

인공지능 음성합성 시스템의 학습용 데이터에 대한 실연자의 권리*

조 병 한**

한양대학교 문화콘텐츠학과 박사과정
한국방송실연자권리협회 기획운영팀장

■ 국문 초록

본 연구의 목적은 인공지능 음성합성 서비스의 사업자가 인공지능의 학습용 데이터를 확보하기 위해 실연자와 계약을 체결하는 과정에서 실연자의 권리가 부당하게 침해되고 있다는 사실을 밝히고, 실연자의 권리를 보호하는 방안을 모색하는 데 있다.

인공지능 음성합성 서비스를 제공하는 사업자는 거래상 지위를 남용하여 실연자에게 통상의 방송녹음 보수 수준에도 미치지 못하는 적은 금액을 일시불로 지급하는 대가로 학습용 음성데이터에 대하여 인격권의 불행사특약을 포함한 영구하고 무제한적인 이용허락을 요구하였다.

이에 대하여 이 글은 실연자의 저작권법상 인격권, 실연자의 일반적 인격권, 부정경쟁방지법에 의한 실연자의 권리보호 방안을 검토하였으며, 한편으로는 공정한 이용허락관행을 정착시키기 위한 방안을 모색하였다.

이 글은 인공지능 음성합성 시스템의 학습용 데이터에 대한 실연자의 권리를 논하는데 중점을 두었다. 그러나 문제의 본질은 인공지능 생성물로 인한 기술적 실업(technological unemployment) 현상이다. 개별 직군이나 사안에 국한된 것이 아니다. 100여 년 전 상업용 음반의 등장으로 가수나 연주자들이 직면하였던 기술적 실업을 타개하기 위해 저작권접권이

* 이 글은 한양대학교 일반대학원 법학과 '지적재산권법 특수연구' 세미나 수업(당교수 박성호)에서 도움을 받아 작성하였습니다.

** bottleone@hanyang.ac.kr

라는 새로운 권리를 창출하였던 역사적 사실을 교훈으로 삼아, 일국(一國) 차원의 해결이 아닌 국제규범 차원의 해결방안을 모색해야할 시점이다.

주제어: 인공지능, 음성합성 시스템, 실연자의 권리, 저작권법, 콘텐츠산업

목 차

- I. 서론
- II. 인공지능 음성합성 시스템의 텍스트·데이터마이닝
- III. 실연자 권리를 이용하는 계약 사례의 문제
- IV. 현행법에 의한 실연자 권리의 보호
 - 1. 논의의 단서
 - 2. 저작권법상 실연자의 인격권에 의한 보호
 - 3. 실연자의 일반적 인격권에 의한 보호
 - 4. 부정경쟁방지법에 의한 보호
- V. 인공지능 생성물에 의한 실연자의 기술적 실업의 문제
 - 1. 권리침해에 대한 현행법상 보호의 한계
 - 2. 실연자 단체에 의한 보호 사례
 - 3. 입법에 의한 보호의 필요성
- VI. 결론

I. 서론

본 연구는 인공지능 음성합성 시스템의 학습과정에서 수반되는 저작물 이용에 있어서의 실연자 권리 침해 사례를 검토하고, 실연자의 권리를 보호하기 위한 방안을 모색하는 것을 그 목적으로 한다.

음성합성 기술이 점차 고도화되면서 다양한 분야에서 기술의 상용화가 이루어지고 있다. 과거에는 음성합성 기술의 수준이 조악하여 네비게이션, 대중교통 등에서 단순한 정보를 전달하는 데 그쳤지만, 오늘날에는 오디오북, 방송, 유튜브, 게임 등 다양한 문화콘텐츠에서 인간의 실연을 대체하여 이용될 정도로 기술이 발달하였다.

그런데 여기서 문제가 있다. 음성합성 시스템이 인간의 실연을 대체하기 시작했다는 것이다. 보호방안이 마련되지 않는 한 음성합성 시스템의 상용화는 인간 실연자에 대한 권리 침해를 수반한다. 예컨대 유명 성우 이아무개의 목소리를 학습한 인공지능 A가 개발된 상황을 가정해보자. 기존의 저작물에 녹음된 이아무개의 음성을 기반으로 학습을 마친 인공

지능 A에게 대사를 입력하고 적절한 분위기와 톤을 선택해주면, 인공지능 A는 이아무개의 목소리인 것처럼 느껴지는 음성을 생동감 있게 구현해낸다. 인공지능 A는 실수도 없이 언제 어디서나 편리하게 음성을 만들어내므로, 사업자들은 점차 이아무개를 고용하는 대신 인공지능 A를 이용할 것이다. 제도적으로 이아무개를 보호하지 않는다면, 머지않아 일자리를 잃은 이아무개는 인공지능 A를 통해 만들어진 자신의 목소리가 널리 이용되는 모습을 비통한 심정으로 보게 될 것이다. 실연자의 목소리가 널리 이용됨에도 불구하고 실연자가 오히려 곤궁한 상황에 처하게 되는 상황은 결코 바람직하지 않으므로 이에 대한 논의가 필요하다.

지금까지 인공지능의 학습에 관한 논의는 텍스트·데이터마이닝(Text and Data Mining, 약칭 TDM) 일반에 대한 저작권의 보호와 저작권권의 제한을 중심으로 이루어져 왔다.¹⁾ 이러한 논의의 유익으로 공정이용 조항의 요건을 충족하는 경우에 TDM의 대상이 되는 저작물에 대해 저작권권을 제한하는 해석론과 더불어 여러 입법적 보완의 방향성이 제기되었다. 전반을 아우르는 논의가 활발히 이루어져 성과를 거두었으므로 이제는 각론으로 구체적인 TDM 사례에 입각하여 권리 침해가 어떻게 일어나는지 그 양상을 다루는 실천적인 논의가 필요하다.

실연자는 흔히 연주자, 가수 등의 청각 실연자와 배우, 코미디언 등의

1) 2019년 6월 7일 발효된 ‘유럽연합의 디지털 단일시장의 저작권에 관한 지침’ 제2조 제2항에 의하면, TDM은 “유형(patterns), 경향(trends) 그리고 상관관계 등의 정보를 생성하기 위해 디지털 형태의 텍스트와 데이터를 분석하는 것을 목적으로 하는 모든 자동화된 분석기술”을 말한다.

TDM 일반에 대한 저작권의 보호와 저작권권의 제한에 관련한 논의로는 박성호 (2020). 텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작권권의 제한. <인권과 정의>, 제494호, 39-69. ; 오승중 (2019). 데이터 마이닝 및 텍스트마이닝과 저작권권의 제한. <홍익법학>, 제20권 제2호, 455-487. ; 윤정운 (2020). 저작권법상 공정이용에 관한 소고: 인공지능의 빅데이터 활용을 포함한 다양한 영역에서의 사례를 중심으로. <Law&Technology>, 제16권 제5호, 3-28. ; 정상조 (2020). 딥러닝에서의 학습데이터와 공정이용. <Law&Technology>, 제16권 제1호, 3-35. ; 차상욱 (2021). 저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점: 텍스트·데이터마이닝(TDM) 면책규정을 중심으로. <經營法律>, 제32권 제1호, 1-55.

시청각 실연자로 구분하는데, 인공지능 음성합성 시스템에 학습용 데이터를 제공하는 성우는 시청각 실연자에 속하는 것으로 보아야 할 것이다.²⁾ 시청각 실연자의 권리에 대하여서는 주로 영상저작물의 특례에 관한 논의와 시청각 실연에 관한 베이징조약에 대한 연구가 이루어져 왔다.³⁾ 그러나 기술의 변화에 따라 발생하는 구체적인 사례에 입각한 논의는 아직 부족한 실정이다.

이에 이 글에서는 오늘날 활발히 이용되기 시작한 인공지능 음성합성 시스템의 학습용 데이터에 논의를 집중하여, 그 과정에서의 실연자 권리 침해 문제를 논하였다. 이 글의 구체적인 연구문제는 세 가지이다. 첫째는 인공지능 음성합성 시스템의 학습용 데이터에 대한 실연자의 권리가 침해되고 있는 실상을 밝히는 것이고, 둘째는 개별 권리 침해 사안에 대하여 현행법상 실연자 권리 보호방안을 모색하는 것이며, 마지막으로 셋째는 근본적인 문제 해결을 위해 실연자의 권리를 향상할 수 있는 방향성과 입법의 필요성을 제시하는 것이다.

-
- 2) 성우의 실연은 본질적으로 음성이므로 성우가 청각 실연자라는 반론이 제기될 수 있으나, 근본적으로는 인간이 행하는 모든 실연은 시청각적으로 행하여지는 실연이며 저작물의 형태와 실연의 범주에 따라 청각 실연자와 시청각 실연자가 구분된다는 사실을 고려한다면 성우를 청각 실연자로 보기 어렵다. 성우는 주로 영상저작물의 제작에 참여하여 연기하기 때문이다. 게다가 스스로의 권의 보호를 위해 텔런트, 코미디언 등의 시청각실연자와 함께 한국방송연기자노동조합, 한국방송실연자권리협회에서 활동하고 있으므로 시청각실연자의 범주에 속한다고 보아야 타당할 것이다.
- 3) 영상저작물의 특례에 관한 논의는 주로 특약이 없는 한 시청각 실연자의 권리가 영상제작자에게 양도된다는 문제에 착목하여 실연자의 권리 보호방안을 논하고 있으며, 시청각 실연에 관한 베이징조약에 대한 논의는 주로 베이징조약의 내용과 조사하고 우리 저작권법과 함께 분석하는 비교법적 연구가 이루어져 왔다. 영상저작물의 특례의 해석에 관하여서는 김기복 (2014). 시청각 실연자(視聽覺 實演者)의 권리보호 방안. <한국콘텐츠학회논문지>, 제14권 제12호; 안효길 (2018). 영상저작물의 이용활성화와 실연자의 보상청구권. <법학논총>, 제42권 제3호 참조. 시청각 실연에 관한 베이징조약에 관하여서는 이동기 (2016). 시청각 실연자의 권리와 그에 관한 베이징 조약 연구. <圓光法學>, 제32권 제3호; 신창환 (2019). 시청각 실연에 관한 국제 보호 규범과 우리나라에서의 발전 방향 연구. <계간 저작권>, 제32권 제3호; 한지영 (2016). 시청각실연자의 법적 보호에 관한 고찰. <계간 저작권>, 제29권 제4호 참조.

