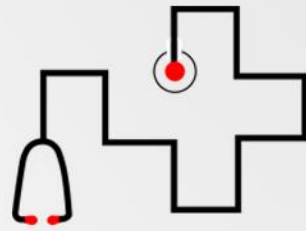




# 1. 프로젝트 소개

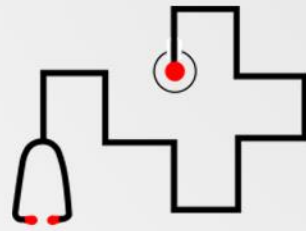
---



- 1) Project Name: **MediChA.I.n**: Medical + BlockChain + A.I.
- 2) Description:
  - 1) 개인 의료 데이터를 블록체인에 저장하고 체인을 통해 접근할 수 있는 이미지를 학습하는 머신 러닝 시스템을 만들어 새로 들어오는 의료 이미지 데이터의 신뢰성을 계산하고 질병 발생 여부를 예측하여 새로운 블록을 생성해 체인을 구성하는 시스템.
  - 2) 초기에는 공공 의료 이미지 데이터로 학습을 진행하고 학습이 완료되면 입력되는 데이터를 블록이 채굴되는 과정에서 분석 및 예측하여 블록에 넣어서 기존 블록체인에 연결해준다.
  - 3) 개인은 의료 데이터 중 본인의 의료 데이터만 접근, 조회 가능하며 의료진은 환자의 데이터를 모두 볼 수 있는 권한을 가졌다.

