

■ 논 단

저작권법상 텍스트·데이터 마이닝(TDM) 면책규정 도입 방향의 검토*

류 시 원

전남대학교 법학전문대학원 부교수, 법학전문박사

요 약 문

현대사회에서 데이터의 중요성이 날로 강조됨에 따라 텍스트·데이터 마이닝(Text and Data Mining, 'TDM') 기술이 인류의 지식과 편익의 증대에 크게 기여할 것임은 분명하다. 그러나 TDM은 대량의 데이터 복제를 수반하기 때문에 저작물과 데이터베이스를 배타적 보호대상으로 삼는 저작권법과 필연적으로 긴장관계에 놓이게 된다. TDM의 유용성을 긍정하는 전제에서 우리 저작권법에 TDM 행위를 위한 개별 권리제한조항을 도입하는 것에 관한 논의가 전개되고 있고, 그 바탕 위에서 우리 저작권법 개정이 시도되고 있다. TDM 면책입법은 전 세계적인 추세이므로 우리 저작권법의 TDM 면책규정의 형태를 만들어가는 데에 해외 입법례가 참조되는 것은 당연하지만, 우리나라의 정책적·법제적 상황에 가장 적합한 해결 방안을 찾기 위해서는 타국 입법의 '결과'만이 아니라 그 '과정과 배경'까지 함께 살펴보고 무엇이 같고 무엇이 다른지를 이해하려는 자세가 필요하다.

이와 같은 견지에서 주요국의 최근 TDM 면책규정의 입법 동향을 그 배경과 함께 소개하였다. 주요 논지는 다음과 같다. TDM 개별 면책규정의 초기 도입 사례이자 비상업적 연구 목적의 이용으로 면책 범위를 제한한 대표적인 입법례로 소개되는 영국 CDPA 제 29A조에 대해 영국 지적재산권청이 비상업적 목적에 관한 제한을 폐기하는 법개정을 시도하고 있다. 사상·감정의 비향수적 이용을 규정하고 있는 일본 저작권법 제30조의4는 미국식 일반 면책규정 도입의 대안으로 '유연한 권리제한' 규정으로서 신설되었으므로 공정 이용 조항에 의한 대처가 갖는 예측가능성의 문제를 해결하기 위해 개별 TDM 면책규정의 도입이 논의되고 있는 우리나라의 상황에서 비판적으로 참조하여야 한다. 최근 저작권법을 개정된 싱가포르의 경우 공정이용 조항이 이미 마련되어 있는 상황에서 예견가능성

* 이 글은 2022. 12. 16. 전남대학교 법학연구소, 인하대학교 법학연구소, 지식재산포럼 공동 주최로 열린 '디지털 혁신을 뒷받침하는 법과 정책' 학술대회에서 발표한 내용을 수정·보완한 것임을 밝힙니다. 심사 과정에서 논문의 개선 및 향후 연구 방향에 있어서 큰 도움이 되는 의견을 주신 익명의 심사위원들께도 깊은 감사의 말씀을 드립니다.

의 제고를 위해 비상업적 이용 등 목적의 제한이 없는 넓은 범위의 TDM 면책규정을 신설하였는바, 그 법개정의 동기 및 산업정책적 상황 등의 유사성에 비추어 싱가포르의 최근 입법례는 우리나라의 법개정에 특히 참고할 가치가 있다.

저작권법상 TDM 면책규정 도입 논의가 정초할 전반적인 방향성에 관해서는 우리나라의 상황에 대한 실증적 분석 및 그에 기초한 정책 목표의 설정이 선행되어야 한다는 점을 강조하였다. 필자가 저작권법 개정안 제43조의 개별 요건에 대한 당부 평가나 구체적인 입법 방안에 관한 의견 제시를 의식적으로 자제한 이유는 우리나라의 상황을 객관적으로 분석한 기초자료가 미비한 상태에서 선부른 입론이 정책 형성에 도움이 되지 않는다는 판단에서였다. 서로 다른 입법례의 일부 요소를 각각 차용할 때 나타날 부작용 혹은 시너지에 대한 결과 예측을 바탕으로 세밀한 입법적 검토가 요구된다. TDM 면책규정 도입에 관한 논의는 아직 성숙하지 않았다. 차제에 바람직한 입법 논의의 방향성을 재검토하고, 해외의 입법 과정과 배경을 참고하여 우리나라의 고유한 법제적·정책적 상황에서 가장 바람직한 결과를 도출할 수 있도록 차분히 다시 논의를 전개할 수 있기를 바란다.

주제어 : 텍스트·데이터 마이닝, TDM, 저작권재산권 제한, 면책, 공정이용, 비향수적 이용, 비향유적 이용, 정보분석, 컴퓨터 분석, 인공지능

〈目 次〉

I. 서 설	IV. TDM 면책규정 도입의 방향성
II. TDM의 기술적·법적 맥락	1. TDM 면책규정 도입에 관한 바람직한 논의 절차
1. TDM의 의의와 적용 범위	2. 각국 입법의 정책적·법제적 배경에 관한 고려
2. TDM과 이를 둘러싼 이해관계의 구조	3. 저작권법 전부개정안에 관한 단견
3. TDM의 유용성 및 법제 정비의 필요성	
III. 비교법적 검토	V. 결 어
1. 주요국의 TDM 면책 법제	
2. 주요국 법제 정리 및 비교	
3. 우리나라 저작권법 전부개정안	

I. 서 설

정보를 디지털 형태로 기록, 가공 및 전달하는 기술의 발달에 힘입어 매일 엄청난 양의 정보가 생산되고 있으며, 이 추세는 사물인터넷의 확산에 따라 가속화할 것으로 전망된다.¹⁾ 상상할 수 있는 수준을 뛰어넘은 정보의 범람 속에서 인간의 능력만으로 수많은 정보 혹은 데이터 속에서 가치를 발견하는 것은 점점 더 어려운 일이 되고 있다.²⁾ 이와 같은 상황에서 컴퓨팅 자원을 활용해 대량의 데이터를 분석하고 그로부터 일정한 패턴이나 구조를 발견해 의미 있는 메타정보를 추출하는 이른바 텍스트·데이터 마이닝(Text and Data Mining, 'TDM') 기술에 대한 관심이 높아지고 있다. 특히 현대적 형태의 TDM 기술은 인공지능(Artificial Intelligence, 'AI') 기계학습 기술과 연계되어 학문연구의 도구로서뿐 아니라 국가 경쟁력을 좌우하는 산업정책적 의미까지 지니게 되었다.

그러나 TDM이 분석의 대상으로 삼는 정보는 문자나 그 밖의 표현 형태로 존재하고, 그에 저작물이 포함될 수 있다. 또, TDM 분석을 위해 접근·이용하게 되는 데이터의 출처가 저작권법에 의해 보호되는 데이터베이스일 수도 있다. 이처럼 TDM 활용에 대한 요구는 저작권 보호제도와 일정한 긴장 아래 놓여 있다. 대량의 데이터를 분석하는 TDM의 속성상 데이터 혹은 저작물의 수집·분석 과정에 수반되는 복제의 양적 크기는 저작권 보호제도와 충돌 가능성을 증가시키는 동시에,³⁾ 데이터의 다양성과 결합해 그러한 문제의 탐지·회피 기회를 줄이는 내재적 한계로 작용한다.⁴⁾ 이에 세계 각국에서는 데이터 주도 경제가 본격적으로 부상한 2000년대 초반 무렵부터 저작권법제에 TDM 면책규정을 도입하는 것에 관한 정책적 논의가 전개되

-
- 1) Bernard Marr, 'How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read', *Forbes*, May 21, 2018.
 - 2) 이러한 경향은 학술연구 분야에서도 찾아볼 수 있다. 2010년을 전후하여 보고된 통계에 의하면, 전 세계에서 매년 150만여 건의 새로운 학술논문이 발행되지만 그중 약 90%는 전혀 인용되지 않으며, 저자와 심사위원, 편집자 외에는 누구에게도 읽히지 않는 논문의 비율도 50%에 이른다고 한다. Thomas Margoni & Martin Kretschmer, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology', *GRUR International*, vol. 71(8), 2022, p. 689.
 - 3) European Commission, 'Impact Assessment on the Modernisation of EU Copyright Rules', Commission Staff Working Document SWD(2016) 301 final Part 1/3, Sep. 2016, p. 104; 정진근 외 8명, "신기술에 대응하는 저작권 보호방안 연구", 한국저작권보호원 연구보고서, 2021. 12., 51-52면.
 - 4) 오승중, "데이터마이닝 및 텍스트마이닝과 저작권재산권의 제한", 『홍익법학』 제20권 제2호, 홍익대학교 법학연구소(2019), 466면 참조.
 - 5) TDM 기술과 저작권 보호제도 간의 긴장관계를 시사하는 최근의 사례로, 국립국어연구원이 우리말 연구에 활용하기 위한 공공 빅데이터의 구축을 목적으로 수 년 전부터 진행하여 2020년부터 서비스한 우리말 말뭉치 사업의 문어(文語) 부문에 참여한 한 출판사에 대한 저작권 침해 문제 제기로 인해 국립국어원의 빅데이터 서비스가 중단된 일이 있다(임인택, "[단독] 국립국어원 '문어 빅데이터 서비스' 저작권 문제로 일부 중단", 한겨레 2022. 8. 24.자 기사). 이 사례는 TDM 분석 과정이 아니라 그 전 단계에서의 데이터 수집 및 TDM 분석에서의 제공을 둘러싸고 저작권 문제가 불거진 일이지는 하나, TDM이 그것이 적용되는 일련의 과정 속에서 현행 저작권 보호제도와 충돌할 수 있음을 보여주는 좋은 사례이다.

어 왔다. 그 결과 실제 규정 도입의 성과를 낸 국가·권역도 상당수에 이른다.

우리나라에서도 2020년과 2021년 사이 TDM 면책규정의 신설에 관한 저작권법 개정안이 단계적으로 성안되었고, 위 시기를 전후하여 학계에서 해외 법제들의 소개와 더불어 우리나라의 TDM 면책규정 도입 방향을 제시한 연구물들이 발표되었다.⁶⁾ 그러나 대부분의 선행연구들은 특정 국가 법제의 소개에 치중하거나 제 외국의 규정들을 평면적으로 제시하는 등의 한계가 있다. 각 국가·권역의 TDM 면책규정이 그 형식이나 면책 범위 제한 등에서 보이는 차이의 정책적 원인과 제도적 배경에 대한 분석에 있어서는 다소간의 부족함이 있다. 또한 선행연구들은 대부분 미국, 영국, 독일, 프랑스, 일본의 관련 법제를 소개하고 있는데, 독일과 프랑스는 2019년에 제정된 유럽연합 디지털 단일시장 저작권 지침(이하 'DSM 지침')⁷⁾의 국내법 전환(national transposition)을 위해 2021년에 TDM 면책에 관한 해당국 저작권법 규정을 크게 개정했고, TDM 개별 면책규정의 초기 도입 사례로 소개되는 영국 저작권법 제29A조 역시 영국 정부가 비상업적 목적에 관한 제한을 없애는 방향으로 그 개정을 추진하는 등⁸⁾ 변화의 움직임이 목격되는데, 최근의 변화된 법제의 내용과 경향을 재검토할 필요가 생겼다.

이에 본고에서는 TDM 면책에 관계된 각국의 규정들을 체계적으로 비교·분석하고, 침해의 예외사유를 구성하는 이용 주체와 목적 등의 제한 요소들(limiting elements)이 지닌 의미를 검토하고자 한다. 더 나아가 타국 TDM 면책규정의 구조와 개별 제한 요소들을 우리나라의 저작권법 개정에 참고하는 데 있어서 고려할 사항들에 대해서도 살펴본다.

II. TDM의 기술적·법적 맥락

1. TDM의 의의와 적용 범위

TDM은 데이터 과학의 분과인 텍스트 마이닝(text mining)과 데이터 마이닝(data mining)을

6) 김용주, "텍스트 및 데이터마이닝을 위한 저작권법 개정 방향", 「법학연구」 제61권 제2호, 부산대학교 법학연구소(2020); 박성호, "텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작재산권의 제한 - 인공지능의 빅데이터 활용을 중심으로 -", 「인권과 정의」 제494호, 대한변호사협회(2020); 차상욱, "저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점 - 텍스트·데이터마이닝(TDM) 면책규정을 중심으로 -", 「경영법률」 제31집 제1호, 한국경영법률학회(2021); 김창화, "데이터 마이닝과 저작권 면책의 범위 및 한계", 「계간 저작권」 제34권 제2호, 한국저작권위원회(2021); 김경숙, "TDM 관련 저작권법 개정안의 비판적 고찰 - 기술적보호조치 관점에서의 비교법적 검토 -", 「경영법률」 제31집 제3호, 한국경영법률학회(2021) 외 다수.

7) Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (hereinafter, 'EU DSM Directive').

8) 후술하는 바와 같이, 영연방에 속한 싱가포르를 영국보다 먼저 상업적 목적이나 영리성에 관계없이 TDM을 허용하는 내용의 포괄적인 면책규정을 도입하는 저작권법 개정을 최근 단행하였다.

통칭하는 용어이다. 전자는 주로 텍스트 데이터⁹⁾로부터 패턴, 경향, 관계 등 의미 있는 정보를 발견하고 해석하는 기술로서¹⁰⁾ 비구조적(unstructured) 데이터를 구조화하는 데에 특징이 있고,¹¹⁾ 후자는 흔히 데이터베이스와 같이 구조화된(structured) 데이터셋에서 유용한 정보를 발견하기 위해 사용되는 전산·통계적 기법을 의미한다는¹²⁾ 점에서 약간의 차이가 있다.¹³⁾ 텍스트 마이닝과 데이터 마이닝의 관계에 관해서는 전자가 후자의 한 분과라는 견해¹⁴⁾를 비롯해 다양한 시각이 존재하지만, 기술 발달에 따라 양자의 경계가 점차 흐려지고 있다. 이에 따라 정책적 논의에서는 양자를 결합하여 ‘텍스트·데이터 마이닝’이라고 칭하는 것이 일반화되었고,¹⁵⁾ 본고에서도 그 약칭인 ‘TDM’이라는 용어를 사용한다.