II. 인공지능 음성합성 시스템의 텍스트·데이터마이닝

음성합성 기술은 대중교통의 안내방송, 내비게이션 음성 등 간단한 정보를 전달하는 용도로 과거에도 존재해왔으나, 최근 인공지능말 알고리즘이 접목되면서 빠르게 발전하고 있다. 지금까지 음성합성 기술의 대부분은 오픈소스로 개발되고 있으며, 대표적인 음성합성 기술로는 네이버 클로바 보이스, 구글의 타코트론(Tacotron), 바이두의 딥보이스(DeepVoice)가 있다.⁴⁾

그중에서도 구글의 타코트론은 딥러닝을 이용한 최초의 종단형(End-to-end) 음성합성 기법이다.⁵⁾ 타코트론은 2017년 구글에 의해 개발된 음성합성 기술로, 기존의 복잡한 처리가 필요했던 음성합성 기술과 달리 다수의 〈텍스트, 오디오〉 쌍을 입력하는 것만으로 완전한 학습이 가능하다.⁶⁾

타코트론은 24.6시간의 음성 데이터를 기반으로 음성합성 능력을 갖추었고, 당시 타코트론 개발팀의 음성 품질 측정 테스트(Mean Opinion Score Test)에서 3.82점을 받았다.⁷⁾ 그리고 같은 해 12월 인코더와 디코더의 모델 구조를 변경하고, 새로운 보코더를 적용한 타코트론2가 개발팀의 MOS 테스트에서 4.526점을 받았는데, 이것은 실제 사람의 음성을 평가한 점수인 4.582점과 거의 유사한 수준이었다. 이후 음성합성 속도를 향상시킨 FastSpeech 1, 2, 2S, SpeedySpeech가 발표되었고, 최근에는

4) 김경호·성혜림·성미경·정찬경·성진택 (2020, 8월). 〈타코트론 모델을 이용한 개인 맞춤형 한국어 음성합성 시스템〉. 2020년 대한전자공학회 하계학술대회, 제주: 롯데호텔, 2089.

5) 타코트론의 설계에 관하여는 Yuxuan Wang et al. (2017). Tacotron: Towards End-to-End Speech Synthesis. *Proc. Interspeech*, 1-10 참조.

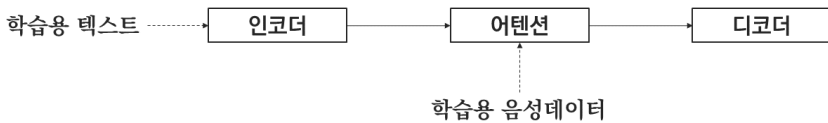
6) 위의 논문, 2.

7) 음성 품질 측정 테스트(MOS)는 음성전화의 통화품질을 구분하는 평가 방법으로, 사람이나 가공된 음성을 수화기를 통해 피평가자에게 들려주고 이들이 평가하는 음성 품질수준을 5등급(우수)에서 1등급(불만족)으로 나누어 표현한다. 음성의 품질을 인지하는 것은 인간의 주관적인 영역이므로 객관적 데이터를 분석하는 방법보다 주로 실제 사람에 의해 품질을 평가하는 방법을 사용한다.

다양한 감정을 표현하는 동시에 화자의 어투까지 모방할 수 있는 음성 합성 시스템 Flowtron, VAE(Variational Autoencoder) Tacotron도 발표되었다.⁸⁾

이처럼 인공지능 음성합성 기술은 빠르게 발전하고 있다. 타코트론은 24.6시간의 학습용 음성 데이터가 필요했지만, 네이버의 인공지능 음성 합성 기술인 클로바 보이스 NES는 약 40분 수준의 음성 녹음만으로 고품질의 음성합성 시스템을 제작할 수 있는 수준에 이르렀으며, 엔씨소프트는 김영하 작가가 녹음한 10분 정도의 음성 데이터로 음성합성 시스템을 만드는 데 성공했다.⁹⁾

주목할 것은 인공지능 음성합성 기술이 얼마나 발전했건 간에, 또 얼마나 적은 음성 데이터를 가지고 시스템을 만들어낼 수 있건 간에 반드시 학습용 데이터를 필요로 한다는 사실에 있다. 학습용 데이터로부터 패턴을 추출하는 텍스트·데이터마이닝은 인공지능의 필요조건이다. [그림 1]은 타코트론의 텍스트·데이터마이닝 과정을 도식화한 것이다.



[그림 1] 타코트론 TDM 도식도

타코트론의 TDM은 다음과 같은 절차를 통해 이루어진다.¹⁰⁾ 먼저 타

8) SungWoo Moon & SungHyun Kim & Yong-Hoon Choi (2022). MIST-Tacotron: End-to-End Emotional Speech Synthesis Using Mel-Spectrogram Image Style Transfer. *IEEE Access*, 10, 25455-25456.
 9) 실제로 결과물을 들어보면 아직까지 적은 시간으로 제작된 음성합성 결과물은 충분한 학습을 거친 결과물보다 질적인 수준이 낮은 것으로 보인다. 김인경 (2020, 2, 10). 녹음 없이 동영상 더빙을? 네이버, ‘클로바더빙’ 출시. <BLOTTER>. URL: <https://www.bloter.net/newsView/blt202002100011> ; 엔씨소프트, AI 김영하, 읽다. URL: <https://blog.ncsoft.com/ai-tech-youngha-kim/>
 10) 김경호, 앞의 글, 2089-2090 ; 조수희 (2019). <딥러닝 기반 보코더의 음성 합성에 대한 음성학적 관찰>. 고려대학교 대학원 석사학위논문, 27-31 참조.

코트론에 학습용 데이터의 텍스트가 입력되면 인코더는 텍스트를 128차원의 벡터값으로 변환한다. 이 과정의 목표는 텍스트의 특징을 잘 표현해내는 것이다. 벡터값이 도출되면 어텐션 과정에서 벡터값의 어떤 부분을 디코더에서 집중할 것인지 판단하고 구분하게 된다. 예컨대 “저작물을 이용하기 위해서는 저작권자의 허락을 받아야한다.”라는 문장을 학습한다고 가정한다면, 이에 대한 음성 데이터는 [저작무를이용하기위해서는저작권자의허라글바다야한다] 따위가 되는데 여기서 전체 텍스트와 전체 음성 데이터를 비교하여 분석한다면 문장의 길이가 길어질수록 학습의 난이도가 높아지고 결과의 질은 떨어질 것이다. 이러한 문제를 해결하기 위하여 어텐션 과정에서는 데이터를 일정 단위로 분리하여 보다 효율적인 학습이 이루어질 수 있도록 하였다. 디코더는 어텐션으로부터 판단된 벡터값을 받아 학습용 음성데이터와 비교하여 학습한다. 학습용 데이터를 기반으로 타코트론은 상술한 일련의 과정을 수없이 거쳐 음성합성을 해내는 인공지능을 갖추게 된다.

그런데 타코트론의 개발팀은 “학습과정에서 우리는 언제나 해당 순서의 실제 음성데이터를 디코더에 입력하였다(During training, we always feed every r -th ground truth frame to the decoder.)”고 하였다.¹¹⁾ 타코트론이 목소리를 내기 위해서는 반드시 실제 음성데이터, 즉 실연이 필요하다라는 것이다.

이를 통해 다음의 사실을 추론할 수 있다. 첫째, 인공지능 음성합성 시스템이 만들어내는 음성의 질은 학습의 대상이 되는 실연의 질과 양의 상관관계를 갖는다.¹²⁾ 인공지능 음성합성 시스템의 학습은 반드시 실제 음성데이터를 통해 이루어지기 때문에 실연의 질의 좋다면 인공지능의 학습결과 역시 좋을 것이고, 실연의 질이 나쁘다면 인공지능의 학습결과는 나쁠 수밖에 없다. 둘째, 인공지능 음성합성 시스템은 텍스트·데이터마이닝 과정에서 실연자의 발음패턴, 발성법, 억양 등 개인을 식별

11) Y. Wang et al. 앞의 글, 5.

12) 여기서 실연의 질은 발음, 발성 등 음성의 표현력을 의미한다.

할 수 있는 표지를 복제한다. 타코트론의 텍스트·데이터마이닝 과정에서 드러난 것처럼 인공지능 음성합성 시스템의 학습은 텍스트를 인코더로 변환한 데이터와 음성데이터를 비교함으로써 이루어진다. 실연자에게는 2차적저작물작성권이 부여되지 않은 바, 실연을 모방한 것에 대하여 권리를 주장할 수 없지만, 인공지능 음성합성 시스템은 기본적으로 실연 그 자체를 복제하는 과정을 통해 실연자의 인격성 중 일부인 목소리를 만들어내는 능력을 갖추게 되므로 실연자의 권리가 미치지 못한다고 볼 수 없을 것이다. 게다가 인공지능 음성합성 결과물은 실연자의 인격적 표지에 해당하므로 부적절한 내용이 음성으로 합성되어 유포되는 경우에 저작권법상 실연자의 인격권 내지는 실연자의 일반적 인격권이 침해될 수 있으며, 음성이 광고 등의 상업적 목적으로 이용되는 경우에는 퍼블리시티권 침해의 소지가 있다. 이에 대한 구체적인 해석은 4장에서 후술한다.

Ⅲ. 실연자 권리를 이용하는 계약 사례의 문제

본 연구는 실연자 권리 침해 사례에 대한 논의를 위해 실제 이용자가 성우에게 제시한 계약서를 검토하였다. 이 계약서는 용어상으로도 문제가 있으며, 실연자에게 불합리한 계약조건을 제시하고 있다.

Voice DB 재산권 독점적 **이용허락** 계약서(이하, ‘계약서’라고 한다)¹³⁾
 저작자 및 저작권 **이용허락자** A(이하 “권리자”이라 함)와 저작권 이

13) 인공지능 음성합성 시스템과 실연자 사이에 체결하는 계약서의 내용은 개개 사례마다 상이하지만 일시불의 대가로 권리를 모두 양도한다는 점에서 대동소이하다. 이같은 계약의 양상을 구체적으로 살피기 위하여 실제로 인공지능 음성합성 시스템 사업자가 실연자에게 제시한 계약서를 한국성우협회를 통해 확인하였다. 실연자는 A, 이용자는 B, 저작물은 C로 칭하였으며, 본 연구에서 논의가 필요한 부분은 굵게 표시하였다.

용자 B(이하 “이용자”라 함)는 아래 저작물 C에 관한 저작권산권 **이용허락**과 관련하여 다음과 같이 계약을 체결한다. 다음

제1조 (계약의 목적) 본 계약은 저작권산권 **이용허락**과 관련하여 권리와 이용자 사이의 권리관계를 명확히 하는 것을 목적으로 한다.

제2조 (계약의 대상) 본 계약의 **이용허락** 대상이 되는 권리는 아래의 저작물(이하 “대상저작물”)에 대한 저작권산권 중 당사자가 합의한 권리로 한다.

저작자: A / 종별: 음성합성 AI Voice DataBase / 권리: 독점 사용권

제3조 (**이용허락** 기간) 대상저작물의 독점사용권은 B에게 **영구적으로 양도**된다.

제4조 (권리자의 의무)

- (1) 권리자는 이용자에게 대상저작물에 대하여 제2조에서 규정한 범위 내에서 **독점적으로 이용하는 것을 허락**한다.
- (2) 권리자는 이용자에게 0년 0월 0일까지 대상저작물의 이용을 위해 필요한 상당한 자료를 인도하여야 한다.

