

기술철학적 관점을 반영한 생성적 인공지능 예술창작 사례 연구

조영각¹ · 김현주^{2*}

¹서울미디어대학원대학교 융합예술디자인학과 객원연구원

²서울미디어대학원대학교 융합예술디자인학과 교수

Generative AI Art on the View of the Philosophy of Technology

Youngkak Cho¹ · Hyun Ju Kim^{2*}

¹Researcher, Dept. of Convergence Art & Design, Seoul Media Institute of Technology, Seoul 07590, Korea

²Professor, Dept. of Convergence Art & Design, Seoul Media Institute of Technology, Seoul 07590, Korea

[요약]

예술가와 인공지능 간의 존재론과 창의성의 문제를 넘어서, 인공지능 예술의 창작 사례와, 그 미적 함의 및 해석에 대한 연구는 아직 충분하지 않은 실정이다. 본 연구는 ‘창발과 복잡계’, ‘순수언어’, ‘공명’ 등 현재 인공지능과 관련한 기술철학적 논의를 소개하고 이를 인공지능 예술 창작 사례에 적용해 봄으로써 인공지능 예술의 고찰을 위한 이론적, 개념적 자원을 공유하고자 한다. 이를 바탕으로 2022년 생성형 인공지능이 본격적으로 국내에 도입되기 전 일부 작가들이 실험적으로 프롬프트엔지니어링을 예술 현장에서 적용한 초기 사례로 <까마귀속 연작> 작품의 제작과정과 의미를 분석한다. 특히 인공지능이 단순한 유행하는 도구가 아닌 예술 창작의 과정에서 인간과 상보적 역할을 하는 지능적 수행체로 바라볼 것을 제안한다. 본 연구는 인공지능 예술이 가지는 미적 함의와 해석에 있어 다양성을 제공하고, 역사적 맥락에서 인공지능 예술의 발전 과정과 사례를 살핌으로써 앞으로 나아갈 방향과 과제를 전망하는데 의의를 둔다.

[Abstract]

The study of AI art creation, as well as the aesthetic implications and interpretations of AI art, remains insufficient, beyond the ontological and creative issues between artists and AI. This research introduces perspective of philosophy of technology related to AI, such as ‘emergence and complex systems,’ ‘pure language,’ and ‘resonance,’ and applies these concepts to cases of AI art creation. By doing so, the study seeks to share theoretical and conceptual resources for contemplating AI art. Based on this foundation, the research analyzes the production process and meaning of <Corvus Series>, an early example of prompt engineering experimentally applied in the art field by some artists before the full-scale introduction of generative AI in Korea in 2022. The study particularly proposes that AI should be positioned as an ‘intelligent agent’ that plays a complementary role with humans in the creative process, rather than merely as a trendy tool. The significance of this research lies in providing diversity in the aesthetic implications and interpretations of AI art, while also examining the developmental process and cases of AI art within a historical context, thus offering insights into future directions and challenges.

색인어 : 인공지능, 인공지능예술, 프롬프트 엔지니어링, 기술철학, 인공지능미학

Keyword : Artificial Intelligence, AI Art, Prompt Engineering, Philosophy of Technology, AI Aesthetics

<http://dx.doi.org/10.9728/dcs.2024.25.9.2525>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 01 July 2024; Revised 05 August 2024

Accepted 29 August 2024

*Corresponding Author; Hyun Ju Kim

Tel: +82-2-6953-7174

E-mail: hjkim@smit.ac.kr

I. 서론

새로운 매체와 디지털 기술은 그 등장부터 비선형적인 특성과 리좀(rhizome) 형태의 다방향성을 내포하여 커뮤니케이션의 형식과 방법에서 근본적인 변화를 불러일으켰다. 이 변화는 예술의 정치적인 지형이 지속해서 변화하는 원인으로 작용하며, 오늘날 더욱 가속도가 붙고 있다. 이러한 상황 속에서 인공지능 기술은 기존의 기술에 대한 관념을 넘어서는 다양한 방식으로 기능하여 새로운 관념 체계를 만들어 가고 있다. 또한 그 변화의 가속도는 무어의 법칙 이상으로 비정상적인 수준으로 상당히 크다.

인공지능이 생성한 이미지가 예술작품으로 만들어지고 전시되며, 또 판매되는 등의 예술에 대한 파생 활동이 일어나면서, 예술계에서 “인공지능이 인간의 창의성을 대체할 수 있는가?”, “예술가는 인공지능의 시대에 어떠한 지위로 살아남을 것인가?” 등등 예술가와 인공지능의 존재론과 창의성의 문제는 충분히 논쟁적인 이슈였다. 한편 인공지능 예술 창작의 사례와, 인공지능 예술 작품이 가지는 미적 함의와 해석에 대한 연구는 아직까지 충분히 이루어지지 못한 실정이다.

본 연구는 인공지능 예술을 기존의 기술철학적 주요 개념을 바탕으로 해석하고 그 미적 함의를 분석하고자 한다. 이는 지금까지 많은 관심 속에 시도되고 있지만 그 결과에 대한 학문적 논의가 부족한 인공지능 예술 창작물이 좀 더 의미 있게 해석될 수 있는 이론적, 개념적 자원을 공유한다는 데 의의가 있다.

우선 과거 인공지능 예술의 역사적 맥락과 발전 단계를 소개하고, 오늘날의 인공지능 예술의 기술적 현황을 살핀다. 또한 ‘창발과 복잡계’, ‘순수언어’, ‘공명’ 등 현재 인공지능과 관련한 기술철학적 논의를 소개하고 이를 인공지능 예술의 미학적 해석에 적용하고자 한다. 이 세 가지 개념은 인공지능 예술의 본질적 특성을 포괄적으로 설명할 수 있는 핵심 이론들로, 각각 인공지능 예술의 예측 불가능성과 새로운 의미 생성, 매체 간 전이 가능성, 그리고 시공간을 초월한 상호작용을 설명하는 데 적합하기에 인공지능 예술의 다층적 의미를 더욱 체계적으로 분석하는데 유용한 개념이라 여겨진다.

이러한 접근은 인공지능을 예술가를 대체하는 위협적인 존재로 보기보다는, 이를 예술가의 창작 활동에 대한 또 다른 차원의 기술이자 미디엄으로써 예술가와 함께 보완적 역할을 주고받을 수 있는 ‘지능적 수행체’(intelligent agent)로 바라볼 것을 제안한다. 즉, 인공지능을 단순한 기계나 인간의 능력을 모방하는 도구 이상을 넘어, 창발적 특성과 학습 능력을 보유하고 있으며, 인간과 복잡한 상호작용을 할 수 있는 수행체로 접근하는 것을 뜻한다. 이로써 프롬프트 엔지니어링 기반 인공지능 예술 작품을 생성하는 과정과 결과가 어떠한 미적 함의를 지니는지 좀 더 근본적인 현상에 접근해 논의가 가능할 것으로 본다.

또한 본 논문에서는 현재의 인공지능 예술이 촉발하게 된

프롬프트 엔지니어링의 기술적인 특성을 분석하고, 그 속성과 방법 등에 관하여 기술한다. 그리고 프롬프트 엔지니어링이 인공지능 예술 창작에 어떻게 적용되었는지, 그 미적 가능성은 어떠한지 선행한 이론적 연구를 적용해 연구자의 <까마귀속 연작> 시리즈의 세 작품에 적용해 분석한다. 특히 <까마귀속 연작>은 2022년 국내에 당시 생성형 인공지능이 본격적으로 도입되기 전 일부 작가들이 실험적으로 프롬프트 엔지니어링을 예술 현장에서 적용한 초기 사례 중 하나로 의미가 있다. 따라서 인공지능 기술이 급격히 발전하는 상황에서, 국내 인공지능 예술의 초기 적용 과정이 어떠한 양상으로 진행되었는지 기록한다는 점에서도 본 분야에 기여할 것으로 본다. 궁극적으로 인공지능 예술이 가지는 미적 함의와 해석에 있어 다양성을 제공하고 향후 인공지능 예술이 나아갈 방향과 과제를 제안하고자 한다.

II. 인공지능 예술의 개념과 발전 양상

인공지능 예술(Artificial intelligence art, AI art)은 인공지능 기술을 바탕으로 하는 다양한 일련의 시도가 예술의 분야에 적용된 예술 장르 혹은 형식, 방법에 해당한다. 이 분야는 근 몇 년간의 성취를 바탕으로 논의되는 경우가 다반사이다. 이는 특정한 변곡점을 넘은 딥러닝(Deep Learning) 분야의 폭발적인 성장과 맞물려 있다. 인공지능(artificial intelligence)을 대범주로 본다면, 그 하위에 기계학습(machine learning)이 자리하고, 그 하위에 딥러닝이 자리한다. 이처럼 한정된 분야의 급속한 발전이 전체인 것으로 오해되곤 한다. 하지만 이전에도 인공지능에 대한 오랜 관심만큼 예술 분야에서도 다양한 접근이 시도되었다. 일례로 20세기 현대 음악 중 알고리즘 작곡의 대표적인 마르코프 체인(Markov Chain) 방식은 수학적 접근과 통계의 상관성을 기본으로 한다. 이는 인공지능 예술 창작의 초기의 한 형태로 바라볼 수 있다. 특정한 알고리즘을 기반으로 하며, 데이터를 근간으로 한다는 기본적인 환경 조건, 마지막으로 출력값의 사용까지 현재의 인공지능 예술이 이루어지는 양태와 매우 흡사한 구조로 되어 있다. 따라서 인공지능 예술을 생성 예술(Generative Art)의 한 갈래로 바라보는 시각도 존재한다.

인공지능이라는 대범주의 기술 영역 자체가 워낙 광범위하고, 빠른 속도로 분화되며 발전하는 분야이기에 이를 영역별로 나누어 구조화하기는 쉽지 않다. 대다수가 동의하는 특정한 분기점은 2016년의 알파고 이슈를 전후로 한 인공지능 기술이 보편적으로 통용되는 지점과 맥락을 같이한다.

이 지점부터는 인공지능 예술은 일반적으로 두 가지 방식으로 이해될 수 있다(그림 1). 첫 번째는 예술을 위한 분석적 접근 방식(Tech for art approach)으로, 예술 창작에 있어 인공지능의 분석 및 예측 기술을 활용하는 방식이고, 두 번째는 생성적 예술의 접근 방식(Art generation approach)으로

새로운 예술을 창작하는 인공지능으로 분류된다. 첫 번째의 경우는 이미지 분류(image classification), 객체 검출(object detection) 등의 분석적 기술이 기존의 예술에 적용되는 경우로 본문에서 다루고자 하는 생성적 인공지능 분야와는 그 궤를 달리하기에, 두 번째 경우를 기준으로 인공지능 예술에 대하여 더 알아보도록 한다.

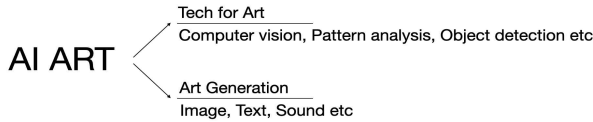


그림 1. 인공지능 예술의 방식
Fig. 1. Two approaches of AI art

여기서도 인공지능이 새로운 창작을 위한 매개의 역할로서 기능하는 객체에 대하여 파악하는 것이 인공지능이 자체적으로 예술을 판단하고 창조해 낸다는 견해는 포함하지 않고 본문을 작성한다.

이 외에도 로보틱스 분야 등 기술을 기반으로 하는 예술 분야에 인공지능의 접목은 가속화되고 있으나, 뚜렷한 발전과 변화되는 현상을 포착할 수 있는 시각 분야를 위주로 본문에서는 다룬다. 그림 2와 같이 인간이 향유하기 위해 만들어 내는 각 예술 분야에 기술의 영역이 포함된 미디어아트가 발전되어 온 이래, 현재는 인공지능과 각 첨단 기술이 예술의 영역에 다양하게 녹아들어, 인공지능 예술이라는 분야는 그 교류의 중심에서 하나의 개념으로 자리 잡아가는 중이다.

