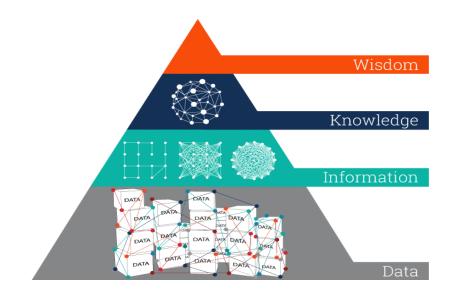
# 4차산업혁명과 관련 기술들

담당교수 : 육계산(陸桂)山

#### Wikipedia 정의

빅데이터는 기존 데이터베이스 관리 도구의 능력을 넘어서는 대량의 정형 또는 비정형 데이터(텍스트 등) 집합까지 포함한 데이터로부터 가치를 추구하고 결과를 분석하는 기술이다. 즉, 빅데이터는 단순히 대량의 데이터 저장소를 의미하는 것이 아니라, 다양한 종류의 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 의미한다.

정보시스템이나 지식관리 영역에서 흔히 사용되는 DIKW 피라미드
 -> 단순히 데이터나 인포메이션을 넘어서 지식과 지혜를 추출해내는 기술이 빅데이터



#### Each step up

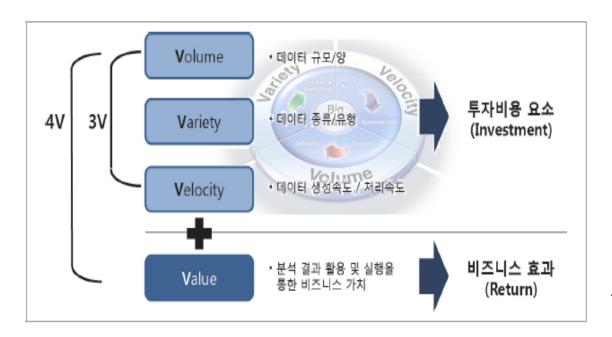
the pyramid answers questions about and adds value to the initial data.

▶ 빅데이터 의미 -> 다음과 같은 요소들을 포함

'빅데이터'란 데이터의 많은 부분을 가져와 분석하고, 여기에서 무엇인가 유용한 것을 만든다는 것을 의미하겠지만 빅데이터의 의미는 그 보다 훨씬 더 넓어서 다음과 같은 요소를 포함한다.

- 대부분의 경우 여러 소스로부터 방대한 양의 데이터를 수집
- 단순하게 데이터의 양만 많은 것이 아니라 그 종류도 다양
  - 대부분의 경우 동시에 여러 종류의 데이터와 시간이 경과하면서 바뀌는 데이터들을 수집
  - 처음부터 구체적인 형식으로 변형시키거나 또는 일관적으로 만들 필요가 없는 데이터들
- 동일한 데이터 풀을 다양한 목적으로 지속적으로 분석할 수 있도록 이 데이터들을 분석
- 이 모든 작업을 신속하게, 때에 따라서는 실시간으로 수행

- 미국의 IT 자문기관 가트너 (Gartner)는 빅데이터를 다음과 같이 정의 빅데이터란 높은 통찰력, 의사결정, 프로세스 자동화를 위해 비용 효과가 높은 혁신적인 정보처리 과정을 요하며, 대용량의 데이터 규모 (high-volume), 빠른 속도 (high-velocity), 높은 다양성 (high-variety)을 지닌 정보 자산이다. (3V 정의)
- V로 설명되는 특징에 대해서는 다양한 견해가 있으나 가트너의 3V 정의 모델이 가장 널리 사용되는 박데이터의 정의가 된 바 있다.
- 이에 더해 IBM은 정확성(Veracity)이라는 요소를 더해 4V로 정의하였고, 브라이언 홉킨즈(Brian Hopkins) 등은 가변성 (Variability)을 추가하여 4V로 정의하기도 하였다.



