|  |  |  |
| --- | --- | --- |
| **발명자 – 인적 사항** ※발명자가 다수인 경우에는 아래의 표를 복사하여 사용하시기 바랍니다. | | |
| 성명 | 국 문 | 박종건 |
| 영 문 | PARK JONGGUN |
| 주민등록번호 | 830101-1408811 | |
| 주민등록상의 주소  **(도로명주소)** | 서울시 동작구 매봉로99 104-401 | |
| 우편번호 | 06916 | |
|  |  | |
|  |  | |

**요약 : 4건 신규 출원**

**1. 다중 페르소나 지원을 위한 효율적 생성 AI 모델 구축 및 사용자 평가 기반 강화학습 시스템**

**2. 대규모 시각 언어 모델 기반의 사용자 행동 분석을 통한 AI 휴먼 기반의 어시스턴트 시스템**

**3. 생성형 인공지능 컨텐츠의 저작 계보(pedigree) 검증 및 거래(transaction) 시스템**

**4. 온라인 기반 생성 인공지능 다중 학습 방법, 서버 및 그 시스템**

**1.특허명 : 다중 페르소나 지원을 위한 효율적 대규모 AI 모델 구축 방법 및 피드백 시스템**

해결하고자 하는 내용

ChatGPT를 시작으로 초 거대 모델 기반의 AI 서비스를 제공하고 운영하는것에 관심이 크게 증가하고 있다. 하지만 초 거대 모델을 운영하기 위해서는 막대한 운영자원이 발생하게 되고, 비즈니스에 따라서 특화된 서비스를 제공해야 하는 경우, 사용자에 따라 맞춤형으로 제공해야 하는 경우 그에 따른 추가 인공지능 모델 서비스를 제공하고자 한다면 더 큰 비용이 발생하게 된다.

본 발명은 최적의 비용으로 대규모 인공지능 서비스 모델을 제공하고, 증가되는 서비스 모델에 대한 비용을 최소화 하여 다양한 사용자의 니즈에 대응하도록 하며 궁극적으로는 최소의 비용으로 최적의 인공지능 서비스를 제공할 수 있는 것을 목적으로 하며, ChatGPT 가 RHLF(Reinforcement Learning from Human Feedback, 사용자 피드백 기반의 강화 학습) 기반으로 훈련이 되고 향상되는 것 처럼 생성된 결과물

**해결하고자 하는 방안**

1. 인공지능 모델이라 하면 Text 입력 기반으로 Text, Image, Sound, Video 등의 멀티미디어를 생성할 수 있는 생성형 인공지능 (Generative Artificial Intelligent) 모델을 의미한다.

2. 인공지능 모델에 따라서 특화된 생성을 목적으로 제작되는 경우가 있다(Full Fine-tuned). 예를 들어 텍스트 생성은 일반적인 질의 응답에 특화된 모델, 소설 작성에 특화된 모델, 대화에 특화된 모델이 있을수 있고, 이미지 생성의 경우 실사, 반실사, 만화등의 화풍이 구분되면 소리의 경우 효과음 특화, 음악 제작특화, 음성 제작 특화등으로 구분될 수 있다.

3. 해당 AI모델에 대해 일종의 페르소나(스타일) 구분이 가능하다. 즉 글씨로 치면 일종의 작가성향 (ex 선생님, 변호사, 소아과 의사 등 직업), 그림의 경우 화풍 (피카소 스타일, 김수근 스타일 등), 소리의 경우 화자나 작곡 스타일 등으로 구분된다.

4. 해당 특화 모델을 모두 올리는 방법이 가장 간단한 방법이지만. 시스템 자원이 많이 소모되고 이는운영비용의 증가로 이어진다. 즉 텍스트 생성의 경우 3개의 모델을 모두 올리게 된다면 1개를 올렸을 때 보다 3배의 자원이 소모된다.

5. 이 문제 해결을 위해 우선 기본내용이 훈련된 기본 AI 모델을 기본으로 GPU 메모리 적재하여 활용가능하게 한다. (일종의 초/중/고 일반 수업을 받은 듯한 모델) Text, Image, Sound, Video 에 대한 특화 모델이 하나씩 시스템에 적용되는 것이다.

6. RHLF 로 훈련된 기본 텍스트 모델이 사용자가 각 모델로 직접 서비스 호출을 하지 않는 상황이라면, 중앙 컨트롤러 역할을 수행한다. 즉 이미지, 영상, 소리 관련된 제작 요청이 오는 경우 기본 모델이 해당 모델을 호출하여 서비스를 수행하는 구조를 가진다.

7. 서비스 입력이 프롬프트 (텍스트 문장)으로 이루어지는 특성을 반영하여 일종의 자동완성 기능과 같은 문장 보완 기능은 별도의 소형 모델로 구축되어 제공된다.

8. 기본 모델을 Full-Finetuning 하지 않고 대규모 모델을 양자화 하여 효과적으로 튜닝하는 QLora 기법을 통하여 해당 모델 별로 튜닝을 수행한다. (텍스트는 의사, 판사, 변호사, 선생님, 변리사 등 직업데이터 기반 튜닝, 이미지는 화가 기반 만화 작가, 사진 작가 등등, 음악은 작곡가 스타일, 음성은 사람 목소리 별 튜닝) 이렇게 튜닝된 모델은 기본 모델이 10GB 에 달하는 반면에 100MB 내외의 작은 형태로 훈련을 수행할 수 있다.

9. QLORA 기법으로 훈련된 모델(adaptor)데이터는 CPU 메모리에 적재하여 대기하는 형태로 준비한다.

10. 사용자 요청 발생시 훈련 모델(adaptor)를 동적으로 GPU 메모리에 적재한후 기본 모델 + 훈련 어댑터 통합 모델 기반으로 사용자의 요청에 대응한다. 즉 변호사 스타일로 대화서비스를 요청하는 경우 기본 모델에 변호사 adaptor 모델을 적용하고 사용자에게 서비스를 제공하는 형태이다.

11.사용자로부터 제공받은 서비스에 대한 만족도를 피드백 받는다. 좋아요/싫어요 형태로 태깅을 하여 해당 결과 데이터를 별도로 보관한다.

12. 모델 특성상 입력되는 프롬프트(prompt) 에 따라 결과값에 대해 영향을 받기 때문에, 상대적으로 긴 프롬프트 중에 긍정적 피드백이 많은 프롬프트는 별도로 보관하며, 타 입력 프롬프트 및 결과값이 유사한 다른 사용자 입력 데이터를 별도로 보관한다.

