



문화정보 이슈리포트

2022-8호(제38호)

# 인공지능아트 산업과 기술동향



이득우

(청강문화산업대학교 융합콘텐츠스쿨)

KCISA 한국문화정보원

## CONTENTS

- I. 인공지능아트 기술 개요
- II. 인공지능아트의 확산
- III. 인공지능아트의 대중화
- IV. 인공지능아트 산업 현황 정리
- V. 문화체육관광분야 시사점

문화정보 이슈리포트  
2022-8호(제38호)

# 인공지능아트 산업과 기술동향

이득우(청강문화산업대학교 융합콘텐츠스쿨)

## 인공지능아트 산업의 흐름

최근 들어 빅테크 기업들이 이미지를 생성하는 인공지능 분야에 주목하고, 관련 기술을 잇달아 발표하고 있다.

엔비디아(NVIDIA)는 자사의 컨퍼런스에서 그림을 그려주면 환경을 자동으로 생성해주는 기술을 선보였고 메타는 자동으로 동영상을 재생하는 기술을 선보였다. 또한 구글 역시 질세라 동영상을 재생하는 기술을 선보였으며, MS는 인공지능 서비스 소스를 공개하는 한편, 앞으로 선보일 서비스에 인공지능아트 생성 기능을 탑재한다고 밝혔다.

그런데 흥미로운 것은 이미지 아트 생성 분야에서 가장 주목받는 기업은 이러한 빅테크기업이 아니라 스타트업이라는 사실이다. 스태빌리티시는 이미지를 생성하는 인공지능 분야에서 급성장하고 있다.

스태빌리티시가 관심을 받고있는 현 상황을 이해하기 위해서는 지난 2년간 발생했던 인공지능아트 서비스에 대한 기술과 동향을 이해해야 한다. 이의 흐름을 정리해보았다.

※ '문화정보 이슈리포트'의 내용은 작성자의 의견으로 한국문화정보원의 공식적인 입장과 다를 수 있습니다.

# I. 인공지능아트 기술 개요

## 1. 오픈AI(OpenAI)의 달리(DALL-E)

- 2021년 1월 오픈AI는 자연어를 입력받아 이미지를 생성하는 인공지능아트서비스 달리를 발표함.<sup>1)</sup>
- 다음 달인 2021년 2월 달리에 사용된 논문 “Zero-Shot Text-to-Image Generation”이 발표됨.<sup>2)</sup>
- 이전까지는 인공지능은 인위적으로 분류된 이미지만 학습하고 유사한 이미지에 대해 정답을 예측했으나, 본 적이 없는 이미지도 스스로 예측할 수 있는 능력을 갖춘 것으로 평가됨.
- 논문 제목의 ‘Zero-Shot’이라는 용어는 인공지능에게 학습되지 않은 이미지를 주었을 때 이를 구분할 수 있는 능력을 의미함.
- 아래는 달리에게 아보카도와 의자의 두 키워드를 던졌더니, 아보카도처럼 생긴 의자를 다양하게 만들어 낸 결과 화면임.

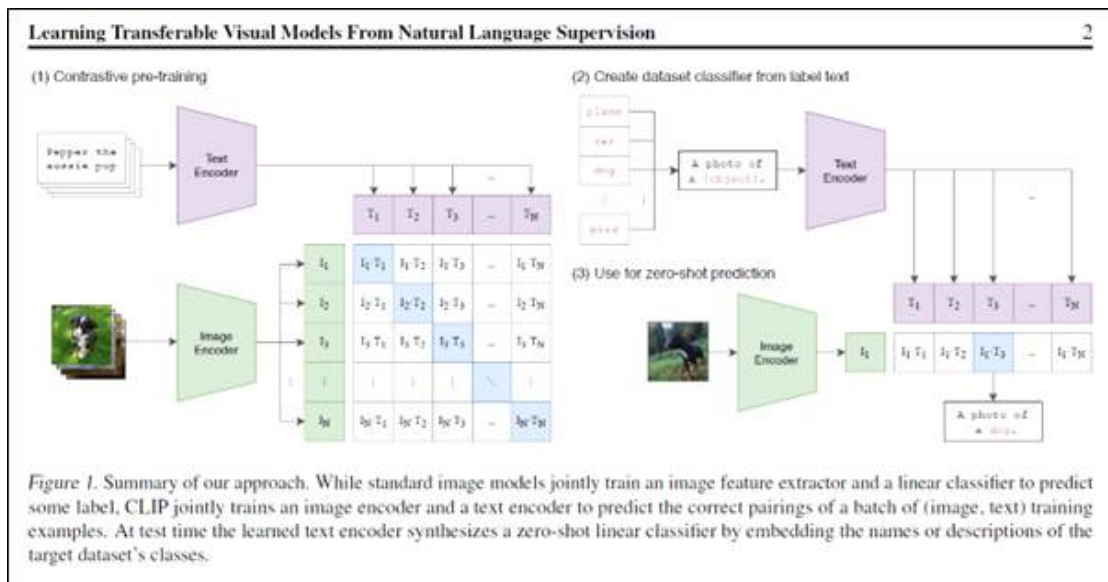


자료: 달리가 만든 아보카도 의자 이미지 ( 출처 : <https://openai.com/blog/dall-e/> )