<빅데이터의 특징을 고려한 분석 요소>

#### ❖ 빅데이터의 특징

- **3V** : 규모(Volume), 다양성(Variety), 속도(Velocity)
- 5V : 규모(Volume), 다양성(Variety), 속도(Velocity), 정확성(Veracity), 가치(Value)
- 빅데이터의 특징은 3V로 요약하는 것이 일반적이다.
- 즉 데이터의 양(Volume), 형태의 다양성(Variety), 데이터의 생산 속도(Velocity)를 의미하는데,
- 최근에는 기존 데이터와 차별화되는 빅데이터의 특성을 5V(Volume, Variety, Velocity, Veracity, Value)로 설명한다.

#### ■ 규모(Volume)

- 물리적인 크기와 개념적인 범위까지 대규모인 데이터 양
- 기업이 관리하고 분석하는 빅 데이터 크기와 양
- 기술, ICT 발전 -> 디지털 정보량 급증 -> 제타바이트(ZB) 시대 진입

#### ■ 다양성(Variety)

- 기존의 구조화된 정형 데이터는 물론 사진, 동영상 등의 비정형 데이터가 포함
- 비정형 데이터, 반정형 데이터 및 원시 데이터를 포함하는 다양한 데이터 유형
- 텍스트 외 음성, 영상 등 비정형 데이터 종류 증가

#### 속도(Velocity)

- 실시간으로 생산되며, 유통 속도 또한 매우 빠른 데이터
- 회사가 데이터를 수신, 저장 및 관리하는 속도(예: 특정 기간 내에 수신된 특정 소셜 미디어 게시물 또는 검색 쿼리 수)
- IoT, 실시간 정보 증가 -> 데이터 생성 및 이동 증가 -> 실시간 데이터 처리 분석 위한 속도의 중요성

#### ■ 신뢰성, 정확성(Veracity)

- 데이터의 원천 및 형태 다양화에도 불구하고 신뢰성이 보장된 데이터
- 종종 경영진들의 신뢰도에 영향을 미치는 데이터 및 정보 자산의 진실 또는 정확성
- 방대한 데이터의 질이 데이터 분석 정확도에 영향을 미침

#### ■ 가치(Value)

- 새로운 가치를 창출할 수 있는 데이터
- 빅데이터가 추구하는 것이 바로 가치
- 빅데이터 분석 -> 통찰력 제공 -> 기업의 현실문제 해결에 도움

#### ■ 빅데이터 분석(Analytics)

- 빅데이터에서 분석이란 사물을 이해하는데 필요한 광의의 분석이나 데이터의 단순 조회 및 단순 리포 팅의 생산 과정이 아닌 데이터에 근간한 통계분석, 트랜드 예측, 최적화 등이 여기에 해당
- 또한 의사결정과 그 Action에 활용하기 위한 데이터의 광범위한 활용, 통계적이며 정량적 측면의 분석, 탐색적 분석 및 예측 모델링, 사실에 근거한 경영을 의미



#### ■ 빅데이터 활용 트랜드

- 빅데이터는 2012는 세계경제포럼에서 발표한 10대 신기술 중 첫 번째로 선정되면서 기업들에게 데이터의 가치와 중요성을 인식하게 하였음
- 이후 많은 기업들이 필수적으로 빅데이터 시스템을 구축하고 방대한 양의 데이터를 수집, 관리하고 있으며, 이를 분석해 의사결정, 비즈니스 모델 확장 등에 필요한 정보를 도출하기 위해 많은 비용과 시간을 투자하고 있음
- 뿐만 아니라 데이터 처리 및 분석 기술의 발달로 빅데이터 활용 분야가 다양해 지고 있으며, 4차 산업 혁명으로 다양한 분야가 융합되면서 데이터 간의 연결을 통한 통합 분석이 중요해 지고 있음
- 현대는 모든 것이 연결되는 초연결(Hyper-Connect) 사회로, 이러한 연결은 더욱 견고해 지고 있음
- 따라서 지금까지는 데이터의 안정적인 관리가 핵심이었다면, 최근에는 데이터간 초연결을 통한 통찰력(Insight)을 발견하는 것으로 변화의 흐름이 가속화 되는 중임

