

분석 자동화 시스템 구축

2022. 11. 9~11 (3일간, 24시간)

정 준 수 PhD

분석 자동화 시스템 구축 교육 일정

일 정	학습 내용	상세 내용
1일 차	과정 소개	분석 자동화 시스템 구축 개요 및 과정 소개
		소프트웨어 로봇과 프로세스 자동화(Robotics Process Automation, RPA) Python 개발환경 구축
	Break	
	Web 자동화	web crawling - requests beautifulsoap - selenium
	실습 예제 1	Amazon web crawling - https://github.com/JSJeong-me/Python_RPA/blob/main/web/beautifulsoup-0.ipynb
	Robot Framework	Browser library - Playwright - https://robotframework-browser.org/
	실습 예제 2	https://robocorp.com/docs/courses/beginners-course/python-robot
	실습 예제 3	Web testing - https://github.com/robotframework/WebDemo#running-demo
2일차	데스크톱 자동화	마우스 자동화 - pyautogui 키보드 자동화 - pyautogui 간단한 메세지 박스 만들기 이미지로 좌표찾기
	App 자동화	Notepad Explorer Calculator MSpaint
	Break	
	Robot Framework	Calculator 자동화 Word 자동화 Power Point 자동화
	실습 예제 1	Web과 App연동
	문서 자동화 1	PDF Excel
	Break	
3일차	문서 자동화 2	한글 문서 자동 작성
	Regular Expression	문서에서 Pattern 찾기
	데이터 분석 사례	시계열 분석 사례

과정 목표 및 개요

✓ 과정 목표

머신러닝·빅데이터 분석 단계별 Pipeline을 소프트웨어 로봇을 사용하여 데이터 수집과 분석 단계별 업무 자동화 및 소프트웨어 로봇을 활용한 효율 개선 방안 학습

✓ 과정 개요

- Web 자동구동 및 데이터 수집/입력 자동화
- App 자동구동 및 업무 수행 자동화
- 문서작성 자동화(ex. PDF, Excel, 한글문서 자동 작성)
- Python 및 Robot Framework 활용
- 프로젝트 실습

Robotic Process Automation(RPA)이란?

- ✓ Robotic Process Automation은 사용자가 PC 및 모바일 화면에서 수행하는 정형화되고 반복적인 업무를 사람의 작업을 모방하는 Smart software가 대신 수행하도록 하여 자동화 하는 Software입니다.
- ✓ 소프트웨어 로봇 - Computer coded software NOT real tangible robot
 - 대상 시스템과 보안에 간섭하거나 변화를 주지 않고 통합됨
 - 현재 사용 중인 Software(예: ERP, 엑셀, 메일 시스템 등)를 기반으로 작동함
- ✓ 사용자의 컴퓨터 화면상의 작업 흐름 모방
 - 어플리케이션을 사용하는 동작(interaction)을 녹화하고 자동화함
 - 사용자의 처리 방식과 동일하게 어플리케이션의 User Interface를 작동시킴