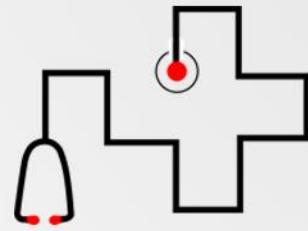
## 2. SW / HW

---

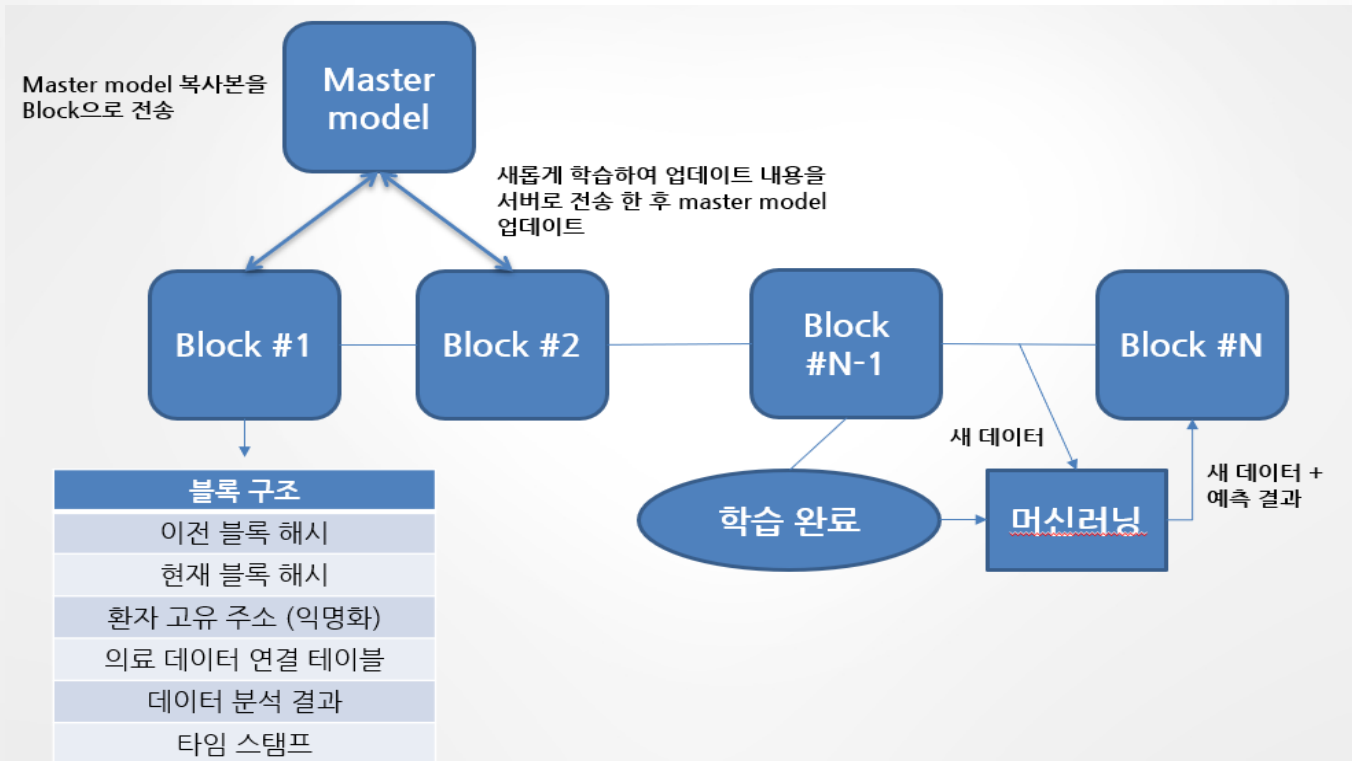


- 1) 새로 만들 SW:
  - 1) 의료 데이터를 보관할 수 있는 블록체인
  - 2) 머신 러닝 모델
  - 3) 간단한 프론트 엔드 사이트
  
- 2) COST SW: 블록체인 플랫폼, Tensorflow
  
- 3) COST HW: 서버 컴퓨터, 클라우드 컴퓨팅 시스템 (AWS EC2, Database 등)

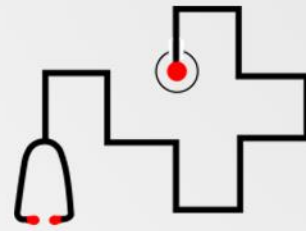
# 3. 최종 산출물



- 1) 정의: 기존 의료 이미지 데이터로 학습된 머신러닝 모델이 새로운 의료 이미지 데이터를 분석하고 그 결과를 포함하여 새로운 블록을 생성하는 시스템
- 2) 구조도와 방식



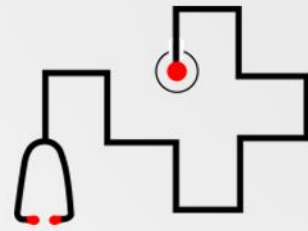
# 4. Alternative Solutions



- 1) **Vytalyx**: 의료데이터를 저장하고 분석하는 데 블록체인과 AI를 사용할 계획을 가진 의료 기술 회사 (URL: <https://vytalyx.io/>)
  
- 2) **MediBlock**: 여러 곳에 분산되어 있는 의료정보를 블록체인을 이용하여 개인이 안전하게 관리할 수 있는 시스템, 블록체인 기술을 이용하여 기존 의료정보 시스템을 환자 중심의 의료정보 시스템으로 전환시키고 개인정보의 탈중앙화를 이루는 것을 목표로 하는 회사(URL: <https://medibloc.org/ko>)
  
- 3) **Reference**:  
Blockchain for AI Review. IEEE Access. (Kaled Salah, Muhammad habib ur Rehman, Nishara Nizamuddin)
  
- 4) **Datasets**:  
histology: The Cancer Imaging Archive  
colon polyps: 강동 경희대학병원 소화기내과, mv-csv(<http://mv.cvc.uab.es/>) 등  
MRI, X-ray: Kaggle(Kaggle.com)