#### ■ 데이터가 필요한 이유

- 정보를 담고 있고, 새롭고 가치 있는 의미를 담은 정보를 재가공 할 수 있어 4차 산업혁명의 대표 기술 로 등장
- 근대사에서 제한된 양의 데이터라도 가장 어려운 도전과제에 대한 해결책을 찾아내는 통찰력을 제공하는데 기여했다.
- 비석이나 파피루스 두루마리, 계몽책자, 인쇄물 등에 기록되었는지에 관계없이 데이터와 데이터의 보 편성 및 중요성은 경제 성장 및 인류 발전의 주요 원동력이었다.
- 21세기에는 이러한 과정이 급속하게 가속화되고 있습니다. 데이터의 양은 더 많아지지만 데이터 저장 비용이 급감하면서 데이터 과학자들은 새로운 기술로 개발된 첨단 도구로 무장하여 다량의 데이터에 서 중요한 통찰력을 얻고 있다.
- 데이터를 처리하는 기술이 점점 다양화되면서 그러한 기술의 영향도 더욱 커지고 있고 기회도 더 많아 지고 있다.

#### ❖ 빅데이터 활용 하기 위한 데이터 분석 사이트

#### • 구글 트렌드

- 글로벌 사이트인 구글에서 운영하는 구글 트렌드는 최근 인기 검색어 확인이 가능하며, 요즘 이슈가 되는 키워드와 관련해 지역별 관심, 관련주제, 관련 검색어 등을 확인할 수 있다.
- 국내 사이트에 비하면 가장 오래된 데이터를 갖고 있다는 것이 장점이다.

#### ■ 네이버 데이터랩

- 네이버 데이터랩은 국내 최대 규모의 포털사이트인 네이버에서 운영 하는 데이터 분석 툴
- 검색 트렌드 및 급상승 검색어 이력, 쇼핑 카테고리별 검색 트렌드 제공 등의 기능이 마련되어 있다.

#### • 카카오데이터 트렌드

- 네이버와 유사한 형태로 트렌드 분석이 제공된다.
- 다음 검색어를 바탕으로 한 정보를 제공하며, 주제어 검색은 불가능하지만 가장 보기 쉬운 유저 인 터페이스를 제공한다는 점이 장점이다.
- 상대적으로 단기간의 트렌드 확인이 필요하거나 키워드 관련 여러 정보가 필요할 때 이용하면 편리하다.

#### ❖ 무료 빅데이터 사이트 모음

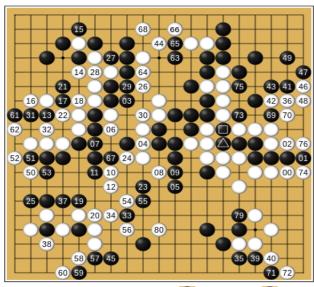
- 네이버데이터랩 : https://datalab.naver.com
- 통계청 : http://kostat.go.kr
- 공공데이터 : https://data.go.kr
- 마이크로 데이터 통합서비스 : https://mdis.kostat.go.kr/mobile/index.do
- 서울 열린데이터 광장 : https://data.seoul.go.kr
- 캐글 : https://www.kaggle.com
- 소셜뱅거 : https://www.socialbakers.com
- 아마존 알렉사 웹분석 : https://www.alexa.com
- 와이즈웹(부분 유료화 됨) : https://www.wiseapp.co.kr
- 인스타 인기태그분석 : https://top-hashtags.com/instagram

