

## 주요 동향(2) : ICT

### 1 미국 AI 스타트업, 텍스트로 음악 생성하는 AI기술 경쟁

⇒ '수노(Suno)'와 '유디오(Udio)', 음악 분야의 생성형 AI 기술 두각

- 하버드大 출신 마이크 술먼이 설립한 AI 기업 '수노(Suno)'는 누구나 음악을 만들 수 있는 미래를 만들겠다는 목표
  - '수노(Suno)'는 2023년 말 마이크로소프트 코파일럿에 플러그인 형태로 AI 음악 모델을 제공한 바 있으며, 지난 4월 더욱 향상된 텍스트 프롬프트에서 음악을 생성하는 모델 'V3' 알파 버전 공개
  - V3는 '사용자 정의 모드'를 통해 자신만의 가사를 쓰고, 제목을 선택하고, 음악 스타일을 지정할 수 있으며, 재생 화면에서 오디오·비디오 형식으로 노래를 공유하거나 다운로드 가능하고 생성된 데이터를 기반으로 요소를 수정하여 노래 리믹스 가능
  - 음악적 지식이 없는 사용자도 텍스트 입력만으로 몇 초 만에 연주와 보컬을 모두 포함하는 완전한 노래를 생성하는 혁신 기능과 뛰어난 품질로 인해 '음악판 소라(Sora)'라는 평가
- 구글 딥마인드 출신으로 구성된 스타트업 유디오(Udio)는 텍스트 프롬프트에서 새로운 고품질 음악 오디오를 생성할 수 있는 음악 생성 AI 앱 '유디오(Udio)' 출시
  - 가사, 스토리, 음악 장르 등을 포함하는 텍스트 프롬프트를 입력하면 40초 안에 노래를 생성
  - '변경(Vary)' 기능을 이용해 무한 편집하거나 길이를 늘일 수 있으며 오디오·비디오 파일을 다운로드해 유디오 커뮤니티는 물론 유튜브와 같은 플랫폼이나 SNS에 공유하는 활용도도 자유로운 편
  - 현재는 베타 기간으로 무료로 사용할 수 있으며 특히 서비스 약관에는 회사가 생성 음악에 대해 어떠한 권리도 주장하지 않으며 상업적 목적으로 사용될 수 있다고 명시
- 스테빌리티 AI(Stability AI)는 오픈소스 기반의 더욱 향상된 '음악 생성 AI' 모델을 지속 업데이트
  - 6월 초 새롭게 공개한 'Stable Audio Open'은 텍스트를 오디오로 변환(text-to-audio)하는 모델로 47초 길이의 고품질 오디오 데이터를 텍스트 프롬프트로 생성할 수 있는 오픈소스 모델

- 드럼 비트, 악기 리프, 환경음, 효과음 등 다양한 오디오 샘플을 생성할 수 있으며 사용자는 자신만의 오디오 데이터로 모델을 미세 조정할 수 있어 창작의 폭을 넓힐 수 있다는 것이 장점

※ “Rock beat played in a treated studio, session drumming on an acoustic kit (스튜디오에서 수록된 록 비트, 어쿠스틱 키트에서의 세션 드럼)”처럼 텍스트로 음악 생성

- 1세대 버전 Stable Audio\*가 고품질의 완전한 트랙을 생성하는 상용 제품으로 최대 3분 길이의 음악 구조를 생성할 수 있는 반면 Stable Audio Open은 짧은 오디오 샘플과 효과음에 특화

\* 디퓨전(Diffusion) 모델에 기반한 음악 생성 AI 모델 ‘스테이블 오디오’ 출시(‘23.9.13.)

- 올해 4월 선보인 ‘Stable Audio 2.0’은 CD 음질 수준인 44.1kHz의 스테레오, 최대 3분 분량 음악 생성 등이 강점

➔ 구글·메타 등 빅테크 기업도 음악과 생성형 AI를 접목한 새로운 혁신 창출

- 주요 빅테크 기업은 글을 넘어 음악 창작 영역에서도 생성형 AI 기술을 결합해 활용도를 높이거나 간단한 명령어만으로 전문가 수준의 음악을 손쉽게 만들어 낼 수 있는 서비스 개발

- 구글은 텍스트-투-뮤직(text to music) 모델 ‘MusicLM’ 발표(‘23.1.)
- 사용자가 ‘우주적 느낌이 나는 레게 톤의 댄스 음악’과 같은 명령어를 입력하면 해당 느낌을 반영하여 짧은 음악을 만들어 주는 모델
- 메타는 AI 개발과 관련해 꾸준히 오픈소스 전략을 구사하며 ‘MusicLM’과 유사한 텍스트-투-뮤직 모델인 ‘MusicGen’을 오픈소스로 출시(‘23.6.)
- 메타가 자체 라이선스를 보유하여 저작권 문제가 없는 2만 시간의 음악을 학습 후 사용자들이 직접 자신의 음악을 학습할 수 있도록 코드까지 공개
- 2023년 8월 초에는 ‘MusicGen’과 효과음 생성 AI인 ‘AudioGen’, AI 음악 압축 모델인 ‘EnCodec’을 모아 음악·소리 관련 AI 모델 패키지 ‘오디오크래프트 (Audiocraft)’를 선보이며 음악 AI 분야 선도자로 부상
- 오픈AI는 주크박스(Jukebox)로 불리는 음악 생성 AI 모델을 2020년에 공개했는데 현재 연구를 완료하고 코드를 공개한 후 추가 작업과 연구는 아직 진척이 없는 상황

출처 : 동아일보 외(2024.5.)

<https://www.donga.com/news/Economy/article/all/20240515/124949073/1>

<https://m.post.naver.com/viewer/postView.naver?volumeNo=37652776&memberNo=31883608>

<https://stability.ai/news/introducing-stable-audio-open>

<https://udioai.ai/ko>

<https://suno.com/blog/v3>