1) <https://openai.com/blog/dall-e/>  
2) <https://arxiv.org/abs/2102.12092>

## 2. 달리 서비스를 받쳐주는 기술 CLIP

- 같은 달 오픈시는 달리와 함께 자연어 기반의 이미지 인식 기술 “Learning Transferable Visual Models From Natural Language Supervision”을 발표했는데 해당 기술은 줄여서 CLIP(Contrastive Image-Language Pretraining)이라고 함.<sup>3)</sup>
- 기존에는 인공지능학습을 위해 이미지 데이터베이스에서 각 이미지들에 레이블(Label)을 붙여 분류하는 작업을 거쳤으나 이 방식은 이미지 데이터베이스의 양이 커질수록 작업량이 커지며, 레이블로부터 의미있는 정보를 추출하기도 쉽지 않다는 단점이 있음.
- 오픈시는 CLIP기술을 발표하면서 이미지에 레이블을 붙이지 않고 이미지와 연관된 텍스트를 자동으로 추출하는 새로운 학습 방법을 제시했는데, CLIP기술은 이전 기술대비 높은 학습 성능을 보여줌.

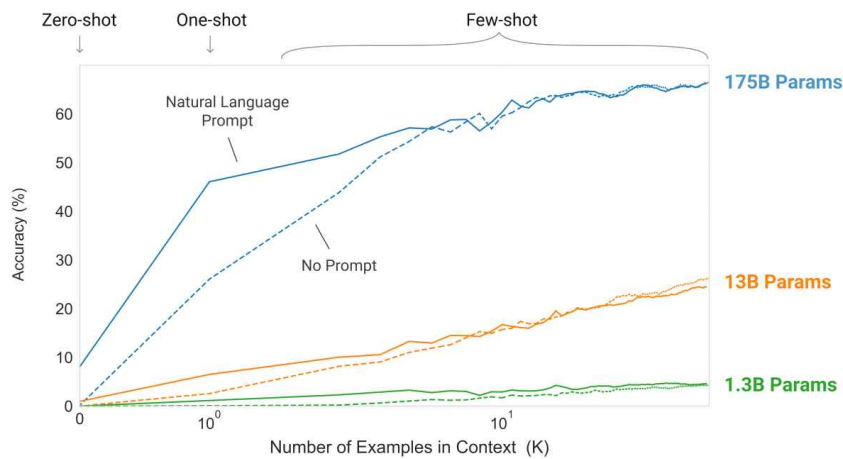


자료: 이미지와 연관된 텍스트를 자동으로 추출하는 CLIP기술 ( 출처 : <https://arxiv.org/abs/2103.00020> )

3) <https://openai.com/blog/clip/>

### 3. 오픈AI의 핵심 서비스 GPT-3

- 오픈AI의 인공지능아트 기술의 내부에는 대형언어모델인 GPT-3모델이 있음.
- 2020년 5월에 발표한 GPT-3의 논문 “Language Models are Few-Shot Learners”에서는 1,750억개의 파라미터를 가진 방대한 신경망 모델에서 책과 텍스트에서 단어와 구절을 연구하고 문장을 생성한 결과를 발표함.<sup>4)</sup>
- 이전 버전인 GPT-2와 비교했을 때 GPT-3에는 특이한 기술적 변화는 없었지만 초대형 모델에서 학습할수록 효과적인 학습이 일어남을 확인함.



**Figure 1.2: Larger models make increasingly efficient use of in-context information.** We show in-context learning performance on a simple task requiring the model to remove random symbols from a word, both with and without a natural language task description (see Sec. 3.9.2). The steeper “in-context learning curves” for large models demonstrate improved ability to learn a task from contextual information. We see qualitatively similar behavior across a wide range of tasks.