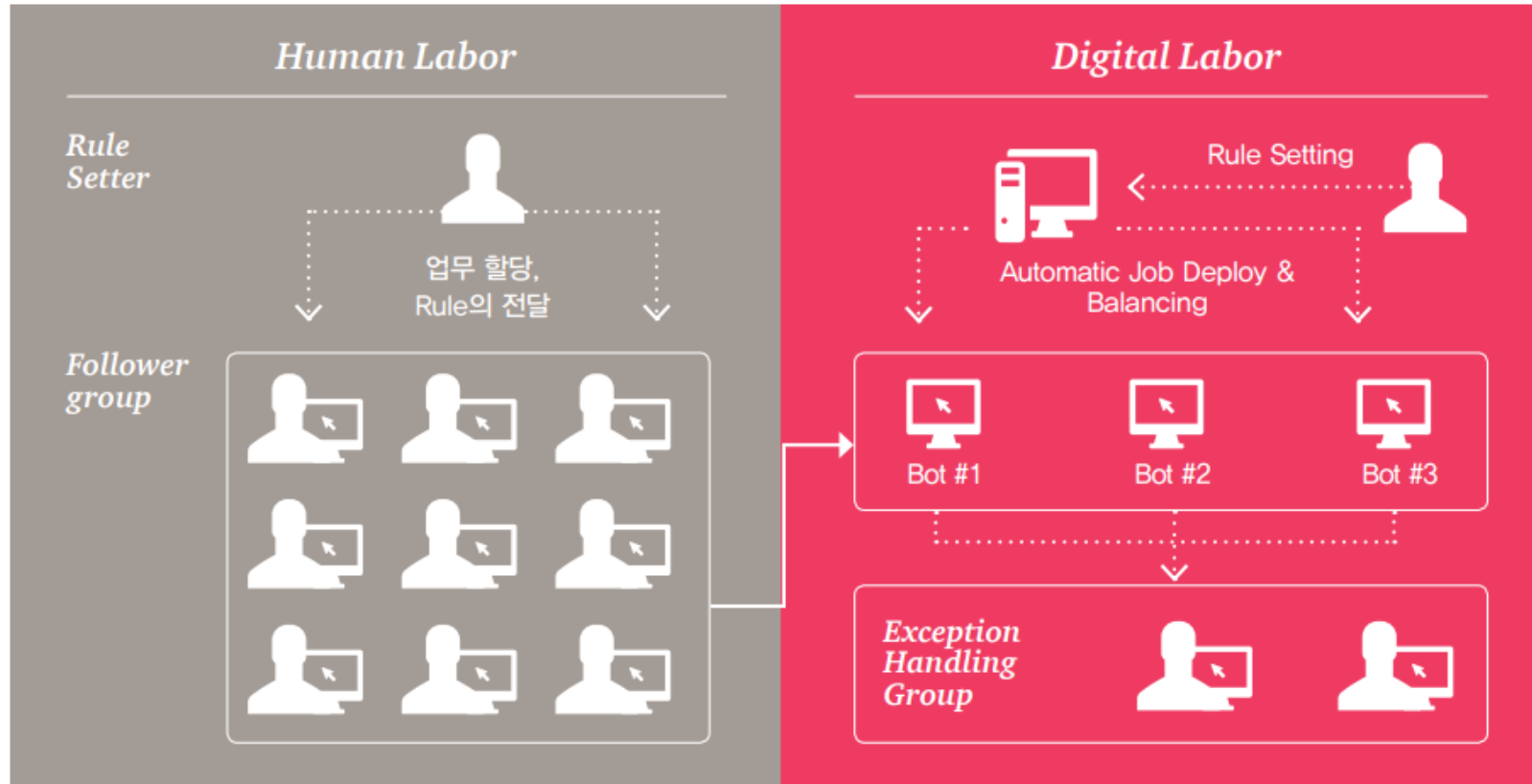
✓ **Cross-functionally and cross-applications 간 작업 수행**

- 특정 기술에 종속되지 않으며 어떠한 어플리케이션과도 함께 사용 가능함
(e.g. ERP, DB, MS Suite, ASCII file, structured PDF, thin clients such as Citrix)
- 자동화 Script와 프로세스 관리를 위해 Central Repository를 사용함

✓ **반복적이고 정해진 규칙으로 프로세스 자동화 가능**

- 동적인 Decision Point/분기점과 순환반복 고리를 가진 Workflow를 구성함
- 자동화 script의 재사용이 가능함

RPA 도입 효과



- 사람이 컴퓨터에서 수행하는 단순 반복적인 작업을 Software가 대신 수행
- 기존 다수의 인력을 1대의 로봇으로 대체

RPA 적용 대상 업무



비용 절감

On-shore > Off-shore > Robot



보안(Security)

데이터 보안 및 Audit trail 관리



업무생산성

24hr/7days 작동, 빠른 처리 속도



Scalability

업무량 증가, 변동에 유연한 대응



업무품질(정확도)