정책 논의에서 TDM은 통상 ‘대량의 텍스트 및 데이터를 고급 알고리즘에 기반하여 읽고 다양한 목적으로 분석, 분류, 정렬, 패턴 인식, 정보 추출을 수행하는 기술’을 지칭한다.¹⁶⁾ 이와 유사하게 영국의 지적재산당국인 IPO(Intellectual Property Office)는 TDM을 ‘패턴, 경향, 기타 유용한 정보를 분석하는 자동화된 분석 기술’로 정의한다.¹⁷⁾ TDM이 적용되는 분야는 주로 AI 기계학습과 관련하여 언급되지만, 그 밖에도 의학 및 제약 연구, 문학 비평, 언어 연구 등의 여러 학문분야에서 다양한 목적으로 활용되며,¹⁸⁾ 인터넷 검색이나 뉴스피드 추천 알고리즘 개발을 위한 AI 학습 등 우리 삶과 밀접한 영역에도 침투해 있다.

TDM 기술의 유용성을 보여주는 사례로, Ted Underwood 등이 2018년에 발표한 ‘영어문학에서 젠더의 변화’라는 연구를 들 수 있다.¹⁹⁾²⁰⁾ 연구자들은 HathiTrust 디지털 도서관 등이 제

9) 이때의 ‘텍스트(text)’는 문자 외에 이미지, 음성, 영상 등을 포괄하는 광의의 데이터 형식을 뜻한다. See Matthew Sag, ‘The New Legal Landscape for Text Mining and Machine Learning’, *Journal of the Copyright Society of the USA*, vol. 66, 2019, pp. 295-296.

10) Jiawei Han, Jian Pei & Hanghang Tong, *Data Mining: Concepts and Techniques* (4th ed.), Morgan Kaufmann, 2022, p. 605.

11) <<https://www.ibm.com/cloud/learn/text-mining>> (마지막 방문: 2022. 9. 4.).

12) See Colleen McCue, *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis*, Butterworth-Heinemann, 2007, p. 25; John A. Bunge & Dean H. Judson, ‘Data Mining’, in *Encyclopedia of Social Measurement* (Kimberly Kempf-Leonard ed.), Elsevier, 2005, p. 617.

13) 텍스트 마이닝과 데이터 마이닝이 분석 대상과 구체적인 기법에 있어 구별된다는 견해로는 다음 문헌 참조. Said A. Salloum, Mostafa Al-Emran & Khaled Shaalan, ‘Mining Social Media Text: Extracting Knowledge from Facebook’, *International Journal of Computing and Digital Systems*, vol. 6(2), 2017, p. 74.

14) *Supra* note 11.

15) Michael W. Carroll, ‘Copyright and the Progress of Science: Why Text and Data Mining Is Lawful’, *UC Davis Law Review*, vol. 53, 2019, p. 899, n. 19.

16) See Sergey Filippov & Paul Hofheinz, ‘Text and Data Mining for Research and Innovation’, *The Lisbon Council*, iss. 20/2016, 2016, p. 2.

17) UK IPO, ‘Exceptions to copyright: Research’, Oct. 2014, p. 6.

18) Sag, *supra* note 9, p. 295. See also Sean Flynn et al., ‘Comment in response to the WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI)’, WIPO/IP/AI/2/GE/201, available at: <https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_global_expert_network_on_copyright_user_rights.pdf> (마지막 방문: 2022. 9. 4.).

공하는 1703년부터 2009년까지의 약 100,000권의 소설 속 말뭉치들과 서지 정보를 분석해 다 음과 같은 결과를 얻어냈다. 첫째로, 지난 170년 동안 문학작품 속 인물들의 등장 빈도와 묘사 등의 측면에서 등장인물의 젠더 구분이 점차 약화되어 왔다는 것이다. 두 번째 분석 결과는 19 세기를 거쳐 1960년대 초반에 이르기까지 문학작품 속 여성 인물의 비중이 꾸준히 감소해 왔 고, 유사한 경향이 여성 작가들의 작품 속에서도 보인다는 점이다. 위의 두 번째 분석 결과는 19~20세기를 거치며 문학작품 속 젠더 균형성이 향상되었으리라는 일반적인 예상을 빗나가는 것인 동시에, 젠더학적 관점에서 첫 번째 분석과 모순되는 것으로 풀이될 수 있다는 점에서 놀 라운 결과였다. 연구자들은 일견 모순된 분석 결과에서 보이는 현상의 원인과 의미를 탐구하 여 논문으로 발표하였다. 위의 연구 결과는 단어와 문장의 경계 확정, 문장의 문법적 구조 분 석, 동일한 등장인물을 지칭하는 상이한 표현을 식별하는 것과 같은 여러 기능을 수행하는 다 층 알고리즘으로 구성된 '원격 읽기(distant reading)'라고 하는 TDM 기술을 대량의 텍스트에 적용해 얻어낸 것이다.²¹⁾ 100,000권이 넘는 책과 그 속에 등장하는 85,000여 명의 인물에 대한 분석은 컴퓨터와 TDM 기술의 도움을 받지 않고서는 수행할 수 없는 규모이고, 그와 같은 대 규모 연구였기에 분석 결과의 신뢰성이 뒷받침될 수 있었음을 짐작하기는 어렵지 않다. 이 사 례는 TDM 기술이 학문과 생활 영역에서 의미 있는 정보를 발견하는 데 있어 갖는 유용성을 보여준다.²²⁾

2. TDM과 이를 둘러싼 이해관계의 구조

TDM은 지금 이 순간에도 발전을 거듭하고 있는 기술 분야이고, 다양한 하위 분석기법을 포 괄하는 용어이다. 따라서 TDM의 방법론을 일률적으로 규정하기는 어렵다. 그러나 TDM 수행 과정을 대략 다음과 같이 개념화하는 것은 가능하다. 이하에서 그리고 글 전체에서 TDM의 대 상으로 언급되는 데이터는 저작물 혹은 데이터베이스의 상당 부분²³⁾을 의미한다. 이러한 점을

19) Ted Underwood, David Bamman & Sabrina Lee, 'The Transformation of Gender in English-Language Fiction', *Journal of Cultural Analytics*, vol. 3(2), 2018 (cited from Sag, *supra* note 9, n. 14).

20) 이하의 서술은 Sag, *supra* note 9, pp. 296-299에 소개된 Ted Underwood 등의 연구 내용을 간략히 정리한 것이다.

21) Sag, *supra* note 9, p. 297.

22) TDM의 유용성을 보여주는 또 다른 사례로, 2016년 Facebook에 게시된 포스팅과 뉴스 채널을 대상으로 텍스트 분석을 수행해 당시 미국 대선 국면과 관련된 언론기관의 활동과 대중의 관심사 등을 연구한 사례가 있다. See Salloum et al., *supra* note 13. 그 밖의 연구 사례로, Sag 교수는 레이노 중후군 환자에게 도움이 되는 어유(魚油)에 관한 1986년의 연구, 헌팅턴병에 관계된 특정 분자들을 다룬 2005년의 연구 사례를 들고 있다. See Sag, *supra* note 9, pp. 295-296.

23) 우리 저작권법의 데이터베이스는 저작물이나 저작물이 아닌 소재를 체계적으로 배열 또는 구성한 편집물로서 그 소재에의 개별적인 접근이나 검색이 가능하도록 한 것을 의미한다(제2조 제17호). 저작권법상 보호되는 데이터베이스의 전부 또는 상당한 부분에 대해서는 복제·배포·방송·전송이 금지되며, 상당한 부분의 복제에 이르지 못하더라도 반복적이거나 특정 목적을 위해 체계적으로 함으로써 데이터베이스의 통상적인 이용과

고려하여 본고에서는 저작권에 관계된 서술 맥락에서 ‘데이터’와 ‘저작물’ 혹은 ‘저작물 등’을 상호 대등한 의미로 교환적으로 사용할 것이다.

TDM이 수행되는 단계를 일별하면, 접근(access), 추출(extraction), 마이닝(mining), 사용(use)으로 나눌 수 있다.²⁴⁾ 이들 단계는 반드시 시간상 순차적으로 배열되거나 상호 명확히 구별되지 않을 수 있다. 또한 모든 단계가 동일한 주체에 의해 수행되지 않을 수도 있다.

첫 번째 단계인 ‘접근’은 디지털 데이터 혹은 유체물에 수록된 아날로그 데이터에 대한 접근을 모두 포함한다. TDM을 위한 데이터가 복제되기 전 단계이지만, 데이터 수집의 전제가 되는 단계라는 의미를 갖는다. 데이터 제공자 자신이 권리자인 경우(원권리자로부터 권리를 양수한 경우 포함)에는 이용자와의 관계가 1:1 대향 구조로 되지만, 원권리자의 허락을 받은 범위에서 저작물 등을 이용자에게 제공하는 전문 데이터베이스 제공기관도 해당 범위 내에서 이용자의 데이터 접근을 통제할 수 있다. 나아가 이들 사이에 저작물 등에 대한 이용자의 접근을 매개하는 도서관 등의 제3의 기관이 개입될 수도 있다.²⁵⁾ 따라서 접근통제 권한을 가진 자는 반드시 저작물의 원권리자로 한정되지 않는다. 다만 이처럼 여러 단계로 구성된 관계 하에서도 권리자는 데이터 제공자가 실시하는 접근통제를 통해 저작물 등에 대한 경제적 보상을 실현할 수 있고, 데이터 제공자와의 계약을 통해 간접적으로 접근통제에 관여할 수 있다. 접근통제의 방법은 크게 둘로 나뉜다. 계약을 통한 통제와, 로그인·인증과 같은 기술적 보호수단(technical protection measure, TPM)에 의한 통제이다. ‘접근’ 단계에서 데이터 제공자는 이용자에게 접근뿐 아니라 TDM의 가부, 범위에도 제한을 가할 수 있다.

다음으로 ‘추출’은 TDM에 활용할 저작물 등의 데이터를 수집하고 분석 목적에 맞게 가공하는 단계이다. 이 단계에서는 다양한 방법에 의한 대량 데이터의 복제가 수반될 수 있다. 수집과정에서의 1차 복제, 1차 복제물을 TDM 처리용 입력 데이터로 가공하는 과정에서의 복제, 백업 혹은 검증 목적의 추가 복제 등이 그 예이다.²⁶⁾ 권리자와의 관계에서 가장 직접적인 법적 제한, 즉 TDM 행위가 저작권자 등의 배타권과 충돌하는 문제가 이 단계에서 발생한다. 유의할 사항은 이 단계의 ‘추출’과 후술하는 ‘마이닝(분석)’이 동일한 주체에 의해 수행되지 않을 수 있다는 점이다. 데이터의 수집과 가공을 전문으로 하는 민간·정부가관이 TDM용 데이터

충돌하거나 그 제작자의 이익을 부당하게 해치는 경우에는 상당한 부분의 복제 등으로 간주된다(제93조 제1, 2항). 따라서 TDM 과정에서 복제 등 이용되는 데이터가 저작물이 아닌 경우에도, 데이터베이스의 상당한 부분(으로 간주될 수 있는 경우)에 해당한다면 저작권법과 마찰을 빚을 수 있다.

24) Sag, *supra* note 9, pp. 346-347.

25) 구글북스 사건에서 이러한 구조를 엿볼 수 있다. 구글북스 프로젝트는 출판사 등 권리자와의 협약인 ‘파트너 프로그램’과 도서관들과의 협약인 ‘도서관 프로젝트’로 구성되었는데, 후자에서 구글은 도서관들과의 협약을 통해 장서를 스캔해 디지털 복제하는 것에 관한 허락을 받았다. 그럼에도 구글은 권리자 단체인 작가조합으로부터 저작권 침해 소송을 당했다. See *Authors Guild, Inc. v. Google, Inc.*, 954 F. Supp. 2d 282, 285-286 (S.D.N.Y. 2013).

26) See Sag, *supra* note 9, pp. 353-357.

베이스를 조성하고, 이를 연구자들에게 제공하는 경우도 상정할 수 있기 때문이다. 이러한 경우에는 ‘추출’ 단계가 추가의 복제 혹은 전송을 수반할 수 있다.²⁷⁾²⁸⁾

앞의 두 단계가 TDM 기법을 적용하기 위한 입력 데이터 형성 과정이었다면, ‘마이닝’은 그 데이터로부터 의미 있는 정보를 얻어내기 위한 본격적인 분석 과정이라는 점에서 TDM 기술의 핵심을 이룬다. 이 단계에서는 컴퓨터가 데이터를 분석하기 위해 읽어들이게 되므로 적어도 일시적 복제는 행해진다. 한편, 대량의 데이터 분석 과정에서 분산 처리를 위해 데이터가 서버 간에 복제될 수도 있다.²⁹⁾ 앞의 ‘추출’ 단계와 연계하여, 분석이 끝난 뒤 저작물 등을 포함하는 입력 데이터를 삭제하는 경우만을 면책 범위에 포섭할 것인지³⁰⁾ 아니면 제한적인 범위에서 보관 또는 전송하는 것까지 허용할 것인지 등이 문제가 될 수 있다.

