

# 2023학년도 2학기 언어데이터과학

## 제3강 Linux 기초

박수민

서울대학교 인문대학 언어학과

2023년 9월 11일 월요일

## 지난 시간에 배운 것

- 1 git 레포지토리 만들기

## 오늘의 목표

- Linux 명령어와 정규표현식을 사용하여 고대 로마 요리에 쓰인 재료를 찾을 수 있다.

## 오늘의 도구

- VS Code > Terminal

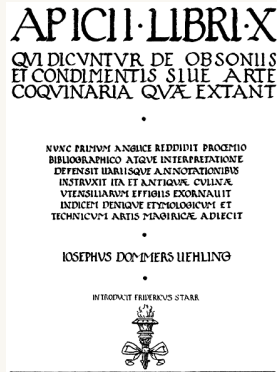
## Apicius, Cooking and Dining in Imperial Rome

- De re coquinaria(5th century?)의 영문 번역
- 고대 로마의 요리책
- <https://www.gutenberg.org/ebooks/29728>

Source: Project Gutenberg

저작권이 만료된 고전을 디지털화하여 공개하는 프로젝트

- <https://www.gutenberg.org/>





<https://www.flickr.com/photos/carolemage/23852992435/>

<https://www.gutenberg.org/ebooks/29728>  
[Plain Text UTF-8] 파일을 저장한다.



Project Gutenberg

About ▾ Search and Browse ▾ Help ▾

Quick search   [Donation](#) [PayPal](#)

[Project Gutenberg](#) > [71,725 free eBooks](#) > [2 by Apicius](#)

## Cookery and Dining in Imperial Rome by Apicius

Read now or download (free!)

Choose how to read this book ?	Size			
<a href="#">Read online (web)</a>	953 kB			
<a href="#">EPUB3 (E-readers incl. Send-to-Kindle)</a>	2.2 MB			
<a href="#">EPUB (older E-readers)</a>	2.2 MB			
<a href="#">EPUB (no images, older E-readers)</a>	408 kB			
<a href="#">Kindle</a>	3.9 MB			
<a href="#">older Kindles</a>	3.9 MB			
<a href="#">Plain Text UTF-8</a>	702 kB			
There may be more files related to this item.				

Project Gutenberg

## TXT 파일의 내용을 살펴보자.

≡ 29728-0.txt ×

```
2922 I
2923
2924
2925 [1] FINE SPICED WINE
2926 | _CONDITUM PARADOXUM_
2927
2928 THE COMPOSITION OF [this] EXCELLENT SPICED WINE [is as follows]. INTO
2929 A COPPER BOWL PUT 6 SEXTARII [1] OF HONEY AND 2 SEXTARII OF WINE; HEAT
2930 ON A SLOW FIRE, CONSTANTLY STIRRING THE MIXTURE WITH A WHIP. AT THE
2931 BOILING POINT ADD A DASH OF COLD WINE, RETIRE FROM STOVE AND SKIM.
2932 REPEAT THIS TWICE OR THREE TIMES, LET IT REST TILL THE NEXT DAY, AND
```

여기서 정보를 어떻게 뽑아낼 것인가?

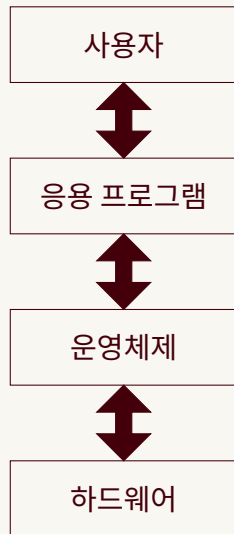
## 운영 체제(Operating system)

컴퓨터 하드웨어 바로 위에 설치되어 사용자 및 다른 모든 소프트웨어와 하드웨어를 연결하는 소프트웨어 계층

## 운영 체제의 목적

- 컴퓨터 시스템을 편리하게 사용할 수 있는 환경을 제공한다.
  - 하드웨어를 직접 다루는 복잡한 부분을 운영체제가 대행한다.
- 컴퓨터 시스템의 자원을 효율적으로 관리한다.
  - 실행 중인 프로그램들에 CPU를 번갈아가면서 할당한다.
  - 실행 중인 프로그램들에 메모리 공간을 적절히 분배한다.

김동현. 디버깅을 통해 배우는 리눅스 커널의 구조와 원리. 위키북스. 2021.  
반효경. 운영체제와 정보기술의 원리. 이화여자대학교 출판부. 2008.



## 운영 체제의 종류

### ■ 데스크톱

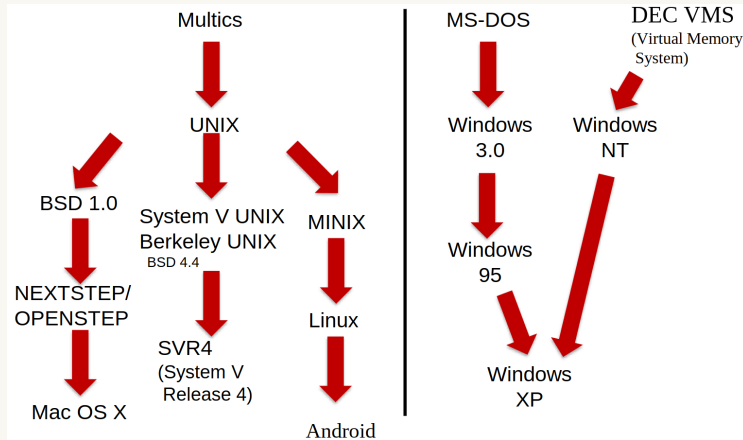
- MS Windows
- Mac OS
- Linux

### ■ 모바일

- Android (Linux 기반)
- iOS

### ■ 슈퍼컴퓨터

- Linux
- Unix



<https://os.korea.ac.kr>



리눅스를 직접 사용해 보자.

## 실습 코드

```
https://github.com/suparklingmin/LingDataSci2023/blob/main/notes/03-20230911.MD
```

## 한계

단어가 책 전체에서 사용된 빈도는 셀 수 있지만 무슨 요리에 나왔는지는 모른다.

## 해결책

하나의 큰 파일을 요리 단위로 작게 분할한다.

## 이번 시간에 배운 것

- 1 Linux 명령어를 활용하여 단어의 사용 빈도 계산하기

## 다음 시간에 배울 것

- 1 awk 명령어를 활용하여 텍스트의 내부 구조 탐색하기