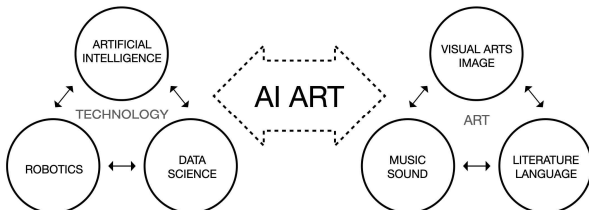


그림 2. 인공지능 예술의 분야 : 기술과 예술 학제
Fig. 2. Areas of AI art: technologies and art disciplines

2-1 생성적 인공지능 예술의 발전

인공지능을 바탕으로 하는 생성 예술(Generative Art)의 역사는 컴퓨터 그래픽의 초기와 컴퓨터 발명으로 거슬러 올라갈 수 있다. 1950년대와 1960년대에 컴퓨터 그래픽은 단순한 패턴과 모양을 생성하는 데 사용되었다. 이 분야의 초기에는 컴퓨터 화면에 렌더링 된 패턴을 만드는 기본 알고리즘을 사용하여 만들어졌다.

가장 대표적인 예는 독일 출신의 미국의 작가 만프레드 모어(Manfred Mohr)이다. 그는 1960년대 후반 알고리즘을 바탕으로 컴퓨터 그래픽을 사용한 이미지로 작품을 선보였다. 이는 컴퓨팅 파워를 사용하고, 이에 대한 소프트웨어적인 구

현에서 알고리즘이 작용하여 특정한 디스플레이를 출력하여 준다는 점에서 인공지능 예술의 원료로 구분될 수 있다.

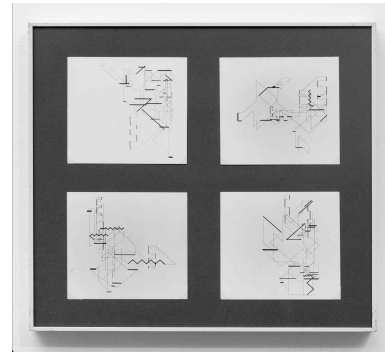


그림 3. 만프레드 모어, P018-mf_11-14-20-21. 사진용지에 라이트빔 플로터를 활용한 드로잉. 12×12cm each. 1969[2]

Fig 3. Manfred Mohr, P018-mf_11-14-20-21. Light beam plotter drawings on photo paper. 12×12cm each. 1969[2]

이외에는 베라 몰나(Vera Molnar), 마이클 놀(Michael Noll) 등이 이 분야에서 선구자적 위치를 차지한다[1],[2].

본격적인 인공지능 예술의 시초라 할 수 있는 예는 1973년 해럴드 코헨(Harold Cohen)이 아론(AARON)이라는 일련의 알고리즘을 개발하여, 컴퓨터가 불규칙성을 사용하여 자유 형태의 이미지를 그릴 수 있도록 한 사례이다[3].



그림 4. 해럴드 코헨이 AARON을 활용해 작업하는 모습, 보스턴 컴퓨터 박물관 1982(컴퓨터 역사 박물관 소장)[3]

Fig. 4. Harold Cohen coloring the forms produced by the AARON at the Computer Museum, Boston, MA, ca. 1982. (Collection of the Computer History Museum)[3]

아론은 특정 물체를 그리도록 프로그래밍 되었으며, 코헨은 지시 중 일부가 이전에 상상하지 못한 형태를 생성한다는 것을 발견할 수 있었다. 이는 기계가 예술적 결정과 같은 것을 할 수 있도록 명령을 설정했다는 것을 발견한 것이다.

1970년대와 1980년대에 AI 생성 예술은 CAD(Computer-Aided Design)에서 더욱 광범위하게 사용되기 시작했다. CAD 소프트웨어를 사용하면 설계자가 컴퓨터에서 3차원 모양을 만들고 조작할 수 있게 되었다. 이를 통해 더욱 복잡하고 사실적인 이미지를 만들 수 있게 되었다. 이 예시가 아론에 해당하며 처음에는 추상 회화를 만들었으며 1980년대와 1990년대를 통해 암석과 식물, 인간의 그림 등을 포함하여 더욱 복잡한 예술로 성장하였다[4].

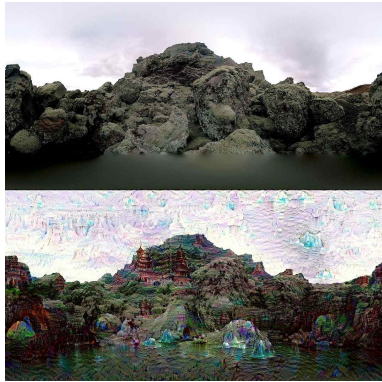


그림 5. 딥드림 생성 이미지[5]
Fig. 5. Generated images of DeepDream[5]

현재적 의미의 인공지능 예술의 기점이 되는 예는 구글의 딥드림(DeepDream)이다[5]. 알렉산더 모드빈체프(Alexander Mordvintsev), 마이크 타이가(Mike Tyka), 크리스 올라(Chris Olah)는 ‘인셉셔니즘(inceptionism): 신경망으로 더 깊이 들어가기’라는 블로그의 글을 통해 딥드림을 공개하였다. 기존에 학습된 인공지능 모델(CNN)에서 의도치 않은 출력값을 출력하도록 오류 역전과 학습을 반복시켜 피쳐 맵에 환각적인 시각효과를 만들어 내는 인공 신경망 기술을 사용했다. 의도치 않은 출력값은 인간이 환각 상태에 빠져있을 때의 시각 경험과 유사하다는 점에서 신경생리학자들의 연구 대상이 되기도 하였다[7].

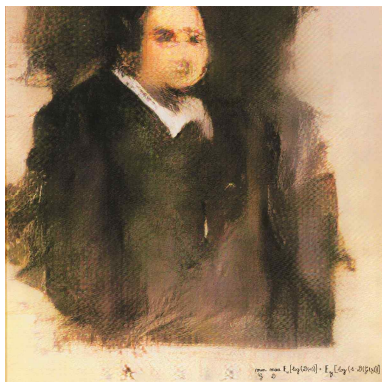


그림 6. 오비어스의 <에드몬드 드 벨라미의 초상화>(2018)[6]
Fig. 6. Obvious's <Portrait of Edmond de Belamy>(2018)[6]

이후 빠른 발전을 통해 심화된 인공지능 예술은 미술 시장에 진출까지 하게 되었다. GAN으로 만들어진 작품으로 미술 시장에서 판매된 최초의 인공지능 작업물로 평가되는 오비어스(Obvious)의 <에드몬드 드 벨라미의 초상화(Portrait of Edmond de Belamy, 2018)>가 등장하게 된 것이다[6].

이 작업은 뉴욕의 크리스티(Christie's)에서 약 43만 2500 달러라는 놀라운 가격에 판매되었다. 선구자적 입장을 가진 특정한 작가가 아닌 일종의 마케팅 수단의 형태로 해당 작품을 미술계에 편입시키려 한다는 비난의 대상이 되기도 하였다[8]. 또한, ING와 마이크로소프트에서 진행한 ‘더 넥스트 렘브란트(The Next Rembrandt)’ 프로젝트는 기존 렘브란트의 화풍을 학습시켜, 그가 생전에 사용한 형태와 동일한 화풍으로 이미지를 생성하도록 구성하였다.

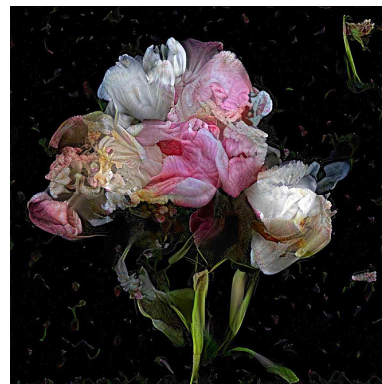


그림 7. 소피아 크레스포의 <신경망 동물원>(2020)[9]
Fig. 7. Sofia Crespo's <Neural Zoo>(2020)[9]

생성적 인공지능 예술의 최신 사례 중 인공지능 생명 개념까지 진보한 흥미로운 예 중 하나는 <신경망 동물원(Neural Zoo, 2020)>이다. 이 프로젝트는 작가 소피아 크레스포(Sofia Crespo)가 진행하며, 시지각을 통해 사변적이고 인공적인 삶을 탐구하는 프로젝트로 현재 지속해서 프로젝트를 진행 중이다. 작가는 인공지능을 통해 본질적으로 완전히 새로운 생물학적 분류 세트를 필요로 하는 상상된 기능을 가진 왜곡된 일련의 가상의 생명체를 형성했다. 이 인공지능 예술은 자연이 제공하는 끝없는 다양성과 함께 작동하지만, 관객들에게는 아직 이러한 개념의 인공지능 예술에 대하여 인식 자체가 제한적일 수밖에 없다.

2-2 인공신경망 기반의 인공지능 예술 창작

현재는 빠른 속도로 인공지능을 통해 예술을 창작하는 다양한 방법이 있다. 인공지능 알고리즘은 매개변수 세트를 기반으로 이미지 또는 비디오를 생성하거나 기존 이미지를 결합 및 변경하여 새로운 이미지를 생성할 수 있다. 신경망의 경우는 특정 예술가의 스타일을 모방한 이미지 또는 비디오를 생성하거나 특정 유형의 예술과 유사한 이미지 또는 비디

오를 생성하는 데 사용할 수 있다.

먼저 심층 신경망 네트워크(Deep Neural Networks)를 이용하여 작품의 스타일을 다른 작품으로 옮기는 방식을 신경망 스타일 전이(NST-Neural Style Transfer) 라고 한다. 2015년 튀링겐 대학교에서 처음 제안된 신경망 스타일 전이의 기본 아이디어는 입력 이미지의 스타일 표현을 얻기 위해 원래의 텍스처 정보를 캡처하도록 설계된 특징 공간(feature map)을 사용한다는 것이다.

이 기능적인 공간은 네트워크의 각 계층에서 필터의 응답 위에서 구축된다. 이는 기능 맵의 공간 범위에 대한 서로 다른 필터의 응답 사이에 상관관계로 구성된다. 이후 여러 계층의 특징과 상관관계를 포함하여, 입력 이미지의 고정된 다중 스케일 표현이 도출된다. 이 표현의 경우는 텍스처 정보는 캡처하지만, 전역 배열은 캡처하지 않는 형태이다. 이를 통해 합성곱 신경망(CNN, Convolutional neural network)에서 콘텐츠와 스타일의 표현이 분리될 수 있다는 것을 실험을 통해 발견했다. 즉, 두 표현이 모두 독립적으로 조작되어 시각적으로 의미 있는 새로운 이미지를 생성할 수 있게 된 것이다. 이 발견은 인공지능 예술에 사용되는 신경망 스타일 전이에 대한 모든 연속적인 방법의 기본 기초가 되었다[10].

위의 신경망 스타일 전이 외에도 인공지능 예술을 창작할 수 있는 다른 알고리즘들이 있다. 이안 굿펠로우(Ian Goodfellow)와 동료들의 논문에서 2014년 제안된 GAN(Generative Adversarial Networks)은 일반적으로 두 개의 신경망이 서로 맞물려 더 나은 학습자로 구성된다.

이미지 분류를 위한 데이터 세트를 보강하기 위해 새 이미지를 생성해야 한다고 가정한다. 두 네트워크 중 하나를 새로운 이미지를 출력하는 심층 네트워크인 생성기(generator)라 한다. 또 다른 네트워크는 판별기(discriminator)라고 하며 입력으로 제공된 이미지가 생성기에 의해 생성된 원본 이미지인지 가짜 이미지인지 분류하는 역할을 한다. 연속적인 반복에서 생성기는 판별자를 속이기 위해 원본 이미지를 더 가깝게 모방하려고 시도하는 반면 판별자는 실제 이미지와 가짜 이미지를 더 잘 구별하려고 시도한다[11]. 이러한 일종의 적대적 게임은 두 네트워크가 모두 훈련하도록 만든다. 훈련의 루프가 완료되면 생성기는 실제 이미지를 출력할 수 있으며, 판별자는 좋은 분류기 모델(classifier model)이 된다. GAN이 활발하게 쓰인 것은 그 작동 원리를 사용하여 다양한 장르에 적용될 수 있었기 때문이다. 글꼴, 일러스트레이션, 만화 및 애니메이션 캐릭터, 스케치 등의 이미지 생성에 특화된 다양한 장르로 파생되어 GAN이 사용된다.