자료: GPT-3논문에서 발표한 거대모델의 학습 능력 ( 출처 : <https://arxiv.org/abs/2005.14165> )

- GPT-3를 사용하면 인간이 쓴 것처럼 보이는 복잡한 문장을 생성하는 것이 가능해짐.
- MS(마이크로소프트), 구글, 페이스북의 빅테크기업도 이와 유사한 자체적인 대형언어모델을 발표했으나 GPT-3가 가장 뛰어난 성능을 보임.
- GPT-3의 다양한 활용 사례들이 선보임. 특히 커뮤니티 서비스 레딧(Reddit)에 GPT-3로 제작한 챗봇(/u/thegentlemetre)을 풀어 유저의 고민을 상담했는데, 인공지능이 답변한 문장에 대해서는 누구도 눈치채지 못함. 다만 단시간에 너무 빈번하고 빠르게 각종 질문에 대한 답변을 게시함으로써 발각됨. <sup>5)</sup>

4) <https://arxiv.org/abs/2005.14165>

5) [https://www.reddit.com/r/GPT3/comments/j5lai1/this\\_user\\_is\\_posting\\_with\\_gpt3\\_uthegentlemetre/](https://www.reddit.com/r/GPT3/comments/j5lai1/this_user_is_posting_with_gpt3_uthegentlemetre/)

#### 4. 디퓨전(Diffusion) 기술의 등장

- 텍스트로부터 이미지를 생성하는 기술도 다양한 방법이 제시되었는데, 2020년 6월에 현재 인공지능아트에서 보편적으로 사용되는 디퓨전(Diffusion) 모델이 발표됨.
- 해당 논문의 제목은 "Denoising Diffusion Probabilistic Models"이며 노이즈로부터 이미지를 생성하는 새로운 방법을 제시함. <sup>6)</sup>
- 디퓨전 기술은 이미지를 무작위로 흐트러트리는 노이즈 단계와 흐트러진 패턴으로부터 이미지를 복원하는 디노이즈 단계로 구분되는데 이를 많은 수의 단계로 나누고 훈련시키면 높은 퀄리티의 이미지를 생성하는 것이 가능해짐.

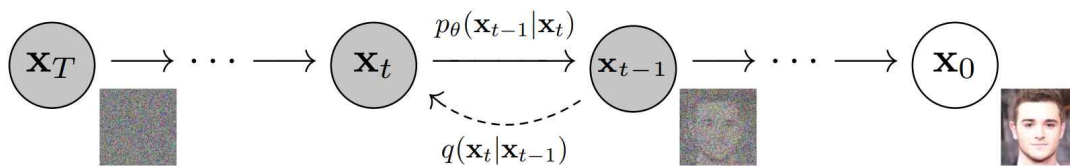
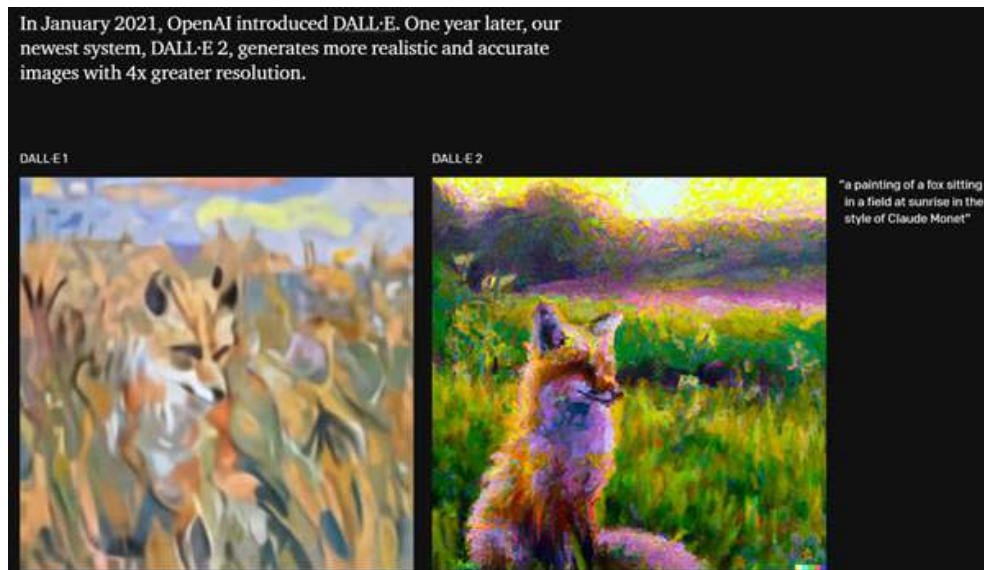


Figure 2: The directed graphical model considered in this work.

자료: 많은 단계의 노이즈와 디노이즈 단계로 구성된 디퓨전 모델 ( 출처 : <https://arxiv.org/pdf/2006.11239.pdf> )

- 오픈AI는 디퓨전 모델에 관련된 연구를 진행해 2021년 11월에 관련 기술인 “GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models”을 발표하고<sup>7)</sup> 2022년 5월 이를 적용한 달리2 서비스를 공개함.



