



## **데이터 라벨링 보상 플랫폼**

**6조 T-San(티끌모아 태산)**

**이정하(팀장), 박상일, 박지희, 윤여환, 이다은, 장태진**

# 목차

---

1. 역할 분담
2. 프로젝트 목표
3. 연구 내용
4. 수정 사항
5. 현재 진행 사항 및 시연 영상
6. 향후 추진 계획
7. 개발 로드맵

# 역할 분담



**이정하 (20171677)**

- 팀장  
- Front-end



**박상일 (20152818)**

- ML이용한 데이터 분석 및 시각화



**박지희 (20162772)**

- 블록체인



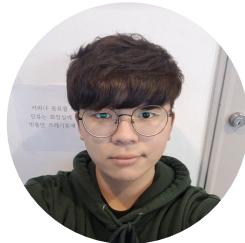
**이다은 (20171661)**

- Back-end



**윤여환 (20171658)**

- Front-end



**장태진 (20171695)**

- 블록체인  
- 문서 벡터화 및 분석 알고리즘 제작

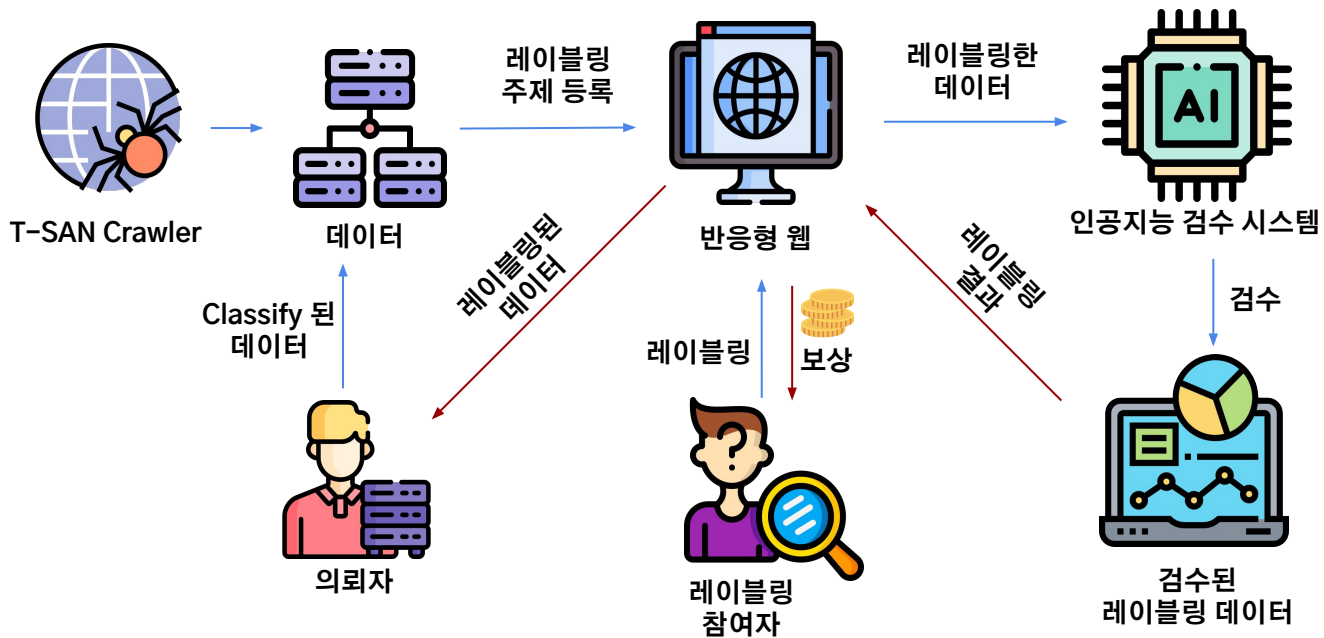