마지막으로 TDM 분석의 결과물을 ‘사용’하는 단계에서는 분석을 통해 추상화된 데이터 혹은 정보만 남게 되므로 분석에 이용된 데이터가 직접 사용되지 않는 때가 많을 것이다. 따라서 저작권 침해 문제가 발생할 개연성은 낮아 보인다. 그러나 적용된 TDM 기법에 따라 처리된 통계적·추상적 정보 외에 말뭉치(corpus) 등의 입력 데이터가 마이닝(분석) 결과물에 일부 남아있을 수도 있다는 보고가 있다.³¹⁾ 이 경우 만약 입력 데이터가 분절적인 형태로 잔존한다면 저작권으로 보호되는 표현에 해당하지 않을 수 있고, TDM 분석 결과는 원저작물의 표현상의 본질적 특징을 감득할 수 없는 ‘창작성의 매몰·퇴색’에 해당해 이론상 저작권 침해가 부정될 수 있다는 견해도 있다.³²⁾ 그러나 이와 달리 분석 결과물에 원저작물이 동일성을 유지한 상태로 상당한 분량 포함되어 있는 경우라면, 그 결과물을 사용하거나 제3자의 사용에 제공하는 것이 형식상 또 다른 저작물·데이터 이용행위를 수반할 수 있을 것이다.³³⁾

27) 서두에서 언급한 국립국어원의 우리말 말뭉치 사업이 그 예이다. 국립국어원은 권리자들로부터 수집한 말뭉치 데이터베이스를 2020년 8월 25일부터 ‘모두의 말뭉치’ 웹사이트(<https://corpus.korean.go.kr/>)를 통해 국어 연구자 및 언어 정보 처리 분야의 개발자 등에게 제공해 왔다.

28) 이러한 경우에 관해 국내에서 제시된 견해 중 ‘중간산출물을 서비스나 거래에 제공하는 최종단계에서는 변형적 이용 내지 비표현적 이용에 해당할 여지가 상대적으로 축소되므로 저작권법상 공정이용 인정의 가능성은 크지 않다’고 하는 견해가 있다. 차상욱, 앞의 글(주 6), 45면 참조.

29) 이러한 일이 생기는 것은 TDM 기술의 중요한 적용분야인 AI 기계학습과 그 밖의 TDM 활용이 대체로 데이터가 중앙집중형으로 처리되기 때문이다. 그러나 데이터 중앙집중과 기기 간 데이터 교환의 최소화를 중심 과제로 하는 새로운 기계학습 분야인 ‘연합학습(federated learning)’의 경우에는 데이터의 수집, 분석, 모델 생성이 로컬의 기기 단위에서만 이루어지므로 기기간 추가적 데이터 복제·전송은 발생하지 않는다. See Mauritz Kop, ‘Machine Learning & EU Data Sharing Practices’, *Stanford-Vienna Transatlantic Antitrust and IPR Developments*, iss. 1/2020, Mar. 2020, p. 12.

30) 2021년 개정 전의 독일 저작권법 제60d조 제3항에서는, 소재의 복제물과 말뭉치(Korpus)를 연구 종료 후 삭제해야 한다고 규정하면서, 도서관과 기록보관소 등의 제한된 기관에 전송하여 영구적으로 보관하는 것은 허용하였다.

31) Margoni & Kretschmer, *supra* note 2, pp. 693-694.

32) 오승종, 앞의 글(주 4), 469면 참조.

33) 이와 같이 분석 결과물에 일부 원데이터가 포함된 상태로의 복제나 전송이 뒤에서 소개하는 일본의 2018년 개정 저작권법과 우리나라의 2021년 1월 저작권법 전부개정안에 공통적으로 포함된 ‘향수’ 내지 ‘향유’ 목적에

이상과 같이 TDM을 둘러싼 법적·경제적 이해관계는 데이터의 생산, 제공, 이용 등에 참여하는 여러 이해관계인들(stakeholders), 나아가 TDM을 통해 얻어진 새로운 정보, 서비스, 관계의 가치를 누리게 될 소비자 내지 공중의 이익까지 포함하는 복잡적이고 중층적인 구조를 이룬다. 정책적 논의의 목적상 저작권 보호체계 내에서 TDM과 관련하여 상충하는 법익을 '권리자 대 이용자'로 단순화하는 것이 불가피한 면이 있더라도, 지나치게 단순화된 관점으로 인해 논의가 왜곡되거나 불충분하게 되는 일은 없어야 한다. 이러한 시각은 우리 저작권법 개정에 관한 논의에 있어서는 물론이고, 그 논의에 참고할 해외 법제를 이해하는 데 있어서도 공통된 바탕이 되어야 할 것이다.

3. TDM의 유용성 및 법제 정비의 필요성

여러 문헌에 소개된 바와 같이 최근 주요국에서 TDM 면책 관련 입법을 논의, 실행하고 있다. 일본은 2009년과 2018년 두 차례에 걸친 저작권법 개정을 통해 AI 개발을 위한 데이터의 수집, 처리 등에 관한 규정을 정비하였고, 유럽의 경우 2014년 영국을 필두로 하여 2016년 프랑스, 2017년 독일이 각각 제한된 범위의 TDM 면책규정을 자국 저작권법제에 도입하였다. 2019년에는 유럽연합이 역내 TDM 관련 규율체계 통일화 내용을 포함하는 DSM 지침을 제정하였다.³⁴⁾

이러한 흐름의 근거에는 TDM의 활용이 경제와 사회 전체의 발전에 기여하고 인간에게 유용한 결과를 가져온다는 인식이 깔려 있다.³⁵⁾ TDM 기술은 각종 연구에 활용되어 새로운 방법으로 과거에 탐구되지 못했던 지식을 발견하고 이로써 인류의 복지를 증진하는 유용한 도구가 된다.³⁶⁾ 그뿐 아니라, 고객 맞춤 서비스 개발과 재정 관리 등의 기업활동, 표절 검사, 대중교통 체계 개선, 플랫폼 서비스 혁신과 같이 공공 및 민간 영역에서 사회 후생을 증대시킬 가능성을 이미 보여주고 있다. 그러나 TDM은 그 수행 과정에서 저작물 등의 복제를 수반하므로 저작권 보호체계와의 마찰을 피할 수 없다고 여겨지고 있다.

해당하지 않는 것으로 해석될 여지가 있을지에 관해서는, 적용된 기술과 결과물의 이용 방법 등 사실관계가 고려되어야 할 것이다.

34) 이들 해외 국가의 TDM 면책규정 도입 내용에 관한 상세한 소개는 차상욱, 앞의 글(주 6) 등 참조.

35) Darren Ang, 'The Web Scraper's World of Copyright Exceptions and Contractual Overrides', *SLR Juris Illuminae*, vol. 13, 2021, p. 7.

36) Thomas Margoni, 'Text and Data Mining in Intellectual Property Law: Towards an Autonomous Classification of Computational Legal Methods', CREATE Working Paper 2020/1, p. 1 ("[T]he impact that TDM may have in the *direction* of scientific enquiry is invaluable. This is because by identifying the correlations and patterns that are often concealed to the eye of a human observer due to the amount, complexity, or variety of data surveyed, TDM allows for the discovery of concepts or the formulation of correlations that would have otherwise remained concealed or undiscovered. Considering this point of view, it can be effectively argued that TDM can create new knowledge from old data.").

앞서 소개한 2018년의 인문학 연구 사례³⁷⁾는 TDM의 유용성과 더불어 저작권법과의 긴장관계를 잘 시사한다. 여기서 특히 주목되는 부분은 이 연구가 HathiTrust 디지털 도서관의 데이터베이스를 이용해 수행되었다는 점이다. 미국 작가협회가 제기한 저작권 침해소송³⁸⁾의 상대방이었던 바로 그 HathiTrust이다. 여러 문헌에 소개되었듯이 이 소송에서는 HathiTrust가 타인의 저작물을 주서버와 미러서버에 복제하여 도서관 데이터베이스를 구축하고 전문 검색(full-text search) 서비스 등을 제공한 행위가 미국 연방저작권법 제107조에 따른 공정이용(fair use)에 해당하는지가 문제되었다. 이에 관해 제2순회항소법원은 제1요소(변형적 이용), 제3요소(전문 검색에 필요한 합리적 범위의 이용), 제4요소(이용자에게 제한된 결과만 제공하고 보안조치가 적용된 점 등에 비추어 출판물 수요를 대체하지 않음)의 평가를 중심으로 HathiTrust의 위 행위가 공정이용에 해당함을 긍정했다.³⁹⁾ 만약 이 소송에서 HathiTrust 측의 저작권 침해가 인정되어 디지털 도서관 서비스가 중단되었다면 위 2018년의 연구는 세상에 나오지 못했을 수도 있다. TDM 적용을 확대하고 그 기반을 조성하는 데 있어서 저작권법제와의 마찰을 해소하는 일의 중요성을 이 사례는 말해준다.

따라서 TDM의 순기능을 긍정하는 전제에서, 우리나라의 현행 저작권법이 TDM에 적합한 환경을 제공해 주는지, 개선할 지점은 없는지를 검토하는 것은 자연스러운 수순이라 할 것이다. 다만 이러한 검토는 저작권법이 TDM 기술의 활용을 위해 극복해야 할 ‘장애물’이라는 인식에서 출발할 것이 아니라 ‘환경 변화에 따른 법제 정비’의 관점에 바탕을 두고 우리나라의 정책 환경 속에서 저작권법의 목적인 권리 보호와 공정한 이용 도모 간의 균형점을 찾아가는 방식으로 이루어져야 한다. 문화체육관광부가 2020년 7월 저작권법 전부개정 추진계획을 발표하면서 그 배경을 “창작과 이용 환경의 변화를 반영”하려는 취지라고 밝힌 것도 동일한 관점에 바탕한 것이다.⁴⁰⁾

주지하는 바와 같이 우리 저작권법에는 이미 일시적 복제에 대한 면책규정인 제35조의2와 저작재산권 제한에 관한 일반조항으로서 제35조의5 공정이용 규정이 마련되어 있다. 그러나 TDM 수행 과정에서 다양하게 일어날 수 있는 저작물 등 복제 양태를 감안했을 때 일시적 복제 규정이 현실의 전체 TDM 단계에서 유효하게 작동할지는 미지수이다.⁴¹⁾ 한편 공정이용 조항은 그 문언의 추상성과 판례 부족으로 인해 TDM 기술을 활용하려는 연구자들의 법적 불안을 제거하기 위한 예견가능성의 제공 측면에서 부족함이 없지 않다. 학계에서도 현행 저작권

37) 이 글의 II.1. 참조.

38) Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014).

39) *Id.* at 94-104.

40) 문화체육관광부, “창작자 권리보호는 강하게, 저작물 이용은 손쉽게 - 문제부, ‘추가보상 청구권’, ‘확대된 집중관리’ 및 인공지능 개발·활용 촉진 등 담은 「저작권법」 전부개정 추진 -”, 문화체육관광부 보도자료, 2020. 7. 1., 1면.

41) 같은 취지로, 오승종, 앞의 글(주 4), 472-473면; 신창환, “인공지능 시대의 저작물 대량 디지털화에 관한 소고”, 「성균관법학」 제31권 제1호, 성균관대학교 법학연구원(2019), 159-160면.

법 제35조의5에 의한 공정이용 인정 가능성과는 별개로 TDM 개별 면책규정을 도입하는 것이 바람직하다는 데에 대체로 견해가 일치하는 것으로 보인다.⁴²⁾ 필자도 큰 틀에서 그러한 입장에 동의하지만, 그 이유는 우리나라의 정책 현실에 부합하면서도 TDM을 둘러싼 상충하는 이익을 세밀하게 조율하기에 현행 법제가 불충분하다고 보기 때문이다.

이하에서는 주요국의 TDM 면책 관련 법제의 특징과 그 도입 배경을 최근의 상황을 중심으로 살펴보고(Ⅲ), 이어서 우리 저작권법에 개별 면책규정을 도입하는 데 있어 고려할 사항들에 대해 논의한다(Ⅳ).

Ⅲ. 비교법적 검토

1. 주요국의 TDM 면책 법제

가. 유럽연합 DSM 지침 - 학술연구 목적 면책규정과 일반 면책규정에 의한 이원적 규율

(1) DSM 지침의 TDM 면책규정

2019년 6월 7일 발효된 DSM 지침은 TDM을 “패턴, 경향, 상관관계 등의 정보를 생성하기 위해 디지털 형식의 텍스트와 데이터를 분석하는 것을 목적으로 하는 자동화된 분석기술”로 정의하고(제2조 제2호), TDM 행위의 면책 요건들을 제3조와 제4조에서 각각 규정한다. 먼저 제3조는 연구기관(research organisation)⁴³⁾과 문화유산기관(cultural heritage institution)⁴⁴⁾이 학술연구⁴⁵⁾ 목적의 TDM을 수행하기 위해 적법하게 접근할 수 있는(have lawful access) 저작물 등을 복제 및 추출하는 경우를 면책 범위로 정하고 있다(제1항). 2016년 9월에 마련된 초안에서는 연구기관만을 면책대상으로 규정하였으나 최종안에서 문화유산기관을 추가한 것이다. 위의 복제물·추출물은 연구결과의 검증을 포함한 학술연구 목적을 위해 적절한 보안수준을 갖춰 저장 및 보유할 수 있다(제2항). 권리자는 저작물 등이 호스팅되는 네트워크와 데이터베이

42) 오승종, 앞의 글(주 4), 477-478면; 김용주, 앞의 글(주 6), 301-304면; 차상욱, 앞의 글(주 6), 43-44면; 김창화, 앞의 글(주 6), 21-22면; 김병일·신현철·안창원, “빅데이터 분석과 데이터 마이닝을 위한 저작권 제한”, 「계간 저작권」 제30권 제1호, 한국저작권위원회(2017), 56면; 이상용, “데이터의 비계약적 이용 - 데이터 마이닝을 위한 저작권 제한을 중심으로 -”, 「강원법학」 제65권, 강원대학교 비교법학연구소(2021), 39면 등 참조.