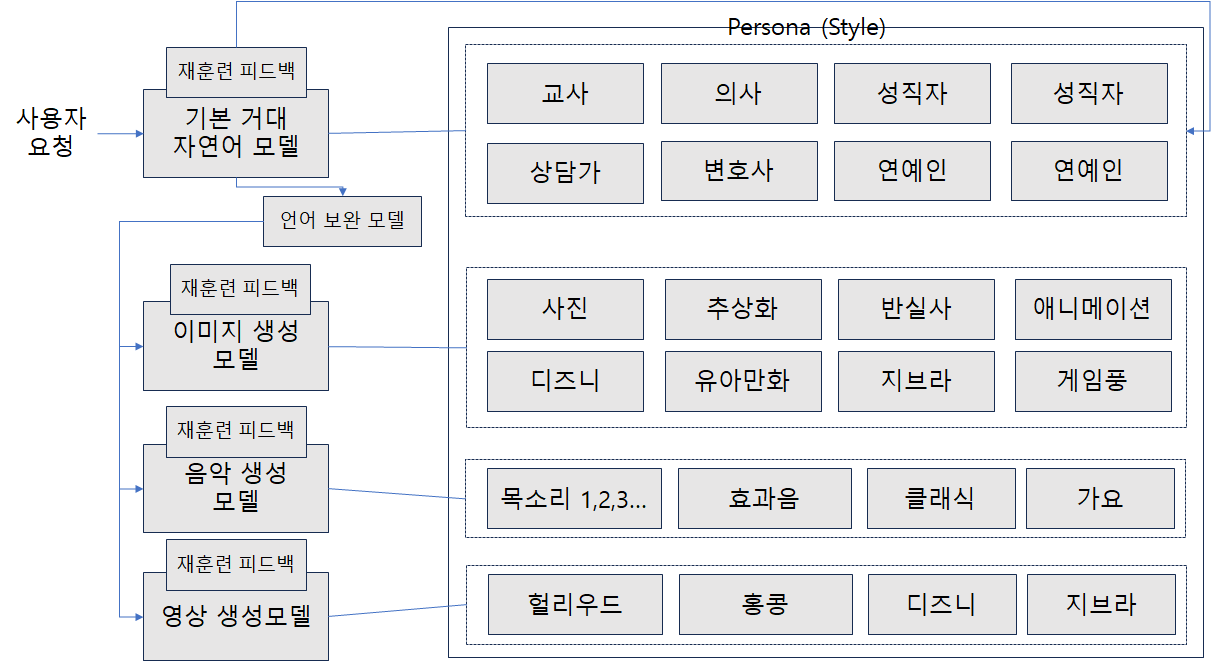
13. 결과값에 대한 긍정/부정 피드백을 기반으로 adaptor 모델 훈련의 재 학습을 수행한다. 피드백을 바탕으로 향후 출력 값이 사용자의 피드백에 좀더 근접하게 반응한다.

14. 프롬프트 결과값에 대해 별도의 모델 튜닝을 수행한다. 일종의 자동완성 기능처럼, 일반 사용자가 프롬프트를 제대로 입력하지 못하여 원하는 결과를 제공받지 못하는 상황을 고려하여, 일반입력-긍정평가 입력 문장으로 훈련된 별도의 소형 모델을 통하여 일종의 프롬프트 보완역할을 수행할 수 있게 한다.

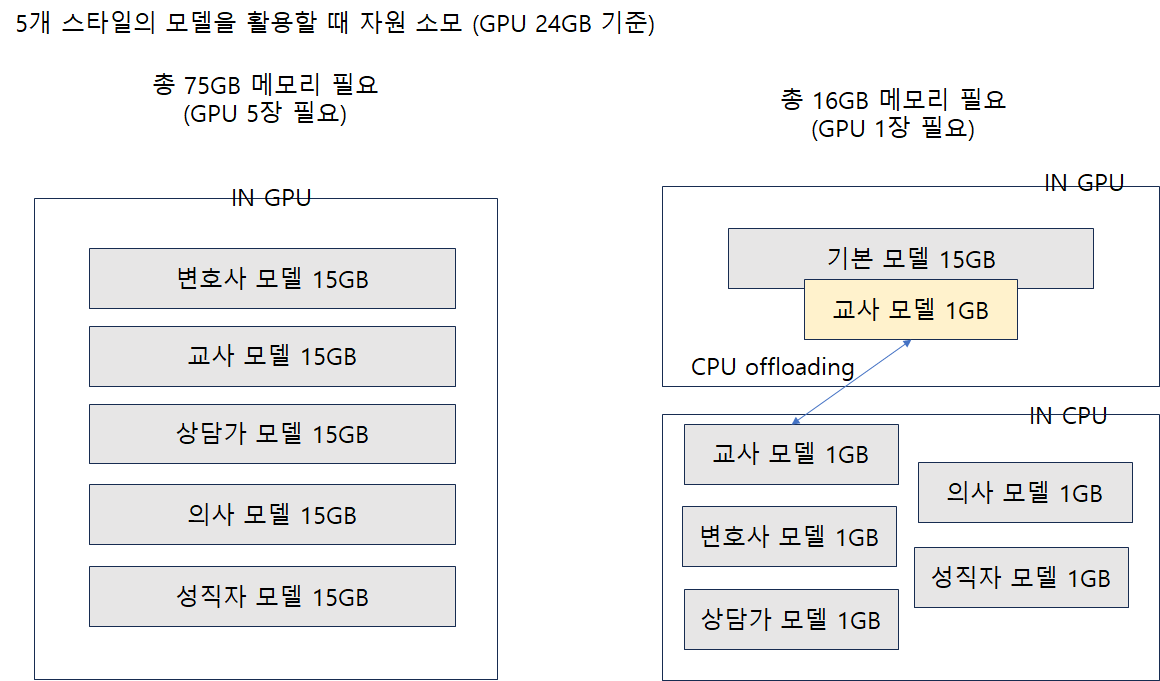
15. 사용자가 다른 스타일로 요청하는 경우 GPU에 적재된 adaptor 모델을 CPU 메모리로 offloading 하고, 다른 스타일 모델을 적재하여 사용자의 요청에 응답한다.

16. 새로운 신규 모델이 추가 되는 경우에도 별도의 시스템 변동 없이 adaptor 모델이 추가되고 이를 동적으로 로딩하는 형태로 활용될 수 있도록 한다..