# 프로젝트 목표

---

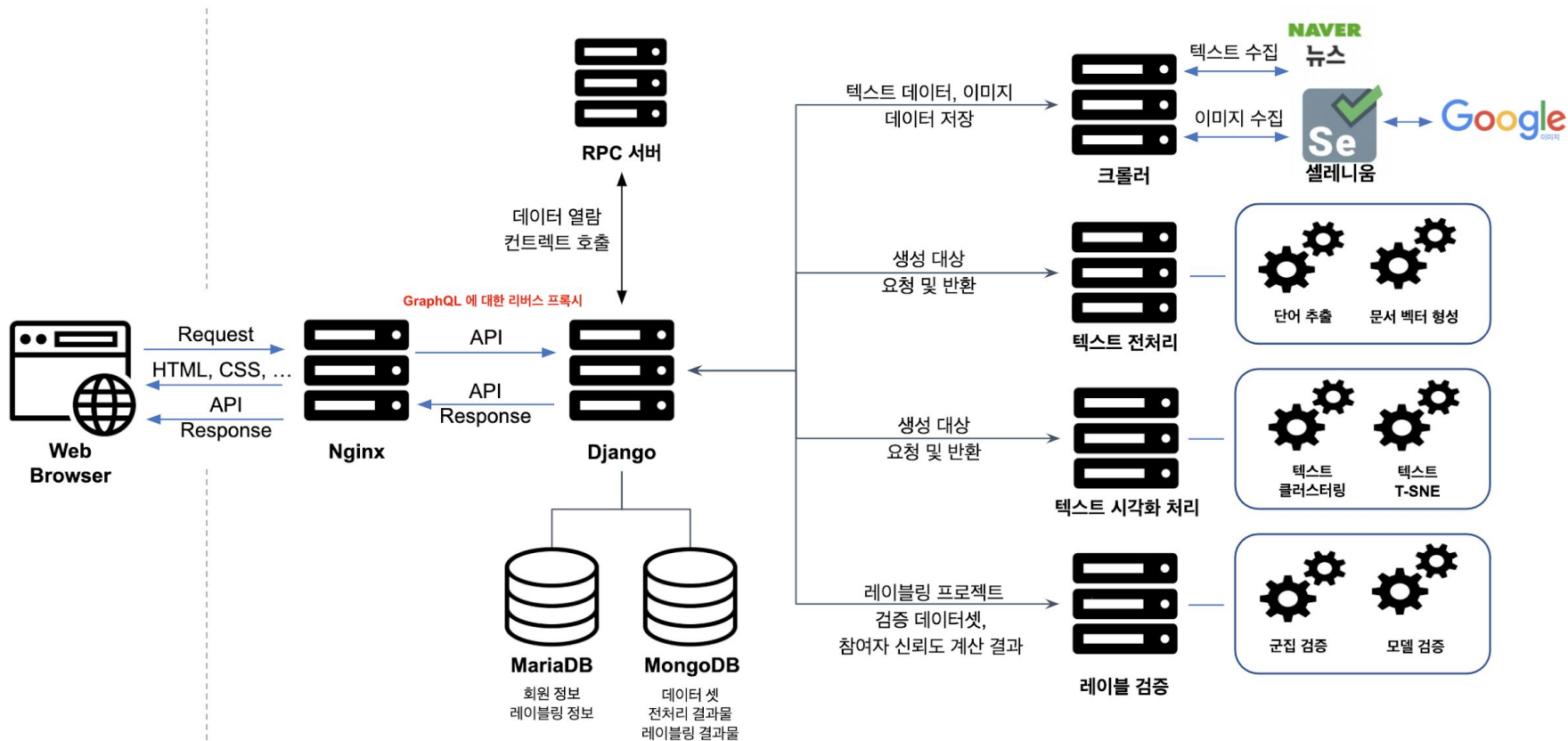


의뢰자에게 받은 데이터, T-SAN이 크롤링한 데이터를 레이블링하여  
**유의미한 정보를 생성**하는 플랫폼을 개발하는 것을 목표로 한다.

# 프로젝트 목표



# 프로젝트 구성도



# 수정 사항

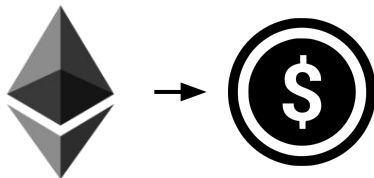
## [ 공공데이터 제공 ]



공공 데이터 레이블링  
프로젝트를 통한 오픈

데이터 플랫폼으로 발전

## [ 보상 제도를 포인트로 확정 ]



보상 제도를 토큰과  
포인트에서 블록체인을  
이용한 포인트로 확정

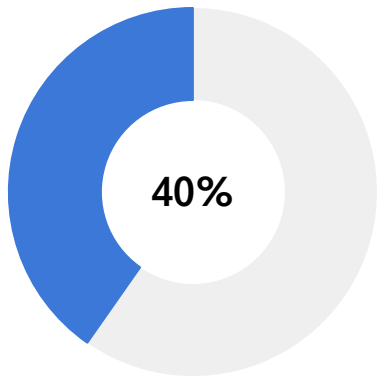
사용자 접근성 향상  
투명한 포인트 관리

## [ CAPTCHA 서비스 제공 ]



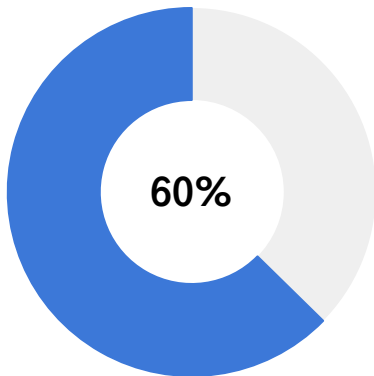
타 웹사이트에서 인증방식으로  
사용할 수 있도록 CAPTCHA  
서비스 제공