43) 학술연구의 수행 또는 학술연구를 수반하는 교육활동을 주된 목표로 하는 대학(그 도서관을 포함) 등의 기관을 뜻하며, 비영리 기반(non-for-profit basis)이거나 모든 수익을 학술연구에 재투자할 것, 혹은 회원국이 인정하는 공익 목적을 추구하는 기관이어야 한다. DSM Directive, Art. 2(1).

44) 공중이 접근 가능한 도서관이나 박물관, 기록보관소(archive) 또는 필름이나 소리 유산을 다루는 기관을 뜻한다. DSM Directive, Art. 2(3).

45) ‘학술연구(scientific research)’는 자연과학과 인문과학 분야를 모두 포함한다. DSM Directive, recital 12.

스의 보안 및 무결성을 확보하기 위한 수단을 그 목적을 넘지 않는 범위에서 적용할 수 있다(제3항).

제4조는, 제3조와 달리 학술연구라는 목적 범위 및 연구기관 등 실행주체에 제한을 두지 않고, 적법하게 접근할 수 있는 저작물 등에 대한 TDM을 수행하기 위한 복제, 추출을 허용한다(제1항). 또한 복제물·추출물 보유의 범위와 조건도 제3조 제2항에 비해 완화된 “TDM 목적을 위해 필요한 범위”로 규정한다(제2항). 다만, 본 조의 TDM 면책은 권리자가 적절한 방법에 의해(온라인을 통해 공중이용 가능한 콘텐츠의 경우에는 기계 판독 가능한 수단 등) 명시적으로 유보하지 않았을 것을 조건으로 한다(제3항). 즉, ‘옵트아웃(opt-out)’에 의한 권리자의 사후 개입 여지를 남겨두었다. 본 제4조는 2016년의 초안에는 포함되어 있지 않았으나 입법 논의 과정에서 추가되었다.

제3조는 계약에 의해 그 적용을 배제하는 것이 금지되는 강행규정이지만, 제4조는 그렇지 않다(제7조 제1항). 제3조와 제4조는 저작물 등에 대한 적법한 접근권한을 요구하는 점에서 공통되는데, 이때의 ‘적법한 접근’이란 오픈액세스 정책이나 구독 등 계약 기타 법적 수단을 통해 가능한 접근과, 온라인에서 자유롭게 이용 가능한 콘텐츠에 대한 접근을 의미한다.⁴⁶⁾ 제3조와 제4조 모두 권리자에 의한 기술적 보호수단(TPM)의 적용이 허용된다.⁴⁷⁾

(2) DSM 지침의 특징 및 비판

DSM 지침의 TDM 면책규정 체계는 서로 다른 적용 범위 및 조건을 가진 제3조와 제4조로 이원화되어 있다. 제3조는 면책의 적용을 받는 주체(연구기관, 문화유산기관)와 TDM 수행의 목적(학술연구)을 제한하지만, 계약이나 사후 유보에 의해 동 조항의 적용을 우회할 수 없는 강력한 효력을 지닌다. 이와 달리, 제4조는 주체나 목적의 제한 없이 모든 TDM을 허용하되 계약 및 옵트아웃에 의한 적용 배제의 길을 열어둔다는 점에서 제3조와 구별된다. 제3조는 ‘특정 TDM 면책조항’으로, 제4조는 ‘일반 TDM 면책조항’으로 부를 수 있는데,⁴⁸⁾ 이와 같은 이원적 체계는 TDM 활용을 둘러싼 다양한 이해관계를 조화시키기 위한 고려에 기반한 것으로 이해된다. 또한 위 이원적 체계는 DSM 지침 초안의 제3조에서 면책대상을 확장하고, 제3조와 독립한 제4조의 별개 면책 범위를 추가한 결과로서, 입법의도 실현의 면에서 긍정적으로 평가할 수 있다.⁴⁹⁾

46) DSM Directive, recital 14.

47) DSM 지침 제7조 제2항 제2문에서 제3조 내지 제6조에 적용되는 것으로 명시한 유럽연합 정보사회지침(Directive 2001/29/EC) 제6조 제4항 제1, 3, 5단락이 권리자에 의해 자율적으로 적용될 수 있는 기술적 보호조치에 관해 규정하고 있다.

48) Charlotte Gerrish & Anders Molander Skavlan, ‘European Copyright Law and the Text and Data Mining Exceptions and Limitations’, *Stockholm Intellectual Property Law Review*, vol. 2(2), 2019, pp. 62, 64.

49) DSM 지침 제3조와 제4조의 적용 범위 및 요건 비교는 다음 자료의 표 참조. Alina Trapova & João Pedro

그러나 이와 같은 의의에도 불구하고 DSM 지침의 TDM 면책규정에 대해서는 발효 이후 현재까지도 그 실효성에 관해 여러 비판이 제기되고 있다. 첫째 비판은 제3조의 적용 범위가 지나치게 협소하다는 것이다. TDM 면책규정은 DSM 지침 제정 논의 초기부터 유럽연합 내 통일적 규율 마련을 위해 강행규정의 도입을 중심으로 논의되었고,⁵⁰⁾ 이러한 취지를 직접적으로 실현하는 규정이 바로 제3조이다. 그런데 제3조는 TDM 면책의 범위를 TDM의 주체, 목적으로 이중 한정(double limitation)하고, 나아가 ‘적법한 접근권한’을 요구한다. 주체·목적의 이중 한정에 관해서는, (i) 타국 법제의 비상업적 목적 제한(non-commercial limitation)에 비해 더욱 강한 제한으로⁵¹⁾ 미국 등 개방형 면책조항 도입국과의 TDM 적용 환경의 간극을 메울 수 없는 점,⁵²⁾ (ii) 이미 주체를 한정한 상황에서 목적을 추가로 제한하여 TDM 프로젝트를 수행하는 연구기관의 법적 불안정과 그에 따른 거래비용 증가를 유발하는 점,⁵³⁾ (iii) 학술연구가 더 이상 전통적인 연구기관에 의해서만 수행되지 않는 연구계의 현실을 반영하지 못하는 점⁵⁴⁾ 등이 문제로 지적된다. 다만, 적어도 DSM 지침은 ‘비상업적 목적’을 TDM 면책의 명시적 요건으로 규정하지 않기 때문에 연구기관 또는 문화유산기관이 학술연구 목적을 위해 민간사업자와 파트너십을 이루거나 민간사업자의 펀딩을 통해 혹은 민간의 개발도구를 이용해 TDM 프로젝트를 진행하는 경우는 TDM 면책 대상에 여전히 포함될 수 있고,⁵⁵⁾ 이와 관련된 구체적인 입법적 선택은 각 회원국의 결단에 맡겨진다고 한다.⁵⁶⁾

제4조에 관해서는 제3조의 적용을 받을 수 없는 개인연구자, 언론사, 정책입안자, 사기업 등에 대해서도 TDM 면책의 여지를 확보한다는 점과 옵트인이 아닌 옵트아웃 방식을 채택했다

Quintais, ‘The UK government moves forward with a text and data mining exception for all purposes’, *Kluwer Copyright Blog*, Aug. 24, 2022, available at: <<http://copyrightblog.kluweriplaw.com/2022/08/24/the-uk-government-moves-forward-with-a-text-and-data-mining-exception-for-all-purposes/>>.

50) European Commission, ‘Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market’, COM/2016/0593 final - 2016/0280 (COD), Explanatory Memorandum, Sep. 14, 2016, p. 8.

51) Margoni & Kretschmer, *supra* note 2, p. 695.

52) Christopher Geiger, Giancarlo Frosio & Oleksandr Bulayenko, ‘Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU’, in *Propiedad Intelectual y Mercado Único Digital Europeo* (C. Saiz Garcia & R. Evangelio Llorca eds., 2019), pp. 50-51.

53) *Id.*, p. 53; Séverine Dusollier, ‘The 2019 Directive on Copyright in the Digital Single Market: Some progress, a few bad choices, and overall a failed ambition’, *Common Market Law Review*, vol. 57(4), 2020, p. 987, *cited from* Christopher Geiger, ‘The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Fundamental Right to Research and the Failed Text and Data Mining Limitations in the CSDM Directive’, PIJIP/TLS Research Paper Series 2021, No. 66, Apr. 2021, p. 9, n. 34.

54) Rossana Ducato & Alain Strowel, ‘Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out’, CRIDES Working Paper Series No. 1/2021, Feb. 2021, p. 12.

55) DSM Directive, recital 11, 12.

56) Martin R.F. Senftleben, ‘Study on EU Copyright and Related Rights and Access to and Reuse of Data’, European Commission, Mar. 2022, pp. 38-39.

는 점이 의의로 평가되기는 하나, 권리자에 의해 동 규정의 적용이 손쉽게 배제될 수 있어 면책사유로서의 실효성에 관해 의문이 제기된다.⁵⁷⁾ 한편, 제3조와 제4조에 공통되는 '적법한 접근권한'은 DSM 지침 제정 과정에서 참조된 영국, 프랑스, 독일 등의 규정에 모두 포함된 요건이었고, DSM 지침 제정의 배경이 된 유럽연합의 정보사회지침과 데이터베이스지침에서도 요구하는 사항이기에 새롭지 않다.⁵⁸⁾ 그러나 이에 관해서도, (i) TDM 연구를 사적 규제(private ordering) 하에 두는 것이라거나, (ii) 대규모 데이터베이스의 라이선스를 확보할 경제력을 갖춘 연구기관 혹은 부유한 회원국을 우대하여 불공정한 경쟁환경을 조성하고⁵⁹⁾ 영세한 연구기관과 스타트업 등에 의한 혁신을 제약한다는 점,⁶⁰⁾ (iii) 복수의 기관이 컨소시엄 형태로 수행하는 연구 프로젝트에 장애가 된다는 점⁶¹⁾ 등의 비판이 있다. 마지막으로 제3, 4조에 TPM 적용을 허용한 점에 대해서는 권리자가 데이터 접근을 사실상 배제할 가능성,⁶²⁾⁶³⁾ 회원국들 간 통일적 규율을 저해할 가능성⁶⁴⁾ 등의 비판이 있다.

(3) 독일, 프랑스 등의 국내법제화

DSM 지침의 발효와 함께 유럽연합 회원국들은 2021년 6월 7일까지 동 지침을 국내법에 반영하는 작업을 완료할 의무 아래 놓이게 되었다. 그에 따라 국내 문헌들에서 TDM 면책에 관한 해외 입법례로 소개되어 온 독일과 프랑스도 저작권법령을 개정하였다.⁶⁵⁾ 개정 내용은 DSM 지침 제3, 4조와 거의 동일하지만, 이를 최소한의 기준으로 삼아 각국 기존 법령의 틀 안에서 국내법제화한 것이기 때문에 그 형식이나 내용에서 약간의 차이가 보인다.

독일은 DSM 지침의 국내법 전환 이행 시한인 2021년 6월 7일에 맞춰 저작권법(UrhG)을 개정, 시행하였다.⁶⁶⁾ DSM 지침 제3조의 내용은 기존의 TDM 면책조항인 제60d조에 반영하였고,

57) Margoni & Kretschmer, *supra* note 2, p. 686.

58) Senftleben, *supra* note 56, p. 43.

59) 특히, TDM 분석에 적용된 데이터의 규모와 다양성이 커질수록 TDM 분석 결과의 신뢰성이 증가한다는 점에서, 경제적 능력에 따라 시장에서의 지위 혹은 연구능력의 격차가 고착될 수 있음을 시사한다.

60) Geiger et al., *supra* note 52, p. 54; Gerrish & Skavlan, *supra* note 48, p. 62.

61) Senftleben, *supra* note 56, pp. 44-45.

62) Margoni & Kretschmer, *supra* note 2, p. 695; Gerrish & Skavlan, *supra* note 48, p. 62-63.

63) 계약에 의한 제3조 적용 배제 금지(DSM 지침 제7조 제1항)를 우회하거나, TDM 면책규정의 적용 가능성을 사전적으로 봉쇄하는 결과가 되는 것을 뜻한다.

64) Geiger et al., *supra* note 52, p. 56.

65) 유럽연합 회원국들의 DSM 지침 국내법 전환 현황은 아래의 웹사이트들에서 확인할 수 있다(각 마지막 방문: 2022. 11. 7.). <<https://eur-lex.europa.eu/legal-content/EN/NIM/?uri=celex:32019L0790>>; <<https://www.create.ac.uk/cdsm-implementation-resource-page/>>; <<https://www.notion.so/DSM-Directive-Implementation-Tracker-361cfae48e814440b353b32692bba879>>.

66) TDM 지침의 국내법 전환에 따른 독일의 2021년 저작권법 개정 내용을 소개한 국내 문헌으로는 최상필, "빅데이터의 분석과 활용을 위한 TDM의 저작권적 쟁점과 입법론", 『계간 저작권』 제35권 제1호, 한국저작권위원회 (2022), 69-71면이 있다.