17. 한 시스템에서 여러대의 GPU가 병렬적으로 동작하는 경우에 특정 서비스에 호출이 많이 이루어진다면, (ex 자연어 생성, 이미지 생성이 각각 GPU에 별도 할당) 해당 시스템에서 자주 호출되는 GPU쪽으로 전원할당을 늘려 처리 성능을 개선한다. 예를들어 총 500W 의 전원을 소비할수 있는데, 자연어에 비해 이미지 요청이 많다면 해당 GPU에 300W, 덜 호출되는 GPU에 200W를 할당해 성능을 개선



기본 모델 및 Adaptor 모델 구성



시스템 구성 비교 (동적 Offloading 지원)

**2. 특허명 : 대규모 시각 언어 모델 기반의 사용자 행동 분석을 통한 AI 휴먼 기반의 어시스턴트 시스템**

**A. 해결하고자 하는 내용**

태블릿이나 PC와 같이 디지털 시스템 혹은 디지털 교과서를 통해 교육받는 환경이 늘어나고 있다. 다만 이러한 시스템들은 단순히 아날로그적 입력을 디지털로 바꾼 수준으로 진행되거나 AI를 도입한 경우에도 단순히 사용자의 얼굴 시선 트래킹을 통해 집중도 여부를 판단하는 수준에 그치고 있어서 상황에 맞는 적절한 피드백을 주는 것이 제한된다는 문제가 있다. 아울러 강사 혹은 교사를 대체하는 수준이 아닌 교육 과정에 있어서 정답 여부나 유사 문제 추천등의 단순 피드백에 기반하고 있어, 강사가 지속적으로 관찰하고 피드백을 주는 교육과정 만큼의 효과를 기대하기 힘들다

이에, 대규모 언어모델 및 시각적 모델을 결합하여, 중급 강사 수준의 지속적인 관찰과 피드백을 디지털 휴먼 기반의 관찰 및 피드백을 통해 교육 성과를 높이고, 적은 수의 강사 문제, 강사의 격차 문제, 초등학교의 경우 남자 선생님을 선호하나 여 선생님 위주 등의 문제 및 보호자의 자녀 수업 문의등의 조치에 대해서도 대응할 수 있도록 한다.

**B. 해결하고자 하는 방법**

1.기본적으로 디지털 학습 시스템은 노트북이나 태블릿처럼 화면과 키보드/터치의 입력장치, 음성 및 카메라 입력 장치를 가지는 시스템을 기반으로 한다.

2.디지털 화면에는 학습용 콘텐츠가 출력되는 부와 디지털 휴먼 (혹은 디지털 Assistant) 가 노출되는 부분으로 구성된다.

3.콘텐츠를 수행하는 방법에는 터치 입력, 키보드 입력, 터치 펜 활용, 모션 동작, 음성 입력등의 방법을 활용하여 수행한다.

4. AI 시스템은 사용자의 수행 성과를 3가지 측면 (정확도, 반응속도, 시선집중)의 정량적인 데이터를 입력받는다.

5. 기본적인 시선 추적을 통해 AI Assistant 는 사용자를 응시하고 있는 방향으로 Tracking 하도록 한다. 이는 별도 외부의 로봇 시스템이 사용자를 관찰하는 것처럼 시선을 응시하고 있어, 사용자의 집중도를 유도하는데 도움을 준다.

6.어시스턴트의 성별은 남/녀 모델이 제공되며, 사용자가 남자인 경우 여자 튜터가 사용자가 여성인 경우 남자 튜터가 제공되어 커뮤니케이션 향상이 가능하도록 한다.

7. 대규모 언어모델은 Large Language and Vision Assistant 모델로 구성되어 GPT-4와 같은 언어모델이긴 하나 시각입력 + 텍스트 입력을 함께 받을수 있는 멀티 모달로 훈련되어 있다.

8. 해당 모델은 초급, 중급, 고급 교사형태로 훈련이 되어 있으며, 사용자의 성과에 따라 AI 휴먼 튜터의 모델이 동적으로 변경될 수 있도록 한다. 컨텐츠에 대한 기본적인 구성 내용 및 추진 목표 및 커리큘럼에 대해 기본 학습이 되어 있다.

9. 사용자의 성과가 떨어지거나 우수한 경우, 또한 일정시간 동안 꾸준하게 교육이 진행되는 경우 대규모 모델에 수행성과 정보를 텍스트 입력으로 제공하고, 현재 컨텐츠 화면 및 사용자 응시 화면을 입력값으로 제공하여, 언어모델로부터 사용자 피드백을 문장으로 생성한다.

10. 생성된 피드백을 바탕으로 가상 AI 휴먼 합성을 통해 사용자에게 교육 진행중인 내용을 피드백으로 제공한다.

11. 교육 과정 수행중 어려운 사항에 대해서는 AI휴먼에게 질문하고 그에대한 답을 받을수 있도록 제공한다. 사용자 입력은 텍스트와 음성입력, 이미지 입력은 현재 수행중인 컨탠츠의 캡쳐된 이미지가 제공된다.

12. 시각기반 언어모델을 통해 사용자의 질문과 참고 이미지를 입력받아 대답을 생성하고 이를 AI휴먼 형태로 생성하여 사람 수준의 즉각적인 피드백을 수행한다.

13. 사용자의 과업 수행 형태 및 질문 수준 및 성취도를 기반으로 AI 튜터모델이 변경된다. 즉 기본모델을 바탕으로 수강자가 과업을 잘 수행한다고 하면 고급 모델을 통해 보다 더 도전적인 과업을 수행하고 그에 따른 피드백을 제공할수 있도록 하며, 사용자가 부진하다고 하면 좀더 쉽고 기초적인 가이드가 가능하도록 초급 모델을 통해 사용자가 지속적인 관심을 가지고 수업에 임할 수 있도록 한다.