# 현재 진행 사항



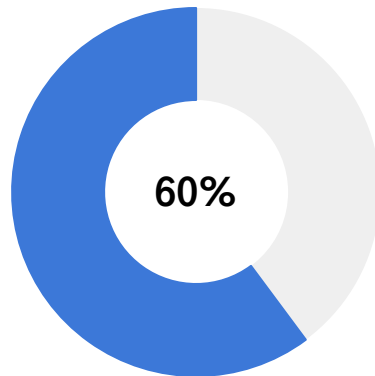
## 레이블링 검증

- ✓ 이미지 분류 모델 Classifier 학습
- ✓ 텍스트 단어 추출 및 문서 벡터화
- ✓ 신뢰도 기반 확률적 임시 레이블링
- ✓ 신뢰도 계산 알고리즘
- ✓ 2차 레이블링 요청 및 검증



## 블록 체인

- ✓ Smart Contract 작성
- ✓ Smart Contract 배포
- ✓ 로컬 프라이빗 네트워크 생성
- ✓ WEB3를 이용한 서버와 연동
- ✓ Testnet RPC 사용



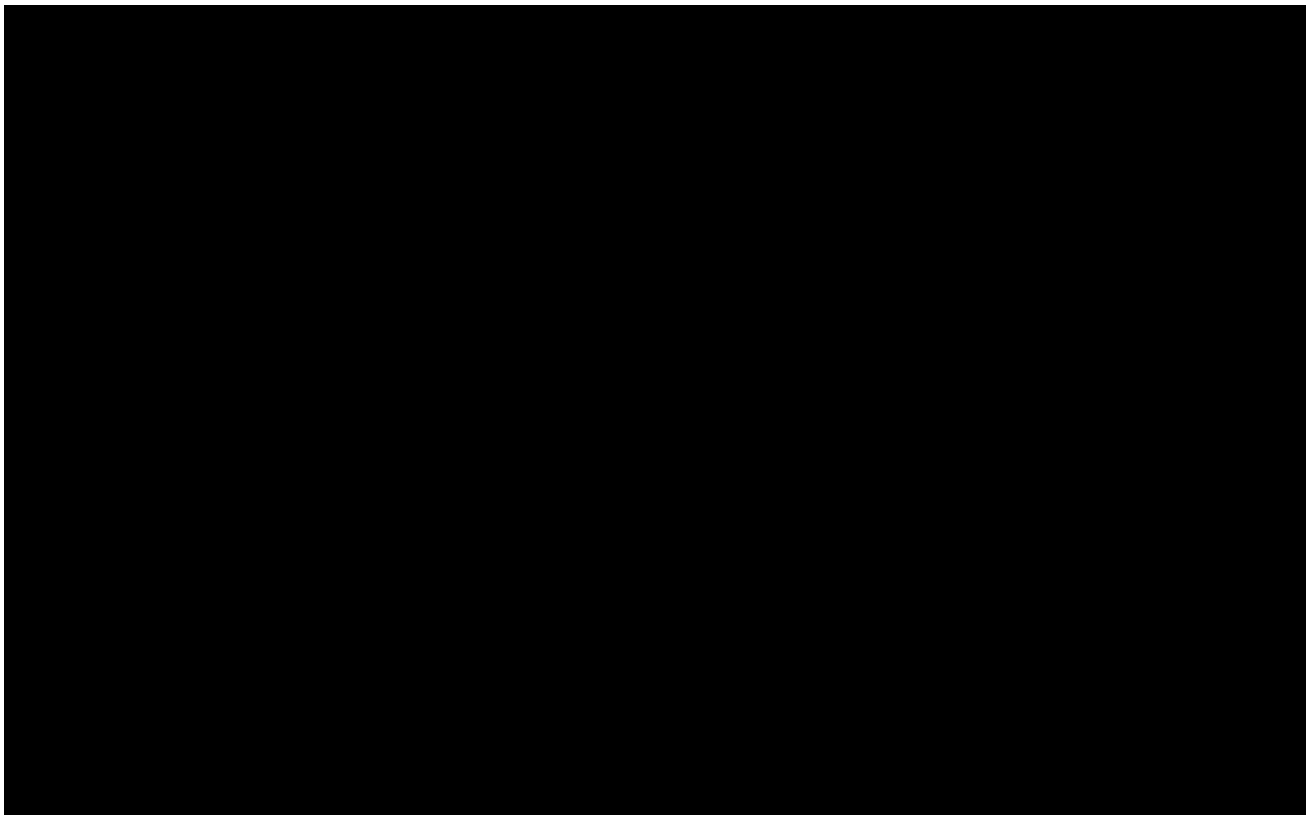
## 웹 어플리케이션

- ✓ UX/UI 디자인 및 구현
- ✓ 웹 서버 구축
- ✓ DB 구축
- ✓ JWT 구현
- ✓ Front-Back API 연동

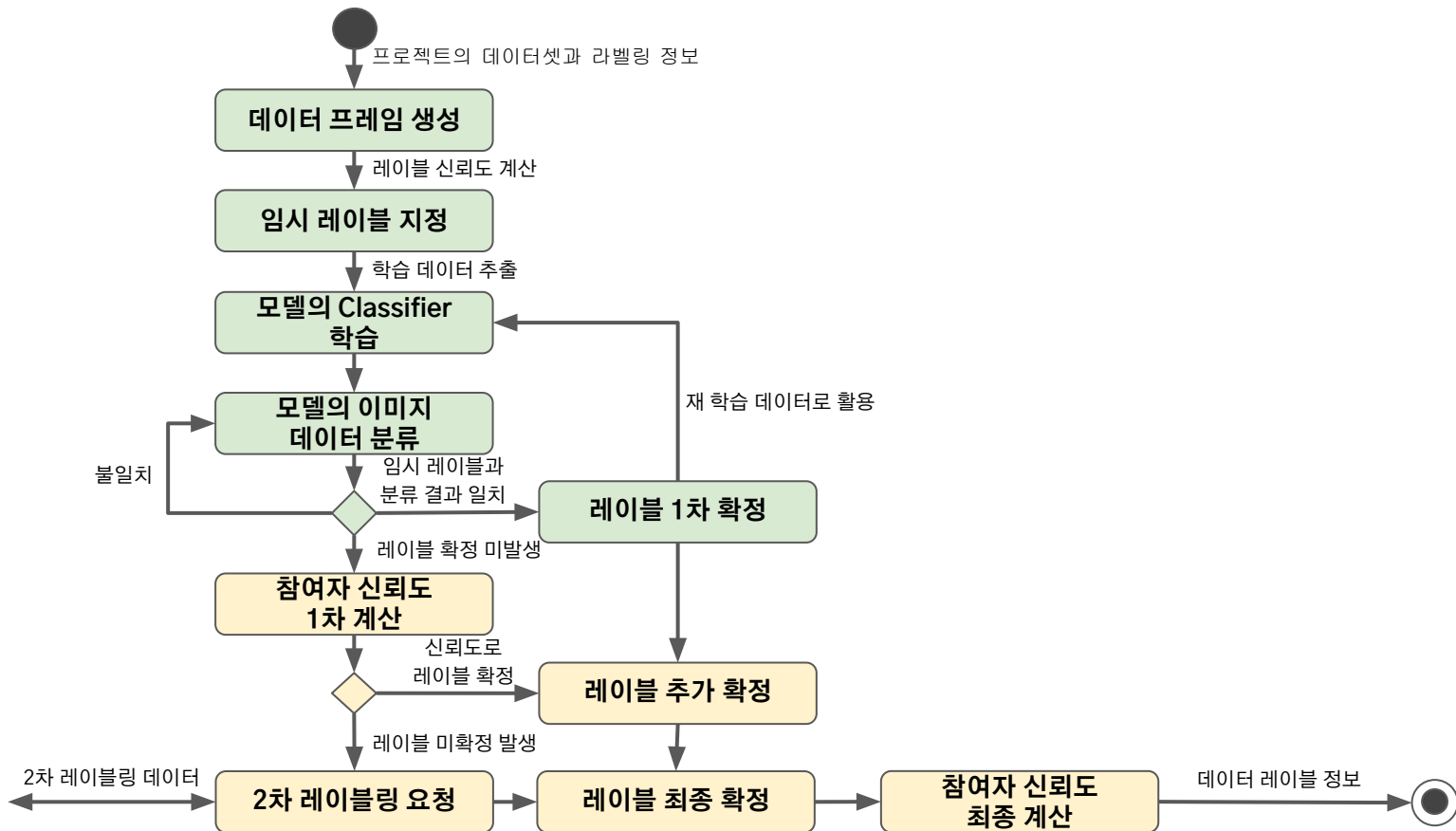


# 시연 영상

---



# 현재 진행 사항: 이미지 레이블 검증 모듈



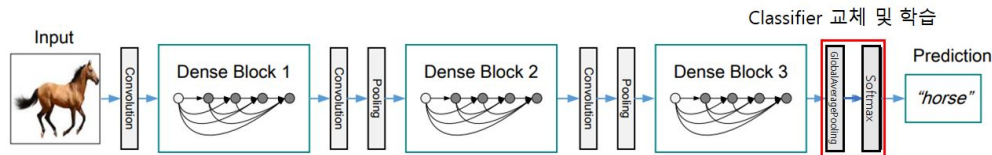
# 현재 진행 사항: 이미지 레이블 검증 모듈

project_id	file_index		path	label	credibility	id
0	0.0	0.0	../furniture-images/val/bed/00000901.jpg	bed	0.92	a
1	0.0	0.0	../furniture-images/val/bed/00000901.jpg	bed	0.83	b
2	0.0	0.0	../furniture-images/val/bed/00000901.jpg	bed	0.85	c
3	0.0	0.0	../furniture-images/val/bed/00000901.jpg	sofa	0.75	d