TDM의 정의 및 일반 면책조항인 DSM 지침 제4조의 내용은 신설조항인 제44b조에 규정했다. 개정 전 규정 및 DSM 지침과 비교해 특기할 점은 다음과 같다. 학술연구 목적의 TDM 면책은, 연구기관에 소정의 영향력을 행사하고 그 연구 결과에 우선적으로 접근할 수 있는 민간사업자와 협력하는 경우에는 적용되지 않는다(제60d조 제2항 제2문).⁶⁷⁾ 특정 TDM 면책조항의 적용을 받는 자로, DSM 지침에서 정한 연구기관, 문화유산기관 외에 “비사업적 목적을 추구하는 개인연구자”를 명시한다(제60d조 제3항 제2호). DSM 지침 내용을 반영해 개정 전 규정의 “비사업적 목적(nicht-kommerzielle Zwecke)”을 DSM 지침상의 주체 및 목적의 제한으로 대체함에 따라 복제에 관해서는 비사업성을 행위의 목적으로 요구하지 않지만, 특정한 목적을 위해 일정 범위에서⁶⁸⁾ 복제물을 공중이용제공하는 것에는 “비사업적 목적의 추구”를 요구하며 그러한 공중이용제공은 해당 목적이 종결되면 지체없이 종료되어야 함을 규정한다(제60d조 제4항). 또한 개정 전의 규정과 달리 학술연구 목적의 TDM 과정에서의 복제를 보상금지급 대상에서 제외하였다(제60h조 제2항 제3호).⁶⁹⁾

프랑스는 DSM 지침 발효 후 단계적으로 지적재산권법(Code de la propriété intellectuelle)을 개정하였다. 언론사 저작권접권에 관한 제15조는 유럽연합 회원국 중 가장 빠르게 도입한 반면,⁷⁰⁾ TDM 면책조항인 제3조와 제4조는 국내법 전환 이행 시한이 지난 2021년 11월 24일에야 이를 반영한 지적재산권법 개정법률을 공포하였다.⁷¹⁾ 프랑스 지적재산권법은 L122-5조 제10호에서 TDM 목적으로 ‘공표된 저작물’을 복제하는 행위를 면책대상으로 정하는 한편, L122-5-3조에서 DSM 지침상의 면책사유들을 규정하고 있다. L122-5-3조 제 I 항에는 TDM의 정의를, 제 II 항과 제 III 항에는 각각 DSM 지침 제3조, 제4조와 거의 동일한 내용을 규정한다. 제 II 항의 제1단락 말미 및 제2단락에서 민간사업자와 비영리 파트너십(partenariat sans but lucratif) 하에서 행하는 TDM이 학술연구 목적의 면책 대상임을 명시하면서 해당 민간사업자나 그 주주, 관련인이 연구결과에 대한 접근 특권을 갖는 경우에는 면책사유가 적용될 수 없다고 정하고 있는데, 이는 DSM 지침 전문 제11, 12항의 내용을 반영한 것이다. 한편 L211-3조 제

67) DSM 지침 전문 제12항의 내용을 반영한 것이다.

68) 공동학술연구를 위한 특정 제한된 범위의 사람들, 학술연구의 품질 모니터링을 위한 독립한 제3자(제60d조 제4항 제1, 2호).

69) 제60h조 제1항에서 “본 절에 따른 이용”에 대해 공정한 보상금을 수취할 권리를 규정하는데, DSM 지침 제4조에 대응하는 내용을 규정한 제44b조는 위 제60h조와 같은 절에 포함되어 있지 않다. 이에 따라 비-학술연구 목적의 TDM 이용도 보상금지급의무의 대상에 포함되지 않는다. 그러나 이 경우는 권리자의 옵트아웃이 가능하므로 라이선스를 통한 경제적 보상 수취 기회는 잔존한다. See Artur-Axel Wandtke & Winfried Bullinger, *Praxiskommentar Urheberrecht* (6th ed.), 2022, UrhG § 44b ¶ 11.

70) AFP, “We can be proud: France becomes first country to adopt EU copyright reform, *THE LOCAL*, Jul. 24, 2019, available at: <<https://www.thelocal.fr/20190724/french-parliament-adopts-eu-copyright-reform/>>.

71) Ordonnance n° 2021-1518 du 24 novembre 2021 complétant la transposition de la directive 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d’auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE.

8호는 TDM 목적의 복제에 대해서는 실연자, 음반제작자, 언론사 등 저작권접권자의 권리도 미치지 않는 것으로 규정하여 TDM 면책규정의 실효성을 확보하고 있다.

독일과 프랑스의 TDM 면책규정은 데이터베이스에 대해서까지 그 적용 범위가 확장되어 있고,⁷²⁾ 이 점은 DSM 지침과 공통된다.

나. 일본 - 표현된 사상·감정의 비향수적 이용 및 경미한 정보해석 이용에 대한 면책

일본은 2009년에 이은 2018년 저작권법 개정을 통해 기계학습 등에 수반되는 데이터 복제 등에 관계된 저작권 침해 면책규정을 이른바 ‘유연한 권리제한 규정’의 체계로 정비하였다. 일본의 2018년 개정 저작권법에서 TDM 면책과 관련이 있는 규정은 제30조의4와 제47조의5이다. 제30조의4는 저작물에 표현된 사상이나 감정을 자신 또는 타인이 향수하는 것을 목적으로 하지 않는 경우 일반에 적용되는 면책규정이다. 여기서 ‘향수(享受)’란 저작물을 인식하고 지적 혹은 정신적 욕구를 충족시키는 이용을 가리킨다. 이와 같이 ‘표현된 사상 또는 감정의 향수’ 여부를 기준으로 일반규정에 가까운 면책 범위를 정한 것은, 저작물의 경제적 가치는 저작물에 표현된 사상·감정 향수의 대가이며 따라서 그 향수를 목적으로 하지 않는 이용은 형식적으로는 저작물의 이용일지라도 저작권법이 보호를 상정하고 있는 본래적 이용형태로 볼 수 없다는 사고에 바탕한 것이라고 한다.⁷³⁾ 제30조의4 각 호는 표현된 사상·감정의 비향수적 이용 형태를 예시하고 있는데, 그중 TDM에 직접 관계된 것이 제2호 ‘정보해석(解析) 용도로의 제공’이다. 정보해석은 “다수의 저작물 또는 기타 대량의 정보로부터 해당 정보를 구성하는 언어, 음, 영상, 기타 요소에 관한 정보를 추출, 비교, 분류 그 밖의 분석을 실시하는 것”으로서, 대량의 소재에 대한 비교, 분석 등을 통해 새로운 정보를 생성하는 TDM의 의미에 정확히 부합한다. 본 조는 미공표 저작물과 영리 목적 이용에도 적용된다.⁷⁴⁾

한편, 일본 저작권법 제47조의5는 컴퓨터를 이용한 정보처리를 통해 새로운 지식·정보를 창출함으로써 저작물의 이용촉진에 이바지하는 행위를 하는 자가 일정한 행위에 부수하여 경미하게 저작물을 이용하는 것을 면책대상으로 규정하고, 그중 하나로 정보해석의 실시 및 결과 제공(제1항 제2호)을 열거한다. 제30조의4에 따른 비향수적 이용에 해당하지 않더라도 TDM 등에 부수한 경미한 이용에 대해서는 저작권 침해가 되지 않도록 정한 것이다.⁷⁵⁾ 본 조의 적용 대상이 되는 저작물은 공중에 제공 혹은 제시된 저작물이며, 개정 전 제47조의7에서 “통계적인”을 삭제하였으므로 정보해석 방법에 제한이 없다. 제2호의 정보해석을 위한 경미한 이용에

72) UrhG, § 87c (1) 4, 5; Code de la propriété intellectuelle, Art. L342-3, 6°.

73) 中山信弘, 「著作権法(第3版)」, 有斐閣, 2020, 382면.

74) 위의 책, 383면.

75) 고수윤, “디지털저작물에 있어 일본의 유연한 권리제한규정(柔軟な權利制限規定)의 시사점 - 비향수이용 규정 (일본 저작권법 제30조의4)을 중심으로 -”, 「계간 저작권」 제32권 제3호, 한국저작권위원회(2019), 39면.

해당하는 경우로는 ① 서적 등에 포함된 대량의 정보 중에서 필요한 요소를 추출해 일정한 특징이나 경향 등을 찾아내는 해석의 결과를 제공하는 행위, ② 특정 정보에 대한 평판이 게재된 블로그, 신문, 잡지 등의 내용을 분석해 그 결과를 제공하는 서비스에서 그와 함께 블로그 등의 일부분을 제공하는 행위, ③ 대량의 논문이나 서적을 디지털화하여 검증 대상 논문과의 문장 일치도를 해석하는 표절 정보제공 서비스에서 표절된 논문의 일부분을 제공하는 행위, ④ 과거의 증례, 치료방법, 약효 등에 관한 다양한 정보에 기반하여 최적의 치료방법을 분석하고 그 결과 제공에 부수하여 문헌 등의 일부분을 제공하는 행위를 들 수 있다고 한다.⁷⁶⁾ 제47조의5 제1항 제2호는 정보해석 자체뿐 아니라 그 결과를 제공할 때의 저작물 이용까지 허용하는 점에서 제30조의4와 차이가 있다.⁷⁷⁾ 또한 제47조의5 제2항은 정보해석 등 제1항 각 호에 열거된 행위의 준비를 위한 정보의 수집, 정리, 제공을 하는 자(구체적 기준은 정령에서 정함)에 대하여 그 준비에 필요하다고 인정되는 한도에서 복제, 공중송신 또는 배포를 허용함으로써, 학습용 데이터셋 제공사업, 정보해석 위탁사업 등 분업 형태의 TDM 프로젝트 진행을 제도적으로 뒷받침하고 있다.⁷⁸⁾

일본 저작권법 제30조의4 및 제47조의5는 이용 방법에 제한을 두지 않아 복제 기타 모든 방법으로의 이용이 허용된다. 다만 정보해석 등 법정된 목적에 “필요하다고 인정되는 한도에서” 이용할 수 있다는 한계와, 저작물의 종류·용도·이용형태에 비추어 저작권자의 이익을 부당하게 해치는 경우에는 면책되지 않는다는 한계를 공통적으로 설정하고 있다. 제30조의4에 따라 작성된 복제물은 양도에 의해 공중에 제공할 수 있으나, 사상이나 감정의 향수를 목적으로 하는 경우에는 그 복제물의 양도 등 모든 이용이 제한된다(제47조의7, 제49조).

다. 영국과 싱가포르 - 목적의 제한이 없는 TDM 면책규정 도입 움직임

(1) 영국

영국의 지적재산권법인 CDPA(Copyright, Design and Patent Act 1988)는 일시적 복제(제28A조), 연구 및 사적학습(제28B조), 비평·검토 등을 위한 공정취급(제30조)을 비롯하여 제28조부터 제76A조까지 침해 면책사유들을 상세하게 규정하고 있다. 그럼에도 불구하고 TDM과 관련된 개별 면책조항을 2014년 6월 1일 도입, 시행하였다. CDPA 제29A조는 (i) 저작물에 대한 적법한 접근권한을 가진 자가 (ii) 비상업적 연구만을 목적으로(for the sole purpose of research for non-commercial purpose) (iii) 컴퓨터 분석(computational analysis)을 위해 행하는 복제에 대하여 (iv) 현실적으로 가능한 범위에서 그 복제물에 충분한 출처명시(sufficient

76) 小倉秀夫·金井重彦(編), 『著作権法コンメンタルⅡ(改訂版)』, 第一法規, 2020, 346-347면.

77) 위의 책, 347면.

78) 위의 책, 348면.

acknowledgement)를 하는 경우에 저작권 침해가 되지 않는 것으로 규정한다(제1항). 단, 그 복제물을 저작권자의 허락 없이 타인에게 양도하거나 전술한 목적 이외로 이용하는 때에는 그러하지 않다(제2항). 또한 본 조에 의해 저작권 침해가 되지 않는 복제를 금지하거나 제한하는 계약 조항은 무효로 한다고 규정함으로써 그것이 강행규정임을 분명히 하고 있다(제5항).

위와 같은 CDPA 제29A조는 TDM 면책규정의 적극적 입법 사례로 평가된다. 그러나 영국 지적재산권청(IPO)은 2021년 10월 29일, 국가 AI 전략에 따라 영국의 AI 개발·활용 촉진을 위한 입법적 개선의 일환으로 지적재산권 보호제도에 관한 공개의견수렴(public consultation) 절차를 개시했다.⁷⁹⁾ 의견수렴 대상은 크게 세 가지로, 컴퓨터 생성 저작물(computer-generated works, 'CGW') 보호, TDM 면책, AI 고안 발명(AI-devised invention)의 특허 보호였는데, 그 중 TDM 면책과 관련해 제시된 선택지는 현행 제도 유지(옵션 0), 현행 면책조항을 유지하면서 표준 라이선스, 확장된 집중관리제도, 윤리강령 등을 통해 라이선스 환경을 개선하는 방안(옵션 1), 상업적 연구로 TDM 면책의 범위를 확장하는 방안(옵션 2), TDM의 목적에 관한 제한을 제거하되 기계적으로 인식 가능한 옵트아웃을 허용하는 방안(옵션 3), 목적을 비롯한 일체의 제한이 제거된 포괄적인 TDM 면책규정을 도입하고 옵트아웃도 허용하지 않는 방안(옵션 4)이었다.⁸⁰⁾

영국 IPO는 수 개월에 걸쳐 다양한 이해당사자들의 의견을 수렴한 뒤 2022년 6월 28일 정부의 입장을 발표했다. CGW와 AI 고안 발명에 관해서는 현행 제도를 유지하되, TDM에 관해서는 목적의 제한이 없는 신규 TDM 면책조항을 도입한다는 것이었다.⁸¹⁾ IPO는 상업적 목적을 포함한 모든 TDM에 대한 저작권 및 데이터베이스 침해책임을 면제하고 옵트아웃을 불허하는 '옵션 4'에 따른 법제 개선이 연구자, AI 개발자, 소기업, 저널리스트, 소비자 및 일반 공중의 이익을 증진할 것이라고 밝혔다. 또한 이 같은 신규 면책조항이 AI 시대 영국 저작권법제의 혁신친화성과 경쟁력을 세계 최고 수준으로 끌어올릴 것이라는 비전을 제시했다.⁸²⁾ 다만, 권리자의 적절한 이익 수취 기회를 보호하기 위해 '적법한 접근(lawful access)' 요건은 유지될 것이며, 시스템의 무결성과 보안을 확보하기 위한 권리자의 조치도 허용할 것이라고 한다.⁸³⁾ 상업적 목적에 관한 제한과 옵트아웃을 배제하는 영국 TDM 면책입법은 유럽연합 DSM 지침의 그것에 비해 연구 및 신기술 투자를 촉진할 것이라는 기대를 받고 있다.⁸⁴⁾ 영국 IPO는 이와 같

79) UK IPO, 'Consultation stage impact assessment on Artificial Intelligence and Intellectual Property', RPC-BEIS-IPO-5101(1), Oct. 29, 2021, pp. 7-8.