2-3 프롬프트 기반 인공지능 예술

1) 프롬프트를 기반으로 한 엔지니어링

인공지능 예술은 그 영역의 특성상 현재 짧은 시간 내에 빠른 속도로 분화하고 있다. 예술 분야 이전에 본래의 분야인

공학에서의 인공지능 기술은 서비스의 형태에 더욱 특화하여 분화하는 중이다. 그 대표적인 분야로는 이미지 검색, 자연어 생성, 음성 인식, 가상 에이전트, 기계학습 플랫폼(machine learning platforms), AI 최적화 하드웨어, 의사 결정 관리, 깊은 학습 플랫폼(deep learning platforms), 로봇 프로세스 자동화, 텍스트 분석 및 자연어 처리(NLP, natural language processing) 등이라 할 수 있다. 특히 자연어 처리 분야와 이미지 생성 분야가 결합하는 과정에서, 생성적 인공지능 창작은 프롬프트 엔지니어링(prompt engineering)을 통해 더욱 대중화가 가속되고 있다.

프롬프트(prompt)는 거대언어모델(LLM)에 제공되는 일련의 지시 사항으로, 이를 통해 LLM을 사용자 맞춤형하거나 기능을 향상 및 개선하기 위해 사용한다[12]. 본래 프롬프트(prompt)란 표준국어대사전의 명사로 등재되기로는 연극 분야에 해당하는 용어로 ‘연극을 공연할 때 관객이 볼 수 없는 곳에서 배우에게 대사나 동작 따위를 일러 주는 일’로 명시되어 있다. 또 이 단어는 정보통신 용어로서의 명사로 ‘시스템이 다음 명령이나 메시지, 또는 다른 사용자의 행동을 받아들일 준비가 되었음을 사용자에게 알려 주는 메시지’라 설명되어 있다. 즉, 인간의 언어 중에서 특정한 의도를 가지고 소통할 목적으로 생산한 모든 인공물을 이르는 용어로 텍스트를 기반으로 그 층위 자체에 어떠한 목적을 가지고 특정한 객체에 설명을 담은 형태이다. 이에 더하여 프롬프트 프로그래밍(prompt programming)이란 프롬프트라는 단어의 의미와 유사한 형태로 자연어를 기반으로 하며, 프로그래밍 언어를 사용하지 않고 적용할 수 있도록 구성된 형태이다. 프롬프트 엔지니어링(prompt engineering)은 LLM이 프롬프트를 통해 프로그래밍이 되는 방식을 의미한다.

프롬프트 엔지니어링은 이미지 분야보다는 우선 언어의 분야에서 먼저 발전을 이루었다. 특히 자연어 처리 분야는 인공지능 분야에서도 가장 많은 연구가 이루어진 분야였다. 이는 곧 딥러닝 분야에서 다양한 생성이 가능한 모델들로 확인되었지만, 2024년 현재 높은 성능 수준으로 평가되는 언어모델은 오픈에이아이(OpenAI)의 GPT-4o(Generative Pre-trained Transformer 4o)와 엔트로픽(Anthropic)의 Claude-3 등이다. 이러한 모델이 등장하는 데는 트랜스포머(Transformer) 모델이 큰 역할을 하였다. 구글(Google)의 2017년 논문에 처음 등장한 이 모델은 문장 속 단어와 같은 선형적인 순차의 데이터 내의 관계를 추적하여 맥락과 의미를 학습하는 신경망이다. 어텐션(attention)이라 불리며 진화를 거듭하는 수학적 기법을 응용해 데이터 요소들이 서로 떨어져 있는 상황에서, 관계에 따라 미묘하게 의미가 달라지는 부분까지 감지하는 것이 핵심이다[13].

2) 텍스트 투 이미지(TEXT-TO-IMAGE)를 위한 인공지능 모델

텍스트 투 이미지(TEXT-TO-IMAGE)는 한글로 변환하면 ‘글에서 이미지로’이다. 이는 인공지능 예술 중에서 시각

분야에 대하여 프롬프트 엔지니어링을 적용한 분야에 대해서 다양한 모델을 장르 형태로 묶어서 부르는 명칭으로 굳어져 가고 있다. 이 명칭에 해당하는 인공지능 모델 들은 각각의 특성을 갖고 있으며, 데이터와 알고리즘에 따라 특화된 결과물을 생성한다. 또한 이때 사용되는 모델들은 거대 인공지능 모델로 분류되며, 그 데이터는 매년 더욱 큰 크기의 내용을 포함하고 있다.

인공지능 분야는 특성상 뛰어난 알고리즘과 학습 방법을 바탕으로 급속한 전개가 이루어지는 경우가 많다. 이 분야의 경우는 클립(CLIP)과 제로 샷(Zero shot)학습 등이 이에 해당한다. 본 논문에서는 공학적 접근을 통한 기술적인 분석은 다루고 있지 않지만, 간략하게 살펴보자면 클립은 언어를 이미지로, 이미지를 언어로 이해할 수 있도록 일종의 ‘연결고리’를 담당하는 딥러닝 모델에 해당된다. 제로 샷 학습은 인공지능 모델이 학습 과정에서 배우지 않은 작업을 수행하는 형태이다. 이 두 가지의 인공지능 모델의 조합으로 일종의 ‘연상’ 능력에 해당하는 기능을 수행할 수 있게 된 것이다. 이 부분이 핵심적인 작용의 원리에 해당한다[13].

텍스트를 바탕으로 이미지를 생성하는 인공지능은 현재 기업과 오픈소스 진영의 경합에서 벗어나 서비스의 영역으로 확대되고 있다. 2022년 당시의 대표적인 생성적 인공지능 이미지의 모델 혹은 애플리케이션은 다음과 같다. 달 리 3 (Dall-E 3), 미드저니(Midjourney), 웬보(Wombo), 아트 브리더(Art breeder), 드림부스(DreamBooth) 등 인공지능 기술 바탕의 새로운 기업과 단체가 있다면, 구글의 이매젠(Imagen), 파티(Parti)와 메타(META)의 메이크 어 신(Make-A-Scene) 등 기존의 거대 IT 기업의 리서치 랩들이 다양한 인공지능 모델을 선보이는 중이었다. 이러한 인공지능 모델들의 가장 중요한 세 가지 시스템 자원(system resource)은 역시 데이터와 알고리즘, 컴퓨팅 파워(GPU)이다. 이 자원적 측면에서 기업이 이윤을 추구하면서 발생하는 불균형의 강도는 더욱 심화하고 있으며, 이를 타개하기 위한 오픈소스 진영의 반발도 만만치 않은 상황이다.

3) 딥러닝 프레임워크, 이미지의 생성과 영상화

인공지능 예술 분야에서 현재는 텍스트 두 이미지를 넘어서 단순히 이미지 낱장을 생성하는 것이 아니라 연속적으로 이미지를 영상의 프레임 개념으로 생성하여 비디오로 전환하는 형태가 등장하였다.

이 형태는 확산(Diffusion) 모델에 해당하며, 특히 구글의 코랩(Colab), 아마존 웹 서비스(AWS, Amazon Web Services)와 같은 가상 컴퓨터(VM: Virtual Machine) 등의 인공지능 학습을 위한 보편적인 컴퓨팅 환경이 개선됨에 따라 더욱 쉽게 접근할 수 있게 되었다. 확산 모델은 기본적으로 GAN의 알고리즘과 동일한 배경을 가지고 있으며, 이 방식에 따라 노이즈와 안티노이즈가 지속적인 관계를 형성하며 이미지를 생성하여 비디오로 전환되는 방식을 따른다. 또한 잠재확산모델(Latent Diffusion Model)을 통해 확산모델의 효율성이 증

대하였고, 고해상도의 영상 제작이 어느 정도 가능해졌다[14].

2022년부터 본격적으로 가시화된 확산모델 기반 생성형 AI 모델은 디스코 디퓨전(Disco Diffusion)과 스테이블 디퓨전(Stable Diffusion)이 대표적이다. 디스코 디퓨전은 본질적으로 확산모델을 기반으로 했기 때문에 추상적이고 예술적인 이미지의 표현에 더 적합하다[15]. 파이썬(Python) 프로그래밍 언어로 제어 가능하고, 간단한 텍스트 설명에서 인공지능이 생성한 이미지를 프레임의 낱장으로 만들어 최종적으로는 영상으로 제작할 수 있도록 구성되어 있다. 스테이블 디퓨전은 잠재확산모델 기반으로 좀 더 사실적인 표현과 고해상도의 이미지 생성이 가능하며, 현재 인공지능 예술 창작에서 많이 사용되고 있는 모델이다. 두 모델의 경우는 깃헙(Github)과 구글의 코랩을 통해 사용 가능 하도록 오픈 소스로 공개되어 있다. 또한 둘 다 고정된 텍스트 인코더(CLIP ViT-L/14)를 사용하는 것이 가능하며, 해당 프레임 워크에서 창작물에 대한 텍스트 프롬프트를 최대한 자연어의 내용과 유사하게 생성할 수 있도록 구성되어 있다. 그 외 스테이블 디퓨전의 GUI 기반 시스템인 컴피UI(CompfyUI)나 또 다른 잠재확산 모델 기반 모델인 런웨이 Gen-2(Runway Gen-2) 또한 2024년 현재 활발히 AI기반 영상 제작에 활용되고 있다.

이미지를 생성하는 방식 외에도 인공지능 기술을 바탕으로 한 예술의 창작은 끊임없이 발전하는 중이다. 음악이 그 대표적인 예이다. 구글의 텐서플로우 마젠타(Tensorflow Magenta) 이후 뚜렷한 프레임워크의 등장은 없는 실정이나, 기업에서 인공지능 알고리즘을 바탕으로 서비스 중인 에이바(AIVA), 부미(Boomy)등의 음악 생성기(music generator)는 보편적으로 통용될 만큼 사용성이 증대되었다.

2-4 인공지능 예술의 미학적 해석과 기술철학적 관점

인공지능 예술에 대한 미학적 해석은 아직까지 인공지능의 존재론과 창의성의 문제를 중심으로 논의가 두드러진다. Aris 등[16]은 인공지능 시대의 예술 작품 창작에서 인간의 역할이 여전히 중요함을 강조하며, AI에 창의성의 우월함을 내어줄 것이라는 우려를 일종의 불필요한 유행으로 바라보고 있다. Hong과 Curran[17]은 이러한 논의를 컴퓨터를 사회적 대리인으로 보는 CASA(Computers Are Social Actors)의 관점과 스키마이론(Schema Theory)에 기반해 해석하며, 인공지능 예술 작품과 예술가의 작품에 대한 인식을 경험적 방식으로 조사한 바 있다. Manovich[18]는 일찍이 인공지능미학(AI Aesthetics)이라는 저서를 통해 인공지능 시대 문화적 변화와 도전을 예고한 바 있다. 그러나 이들 연구는 핵심적으로 앞서 설명한 생성형 인공지능을 활용한 예술 작품이 공통적으로 지닐 수 있는 해석적 가능성에까지 구체적으로 도달하지는 못했다. 이에 본 논문에서는 기술철학적 관점을 차용한 미학적 해석을 제안한다.

일반적으로 기술철학은 크게 3종류의 흐름으로 나눌 수 있다. 그 구분은 기술의 존재론, 기술의 인간학, 기술에 대한 가

치론적 논의이다[19]. 인공지능 예술은 최첨단 기술인 인공지능을 사용하는데 있어서 기술적 철학적인 논의를 바탕으로 그 의의의 타당성을 살펴보는 것이 가능하다. 즉 기술에 대한 본의를 역추적하는 방법을 통하여 인공지능이 형성하는 그 의의에 대하여 논하고, 이 논의가 예술로서 자리할 수 있는 의의의 지점 들을 포함하는가를 짚어 보는 것이 도움이 될 것이다. 특히 마누엘 데란다(Manuel Delanda)의 창발(emergence)과 복잡계(complex system) 개념, 벤야민(Benjamin)의 순수언어(pure language) 개념, 그리고 셸드레이크(Sheldrake)와 이재현 등의 공명(resonance) 개념은 기술철학에서 출발했지만 인공지능 예술을 폭넓게 해석하는데 유용해 보인다.