자료 : 달리1과 달리2의 결과 비교 ( 출처 : <https://openai.com/dall-e-2/> )

6) <https://arxiv.org/abs/2006.11239>

7) <https://arxiv.org/abs/2112.10741>



## II. 인공지능아트 의 확산

### 1. 오픈AI와 마이크로소프트와의 파트너십

- 오픈AI는 2015년 12월 인공지능을 활용해 글로벌, 인류 문제를 해결하는 목적을 가진 비영리단체로 출발함.
- 하지만 2019년 오픈AI는 영리를 추구하는 법인으로 변경하고<sup>8)</sup> 그 해 마이크로소프트로부터 10억달러의 투자를 유치함.<sup>9)</sup>
- 2020년 9월, 오픈AI는 자사 기술 GPT-3의 독점 라이선싱 모델을 마이크로소프트에 부여함.
- 이는 이미 초대형화된 GPT-3 기술을 안정적으로 뒷받침할 하드웨어 지원과 회사를 운영하기 위한 재원을 마련하기 위함으로 파악됨.
- 2019년부터 영리로 전환된 OpenAI의 횡보에 많은 이들이 우려를 표현함. 오픈AI 초기 투자자이며, 이전부터 오픈AI가 더욱 개방되어야 한다고 비판했던 테슬라 대표 일론 머스크(Elon Musk)는 독점 라이선싱 모델 발표에 대해 오픈AI가 오픈과 정반대의 행보를 보이고 있으며 마이크로소프트에 장악됐다고 비판함.



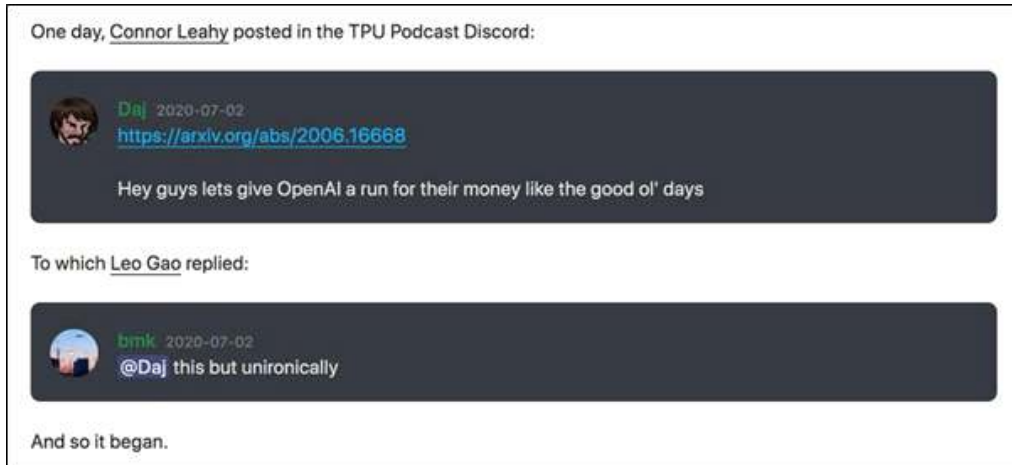
자료 : 일론 머스크의 비판 트윗 ( 출처 : <https://twitter.com/elonmusk/status/1309052632850468864>)