Human Error 제거



Compliance

정해진 규칙에 따라서만 업무를 처리

RPA 적용 가능 영역

- ✓ 현재 '구조화된 데이터의 Input, 정해진 규칙에 따른 처리' 중심으로 적용되고 있으나, Data 인식과 예외사항 처리 능력 발달에 따라 적용범위와 영향력 확대 적용
- OCR(Optical Character Reader) 기술적용을 통한 서류 인식 및 처리
- 자가 학습을 통한 예외 사항 처리규칙의 지속적 개선(AI, Machine Learning 등)
- 머신러닝 데이터 수집 · 분석 · 배포 – Pipeline 자동화
- 룰 기반의 사용자 분석(User Behavior Analytics)
- 네트워크 보안 - SOAR (Security Orchestration, Automation and Response)

RPA 적용 가능 영역



HR Services

- 인사절차
- 근태관리
- 인사행정
- 복리후생
- 대외보고
- 급여
- 데이터 입력
- 교육
- 채용



IT

- 시스템 설치
- FTP 다운/업로드, 백업
- 파일관리
- 이메일 관련작업
- 폴더 동기화
- 배치 작업
- 서버 모니터링
- Application 모니터링



Supply chain

- 재고관리
- 수요/공급계획
- 계약 관리
- 건적 관리
- 작업 지시관리
- 수송관리
- 반품처리



Finance

- 회계전표/증빙관리
- 계정 대사
- 채무관리, 지급
- 고객 청구, 채권관리
- 출장 등 경비 정산
- 수익/비용 배부
- 고정자산 회계



Others

- 문제 해결
- 문서 및 자료관리
- 일정 관리
- Workflow 관리
- CRM
- PC 통합
- 부정적발
- 테스트
- 문서 형식 편집

RPA의 장점

✓ 구현의 신속성과 용이성

- 대규모 ERP 및 타 IT 시스템 투자 대비 소규모 투자
- 신속한 Go-Live(월 단위가 아닌 주 단위)
- 시스템 설계에 많은 노력이 소요되지 않음

✓ 제약 받지 않는 시스템 구조

- Non-invasive: 기존 IT 시스템의 변경이나 신규 인터페이스 개발 불필요
(기존 시스템 기반 위에서 작동)

✓ Cross-function 적용 가능

- 여러 종류의 시스템, 어플리케이션, 커뮤니케이션 수단 간 인터페이스

Web 자동화 – 자동화의 원리 이해

크롤링에는 대표적으로 세 가지 라이브러리

- requests or urllib(beautifulsoup)
- Selenium
- Scrapy

https://github.com/JSJeong-me/Python_RPA/tree/main/web

App 자동화 – 자동화의 원리 이해

App 자동화 예제

- Notepad
- Explorer
- MSpaint
- Calculator
- PPT

https://github.com/JSJeong-me/Python_RPA/tree/main/pywinauto

문서처리 자동화 – 자동화의 원리 이해

문서 자동화

- Pdf
- Excel
- 한글파일
- Regular expression

https://github.com/JSJeong-me/Python_RPA/tree/main/excel

Robot framework

Robot Framework는 일반적인 오픈 소스 자동화 프레임워크입니다.
테스트 자동화 및 RPA(로봇 프로세스 자동화)에 사용할 수 있습니다.

- **Browser**
- **Desktop**
- **Word**
- **Excel**
- **PDF**

[**https://github.com/robocorp/rpaframework**](https://github.com/robocorp/rpaframework)

Robot framework - examples

<https://robotframework.org/#getting-started>

Robot framework – WebDemo

[**https://github.com/robotframework/WebDemo**](https://github.com/robotframework/WebDemo)

별첨: Python 프로그램 언어와 데이터 추상화의 이해

컴퓨터 프로그램이란?

알고리즘의 추상화

★ 생각해 보기

질문)

- 1 프로그램이 먼저 나왔을까요?
- 2 컴퓨터가 먼저 나왔을까요?

객체 지향 프로그래밍 object-oriented programming (OOP)

OOP focuses on the objects that developers want to manipulate rather than the logic required to manipulate them.