80) *Id.*, pp. 13-14.

81) UK IPO, 'Consultation outcome - Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation', Jun. 28, 2022, paras. 84-87, available at: <<https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation>>.

82) *Id.*, paras. 58-62.

83) *Id.*, para. 61.

은 과감한 계획이 브렉시트(Brexit)에 의해 확보한 입법적 융통성을 최대한 발휘한 결과라고 자평하나,⁸⁵⁾ 달리 보면 브렉시트 이후 유럽 권역 내 타 국가들과의, 나아가 전 세계적인 AI 규범 경쟁에서 헤게모니를 선취하려는 시도로 평가할 여지도 있다.

(2) 싱가포르

영국의 포괄적 TDM 면책규정 도입 계획에는 그보다 한 발 앞서 진행된 싱가포르의 입법례가 직간접적으로 참고가 된 것으로 보인다. 싱가포르 정부는 디지털 시대의 권리 보호와 창작물 이용환경 개선을 위한 논의를 2016년 경부터 시작했고, 약 3년 간의 집중적인 검토 및 공개 의견수렴을 거쳐 2019년 1월 17일 TDM 개별 면책규정의 도입을 포함한 다방면에 걸친 저작권법 개정 계획을 발표했다.⁸⁶⁾ 2021년 9월 싱가포르 의회를 통과하여 2021년 11월 21일부터 시행된 개정 저작권법은 “컴퓨터 데이터 분석(computational data analysis)”을 저작물이나 실연의 복제물(이하 ‘소재’)로부터 정보 또는 데이터를 식별, 추출, 분석하는 컴퓨터프로그램의 사용, 및 특정 정보나 데이터 유형에 관계된 컴퓨터프로그램의 기능을 개선하기 위해 그 유형의 예시로서 소재를 이용하는 행위로 정의하고(제243조), 이러한 의미의 컴퓨터 데이터 분석 및 그 준비작업을 자유이용이 허용되는 행위로 규정한다(제244조 제2항 제a호). 제244조 제2항은 동 조의 면책대상이 되기 위해 요구되는 추가 요건들을 규정한다. 컴퓨터 데이터 분석 이외의 목적으로 복제물이 이용되어서는 안 되며(제b호), 복제물을 송신 등의 방법으로 제3자에게 제공하는 것이 분석 결과의 검증 혹은 해당 목적에 관계된 합동 연구나 학습을 위한 범위로 제한되어야 하고(제c호), 행위자가 소재에 대한 적법한 접근권한(lawful access)을 가져야 하며(제d호), 소재가 침해물이 아닌 경우, 또는 침해물이더라도 행위자에게 악의·과실이 없는 경우이거나 침해물이 전술한 목적에 필수적이며 그 이외의 목적으로 이용하지 않는 경우여야 한다(제e호).⁸⁷⁾

제244조 제2항에 따라 허용되는 행위에는 저작권법에 의해 보호되는 소재의 복제 외에 그 복제물을 저장하고 보유하는 행위가 포함되며(제3항), 위 복제물은 컴퓨터 데이터 분석 결과의 검증 또는 해당 목적에 관계된 합동 연구·학습을 위한 범위 안에서 제3자에게 송신 등의 방법으로 제공하는 것이 허용된다(제4항). 컴퓨터 데이터 분석을 위한 복제 및 그 제공은 ‘발행(publication)’으로 간주되지 않고(제60조 제1항 제h호, 제244조 제5항 제a호), 허용되는 컴퓨터

84) Trapova & Quintais, *supra* note 49.

85) UK IPO, *supra* note 81, para. 60.

86) Ministry of Law Singapore, ‘Singapore Copyright Review - Enhancing Creators’ Rights and Users’ Access to Copyrighted Works’, Press Releases, Jan. 17, 2019, available at: <<https://www.mlaw.gov.sg/news/press-releases/singapore-copyright-review-report-2019>>.

87) 2021년 11월 21일 시행된 싱가포르 개정 저작권법의 원문은 다음 웹사이트에서 확인 가능하다.

<<https://sso.agc.gov.sg/Acts-Supp/22-2021/>> (마지막 방문: 2022. 11. 8.).

데이터 분석을 배제하거나 제한하는 계약 조건은 무효이다(제187조 제1항 제c호).

싱가포르 저작권법에서 주목할 점은 영국 CDPA와 유사하게 제5장(제183조 내지 제303조)에서 다양한 종류의 저작권 침해의 면책사유들을 규정하며, 그 중 특히 제2절(제190조 내지 제194조)에서 미국 저작권법 제107조와 유사한 공정이용(fair use)에 대해 규정하고 있다는 점이다. 이처럼 포괄적, 개별적 면책규정들이 모두 갖추어져 있음에도 불구하고 TDM 개별 면책조항을, 그것도 상업적/비상업적 목적의 구분이나 주체의 제한이 없는 상당히 포괄적인 강행규정으로 도입하였다는 점이 특기할 만하다. 공정이용 규정이 사안별로 적용될 수밖에 없고 그 적용 범위에 관한 명확한 판례가 부재한 상황에서 데이터 분석을 필요로 하는 산업계에 안정감을 주기 위한 취지였다고 한다.⁸⁸⁾

라. 미국 - 공정이용(fair use) 조항에 의한 사법적 해결

미국은 TDM 관련 개별 면책조항을 마련하지 않고 공정이용(fair use) 조항의 해석을 통해 탄력적이고 유연한 해결을 도모하고 있다. 미국 연방저작권법 제107조의 내용 및 관련 사례들은 국내외 다수 문헌들에서 이미 충분히 소개하고 있으므로 중복 해설은 생략한다. 다만 자주 언급되는 사례들은 TDM 행위 자체에 관한 사안이라기보다 대체로 TDM과 관련성이 있는 행위 혹은 TDM 분석과 유사한 성격의 행위에 관한 것이라는 점을 유념하여야 한다는 점을 지적해 두며,⁸⁹⁾ 아울러 주요 참고 사례로 언급되는 구글북스 판결⁹⁰⁾에서 TDM에 관해 직접 언급하는 대목이 있어 소개하고자 한다. 도서관들의 장서를 디지털화하여 검색 이용자에게 제공하는 구글북스 프로젝트에 관해 미국 작가협회(Authors Guild) 측이 제기한 저작권 침해 소송에서 제2순회항소법원은 구글이 해당 프로젝트를 위해 디지털 사본을 서버에 저장하고 검색 이용에 제공한 행위에 대해 공정이용으로 판단하면서, 사건의 배경 부분에서 “[구글의] 검색 엔진이 텍스트 마이닝 혹은 데이터 마이닝으로 알려진 새로운 연구 기법을 가능케 한다”고 언급하였다.⁹¹⁾ 이러한 판시는 TDM 분석을 침해적 이용으로 보는 견지에서 나오기 어려운 판시로, TDM 분석행위 그 자체에 관해서도 미국 연방법원이 공정이용 해당성을 적극적으로 인정할 가능성을 시사한다.

학계에서도 TDM 분석 목적의 저작물 복제 등이 공정이용으로 판단될 가능성을 긍정하는 것이 대체적인 견해이며, 주로 제1요소(저작물 이용의 목적과 성격)에 관한 분석에 있어 TDM 과정에서의 저작물 이용이 새로운 정보나 가치 또는 의미를 창출하는 이른바 변형적 이용

88) 홍승기, “데이터마이닝 면책 입법 방향에 대한 의문”, 『경영법률』 제32집 제4호, 한국경영법률학회(2022), 29-30면.

89) 미국 연방저작권법의 공정이용 조항의 내용과 TDM 면책에 관해 시사점을 제공하는 판례들을 소개한 국내 문헌으로, 박성호, 앞의 글(주 6), 55-63면 참조.

90) Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).

91) *Id.* at 209.

(transformative use)에 해당하고, 제4요소(잠재적 시장에 미치는 영향)에 있어서도 TDM의 결과가 원저작물의 수요를 대체하기가 어렵다는 점을 근거로 한다.⁹²⁾⁹³⁾ 일각에서는 TDM의 성격을, 복제된 표현을 인간이 표현으로서 향수, 감상, 이해하게 하는 것을 목적으로 하지 않는다는 의미의 '비표현적 이용(non-expressive use)'으로 규정하고, 그러한 비표현적 이용이 공정이용 분석의 제1요소인 '이용의 목적·성격'에 관계된 변형적 이용을 범주화한 지표로서 TDM 행위의 공정이용적 면모를 묘사한다는 견해가 있다.⁹⁴⁾⁹⁵⁾ '비표현적 이용'은 '비소비적 이용(non-consumptive use)'으로도 불리는데,⁹⁶⁾ 이는 변형적 이용의 한 속성을 생산적 이용(productive use)으로 표현하는 것과 일맥상통한다. 비표현적 이용 혹은 비소비적 이용이라는 기준은 기계학습 또는 TDM이 통상적인 저작물 이용과 구분되는 변형성의 일면을 포착한 것이지만, 컴퓨터 분석 목적으로 저작물을 데이터베이스화하여 계속 보관하는 경우를 해명하는 데에 한계가 있다는 비판,⁹⁷⁾ 대량의 표현 데이터를 학습하여 동종의 표현을 만들어내는 생성형 AI(generative AI)⁹⁸⁾와 같은 '표현적' 이용례를 외면한 것이라는 비판, 지식 또는 정보의 획득을 위한 생산적 이용은 저작물의 '소비'를 전제로 하는 점에서 비소비적 이용은 불충분한 기준이라는 비판⁹⁹⁾ 등이 존재한다.

2. 주요국 법제 정리 및 비교

이상 살펴본 바와 같이 TDM 면책 관련 법제를 마련한 주요국들은 권리자-이용자 관계를 포

92) Carroll, *supra* note 15, pp. 935-959; Jessica L. Gilotte, 'Copyright Infringement in AI-Generated Artworks', *U.C. Davis Law Review*, vol. 53, 2020, pp. 2679-2690.

93) TDM 사안에서 공정이용 판단의 제2요소(저작물의 성격)와 제3요소(이용의 양과 상당성)가 큰 역할을 하지 못하는 이유는, TDM이 적용되는 연구분야에 따라 이용되는 저작물의 성격이 각양각색일 수 있고, TDM 기술의 특성상 대량의 데이터가 분석 목적으로 이용되는 것이 불가피하기 때문이다.

94) Sag, *supra* note 9, pp. 301-302, 327.

95) '비표현적 이용'이 변형적 이용의 하위개념으로 논의되는 것과 달리, Lemley 교수 등은 기계학습을 위한 훈련데이터의 이용 등을 읽기(reading)와 대비되는 학습(learning)의 성격을 가지고 있다는 점에 착안하여, 저작물의 보호되는 요소를 보호되지 않는 요소의 학습 등을 목적으로 하는 '공정한 이용(fair learning)' 개념을 공정이용 제1요소의 평가에 있어 고려할 것을 제안한다. 이러한 관점에 의할 때 변형적 이용에 매몰되어 있는 공정이용의 이용 목적 분석을 다면화하여 그 판단에 타당성을 기할 수 있다는 것이다. See Mark A. Lemley & Bryan Casey, 'Fair Learning', *Texas Law Review*, vol. 99, 2021, pp. 753, 776-779, 782-785.

96) Sag, *supra* note 9, p. 321.

97) Carroll, *supra* note 15, p. 937.

98) 생성형 AI란 텍스트, 오디오, 이미지 등의 데이터를 학습하여 새로운 콘텐츠를 생성하는 인공지능 기술을 뜻하며, 구체적인 방법론으로 생성형 적대 신경망(generative adversarial network, 'GAN') 학습모델이 연구되고 있다. GAN 모델을 자연어처리 및 미술·음악·어문·무용 등 분야의 '인공지능 창작'에 적용한 사례들에 대한 소개는, 김혜창 외 3명, "인공지능 및 데이터베이스 등 기술변화에 따른 저작권법 개선 방안 연구", 한국저작권위원회 연구보고서, 2021. 12., 15-22면 참조.

99) Benjamin L.W. Sobel, 'Artificial Intelligence's Fair Use Crisis', *Columbia Journal of Law & the Arts*, vol. 41, 2017, pp. 71-75.

합한 TDM 생태계 내의 다양한 이해관계를 조정하기 위한 균형점을 저마다 달리 정하고, 각자의 방식으로 이를 구현하고 있다. 유럽연합의 DSM 지침은 학술연구 목적의 TDM과 그 밖의 TDM을 구분해 이원적으로 규율하면서, 전자에 대해서는 제3조에서 계약 또는 기술에 의한 적용 배제를 금지하여 권리제한의 실효성을 확보하는 한편, 후자의 경우 제4조의 일반규정을 통해 목적이나 주체의 제한 없이 면책을 넓게 인정하면서도 옵트아웃을 허용하여 이익 조정을 꾀하고 있다. 이와 같은 이원적 규율은 DSM 지침 제정 과정에서 TDM 이용환경이 상이한 회원국들의 상반된 이해관계가 반영된 논의의 결과로서, 초안에 비해 TDM 면책의 범위를 일부 확장했음에도 불구하고 제3조의 주체 및 목적상의 제한이 면책의 범위를 크게 축소하는 점, 제4조의 옵트아웃이 면책 적용의 적법한 우회 수단이 되는 점 등에서 여전히 TDM 면책규정으로는 ‘제한적(restrictive)’이라고 평가된다.¹⁰⁰⁾ 제3조와 제4조가 일정한 범위에서 TDM 분석 결과물의 보유를 허용하는 것은 TDM 이용 현실을 적절히 반영한 조치라고 할 수 있지만, TDM을 위한 제3자의 데이터셋 제공 등을 면책대상으로 명시하지 않은 것은 TDM 면책규정으로서의 의의를 반감시키는 요소라 할 수 있다.