# 인공지능(AI)

- 2016년 3월 9일 ~ 15일 알파고 대 이세돌 9단간의 혹은 딥마인드 챌린지 매치(Google Deepmind Challenge Match)가 대한민국의 포 시즌스 호텔에서 개최
- 최고의 인공지능 프로그램과 바둑 분야 최고의 인간 실력자와의 대결
- 최종 결과는 인간 대표인 이세돌 9단의 1-4 패배로 인공지능의 실체를 세계에 알려준 사건
- 2016년 한 해 동안 4차 산업혁명을 대표하는 키워드 중 가장 많은 관심을 받은 키워드
- 알파고(AlphaGo)는 2014년 구글에 인수된 딥마인드(DeepMind)에서 개발한 머신러닝 기반의 인공지능 바둑 프로그램으로 기존의 바둑 프로그램들과는 달리 자기 자신과의 대국을 통해 스스로 학습할 수 있 다는 특징을 가지고 있음



<방송에 소개되었던 이세돌 9단과 알파고의 바둑 대결 소개 영상 및 이세돌 9단이 승리한 4국 모습(백이 이세돌9단)>



100수 ~ 180수 (177 자리에 📵, 178 자리에 🛆)

# 인공지능(AI)

■ 인공지능 정의(AI):

인공 + 지능

**Artificial + Intelligence** 

- 데이터 분석에 특화된 과학분야에 활용되는 학술용어
- 인공지능(Al, Artificial Intelligence)이란? 인간이 가지고 있는 지적 능력을 컴퓨터에서 구현하는 다양한 기술이나 소프트웨어, 컴퓨터 시스템 등을 통틀어 일컫는다.
- 1956년 미국 디트머스 대학에서 열린 회의에서 존 매카시는 인공지능이란 용어를 언급하며, "기계를 인간 행동의 지식 에서와 같이 행동하게 만드는 것"이라고 정의 내림
- 영국의 수학자인 앨런 튜링은 컴퓨터가 사람처럼 생각할 수 있다는 견해를 제시하며, 대화를 나누어 상대방이 컴퓨터인지 사람인지를 구별할 수 없다면 그 컴퓨터는 사고할 수 있는 것으로 간주해야 한 다고 주장

### 인공지능(AI)

#### ■ 인공지능과 미래

- 2000년대까지 인공지능은 주로 인간이 미리 수집하여 만들어 놓은 지식을 기계가 학습하는 방식으로 구현
  - -> 즉, 각 분야에서 필요한 지식을 미리 수집하여 그것을 전문가가 정교하게 모델링하고, 그 결과를 기계에 학습시키는 것으로 특정 분야의 인공지능을 만들 수 있었음
- 엄청난 양의 데이터를 활용할 수 있게 해주는 빅데이터 기술과 컴퓨터가 스스로 학습하는 방식인 머신러닝을 통해 조금씩 발전해온 인공지능은 인간의 뇌를 모방한 신경망 네트워크 구조로 이루어지는 혁신적인 딥러닝 알고리즘의 등장으로 기존보다 압도적인 성능을 보여주는 인공지능을 생산해 내기 시작함
- 인공지능이 활용되는 분야 : 여러 학문을 포괄하는 광범위한 분야
  - 인공지능은 인간이 할 수 없었던 광범위한 데이터를 분석하여 그동안 해결하지 못했던 많은 문제들을 해결
  - 인공지능은 현재 데이터과학이라는 이름으로 분야, 학문에 관계없이 정치, 경제, 의료, 예술 등 모든 곳에 활용되고 있음
  - -> 컴퓨터공학, 데이터 분석 및 통계, 엔지니어링, 마이닝을 통한 번역(언어학) 의학 등 여러 학문에서 축적된 데이터와 자료를 통해 새로운 인사이트를 발굴하는데 인공지능이 활용되고 있다.

