

# 빅데이터분석개론 Curriculum



컴퓨터공학과  
이 강윤 (keylee@gachon.ac.kr)



# 가천대학교 컴퓨터공학과

- <https://www.gachon.ac.kr/kor/index.do>
- [가천대학교, 2024년 중앙일보 대학평가 '25위' -역대 최고](#)
- [가천대, 'THE 세계대학평가' 첫 참여 '국내 19위'](#)
- [컴퓨터공학과](#)

【서식 3】

2027 Top 10, 2032 Global Top 100		
 아름다운 인재의 샘 가천대학교	업무처리요청서	



이강윤 교수 (IT융합대학 컴퓨터공학과)

- 클라우드공학과 학과장
- 컴퓨터공학과 외국인 학과장
- 국제대학 컴퓨터공학과 학과장
- 가천인공지능기술원장

연구실 : AI공학과 312호,

연락처 : 031-750-8902 / 010-4997-7632

[keylee@gachon.ac.kr](mailto:keylee@gachon.ac.kr)

Cognitive Computing Lab (CCL)

AI공학과 103호, 031-750-8598

<https://sites.google.com/view/keylee>

#### [학력]

연세대 전자공학 학사

연세대 전자계산 석사

숭실대 IT정책경영 공학박사

#### [경력]

가천인공지능기술원장

전 한국 IBM 왓슨사업본부장

전 한국 IBM 연구소장

전 IBM 유비쿼터스컴퓨팅 연구소장

전 한국 IBM 최고기술자/상무

#### [현재]

인터넷정보학회 이사

행정안전부 지능형서비스 민간위원,

국민권익위원회 자문위원

#### [전문분야]

인공지능, 디지털혁신, IoT, 빅데이터,  
클라우드, 연합학습

# Computer Engineering

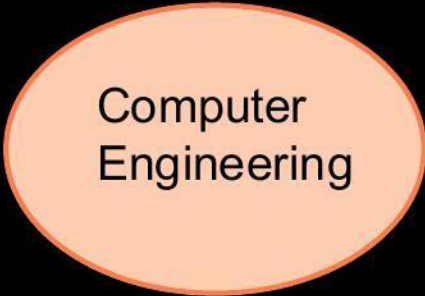
---

## General

- . Introduction
- . Computer Architecture
- . Algorithm/Math
- . SW Engineering
- . G-Project
- . P-Project

## OS

- . Linux
- . Embedded
- . Network
- . Security



Computer  
Engineering

## Programming

- . C, C++
- . Java, Python
- . Server, Web, Mobile

## Data

- . Database
- . Statistics
- . AI, IoT, Cloud
- . BigData
  - 데이터마이닝
  - 데이터분석처리
  - 빅데이터분석개론



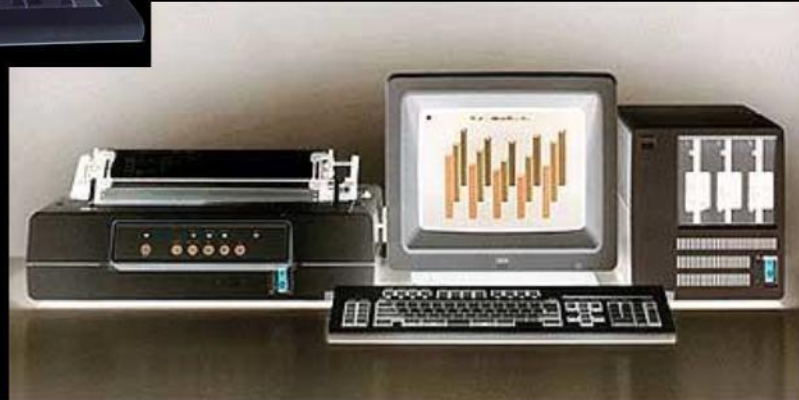
# Host and Terminals (1960 ~)