한편, 영국이 준비중이고 싱가포르에서 이미 실행한 입법의 경우, 상업적/비상업적 목적을 구분하지 않고 옵트아웃조차 허용하지 않는 강행규정인 점에서 유럽연합 DSM 지침에 비해 면책의 범위가 한층 넓다. 영국이 2014년에 비상업적 연구 목적에 국한된 개별 TDM 면책규정을 도입한 지 채 10년이 되지 않아 이 같은 추가 입법을 시도하는 것은 기술 발전과 이용상황 변화의 속도를 시사한다. 다만 싱가포르의 입법례 및 그와 유사한 영국의 입법 계획은 ‘적법한 접근’ 요건을 잔존시킴으로써 권리자 이익 보호의 요구를 반영한 점에서 유럽연합 DSM 지침과 공통된다. 싱가포르 저작권법 제244조는 제3, 4항에서 컴퓨터 데이터 분석 과정에서의 복제뿐 아니라 그 복제물의 저장·보유, 나아가 일정한 목적 혹은 범위에서 제3자에게 제공하는 행위도 허용한다. 영국의 경우 입법안이 아직 구체화되지 않았고 현행 CDPA 제29A조 제2항이 복제물의 이전을 금지하고 있어, 법 개정을 통해 TDM에 후속하는 복제물의 보관·전송에 대해 면책 범위를 확장할지 현재로서 예단하기는 어렵다.

일본은 TDM 면책규정이 이원화되어 있다는 면에서 유럽연합 DSM 지침과 일견 유사해 보이지만 그 성격이 조금 다르다. 일본 저작권법 제30조의4와 제47조의5는, 이용의 목적이나 주체에 따라 서로 구분되는 면책 범위를 설정하여 관련 이해관계를 조정하고자 하는 DSM 지침의 이원화와는 다르게, 전자는 정보해석 그 자체를 위한 비향수적 이용을, 후자는 정보해석에 부수하는 경미한 이용을 각기 허용함으로써 면책의 공백을 최소화하기 위한 이원화로서의 성격을 갖기 때문이다. 다시 말해, TDM 면책의 중심이 되는 규정을 유럽연합 DSM 지침은 제3조로, 일본 저작권법은 제30조의4로 본다면, 전자에서 제4조는 면책 범위를 다소 넓히면서 옵

100) Trapova & Quintais, *supra* note 49.

트아웃을 통해 실효성을 제한하는 일종의 타협안인 데 반해, 후자에서 제47조의5의 역할은 제30조의4에 따른 이용에 부수하는 경미한 이용까지도 포섭함으로써 TDM 활동 전반에 대한 면책의 실효성을 배가시키기 위한 규정이라는 점에서 양자의 이원화는 성격이 서로 다르다. 한편 일본 저작권법의 위 두 조항은 형식상 공정이용과 같은 일반 면책조항은 아니지만, 그중 제30조의4의 면책요건을 저작물에 표현된 사상·감정을 향수하는 것을 목적으로 하지 않는 이용이라는 추상적인 문언으로 규정하여 일반규정에 근접한 매우 넓은 면책 범위를 갖는다. 또한 위 두 조항은 공통적으로 ‘적법한 접근’을 요구하지 않고 이용 방법에 제한이 없어, 정보해석 과정 전후의 이용행위의 허부는 결국 정보해석 목적상 “필요하다고 인정되는 한도”의 해석에 맡겨질 것으로 생각된다. 요컨대 일본 저작권법은 주체, 목적 및 이용 방법에 제한이 없고 저작권자에 대한 보상의무도 없다는 특징을 지니며, 이 점 때문에 개별 면책규정을 둔 법제 중 TDM에 가장 우호적인 입법례로 평가되기도 한다.¹⁰¹⁾¹⁰²⁾

개별 면책규정을 둔 타 국가·권역과 달리 미국은 ‘공정이용(fair use)’ 조항의 해석을 통해 유연하고 탄력적으로 TDM 관련 저작물 이용을 규율한다. 이용 방법이나 주체, 목적에 관한 명시적 제한이 없음은 물론, 적법한 접근에 관한 요건, 보상의무에 관한 규정도 없다. 다만 이용 목적이 공정이용 분석의 제1요소로서 고려될 뿐인데, 상업적 혹은 영리적 목적이 연방저작권법 제107조에 고려사항으로 명시되어 있기는 하나(“including whether such use is of commercial nature or is for nonprofit educational purposes”) 변형적 이용 사안에서 상업성이나 영리성의 여하는 공정이용의 인정에 있어 결정적인 역할을 하지 못하는 것으로 받아들여지기 때문에¹⁰³⁾ TDM의 속성에 비추어 이용 목적의 변형성을 인정할 수 있음을 전제로 한다면 상업적 혹은 영리적 목적의 유무가 TDM 이용의 공정이용 해당 여부를 실질적으로 좌우하기는 어려울 것이다.

아래 표는 주요국의 TDM 면책 관련 법제의 내용을 비교한 것이다(영국 CDPA 제29A조는 개정될 가능성이 있으므로 생략하였다).

101) 안효철, “빅데이터 활용과 인공지능 개발을 위한 Text and Data Mining”, 법률신문 2018. 12. 24.자 기사.

102) 일본 저작권법 제30조의4는 (i) 비상업적 목적과 상업적 목적의 TDM 활동을 포괄하고 연구기관뿐만 아니라 영리기업의 행위에도 적용되는 점, (ii) 권리자의 유보(오프아웃)에 의한 적용 배제를 예정하고 있지 않은 점, (iii) 모든 이용 방법을 허용하며 컴퓨터에 의하지 않은 데이터 분석까지도 포함하는 점, (iv) 적법한 접근 요건을 두지 않은 점에서 유럽연합 DSM 지침상의 그것에 비해 TDM 면책의 범위가 상당히 넓다고 평가된다. ‘일본은 기계학습과 TDM의 천국’이라고까지 표현할 수 있을 정도라고 한다. Tatsuhiko Ueno, ‘The Flexible Copyright Exception for ‘Non-Enjoyment’ Purposes - Recent Amendment in Japan and Its Implication’, *GRUR International*, vol. 70(2), 2021, p. 149.

103) Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183, 1204 (2021).

	유럽연합 DSM 지침		싱가포르 저작권법	일본 저작권법	미국 연방저작권법
	제3조	제4조	제244조	제30조의4	제107조
목적 한정	학술연구	X	X	X	X
주체 한정	연구기관, 문화예술기관	X	X	X	X
상업적 목적 이용 가능성	△	O	O	O	O
이용 방법	복제, 추출	복제, 추출	복제·준비작 업, 결과 검 증·합동연구 목적 송신	모든 이용 방법(사 상·감정 향수 목적 으로 하지 않는 경우 공중에 양도 가능)	모든 이용 방법
결과물 보관 요건	결과 검증 등 과학연구목적 & 보안조치	TDM 목적	-	-	-
옴트아웃 가능성	X	O	X	X	X
계약에 의한 배제 가능성	X	O	X	-	-
적법한 접근 요건	O	O	O	X	X
기타	-	-	침해물도 소 정 범위에서 이용 가능	‘저작물에 표현된 사상·감정의 비향 수적 이용’ 요건	각 Fair Use 고 려요소의 형량

3. 우리나라 저작권법 전부개정안

우리나라의 저작권법 소관부처인 문화체육관광부는 2020년 2월 ‘저작권법 비전 2030’을 발표하면서 저작권법 전부개정 추진 계획을 밝히고, 학계 전문가 등으로 구성된 연구반에서 개정 논의를 진행하여¹⁰⁴⁾ 2020년 9월 1일과 같은 해 11월 2일 순차로 TDM 면책규정(안)을 제안했다.¹⁰⁵⁾ 이후 위 11월 2일자 문화체육관광부 개정안의 TDM 면책조항을 일부 수정한 내용이

104) 문화체육관광부, 앞의 자료(주 40), 1면.

105) 문화체육관광부의 2020년 9월 1일자 및 11월 2일자 저작권법 개정안에 포함된 TDM 면책규정의 내용을 비롯해 입법 논의를 위한 선행연구 및 각 단계별 개정안에 대하여 공청회 등을 통해 제기된 의견들에 관해서는, 홍승기, 앞의 글(주 88), 4-10면; 유현우, “인공지능 시대 대응을 위한 「저작권법」 전부개정안 검토 - 정보분석을

도종환 의원이 대표발의한 저작권법 전부개정안¹⁰⁶⁾ 제43조에 다음과 같이 포함되어 국회에 계류중이다.

제43조(정보분석을 위한 복제·전송) ① 컴퓨터를 이용한 자동화 분석기술을 통해 다수의 저작물을 포함한 대량의 정보를 분석(규칙, 구조, 경향, 상관관계 등의 정보를 추출하는것)하여 추가적인 정보 또는 가치를 생성하기 위한 것으로 저작물에 표현된 사상이나 감정을 향유하지 않는 경우에는 필요한 한도 안에서 저작물을 복제·전송할 수 있다. 다만, 해당 저작물에 적법하게 접근할 수 있는 경우에 한정한다.

② 제1항에 따라 만들어진 복제물은 정보분석을 위하여 필요한 한도에서 보관할 수 있다.

위 개정안 제43조에서 면책 범위의 외연을 구체화하는 “컴퓨터를 이용한 자동화 분석기술” 및 “규칙, 구조, 경향, 상관관계 등의 정보를 추출”이라는 문언은 유럽연합 DSM 지침 제2조 제2항의 TDM의 정의에서 가져온 것으로 보이고, “다수의 저작물을 포함한 대량의 정보를 분석하여 추가적인 정보 또는 가치를 생성하기 위한 것”은 일본 저작권법 제30조의4 제2호에서 ‘정보해석’의 정의로 제시하는 “다수의 저작물 및 기타 대량의 정보에서 해당 정보를 구성하는 언어, 소리, 영상 기타 요소와 관련된 정보를 추출하여 비교, 분류 및 기타 해석을 하는 것”이라는 표현과 흡사하다. 또한 이용의 양태 및 범위를 한정하는 “저작물에 표현된 사상이나 감정을 향유하지 않는 경우에는 필요한 한도 안에서” 역시 일본 저작권법 제30조의4 본문을 거의 그대로 차용한 것인데, 다만 이용의 방법에 관해 아무런 제한을 두지 않는 일본 규정과 달리 복제·전송으로 한정하고 있다. 개정안 제43조 제1항 단서에서 요구하는 ‘적법한 접근’은 유럽연합 DSM 지침 제3, 4조와 현행 영국 CDPA 제29A조, 싱가포르 저작권법상의 요건과 동일하며, 제43조 제2항에서 복제물을 정보분석을 위해 필요한 한도에서 보관할 수 있도록 한 것은 DSM 지침 제4조 제2항을 참조한 것으로 보인다.

이처럼 2021년 1월 15일 발의된 저작권법 전부개정안 제43조는 TDM 이용에 관한 면책의 범위 및 요건을, 일부는 유럽연합 DSM 지침에서, 또 일부는 일본 저작권법 제30조의4에서 차용하여, 그 면책 범위를 DSM 지침에 비해서는 넓게, 그러나 일본, 싱가포르(및 영국 CDPA 제29A조의 개정입법 계획)에 비해서는 좁게 설정하고 있다. 이러한 개정안 제43조에 관해 (i) ‘사상 또는 감정의 비향유적 이용’의 의미가 불명확하고¹⁰⁷⁾ 그 해당 여부가 저작물의 표현형식에 따라 달라질 수 있다는 취지¹⁰⁸⁾의 비판이 있고, (ii) 상업적 이용까지 면책하는 점에 관해서

위한 복제·전송 허용 규정을 중심으로”, 『AIIA』 제44호, 지능정보산업협회(2021. 3.), 10-18면 참조.

106) 도종환 의원 등 13인 발의 저작권법 전부개정안(의안번호 제2107440호, 발의일 2021. 1. 15.).

107) 차상욱, 앞의 글(주 6), 42면.

108) 박성호 교수는 위 저작권법 전부개정안 제43조와 유사한 내용을 담은 2020년 9월 1일자 및 11월 2일자 각

는 저작권자와 이용자 간의 법익균형 및 국제적 조화 면에서 바람직하지 않다는 비판이 있다.¹⁰⁹⁾ (iii) 적법한 접근권한 요건에 관해서는, TDM 개별 면책규정 도입의 실익을 해한다거나¹¹⁰⁾ 접근통제 문제로 이어질 우려 내지 사물인터넷 등을 통한 대량 데이터 수집의 경우에는 그 데이터가 적법한 권한 있는 자에 의해 업로드된 것인지 확인하는 데에 현실적인 한계가 있다는 등¹¹¹⁾의 비판이 제기된다.¹¹²⁾

IV. TDM 면책규정 도입의 방향성

1. TDM 면책규정 도입에 관한 바람직한 논의 절차

전술한 바와 같이 우리나라 저작권법에 도입하기 위해 제안된 TDM 면책규정은 유럽연합 DSM 지침과 일본 저작권법 제30조의4의 특징을 뒤섞은 것이다. 초안에서 2차, 3차 수정을 거쳐 도종환 의원 등이 발의한 전부개정안 제43조에 이르는 과정에서 DSM 지침과의 유사성이 약화되고 일본 저작권법 규정 내용에 근접하게 되었는데,¹¹³⁾ 이와 관련하여 우리나라의 상황이 면밀하게 분석되거나 논의된 흔적은 발견되지 않는다. 그러나 TDM 면책규정의 도입을 현대적 저작물 이용환경에서 다양한 이해관계의 균형점을 탐색한다는 관점에서 바라본다면,¹¹⁴⁾ 타국 법제의 요소들을 차용하기에 앞서 우리나라의 연구계와 산업계에서의 각각의 TDM 기술의 발전상황과 수요, 면책의 범위나 수준이 국가 경쟁력과 시장 경쟁에 미칠 영향 등에 관한 조사·분석이 선행되고, 이를 바탕으로 우리나라의 산업적·정책적 상황에 가장 적합한 균형점이 탐색되고 설정되어야 한다. 영국은 현행 CDPA 제29A조의 비상업적 연구 목적 제한을 삭

문화체육관광부 개정안의 TDM 면책규정에 대하여, ‘사상 또는 감정의 비향수적 이용’의 해석·적용에 관해 일본에서조차 모호한 점이 많다는 비판이 제기되고 있으며, 컴퓨터프로그램과 같은 기능적 저작물에서는 프로그램의 실행 자체가 프로그램의 기능을 향유하는 것을 목적으로 하는 점에서 비향유에 해당하기 어려울 것이라고 지적한다. 박성호, “저작권법 전면 개정안 주요내용의 비판적 고찰”, 『정보법학』 제24권 제3호, 한국정보법학회(2020), 15-17면.