4	0.0	0.0	../furniture-images/val/bed/00000901.jpg	sofa	0.52	e
5	0.0	1.0	../furniture-images/val/chair/00000011.jpg	chair	0.92	a
6	0.0	1.0	../furniture-images/val/chair/00000011.jpg	swivelchair	0.83	b
7	0.0	1.0	../furniture-images/val/chair/00000011.jpg	chair	0.85	c
8	0.0	1.0	../furniture-images/val/chair/00000011.jpg	chair	0.75	d
9	0.0	1.0	../furniture-images/val/chair/00000011.jpg	swivelchair	0.52	e
10	0.0	2.0	../furniture-images/val/sofa/00000203.jpg	sofa	0.92	a
11	0.0	2.0	../furniture-images/val/sofa/00000203.jpg	sofa	0.83	b
12	0.0	2.0	../furniture-images/val/sofa/00000203.jpg	sofa	0.85	c
13	0.0	2.0	../furniture-images/val/sofa/00000203.jpg	bed	0.75	d
14	0.0	2.0	../furniture-images/val/sofa/00000203.jpg	bed	0.52	e
15	0.0	3.0	../furniture-images/val/swivelchair/00000015.jpg	swivelchair	0.92	a
16	0.0	3.0	../furniture-images/val/swivelchair/00000015.jpg	chair	0.83	b
17	0.0	3.0	../furniture-images/val/swivelchair/00000015.jpg	chair	0.85	c
18	0.0	3.0	../furniture-images/val/swivelchair/00000015.jpg	swivelchair	0.75	d
19	0.0	3.0	../furniture-images/val/swivelchair/00000015.jpg	swivelchair	0.52	e
20	0.0	4.0	../furniture-images/val/table/00000379.jpg	table	0.92	a
21	0.0	4.0	../furniture-images/val/table/00000379.jpg	table	0.83	b
22	0.0	4.0	../furniture-images/val/table/00000379.jpg	table	0.85	c
23	0.0	4.0	../furniture-images/val/table/00000379.jpg	table	0.75	d
24	0.0	4.0	../furniture-images/val/table/00000379.jpg	bed	0.52	e

[데이터 프레임 생성 예시]

file_index	path	label_temp
0	../furniture-images/val/bed/00000901.jpg	bed
1	../furniture-images/val/chair/00000011.jpg	chair
2	../furniture-images/val/sofa/00000203.jpg	sofa
3	../furniture-images/val/swivelchair/00000015.jpg	chair
4	../furniture-images/val/table/00000379.jpg	table

[임시 레이블 지정 예시]



[모델 구조와 Classifier 교체]