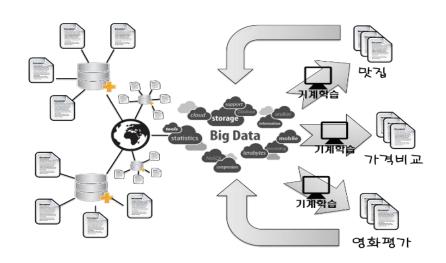
- 인공지능의 분류: 인공지능이라는 용어는 데이터 과학에서 가장 큰 범주
  - 인공지능의 세분야는 머신러닝(기계학습)과 딥러닝이 포함되어 있음
  - 머신러닝과 딥러닝은 인공지능 AI와 분명히 차별화된 분야지만 광의의 의미에서 인공지능에 포함될 수 있다.
  - 데이터 과학에서 인공지능을 바라보는 시각
    "A program that can sense, reason, act and adapt"
    감지하고 추론하고 수행하고 적용하는 프로그램

#### ■ 인공지능의 범주

- 인공지능의 세분야는 머신러닝(기계학습)과 딥러닝이 포함되어 있음
- 인공지능 AI(광의의 개념) > 머신러닝(협의의 개념) > 딥러닝(최협의 개념)



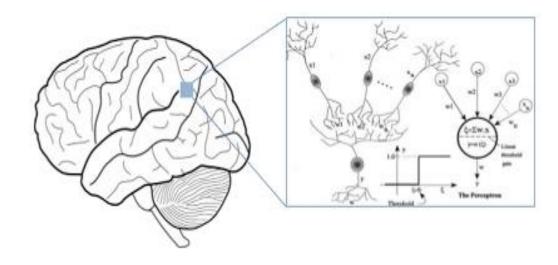
- 인공지능(AI)이란: "인간이 지닌 지적 능력의 일부 또는 전체를 인공적으로 구현한 것"을 의미
- 머신러닝(Machine Learning, 기계학습)
  - 컴퓨터를 인간처럼 학습시킴으로써 컴퓨터가 새로운 규칙을 생성할 수 있지 않을까 하는 시도에서 시작되었다.
  - 머신러닝이란 컴퓨터가 스스로 학습할 수 있도록 도와주는 알고리즘이나 기술을 개발하는 분야
  - 많은 관련 데이터를 입력해주면 관련 프로그램이 스스로 규칙을 학습하는 기계 학습을 의미
  - 머신러닝 분야는 인공지능의 한 분야로 딥러닝을 포함하고 있는 분야
  - 핵심 기술
    - \_ 표현(Representation)
    - 일반화(Generalization)



<기계 학습 활용 예>

#### ■ 인공 신경망(Artificial Neural Network)

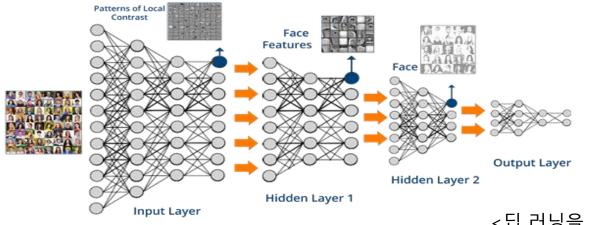
- 인간의 뉴런 구조를 본떠 만든 기계 학습 모델
- 생물학의 신경망(동물의 중추신경계 중 특히 뇌)에서 영감을 얻은 통계학적 학습 알고리즘
- 컴퓨터 비전 또는 음성 인식과 같은 범위의 문제 해결에 이용



<인공 신경망 구조>

#### ■ 딥 러닝(Deep Learning)

- 딥러닝(심층학습)이란 머신러닝(기계학습)의 한 부분으로, 오랜 시간 축적된 데이터를 바탕으로 생산성을 높일 수 있는 알고리즘을 찾아가는 것
  - -> 여러 비선형 변환기법의 조합을 통해 높은 수준의 추상화를 시도하는 기계 학습 알고리즘의 집합
- 딥러닝은 인간의 인지능력(신경망: neural network)을 모방해서 만드는 게 핵심 아이디어
- 딥러닝은 기존의 기계학습에서 인공신경망(ANN: Artificial Neural Network) 개념을 적용해서 통계학적 인 학습 알고리즘을 구현해 낸 것
- 딥러닝 분야는 머신러닝 모델의 한 종류이므로 인공지능과 머신러닝에 포함되는 분야
- 컴퓨터 비전, 음성인식, 자연어 처리 및 음성/신호처리 분야에 적용 (지진 예측과 자율 주행 자동차 운전 등에서 적극 활용 중)