---



# IBM PC and Clones (1980 ~)

---



# Internet / Mobile (2000 ~)

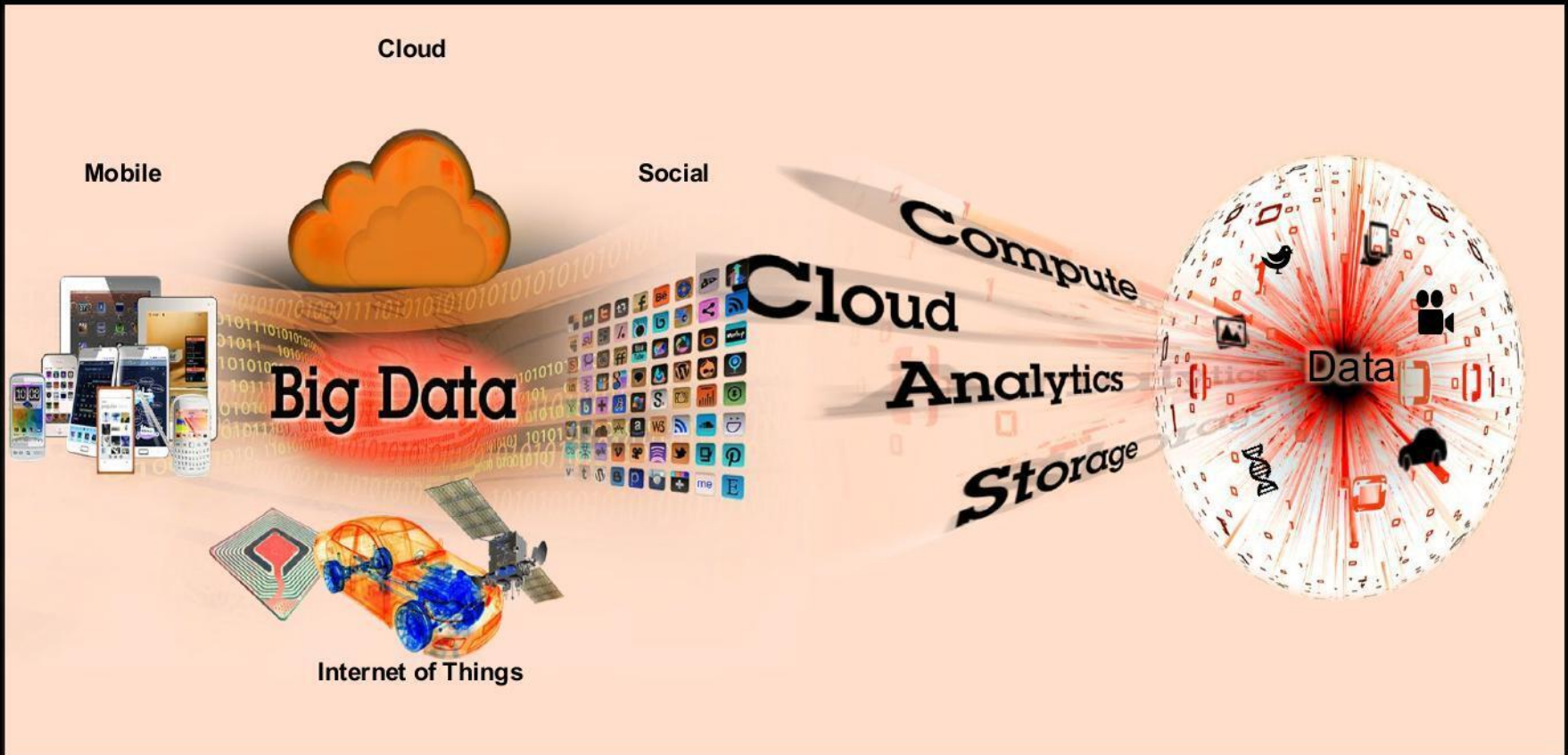