제5조 (이용자의 권리 및 의무)

- (1) B는 관례적으로 저작자 및 저작권산권자의 성명 등 표시를 허용하는 대상저작물을 이용하는 경우, 그 저작자 및 저작권산권자의 성명 등을 표시하여야 한다.
- (2) B는 대상저작물의 이용함에 있어서 저작권인격권을 침해하지 아니한다.

제6조 (확인 및 보증)

- (1) A는 B에게 다음 각 호의 사항을 확인하고 보증한다.
 1. 대상저작물의 **저작권이용허락**을 체결하는데 필요한 권리 및 권

한을 적법하게 보유하고 있다는 것

2. 대상저작물의 내용이 제3자의 저작권, 상표권, 인격권을 비롯한 일체의 권리를 침해하지 아니한다는 것
 3. 본 계약을 체결하기 전 대상저작물은 제3자에게 저작재산권이 양도되거나, 이용허락되었거나, 질권이 설정되는 등 이용자의 독점적 이용권을 제한하는 어떠한 부담도 존재하지 아니한다는 것
 4. 본 계약에 따른 **이용허락** 기간 동안 제3자에게 대상저작물에 대한 저작재산권의 양도, 이용허락, 질권의 설정 등 이용자의 독점적 이용권을 침해하는 어떠한 행위도 하지 아니한다는 것
- (2) B는 A에게 다음 각호의 사항을 확인하고 보증한다.
1. 대상저작물 이용허락권을 권리자의 동의 없이 제3자에게 양도하거나 재이용을 허락하지 아니한다는 것
 2. 대상저작물을 제3자의 명예권을 비롯한 인격적 권리를 침해하는 방식으로 이용하지 아니할 것

제7조 (계약내용의 변경) 본 계약 내용 중 일부를 변경할 필요가 있는 경우에는 권리자와 이용자의 서면합의에 의하여 변경할 수 있으며, 그 서면합의에서 달리 정함이 없는 한, 변경된 사항은 그 다음날부터 효력을 가진다.

제8조 (계약의 해지)

- (1) 당사자는 천재지변 또는 기타 불가항력으로 계약을 유지할 수 없는 경우에 본 계약을 해지할 수 있다.
- (2) 당사자는 상대방이 정당한 이유 없이 본 계약을 위반하는 경우에 상당한 기간을 정하여 상대방에게 그 시정을 최고하고, 상대방이 그 기간이 지나도록 이행하지 아니하는 경우에는 계약을 해지할 수 있다. 다만, 상대방이 명백한 시정 거부 의사를 표시하였거나 위반 사항의 성격상 시정이 불가능하다는 것이 명백히 인정되는 경우에는 위와 같은 최고 없이 계약을 해지할 수 있다.
- (3) 본 계약에 대한 해지권의 행사는 상대방에 대한 손해배상청구권

행사에 영향을 미치지 아니한다.

제9조 (손해배상) 당사자가 정당한 이유 없이 본 계약을 위반하는 경우, 그로 인하여 상대방에게 발생한 모든 손해를 배상할 책임이 있다. 다만, 제8조 제1항의 사유로 본 계약을 이행하지 못한 경우에는 손해배상책임을 면한다.

제10조 (비용의 부담) 계약 체결에 따른 비용은 당사자가 동등하게 부담한다.

제11조 (분쟁해결)

- (1) 본 계약에서 발생하는 모든 분쟁은 권리자와 이용자가 상호 원만한 합의에 이르도록 노력하여야 하며, 분쟁이 원만히 해결되지 않는 경우에는 소제기에 앞서 한국저작권위원회에 조정을 신청할 수 있다.
- (2) 제1항에 따라 해결되지 아니할 때에는 대한민국의 민사소송법 등에 따른 관할법원에서의 소송에 의해 해결토록 한다.

제12조 (비밀유지) 양 당사자는 본 계약의 체결 및 이행과정에서 알게 된 상대방에 관한 정보, 본 계약의 내용 및 대상저작물의 내용을, 상대방의 서면에 의한 승낙 없이 제3자에게 공개하여서는 아니 된다.

제13조 (기타부속합의)

- (1) 권리자와 이용자는 본 계약의 내용을 보충하거나, 이 계약에서 정하지 아니한 사항을 규정하기 위하여 부속합의서를 작성할 수 있다.
- (2) 제1항에 따른 부속 합의는 본 계약의 내용과 배치되거나 위반하지 않는 범위 내에서 유효하다.

제14조 (계약의 해석 및 보완) 본 계약서에서 명시되어 있지 아니하거나 해석상 이견이 있을 경우에는 저작권법, 민법 등을 준용하고 사회 통념과 조리에 맞게 해결한다.

제15조 (계약 효력 발생일) 본 계약의 효력은 계약 체결일로부터 발생한다.

먼저 위 계약서는 용어상의 문제를 지닌다.¹⁴⁾ 가장 큰 문제는 ‘이용허락’과 ‘양도’를 혼용하고 있다는 것이다. 해당 계약서는 제목과 본문에서 ‘이용허락’에 대한 계약서임을 명시하고 있지만, 계약서 제3조에서는 독점사용권을 영구적으로 ‘양도’한다고 규정함으로써 용어상의 혼란을 일으키고 있다. 저작(인접)물의 이용허락은 저작(인접)권자가 제3자에게 저작(인접)물의 이용을 허락하고 그 대가를 받는 것으로서 저작(인접)권이 저작(인접)권자에게 유보된다는 점에서 양도와는 차이가 있다.¹⁵⁾ 만약 계약서의 성격이 양도계약이라면 실연자에게 저작권법상 실연자의 재산권이 없다고 보아야 할 것이다.

그러나 본 계약서 제3조에서 언급한 ‘독점사용권’은 배타적 이용허락으로 해석하는 것이 타당하다. 그 까닭은 ①계약서의 제목과 본문에서 이용허락 계약임을 명시하고 있으며, ②제1조에서도 본 계약이 이용허락 계약임을 명확히 하였으며, ③제3조의 제목이 ‘이용허락 기간’이며, ④제4조의 제1항에서도 이용자가 이용을 허락한다고 규정하였기 때문이다. 특히 제3조의 제목이 ‘이용허락 기간’임을 고려하면 조문의 목적이 영구히 이용허락이 이루어진 것을 규정하는 데 있으며, 문서상 이를 표현하는 과정에서 저작인접물을 배타적으로 이용할 수 있는 권리를 ‘독점사용권’으로 표현하고 이를 영구히 이용자에게 귀속시키기 위해 ‘양도’라는 표현을 사용한 것으로 볼 수 있다. 계약서의 맥락을 고려하였을 때 본 계약은 명백한 이용허락 계약임에도 개별 조문에서 권리를 양도한다는

14) 실연자인 성우에게는 실연의 결과로서 저작인접권이 발생함에도 계약서에서는 ‘저작자’ ‘저작권’ ‘저작재산권’ ‘저작인격권’ ‘저작물’이라는 용어를 사용하고 있다. 이는 모두 ‘저작인접권자’ ‘저작인접권(또는 실연자의 재산권)’ ‘실연자의 인격권’ ‘저작인접물’이라는 용어로 변경되어야 할 것이다.

15) 노현숙·구천을 (2016). 저작재산권 양도와 저작물 이용허락의 비교 고찰. <江原法學>, 제48권, 246.

표현을 사용하여 혼동의 여지를 두었다.

다음으로 계약조건의 문제는 보다 본격적인 권리 침해의 문제이다. 저작권 계약에서는 경제적 약자의 지위에 있는 자들이 권리를 주장하지 못하고 침묵함으로써 유지되어 온 관행이 있는데, 상기의 계약 역시 이러한 관행을 잘 보여준다. 해당 계약서는 대상저작인접물의 이용허락기간을 영구히 하는 동시에 이용의 범위를 무제한적으로 허락하도록 하였다. 본 계약서의 부속합의서는 이용자가 대상저작인접물을 어떠한 목적으로도 ‘활용’할 수 있는 권리를 갖는다고 정의하고, ‘활용’을 제2조 제5항에서 다음과 같이 규정하였다.

보이스 데이터 또는 결과물의 “활용”이란 보이스 데이터 또는 결과물을 ‘이용자’가 판매·개발·연구 등 상업적·비상업적 목적 여부를 불문하고 어떠한 목적으로든 이용·저장·편집·배포·가공하는 등 서비스 제공에 필요한 형태로 사용하는 일체의 행위와 ‘이용자’의 서비스를 이용하는 이용자가 보이스 데이터 또는 결과물을 이용·저장·편집·배포·가공하는 등 서비스 이용에 필요한 형태로 사용하는 일체의 행위를 모두 포함한다. 또한 ‘이용자’ 또는 ‘이용자’의 서비스를 이용하는 이용자의 보이스 데이터 또는 결과물의 활용을 통해 배포된 보이스 데이터 또는 결과물을 제3자가 활용하는 것까지도 위 “활용”의 범위에 포함되는 것으로 본다. 단, 보이스 데이터 자체를 제3자에게 재판매하는 행위는 “활용”에 포함되는 것으로 보지 않는다. 명확히 하자면, ‘이용자’는 보이스 데이터를 활용하여 연기용 TTS/VC 기술 또는 다른 기술을 개발하고 이를 활용한 서비스 등을 개발할 수 있으며, 개발한 기술이나 서비스를 통해 가공된 결과물을 직접 활용하거나 제3자가 활용하도록 허락할 수 있으나, 가공되지 않은 보이스 데이터 전부 또는 일부를 제3자에게 재판매하는 행위는 허용되지 아니한다.¹⁶⁾

계약서의 부속합의서의 요지는 이용자의 대상저작인접물 이용에 어떠

16) 이용자의 명칭은 ‘이용자’로 치환하였다.

한 제한도 없다는 것이며, 이용자의 서비스를 이용하는 고객의 이용과 그로부터 파생하는 제3자의 이용까지도 이용허락의 범위에 포함한다는 것이다. 이는 바꾸어 말하자면, 인공지능 음성합성 시스템이 만들어내는 모든 결과물에 대해서 실연자의 권리를 박탈한다는 것이라 해석할 수 있다. 계약을 체결한 실연자는 자신의 목소리가 자신의 의사와 관계없이 유통되어도 어떠한 권리도 주장할 수 없다. 이용자의 고객이 만든 음성을 제3자가 활용하는 것에 대해서도 말이다.

물론 아무리 범위가 넓고 기간에 한정이 없는 계약이라도 정당한 대가에 의거한 것이라면 계약에 의거한 이용이 실연자의 권리를 침해한다고 볼 수는 없다. 하지만 문제가 되는 것은 이용허락에 대한 합리적인 보상의 문제이다.