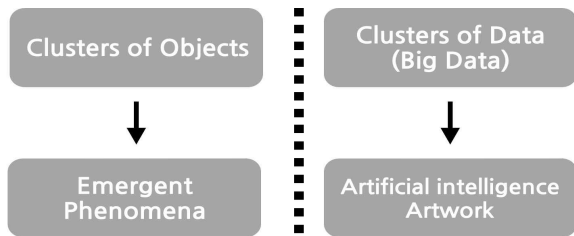


그림 8. 창발적 현상과 인공지능 예술작품 비교
 Fig. 8. Comparison of emergence and AI Art

1) 창발(Emergence) 그리고 복잡계(Complex System)

사전적 의미로 창발(創發, emergence)은 ‘떠오름’이라는 의미가 있으며, 영어로는 ‘갑자기 솟아나는 특성’(emergent property) 또는 창출이라 한다. 데란다는 시스템이 개별 구성 요소로는 직접적으로 환원될 수 없는 특성이나 행동을 획득하는 과정을 창발이라 규정했다[21]. 이는 생물학적 유기체에서 사회 구조에 이르기까지, 복잡한 시스템이 어떻게 예측할 수 없는 새로운 행동과 특성을 나타내는지 이해하는데 중요하게 참조할 만하다.

인공지능 예술은 대규모 데이터에 대한 학습에 기인하지만, 결과적으로 새로운 생성물이 도출된다. 그리고 데이터에서 생성의 결과물로 도달하는 과정은 설명이 가능하지 못한, 즉 환원 불가능한 비가역적 과정이다. 이렇게 일련의 빅데이터는 군집을 이루는 대상체와 같은 선상에서 바라볼 수 있고, 알고리즘의 작용으로 인해 도출된 결과는 창발과 맞닿을 수 있다.

이러한 인공지능의 창발과 함께 살펴봐야 하는 개념은 복잡계(複雜系, complex system, complexity system)이다. 소규모의 네트워크에서 현재는 거대한 데이터를 바탕으로 한 네트워크의 한 형태로 등장한 인공지능 모델은 복잡계와 유사함을 가지고 있으며, 일정 부분은 그 안에 속한 특징이 맞물린다. 복잡계는 통상적으로 물리학, 경제학, 생명현상 등에 대하여 사용되는 용어로서 “완전한 질서나 완전한 무질서를 보이지 않고, 그사이에 존재하는 계로써, 수많은 요소로 구성되어 있으며 그들 사이의 상호작용 때문에 집단 성질이 떠오르는 다체문제”이다[20].

일반적인 기술 시스템적 접근으로 본다면 인공지능 기술은 인간의 지적 능력을 모방한 수학적 원리를 컴퓨팅 시스템에 의해 연산하여 생성, 적용, 구현하는 시스템에 해당한다. 그러나 현재 이 장치는 단순한 연산이 아닌 복잡성(complexity)을 내포하는 단계로 진화하고 있다. 물론 자신의 진화가 아닌 인간의 개입으로 인해 그 관계성이 더욱 단단해지고, 넓은 형태로 변모하고 있기에 인간이 인식하기에는 진화에 가까운 형태로 인지하게 되는 것이다.

마누엘 데란다의 경우는 복잡계와 사회의 상관관계를 질 들뢰즈(Gilles Deleuze)의 연구에서부터 찾았다. 그는 들뢰즈의 배치(Agencement)이론을 설명하며 “전체가 외재성의 관계들(Relations of exteriority)을 통해 이루어진다”라고 밝힌다. 또한 “무엇보다도 이러한 관계들이 함축하는 것은 배치를 구성하는 부분 들이 그 배치로부터 떨어져 나와서 상호작용이 다른 배치에 접속(Plug)될 수도 있다는 것”이라 한다[21]. 또한, 데란다는 인터뷰를 통해 “모든 문화가 그 자신의 세계 안에 있음을 근거로 경험의 언어적 성격을 거부하는 것은 공유된 인간 경험이라는 개념으로 가닿는다. 그리고 여기서 변화란 언어학적 개념의 의미작용(signification)에서의 차이가 아니라 화용론적 개념으로서의 함축성(significance)으로부터 나오는 것이다”라고 말한다[22].

이와 같은 데란다의 관점에서 인공지능을 바라보게 되면 그 존재가 인정될 수 있으나, 객체의 신체화가 이루어지지 않은 일종의 관념적인 대상이라 볼 수 있다. 또한, 이 관념 안에는 애초에 무수히 많은 표상과 기호 등의 의미가 연결되어 있음을 추론할 수 있다. 이는 딥러닝이라 불리는 인공지능경망을 통한 연산 작용을 넘어 프롬프트 엔지니어링의 관점으로 더욱 깊게 상관성을 풀이할 수 있다. 이는 곧 자연어가 기존에 가지고 있는 의미와 이미지가 본래 내포하는 함의가 의미작용으로 일어나는 가운데 예술로서 창출될 수 있는 근거가 된다. 물론 현재의 인공지능경망이 스스로 사고하는 존재는 아니기에 인간의 개입이라는 조건으로 이는 성립된다.

2) 순수언어(Pure Language)

국내 기술철학자인 이재현은 그의 인공지능 기술비평에서 구글의 언어모델을 발터 벤야민의 순수언어와 연결해 통찰력 있게 해석한다. 벤야민은 순수언어를 통해 기존의 번역이 갖던 직접성이 아닌 내재적인 가능성에 무게를 두고 번역 가능성(Translatability)을 주장하였다. 이는 인간과 무관한 특정한 결과물이 내재한 그 형식 자체가 의미가 있음을 뜻하는 바이다[23]. 원작이라는 대상체가 있고 이에 대한 부차적인 형태로서의 번역이 아니라 원작이 내재한 그 가능성 자체가 선행성을 바탕으로 연계되어 있음을 말하는 것이다. 이는 ‘관계 표상물(Relational representations)’이라는 단어로 정리될 수 있으며, ‘가상적 네트워크’ 또한 ‘번역’을 계기로 현실화가 가능함을 의미한다[24].

2017년 구글의 번역이 GNMT(Google Neural Machine Translation)를 기점으로 중간언어 번역이 아닌 문장이 문장

으로 번역되는 대변화가 있었다. 기존의 언어적인 변환은 매개체 역할을 하는 중간언어가 존재하고, 이를 통과하여 전해지는 방식이었다. 그러나 GNMT를 통해 직접적인 전환이 이루어진다는 점이 발견되고, 이는 문장이 내포하는 의미가 서로 맞닿았기에 가능한 변환임을 확인할 수 있다[25].

인공지능 창작의 경우도 데이터를 근간으로 하여 그 내재하여 있는 선형의 결과를 도출해 내는 형태로 볼 수 있다. 수학적 알고리즘의 공식으로 변형하고 되짚고 해체하여도 그 본류의 데이터는 이미 구조화되어 있으며, 이 구조화된 상황에서 맥락이라는 것이 발생하여, 인공지능 창작의 결과가 나타나는 것이다. 이를 더 풀어서 논의하자면 맥락이라는 어문적 접근에서 사용되는 방식이 현재의 인공지능 알고리즘 속에서도 작동됨을 설명할 수 있다. 이는 소쉬르가 기표와 기의로 설명한 기호의 정의처럼, 동일한 의미의 특정한 표상만을 생성해 내는 것이 아니라 동일한 의미의 표상이 연계되어 또 다른 표상을 생성하는 형태가 바로 현재의 인공지능 알고리즘의 구조인 것이다. 이 형태는 기존의 인공지능 알고리즘이 갖는 한계를 넘어서 프롬프트 엔지니어링으로의 발전으로 더욱 확장된 형태로서 지능의 형태에 가까워지고 있다.

3) 공명(Resonance)

인간은 역사를 통해 매개되어있는 지적 생명체로서, 과거의 유산이 현재로 이어지고, 다음으로는 미래로 연결될 것을 알고 있다. 이 관점에서 현재의 인간은 창작에 대한 욕구가 현실점에서 기술과 미디어의 발전을 통해 또 다른 도전을 받고 있다. 위에 기술된 첫 관점으로 바라본다면 인간은 과거에도 매체와 기술의 발전에 의한 위협을 받았고, 미래에도 동일할 것이다. 낯선 것에 대한 두려움이다. 이 두려움을 넘어 이 변화 전체에 대한 행간을 읽게 된다면, 이 두려움에서는 해방될 수 있을 것이다. 쉘드레이크의 형태 공명(Morphic resonance)[26]과 이재현의 공명[27] 개념은 이러한 두려움을 넘어설 행간의 유사성에 대한 관점이다. 인공지능에 대한 위협이 단순히 특정한 기술에 대한 무지에서 비롯된 것이 아닌 역사적이고 반복적인 순환의 일종으로 이해하고 대처할 수 있음을 알 수 있다. 또한, 쉘드레이크는 “집단기억(collective memory)이라는 용어를 통해, 기술적 장치나 관념의 반복적 등장은 형상 장을 공유라는 특정 시기와 장소의 사람들, 특히 기술자들이 형태 공명을 하게 됨에 따라 나타나게 된 것”이라고 설명한다[27].

특히 이 두 개념 이에 인공지능을 과거에서 찾아보자면, 아마도 주술적인 환상에 대한 인간의 믿음과 연관될 수 있다. 인간이 믿고자 한 환상과 미신은 환영(illusion)적인 성격이며, 이 환영으로 인해 운명을 결정짓는 형태를 과거를 통해 우리는 알 수 있다. 인공지능 예술의 환영적인 성격이 이에 해당한다고 본다. 이는 카를 융이 말한 “정신적 본유 구조(innate psychic structures)”와 맞닿는다. 융 또한 임마누엘 칸트의 영향을 받아, 선형적 범주로서의 원형이 존재한다고 보았다[27]. 방대한 데이터는 과거에서 출발하여 인간이

가늠할 수 없는 수준으로 뻗어 나가고 있고, 이를 학습한 인공지능의 알고리즘 또한 계층적 단계 속에서 신경망과 같은 원리로 작동되어 인간이 상상하던 상황을 넘는 결과를 보여주고 있다. 그러나 그 결과들은 인간이 만들어 낸 데이터에서 현재의 결과로써 공명하고 있다.

덧붙이자면 인공지능이 만들어 내는 완료되지 못한 노이즈와 글리치한 영역은 인간이 인지에서 놓친 꿈 혹은 기억과도 같은 상상을 하게 만드는 것이 우연의 일치가 아닐 수도 있다. 앞서 설명한 딥드림의 경우는 신경과학 분야에서 연구 중인 환각 상태의 인간 뇌의 작용과 인공지능이 작동되는 원리의 상관성이 엔트로피(Entropy)의 작용이라 밝힌 바 있다[28]. 즉, 이것이 대상과 객체에 대한 공명일 수도 있는 것이다. 인간이 가진 무의식의 영역이 인공지능 특히 인공신경망이 결과를 도출하는 과정의 연속과 매우 흡사하기 때문에 이와 같은 결과가 생성되지 않는가 추론된다.

위에서 살펴본 바와 같이 인공지능을 바탕으로 한 예술 분야는 아직 명확한 사조로서 정립되었다고 보기는 어렵다. 그러나 인공지능이라는 새로운 기술이 예술에 영향을 주어 하나의 흐름으로 자리 잡아가는 흐름은 부정할 수 없는 상황이다. 또한, 기존의 기술철학적 관점을 바탕으로 살펴볼 때 이 분야의 의미와 의의가 정립될 수 있는 여지는 여러 관점에서 찾아낼 수 있었다. 특히 인공지능 예술에서 현재의 흐름으로 볼 수 있는 다양한 인공신경망이 연결되어 사용될 앞으로의 상황을 바라볼 때, 발터 벤야민의 순수언어와 같이 특정한 대상이 수반하고 있는 본질적인 의미에 대하여 매체와 기술의 관점에서 이해할 수 있는 지점도 찾아낼 수 있었다. 수많은 데이터를 바탕으로 창출되는 결과에 대하여 창발적인 현현을 찾아볼 수 있었고, 기존에 보았던 특정한 형태가 언어적인 맥락을 포함한 유사성과 반복성을 통해 이루어지는 미디어 공명과도 한 맥락을 같이 함을 알 수 있었다. 이상에 기술한 기술철학적 개념과 관점을 적용해 연구자의 인공지능 예술 창작 사례인<까마귀속 연작>의 과정을 설명하고 분석하고자 한다.