---



iPhone 2007



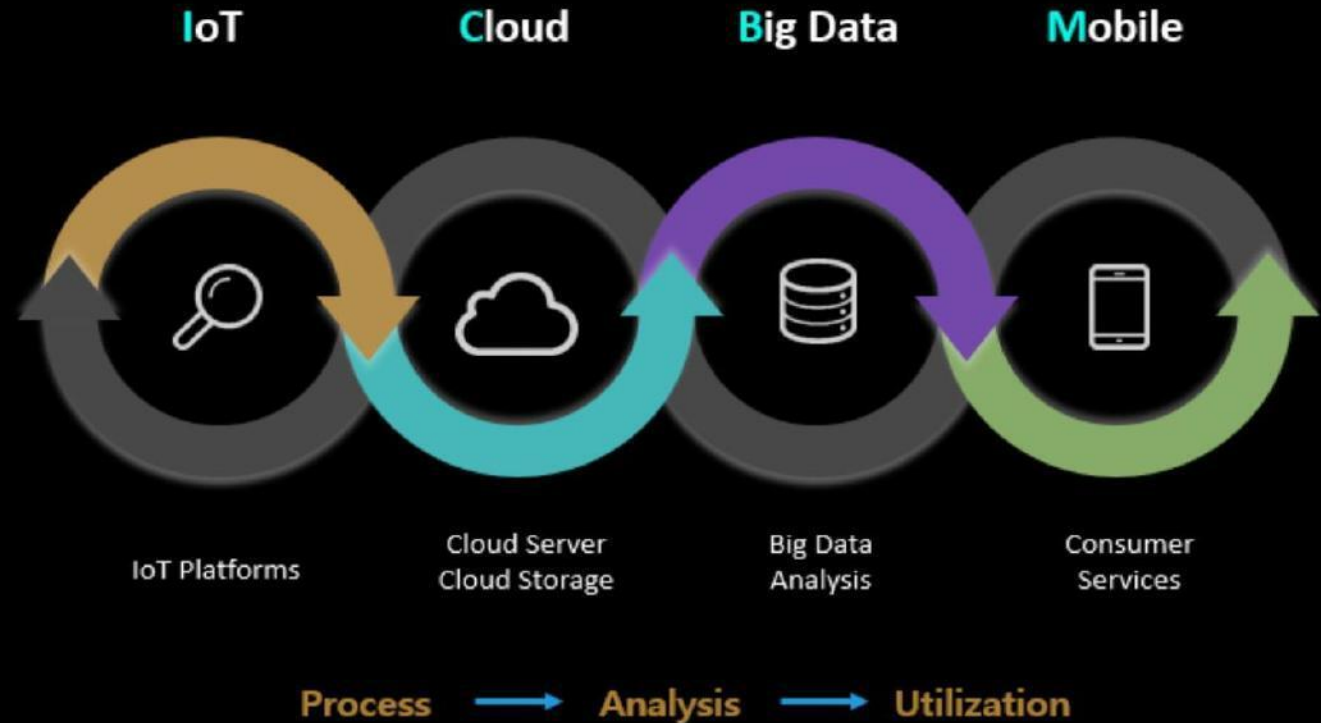
## Data and AI (2010~)