### [모델의 Classifier 학습 과정]



# 현재 진행 사항: 이미지 레이블 검증 모듈

	path	label	predicted_label
0	../dataset1/train/man/face_1000.jpg	man	0
1	../dataset1/train/man/face_1001.jpg	man	0
2	../dataset1/train/man/face_1003.jpg	man	0
4	../dataset1/train/man/face_1005.jpg	man	0
5	../dataset1/train/man/face_1006.jpg	man	0
...	...	...	...
2247	../dataset1/test/woman/face_67.jpg	woman	0
2252	../dataset1/test/woman/face_71.jpg	woman	0
2255	../dataset1/test/woman/face_74.jpg	woman	0
2256	../dataset1/test/woman/face_76.jpg	woman	0
2259	../dataset1/test/woman/face_8.jpg	woman	0

	path	label	predicted_label
3	../dataset1/train/man/face_1004.jpg	man	1
8	../dataset1/train/man/face_1011.jpg	man	1
43	../dataset1/train/man/face_1059.jpg	man	1
166	../dataset1/train/man/face_1217.jpg	man	1
172	../dataset1/train/man/face_1231.jpg	man	1
...	...	...	...
2275	../dataset1/test/woman/face_95.jpg	woman	1
2276	../dataset1/test/woman/face_96.jpg	woman	1
2277	../dataset1/test/woman/face_97.jpg	woman	1
2278	../dataset1/test/woman/face_98.jpg	woman	1
2279	../dataset1/test/woman/face_99.jpg	woman	1

	path	label	predicted_label
0	../dataset1/train/man/face_1004.jpg	man	1
1	../dataset1/train/man/face_1011.jpg	man	1
2	../dataset1/train/man/face_1059.jpg	man	1
3	../dataset1/train/man/face_1217.jpg	man	1
4	../dataset1/train/man/face_1231.jpg	man	1
...	...	...	...
96	../dataset1/test/woman/face_67.jpg	woman	0
97	../dataset1/test/woman/face_71.jpg	woman	0
98	../dataset1/test/woman/face_74.jpg	woman	0
99	../dataset1/test/woman/face_76.jpg	woman	0
100	../dataset1/test/woman/face_8.jpg	woman	0

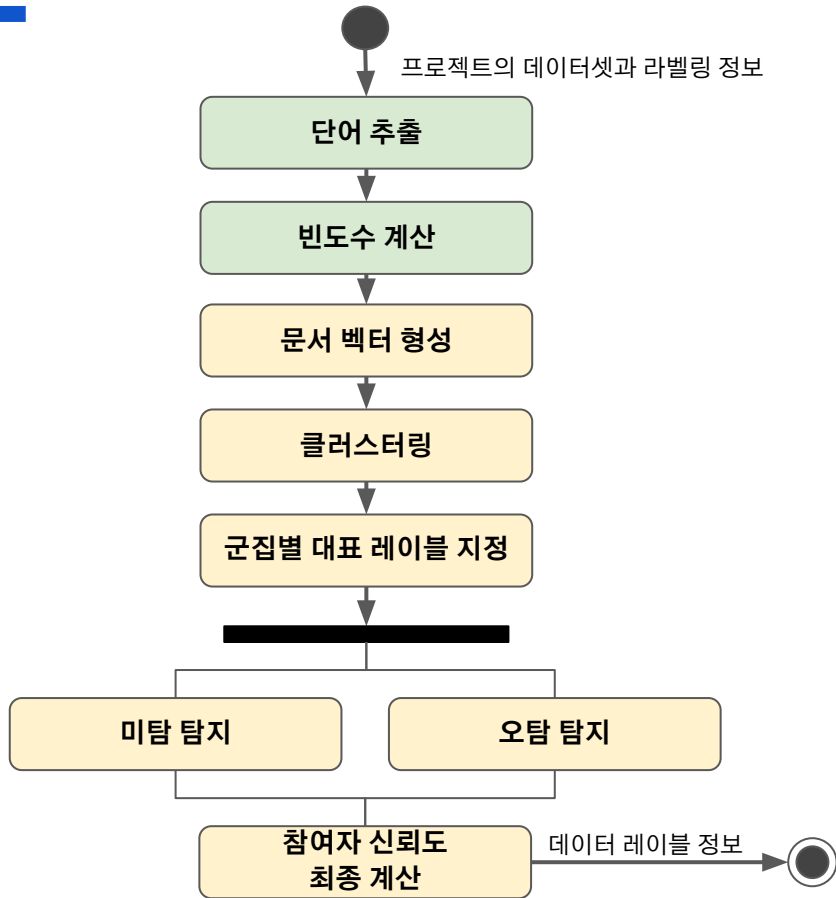
[('man', 1121), ('woman', 82)]  
man

[('woman', 1058), ('man', 19)]  
woman

[모델의 데이터 분류]

[오분류 데이터]