#### ■ 인공지능과 미래

- 인공지능은 인간이 할 수 없었던 광범위한 데이터를 분석하여 그동안 해결하지 못했던 많은 문제들을 해결
- 인공지능은 현재 데이터과학이라는 이름으로 분야, 학문에 관계없이 정치, 경제, 의료, 예술 등 모든 곳에 활용되고 있다.

#### 인지 컴퓨팅(Cognitive Computing)

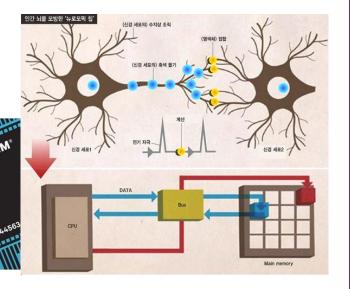
- 지각 컴퓨팅(Perceptual Computing)이라고도 함
- 인간과 같은 감각을 가지고 사용자의 의도를 파악할 수 있는 컴퓨팅

#### ■ 뉴로모픽 컴퓨팅(Neuromorphic Computing)

- 인공 신경망을 하드웨어적으로 구현한 것
- 딥러닝의 소프트웨어적인 구현을 하드웨어적으로 구현



<인지 컴퓨팅>



<뉴로모픽 칩인 퀄컴사 '제로스(Zeroth)'>

- 자율주행 자동차(Autonomous Vehicle)
  - 운전자의 개입 없이 주변 환경을 스스로 인식하고 주행 상황을 판단해 차량을 제어함으로써
     스스로 정해진 목적지까지 주행하는 자동차
  - 아직은 자율 주행 기술의 안전성이 검증을 받고 있는 단계이며 향후 대중화가 될 것임









### 인공지능(AI)과 로봇의 관계?

- 로봇(Robot)이란: "인간과 유사한 모습과 기능을 가진 기계 또는 무엇인가 스스로 작업하는 능력을 가진 기계"를 의미
- 산업용 로봇: 제조업 공장에서 조립,용접,핸들링 등의 자동화된 작업을 수행지능형 로봇: 환경을 인식하고 스스로 판단하는 기능을 가진 로봇
   안드로이드: 사람과 닮은 모습을 한 로봇
- 1920년 체코슬로바키아의 극작가 카렐 차페크(Karel Capek)의 희곡 R.U.R(Rosuum's Universal Robits)에서 처음으로 사용
- 어원은 체코어의 노동을 의미하는 단어 'Robota'에서 유래

# 인공지능(AI)과 로봇의 관계?

#### ■ 산업현장에서의 로봇

• 인간들이 해오던 단조로운 반복 작업이나 따분한 작업, 불쾌한 작업 등을 대신





#### ■ 위험한 일을 대신하는 로봇

- 방사선 물질을 취급하거나, 유독 화학물질 등을 취급
- 너무 덥거나, 추운 환경에서 작업 등을 대신





# 인공지능(AI)과 로봇의 관계?

#### ■ 지능형 로봇

- 외부 환경을 인식하고 스스로 상황을 판단하여 자율적으로 동작하는 로봇
- 상황판단 기능과 자율동작 기능이 추가됨







<다양한 로봇 제품들 (좌로부터 KAIST 휴보, KIST 키보, 일본 아시모, 중국 안내 로봇 지아지아)>

# Tip: 인공지능(AI) 소프트웨어

#### ■ 파이썬(Python)

- 1991년 프로그래머 귀도 반 로섬(Guido van Rossum)이 발표한 언어
- 플랫폼 독립적이며 인터프리터식, 객체지향적, 동적 타이핑 대화형 언어
- 파이썬 소프트웨어 재단이 관리하고 있는 개방형, 공동체 기반의 C언어로 구현된 프로그래밍 언어
- 라이브러리가 풍부하여 대학, 교육기관, 연구기관 및 산업계에서 이용



