상스의 초석이 되었으며, 푸어드니어(Fourdrinier) 형제가 1806년 종이를 대량생산하게 되는 초지기(Paper Machine)를 발명하면서 저렴한 비용으로 일반대중도 아이디어를 문자로 종이에 기록하여 널리 알릴 수 있는 토대가 완성되었다. 여기에

더해 1922년 독일에서 ISO 216을 통해 종이에 대한 국제 표준이 자리잡으면서 미국을 비롯한 일부 국가를 제외하고 전세계적으로 문자를 담아낸 매체에 대한 표준이 확립되었다.

2.3 문자

언어를 기록하기 위한 상징체계로 문자(□□, writing system)가 발명되었다. 문자는 언어중에서도 음성언어를 기록하기 위해 생겨난 것으로 알려져 있다. 문자는 정보의 저장과 전달에 신뢰성이 담보된다는 점에서 그 유용성이 크다.

문자는 크게 표음문자(Phonographic)와 표의문자(Logographic)로 나뉜다. 표의문자의 대표적인 문자로 중국 한문과 이집트 상형문자가 꼽힌다. 예를 들어 □(나무 목)은 목이라는 소리를 내지만 뜻은 나무를 갖고 있다. 표음문자는 한글을 비롯한 라틴계열 대부분의 언어가 여기에 포함된다.

2.3.1 알파벳

알파벳(Alphabet) 문자는 이집트 상형문자에서 시작되었다고 알려져 있다. 이후 상형문자는 페니키아 상인이 소리를 기반으로 의미를 전달하는 방식으로 진화되었고 그리스를 거쳐 알파벳 뒤쪽 문자가 추가되고 로마시대에 오늘날 알려진 영어 대문자가 완성되고 이후 소문자가 손으로 필사되면서 발명되고 다양한 글꼴이 생겨나게 되었다. 구텐베르그가 활자인쇄술을 발명되면서 현재와 같은 알파벳 문자가 완성되었다.

Evolution of the Alphabet | Earliest Forms to Modern Latin Script [유튜브](#)

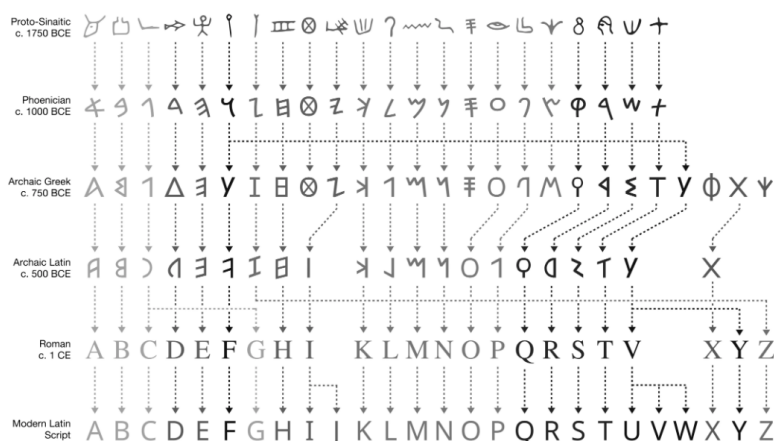
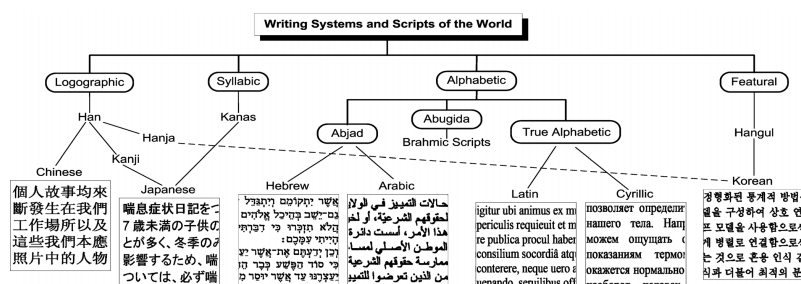


그림 2.2: 알파벳 진화 역사

2.3.2 문자 시스템

앞서 문자를 크게 표음문자와 표의문자로 구분하였다. 이를 좀더 상세히 세분화하면 전세계 문자는 크게 다음과 같이 4개로 구분된다. (Ghosh et al., 2010) 표어문자의 대표적인 문자가 중국의 한자, 음절문자의 대표적 문자가 일본어, 자질문자의 대표 문자가 한글이다. 그외 음소문자는 로마자로 대표되는 영문자가 있다.