# 현재 진행 사항: 텍스트 레이블 검증 모듈



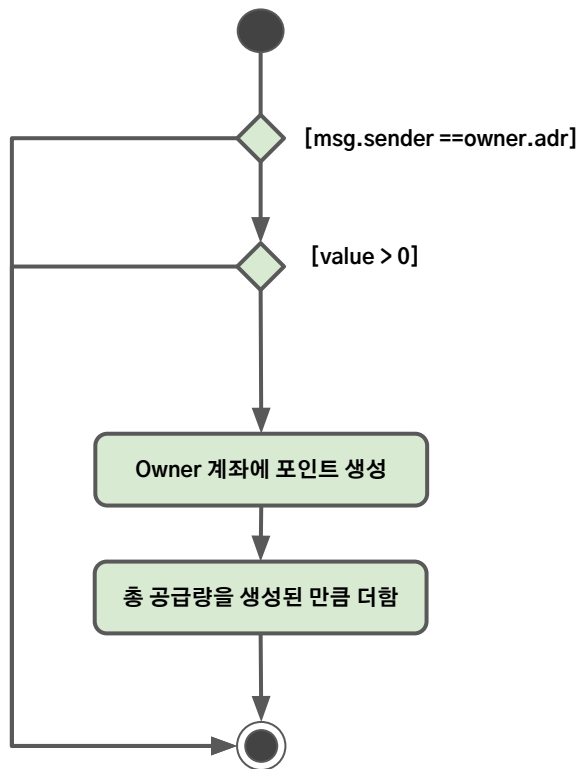
# 현재 진행 사항: 포인트 보상



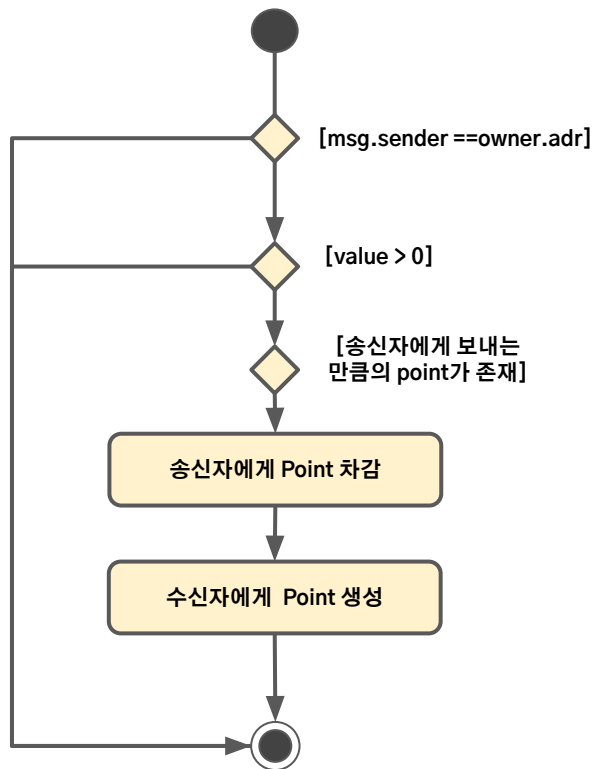
- ERC20 표준 토큰 인터페이스 기반
- 포인트 생성, 전송, 소멸 등의 기능 추가
- 기존 토큰 방식은 사용자 모두가 Ethereum Address를 의무적으로 보유
- 새롭게 도입한 방식은 서비스 적용과 익명성을 지키기 위해 이더리움 주소가 아닌 사용자 고유 번호로 처리

# 현재 진행 사항: Solidity 소스 코드 주요 기능

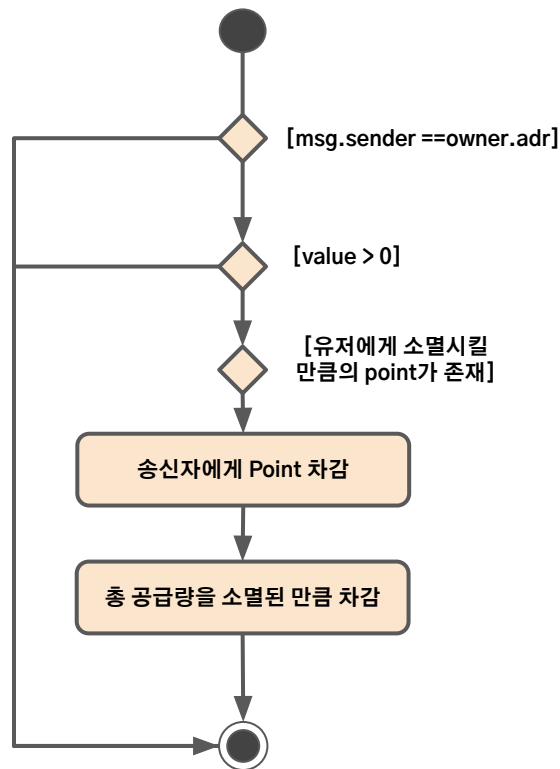
[ 포인트 생성 ]