실제 인공지능 음성합성 시스템을 제작하기 위해서는 성우 1인당 20시간에서 30시간 분량의 음성데이터가 소요된다. 1시간 분량의 음성데이터를 녹음하는 데 평균 2시간의 음성작업이 필요하다는 점을 고려하였을 때, 성우 1인당 적게는 40시간에서 많게는 60시간의 음성작업이 필요한 것이다. 성우는 작업에 대한 대가로 음성데이터 1시간 당 15만원에서 30만원 수준의 보수를 받는다. 그렇다면 성우는 자신의 목소리를 합성해내는 인공지능 음성합성 시스템을 제작하는데 최소 300만원에서 최대 900만원의 보수를 받게 된다.¹⁷⁾ 거래는 일시불을 지급하고 권리일체를 영구히 이용허락함으로써 사실상 양도하는 형태로 이루어진다.

성우는 직업 특성상 고정 급여가 없어서 방송사 공채로 활동해도 안정적인 소득이 보장되지 않고, 경제적인 측면에서 안정성이 취약하여 소규모 자영업과 병행하는 사례가 많다.¹⁸⁾ 한국콘텐츠진흥원의 실태조사에 따르면, 성우를 포함한 연기자의 월평균 개인 소득의 중앙값은 180만원

17) 이 글의 인공지능 음성학습 시스템의 음성데이터 녹음시간, 보수에 대한 언급은 2022년 3월 23일 사단법인 한국성우협회 최재호 사무총장을 대상으로 진행된 인터뷰에 의한 것이다. 인터뷰에 응해주신 최재호 사무총장께 감사드린다.

18) 이은숙 (2021). <2021 대중문화예술산업 실태조사>. 나주: 한국콘텐츠진흥원, 278.

에 불과하며 그중에서도 대중문화예술 관련 소득은 100만원에 불과하였다.¹⁹⁾ 이러한 처지에 놓인 성우는 계약 지위상 매우 취약하여 미래에 발생할 수 있는 자신의 권리를 제한받으면서도 울며 겨자 먹기로 계약을 체결하게 된다. 더군다나 계약 체결 과정에서 저작권접권 문제를 제기하면 이후 녹음 작업에서 배제되는 사례도 빈번히 발생하기 때문에 불합리한 계약의 내용을 변경해달라고 요구하기도 어렵다.²⁰⁾

지상파방송사의 성우출연조건표에 의하면 60분 방송물에 대한 성우의 보수는 30만 원에서 40만 원 수준이다. 게다가 성우는 한국방송실연자 권리협회와 방송사가 체결한 특약에 따라 방송프로그램이 재방송·복제·전송되는 경우에 대해 저작권접권 사용료도 받게 된다. 방송프로그램 녹음에 대한 보수의 수준이 인공지능 음성학습 시스템의 학습용 음성데이터에 대한 보수보다 높은 것이다. 게다가 그 활용도를 비교한다면 인공지능 음성합성 시스템의 학습용 음성데이터에 대한 보수는 심각한 수준으로 적은 상황이다. 그야말로 계약지위를 악용한 불공정한 계약이다.²¹⁾ 이러한 계약행태에 대한 개선방안이 마련되지 않으면 많은 성우가 부당한 조건으로 자신의 목소리를 빼앗기게 될 것이다.

인공지능 음성합성 시스템이 시장에서 서비스를 시작하면 실연자는 시장에서 영구적으로 자신의 목소리와 경쟁해야 하며, 최악의 경우에는 영원히 성우로서 일자리를 잃을 수도 있다. 인공지능 음성합성 시스템은 비용을 요구하지 않으면서도 쉬지 않고 일관된 품질의 음성을 합성해낼 수 있기에, 시장에서 실연자를 대체해나갈 것이다. 그렇다면 실연자는 통상의 방송녹음보다도 적은 수준의 일시불의 대가를 지급받은 후 일자

19) 위의 글, 201-202.

20) 한동현 (2021, 12, 6). 성우가 목소리 노예?...IT기업, 음성 저작권 영구 양도 요구. <서울와이어>. URL: <https://www.seoulwire.com/news/articleView.html?idxno=456552>

21) 민법 제104조는 당사자의 공박, 경솔 또는 무경험으로 인해 현저하게 공정을 잃은 법률행위는 무효에 해당한다고 하였다. 그러므로 일부 사례에 한하여 실연자의 공박함을 입증한다면 계약을 무효화할 수 있을 것이다. 다만 본 연구는 보다 일반적인 대응방안에 논의를 집중하고자 하였다.

리를 잃어버리게 된다. 결국 성우를 직업으로 가지려는 사람이 적어질 것이고, 장기적으로는 시장의 인력 부족으로 인해 실연의 수준이 전반적으로 하락할 것이다. 시대의 흐름에 따라 대중이 요구하는 실연의 스타일이 변화하므로 새로운 스타일의 음성데이터를 기반으로 인공지능 음성학습 시스템을 지속적으로 제작해야 할 것인데, 인공지능 음성학습 시스템 역시 인간의 음성 데이터를 기반으로 학습하는 바, 인공지능 음성학습 시스템도 당대 실연의 전반적인 수준 저하에서 자유로울 수 없다.

실연의 수준 저하는 콘텐츠의 소구력을 저하할 것이고, 이는 생산-유통-소비의 콘텐츠산업 가치사슬 전반에 연쇄적인 영향을 미쳐 결과적으로는 산업생태계의 선순환 구조를 훼손하게 된다. 디지털 전환에 따라 플랫폼 등 콘텐츠 유통에 참여하는 사업자들에게 가치가 집중됨에 따라 콘텐츠 생산자에게 배분되는 부가가치가 줄어드는 ‘유통 사업자로의 가치이전(Value Transfer) 현상’에 대한 우려는 오래전부터 존재해왔다.²²⁾ 인공지능 음성합성 시스템의 실연자에 대한 불공정계약 문제도 이와 맥락을 함께한다. 생산자와 소비자 사이에 대규모 유통 사업자가 포진하면서 그들에 의해 부가가치가 좌우되어 소비자가 지불한 대가가 정작 콘텐츠 생산자에게 도달하지 못하는 것이다. 대규모 유통 사업자를 중심으로 한 가치이전 현상은 콘텐츠산업은 규모를 키워왔지만, 지속적인 성장을 위해서는 반드시 풍부하고 다양한 콘텐츠가 있어야 하며, 콘텐츠 생산자에게 적절한 보상이 주어지지 않는 왜곡된 구조는 장기적으로 산업생태계 전체의 성장을 저해할 수밖에 없다.²³⁾ 이제 상용화를 시작하는

22) 김영재 (2014). 디지털 미디어 환경의 콘텐츠 산업 진흥기금. <한국콘텐츠학회 논문지>, 제14권 제2호, 148.

23) “올바른 문화콘텐츠 생태계가 발현되고 육성되려면 무엇보다도 그 스스로 외부 환경에 대응하여 자기조직화를 이룰 수 있는 패턴과 구조, 그리고 과정을 확보하여야 한다. 패턴은 문화콘텐츠 생태계를 구성하는 각 요소들 간의 관계 설정이며, 구조는 그들 간의 상호 관계, 그리고 과정은 관계간의 안정성과 효율성, 그리고 지속성을 말한다. 특히 이러한 요소들은 문화콘텐츠 생태계의 긍정적 발전, 즉 진화적 상황의 도출이란 결과를 획득할 수 있는 구조의 구축이 무엇보다 중요하다. 이런 관점에서 볼 때 문화콘텐츠 생태계는 기존의 유통 중심의 관점, 기술 중심의 관점, 경제성 중심 관점에서 벗어나 생산과 소비 중심의 순환의 관점으로서의

단계에 있는 인공지능 음성합성 서비스 역시 마찬가지일 것이다. 인공지능 음성합성 시스템은 음성합성 기술을 대중화함으로써 개개 콘텐츠 생산자의 창작활동을 촉진한다는 점에서 커다란 의의가 있지만, 실연자에 대한 불공정한 계약을 기반으로 사업을 확장해나간다면 콘텐츠산업의 가치사슬 전반에 부정적인 영향을 미치게 될 것이다. 장기적이고 지속적인 산업의 공진화(共進化)를 위해서는 공정한 계약 관행이 갖추어져야 한다.

IV. 현행법에 의한 실연자 권리의 보호

1. 논의의 단서

인공지능 음성합성 시스템의 학습용 데이터에 관한 불공정한 보상체계는 미시적으로는 개별 실연자의 인격권을 침해하고, 거시적으로는 전체 실연의 수준을 저하시킴으로써 산업의 선순환구조를 파괴한다. 그러므로 인공지능 음성합성 시스템의 학습용 데이터 제공 계약에 있어서는 실연자를 보호하고 산업의 선순환구조를 지탱하기 위해 실연자의 인격권을 강하게 보호하는 합목적적 해석이 필요하다. 인공지능 음성합성 시스템의 불공정한 계약 관행에 의해 실연자의 재산권은 일체 무제한적으로 이용허락 되고 있으므로, 이 글에서는 실연자의 인격권 내지는 실연자의 일반적 인격권에 의한 실연자 권리 보호방안과 인공지능 음성합성 시스템의 학습용 데이터 이용허락 계약이 부정경쟁방지법상 부정경쟁행

재편을 필요로 하고 있다. 특히 중요한 것은 1차 생산자 개체 집단군에 대한 관심과 배려의 강화이다. 그들이 보다 편리하게 생산 체계에 참여하고, 그들에 대한 보상폭을 확대해 나가는 것은 문화콘텐츠 생태계에 가장 근간을 이루는 사항이 되고 있다.”

류준호·윤승금 (2010). 생태계 관점에서의 문화콘텐츠 산업 구성 및 구조. <한국콘텐츠학회 논문지>, 제10권 제4호, 337.

위에 해당하는지 여부를 중심으로 문제를 검토하였다.²⁴⁾

2. 저작권법상 실연자의 인격권에 의한 보호

저작권법이 실연자에게 부여한 인격권은 동일성유지권과 성명표시권이다. 인공지능 음성합성 시스템이 사업을 개시한다면 동일성유지권과 성명표시권 모두에 있어서 권리침해가 이루어질 가능성이 있다. 인공지능 음성합성 시스템은 새로운 음성을 만들어낼 때마다 실연의 동일성을 훼손하게 된다. 인공지능 음성합성 시스템이 만들어내는 음성은 실연자의 목소리로 텍스트를 발음한 것으로, 텍스트·데이터마이닝을 통해 학습한 원 실연의 동일성을 왜곡하는 경우도 발생할 수 있기 때문이다. 실연의 동일성을 훼손하는 왜곡의 전형적인 예로는 그로 인해 실연자의 명예를 훼손하는 경우이다. 2장에서 논한 바 인공지능 음성합성 시스템은 실연자의 목소리, 즉 개인의 고유성을 식별할 수 있는 표지 그 자체를 복제하는 것이며 그로 인해 인공지능 음성합성 시스템이 만들어낸 음성을 통해서도 실연자가 행하는 실연의 특성을 감지할 수 있으므로, 동일성유지권 침해의 성립요건 중 하나인 ‘표현형식상 실연의 특성이 감지될 수 있는 정도의 개변’에 해당한다.²⁵⁾ 더불어 성명표시권의 침해도 발생할 수 있다. 계약에서 명시한 ‘활용’의 범위가 이용자와 이용자의 고객이

24) 저작권법상 실연자는 인공지능 음성합성 시스템의 학습용 데이터에 대하여 성명표시권·동일성유지권의 인격권과 복제권·배포권·대여권·공연권·방송권·전송권의 재산권을 가진다. 그러나 인공지능 음성합성 시스템의 사업자는 앞서 살핀 사례에서처럼 권리 일체의 양도 내지는 무제한적 이용허락을 요구하고 있고 계약지위상 약자인 실연자는 사업자의 요구를 거절하지 못하는 실정이기 때문에 저작권법상 실연자의 재산권에 기반하여 보호방안을 모색하는 것은 부적절하다.