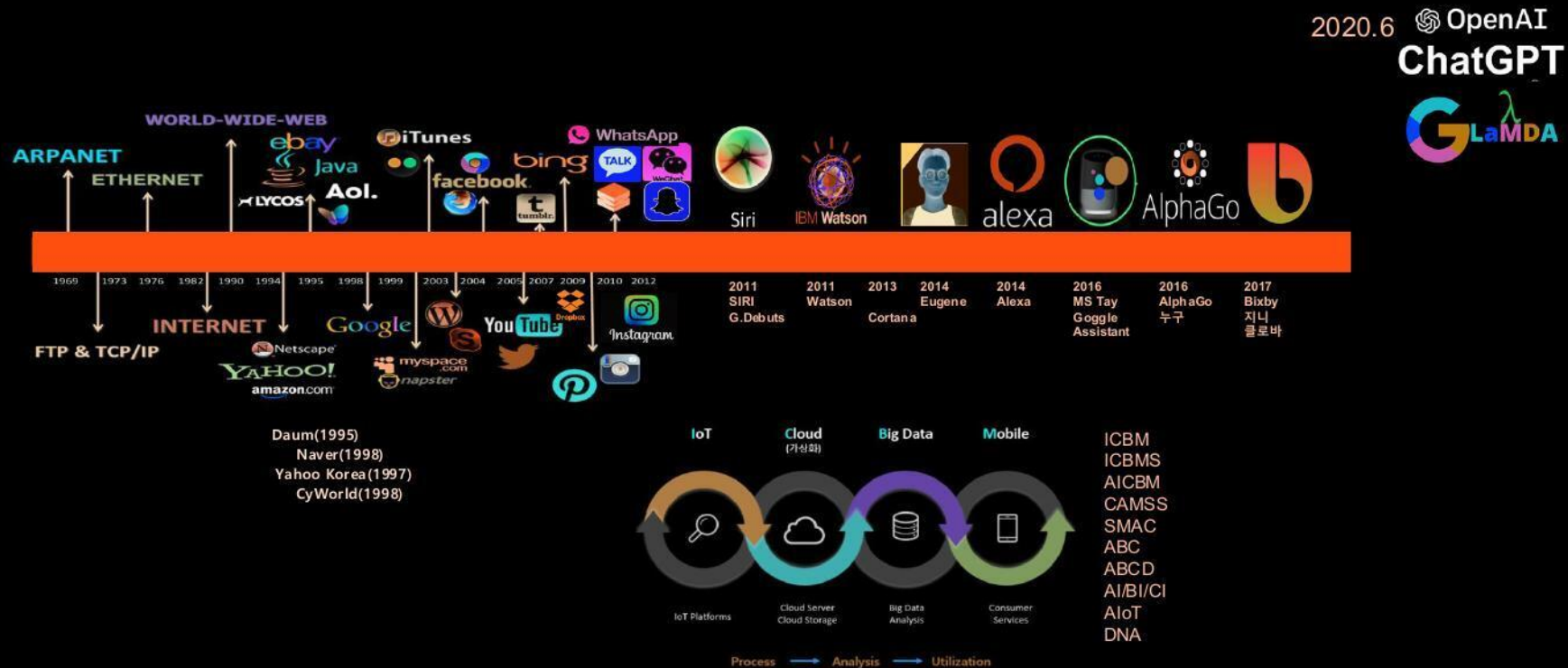


# IT Trends

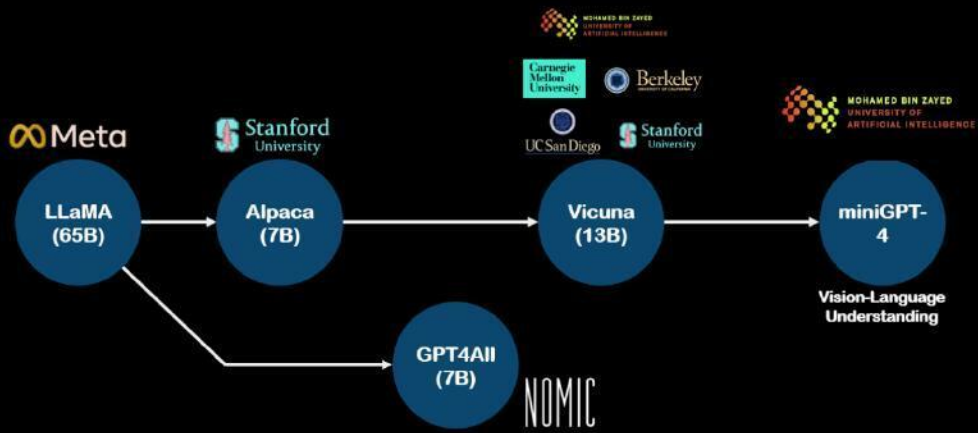
- ICBM
- ICBMS
- AICBM
- CAMSS
- SMAC
- ABC
- ABCD
- AI/BI/CI
- DNA



# Internet Timeline



# Generative AI



COST/RESOURCE	LLAMA 2	GPT-4	BARD
Computational Resources Required	Med	High	High
Model Training Cost	Low	High	High
Accessibility Cost	Free	Paid	Paid

# Future Outlook

---

- "I think there is a world market for maybe five computers."  
Thomas Watson, president of IBM, 1943
- Here's the legend: at a computer trade show in 1981, Bill Gates supposedly uttered this statement, in defense of the just-introduced IBM PC's 640KB usable RAM limit:  
"640K ought to be enough for anybody."