[ 포인트 전송 ]



[ 포인트 소멸 ]





# 현재 진행 사항: geth를 이용한 네트워크 구성

```
il> DAO: <nil> DAOSupport: false EIP150: <nil> EIP155: <nil> EIP158: <nil>}
[0419 21:20:45.243228 core/blockchain.go:219] Last header: #75 [aa8759c6...] TD=10020026
[0419 21:20:45.243261 core/blockchain.go:220] Last block: #75 [aa8759c6...] TD=10020026
[0419 21:20:45.243279 core/blockchain.go:221] Fast block: #75 [aa8759c6...] TD=10020026
[0419 21:20:45.245580 p2p/server.go:342] Starting Server
[0419 21:20:45.245726 p2p/server.go:610] Listening on [::]:30303
[0419 21:20:45.248789 node/node.go:411] HTTP endpoint opened: http://localhost:8545
[0419 21:20:45.248865 node/node.go:341] IPC endpoint opened: /home/kobot-b/ethereum/test_data/geth.ipc
[0419 21:24:34.138538 eth/backend.go:475] Automatic pregeneration of ethash DAG ON (ethash dir: /home/kobot-b/.ethash)
[0419 21:24:34.138604 eth/backend.go:482] checking DAG (ethash dir: /home/kobot-b/.ethash)
[0419 21:24:34.138661 miner/miner.go:136] Starting mining operation (CPU=4 TOT=5)
[0419 21:24:34.139169 miner/worker.go:516] commit new work on block 76 with 0 tx & 0 uncles. Took 456.198µs
[0419 21:24:34.139270 vendor/github.com/ethereum/ethash/ethash.go:259] Generating DAG for epoch 0 (size 1073739904) (0000000000000000000000000000000000000000000000000000000000000000)
[0419 21:24:35.120117 vendor/github.com/ethereum/ethash/ethash.go:276] Done generating DAG for epoch 0, it took 980.860959ms
```

[ 프라이빗 네트워크 RPC 서버 활성화 ]

```
Welcome to the Geth JavaScript console!

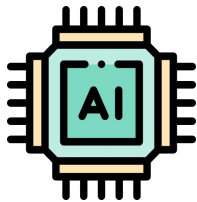
instance: Geth/v1.5.5-stable-ff07d548/linux/go1.6.2
coinbase: 0xaf549ddf419cd794836cfdbbe2ebb5791766597c
at block: 75 (Fri, 17 Apr 2020 20:26:09 KST)
datadir: /home/kobot-b/ethereum/test_data
modules: admin:1.0 debug:1.0 eth:1.0 miner:1.0 net:1.0 personal:1.0 rpc:1.0 txp
ool:1.0 web3:1.0

> eth.accounts
["0xaf549ddf419cd794836cfdbbe2ebb5791766597c"]
> eth.getBalance(eth.accounts[0])
412656250000000000000000
> miner.start()
true
> miner.stop()
true
>
```

[ geth로 attach ]

# 향후 추진 계획

## [ 검수 시스템 ]



- 참여자의 신뢰도 관리 알고리즘 개발
- 모델 Fine-Tuning
- 2차 레이블링 및 검증
- AWS 이전 및 백엔드 서버와의 연결

## [ 보상 시스템 ]



- 블록체인을 통한 투명성 향상
- 보안을 위해 SafeMath 사용
- Web3를 통한 서버와 블록체인 네트워크간의 상호작용

## [ 웹 어플리케이션 ]



- 웹 Front-Back API 연동
- 반응형 웹 구현 및 UX/UI 검수
- 모듈간의 API 구현

# 개발 로드맵

항목	세부 내용	3월	4월	5월	6월
설계 및 구현	UX/UI 설계 및 개발				
	DB 설계 및 개발				
	문서 벡터 데몬 설계 및 구현				
	Crawling 데몬 설계 및 구현				
	블록체인 모듈 설계 및 구현				
	스마트 컨트랙트 설계 및 구현				
	클러스터링 데몬 설계 및 구현				
	T-SNE 데몬 설계 및 구현				
테스트	시스템 테스트				
최종 발표	발표 준비 및 마무리				



**감사합니다**