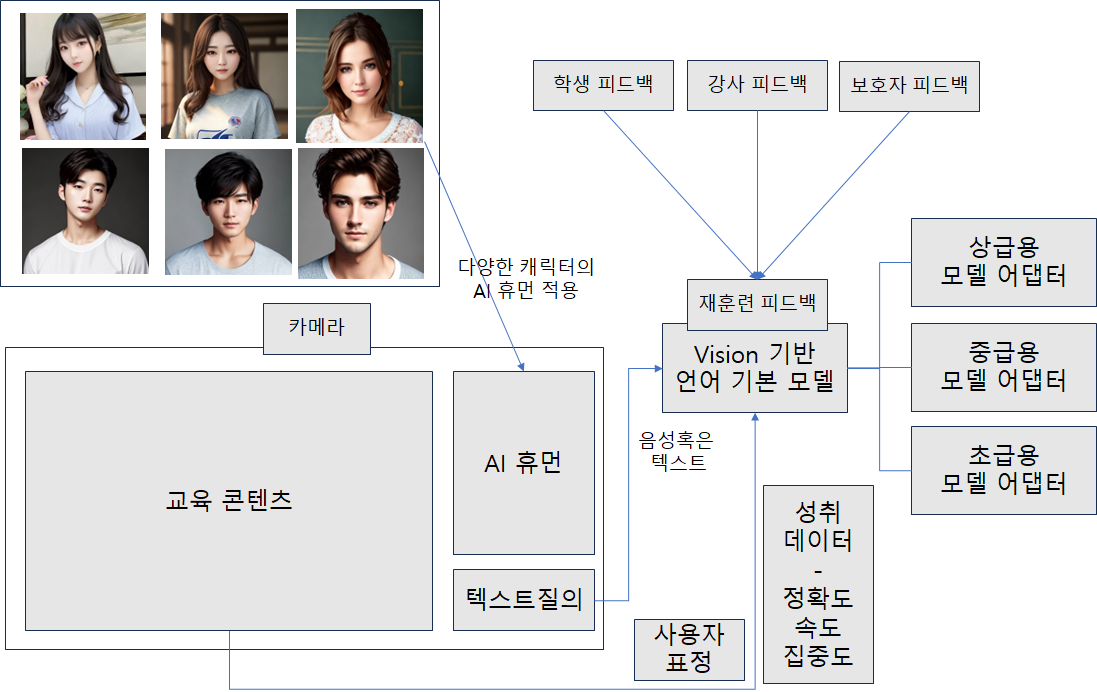
14.학습 태도에 따라 AI 휴먼의 연령대가 변경된다. 기본 모델은 젊은 선생님을 기준으로 하나. 학습자  
태도에 문제가 발견되는 경우 중장년의 연령대나 성별 교체등이 AI 기반으로 자동으로 수행될 수도 있고 주 강사의 조정에 따라 변경될 수도 있다.

15. 사용자의 기본 정보 및 교육 성과 정보가 입력되어 있어, 보호자에 대한 학업 성취 피드백이나 자문 지원이 가능하다. 즉, 보호자의 어시스턴트 접근(사용자 정보 입력.. 로그인)을 통해 학업 성취에 대한 질의나 개선점 및 피드백 요청에 대해 어시스턴트가 대응한다.

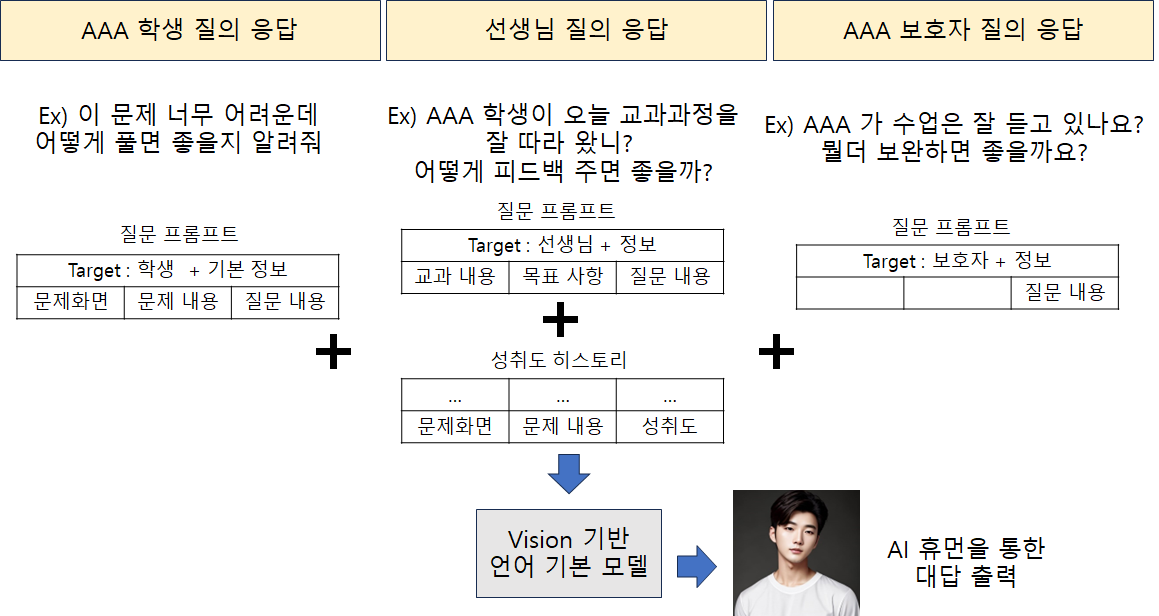
16. 학생 및 보호자에게 제공된 피드백은 실제 강사에게도 전달되어, 교육 성과와 함께 판단 및 사람의 개입을 통한 보조 피드백을 제공하는데 활용한다. 또한 실제 강사 역시 AI 어시스턴트에 질의를 통해 실제 교육 방향 개선을 진행하는데 도움을 제공한다.

17. 수강생의 만족도, 보호자의 만족도, 선생님(실제사람)의 만족도등 3가지 각도의 피드백 시스템(Reinforcement Learning From Human Feedback RLHF) 으로 연동하여, 해당 언어모델이 재훈련을 통하여 성능이 지속적으로 업그레이드 되도록 제공한다.

18. 시스템에 존재하는 각 사용자의 다양한 Assistant 는 사용자의 public/private 설정에 따라 public 인 경우 사용하지 않는 시간대에 타 사용자의 반대되는 성향의 public assistant 와 시스템적 대화를 통해 상호 서비스를 업그레이드 한다 (유전적 법칙)



기본 시스템 구조



사용 주체별 정보 입력 및 그에 따른 응답 구조

**3.특허명 : 생성형 인공지능 컨텐츠의 저작 계보(pedigree) 검증 및 거래(transaction) 시스템**

발명 목적

생성형 인공지능을 이용하여 문장, 이미지, 음원, 비디오등 다양한 멀티미디어 요소를 생성하고 활용하는 빈도가 크게 증가하고 있다. 사람의 저작물은 저작권의 형태로 보호를 받지만, AI를 활용한 디지털저작물은 해석이 분분한 상황이다. (미국에서는 사람의 창의적인 행위, 즉 해당 생성물을 만들기 위한 프롬프트 엔지니어링 노력을 인정해 주는 등의 일부 사례 등) 생성형 AI 가 동일한 프롬프트를 입력하더라도 동일한 생성물이 나올확률이 극히 적다는 면에서 하나의 생성물은 유일성이 보장이 된다. 따라서 마치 우리가 파워포인터를 통해 작업을 한 자료의 경우에도 대외비나 일종의 저작권을 보장 받는것처럼 이러한 AI생성물에 대해서도 그 저작자에 대한 절차와 행위에 대한 인증 및 해당 디지털 저작물에 대한 유일성을 보장하는 시스템이 존재해야 해당 저작물의 저작권인정 차원이 아닌 활용차원에서 활용의 유일성 보장 및 거래 보장이 가능해 질 수 있으므로, 이에 따른 저작물 이력 추적 및 유일성 보장, 거리 지원 시스템을 기술한다.