25) “한편, 동일성유지권은 타인의 저작물의 창작성이 인정되는 표현 부분에 삭제나 추가 또는 변경 등의 왜곡행위가 이루어진 경우에만 발생하므로 동일성유지권 침해가 성립하기 위해서는 타인의 저작물의 표현형식상의 창작적 특성이 감지될 정도로 이용되어야 한다.”(박성호 (2017). <저작권법>(제2판). 서울: 박영사, 292 참조). 저작자에게 인정되는 동일성유지권의 경우를 유추한다면 실연자의 동일성유지권 침해가 성립하기 위해서는 타인의 실연물의 표현형식상 실연의 특성이 감지될 정도로 이용되어야 할 것이다.

인공지능 음성합성 시스템을 이용하여 만들어진 결과물을 제3자가 이용하는 것까지 포괄하기 때문이다. 그 과정에서 모두 실연자의 성명을 표시하거나, 표시하지 않는 것은 결국 음성합성 결과물을 이용해서 콘텐츠를 창작하는 제3의 창작자의 몫이 될 것이다. 이때 실연자의 목소리와 동일성이 감지되는 결과물이 만들어져 이용되면서 성명표시가 누락되거나 다른 사람의 이름이 표시된다면 성명표시권 침해 문제가 발생할 수 있을 것이다.

실연자의 인격권 침해의 중점은 실연자의 명예를 훼손하는 데 있다고 할 수 있는데, 음성합성 시스템이 실연자의 목소리와 동일성이 감지되는 결과물을 만들어냈다는 사실만으로는 실연자의 인격권을 침해하였다고 보기는 어려울 것이다. 그러나 누구나 텍스트만 입력하면 음성합성 시스템을 통해 쉽게 음성을 만들어낼 수 있으므로 일부 고객이 부적절한 음성을 제작하고 유포함으로써 실연자의 명예를 침해하는 등 실연자의 인격권을 침해하는 사례가 발생할 수 있다. 이러한 경우에는 저작권법 제123조에 따른 침해금지청구와 제125조에 따른 손해배상청구를 할 수 있다.

그런데 계약서의 부속합의서 제3조 제2항은 다음과 같은 불행사특약을 포함하여 상기와 같은 권리 침해 문제를 해결하고자 하였다.

성우 및 ‘녹음스튜디오’는 제2조 제5항에 따른 보이스 데이터 또는 결과물의 활용에 대해 저작권인격권, 퍼블리시티권, 성명권, 초상권 등을 포함한 어떠한 권리도 주장하지 않을 것에 동의한다.²⁶⁾

성우가 저작권인격권(실연자의 인격권)을 포함한 어떠한 권리도 주장하지 않을 것에 동의한다는 것이다. 만일 불행사특약이 유효하다면 실연자

26) 스튜디오의 명칭을 ‘녹음스튜디오’로 치환하였다. 저작권법 제66조 및 제67조에 의해 실연자에게 부여된 인격권은 ‘실연자의 인격권’이나 본 계약서는 이를 ‘저작권인격권’으로 잘못 표기한 것으로 보인다.

는 실연자의 인격권이 침해되는 우려에 대해 예방조치를 청구할 수 없을 것이다. 그런데 저작권법상 실연자의 인격권은 저작인격권과 마찬가지로 실연자의 일신에 전속하는 것(저작권법 제68조)으로서 원칙적으로 불행사 또는 포기 특약은 무효이다.²⁷⁾ 학설에서는 인격권의 불행사특약을 유효한 것으로 보아야 한다는 주장이 있으나, 이 경우에도 구체적인 범위를 약정한 제한적인 범위 내에서만 불행사특약이 유효하다는 조건을 언급하고 있다.²⁸⁾ 본 계약서는 ‘결과물의 활용’에 대해 저작인격권(실연자의 인격권) 불행사를 약정하였는데, ‘활용’의 범위가 상업적·비상업적 목적을 불문하고 어떠한 목적으로든 서비스 제공에 필요한 형태로 사용하는 일체의 행위와 서비스를 이용하는 고객과 고객이 서비스를 이용한 결과물을 제3자가 이용하는 것까지 포괄하므로 사실상 범위의 제한이 없다고 보아야 할 것이다. 그러므로 이 글에서는 해당 특약을 무효로 보아야 한다고 생각한다.

한편 실연자의 인격권 침해 문제를 검토함에 있어서는 실연자의 저작인접권의 보호범위에 관하여 살펴볼 필요가 있다. 서울고등법원은 “실연자가 법 제63조 등에 의하여 가지는 저작인접권은 실연자가 특정 시점에서 실제로 행한 실연 그 자체를 녹음·녹화 또는 사진촬영하는 등 복제할 수 있는 권리일 뿐, 그 실연과 유사한 다른 실연에 대하여는 권리

27) “동일성유지권의 양도나 포기가 무효라면, 이 권리를 행사하지 않는다는 합의는 유효인가 하는 의문점이 생긴다. 실무상 동일성유지권을 행사하지 않기로 합의하는 경우가 있을 수 있기 때문이다. 일반적으로 인격권을 행사하지 않기로 하는 취지의 합의는 인격을 부정하는 것이 된다는 점에서 포기의 경우와 마찬가지로 무효라고 보아야 할 것이다.”

박성호, 앞의 책, 305.

“저작인격권은 인격권으로서의 성질을 가지기 때문에 포기할 수 없으며 따라서 원칙적으로 저작인격권 포기 합의는 무효라고 해석하여야 한다. 이와 같이 ‘인격권의 포기’는 인격권의 본질에 부합하지 않는다는 인식 때문에 실무적으로는 저작인격권의 ‘불행사 특약’ 또는 ‘개변의 동의’ 등의 용어가 많이 사용되고 있으나 그 실체는 크게 다르지 않은 것 같다.”

오승중 (2020). <저작권법>(제5판). 서울: 박영사, 490.

28) 이와 관련하여 김근우 (2016). 실연자의 범위와 실연자의 인격권에 관한 고찰. <경희법학>, 제48권, 163. ; 이해완 (2019). <저작권법>(제4판). 서울: 박영사, 527.

가 미치지 않는다”고 판시하였다.²⁹⁾ 이를 인공지능 음성합성 시스템이 만들어낸 음성합성물에 그대로 적용한다면, 실연자는 인공지능 음성합성 시스템이 만들어내는 음성합성물이 실연자의 명예를 훼손한다고 할지라도 음성합성물은 실연자가 행한 실연 그 자체가 아니므로 그에 대해 저작권법상 실연자의 인격권 침해를 주장할 수 없을 것이다. 그러나 인공지능 음성합성 시스템이 만들어낸 음성합성물이 여타의 일반적인 ‘실연과 유사한 다른 실연’과 동일시될 수 있을지는 의문이다. 일반적으로 인간에 의해 모방된 실연은 대체로 원 실연을 모방한 실연이라는 점을 제3자가 인지할 수 있으며 제 아무리 원 실연과 유사하다고 할지라도 원 실연을 대체할 수 있을 정도의 시장 지위를 가지지 못하지만, 인공지능 음성합성 시스템의 음성합성물은 실제 실연자가 발화한 것과 제3자로서는 구분하기 어려우며 신속성·경제성·정확성 등 여러 가지 면에서 실연자를 압도하는 시장 경쟁력을 갖추고 있기 때문이다. 그러한 점에서 인공지능 음성합성 시스템이 만들어낸 음성합성물도 실연자의 저작권접권의 보호범위에 포함되어야 한다고 생각한다.

3. 실연자의 일반적 인격권에 의한 보호

‘실연자의 저작권접권은 실연 그 자체에 국한된다’는 견해를 적용한다면, 인공지능 음성합성 시스템이 만들어낸 음성합성물에 대해 실연자는 저작권접권을 주장할 수 없게 될 수도 있다. 그러한 경우에는 실연자의 인격권이 아닌 인간이라면 누구에게나 보장되는 일반적 인격권에 의해 실연자의 보호를 도모할 수 있을 것이다.

대법원은 “저작권법은 공표권(제11조), 성명표시권(제12조), 동일성유지권(제13조) 등의 저작권접권을 특별히 규정하고 있으나, 작가가 자신의 저작물에 대해서 가지는 인격적 이익에 대한 권리가 위와 같은 저작

²⁹⁾ 서울고등법원 2007. 5. 22. 선고 2006나47785 판결

권법 규정에 해당하는 경우로만 한정된다고 할 수는 없으므로 저작물의 단순한 변경을 넘어서 폐기 행위로 인하여 저작자의 인격적 법익 침해가 발생한 경우에는 위와 같은 동일성유지권 침해의 성립 여부와는 별개로 저작자의 일반적 인격권을 침해한 위법한 행위가 될 수 있다”고 판시한 바 있다.³⁰⁾ 예컨대 인공지능 음성합성 시스템의 이용자가 음란한 내용이나 폭력적인 내용 등을 음성합성물로 제작하여 유포할 경우 제3자로서는 실연자의 음성으로 접하게 되므로 실연자는 일반적 인격권이 침해되었다고 주장하여 보호를 받을 수 있을 것이다. 이는 민법상 불법행위에 해당하며 부적절한 음성합성물을 제작하고 유포한 이용자뿐만 아니라, 이를 방조한 인공지능 음성합성 서비스 제공자에게도 민법 제 760조 제3항에 따라 방조에 의한 공동불법행위자로 책임을 물을 수 있을 것이다.