1955



1975



1995



2015



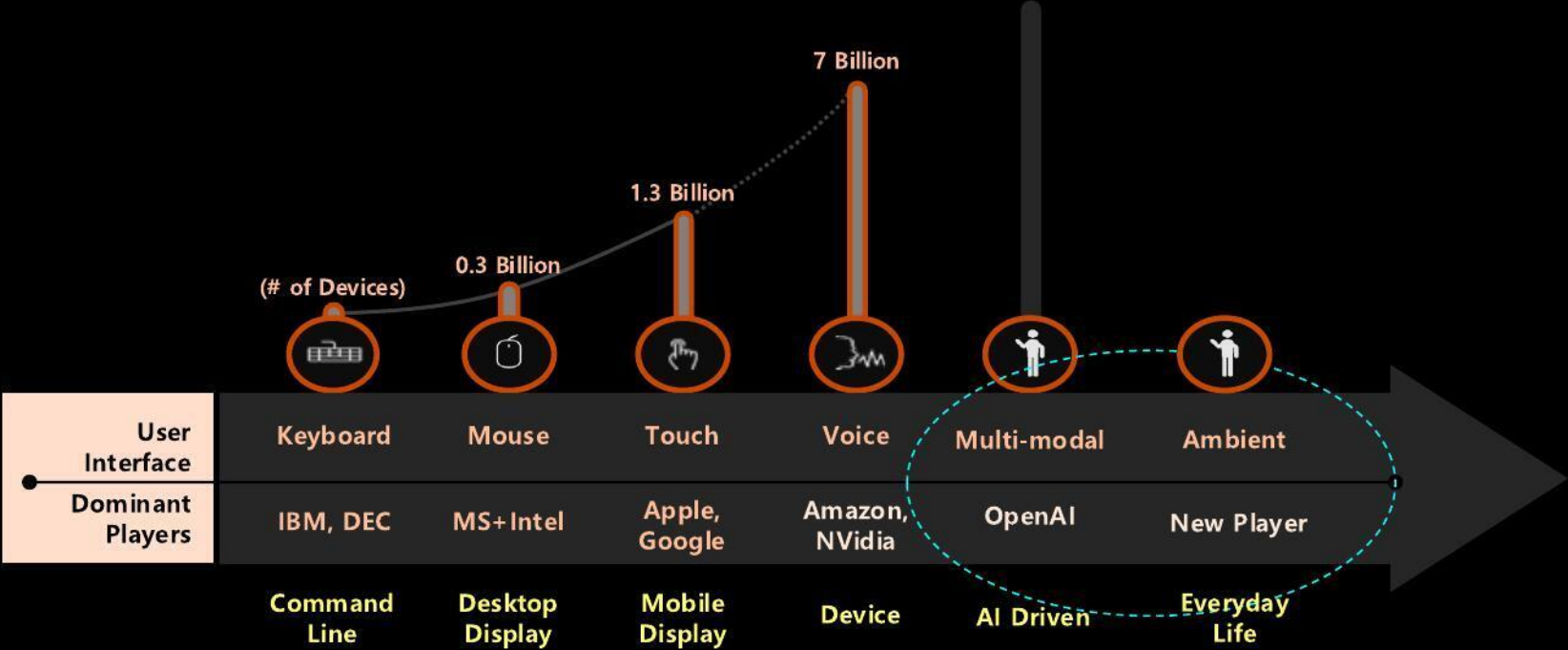
2035



2055



# New Player with New UX



# 한국경제, GDP순위

## OECD 전망 기준 명목 GDP 순위

순위 국가 ( ): 성장률실적, % 2020년 12월 대비 2021년 3월 기준



※ 성장률: OECD 중간경제전망  
디스플레이터, 환율: OECD 경제전망

자료/OECD

연암뉴스

## 2022년 기준 주요 국가 1인당 GDP (단위: 달러)



명목 GDP :  
한 나라에서 재화와 서비스가 얼마만큼 생산됐는지를 보여주는 지표로, 시장가격(당해연도 가격)을 기준으로 집계한다.

## 국제비교(명목 GDP)

2022년 우리나라의 경제규모는 13위 수준인 것으로 보임

(단위: 억USD달러, 시장환율 적용)

국가명	순위	금액	한국대비규모 (한국=100)	국가명	순위	금액	한국대비규모 (한국=100)
미국	1	254,627	1522	브라질	11	18,747	112
중국	2	178,760	1068	호주	12	17,023	102
일본	3	42,256	253	대한민국	13	16,733	100
독일	4	40,752	244	스페인	14	15,207	91
영국	5	30,798	184	멕시코	15	14,597	87
인도	6	30,096	180	인도네시아	16	13,340	80
프랑스	7	27,791	166	사우디아라비아	17	11,081	66
캐나다	8	21,436	128	네덜란드	18	11,037	66
러시아	9	20,503	123	튀르키예	19	8,855	53
이탈리아	10	20,105	120	스위스	20	8,054	48

주 : 1) 대한민국, 미국, 중국, 일본, 독일, 인도, 영국, 프랑스, 이탈리아, 호주 외에는 OECD 및 IMF 전망 기준  
자료: UN(2023.6.1일 기준), OECD, IMF





# Technology Company

## Global Consulting Company

McKinsey  
& Company

BCG

BAIN & COMPANY

EY  
Building a better  
working world

McKinsey & Company, BCG, Bain & Company,  
EY, PwC, KPMG, Deloitte, Accenture

pwc

KPMG

Deloitte

accenture

## IT Company

NAVER

kakao  
LINE

coupang

배달의민족

당근마켓

toss

SAMSUNG SDS

LG CNS

SK  
주식회사  
C&C

SAMSUNG

SK

HYUNDAI

LG

## THE MOST INNOVATIVE COMPANIES IN 2023

Boston Consulting Group (BCG) has released the latest edition of their Most Innovative Companies ranking.

### Industry

Technology Transport & Energy Healthcare Consumer Goods & Services Telecommunications

Rank	Change in Rank (1 year)		
1	—	Apple	11 Pfizer
2	▲3	TESLA	12 Johnson & Johnson
3	—	amazon.	13 Returned SPACEX
4	—	Alphabet	14 ▲1 NVIDIA
5	▼3	Microsoft	15 Returned ExxonMobil
6	▲1	moderna	16 ▼5 Meta
7	▼1	SAMSUNG	17 ▼5 Nike
8	—	HUAWEI	18 ▼8 IBM
9	Returned	BYD	19 ▲18 3M
10	▲10	SIEMENS	20 Returned TATA