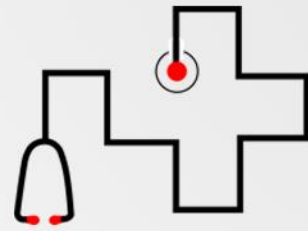
# 5. Project Justification

---



- 1) 현 상황: 블록체인을 이용해 의료 데이터를 저장하려는 의도는 많은 곳에서 일어나고 있지만, 아직 실용성이 뚜렷하지 않은 상황이다.
- 2) 차이점 도출
  - 1) 이상한 데이터(상이한 신체 부위 또는 의료 데이터가 아닌 경우)가 들어올 경우 머신러닝으로 거르거나, 경고 메시지를 보내줄 수 있음.
  - 2) 블록 별로 랜덤으로 Validation Set을 지정하는 등 블록에 저장된 이미지에 따라 다른 알고리즘을 세울 수 있다.
  - 3) 공공 데이터 사용을 통한 초기 블록 데이터 구축을 할 수 있다.
  - 4) 블록 생성 알고리즘 변경 (하드포크) 를 통해 머신러닝이 학습이 완료된 시점 부터의 블록에 대한 조건을 변경할 수 있다.

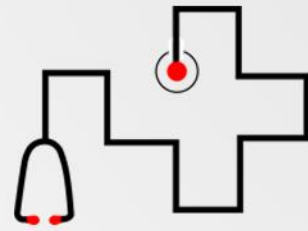
# 6. Risk Analysis & Reduction Plan



- 1) 데이터 보관을 할 수 있는 블록체인 구축에 대한 고민
  - 직접 구축할 것인지 또는 플랫폼을 사용할 것인지에 대한 고민
  - 다양한 프라이빗 블록체인 플랫폼 중 자료가 많고 활용도가 많은 '하이퍼레저 패브릭' 공부 및 구축 실습을 통한 테스트
  
- 2) 블록체인에 이미지를 담을 경우 잘 돌아갈 것인지 고민, 데이터 크기 고려
  - 데이터를 압축하는 방식을 고려해봄
  - 이미지를 직접 담지 않고, 다른 방식을 고려해봄. (링크 연결 또는 IPFS 와 같은 분산 시스템 사용)
  
- 3) 딥러닝 모델 선정에 대한 고민 사항
  - 객체 검출 방법에 따라서 model의 종류가 달라짐.
    - 데이터의 종류에 따라 적합한 객체 검출 방법에 대해 대학병원 교수님께 자문을 구할 예정
  - Model 종류에 따라 성능이 모두 다름.
    - 다양한 종류의 model 을 훈련 및 검증 후 model 선정 및 ensemble

# 6. Risk Analysis & Reduction Plan

---



## 4) 이미지에 대한 고민

- 종류에 대한 고민
  - 블록체인에 올라갈 수 있어야 하기 때문에 크기가 너무 크면 안됨.
- 이미지 처리에 대한 고민
  - 분리 저장된 이미지를 합쳐서 학습할 수 있는 이미지로 만들어야 함.

## 5) 학습을 멈추고 예측을 하는 시점

- Block 별로 Validation, train 블록을 지정하여 train
  - Train loss와 validation loss를 모니터링 하여 중단 시점을 정하는 것이 중요.

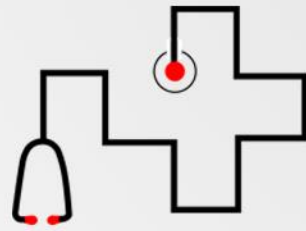
## 6) Visualization (프론트 엔드)

- 간단한 Html, css, Javascript 공부: 데이터 조회, 데이터 입력하는 경우의 화면상 시각화 필요
- 새 블록이 예측 결과와 함께 올라가는 것을 보여줄 수 있는 정도.



# 7. Success Criteria

---



- 1) 블록체인에 성공적으로 의료 데이터가 담기는 것을 확인
- 2) 블록체인의 정합성 확인 - 이전 해시가 현재 블록 해시에 잘 기록이 되어 있는지, 이전 블록 변조 시 체인의 정합성이 깨짐이 확인 되는지
- 3) 머신러닝을 통해서 의료 이미지에 대한 학습과 결과 확인
- 4) 데이터 Set 학습 결과 확인