IV. <까마귀속 연작(Corvus Series)>작품 연구

<까마귀속 연작>은 2022년 국내에서는 프롬프트 기반 생성형 인공지능 창작 방식이 아직 널리 알려지기 전에 진행된 일련의 작품으로 기존의 문학 작품을 기반으로 제작한 <Crow Eye's View(오감도)>(49'25"), <The Raven(갈까마귀)>(48'), <Kafka's The Castle(카프카의 성)>(35') 세계의 중/장편 분량의 인공지능 기반 영상 작품을 포함한다. 이 연작은 인간의 행동을 모방하는 까마귀에서 모티브를 얻고, 순환적인 공명적 접근을 바탕으로 인공지능이 인간의 영역으로 다가오는 상황을 프롬프트 엔지니어링을 바탕으로 문학 작품들을 해석하는 시도로 출발하였다. 이 시도의 접점은 인간과 유사한 형태의 ‘학습’ 능력을 통해 오랜 기간 성체가

되어가는 까마귀의 성장 과정이 마치 인공지능의 발전 과정과 흡사하다는 점에서 착안하게 되었다. 또한, 인지적 측면에서 뛰어난 까마귀의 행동 양태가 인간적인 발전과는 또 다른 양태를 띄기에 인공지능과의 접점을 생각해 보게 되었다.

까마귀속이란, 유전적 혹은 계통적으로 밀접한 관계를 갖는 근연종들의 그룹을 정리할 때 쓰이는 용어인 속(屬, Genus)을 기준으로 우리가 흔히 알고 있는 까마귀와 그 유사 개체 45종이 모인 분류 기준이다. 이 분류를 하나의 맥락적 기준으로 삼고서, 동시대의 인공지능이 여러 신경망 알고리즘과 상관관계에 있는 링크를 통해 이미지를 생성하듯이, 까마귀 속과 상관되지만, 본래는 전혀 다른 문학 작품들 사이에서 데이터를 수집하고 이를 인공지능을 통해 다시 변환하여 각각의 작업으로 구현한 형태이다. 이에 대한 설명을 위해 인공지능 생성 영상 - 시적 드로잉 시리즈라는 부제와 시리즈의 표기를 덧붙인다.

이 까마귀속을 기준으로 데이터로 수집한 문학 작품을 살펴보면 다음과 같다.

먼저 일제강점기의 문학가인 이상의 '오감도'는 한글로 변환하면 '까마귀 눈으로 보기(Crow's eye view)'에 해당한다. 조감도라는 단어로 알려진 새의 시선으로 바라보는 것을 넘어 까마귀라는 존재 특유의 영험함에 빚든 시의 연작 타이틀로 보인다. 다음으로 미국의 대문호인 에드거 앨런 포의 '갈까마귀(The Raven)'이다. 북미 지역의 대표적인 새인 갈까마귀는 세부적으로 말하자면 큰까마귀로 분류될 수 있으나, 한국에서는 갈까마귀라는 번역명이 더 많이 쓰인다. 마지막으로 체코의 대문호 프란츠 카프카의 '성'이다. 카프카의 경우는 체코어로 작가의 성씨인 '카프카'라는 단어가 체코어 단어 'Kavka'의 이 표기로 추정되며, 이는 '검은 까마귀'를 뜻한다.

위의 세 작품과 연관이 있는 까마귀속이라는 키워드를 통해 작품들의 기본적인 컨셉을 정립하고, 각 문학 작품이 풍기는 특유의 분위기 등을 실제 작품 제작에 반영하여 제 2 저작물의 형태를 기본적인 골자로 제작하였다. 특히 위의 세 작품의 경우는 본래의 내용들이 명확하고 구체적인 의미를 담고 있지 않은 점에서 프롬프트 엔지니어링에 사용되는 인공지능 모델들이 표현할 수 있는 범위의 한계가 제한되지 않을 것이라는 예측을 바탕으로 선정하였다. 이상(아방가르드), 에드거 앨런 포(낭만주의), 카프카(실존주의)는 각 장르를 대표하는 거장들의 작품인 점도 제 2 저작물의 관점에서 훌륭한 소재이자 데이터로 작용할 것으로 파악하였다. 이 연작은 모두 동일한 프로세스를 근간으로 하되, 각기 요소별 변경 사항을 적용하여 각각의 작품별로도 독자적인 전시가 가능한 형태로 프롬프트 엔지니어링을 통해 제작하였다. 디스코 디퓨전(Disco Diffusion)과 스테이블 디퓨전(Stable Diffusion)은 각각 추상적이고 예술적인 이미지의 표현에 더 적합하며 사실적인 표현과 고해상도의 이미지 생성이 가능해[15] 각 문학 작품의 특성과 분위기를 시각적으로 표현하는 데 적합한 것으로 판단하여 사용하였다. 또한 음악 생성을 위해 사용된 부미(Boomy)와 에이바(AIVA)는 인공지능 기술을 바탕으로 특정

장르나 스타일의 음악을 생성하는 데 특화되어 있어 선택하여 사용하였다.

이 작품들은 2022년 11월 21일부터 12월 23일까지 한국수출입은행의 금고미술관에서 조영각 개인전 <The unexpected syntax : Request and Response/ 예기치 않은 구문 : 요청 및 응답>을 통해 전시되었다.

4-1 작품 <Crow Eye's View(오감도)> 제작 연구

1) 작품 <Crow Eye's View(오감도)> 기획 의도

이 작업은 언어로 구성된 문학 작품을 다양한 신경망의 작용을 거치는 프로세스를 통해 구성하는 첫 번째 시도이다.

각 시의 문장 속 텍스트를 근간으로 데이터를 구성하여 영상으로 생성해 낸다. 음악의 경우도 동일 데이터를 파라미터로 적용하여, 로파이(Lo-fi)의 사운드와 국악의 십이율(十二律) 구성을 반영한 인공지능 알고리즘 기반의 생성기로 제작한다. 이는 과거의 문학 작품을 인공지능이 재해석하여 현대의 작업으로 전환되는 일종의 제 2 저작물이다. 이에 시가 쓰인 당시의 배경인 조선말, 일제강점기에 대한 화풍의 해석을 추가하여 과거의 문학 작품이 영상으로 제작되었다는 가정에서 애니메이션으로 구현되는 과정을 담았다. 더불어, 원문인 오감도는 국내 문학 중 난해한 시로 평가되는 만큼 다양한 관점의 텍스트로 구성되어 있다. 이는 인공지능이 텍스트에 기반을 둔 이미지를 만들어 내고, 나머지 유추하지 못한 이미지

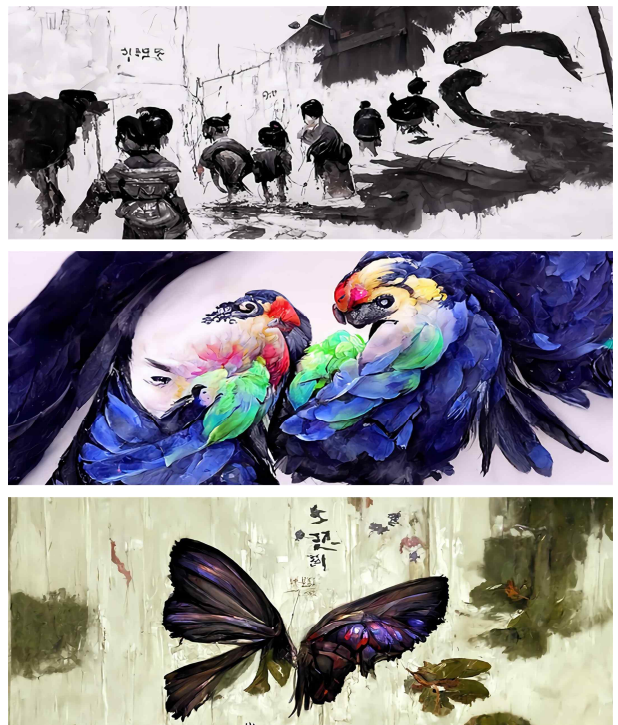


그림 9. 작품 <오감도>의 생성이미지 - 1연(상), 6연(중), 10연(하)
Fig. 9. AI generated images of Crow Eye's View - stanza 1(above), stanza 6(middle), stanza 10(below)

에 대하여 데이터를 무작위적 링크로 생성하는 방식이 효과적인 표현을 구축하는지 확인하는 기회이다.

2) 작품 <Crow Eye’s View(오감도)> 구현

오감도 총 15편의 시를 표준 한국어로 1차 변환하고, 이를 다시 영문으로 재번역하여 텍스트를 데이터베이스로 구축하였다(표 1). 이 변환된 텍스트는 구글 코랩을 환경에서 각각의 행별로 프롬프트로 적용되었다. 각 프롬프트는 디스코 디퓨전(Disco Diffusion) 4.1 버전의 프레임워크를 사용하여, 60프레임의 이미지로 생성되도록 구성한다(그림 10, 그림 11).

표 1. 프롬프트 생성을 위한 <오감도> 의 원문, 한국어, 영어 번역 예시

Table 1. Example texts of <Crow Eye's View> to generate prompt original, Korean and English version

| Lee Sang's Crow Eye's View | | Text |
|-----------------------------|---------------------|--|
| Stanza 1(1st and 2nd lines) | Original | 十三人の兒孩が道路로疾走하오. (길은막달은골목이適當하오.) |
| | Contemporary Korean | 13인의아해가도로로질주하오. (길은막다른골목이적당하오.) |
| | English translation | Thirteen children are speeding down the road. (A dead end is appropriate for the road.) |

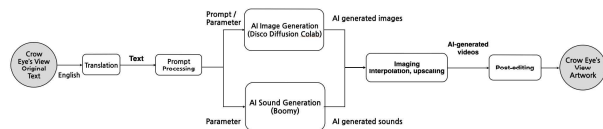


그림 10. 작품 <오감도> 제작 순서도
Fig. 10. Production flowchart of <Crow Eye's View>



*This study utilizes Korean text data to derive results.
그림 11. 작품 <오감도> 제작 과정
Fig. 11. Production process of <Crow Eye's View>

이때 사용된 인공지능 모델은 디퓨전 모델(Diffusion: 256×256 diffusion uncond)과 클립(CLIP: ViTB32, ViTB14, RN50×4)이며, 영상의 좌표계를 설정하여 공간감과 시적 상상을 영상화 하는데 초점을 맞추었다. 사운드의 경우는 프롬프트에서 사용된 어절들에 맞추어, 음악의 구성요소인 템포, 멜로디, 코드, 베이스, 드럼에 해당 값으로 입력하여 부미(Boomy)를 통해 생성하였다.

3) 작품 <Crow Eye’s View(오감도)> 전시 및 작품 분석

총 15편의 시에 대하여 제작된 영상은 전체 49분 25초 분량이며, 기존 시의 난해함 만큼 복잡다단한 영상의 형태로 구현되었다. 특히 디스코 디퓨전 4.1 버전의 모델을 사용할 경우, 완전한 문장으로 결과를 도출하려 하였음에도 아무런 형상이 출력되지 않는 노이즈에 머문 이미지도 약 절반 이상이 확인되었다. 이에 터치 디자이너(Touch Designer) 프로그램을 사용하여 잔상효과(feedback)를 적용해 연속되는 이미지가 끊어지지 않고 영상에서 좀 더 자연스럽게 이어질 수 있도록 하였다(그림 12).