4. 부정경쟁방지법에 의한 보호

인공지능 음성합성 시스템이 만들어낸 음성합성물을 광고 등의 상업적 목적으로 사용하는 경우 이것이 이른바 퍼블리시티의 부정사용으로, 부정경쟁방지법 제2조 제1호 타목 내지는 파목에 해당하는 부정경쟁행위에 해당하는지 여부를 검토해 볼 수 있을 것이다.³¹⁾ 부정경쟁방지법 제2조 제1호 타목은 ‘공정한 상거래 관행이나 경쟁질서에 반하는 방법으로 자신의 영업을 위하여 무단으로 사용함’을 부정경쟁행위의 요건으로

30) 대법원 2015. 8. 27. 선고 2012다204587 판결

31) 부정경쟁방지법 제2조 제1호 타목은 “국내에 널리 인식되고 경제적 가치를 가진” 것을 요건으로 하였으므로 이 글에서 다루는 인공지능 음성합성 시스템이 만들어낸 음성합성물을 광고 등의 상업적 목적에 사용하는 경우 실연자의 유명세에 따라 요건을 충족하지 못하는 경우가 있겠으나, 국내에 널리 인식되지 못하는 실연자와 인공지능 음성합성 시스템 사업자 사이의 계약이라 할지라도 인공지능 음성합성 시스템의 학습용 데이터는 실연자의 상당한 노력으로 만들어지므로 부정경쟁방지법의 보충적 일반조항인 제2조 제1호 파목의 요건을 충족 여부에 대해서는 검토해 볼 수 있을 것이다.

하는데, 요건의 구체적인 법리에 관하여 방탄소년단 멤버의 사진을 대량으로 수록한 부록과 포토카드 등을 제작하여 판매하는 행위에 대한 대법원의 결정이 이루어진 바 있다.³²⁾ 대법원은 요건에 해당하기 위해서는 “(①)권리자와 침해자가 경쟁 관계에 있거나 가까운 장래에 경쟁관계에 놓일 가능성이 있는지, (②)권리자가 주장하는 성과 등이 포함된 산업분야의 상거래 관행이나 경쟁질서의 내용과 그 내용이 공정한지 여부, (③)위와 같은 성과 등이 침해자의 상품이나 서비스에 의해 시장에서 대체될 가능성, (④)수요자나 거래자들에게 성과 등이 어느 정도 알려졌는지, (⑤)수요자나 거래자들의 혼동가능성 등을 종합적으로 고려해야 한다”고 판시하였다.³³⁾ 표면적으로 보면, 인공지능 음성합성 시스템의 학습용 데이터에 대한 인공지능 음성합성 시스템 사업자와 실연자 사이의 계약은 대법원이 언급한 5가지 요건 중 적어도 4가지 요건에 부합한다고 볼 수 있으며, 경우에 따라서는 5가지 요건을 모두 충족한다고 볼 수도 있을 것이다. (①)인공지능 음성합성 시스템이 대중을 상대로 서비스를 시작하면 실연자는 즉각적으로 시장에서 경쟁관계에 놓이게 되며, (②)방송 실연에 대해서는 실연자가 저작권접권 사용료를 주기적으로 분배받고 있는데 반하여 인공지능 음성합성 시스템에 대해서는 방송 실연의 보수보다도 보수의 수준이 적음에도 불구하고 인공지능 음성합성 시스템 사업자는 실연자로부터 일시불에 의해 권리 일체를 양도받고자 하였으며, (③)인공지능 음성합성 시스템은 실연자의 음성을 모방한 것이므로 시장에서 실연자를 완벽하게 대체할 수 있으며, (⑤)수요자나 거래

32) 대법원 2020. 3. 26. 선고 2019마6525 결정.

33) 대법원의 결정은 현 부정경쟁방지법 제2조 제1호 타목의 입법이 이루어지기 이전에 보충적 일반규정인 “그 밖에 타인의 상당한 투자나 노력으로 만들어진 성과 등을 공정한 상거래 관행이나 경쟁질서에 반하는 방법으로 자신의 영업을 위하여 무단으로 사용함으로써 타인의 경제적 이익을 침해하는 행위(현 부정경쟁방지법 제2조 제1호 파목)”에 기반하여 이루어진 것이지만 ‘공정한 상거래 관행이나 경쟁질서에 반하는 방법으로 자신의 영업을 위하여 무단으로 사용함’ 요건은 현 부정경쟁방지법 제2조 제1호 타목과 파목에 공통되게 존재하므로 대법원이 판시한 법리를 제2조 제1호 타목의 법리를 판단하는 데 있어서 참고할 수 있다.

자가 인공지능 음성합성 시스템이 합성해낸 음성과 실연자의 음성을 구별하기 어렵다. 다만 ④수요자나 거래자들에게 성과 등이 얼마나 알려졌는지 요건은 실연자의 대중적 인지도의 수준에 따라 충족되지 않을 수도 있다.

그러나 위 요건의 충족 여부에도 불구하고, 실연자의 음성 데이터를 학습한 인공지능이 실연자와 음성합성 시스템 사업자 간의 계약에 의해 음성 데이터를 생성한 것이라면 그 계약이 무효로 판명되지 않는 한 계약조건이 아무리 불합리한 것이라 하더라도 부정경쟁방지법상의 부정경쟁행위라고 주장하기는 어려울 것이다. 부정경쟁행위에 해당하기 위해서는 행위가 ‘무단으로’ 이루어져야하기 때문이다. 따라서 대법원이 판시한 위 다섯 가지 요건이 부정경쟁방지법 제2조 제2호 타목의 요건에 해당하는지 여부를 종합적으로 판단하기 위해서는 그에 앞서 인공지능 음성합성 시스템의 학습용 데이터에 대한 인공지능 음성합성 시스템 사업자와 실연자 사이의 계약이 불공정거래행위로서 무효라는 점이 전제되어야 할 것이다.

V. 인공지능 생성물에 의한 실연자의 기술적 실업의 문제

1. 권리침해에 대한 현행법상 보호의 한계

앞서 인공지능 음성합성 시스템의 학습용 데이터 이용허락 계약이 실연자의 인격권 또는 실연자의 일반적 인격권 침해에 해당할 수 있으며, 계약이 불공정행위로서 무효라면 부정경쟁방지법상 부정경쟁행위에도 해당한다는 것을 밝혔다. 그런데 설령 저작권법상 실연자의 인격권 내지는 실연자의 일반적 인격권 침해가 성립하고 또한 인공지능 음성합성 시스템 사업자와 실연자 간의 계약이 무효로 판명되어 실연자의 실연이

부정경쟁방지법에 의한 보호대상이 된다고 가정하더라도 실연자는 이미 발생한 권리침해에 대해서 손해배상청구를 하거나 금지청구를 할 수 있을 뿐, 소정의 일시불로 권리를 영구적으로 이용을 허락받는 불공정한 계약관행은 변하지 않을 것이다. 관행에 특별한 이변이 없다면 인공지능 기술 발전의 추이에 따라 성우는 인공지능 생성물로 인한 기술적 실업 현상을 전격적으로 맞이하는 첫 번째 사례가 될 것으로 보인다. 앞서 계약의 개별 권리 침해 사안에 대하여 현행법상 보호방안을 모색하였지만, 실연자 모두가 직면한 보편적인 문제는 인공지능 생성물에 의한 실연자의 기술적 실업 현상이다.

그러므로 근본적으로 문제를 해결하기 위해서는 처음 계약 체결단계에서부터 모든 실연자가 자신의 권리가 제대로 보호받을 수 있도록 계약 내용을 면밀히 살피는 노력을 해야 할 것이다. 즉 실연자는 인공지능 음성합성 시스템에 대한 학습용 음성데이터 계약 체결과정에서 계약 내용에 이의를 제기해야 할 것이다. 개별 실연자가 사업자를 대상으로 이의를 제기하는 것은 어려운 일이지만 자신의 권리를 스스로 지킬 수 있도록 노력해야 한다. 계약체결 과정에서 권리를 주장하다 보면 때로는 해당 작업에서 배제될 위험을 감수해야 하겠지만, 모든 실연자가 단결하여 그러한 위험을 감수하는 과정에서 실연자 모두의 권리를 지킬 수 있는 단체교섭력이 배양되기 마련인 것이다. 물론 계약 지위상 차이로 인해 개개 실연자의 차원에서 실천적인 해결방안을 모색하는 것이 어려운 일임에 틀림없지만 누군가가 아니라 나부터 희생을 각오하고 앞장서야 한다. 그것이 바로 ‘권리를 위한 투쟁의 길’이다. 이와 관련하여 이 글에서는 노동조합과 신탁관리단체를 통한 교섭 및 저작권접권 사용료 징수·분배의 성공적인 과거사례를 살펴보고 이러한 예전의 성공사례가 인공지능 음성합성 시스템의 학습용 데이터에 대한 실연자의 권리를 위해서도 활용 가능한 대안이 될 수 있는지 그 가능성을 살펴보고자 한다.

2. 실연자 단체에 의한 보호 사례

전 세계적으로 시청각실연자가 스스로의 권리를 보호하는 형태는 한국, 미국, 영국, 프랑스, 일본 등 각국의 법제도의 차이에도 불구하고 대부분 비슷하다.³⁴⁾ 요컨대 시청각실연자가 스스로 단체를 구성하여 단결함으로써 교섭력을 갖추어 제작자를 상대로 실연의 이차적인 이용에 관하여 사용료 징수 계약을 체결하는 것이다. 그중에서 주목해볼만한 사례가 미국의 영화배우조합과 텔레비전 라디오 예술인조합(Screen Actors Guild - American Federation Television & Radio Artists, 이하 SAG-AFTRA)이다.³⁵⁾ 미국 저작권법에는 저작인접권이라는 개념이 없어서 시청각실연자의 권리를 보호하는 저작권법적 근거가 매우 취약하지만, SAG-AFTRA가 제작자를 상대로 계약을 체결함으로써 시청각실연자의 권리를 보호하는 수준이 매우 높다. SAG-AFTRA는 제작자와 계약을 통해 연기자가 제작사와 일을 하게 되면 자동적으로 노동조합에 가입되도록 제도를 구축한 한편, 노동조합의 규정으로 “어떤 노조원도 최소한의 기본조건을 완전하게 집행하지 않는 제작자들을 위하여 배우로 노동계약을 체결할 수 없고 배우로서 일할 수 없다”고 명시하여 노동조합의 단결력을 극대화하였고 상당한 수준의 교섭력을 갖추게 되었다.³⁶⁾ 그로인해 미국의 시청각실연자는 SAG-AFTRA의 단체 협약을 기반으로 근로조건을 보장받고, 이차적인 이용에 대한 대가(residuals)를 지급받으며, 심지어는 제작

34) 김기복 (2020, 12, 23). 방송연기자노동조합과 방송실연자권리협회는 한 뿌리. <한국방송연기자노동조합신문>, 5면.

35) 영화배우조합(Screen Actors Guild)과 텔레비전 라디오 예술인조합(American Federation Television & Radio Artists)은 2012년 3월 30일부로 하나의 조합(SAG-AFTRA)으로 합병하였다.
SAG-AFTRA, Our History 2010s. URL: <https://www.sagaftra.org/about/our-history/2010s>

36) 김인철 (2012). 영상저작물에서 미국실연자들의 권리 보호. <가천법학>, 제5권 제1호, 458-459 참조.
SAG-AFTRA의 규정에 관하여는 각주 38번에 후술한다.

자로부터 자금을 수령하여 조성한 기금으로부터 연금과 건강보험 혜택을 받고 있다.