8) <https://openai.com/blog/openai-lp/>

9) <https://openai.com/blog/microsoft/>

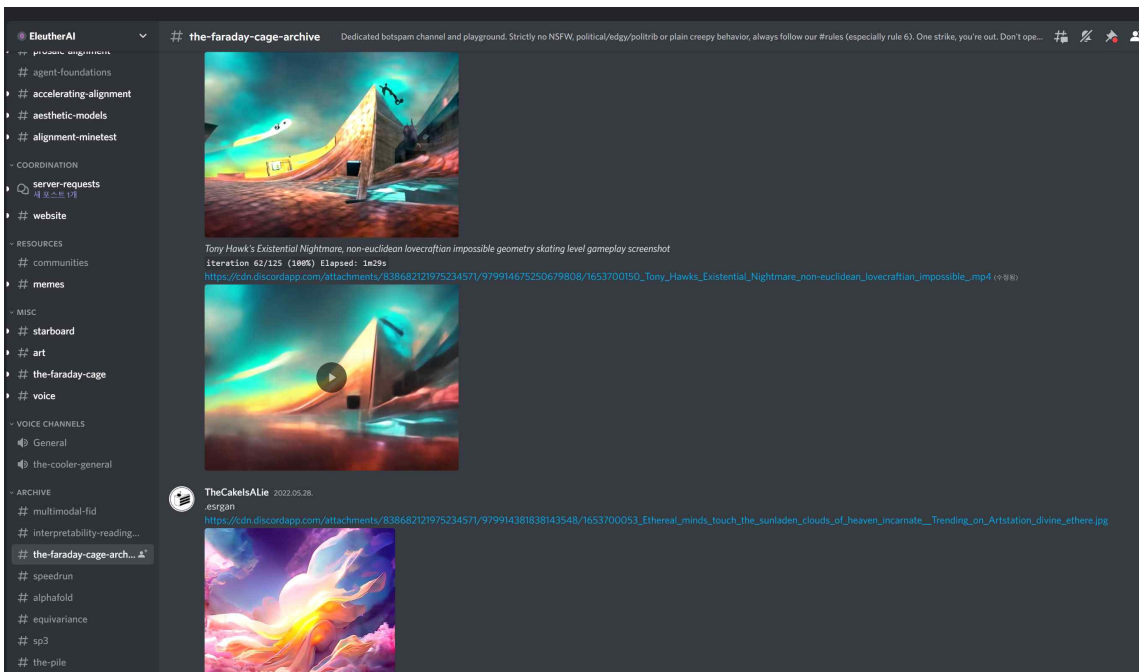
## 2. 개방을 추구하는 인공지능 해커 커뮤니티 일루더시(EleutherAI)의 탄생

- 오픈시의 행보에 못미더운 반응을 보였던 업계와 학계 관계자들이 모여 온라인 커뮤니티 서비스 디스코드(Discord)에서 일루더시라는 이름의 해커 커뮤니티를 조직함.



자료 : 일루더시 커뮤니티의 탄생 ( 출처 : <https://www.eleuther.ai/faq/> )

- 일루더시는 오픈시의 GPT-3를 대체하는 GPT-Neo<sup>10)</sup>, GPT-NeoX<sup>11)</sup>의 프로젝트를 런칭함.
- 해당 커뮤니티는 17,000여명이 가입할만큼 관심을 받고 있으며 GPT 프로젝트를 진행하면서 #the-paraday-cage 라는 채널을 통해 인공지능아트에 대한 테스트도 이루어지고 있었음.



자료 : 디스코드 기반의 자유 커뮤니티 일루더시와 인공지능아트 실험 ( 출처 : <https://discord.com/invite/zBGx3azzUn> )

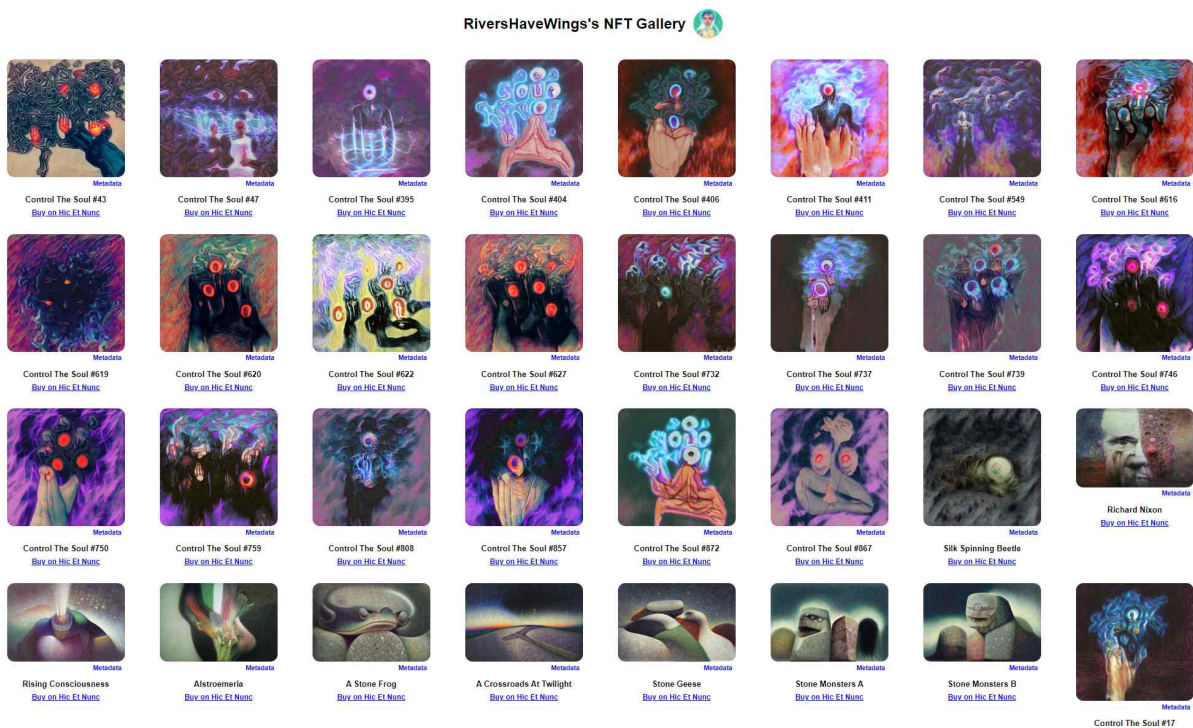
10) <https://github.com/EleutherAI/gpt-neo>