# Global Tech Analysis Research Firms

---

**Gartner**



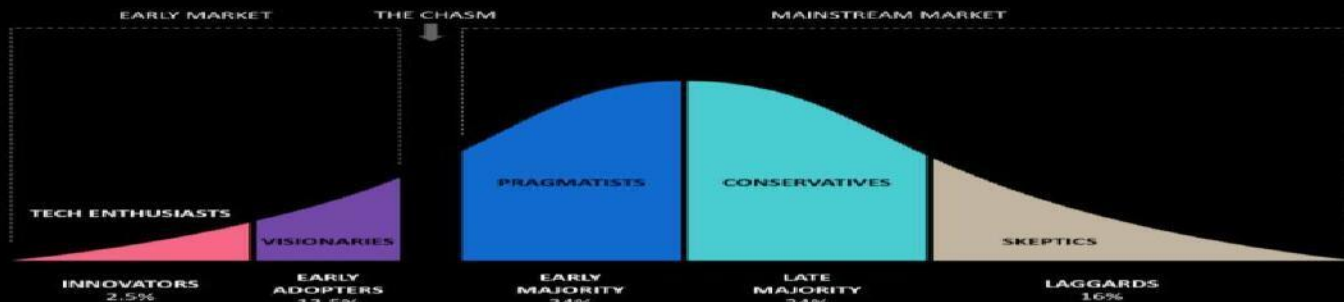
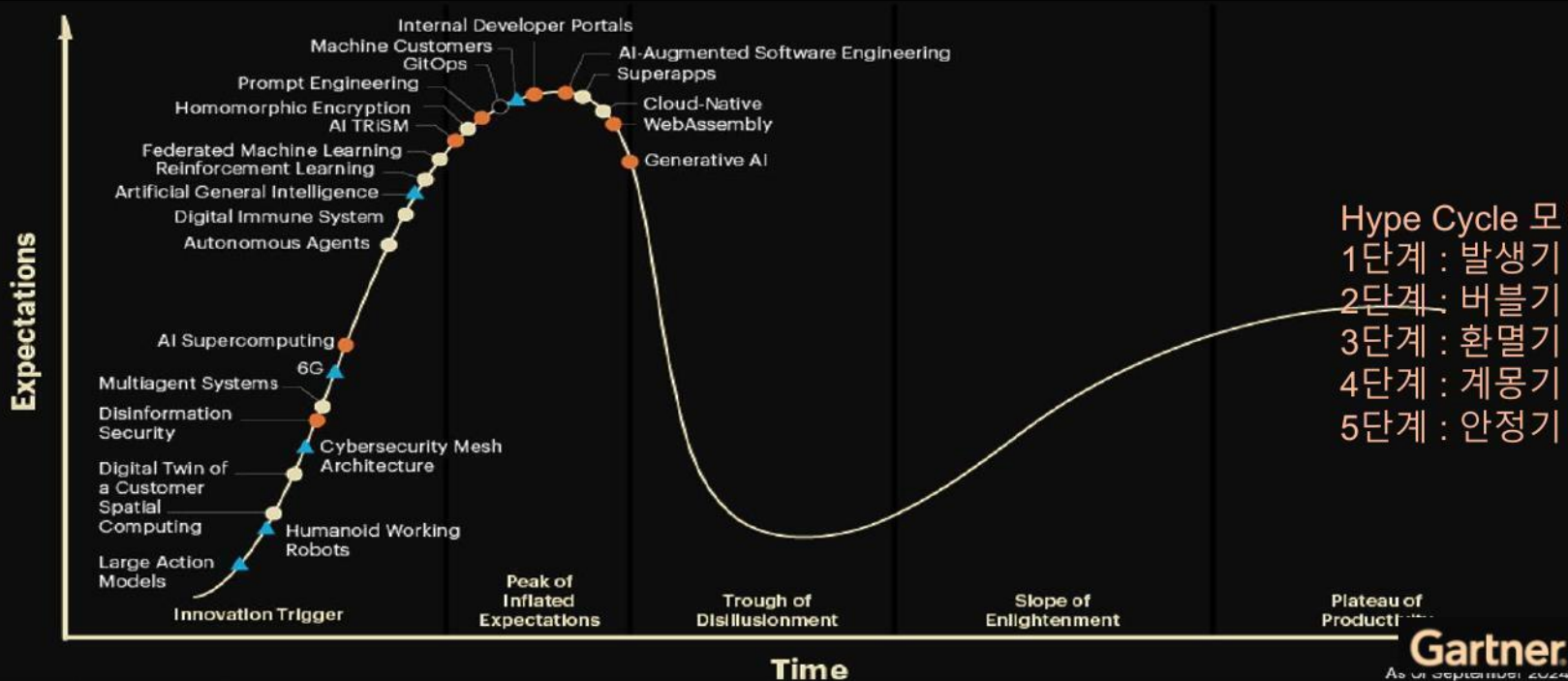
**FORRESTER®**