우리나라는 근로조건에 관하여 한국방송연기자노동조합에서 방송사와 단체 교섭을 진행해오고 있으며, 저작권접권의 신탁관리에 관하여서는 한국방송실연자권리협회에서 방송사와 특약을 체결하여 방송프로그램의 이차적인 사용에 대한 사용료를 징수·분배하고 있다.³⁷⁾ 그런데 우리나라의 저작권법은 영상저작물에 대한 실연자의 권리를 특약이 없는 한 영상제작자에게 양도된 것으로 추정하고 특약체결을 강제하는 제도를 마련하지 않았다. 그로인해 저작권법상 시청각실연자의 권리 보호수준이 매우 미약하다. 오늘날 시청각실연자는 오로지 한국방송실연자권리협회의 회원으로서 협회가 자체적인 협상력을 기반으로 방송사와 체결한 특약에 의해 저작권접권 사용료를 지급받고 있는데, 협회의 협상 요구를 영상제작자가 거부하더라도 현행법상 대응할 방법이 없기 때문에 미디어환경 변화에 따라 등장하는 새로운 저작권접권 이용형태에 대하여 협회가 특약을 체결하는 데 많은 어려움을 겪고 있다.

문제는 인공지능 음성합성 시스템이 만들어낸 음성합성물의 이용으로 발생하는 부가가치에 대해 주기적으로 분배를 요구하는 경우에도 위와

37) 1994년 1월 7일 저작권법 일부개정 이전에는 영상저작물 특례 규정에 ‘특약이 없는 한’이라는 문구가 존재하지 않았기 때문에 시청각실연자가 저작권접권을 주장할 수 있는 여지가 없었다. 1994년 1월 7일 저작권법 일부개정으로 구 저작권법 제75조 제3항(현 제100조 제3항)에 “특약이 없는 한 권리가 양도된 것으로 본다”는 조문의 입법이 이루어졌으나, 방송사의 반발로 해당 조문은 5년간의 유예를 거쳐 1999년 1월 7일부터 효력을 가지게 되었다. 당시 한국방송연기자노동조합은 조합원으로부터 ‘실연자 권리 위임장’을 수령하여 방송사업자에게 실연자의 저작권접권에 대한 특약을 요구하였으나, 방송사는 노동조합이 저작권 신탁관리업 허가를 받지 못했다는 이유로 특약 체결을 거부하였다. 이후 노동조합은 저작권 신탁관리업 허가를 받고자 하였으나 당시 문화관광부는 노동법에 의해 설립한 노동조합에 저작권법에 기초한 저작권 신탁관리업무를 허가할 수 없다고 하였다. 이에 노동조합에서는 2001년 제3의 단체인 한국방송실연자협회를 설립하였으며, 2002년 한국방송실연자협회는 문화관광부로부터 저작권 신탁관리업 허가를 받아 정식으로 출범하게 되었다.

사단법인 한국방송실연자협회 (2011). <社團法人 韓國放送實演者協會 10年史>. 서울: 사단법인 한국방송실연자협회, 21-26 요약.

같은 성공사례를 활용할 수 있을 것인가 하는 점이다. SAG-AFTRA의 성공사례를 모방하여 한국방송연기자노동조합의 교섭력과 한국방송실연자권리협회의 신탁관리 역량을 기반으로 하는 방안을 생각해볼 수 있겠으나, 이 방안이 설득력을 갖추기 위해서는 사업자가 노동조합의 요구를 도저히 거부할 수 없을 정도로 노동조합이 강력한 교섭력을 갖추어야 한다.³⁸⁾ 노동조합이 충분한 교섭력을 갖추어 사업자로 하여금 한국방송실연자권리협회와 부가가치의 주기적인 분배에 관한 계약을 체결한다면 한국방송실연자권리협회는 기존의 신탁관리역량을 통해 최소한의 비용으로 부가가치를 징수하고 이를 개개 실연자에게 분배할 수 있을 것이다.³⁹⁾

38) 노동조합이 강력한 교섭력을 갖추기 위해 일차적으로는 실연자 사이에서 ‘모두의 권리를 위해 개개인이 노동조합을 중심으로 강하게 단결해야한다’는 인식이 뿌리 깊게 자리잡아야하며, 노동조합 차원에서 더 많은 사람들이 조합에 가입하고, 이탈하지 않도록 제도를 마련해야한다. 예컨대 SAG-AFTRA는 조합원에게 건강보험, 퇴직연금, 재사용료(residual), 캐스팅 워크숍 등 충분한 혜택을 제공하는 한편, “조합원은 전세계 어디에서도 SAG-AFTRA와 계약을 체결하지 않은 고용주를 위해 일할 수 없다(No member shall render any services or make an agreement to perform services for any employer who has not executed a basic minimum agreement with the union, which is in full force and effect, in any jurisdiction in which there is a SAG-AFTRA national collective bargaining agreement in place. This provision applies worldwide.)”라는 국제규정 1(Global Rule One)을 명시하고 있다. 국제규정 1을 위반한 조합원은 견책(reprimands)에서 벌금(fines), 제명(expulsion)까지 이르는 징계를 받게 된다.

SAG-AFTRA, Global Rule One. URL: <https://www.sagaftra.org/contracts-industry-resources/global-rule-one>

39) 다만 한국방송실연자권리협회의 신탁관리범위가 ‘방송실연자의 권리’로 한정되어 있기 때문에 신탁관리범위의 조정이 이루어지지 않는 한 협회는 회원이 인공지능 음성합성 시스템의 학습용 데이터에 대하여 가지는 권리를 신탁할 수 없다. 문화체육관광부에 의해 한국방송실연자권리협회의 신탁관리범위는 ‘방송실연자의 권리’로 한정되어 있고 저작권법상 방송은 ‘수신의 동시성’과 ‘일방향성’을 토대로 정의되어 있기 때문이다. 신탁관리의 범위를 정의함에 있어서 저작물을 송신하는 방법인 저작권법상 ‘방송’ 개념을 사용하는 것이 적절한 것인지 의문이 있으나, 축자적으로 해석한다면 인공지능 음성합성 시스템의 학습용 데이터는 저작권법상 방송의 범주에 포함되지 못하므로 한국방송실연자권리협회가 인공지능 음성합성 시스템의 학습용 데이터를 신탁관리함에 제한이 있다. 방송 분야 저작권 신탁관리업의 신탁관리범위의 적절성에 관하여 조병한 (2021). 방송콘텐츠 개념의 제도적 정립을 위한 시론. <한국콘텐츠학회 논문지>, 제21권 제9호, 479 참조.

3. 입법에 의한 보호의 필요성

음성 데이터를 제공한 실연자로서의 성우들이 인공지능이 생성한 음성합성물의 부가가치에 대해 분배를 요구하는 것은 기술발전으로 인해 새롭게 등장한 문제이기 때문에 사안에 국한하여 간단히 답하기 어렵다.

이 문제는 실연자의 권리에 국한한 것으로서, 문제를 해결하기 위하여 이 글에서 다룬 것처럼 저작권법상 실연자의 인격권 내지는 실연자의 일반적 인격권과 부정경쟁방지법에 의한 보호, 나아가서 실연자의 강력한 단결력에서 비롯하는 한국방송연기자노동조합의 단체교섭력을 통한 권리 보호방안을 생각해볼 수 있다. 그러나 이것은 수면에 드러난 빙산의 일각이며 앞으로 인공지능 기술과 인공지능의 생성물이 다른 분야에서도 상용화된다면 유사한 문제가 슬하게 야기될 것이다.

그러므로 궁극적으로는 인공지능 음성합성 시스템의 학습용 데이터에 대한 실연자의 권리 침해 문제가 인류에게 다가올 인공지능 생성물로 인한 기술적 실업 현상의 단초라는 점에 착목하여 일국(一國) 차원의 해결이 아닌 국제규범 차원의 해결방안을 모색해야 할 것이다. 지금부터 100여년 전 상업용 음반의 등장으로 가수나 연주자들이 직면하였던 기술적 실업을 타개하기 위해 저작권접권이라는 새로운 권리를 창출하였던 역사적 사실을 교훈 삼아 인공지능 생성물의 등장으로 인한 기술적 실업문제를 해결하기 위해 새로운 방안을 마련해야 할 시점이다.⁴⁰⁾

40) “녹음 기술 등의 발달에 따라 음악업계에 종사하는 노동자들에게 ‘기술적 실업’이 발생함으로써 노동의 권리가 침해되는 경우가 발생하는데, 대표적인 것이 판매용 음반의 방송사용으로 인하여 가수나 연주자들이 각종 음악회에서 ‘라이브’로 노래를 부르고 연주하거나 또는 방송에 출연할 기회를 상실하게 되어 발생한 실업 문제이다. 이에 대하여 국제노동기구(ILO)는 가수나 연주자에 대한 보상을 제창하였고, ‘실연자·음반제작자 및 방송사업자의 보호를 위한 국제협약’(1961년, 일명 ‘로마협약’)에서는 이를 반영하여 제12조에 방송사업자가 실연자인 가수나 연주가, 음반제작자 등에게 보상하도록 하는 규정을 마련하였다.” 박성호 (2014). 음악저작권의 생성과 발전에 관한 역사적 고찰. <법학논총>, 제28권 제3호, 51-52.

VI. 결론

메아리의 영문 에코(Echo)는 그리스 신화에 등장하는 요정(Nymph)의 이름이다. 에코는 목축의 신 판(Pan)의 사랑을 받아주지 않았다는 이유로 판이 보낸 양치기들에 의해 몸이 찢겨 들판에 뿌려졌는데, 이를 본 가이아가 그녀의 시신을 보존하여 목소리만 남게 되었다.⁴¹⁾ 오늘날 인공지능 음성합성 시스템의 불합리한 계약조건으로 인해 실연자는 에코가 될 운명 앞에 놓였다. 판의 만행은 불공정한 계약 조항에 비견된다. 메아리는 콘텐츠산업 생태계의 선순환 구조를 서서히 망가뜨리며 공허하게 울릴 것이다.

인공지능 음성합성 시스템은 실연자가 제공한 학습용 음성데이터를 기반으로 발음패턴, 발성법, 억양 등 목소리를 내는 방식을 복제하였다. 이는 개인을 식별하는 고유한 표지에 해당하므로, 인공지능 음성합성 시스템의 텍스트·데이터마이닝은 실연자의 지능의 일부를 디지털화하는 것으로 해석할 수 있었다.

그러나 인공지능 음성합성 시스템의 학습용 음성데이터는 부적절한 계약에 의해 제공되고 있었다. 이용자는 통상의 방송 녹음에 대한 보수 수준에도 미치지 못하는 일시불을 대가로 자신과 자신의 고객의 음성합성 결과를 제3자가 이용하는 것까지 무제한적인 이용허락 계약을 요구하였고, 인격권마저 행사할 수 없도록 제한하였다.