- 표어문자(logographic)
- 음절문자(syllabic)
- 음소문자(alphabetic)
- 자질문자(featural)



종류	기호 표현	예
표어 문자	형태소	중국의 한자
음절 문자	음절 또는 모라	일본의 가나 (문자)
음소 문자	낱소리 (자음 또는 모음)	로마자, 키릴문자, 그리스 문자
아부기다	낱소리 (자음+모음)	데바나가리 문자, 그으즈 문자
아브자드	낱소리 (자음)	아랍 문자, 히브리 문자
자질 문자	음성 자질	한국의 한글

2.3.3 글쓰기 방향

대부분의 문자를 왼쪽에서 오른쪽, 위에서 아래 방향으로 작성하는 방식을 취하고 있으나 이슬람권에서는 오른쪽에서 왼쪽으로 작성을 하고 있고 몽골은 위에서 아래로 작성하고 있다.

2.3.4 표의 문자

한자와 같은 표의 문자가 여러가지 단점이 있는 것처럼 보이기도 하지만 표의문자가 갖는 장점으로 인해 그 중요성이 더욱 부각되고

있다. 표의문자가 갖는 장점은 문자 자체 의미가 부여되기 때문에 전세계 누구도 문자만 보면 그 뜻을 유추할 수 있다는 점이다.

- 그림문자(Pictographic): □, □, □, ...
- 표의문자(Ideographic): □ 금지와 같은 의미를 갖는 문자로 □ □ 교통 표지판, ► 전자제품 사용
- 추상 어표(Abstract Logographic): 한자 사람 인(人), %, &, ...

그림문자와 표의문자를 결합하게 되면 다음과 같이 흡연을 금지하는 의미를 명확히 전달할 수 있다.



그림 2.3: 흡연과 금지가 결합된 표의문자

2.4 저작

점토판에 썰기를 사용해서 문자를 활용하여 갖고 있던 아이디어(세금, 계약, 신화 등)를 작성하던 방식이 중세시대 종이가 발명되며 동양에서는 붓과 머루, 서양에서는 잉크와 새깃털펜으로 저작을 하였다. 이후 타자기가 발명되어 개인 저작이 가능하게 되었으며 이후 전동타자기도 등장하여 아이디어를 문서로 남기는 것이 훨씬 비용도 저렴해지고 품질도 좋아지고 출판까지 과정도 신속해졌다. 제2차 세계대전이 남긴 에니악 컴퓨터는 민수로 전환되어 IBM에서 1964년 워드프레스서를 시장에 출시하였고 1969년 IBM에서 저장장치도 출시되어 기계식 저작방식이 디지털 저작방식으로 넘어가는 기반이 다져졌다. 1980년대 전후로 컴퓨터를 기반으로 두가지 저작방식 즉, 마이크로소프트 워드 1.0 혹은 아래한글과 같은 위지윅(WYSWHG) 방식과 문서를 구조적으로 작성하여 컴파일하는 위지윅(WYSIWYM) 방식이 일반화되면서 현재에 이르고 있다.



Part II

디지털 글쓰기

서문

Part III

빛에듀(Bit-Edu)

3 Summary

In summary, this book has no content whatsoever.

1 + 1

[1] 2

참고문헌

Ghosh, D., Dube, T., & Adamane, S. (2010). Script recognition—a review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32, 2142–2161.
<https://doi.org/10.1109/TPAMI.2010.30>