11) <https://github.com/EleutherAI/gpt-neox>



### 3. 인공지능아트 프로그램의 개방과 연구

- 일반인이 인공지능 아트를 직접 생성하기 위해서는 인공지능에 대한 기술 배경을 이해하고, 이를 파이썬 스크립트로 구현한 후 스크립트를 높은 성능의 컴퓨터로 실행해야 하는 등 여러 장벽이 존재했음.
- 하지만 인공지능기술과 스크립트, 클라우드 환경을 이해하는 몇몇 해커 아티스트들은 구글에서 제공하는 클라우드 컴퓨팅인 코랩(Colab)서비스를 활용해 인공지능아트를 생성하는 프로그램을 제작하기 시작함.
- 이러한 해커 아티스트 중 대표적인 사람으로는 일루더시의 일원인 캐서린 크로슨(Catherine Crowson)을 들 수 있음. 캐서린은 코랩 환경에서 오픈시의 CLIP기술과 디퓨전(Diffusion)기술을 조합해 텍스트를 입력하면 이미지가 생성되는 프로그램을 만들고 모든 사람에게 개방함.
- 캐서린은 자신의 인공지능이미지 실험 과정에서 만들어진 이미지를 모아 NFT도 제작하는 등 다양한 활동을 하고 있으며, 트위터 아이디 @RiversHaveWing<sup>12)</sup>을 통해 다양한 인공지능아트를 선보이고 있음.



자료 : 캐서린크로슨의 NFT갤러리 ( 출처 : <https://chainbreakers.kath.io/> )

12) <https://twitter.com/RiversHaveWings>

#### 4. 인공지능아트 프로그램 디스코 디퓨전(Disco Diffusion)의 등장

- 캐서린의 프로그램은 이후 솜나이(Somnai)라는 또 다른 해커 아티스트에 의해 발전하게 됨. 솜나이는 2021년 10월, 캐서린의 프로그램을 기반으로 퀄리티를 높이고 애니메이션 생성과 같은 다양한 기능을 추가해 디스코 디퓨전 v1을 공개함.
- 솜나이는 디스코 디퓨전을 “AI 예술 및 애니메이션 생성을 위한 노트북, 모델 그리고 기술의 프랑켄슈타인 아말감(An amalgamation of notebooks, models and techniques for the generation of AI Art and Animations.)” 이라는 짧은 문장으로 소개함.
- 솜나이 역시 자신의 실험 결과물을 트위터 계정 @Somnai<sup>13)</sup>에 공개함. 이 중에서 디스코 디퓨전을 사용해 생성한 “Journey Across the Kingdom”이라는 애니메이션은 인공지능에 관심있는 아티스트들에게 큰 관심을 불러일으킴.



자료 : Somnai의 AI 애니메이션 Journey Across the Kingdom (출처: [https://twitter.com/Somnai\\_dreams/status/1477525429971939332](https://twitter.com/Somnai_dreams/status/1477525429971939332))

- 구글 계정만 있으면 누구나 솜나이의 디스코디퓨전을 실행해 인공지능이미지를 얻을 수 있었음. 디스코 디퓨전은 인공지능아트를 대중화시킨 촉매제 역할을 하였고, 진보적인 아티스트들은 자신의 작품 활동에 디스코디퓨전을 활용해 인공지능 결과물을 접목함.

13) [https://twitter.com/Somnai\\_dreams](https://twitter.com/Somnai_dreams)