또한 영상의 분위기 경우는 기존의 시가 내포하는 분위기가 더욱 극적으로 연출될 수 있도록, 원작이 작성된 시기의 화풍 즉 프롬프트로는 “Oriental ink painting, 20th century”(동양의 수묵화, 20세기)을 지정하였다. 이에 따라 15편 중 동양화의 특징에 편향된 스타일로 생성된 결과들이 도출되었다.

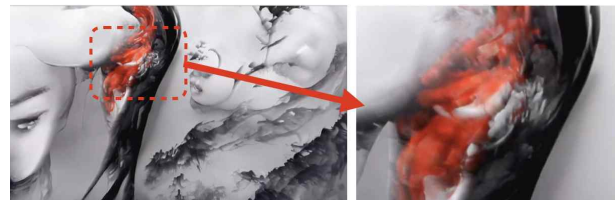


그림 12. 작품 <오감도>에서 자연스러운 전환을 위한 잔상효과(feedback)로 생성된 이미지(좌)와 세부 모습(우)
Fig. 12. The image(left) generated by applying Touch Designer's feedback effects for smooth transitions in the work <Crow Eye's View> and its detail(right)

인공지능이 생성한 이미지를 살펴본다면 특정 단어에 대한 구체적인 이미지 생성이 뚜렷하지만, 전혀 반영되지 못한 단어도 존재하였다. 자세한 설명과 부사 등의 표현적인 측면은 이미지에 표현되지 않은 부분이 상당하였다. 전체 작품의 분위기를 위해 사용되는 프롬프트의 영역은 큰 변화 없이 대다수 적용되었다.

본 작품은 시인 이상의 난해한 시 연작인 <오감도>를 인공지능을 통해 시각화한 작업이다. 앞서 기술철학적 개념을 적용해 해석해 보자면, 먼저 복잡계 이론의 관점에서 볼 때, 원문의 난해함이 인공지능의 처리 과정을 거쳐 예측 불가능한

시각적 결과물로 나타나는 현상은 창발의 개념을 잘 보여준다. 앞서 언급한 노이즈에 머문 이미지도 약 절반 이상이 확인되었다는 점은 이러한 창발적 특성을 단적으로 드러낸다. 또한, 한국어 원문을 영어로 번역한 후 다시 이미지로 변환하는 과정은 발터 벤야민의 순수언어 개념을 실현한다. 이는 원작의 본질적 의미가 다른 형태의 언어로 전이되는 과정을 보여주며, 기술을 매개로 한 새로운 형태의 번역 가능성을 제시한다. 더불어, 일체감점기의 문학 작품을 현대 기술로 재해석하는 이 작업은 공명 개념을 구현한다. 또한 원작이 작성된 시기의 화풍(동양의 수묵화, 20세기)을 지정한 점은 과거와 현재의 기술, 예술적 요소들이 공명하며 새로운 의미를 창출하는 과정을 보여준다.

4-2 작품 <The Raven(갈까마귀)> 제작 연구

1) 작품 <The Raven(갈까마귀)> 기획 의도

이 작업은 언어로 구성된 문학 작품을 다양한 신경망의 작용을 거치는 프로세스를 통해 구성하는 두 번째 시도이다.

갈까마귀는 에드거 앨런 포가 1845년 2월에 발표한 미국의 대표적인 시 중 하나이다. 시 속의 화자는 사랑하던 ‘레노어(Lenore)’라는 이름을 가진 여인을 잃고 슬픔에 잠겨 있을 때, 화자의 집으로 날아 들어온 까마귀에게 그 감정을 투사한다. 화자는 까마귀(Raven)에게 ‘두 번 다시는(nevermore)’이라는 이름을 지어주고, 까마귀에게 끊임없이 질문을 던진



그림 13. 작품<갈까마귀> 생성이미지- 2연(상), 6연(중), 15연(하)

Fig. 13. AI generated images of The Raven - stanza 2(above), stanza 6(middle), stanza 15(below)

다. 하지만 까마귀는 자신의 이름과도 같은 ‘두 번 다시는’이라는 대답만 반복할 뿐이다. 이 내용의 시 전문을 행 단위로 나누어 문장별로 프레임 생성하는 형태로 영상 작업은 제작되었다.

2) 작품 <The Raven(갈까마귀)> 구현

미국의 문학가 에드거 앨런 포의 갈까마귀 텍스트를 데이터베이스로 구축하였다. 텍스트는 각 연으로 구별하여 프롬프트로 적용되며, 각 프롬프트는 디스코 디퓨전(Disco Diffusion) 5.4 버전의 프레임워크를 사용하여, 60프레임으로 구성한다. 특히 이 경우는 각 연을 하나의 영상으로 분리하고, 최종적으로는 하나의 시가 다수의 영상으로 출력되도록 하였다.

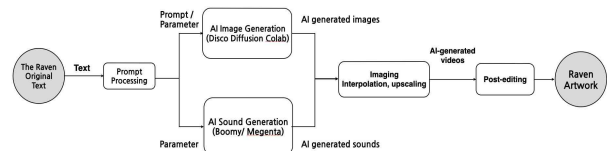


그림 14. 갈까마귀 제작 순서도

Fig. 14. Production flowchart of <The Raven>

Ah, distinctly I remember it was in the bleak December; And each separate dying ember wrought its ghost upon the floor.

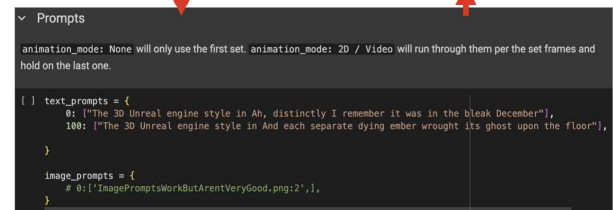


그림 15. 갈까마귀 제작 과정

Fig. 15. Production process of <The Raven>

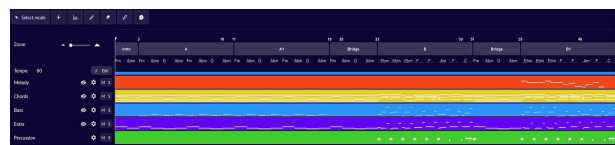


그림 16. 갈까마귀 사운드 미디 생성 AIVA 결과 이미지

Fig 16. Sound MIDI Generation AIVA result of <The Raven>

표 2. 프롬프트 생성을 위한 작품 <갈까마귀>의 원문 예시 - 2연(상), 6연(중), 15연(하)

Table 2. Example texts of <The Raven> to generate prompt - stanza 2(above), stanza 6(middle), stanza 15(below)

| Poe's The Raven | Original Text |
|-----------------|---|
| stanza 2 | Ah, distinctly I remember it was in the bleak December; And each separate dying ember wrought its ghost upon the floor. Eagerly I wished the morrow;—vainly I had sought to borrow From my books surcease of sorrow—sorrow for the lost Lenore—For the rare and radiant maiden whom the angels name Lenore—Nameless herefor evermore. |
| stanza 6 | Back into the chamber turning, all my soul within me burning, Soon again I heard a tapping somewhat louder than before. "Surely," said I, "surely that is something at my window lattice; Let me see, then, what thateat is, and this mystery explore—Let my heart be still a moment and this mystery explore; —'Tis the wind and nothing more!" |
| stanza 15 | "Prophet!" said I, "thing of evil!—prophet still, if bird or devil!—Whether Tempter sent, or whether tempest tossed thee here ashore, Desolate yet all undaunted, on this desert land enchanted—On this home by Horror haunted—tell me truly, I implore—Is there—is there balm in Gilead?—tell me—tell me, I implore!" Quoth the Raven "Nevermore." |

이때 사용된 인공지능 모델은 디퓨전 모델(Diffusion: 512×512 diffusion uncond finetune 008100)과 클립(CLIP: ViTB32, ViTB16, RN50×16)이며, 3D 언리얼 엔진을 사용한 효과와 같은 공간적 설정으로 영상화 하는데 초점을 맞추었다. 이는 프롬프트에 "Like a 3D Unreal 5 engine"의 문장을 통해 구현을 유도했는데, 이 방식을 통해 실제 3D 엔진을 사용하지 않고도, AI 모델이 언리얼 엔진 5의 특징적인 조명, 텍스처, 깊이감 등을 모방한 이미지를 생성하도록 했다. 이는 2D 이미지 안에서 3D적인 공간감을 구현하는 효과적인 방법이었다.

사운드의 경우는 구글의 텐서플로우 마젠타(Magenta)의 톤 변환기(Tone Transfer), 신디사이징(NSynth Super)를 주로 사용하고, 명상 음악 부류의 멜로디를 부미(Boomy)를 통해 생성하였다. 예를 들어, 프롬프트로 적용되는 1연의 텍스트의 감정을 분석(Naver sentiment movie corpus v1.0)하여 "Nevermore"의 반복을 -0.8/1.0 척도의 우울함으로 수치화하였다. 이를 바탕으로 80 BPM의 템포와 A 단조 음계를 선택하여 음악적 기반을 마련하고, 사운드 생성 과정에서는 구글의 텐서플로우 마젠타 기술을 활용하였다. 톤 트랜스퍼를 통해 오픈소스 사운드인 까마귀 울음소리를 피아노 음색으로 변환하여 주 멜로디로 사용하였고, NSynth Super를 이용해 50-200Hz 주파수 범위의 저음 패드를 생성하였다. 또한, Boomy를 활용하여 80 BPM, A 단조 기반의 멜로디를 추가로 생성하였다. 이러한 과정을 통해 최종적으로 멜로디와 패드를 혼합한 트랙으로 구성된 에이바(AIVA) MIDI 파일을 생성 및 사용하였다. 생성된 사운드는 '갈까마귀' 영상의 배경

음악으로 활용되었으며, 특히 "Nevermore"가 등장하는 장면에서는 반복적인 패턴을 통해 시의 핵심 메시지를 강조한다.

3) 작품 <The Raven(갈까마귀)> 전시 및 작품분석

총 18연의 시로 제작된 영상은 전체 48분 분량이며, 기존의 시에서 느껴지는 분위기와 유사하게 어두운 배경의 구체적인 상황을 보여주는 3D 애니메이션 영상 형태로 최종 구현되었다. 각 영상은 연에 해당하는 텍스트를 바탕으로 하나의 챕터를 구현할 수 있도록 개별적으로 생성 및 영상화되었다. 이 경우는 이미지의 구현에 있어 더욱 큰 인공지능 모델을 사용하여 구체적이고 뚜렷한 형상이 대부분을 차지한다. 프롬프트 내의 구도 변환을 위한 시점이 다양하게 적용되었음에도, 영상은 각 연의 내용이 분명히 파악되도록 구성되었다.

<갈까마귀> 영상은 에드거 앨런 포의 시를 인공지능을 통해 재해석한 작업으로, 기술철학적 측면에서 다양한 의미를 내포하고 있다. 우선, 프롬프트 엔지니어링의 철학적 의미가 두드러진다. 각 연을 개별적인 프롬프트로 사용하여 영상을 생성하는 방식은 마누엘 데란다가 언급한 "화용론적 개념으로서의 함축성"을 실제로 구현한 사례다. 이는 인간의 언어적 지시와 인공지능의 해석 사이의 복잡한 상호작용을 보여준다. 기술의 존재론적 측면에서, '3D 언리얼 엔진'이라는 프롬프트를 활용한 공간적 설정은 기술이 단순한 도구를 넘어 새로운 예술적 공간을 창조하는 존재로 기능함을 보여준다. 이는 기술이 예술 창작의 본질적인 부분으로 통합되는 과정을 드러낸다. 또한, 19세기 미국 문학 작품을 현대적 영상 언어로 '번역'하는 과정은 벤야민의 순수언어 개념을 구현한다. 각 연의 내용이 분명히 파악되도록 구성된 점은 원작의 본질이 새로운 매체를 통해 전달되고 있음을 보여주며, 이는 기술을 통한 예술적 번역의 새로운 가능성을 제시한다.