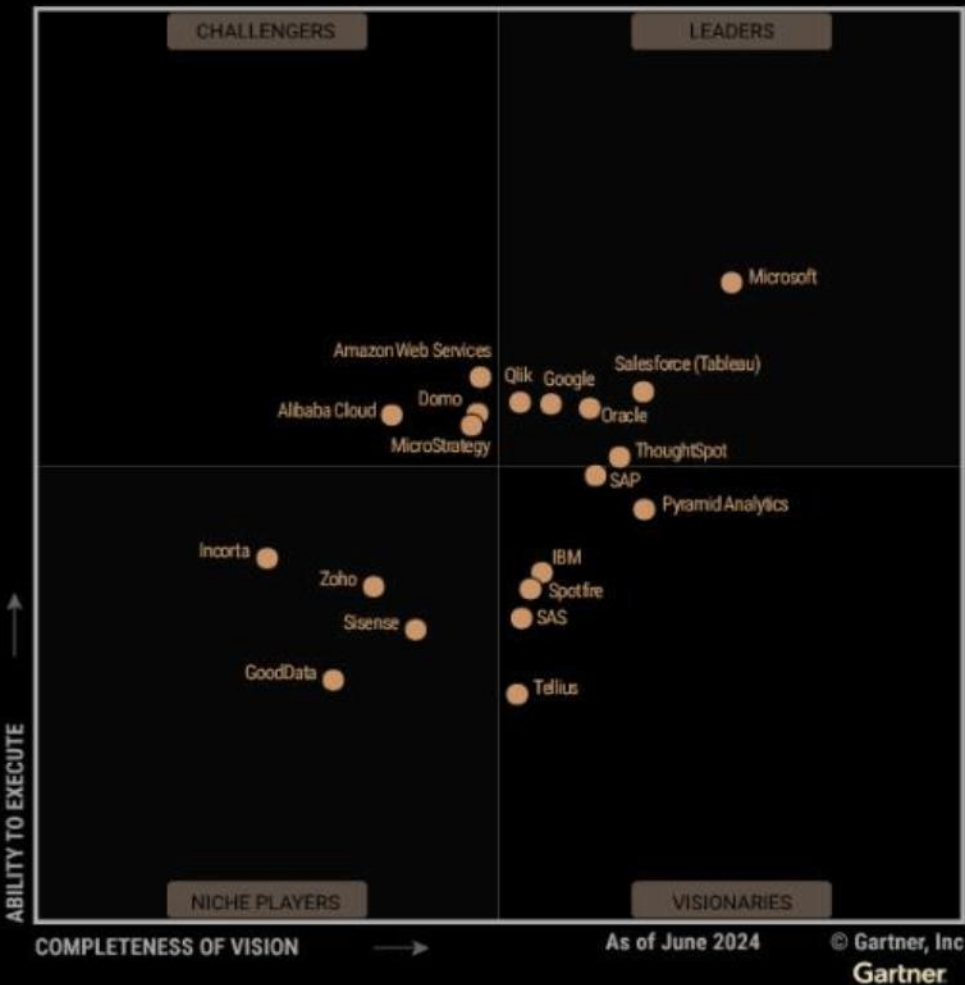
**F R O S T & S U L L I V A N**

**ovum**

# Hype Cycle for Emerging Technologies, 2024

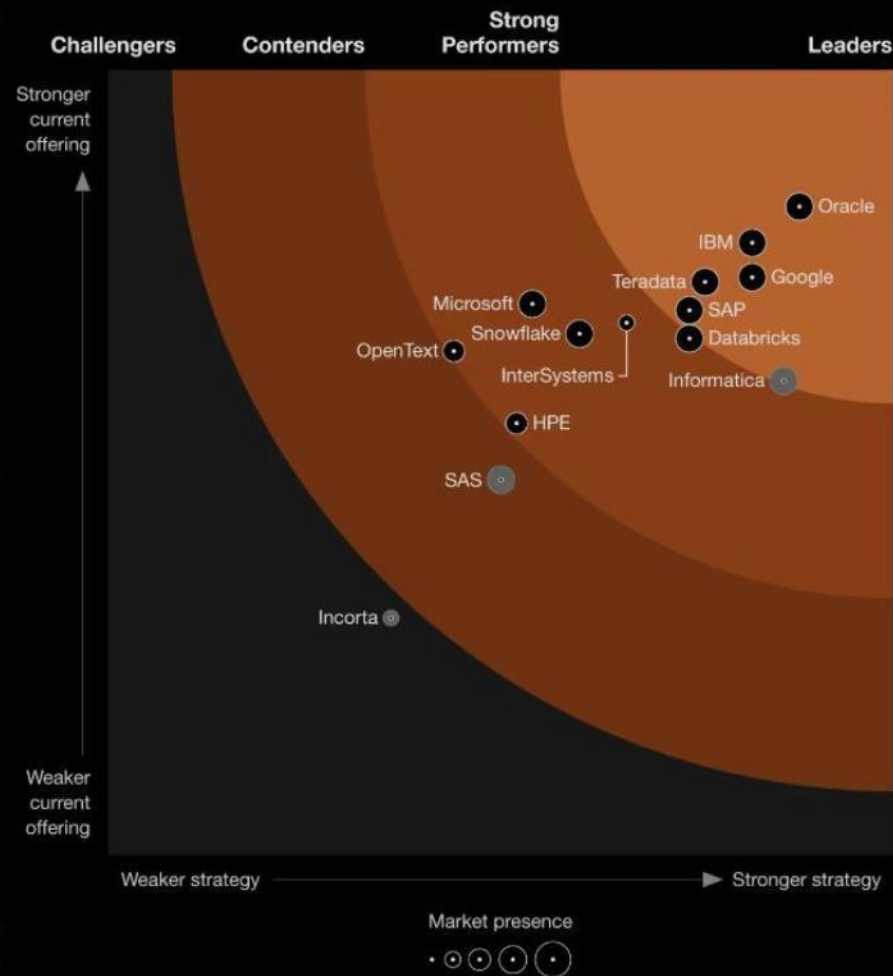


# 2024 Gartner® Magic Quadrant for Analytics & Business Intelligence Platforms



# The Forrester Wave: Data Management for Analytics, Q1 2023

— Forrester Research



# 수업시간표

교과목명	빅데이터분석개론		학수번호	14798001		이수	전선	학점	3
강의시간	수6,수7,수8		강의실 - AI공학관 508						
선수과목									
교수소속	IT융합대학 컴퓨터공학과		교수성명	이강윤	연락처	031-750-8902			
e-mail	keylee@gachon.ac.kr		연구실	AI공학관 312호	지도상담시간	수업후 / 메일로지정			
홈페이지	<a href="http://sites.google.com/view/keylee">http://sites.google.com/view/keylee</a>								
화상강의주소	<a href="http://gachon.webex.com/meet/keylee">http://gachon.webex.com/meet/keylee</a>								