발명 내용

1. 사용자는 텍스트 프롬프트 입력을 기본으로 하며, 보조 자원으로 이미지, 오디오, 비디오를 입력 값으로 제공하고 텍스트, 이미지, 소리, 영상 생성 모델을 통하여 창작물을 만들어 낼 수 있다.

2.사용자는 생성된 결과물을 다시 입력값으로 활용할 수 있다. 즉, 텍스트 ,이미지, 오디오, 비디오등의 결과물을 중간 결과 및 신규 입력형태로 재 활용하여 결과물을 만들 수 있는것이다.

3.해당 결과물이 저장될때는 해당 사용자의 id 겂을 기반으로 하여 이력과정이 저장된다. 입력 미디어의 원본을 hash 로 변경하여 입력값에 대한 유일 값을 생성해낸다.

4. 생성된 결과물에 대해서도 hash 값으로 변경하여 유일 값을 생성한다. 입력값과 출력값을 키로 보관을 하여 정보를 저장한다.

5. 미디어 특상상 변조가 될 수 있으므로, Transformer 기반의 AI기술로 해당 미디어에서 특징(Feature)를 벡터 추출(extraction) 한다.

6. 입력값 hash와 출력값 hash, 특징 벡터추출 데이터, 저장일시, 주 사용자를 중앙 인증 서버에 저장한다.

7. 최초 생성자는 해당 미디어에 대한 1차 소유권을 부여 받는다. 족보(pedigree)에 최초 저작자로 표현된다. 첫 계보 생성시에 사용자는 입력/출력에 대한 해당 미디어의 가치 (비용)을 지정한다.

8. 모든 미디어에는 watermark 가 부착되며(사본), 부착안된 원본도 별도로 보관된다. 영상/이미지에는 qr코드 (pedigree 접속 가능한 url 정보) 가 표현되며, 텍스트에는 마지막에 pedigree 접속 가능한 짧은 url 이 별도로 첨부된다.

9. 디지털 생성물 간의 거래는 가상화폐(Virtual Currency – 게임머니, 싸이월드 도토리) 를 이용하여 거래가 이루어지며 실제 화폐의 전환을 통해 이루어 진다.

10. 거래를 위해서는 해당 비용 만큼을 실물 화폐의 입금을 통해 변경된다.

(거래를 통해 최종 소유하는 virtual currency 를 실제 비용으로 환급할 수 있다.)

11. transaction 의 대상은 입력조건과 결과물이다. 입력조건은 ai 모델 종류, 입력 프롬프트, 생성을 위한 조건값이고, 결과물은 해당 입력 조건을 통해 생성된 결과물이다.

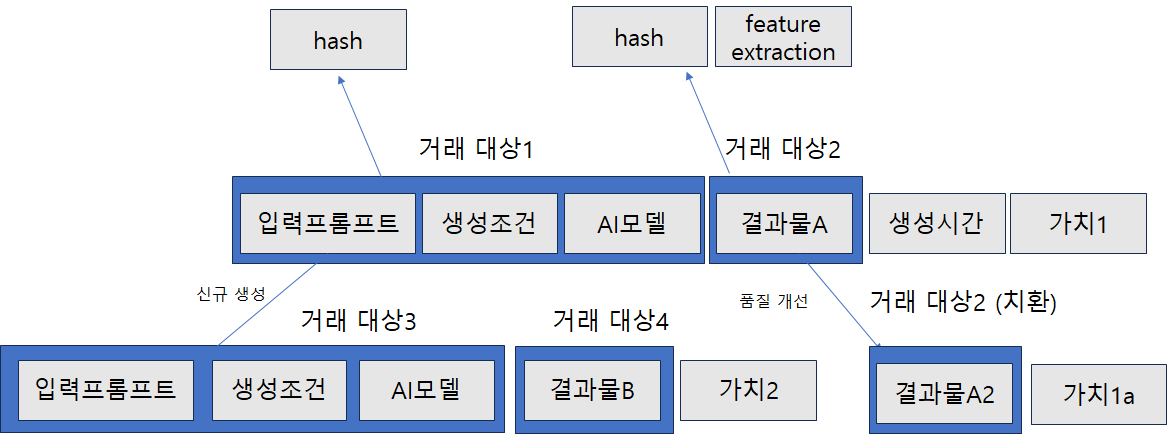
12. 만일 타 사용자에게 해당 생성 데이터를 넘기는 경우 족보에 해당 사용자의 정보가 기록되고, 족보의 마지막 사용자가 현 소유주가 된다. 입력/결과물에 대해 함께 transaction 이 발생할수 있고, 입력물이나 결과물만 transaction 이 일어날수도 있다. 매 Transaction 마다 가치가 변경될 수 있다.

13. 입력물이나 결과물 둘중 하나만 소유한 사용자는 계보 추적을 통해, 나머지 1개에 대한 소유권을 해당 사용자와의 transaction 을 통해서 거래받을 수 있다.

14. 입력 데이터를 기반으로 새로운 생성물을 만들어 낼수 있으며, 그 생성물에 대해서는 마지막 생성자가 소유권을 가지게 되고 비용을 산정할 수 있다.

15. 무단으로 해당 디지털 데이터를 복사 하게 되는 경우, 1차로 hash 기반의 비교를 진행한다. 미디어 자체의 형식을 바꾸면 hash 값이 달라지게 되므로 2차로 extractnion 기반의 벡터 데이터 값을 비교하여, 15% 미만의 일치도일 때 복제품이 아니다로 판단하고, 그 이상의 경우 복제데이터로 의심 추정을 하게 된다.