Class(클래스)

객체를 정의해 놓은 것
객체를 정의하는 틀 또는 설계도





Object(객체)

클래스의 인스턴스
설계도를 통해 만들어진 것

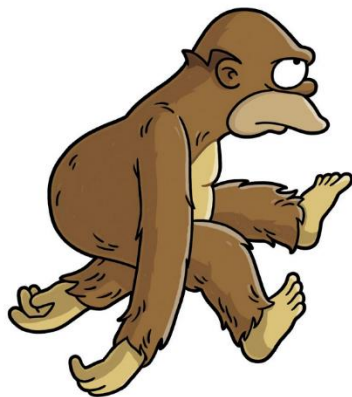
함수형 언어(Functional program)란?

수학적 기호

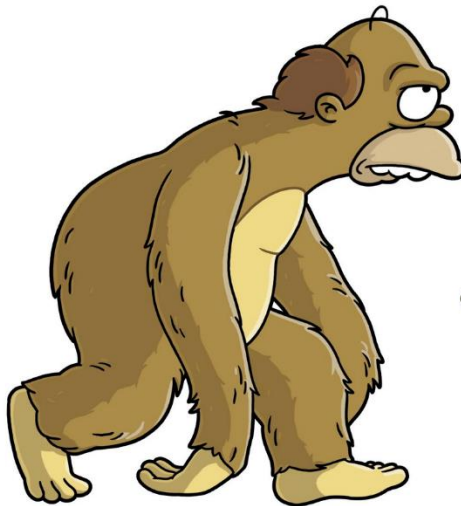
NLP



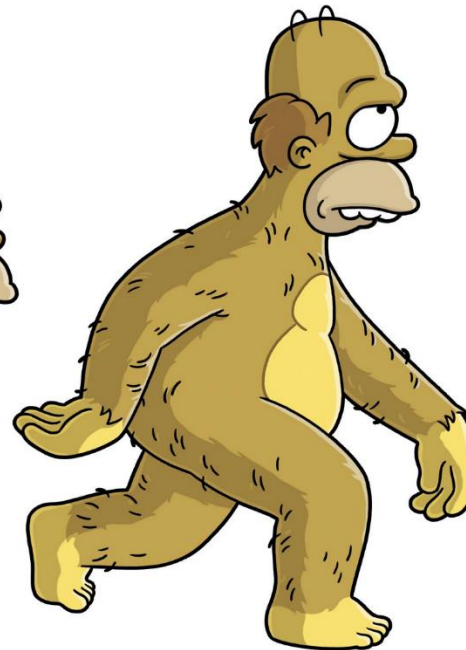
MACHINE



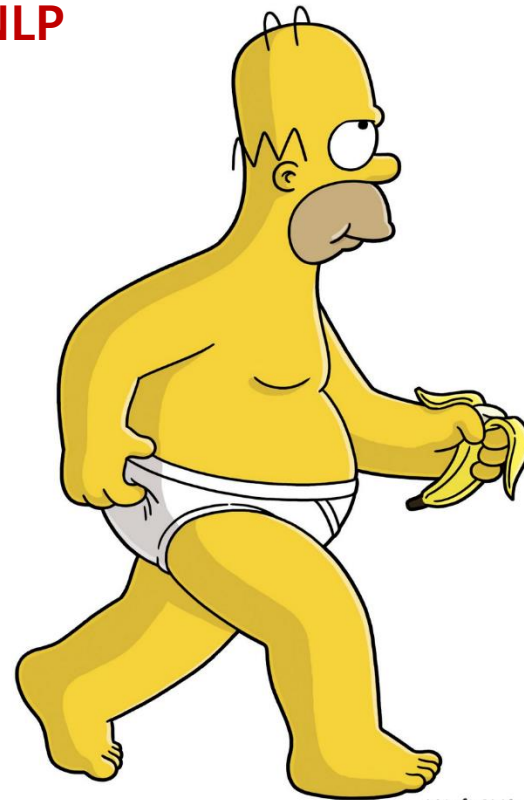
ASSEMBLY



PROCEDURAL



OBJECT ORIENTED



FUNCTIONAL

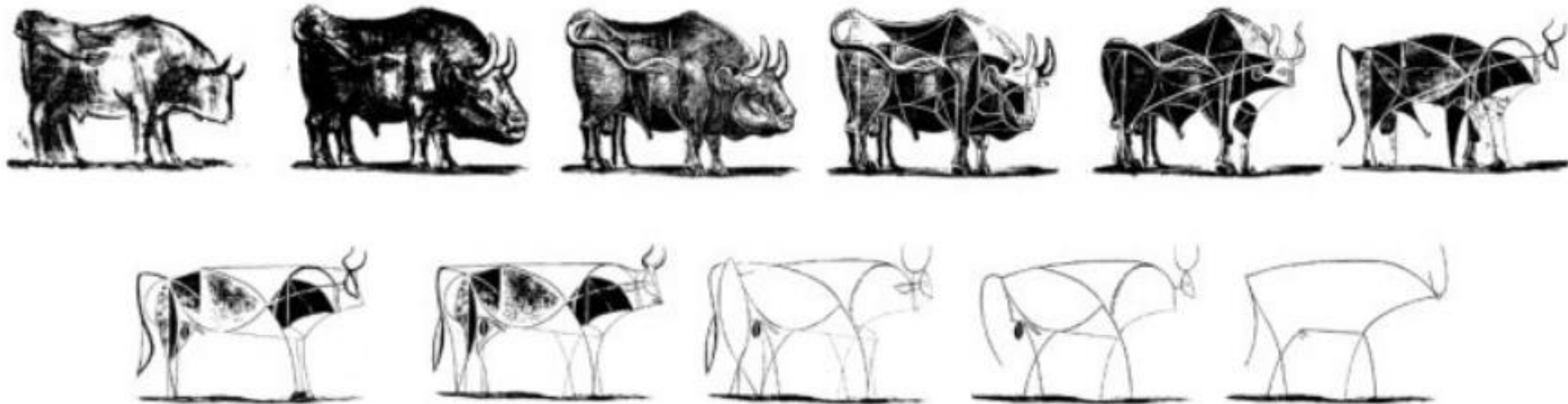
1950년대

1980년대

추상화(Abstraction)

컴퓨터 과학에서 추상화(abstraction)는 복잡한 자료, 모듈, 시스템 등으로부터 핵심적인 개념 또는 기능을 간추려 내는 것을 말한다.

The Evolution of Picasso Bull



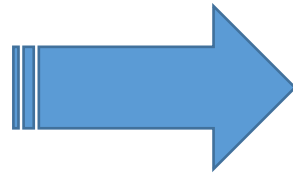




현금(Money)의 추상화(Abstraction) 사례



현금(Money)



Data



인터넷 뱅킹

Features: 교환가치 + 은행의 기능



신용카드

Features: 교환가치 + 신용



티머니

Features: 교환가치 + 버스 지하철 갈아탈 수 있다

현금(Money)의 추상화(Abstraction) – Featuring 해보기



현금과 다른 특징 분석



추상적 자료형(abstract data type, ADT)

컴퓨터 과학에서 자료들과 그 자료들에 대한 연산들을 명기한 것이다

1. 구조(Structure)
2. 연산(Operation)

Python 자료형이란?