# 강의 개요

---

- 빅데이터 강의는 방대한 빅데이터의 개념과 기술을 이해하고 빅데이터 플랫폼 기반에서 빅데이터를 수집 저장 처리 분석 및 표현하는 기술을 구현하는 수업입니다. 이 수업에서는 강의를 통하여 빅데이터에 관한 전반을 이해하고 파이썬 빅데이터 실습을 통하여 빅데이터 시스템의 동작 구현을 해보는 과정으로 진행하고 AWS 클라우드를 활용한 빅데이터를 실습합니다.

1. 데이터 테크놀로지의 발전 방향
2. 빅데이터 컴퓨팅 기술
3. 파이썬 빅데이터 분석
4. AWS 클라우드 빅데이터 활용

- 이 강의의 목표는 빅데이터의 개념과 기술을 이해하고 빅데이터 플랫폼의 컴퓨팅 기술을 이해하고 파이썬 빅데이터 분석 실습 프로젝트를 통하여 직접 개발하고 클라우드 기반으로 운영할 수 있는 기술을 배우고 익히는 것이 최종적인 학습 목표입니다. 본과목은 초반부 이론 강의와 중반부 파이썬 프로젝트 실습, 후반부 클라우드 빅데이터의 활용으로 구성된 과목입니다.

# 강의 교재

구분	교재명	저자	출판사	출판연도
주교재	강의자료			
주교재	데이터과학기반의 파이썬 빅데이터 분석	이지영	한빛아카데미	2024
주교재	AWS Cloud Site - AWS Acdemy			



# Course Agenda

Weeks	기 간	수 업 내 용 및 학 습 활 동
1	3/5	강의소개, 강의일정소개, 평가방법소개
2	3/12	데이터 테크노로지 혁신/ 빅데이터/ 클라우드
3	3/19	빅데이터 컴퓨팅 기술 소개 (수집 및 통합, 저장 및 관리)
4	3/26	빅데이터 기업 사례와 빅데이터 산업의 이해
5	4/2	데이터 과학 기반의 빅데이터 분석/실습환경 준비
6	4/9	빅데이터분석 - 파이썬 크롤링 (API 이용) 실습
7	4/16	빅데이터분석 - 파이썬 크롤링 (라이브러리 이용) 실습
8	4/23	중간고사
9	4/30	통계분석 실습
10	5/7	텍스트빈도 분석 실습
11	5/14	지리정보 분석 실습
12	5/21	AWS Academy 클라우드 서비스
13	5/28	AWS Academy 빅데이터 서비스
14	6/4	AWS Academy EC2/ECS
15	6/11	기말고사

# 데이터 과학 기반의 파이썬 빅데이터 분석(2판)

## ■ 총 18개 프로젝트

01 와인 품질 등급 예측하기

10 센서 데이터로 움직임 분류하기

02 타이타닉호 생존율 분석하기

11 타겟 마케팅을 위한 소비자 군집 분석하기

03 영문 문서 제목의 키워드 분석하기

12 영화 리뷰 데이터로 감성 분석 모델링하기

04 한글 뉴스 기사의 키워드 분석하기

13 챗GPT 뉴스 텍스트의 감성 분석하기 **NEW**

05 지리 정보 분석 후 맵 생성하기

14 뉴스 텍스트에서 챗GPT 토픽 분석하기 **NEW**

06 행정구역별 의료기관 현황 분석하기

15 주가 시계열 분석하기 **NEW**

07 항목에 따른 자동차 연비 예측하기

16 100일 후의 주가 예측하기 **NEW**

08 대기오염 데이터와 미세먼지의 연관성 분석하기

17 숫자 이미지 분류하기 **NEW**

09 특징 데이터로 유방암 진단하기

18 강아지 품종 분류하기 **NEW**

## ■ 최신 라이브러리 적용하여 모든 코드 업데이트



# 평가와 수업 진행

평가요소	성적평가방법	비율
출석	강의참여(결석1일=1점감점, 지각2회 = 1점 감점)	20
중간고사	필기시험	30
기말고사	필기 / 실기시험	30
레포트	과제 / 실습 / 발표 점수	20
그룹프로젝트		0
기타		0
합계		100

Cyber Campus에 강의교제 및 수업 관련 정보 공유

Chance favors the prepared mind. (루이 파스퇴르)

Gachon  
Cognitive Computing Lab