4-3 작품 <Kafka's The Castle(카프카의 성)> 제작 연구

1) 작품 <Kafka's The Castle(카프카의 성)> 기획 의도

이 작업은 언어로 구성된 문학 작품을 다양한 신경망의 적용을 거치는 프로세스를 통해 구성하는 세 번째 시도이다. 미완의 원작 소설의 방대한 내용을 줄거리 형태로 축약하고, 이 텍스트를 바탕으로 인공지능 언어모델이 미완의 부분을 추가로 작성했다. 기존 줄거리와 인공지능에 의해 추가로 작성된 문장들은 프롬프트 엔지니어링과 구글 마젠타 등의 딥러닝 프레임워크를 사용하여 영상으로 제작했다. 이에 원작 소설의 배경이 되는 시점을 22세기의 가상의 배경으로 추가하여 과거의 문학 작품이 미래의 시점에서 작성되었을 때, 만들어질 수 있는 애니메이션으로 가정되는 과정을 담았다.

2) 작품 <Kafka's The Castle(카프카의 성)> 구현

소설 성의 원문을 바탕으로 줄거리를 각색하여 인공지능 언어모델의 데이터로 활용될 수 있도록 1차 가공하였다. 가공된 텍스트는 OPEN AI의 언어모델인 GPT-3를 통해 미완의 텍스트를 추가했다.



그림 17. 카프카의 성 생성이미지 위-챗터1, 중-챗터3, 하-챗터5
 Fig. 17. AI generated images of <Kafka's The Castle> - chapter.1(above), chapter.3(middle), chapter.5(below)

이는 본 작품 연구가 수행된 2022년 당시 GPT-3는 가장 선진적인 공개 언어모델이었고, API를 통한 접근성이 용이하여 예술 창작 과정에서의 실험적 사용에 적합했기 때문이다. 당시 GPT-3는 문학적 텍스트 생성에 있어 우수한 성능을 보여, 카프카의 미완성 작품을 보완하는 본 연구의 목적에 부합했다. 현재 더 발전된 버전의 GPT 모델이 존재하지만, 본 연구의 역사적 맥락과 당시의 기술적 환경을 고려할 때 GPT-3를 사용한 본 작품의 기록과 분석은 급변하는 AI 기술의 발전 속에서 이를 이용한 예술적 실천의 과정을 기록한다는 점에서 의미 있다 할 수 있다.

우선 원문의 줄거리에 해당하는 부분을 위의 인공지능 언어 모델에 입력하고, 그 줄거리를 바탕으로 결과물이 생성되도록 구성했다. 각 프롬프트는 디스코 디퓨전 5.6 버전의 프레임워크를 사용하여, 60 프레임의 이미지로 생성되도록 구성했다. 이때 사용된 인공지능 모델은 디퓨전 모델(Diffusion: 512×512 diffusion uncond finetune 008100)과 클립 (CLIP: ViTB32, RN101, RN50×64)이며, 영상이 네러티브를 포함할 수 있도록 초점을 맞추었다. 특히 22세기라는 가상 시점을 반영하기 위해 “Like the scene in the SCI-FI movie, cyber punk(공상과학 영화, 사이버펑크에 등장하는 장면처럼)”와 같은 프롬프트도 함께 사용했다.

주요 파라미터 설정으로는 템포를 125-140 BPM 사이로 조절하여 긴장감을 나타냈고, 화성 진행에 20%의 불협화음을 포함시켜 불안정한 분위기를 조성했다. 리듬 패턴은 4/4박자를 기본으로 하되, 30% 빈도로 불규칙한 악센트를 추가했

다. 에이바(AIVA)를 활용하여 이러한 파라미터를 바탕으로 앰비언트 테크노 스타일의 음악을 생성했다. 생성된 사운드는 소설의 주요 장면에 따라 템포와 음색을 조절하여 적용했다. 예를 들어, K가 마을에 도착하는 장면에서는 낮은 템포와 미스터리한 패드 사운드를, 성을 향해 가는 장면에서는 템포를 높이고 전자음 비율을 증가시켰다. 이러한 과정을 통해 생성된 MIDI 파일은 소설의 내러티브 구조를 따라 변화하는 사운드스케이프를 형성하며, 카프카의 '성'이 가진 복잡하고 미스터리한 분위기를 AI 기술을 활용하여 청각적으로 재해석했다.

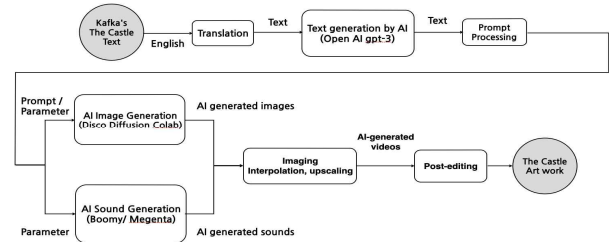


그림 18. 카프카의 성 제작 순서도
 Fig. 18. Production flowchart of <Kafka's The Castle>

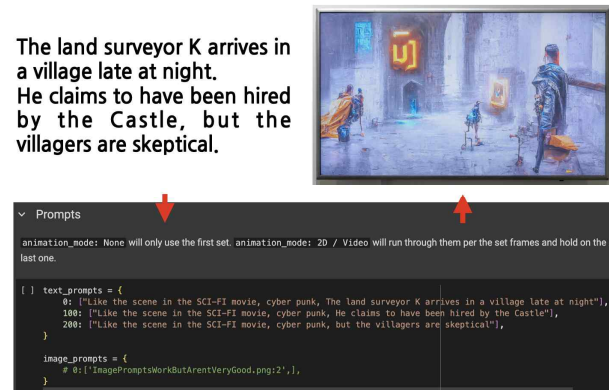


그림 19. 카프카의 성 제작 과정
 Fig. 19. Production process of <Kafka's The Castle>

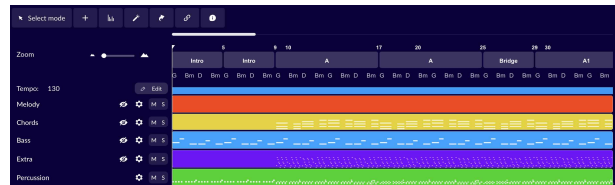


그림 20. 카프카의 성 사운드 미디 생성 AIVA 결과 이미지
 Fig. 20. Sound MIDI Generation AIVA result of <Kafka's The Castle>

표 3. 프롬프트를 위한 카프카의 성의 한국어 요약, 영어 번역 예시 - 1연

Table 3. Example of Korean summary and english translations of Kafka's *The Castle* - stanza 1 for prompt

| Kafka's <The Castle> | | Text |
|----------------------|------------------|---|
| Chapter1 | Korean (summary) | 측량사 K는 늦은 밤 한 마을에 도착합니다. 그는 성에서 그를 고용했다고 주장하지만, 마을 사람들은 의심스러워합니다. K는 여관에서 묵게 되고, 그곳에서 성의 관리인 중 한 명인 슈바르처를 만납니다. K는 자신의 고용에 대해 확인하려 하지만, 상황은 점점 더 복잡해집니다. |
| | English prompt | The land surveyor K arrives in a village late at night. He claims to have been hired by the Castle, but the villagers are skeptical. K stays at an inn, where he meets Schwarzer, one of the Castle's officials. K tries to confirm his employment, but the situation becomes increasingly complex. |

*This study utilizes Korean text data to derive results.

3) 작품 <Kafka's The Castle(카프카의 성)> 전시 및 작품 분석

제작된 영상은 전체 35분 분량이며, 기존의 소설이 갖는 분위기에서 추가된 텍스트로 구성된 내용은 SCI-FI 타입의 영상으로 최종 구현되었다. GPT-3로 구현된 텍스트 일부는 다음과 같다.

“K는 계속 학교 행정관으로 일했고 계속해서 열악한 대우를 받았다. 하지만 프리다와 함께 할 수 있어서 만족했다. 어느 날 K는 건강이 나빠지기 시작했고 침대에 누워 있었다. 프리다는 그를 간호했지만, 그의 상태를 돕기 위해 그녀는 아무 것도 할 수 없었다. 그런 다음 K는 성 관리들에게 마을에 머물 수 있는지 물었지만, 그의 요청은 거부되었다. 그러나 그들은 상황을 감안할 때 그가 머물고 일하는 것을 허용했다. K의 건강은 계속 악화하였고 결국 영양실조로 사망했다.”

위의 텍스트는 원작의 내용과 맥락을 같이 하며, 일단락되는 마지막 챕터 부분을 작성하여 새로운 결말로의 가능성을 타진할 수 있었다. 또한, 음악과 영상의 속도감을 다르게 구분하여 개별 이미지가 프레임별로 인지될 수 있는 장치적 구성을 진행하였다. 영상은 30프레임을 기준으로 만들어진 프롬프트 결과물을 120프레임까지 보간한 다음, 이를 다시 30프레임으로 전환하여 4배로 영상의 분량을 늘였고, 음악은 125-140 bpm 사이의 앰비언트 테크노 음악을 제작하여 영상의 흐름과 상이한 빠른 비트의 음악을 배치하였다. 이에 따라 일반적인 영상 관람의 형태인 전체 영상을 모두 관람하는 형태가 아닌, 일부의 영상을 관람하여도 해당 원문의 분위기가 영상 및 음악을 통해 느낄 수 있도록 표현하였다.

인공지능이 생성한 이미지는 세세한 부분의 내용이 이미지에 적용되었으나, 큰 흐름의 중요 텍스트만 반영된 형태에 가까워 한계가 지적된다. 또한, 이 경우에 프레임과 프레임 사이의 급격한 변환을 줄이고자 보간(interpolation) 프로그램과

업스케일링(up-scaling) 프로그램을 적용하여 전시에 적합한 영상으로 보정해야 했다.

<카프카의 성> 영상은 프란츠 카프카의 미완성 소설을 인공지능을 통해 완성하고 시각화한 작업으로, GPT-3를 통한 미완성 텍스트의 생성과 이를 바탕으로 한 이미지 생성 과정은 창발적 현상의 면모를 띠고 있다. 또한 20세기 초 유럽 문학 작품을 미래적 배경으로 재해석하고 이를 시각적으로 구성함으로써 과거의 서사와 미래적 이미지가 공명하도록 했다. 앰비언트 테크노 음악의 사용은 이러한 시간적, 공간적 공명을 청각적으로도 구현하고자 한 것으로 볼 수 있다. <카프카의 성>은 인공지능 기술이 이미지와 텍스트, 사운드 등 멀티모달한 형태의 창작 과정에 적극적으로 개입된 사례라 할 수 있다.

V. 결 론

이상 까마귀속 연작(Corvus Series)은 프롬프트 엔지니어링을 통해 과거의 유산이자 문학적 가치가 있는 어문에 대한 인공지능의 생성 이미지 결과를 바탕으로 제작되었다. 각 문장에서 내포하는 본작의 함의를 찾아내는 의미론적 해석보다는 각 문장과 단어에 대하여 직접적이고 구체적인 이미지를 만들어 내고, 비정형화된 이미지를 생성하는 형태로 결과가 도출되었다.

기술철학에 기반한 해석적 관점에서 보자면, 작품에서 프롬프트 엔지니어링은 세상에 존재하지 않는 것으로 보이는 부분이 함축적인 의미가 있고 창발적인 결과를 도출하는 것으로 판단된다. 그리고 순수언어의 영역처럼 방대한 데이터 중 미지의 영역이 노이즈로 표현되고 그치는 것이 아닌 어문적인 비정형적 링크들의 상관관계에서 또 다른 맥락이 생성된다는 점이 특히 사항이다. 또한, 기존 문학 작품의 데이터를 근간으로 하여 알고리즘이 작동하기에 과거의 유산과 함께 재현적인 의미로 공명하는 부분들이 존재하는 것을 확인할 수 있었다.

또한 본 연구는 회화를 기준으로 한 시각 예술이 신체적 발현으로 행위를 통해 결과를 만드는 것이 아닌 사고의 귀결인 언어를 통해 생성된 결과를 선택하는 것이 앞으로 다가올 시대의 새로운 예술 창작 방식 일 수 있다는 가능성을 탐색한 과정이었다. 그리고 앞으로 가장 인간적인 예술의 방식이 무엇인가에 대한 인공지능 바탕의 초기의 모의실험으로서 이 작품들이 의의를 갖는다. 또한, 생성된 결과물을 선택하는 과정이 큐레이션의 일종으로 작용하리라는 예상과 함께, 인간의 이성과 감성이 동시에 작용한 결과로서 인공지능이 생성한 결과가 의의를 지닐 것으로 예상된다. 덧붙여, 창작자의 의도와 함께 특정한 지능적인 작용을 바탕으로 협력체로 기능한다는 점에서 도구를 넘어선 대상으로 수행의 주체적인 역할을 일정 부분 담당한다는 점이 인상적인 발전 가능성으로 분

석되었다.