이에 대하여 이 글에서는 불공정한 계약 관행이 실연자 개인의 권익을 침해하는 것을 넘어 산업생태계의 선순환 구조를 해치는 데 이르기 때문에 실연자의 인격권을 강하게 보호하는 합목적적 해석이 필요함을

41) 에코에 관한 이야기는 오비디우스(Ovid)의 『변신이야기』에 실린 이야기와 롱구스(Longus)의 『다프니스와 클로에』에 실린 이야기가 서로 다르다. 널리 알려진 나르키소스와 에코 이야기는 오비디우스에 의한 것이며, 이 글이 인용한 이야기는 롱구스에 의한 것이다.

Longus(unknown). *Δάφνις και Χλόη*. George. Thornley (trans.) (1916). *Daphnis&Chloe*. London: William Heinemann, 161-163.

밝혔다. 계약에서는 재산권을 무제한적으로 이용허락하고 인격권 불행사 특약을 명시하였으나, 이 글은 계약에 포함된 인격권 불행사특약이 범위에 제한이 없는 무제한적 이용허락을 전제로 하기 때문에 무효라고 보았다. 그러므로 실연자의 인격권이 침해된다면 실연자는 저작권법상 실연자의 인격권 내지는 실연자의 일반적 인격권으로 손해배상청구를 하거나 금지청구를 할 수 있다. 더불어 해당 계약이 불공정계약으로 무효라면 부정경쟁방지법에 의한 보호대상이 될 수도 있다.

다만 인격권에 의한 실연자의 보호가 이루어진다고 할지라도 개별 침해사안에 대하여 향후 손해배상청구나 금지청구를 할 수 있을 뿐, 소정의 일시불로 권리 일체를 무제한적으로 이용허락하도록 강요하는 계약 관행은 변화하지 않을 것이다. 따라서 공정한 이용허락계약을 체결하기 위해서는 실연자가 계약과정에서 상기의 논리를 기반으로 권리를 쟁취해야 한다. 실연자 개인의 차원에서 이러한 권리를 주장하다가 이후의 작업에서 배제될 우려가 있으나, 모든 실연자가 그러한 우려를 감수하고 권리를 주장할 때에서야 모두의 권리를 지킬 수 있을 것이다.

이에 이 글에서는 강력한 단결력으로 스스로 권리를 보호하는 미국 SAG-AFTRA의 사례를 기초로, 한국방송연기자노동조합이 단체교섭력을 배양하여 노동조합과 신탁관리단체의 연합을 통해 권리를 보호하는 방안을 제시하였다. 다만 궁극적으로는 인공지능 생성물로 인한 기술적 실업 현상에 대한 일국 차원의 대응이 아닌, 국제규범 차원의 해결방안을 모색해야 할 것이다.

요컨대 인간은 기술을 만들지만, 기술은 다시 인간의 삶을 변화시킨다. 인간이 만든 인공지능은 인간의 노동력을 대체하여, 우리의 일자리를 대신 차지할 것이다. 그러나 그것이 꼭 부정적인 것만은 아니다. 경제적 자유가 전제된다면 일자리의 상실이란 노동으로부터의 해방을 의미한다. 인공지능의 창작활동에서 발생한 재화가 인공지능으로 인해 일자리를 잃은 사람들에게 분배될 수 있다면, 인간은 더 이상 생계를 위한 노동에 시간을 빼앗기지 않은 채로 인공지능이 만든 창작물에 대한 권리를 향유

하며 스스로의 삶을 조금 더 풍요롭게 만들어나갈 수 있을 것이다. 결국 노동은 기계가 하게 될 것이며, 사람들은 내재적인 가치를 창출하고 공유된 사회 공동체 의식을 재활성화하기 위해 해방되어야 한다.⁴²⁾

⁴²⁾ Jeremy Rifkin (1996). *The End of Work*. 이영호 (역) (2005). <노동의 종말>(제2판), 서울: 민음사, 45.

■ 참고 문헌

- 김경호·성혜림·성미경·정찬경·성진택 (2020년 8월). <타코트론 모델을 이용한 개인 맞춤형 한국어 음성합성 시스템>. 2020년 대한전자공학회 하계학술대회, 제주: 롯데호텔, 2089-2091.
- 김기복 (2014). 시청각 실연자(視聽覺 實演者)의 권리보호 방안. <한국콘텐츠학회논문지>, 제14권 제12호.
- _____ (2020, 12, 23). 방송연기자노동조합과 방송실연자권리협회는 한 뿌리. <한국방송연기자노동조합신문>, 5면.
- 김근우 (2016). 실연자의 범위와 실연자의 인격권에 관한 고찰. <경희법학>, 제48권, 151-168.
- 김영재 (2014). 디지털 미디어 환경의 콘텐츠 산업 진흥기금. <한국콘텐츠학회 논문지>, 제14권 제2호, 146-160.
- 김인경 (2020, 2, 10). 녹음 없이 동영상 더빙을? 네이버, ‘클로바더빙’ 출시. <BLOTTER>. URL: <https://www.bloter.net/newsView/blt20200210011>
- 김인철 (2012). 영상저작물에서 미국실연자들의 권리 보호. <가천법학>, 제5권 제1호, 439-472.
- 노현숙·구천을 (2016). 저작재산권 양도와 저작물 이용허락의 비교 고찰. <江原法學>, 제48권, 243-273.
- 류준호·윤승금 (2010). 생태계 관점에서의 문화콘텐츠 산업 구성 및 구조. <한국콘텐츠학회 논문지>, 제10권 제4호, 327-339.
- 박성호 (2014). 음악저작권의 생성과 발전에 관한 역사적 고찰. <법학논총>, 제28권 제3호, 33-56.
- _____ (2017). <저작권법>(제2판). 서울: 박영사.
- _____ (2020). 텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작재산권의 제한. <인권과 정의>, 제494호, 39-69.
- 사단법인 한국방송실연자협회 (2011). <社團法人 韓國放送實演者協會 10年史>. 서울: 사단법인 한국방송실연자협회.
- 안효질 (2018). 영상저작물의 이용활성화와 실연자의 보상청구권. <법학논총>, 제42권 제3호.
- 엔씨소프트. AI 김영하, 읽다. URL: <https://blog.ncsoft.com/ai-tech-young>

a-kim/

- 오승종 (2019). 데이터마이닝 및 텍스트마이닝과 저작권산권의 제한. <홍익 법학>, 제20권 제2호, 455-487.
- _____ (2020). <저작권법>(제5판). 서울: 박영사.
- 윤정운 (2020). 저작권법상 공정이용에 관한 소고: 인공지능의 빅데이터 활용을 포함한 다양한 영역에서의 사례를 중심으로. <Law&Technology>, 제16권 제5호, 3-28.
- 이동기 (2016). 시청각 실연자의 권리와 그에 관한 베이징 조약 연구. <圓光法學>, 제32권 제3호.
- 이은숙 (2021). <2021 대중문화예술평산업 실태조사>. 나주: 한국콘텐츠진흥원.
- 이해완 (2019). <저작권법>(제4판). 서울: 박영사.
- 신창환 (2019). 시청각 실연에 관한 국제 보호 규범과 우리나라에서의 발전 방향 연구. <계간 저작권>, 제32권 제3호.
- 정상조 (2020). 딥러닝에서의 학습데이터와 공정이용. <Law&Technology>, 제16권 제1호, 3-35.
- 조수희 (2019). <딥러닝 기반 보코더의 음성 합성에 대한 음성학적 관찰>. 고려대학교 대학원 석사학위논문.
- 조병환 (2021). 방송콘텐츠 개념의 제도적 정립을 위한 시론. <한국콘텐츠 학회 논문지>, 제21권 제9호, 477-492.
- 차상욱 (2021). 저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점: 텍스트·데이터마이닝(TDM) 면책규정을 중심으로. <經營法律>, 제32권 제1호, 1-55.
- 한동현 (2021, 12, 6.). 성우가 목소리 노예?...IT기업, 음성 저작권 영구 양도 요구. <서울와이어>. URL: <https://www.seoulwire.com/news/articleView.html?idxno=456552>
- 한지영 (2016). 시청각실연자의 법적 보호에 관한 고찰. <계간 저작권>, 제29권 제4호.
- 대법원 2020. 3. 26. 선고 2019마6525 결정.
- 대법원 2018. 10. 12. 선고 2015두38092 판결.
- 대법원 2015. 8. 27. 선고 2012다204587 판결.
- 서울고등법원 2007. 5. 22. 선고 2006나47785 판결.

- Jeremy Rifkin (1996). *The End of Work*. 이영호 (역) (2005). <노동의 종말>(제2판), 서울: 민음사.
- Longus(unknown). *Δάφνις καὶ Χλόη*. George. Thornley (trans.) (1916). *Daphnis&Chloe*. London: William Heinemann.
- SAG-AFTRA. Global Rule One. URL: <https://www.sagaftra.org/contracts-industry-resources/global-rule-one>
- SAG-AFTRA. Our History 2010s. URL: <https://www.sagaftra.org/about/our-history/2010s>
- SungWoo Moon & SungHyun Kim & Yong-Hoon Choi (2022). MIST-Tacotron: End-to-End Emotional Speech Synthesis Using Mel-Spectrogram Image Style Transfer. *IEEE Access*, 10, 25455-25463.
- Yuxuan Wang & RJ Skerry-Ryan & Daisy Stanton & Yonghui Wu & Ron J. Weiss & Navdeep Jaitly & Zongheng Yang & Ying Xiao & Zhifeng Chen & Samy Bengio & Quoc Le & Yannis Agiomyrgiannakis & Rob Clark & Rif A. Saurous (2017). Tacotron: Towards End-to-End Speech Synthesis. *Proc. Interspeech*, 1-10.

ABSTRACT

Performers' Rights to Learning Data in Artificial Intelligence Voice Synthesis Systems

Cho, Byeong Han

PhD Candidate, Hanyang University, Department of Culture Contents
Head of Policy Department, Korea Broadcasting Performers' Rights Association

This study aims to reveal that performers' rights are unfairly violated in the contract process to secure learning data for artificial intelligence voice synthesis systems and seeks ways to protect performers' rights.

Users of artificial intelligence learning voice data are demanding permanent and unlimited permission to use learning voice data in exchange for paying performers a small amount less than regular broadcast recording fees.

In response, this study reviewed the performers' rights under the Copyright Act, their general personal rights, and their rights under the Unfair Competition Prevention Act. In addition, it sought ways to establish fair license practices.

Although this study only discusses the performers' rights to artificial intelligence speech synthesis systems' learning data, this problem also started technological unemployment caused by artificial intelligence products. It is time to learn from the historical fact that the emergence of commercial records over 100 years ago created a new right called neighboring rights to overcome technological unemployment faced by singers and performers and find a solution through discussions at the

international level.

Keywords: Artificial Intelligence, Voice Synthesis System, Performers' Rights,
Copyright Act, Content Industry

[논문투고일 2022. 6. 16. 논문수정일 2022. 7. 24. 게재확정일 2022. 7. 27.]