### III. 인공지능아트의 대중화

#### 1. 미드저니(Midjourney) 서비스의 등장

- 디스코 디퓨전이 만들어내는 결과물의 품질은 상당히 높았기에 아티스트들에게 큰 관심을 끌었으나 구글 코랩 환경은 일반인들이 사용하기에는 복잡한 인터페이스를 가지고 있었음.
- 2022년 7월, 디스코드에서 손쉽게 인공지능아트를 생성할 수 있는 미드저니라는 서비스가 새롭게 선보여짐.
- 미드저니는 디스코드에서 프롬프트만 입력하면 4개의 이미지를 제시해주고 사용자가 누른 버튼에 따라 이를 더 발전시킬 수 있는 단순한 인터페이스를 제공함.
- 미드저니는 최초 가입시 무료로 사용할 수 있는 체험 서비스를 제공했으며, 디스코드 채널 내에서 사용자들이 이미지를 생성하는데 사용한 텍스트를 보여줌으로서 일반인들이 편리하게 인공지능아트를 사용하는데 큰 역할을 함.
- 2022년 10월 현재 미드저니에 가입된 사용자 수는 320만명을 넘어섰는데, 이는 디스코드 서비스 역사상 가장 큰 커뮤니티 유저 수임. <sup>14)</sup>
- 미드저니를 통해 제작된 인공지능아트 '스페이스 오페라 극장(Theatre D'opera Spatial)'이 미술대회에서 1위를 차지하면서 사회적인 관심과 토론이 발전하기 시작함.



자료 : 인공지능아트 '스페이스 오페라 극장(Theatre D'opera Spatial)' ( 출처 : 콜로라도 스테이트 페어 인스타그램 <https://www.instagram.com/p/Ch-6osFfoM/> )

14) <https://netcord.site/largest-discord-server-breaks-all-records/>

## 2 스테빌리티 AI(Stability AI)의 등장

- 미드저니는 인공지능아트서비스를 재빠르게 공개해 상업적 성공을 거두었지만, 최초 인공지능아트 연구가 활성화되는데 가장 큰 촉매제 역할을 했던 개방성과는 거리가 있었음.
- 인공지능산업이 지속적으로 발전하기 위해서는 우수한 연구자 풀과 검증된 대형 이미지 데이터베이스 구축이 필수적인데, 빅테크기업이 아닌 이상 이를 모두 감당하기에는 어려움이 있음. 이러한 문제를 근본적으로 해결하기 위해 일루더시가 추구하던 개방을 중심 슬로건으로 생태계를 구축하는 스테빌리티시라는 스타트업이 등장함.
- 2022년 8월 스테빌리티 시는 스테이블 디퓨전(Stable Diffusion)이라는 새로운 인공지능아트 서비스를 발표<sup>15)</sup>하고 자사의 기술을 공개함.
- 스테빌리티 시는 스테이블 디퓨전을 기반으로 일반인들도 손쉽게 사용할 수 있는 프로그램 드림스튜디오(Dream Studio)도 함께 선보임.
- 스테빌리티 시의 슬로건은 “AI by the people, for the people(사람에 의한, 사람을 위한 AI)” 임. 스테빌리티 시는 일루더시와 협력을 맺고 있으며, 인공지능에 활용이 가능한 오픈 이미지 데이터베이스 서비스 라이온(LAION)과도 긴밀히 협력하고 지원하고 있음.
- 라이온 서비스는 2021년도 8월에 4억 크기의 데이터셋인 LAION-400M을 공개했는데, 2022년 3월에는 해당 데이터 셋의 크기가 50억으로 확대된 LAION-5B를 공개할 정도로 급속히 발전하고 있음.<sup>16)</sup>



자료 : 미드저니, 달리2, 스테이블디퓨전의 결과를 비교한 포스팅 ( 출처 : <https://twitter.com/fabianstelzer/status/1561019187451011074> )

15) 스테빌리티 AI 공식 블로그 (<https://stability.ai/blog/stable-diffusion-public-release>)

16) LAION 공식 블로그 (<https://laion.ai/blog/laion-5b/>)



### 3 스테이블 디퓨전 기반의 파생 서비스 등장

- 스테빌리티 시가 공개한 스테이블 디퓨전을 기반으로 다양한 스타일의 인공지능아트를 생성하는 파생 서비스들이 등장함.
- 대표적인 스테이블 디퓨전의 파생 서비스로는 미국 기업 Anlatan이 공개한 NovelAI를 들 수 있음.
- 2021년 6월에 서비스를 시작한 NovelAI는 이름 그대로 소설을 자동 생성하는 AI서비스 였으나 2022년 10월에 이미지 생성 기능을 발표함.<sup>17)</sup>
- 스테이블 디퓨전의 모델과 학습 프로세스를 변경해 제작한 NovelAI 서비스는 만화풍의 2차원 일러스트에 특화되어 있으며, 스테이블 디퓨전보다 훨씬 높은 정밀도의 일러스트의 생성이 가능한 것으로 평가받고 있음.