이 시리즈의 한계는 작가 본인의 독창성보다는 원작에 대한 해석이자, 재현적인 의미가 구체화 될 수 있도록 인공지능이 사용되어, 2차 저작물을 1차 목표로 한다는 점에서 일반적인 예술 작품과는 궤를 달리하는 것이다. 이를 발판 삼아 프롬프트의 재료가 될 수 있는 작품의 세계관을 담을 수 있는 근거를 마련하려는 고민과 함께, 동시대의 담론에 대한 깊이 있는 논의를 통해 작품의 제작 방향을 설정하는 것이 더욱 나은 결과로 이어질 것이라 예상된다. 또한 본 작품 연구는 2022년 생성 AI 분야에서 급격한 시각변동이 일어나던 시기의 초반에 진행된 작품 시리즈로서, 일부 사용 기술은 당시의 기술적 환경을 반영하고 있다. 그러나 이러한 한계에도 불구하고 인공지능 예술의 역사적 맥락과 작가주의적 상상력이 발현된 인공지능 예술 작품 사례로서 의미가 있다고 본다.

예술에서 창작 영역의 확장은 기존의 태도와 상용 되지 못하여, 인정받기 어려운 경우가 많다. 낯선 재료와 도구를 사용하여 예술로의 새로운 도전은 항상 경시되거나 대우받지 못하는 경우가 많지만, 중국에는 그 시도와 의미가 예술사적으로 반드시 진화를 위한 하나의 단계로 일컬어지는 경우가 후대에서 정리되곤 한다. 현재의 인공지능 예술 또한 앞서 설명한 바와 같은 진행 과정을 겪고 있다고 판단된다. 기술적인 발전은 끊임없이 팽창하고 있으며, 이를 미디어아트에 국한되는 것이 아닌 예술 전 분야에 다양한 접근과 시도를 통해 보편적 사용성을 갖게 되리라 생각한다.

기술적인 발전은 결국 현재의 인공지능 기술이 나아가는 멀티모달의 형태, 즉 다양한 알고리즘의 상관관계가 더욱 접점을 갖게 될 것이며, 더욱 많은 데이터가 이 알고리즘의 영향으로 인간의 필요 이상의 결과를 도출하게 될 것으로 판단된다. 이 현상에 대해서는 단순히 사용자의 입장이 아닌 인공지능이라는 타적 존재에 대한 수용과 성찰을 인간 스스로 먼저 행하여야 할 것이다. 이에 연구된 방향을 통해 인공지능 예술이 나아가야 할 영역이 어떤 유형과 결과로 나타나고 있으며, 또한 멀티 모달과 같은 다층적이고 관계성이 짙은 개념으로 발전되리라는 것을 짐작할 수 있었다.

특히 인공지능 예술 창작 분야는 기존의 데이터 마이닝에서 출발하던 연구 방법을 넘어, 방대한 빅데이터와 거대한 컴퓨팅 파워를 근간으로 하는 거대 인공지능 모델의 등장에 따라 다양한 분야에서 분화될 방법으로 사용될 것으로 보인다. 또한, 거대 인공지능 모델들의 연계, 즉 멀티모달이 더욱 강화되는 상황에서 인공지능을 이용한 창작은 더욱 서비스 차원과 유사해질 수도 있다. 하지만 창작자들이 이 모델들을 어떻게 사용하며, 그 이면에 있는 새로운 사용 방법과 담론에 대한 치열한 고민을 통해 새로운 인공지능 예술 창작의 장이 또 열리지 않을까 기대한다. 거대 자본을 바탕으로 등장한 인공지능 기술의 끝없는 발전이야말로, 창작자가 쉽 없이 탐구하고 연구해야 할 경쟁 상대가 될 수도 있다. 여기에 더하여 본문에서 살펴본 바와 같이 단순하게 인간이 인공지능을 인격으로 대하지 않고, 데이터와 알고리즘에 의해 생겨나는 매체

적인 기능을 포함하는 기술적 대상 또는 타자적 비인간의 존재로서 인정이 필요하다고 생각된다. 이를 본 저자는 ‘기계적’ 또는 ‘지능적 수행체’로 칭하며, 인격과 분리된 대상으로 바라보는 용어에 대하여 앞으로 연구를 지속할 예정이다. 이는 인공지능의 복합적이고 특수한 특성을 고려한 것이다. 인공지능은 단순한 기계나 인간의 모방을 넘어, 창발적 특성과 학습 능력을 보유하고 있으며, 인간과 복잡한 상호작용을 할 수 있다. 또한 독자적 창작 능력이 내포되어 있어, 예술 창작에 있어 새로운 차원을 열어주는 역할을 수행한다. 이러한 특성들은 인공지능에 대한 윤리적, 법적 고려사항을 필요로 하며, 지능과 의식에 대한 기존의 철학적 개념에 도전한다. 따라서 인공지능을 예술 창작의 보완적 파트너이자 '지능적 수행체'로 인식하는 것은, 그 능력과 한계를 정확히 파악하고 인간 예술가와의 시너지를 최대화할 수 있는 타당한 접근 방식이다. 이러한 관점은 인공지능 예술에 대한 더 풍부하고 다양한 논의를 가능케 하며, 미래 예술의 새로운 가능성을 탐구하는 데 기여할 수 있을 것이다.

감사의 글

2020년 대한민국 교육부와 한국연구재단의 일반공동 연구 지원사업의 지원을 받아 수행된 연구이다(NRF-2020S1A5 A2A03045921).

참고문헌

- [1] M. Rush, *New Media in Art*, 2nd ed. London, UK: Thames & Hudson, pp. 192-197, 2005.
- [2] DAM Museum. Early Algorithms [Internet]. Available: https://dam.org/museum/artists_ui/artists/mohr-manfred/early-algorithms/.
- [3] CHM (Computer History Museum). Harold Cohen and AARON—A 40-Year Collaboration [Internet]. Available: <https://computerhistory.org/blog/harold-cohen-and-aaron-a-40-year-collaboration/>.
- [4] A. I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity*, Cambridge, MA: MIT Press, p. 116, 2019.
- [5] Google Research. DeepDream [Internet]. Available: <https://ai.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html/>.
- [6] Obvious Art. Edmond De Belamy [Internet]. Available: <https://obvious-art.com/portfolio/edmond-de-belamy/>.
- [7] A. I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity*, Cambridge, MA: MIT Press, pp. 153-157, 2019.

- [8] A. I. Miller, *The Artist in the Machine: The World of AI-Powered Creativity*, Cambridge, MA: MIT Press, pp. 233-241, 2019.
- [9] Neural Zoo. Sofia Crespo's Work [Internet]. Available: <https://neuralzoo.com/>.
- [10] L. A. Gatys, A. S. Ecker, and M. Bethge, "A Neural Algorithm of Artistic Style," arXiv:1508.06576, 2015. <https://doi.org/10.48550/arXiv.1508.06576>
- [11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, ... and Y. Bengio, "Generative Adversarial Networks," *Communications of the ACM*, Vol. 63, No. 11, pp. 139-144, November 2020. <https://doi.org/10.1145/3422622>
- [12] J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, ... and D. C. Schmidt, "A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT," arXiv:2302.11382, February 2023.
- [13] R. Socher, M. Ganjoo, H. Sridhar, O. Bastani, C. D. Manning, C. D. Manning, and A. Y. Ng, "Zero-Shot Learning through Cross-Modal Transfer," arXiv:1301.3666, 2013. <https://doi.org/10.48550/arXiv.1301.3666>
- [14] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," in *Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans: LA, pp. 10674-10685, June 2022. <https://doi.org/10.1109/CVPR52688.2022.01042>
- [15] NewcastAI. Stable Diffusion vs Disco Diffusion [Internet]. Available: <https://chengh.medium.com/stable-diffusion-vs-disco-diffusion-99e3e8957c0d>.
- [16] S. Aris, B. Aeini, and S. Nosrati, "A Digital Aesthetics? Artificial Intelligence and the Future of the Art," *Journal of Cyberspace Studies*, Vol. 7, No. 2, pp. 219-236, July 2023.
- [17] J.-W. Hong and N. M. Curran, "Artificial Intelligence, Artists, and Art: Attitudes toward Artwork Produced by Humans vs. Artificial Intelligence," *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 15, No. 2s, 58, April 2019. <https://doi.org/10.1145/3326337>
- [18] L. Manovich, *AI Aesthetics*, Moscow, Russia: Strelka Press, 2018.
- [19] J.-Y. Goffi, *La Philosophie de la Technique*, 2nd ed. France, Paris: Presses Universitaires de France, pp. 48-49, 1996.
- [20] M. Choi and H. Park, "Introduction of Complex System," *Physics & High Technology*, Vol. 16, No. 10, pp. 1-6, October 2007.
- [21] M. DeLanda, *A New Philosophy of Society: Assemblage Theory and Social Complexity*, London, UK: Continuum, pp. 25-28, 2006.
- [22] R. Dolphijn and I. van der Tuin, *New Materialism: Interviews & Cartographies*, Ann Arbor, MI: Open Humanities Press, p. 64, 2012. <https://doi.org/10.3998/ohp.11515701.0001.001>
- [23] W. Benjamin, *Über Sprache Überhaupt und über die Sprache des Menschen*, Göttingen, Germany: LIWI Literatur- und Wissenschaftsverlag, p. 79, 2019.
- [24] J. H. Lee, *Critique of AI Technology*, Seoul: Communication Books, p. 56, 2019.
- [25] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, ... and J. Dean, "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation," arXiv:1609.08144, 2016. <https://doi.org/10.48550/arXiv.1609.08144>
- [26] R. Sheldrake, *The Science Delusion: Freeing the Spirit of Enquiry*, London, UK: Coronet, p. 241, 2012.
- [27] J. H. Lee, *Resonance: Critique of Media Technology*, Seoul: Communication Books, pp. 10-18, 2019.
- [28] A. Greco, G. Gallitto, M. D'Alessandro, and C. Rastelli, "Increased Entropic Brain Dynamics during DeepDream-Induced Altered Perceptual Phenomenology," *Entropy*, Vol. 23, No. 7, 839, June 2021. <https://doi.org/10.3390/e23070839>



조영각(Youngkak Cho)

2012년 : 추계예술대학교 미술대학 서양화과 (미술학사)

2017년 : 연세대학교 커뮤니케이션대학원 (석사수료)

2023년 : 서울미디어대학원대학교 융합미디어학과 융합예술디자인전공 (미디어콘텐츠학석사)

2015년~2018년: 아트센터나비미술관 랩장

2018년~2022년: 아이아이컴바인드 브랜딩본부 팀장(브랜드: 젠틀몬스터)

2022년~현 재: 스튜디오 신매체 대표

2023년~현 재: 유한책임회사 매드제너레이터 선임연구원

2023년~현 재: 삼육대학교 일반대학원 통합예술학과 겸임교수

※ 관심분야 : 인공지능(Artificial Intelligence), 프롬프트 엔지니어링(Prompt Engineering), 로보틱 아트(Robotic Art), 뉴미디어 아트(New Media Art) 등



김현주(Hyun Ju Kim)

1996년 : 포항공대 산업공학과 (BS)

2004년 : 미국 시라쿠스대학교 (MFA-Computer Art)

2016년 : 서울대학교 언론정보학과 (박사수료 - 디지털미디어문화)

2005년~2009년: 매사추세츠대학 로웰 조교수

2010년~현 재: 서울미디어대학원대학교 융합예술디자인학과 부교수, 확장미디어스튜디오 디렉터

※ 관심분야 : 미디어 아트(Media art), 미디어 미학(Media aesthetics), 미디어 기술(Media technology)