### Tip: 인공지능(AI) 소프트웨어

#### ■ 텐서플로(TenbsorFlow)

- 다양한 작업에 대해 데이터 흐름 프로그래밍을 위한 오픈소스 소프트웨어 라이브러리
- 심볼릭 수학 라이브러리이며, 뉴럴 네트워크 같은 기계학습(Machine Learning) 응용 프로그램에 사용
- 구글 내 연구와 제품개발을 위한 목적으로 구글 브레인팀이 개발
- 2015년 11월 아파치 2.0 오픈소스 라이선스로 공개
- 파이썬 API를 제공하며, C/C++ API도 같이 제공



- 컴퓨터 등을 사용하여 인공적인 기술로 만들어낸 실제와 유사하지만 실제가 아닌 어떤 특정한 환경 이나 상황 혹은 그 기술 자체를 의미함
- 사용자의 오감(시각, 청각, 미각, 후각 그리고 촉각)을 자극하며, 실제하는 장치나 디바이스를 이용하여 가상현실 속에 구현된 것들과 상호 작용이 가능
- 프랑스의 극작가 이자 시인, 배우이며 연출가인 앙토냉 아르토(Antonin Artaud)의 저서에서 '버추얼 리얼리티(가상현실의 영문표기)'를 사용한 데서 유래
- 2015년 삼성 기어 VR, 2016년 오큘러스 리프트, HTC Vive, 플레이스테이션 VR 등의 주변기기들의 등장으로 대중의 수요와 관심이 증폭



<대표적인 가상현실(VR) 영화들>



<가상현실(VR) 체험을 위한 기기들 (SONY PlayStation VR 및 삼성 기어 VR SM-R325>

- 제시 방식에 의한 분류
  - 컴퓨터 등이 만들어 낸 가상의 세계를 제시하는 타입
    - 3D 컴퓨터 게임이 대표적인 예
  - 현실의 세계를 제시하는 타입
    - 현실 세계의 정보가 가상현실 시스템 기기를 통해 전달되는 경우
    - 증강현실(Augmented Reality, AR), 복합현실(Mixed Reality, MR)이 대표적



<슈퍼배드 게임 화면 모습>



<증강현실 기술을 활용한 하나은행 '하나머니GO' 앱의 모습>

#### ■ 시스템 환경에 따른 분류

- 몰입형 가상현실(Immersive VR)
  - HMD(Head Mounted Display), 데이터 장갑(Data Glove), 데이터 의류(Data Suit) 등의 특수 장비를 통해 사용자가 실제로 보고, 만지는 것과 같은 감각적인 효과를 느낄 수 있게 해 환경에 몰입하도록 하는 시스템
- 원거리 로보틱스(Tele-Robotics)
  - 몰입형 가상현실 시스템 + 로봇의 형태로서, 로봇을 이용하여 원 거리에 있는 공간에 사용자가 존재하는 듯한 효과를 주는 시스템
- 데스크톱 가상현실(Desktop VR)
  - 일반 컴퓨터 모니터에 간단한 입체안경, 조이스틱 등만 갖추면 책상 위에서 쉽게 만날 수 있는 가 상현실 시스템
- 삼인칭 가상현실(Third Person VR)
  - 비디오 카메라로 촬영된 자신의 모습을 컴퓨터가 만들어내는 가상의 공간에 나타나게 하여 자신이 직접 그 공간에 존재하는 것처럼 느끼게 하는 시스템
  - 주로 오락용으로 많이 적용하고 있으며 Xbox 키넥트(KINECT)가 대표적