자료 : 노블AI가 만들어내는 인공지능아트 결과물 ( 출처 : <https://blog.novelai.net/image-generation-announcement-807b3cf0afec> )

- NovelAI의 기술은 기존 CLIP 모델에서 이미지를 학습할 때 화면 중앙을 기준으로 이미지를 1:1의 가로세로비로 자르기 때문에 발생하는 문제를 개선하고 소설 생성을 위해 자사가 개발한 모델인 하이퍼네트워크(HyperNetworks)을 스테이블 디퓨전에 응용함으로써 이미지 생성의 정확도를 높였다고 블로그에 밝힘.<sup>18)</sup>

17) 노블 AI 공식 블로그 (<https://blog.novelai.net/image-generation-announcement-807b3cf0afec>)

18) 노블 AI 공식 블로그 (<https://blog.novelai.net/novelai-improvements-on-stable-diffusion-e10d38db82ac>)

## IV. 결론

### 1. 생태계의 중요성

- 스태빌리티 AI가 현재 가장 주목받고 있는 이유는 기술의 개방이라는 명분을 중심으로 학계와 커뮤니티에 큰 호응을 얻고 있기 때문임.
- 연구자 집단, 학습을 위한 거대 데이터 셋, 모델을 운용할 수 있는 하드웨어 장비의 세 가지가 모두 만족되는 환경이 구축되어야 함

### 2. 다양한 직군의 참여

- 인공지능아트 산업을 연구하는데 있어서 아티스트들의 참여도 중요함. 트위터프로필을 통해 디스코디퓨전을 제작한 솜나이는 현재 미드저니에서 근무하고 있으며, 최초 모델을 설계한 캐서린크로손은 스태빌리티AI에서 근무하는 것으로 보여짐.
- 오래전부터 구글에서 AI와 아트를 접목한 연구자인 David Ha가 최근 스태빌리티 AI로 둠지를 옮긴다고 공식 발표함.<sup>19)</sup>

### 3. 향후 발전 가능성

- 디스코디퓨전을 시작으로 대중들이 사용할 수 있는 인공지능아트 서비스가 선보이고 대중화되기까지 기간은 1년도 걸리지 않았음.
- 오픈 생태계를 표방하는 스태빌리티 AI는 2022년 10월 약 1억달러의 자금을 유치했다고 밝힘. 2022년 10월 스태빌리티 AI의 회사가치는 10억 달러로 평가됨<sup>20)</sup>
- 인공지능아트 기술은 큰 틀에서 CLIP과 디퓨전으로 정리되고있지만, NovelAI의 사례에서 볼 수 있듯이 이를 개량해서 특화시킬 수 있는 응용사례들은 앞으로 많이 등장할 것으로 예상됨.
- 최근들어 엔비디아, 메타, 구글등의 빅테크들이 급하게 인공지능아트 관련된 다양한 기술을 발표하고 있음. 마이크로소프트 역시 오픈AI의 달리2 기능을 자사 서비스에 탑재시킨다고 발표함. <sup>21)</sup>
- 현재 기술과 빅테크 기업들이 발표하는 기술이 융합되면 빠른 시간에 2D에서 3D로의 확장도 가능할 것으로 예상됨.

19) David Ha 트윗 ( <https://twitter.com/hardmaru/status/1579350954537017344>)

20) <https://techcrunch.com/2022/10/17/stability-ai-the-startup-behind-stable-diffusion-raises-101m/>

21) <https://techcrunch.com/2022/10/12/microsoft-brings-dall-e-2-to-the-masses-with-designer-and-image-creator/>



## V. 문화체육관광분야 시사점

- 구글의 딥마인드가 개발한 인공지능 알파고와 이세돌 9단의 대국이 있은지도 벌써 6년이 흘렀는데 그 사이 인공지능이 바둑 분야에 끼친 영향은 상당함. 2022년 10월에 있었던 인터뷰에서 이창호 9단 역시 인공지능을 활용한 바둑을 공부하기 시작했다고 밝힘.<sup>22)</sup>
- 인공지능아트는 향후 발전 가능성이 풍부하고, 대형 투자가 이루어지는 유망 기술 분야이므로, 다양한 문화 산업에 지대한 영향을 끼칠 것으로 예상됨.
- 현재 인공지능아트 학습에 사용하는 이미지 데이터셋은 서양에서 만들어진 자료가 대부분이다 보니 한국적인 이미지를 생성하기가 불가능한 상황임.
- NovelAI의 사례에서 볼 수 있듯이 정교하게 잘 분류된 이미지 데이터셋을 준비해 학습시키면 다양한 모델을 사용해 다양한 스타일의 이미지를 제작할 수 있음.
- 향후 인공지능아트 시대를 대비하기 위해, 한국형 이미지 데이터베이스 구축 방안을 모색하고, 아티스트들이 향후 작업에서 인공지능을 잘 활용할 수 있도록 체계적인 커리큘럼과 교육을 준비하는 것이 필요할 것으로 생각됨.

22) 타이젬TV [특별한 만남] 이창호 9단이 바라보는 인공지능은? <https://www.youtube.com/watch?v=5Wqg-5gydEM>