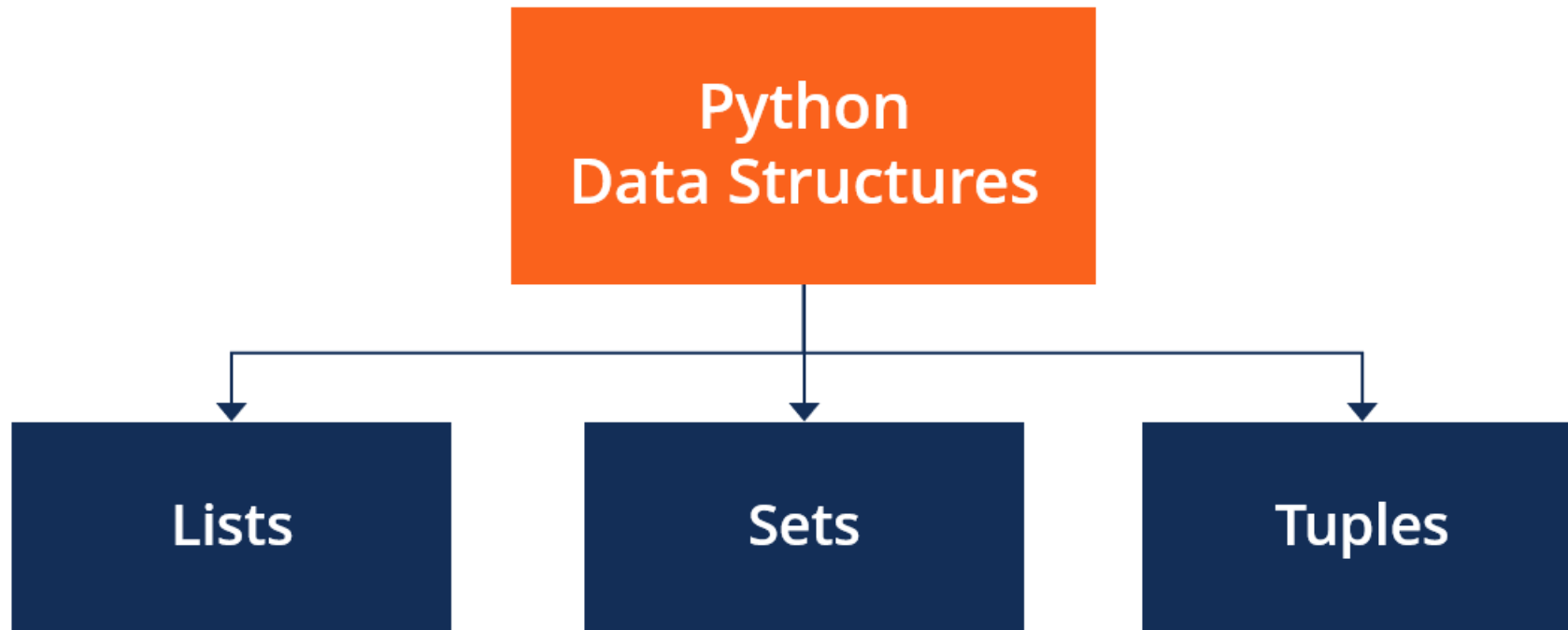


- 일반적으로 프로그램은 연산을 수행하며, 연산을 하기 위해서는 값들을 임시로 혹은 영구적으로 어딘가에 저장해 놔야 함
- 이러한 값들을 저장하는 저장소를 변수라고 함
- 변수의 유형을 자료형이라고 함

Python 자료의 종류

- 수치 - int float complex 등
- 문자형
- 리스트 - 리스트는 쉽게 값들의 나열
- 세트 - 집합의 개념으로 순서가 없음
- 튜플 - tuple은 리스트와 유사하나, 읽기 전용임
- 딕셔너리 - 키와 값의 쌍으로 이루어져 있음
- 부울(bool) - 참(True)과 거짓(False)을 나타내는 자료형

Abstract Data Type (ADT) In Python



정 준 수 / Ph.D (jsjeong@hansung.ac.kr)

- 前) 삼성전자 연구원
- 前) 삼성의료원 (삼성생명과학연구소)
- 前) 삼성SDS (정보기술연구소)
- 現) (사)한국인공지능협회, AI, 머신러닝 강의
- 現) 한국소프트웨어산업협회, AI, 머신러닝 강의
- 現) 서울디지털재단, AI 자문위원
- 現) 한성대학교 교수(겸)
- 전문분야: Computer Vision, 머신러닝(ML), RPA
- <https://github.com/JSJeong-me/>