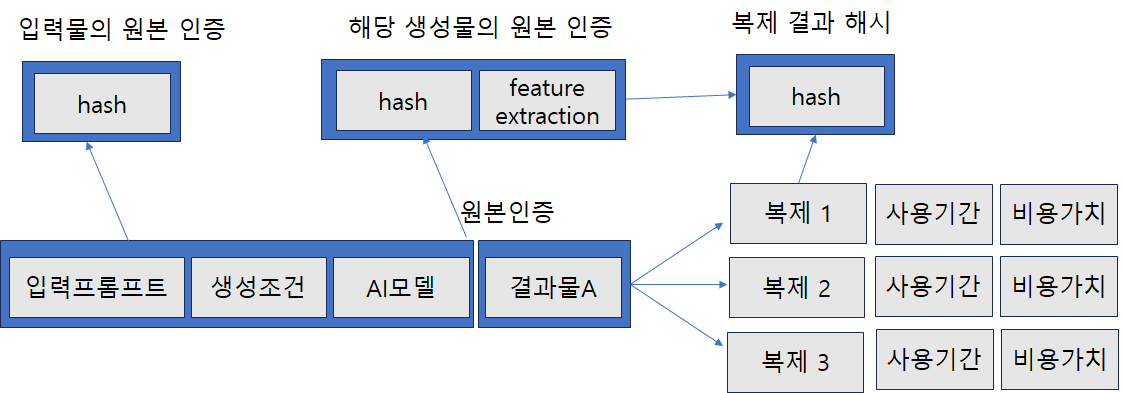
16.최종 소유자는 가치를 업그레이드 해서 기존 생성물을 치환할 수 있다. 예를 들어 이미지나 영상, 오디오의 경우 고화질/고음질로 스케일 업 하는 경우이며 이때는 기존의 결과물을 폐기하고 새로운 생성물로 원본이 변경된다. 변경시에 계보에 기록이 되며 가치는 최종 사용자에 의해 변경 가능하다.





17.여러 개의 생성물을 결합하여 새로운 미디어(영상 등)을 제작하는 경우, 입력대상의 개별 생성물과 최종 결과물에 대해서 거래 가능한 고유값이 제공된다.

18.융합 생성물의 경우 컨텐츠를 생성하는데 입력된 모든 프롬프트와 부분 결과물, 그리고 최종 결과물에 대해 전체 거래가 가능하지만, 부분 거래 (최종 파생 영상만 거래)도 가능하다. 이 경우에 만일 부분적으로 일부만 (텍스트, 이미지, 음악, 삽화 영상등) 활용하기 위해서는 해당 소유자와의 거래가 이루어 져야 전체에 대한 권한이 부여되게 된다.



19. 생성 결과물에 대해서는 복제물에 대해서 거래가 가능하다. 다만 해당 복제물에 대해서는 정해진 기간에 대해 가치(비용)을 지불하고 해당 기간만 활용이 가능하다.

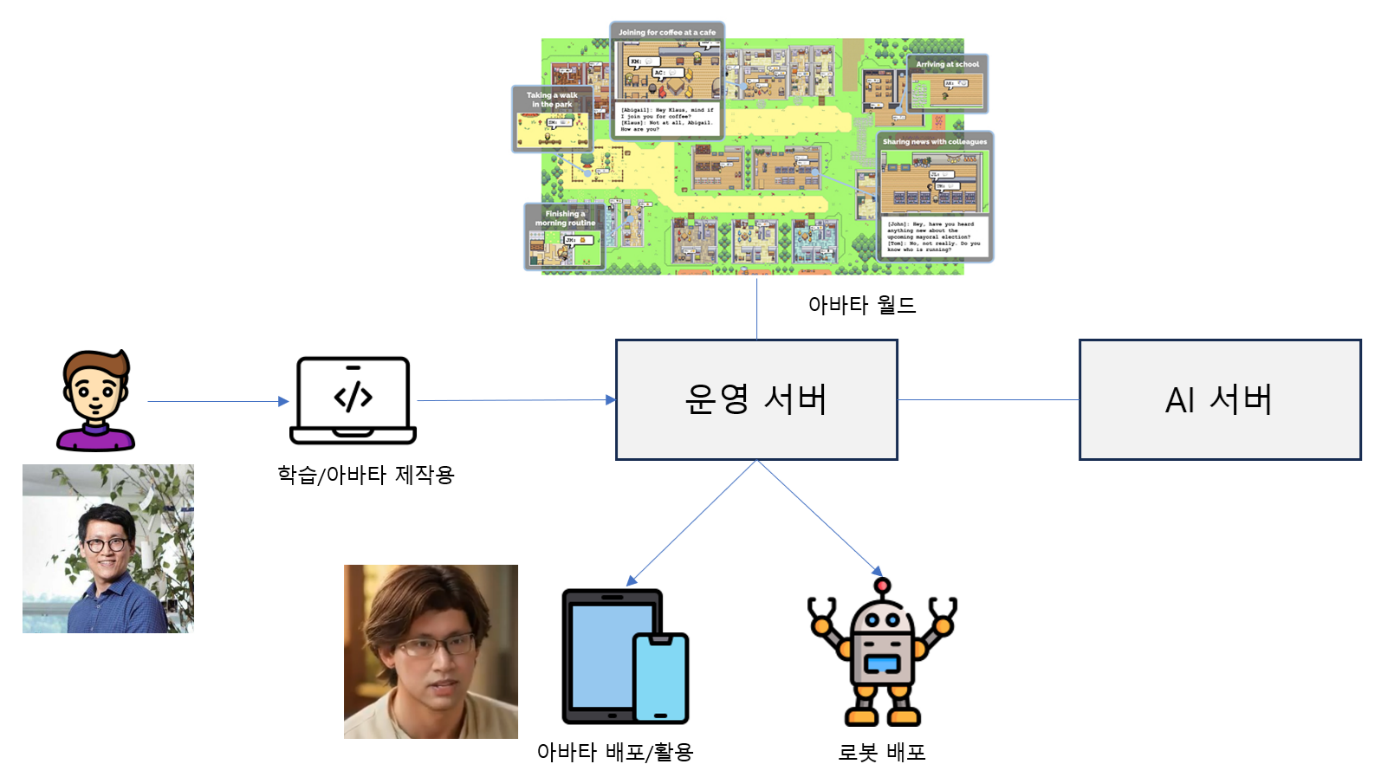
20. 복제물은 원본 인증 정보와 복제 시점을 활용하여, 복제 해시가 생성되고, 이 정보는 원본의 복제 목록에 보관되며, 사용 기간이 만료되는 경우 해당 복제본의 사용권이 파기되어 사용할 수 없다.

**4. 온라인 기반 생성 인공지능 기반 페르소나 학습 및 저작 방법과 배포 및 운영 시스템**

ChatGPT 로 촉발된 생성 인공지능은 대화 기반의 텍스트를 생성하는 것에 머무르는 것이 아닌 목소리, 음악, 효과음, 이미지, 영상, 코딩에 이르는 전 영역으로 급속도로 확대되고 있다. 이에 따라 기존 SW 코딩을 배우는 트렌드에서 생성AI 기반으로 생각했던 서비스를 구축하는 것이 가능해 진 상황이다. SW 코딩을 배우기 위한 온라인 플랫폼을 통해 더욱 크게 확산될 수 있었던 것 처럼, 온라인을 통해 AI활용 기법을 배우고, 배운 내용을 실습해 보며, 이를 마지막으로 공유 및 배포 할 수 있는 시스템을 구성하여 손쉽게 배우고 실습하고 공유할 수 있는 환경을 제공한다.

보다 발전적으로 최종단계를 통해 만들어지는 산출물은 일종의 가상 아바타로서, 가상 아바타를 통해 모바일을 통해 커뮤니케이션을 지원하고, 이 아바타는 가상 세계(일종의 메타버스)에 업로드 되어, 스스로 성장할 수 있도록 구성되어 모바일을 통해 생성된 가상 아바타는 가상 세계에서 타 가상 아바타와 대화 및 인터넷을 통한 자가 발전이 가능하다. 사용자는 타 아바타와 merge 를 통해 새로운 스타일로 생성할 수도 있으며, 물리적 디바이스에 배포(로봇) 하여 활용하는 것도 가능하다)

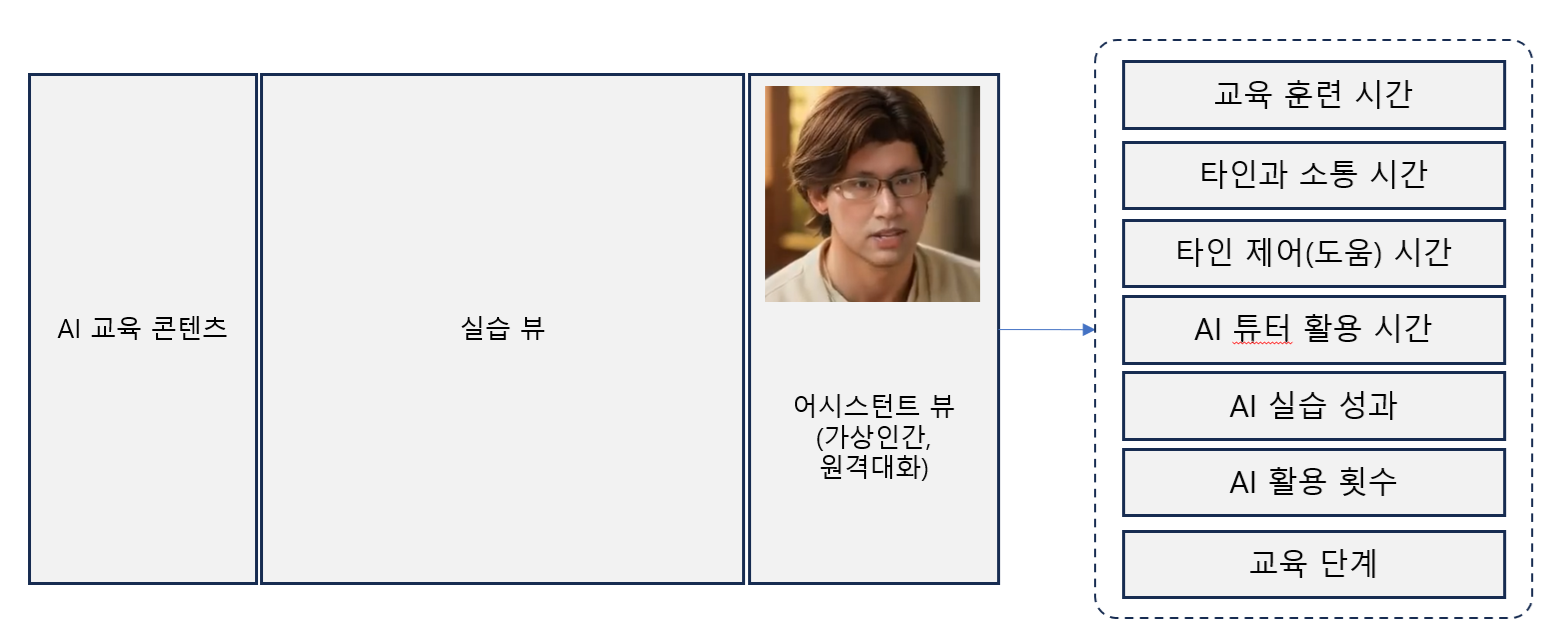
1. 생성형 인공지능은 일반 PC에서 구동하기에는 성능적/비용적 제한으로 인해 GPU가 탑재된 AI 서버가 있고, 교육 콘텐츠와 수행을 위한 교육서버, 교육이후 만들어낸 가상 아바타를 배포하고 활동하게 될 운영서버 그리고 여기에 접속하여 교육 및 실습을 수행하는 PC (혹은 태블릿)와 실제 활용/확인 목적의 모바일 디바이스(혹은 태블릿)으로 크게 구성 된다.



2. 온라인 학습 시스템에서는 AI를 통해서 배우는 단계, 사람에 의해 원격으로 배울수 있는 단계, 오프라인에서 배움에 도구를 활용하게 되는 단계로 나뉘게 되며, AI활용 학습에 사용되는 컨텐츠는 교육 서버에 존재하며, 텍스트/음향/이미지/영상/아바타 의 구분에 따라 온라인 상 학습을 진행할 수 있다.

3 교육 서버에는 텍스트/음향/이미지/영상/아바타 등의 구분에 맞춘 컨텐츠가 튜터에 의해 입력되고, AI어시스턴트에 의해 설명될 내용을 훈련시킬 수 있다. 교육 내용에는 텍스트, 이미지, 영상등의 컨텐츠와 해당 교육에 특화하여 AI 어시스턴트가 동작될수 있도록 시스템 프롬프트(기본적으로 AI 어시스턴트의 페르소나), 기본 질문예시, 옵션으로 목소리 및 가상 얼굴등을 지정할 수 있다.

4. 학습 공간에서는 컨텐츠 뷰, 실습 뷰, 어시스턴트 뷰로 나뉘며, AI 어시스턴트에 의해 컨텐츠 뷰의 주요 교육 컨텐츠를 설명해 주고 이를 실습하게 해 준다. 만일 대상자가 실습에 대한 어려움이 있는 경우 AI 어시스턴트에게 문의하는 경우 실습 뷰에서 실습할 내용을 예시형태로 수행하여 학습자의 이해를 돕는 역할을 한다.



[ 교육 환경 구성 ]

5. 교육 과정을 실습하는 경우, 실습 컨텐츠는 교육서버로부터 전달받으나, 실습 후 내용은 운영 서버로 저장되나, 실습시에 사용되는 AI는 AI서버로부터 호출되어 활용된다.

6. 교육과정 실습은 생성형 인공지능 특성상 타이핑에 의한 텍스트 입력을 기본으로 하나, 음성/소리를 마이크로 입력하여 활용하는 방법과 카메라를 통해 텍스트나 사물/이미지 인식 및 사용자의 얼굴을 입력하여 활용하는 방식, USB메모리나 무선네트워크(Iot,WIFI,BT,RFID 등) 장치를 통해 입력받는 방식으로 구분된다.

7. 어시스턴트 뷰는 기본적으로 AI어시스턴트와의 대화를 기본으로 하나, 온라인 상으로 다중 학습이 진행되고 있는 경우에는 인간 튜터 혹은 타 공동학습자에게 문의가 가능하며, 만일, 단독으로 타 사용자에게 문의하는 경우에는 원격으로 제어권을 넘겨, 해당 사용자로부터 원격으로 제어할 수 있도록 지원한다.

8. 훈련에 참여한 시간, 타인에게 대화한 시간, 타인을 도와준 시간(제어권), 튜터를 활용한 시간, 교육 단계 완성도, AI활용 횟수 및 시간 등의 성과 내용을 교육서버로 전송하여, 사용자의 교육 이행 수준뿐만 아니라 사람 혹은 AI튜터로부터 도움을 받고 도움을 준 사회적 영역에 대한 정보를 보관한다.

9. 학습자가 실습을 통해 생성해 낸 일반적인 결과물 (텍스트, 이미지, 영상, 오디오 등등)은 운영서버에 저장되고 이 저장된 결과물은 고유 url 을 부여받는다. 해당 url을 바탕으로 SNS에 공유하거나 QR코드 형태 출력을 통해 타인에게 전송할 수 있다.

10.학습자가 실습을 통해 생성해 내는 아바타(얼굴 + 대화 페르소나 + 목소리 등을 포함하는 가상 인간)의 경우, 별도의 가상 공간에 배포가 된다. 생성한 사용자의 설정에 따라 자가 학습 금지 (만들어진 형태로만 PRIVATE 하게 존재) 와 타인이 만든 AI 아바타와 서로 가상공간내에서 커뮤니케이션할 수 있는 공유 학습 모드로 구분이 된다.

11. 공유 학습 모드의 경우 일종의 게임과 같은 가상 공간 내에서 아바타 끼리 서로 소통하면서 강화 학습을 통해 발전 할 수 있는 방법. 외부 인터넷 연결 (RSS/API 데이터 연계)을 통해 스스로 학습 발전할 수 있는 방법, 마지막으로 사람과 대화를 통해 발전할수 있는 방법 3가지를 지원한다.

12. 아바타가 모여 있는 가상 공간은 아바타 타운으로, 앱 스토어와 같이 사용자가 아바타를 선택할 수 있는 측면의 스토어와 같은 기능. 아바타 간에 커뮤니케이션 및 활동(학습)이 이루어질 수 있는 가상 세계로서의 기능으로 구분이 된다.



[ 아바타 타운 예시 ]

13. 만들어진 아바타는 고유 url 을 부여받고, 이 주소를 통해 해당 아바타와 외부 사용자가 1대1대화가 가능하도록 제공받는다. 첫 생성자의 설정에 따라 외부 사용자와의 대화를 통해 학습이 가능하게 할 것인지 (read/write 모드) 아니면 정보 제공만 가능하게 할 것인지 (read 모드)로 설정할 수 있다.

14. 아바타를 통해서 커뮤니케이션을 위한 입력방식에는 3가지가 있다. 첫번째는 텍스트 채팅 입력을 통한 커뮤니케이션과 음성 인식에 의한 방식으로 기본 입력이 결정되며, 옵션으로 카메라를 활성화 하여 사용자의 얼굴과 주변 상황 정보를 입력받는 방식으로 구분될수 있다. 텍스트 입력과 달리 음성 입력을 받는 경우 STT 로 텍스트로 변환하는 것 뿐만 아니라 사용자의 목소리로부터 감정정보를 전달받고, 카메라가 활성화 되는 경우 사용자의 얼굴 표정, 주변 상황등을 추가적으로 입력받아 대화를 생성하는 입력값으로 활용된다.

15. 아바타와의 대화시에 사용자는 좋아요/싫어요의 피드백을 제공할 수 있다. 이는 아바타가 예기치 않는 동작을 하는 경우, 학습자에 의해서 훈련 조정을 할 수 있는 정보를 제공한다.

16. 학습자는 생성된 아바타에 접속하여 자신 및 외부 사용자와의 대화 내용 및 외부 인터넷 연결을 통해 학습된 사항에 대한 확인이 가능하다. 싫어요 피드백이 발생한 내용을 확인하여 해당 싫어요 상황이 발생하는 것에 연계된 학습 데이터 (대화 내용/외부 인터넷 연계정보)를 확인하고 해당 학습데이터의 차단/변경 및 대화 내역에 남지 않는 테스트 모드를 통해 싫어요가 발생하는 대화를 반복하여 문제가 해결되는지 확인할 수 있다.

17. 온라인 상으로 훈련된 가상 캐릭터를 물리 로봇에 탑재할 수 있도록 제공한다. 기본적으로 사용되고 있는 모델을 교체하여 생성된 아바타 모델로 바꾸는 경우, 로봇의 대화 및 행동 패턴과 합성 목소리가 자신의 목소리로 바뀌게 되어 동작한다. 로봇에 입혀지게 되는 경우에도 동일하게 커뮤니케이션이 가능하도록 제공 받는다.

18. 아바타는 총 5가지의 커뮤니케이션 상태로, 학습자/외부사용자/아바타간 온라인 대화/디바이스 탑재(로봇)/디바이스 간 대화 로 구분되며, 해당 상태를 플래그 형태로 기록에 보관 된다.

19. 아바타 자체도 다양한 형태로 복제가 가능하다. 원본을 그대로 유지하거나 원본으로부터 일부 변환을 하여 사용할 수 있다. 예를 들어 얼굴 스타일을 1990년대 스타일이나 영화 슈퍼히어로 스타일로 변경 하여 적용할수 있고, 복제된 목소리도 훈련을 변경하여 톤을 수정한 음성으로 활용할 수 있다. 훈련에 사용된 대화를 포함된 데이터를 조정하여 해당 캐릭터를 파생화 하여 분화할 수 있다. 얼굴에 대한 소유권이 주 학습자에 부여되기 때문에 원칙적으로 복제 및 파생은 주 사용자에 의해 가능하다.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 학습자 (실제얼굴) | 복제 아바타 1 | 복제 아바타 2 |

20. 만일 타인의 아바타라도 복제 권한이 풀려 있는 경우에 이를 활용하여 파생이 가능하며, 아바타와 아바타 간의 merge 와 merge 비율을 조정하여 해당 아바타의 페르소나의 스타일의 조정이 가능하다. (ex 이미지에서는 애니스타일과 실사스타일을 조합하면 세미 스타일이 나오듯이, 일종의 교배와 같다고 보시면 됩니다.)