109) 정진근 외 8명, 앞의 글(주 3), 57-58면; 차상욱, 앞의 글(주 6), 34-36면.

110) 최상필, 앞의 글(주 66), 83면.

111) 차상욱, 앞의 글(주 6), 37-38면.

112) 그 밖에, ‘컴퓨터를 이용한 자동화 분석기술을 통한 대량의 정보 분석’ 요건에 관해 수작업으로 진행되는 데이터 전처리 과정에서의 복제가 위 요건을 충족할 수 있을지 의문이라는 비판(전웅준, “인공지능 관련 저작권 침해에 관한 시론”, 『경영법률』 제31집 제4호, 한국경영법률학회(2021), 269면), ‘추가적인 정보 또는 가치를 생성하기 위하여’라는 표현의 모호성으로 인해 소송에서의 증명과 판단에 대해 우려되는 점이 있다는 견해(이규호, “인공지능 학습용 데이터세트에 대한 저작권법과 부정경쟁방지법상 보호와 그 한계”, 『인권과 정의』 제494호, 대한변호사협회(2020), 101면)가 있다.

113) 박성호, 앞의 글(주 108), 13-16면 참조.

114) 오승중, 앞의 글(주 4), 477면 참조.

제하는 입법 계획을 도출하기에 앞서 자국 AI 기술의 개발·투자현황을 비용·편익 분석 방법 등을 통해 진단하고, 정책 목표를 설정한 뒤 법제 개선의 여러 옵션들에 대한 의견수렴과 검토를 거쳤다.¹¹⁵⁾ 영국 IPO는 의견수렴절차를 개시하면서 2021. 10. 29.에 발간한 영향평가서(Impact Assessment)에서 타 국가들과 비교할 때 영국의 AI 기술 채용 및 투자 현황, AI 기술이 영국 경제에 미치는 영향, AI 기술과 IP 보호제도 간 상관관계에 관한 경제적인 분석을 실시하였다. 나아가 이러한 사전 분석의 근거로 OECD 보고서, 회계법인과 컨설팅펌의 보고서, 유럽 내 전문조사기관의 보고서 등을 인용하여 제시한 점, 그에 기초해 나름의 정책 목표를 설정한 뒤 그 범위 내에서 의견수렴을 실시함으로써 의견수렴의 정책적 근거와 목표 및 범위를 구체화한 점이 참고가 된다.¹¹⁶⁾

우리나라에서는 TDM 면책규정 도입의 정책적 동기를 형성하기 위한 구체적 상황 진단이 생략되고 이론적·추상적 논의만이 범람하는 것으로 보인다. 그러나 우리나라의 실태에 대한 정확한 진단과 그에 기초한 정책 목표 설정이 부재한 상황에서 법학자들이 각자의 가치체계에 기대어 결과 예측 없이 ‘규범적 균형점’에 관한 주관을 피력하는 것이 과연 타당한 결론의 도출에 도움이 될지는 재고해 볼 일이다. 요컨대, TDM 기술에 관한 우리나라의 현실이 어떠한지가 사전에 실증적으로 파악되어야만 정책적 관점에서 필요한 면책의 ‘수준’을 정할 수 있다. 그런 연후에 해외 법제의 어느 요소는 적극적으로 참조하고 어느 요소는 덜 참조할지, 입법안의 세부 내용이 현행법의 체계와 어떻게 충돌하고 어떻게 수정해야 할지 등에 관한 논의를 전개하는 것이 순서일 것이다.

위와 같은 과정을 통해 우리나라의 상황에서 이해관계 조정의 최적 균형점을 발견했다면, 그 구체적 실현 방안에 관한 논의에 있어서 TDM 면책규정을 도입하려는 취지를 유념해야 한다. 우리 저작권법에는 포괄적·일반적 저작재산권 제한사유를 담은 제35조의5 공정이용 규정이 존재한다. 그럼에도 불구하고 개별 면책규정 도입을 논의하고 시도하게 된 것은 위 조항이 현실에서 요구되는 예견가능성을 제공하기에 부족함이 있기 때문이다. 이러한 입법 논의의 배경에 비추어 볼 때, 저작권법 전부개정안 제43조의 ‘비향유적 이용’과 같은 불확정 개념표지가 가져올 결과와, “필요한 한도”에 관한 예시가 부재하여 그에 관한 판단이 전적으로 사법부의 해석에 맡겨져 있는 점이 예견가능성의 제고를 주된 동기로 하여 그 도입이 논의되고 있는 TDM 면책규정의 실효성을 저해하지는 않을지에 대하여 보다 진지한 고민이 필요하다.

115) 영국 IPO는 CIPA 제29A조 개정 입법 계획을 밝힌 이후에도 자국의 기술 발전 및 변화 상황에 관한 진단을 계속 진행하여 그 결과를 공개하고 있다. UK IPO, “Research and analysis: Intellectual property and investment in Artificial Intelligence”, Jul. 4, 2022, available at: <<https://www.gov.uk/government/publications/intellectual-property-and-investment-in-artificial-intelligence/intellectual-property-and-investment-in-artificial-intelligence#appendix>>.

116) UK IPO, *supra* note 79, pp. 3-12.

2. 각국 입법의 정책적·법제적 배경에 관한 고려

TDM 면책에 관한 입법례들은 저마다 조금씩 다른 형식과 내용을 가지고 있어, 전반적인 면책 수준의 비교는 가능하지만 어느 법제가 다른 것에 비해 절대적으로 더 폐쇄적이라거나 개방적이라고 말하는 데에는 주의가 필요하다. 일본과 유럽연합의 법제만 비교해 보더라도, 일본의 법제는 이용 방법의 제한이 없는 점, 오프아웃의 불허, 컴퓨터에 의한 분석으로 제한되지 않는 점, 적법한 접근권한을 요구하지 않는 점 등에서 유럽연합 DSM 지침에 비해 상대적으로 넓은 면책 범위를 제공하지만, DSM 지침에는 없는 ‘사상 또는 감정의 비향수적 이용’이라는 요건을 두고 있다. 이처럼 법제 간 차이가 나타나는 것은 해당 국가·권역의 고유한 산업적, 정책적, 그리고 법제적 상황이 빚어낸 결과라고 할 수 있다. 따라서 이러한 배경적 상황을 이해하는 것은 우리나라의 법제에 각 입법례의 특정 요소를 어떤 방식으로 참조할 것인지를 정하는 작업의 전제가 된다. 이하에서는 각 국가·권역의 입법 배경이 우리나라의 TDM 면책규정 도입 논의에 시사하는 바에 관해 간략히 논한다.

가. 유럽연합 DSM 지침

유럽연합 DSM 지침이 전반적으로 매우 신중하게 TDM 면책의 요건을 정한 것은 DSM 지침이 본래 유럽연합 역내시장의 교역 장벽을 제거하고 TDM 기술 활용에 관한 각 회원국 규율의 통일·조화를 위해 마련되었다는 점에 기인한다.¹¹⁷⁾ 그 결과 DSM 지침은 TDM 기술수준 및 활용상황 등이 서로 다른 회원국들 간의 이해를 적당히 조정할 타협안이 되는 것이 불가피하였다.¹¹⁸⁾ 더불어, DSM 지침이 TDM 면책에 관한 제3, 4조 외에 권리 확충에 관한 내용을 포함하는 점도 면책 범위의 확장을 제한하였는데, 일 예로 DSM 지침 제15조에 의해 저작권접권을 부여받게 된 언론사는 TDM 분석에 적합한 텍스트 데이터인 뉴스 기사를 대량으로 보유·제공하는 기관으로서 TDM 면책규정의 수혜자와 반대되는 이해관계를 갖는 자이다. DSM 지침의 전술한 입법 배경 및 그로 인해 전반적으로 TDM 면책 범위가 제한된 점은, 유럽연합과 같은 체제적 배경을 가지지 않은 우리나라에서 면책사유를 유럽연합에 비해 더 완화하는 입법을 시도하는 데 있어 장애요소가 상대적으로 적을 수 있음을 시사한다.¹¹⁹⁾

117) Federico Ferri, ‘The dark side(s) of the EU Directive on copyright and related rights in the Digital Single Market’, *China-EU Law Journal*, vol. 7, 2021, pp. 23, 26.

118) *Id.*, p. 26.

119) 유럽 국가들의 입법은 경직된 데이터베이스 지침의 한계를 배경으로 하며 그와 같은 한계가 최근 DSM 지침을 통해 극복되었다고 지적하면서, “예초에 그러한 한계가 없고 이미 매우 유연한 공정이용 규정을 두고 있는 우리나라에서는 참고하기에 적당하지 않아 보인다”고 주장하는 견해도 있다. 이는 필자와는 조금 다른 견해이기는 하나, TDM 면책 관련 타국 법제의 내용뿐 아니라 배경까지 고려하였다는 점에서 참고가 된다. 김기정·송석현, “딥러닝과 데이터마이닝에 있어서의 데이터베이스권의 범위와 제한에 관한 고찰”, 『정보

환언하면, 유럽연합 DSM 지침이 회원국들의 다양한 이해관계를 적당히 반영한 결과로서 갖는 ‘제한적인(restrictive)’ 면책입법이라는 점은, DSM 지침의 내용을 차용함에 있어 비판적인 시각을 견지해야 할 필요성을 시사한다. 특히, DSM 지침에 대하여 가해지는 비판 중 면책 범위의 제한이 경제력에 따른 연구주체 혹은 회원국 간의 불공정한 경쟁환경을 조성할 것이라는 비판에 유념할 필요가 있는데, 우리나라에 유사한 내용의 TDM 면책규정을 도입할 경우 다량의 데이터에 적법한 접근이 가능한 글로벌 빅테크 기업들에 비해 국내 사업자에게 더 불리한 경쟁환경이 조성될 가능성도 산업정책의 관점에서 고려되어야 하기 때문이다.

나. 일본의 ‘비향수적 이용’ 규정

한편, 우리 저작권법 전부개정안 제43조의 기초를 제공한 일본 저작권법에서 제30조의4는 우리나라와 전혀 다른 배경에서 신설되었음에 주목할 필요가 있다. 일본은 우리나라의 공정이용 조항과 같은 일반적 면책규정이 존재하지 않는 상황에서 일반규정을 도입할 것인지를 검토하여 차선택으로 ‘유연한 권리제한 규정’을 도입했다. 일본의 2018년 개정 저작권법에 신설된 제30조의4 등은 개별 면책조항의 틀을 유지하면서도 미국식 공정이용 조항에 근접한 역할을 의도하고 도입한 조항이다. 이처럼 면책사유의 ‘유연화’를 추구한 일본의 법개정은, 저작권법 제35조의5 공정이용 조항에 내재된 예견가능성 문제를 해결하기 위해 TDM 분야 특정 면책조항의 도입을 논의하고 있는 우리나라의 상황과 정반대의 방향성을 띤다. 따라서 우리 저작권법에 일본법 제30조의4와 유사한 내용의 면책규정을 신설하는 것이 당초 의도한 입법목적을 달성할 수 있을지 신중하게 따져 보아야 한다.

특히 일본 저작권법 제30조의4의 중심 요건인 ‘비향수적 이용’은 저작권의 외적 한계라기보다 표현의 자유와의 균형을 달성하기 위한 내적 한계가 된다고 평가되며,¹²⁰⁾ 일본이 그와 같은 내적 제한을 법제화한 이유는 바로 변형적 이용 등의 심사를 통해 탄력적 운용이 가능한 공정이용과 같은 일반조항이 부재하기 때문이라고 할 수 있다. 이 점은 제35조의5 공정이용 조항이 유사한 기능을 이미 수행하고 있는 우리 저작권법에서 ‘비향유적 이용’이라는 추상적 요건을 굳이 추가하여야 하는지에 관한 의문을 불러일으킨다. 비향수적 이용에 관해서는 일본 내에서도 해석·적용의 모호성이 문제로 지적되고 있다고 보고되는데,¹²¹⁾ 우리나라의 TDM 면책규정이 예견가능성의 확보를 목표로 도입이 추진되고 있으며 그 관건은 법해석상의 모호성을 최소화하는 데 있다는 점에서 일본법에 대해 가해지는 위와 같은 비판을 가볍게 볼 수 없다. 한편 일본법의 ‘향수’와 우리법 개정안의 ‘향유’의 대상이 “표현”이 아닌 “표현된 사상 또는 감정”으

법학」 제25권 제2호, 한국정보법학회(2021), 216면.

120) Ueno, *supra* note 102, p. 151.

121) 박성호, 앞의 글(주 108), 16면; 신현철, “제4차 산업혁명시대의 일본 저작권법 개정과 시사점 - ‘유연성 있는 권리제한규정’을 중심으로 -”, 『법학논총』 제26집 제3호, 조선대학교 법학연구원(2019), 18-19면, 각주 46 참조.

로 되어 있는 점도 그냥 지나칠 수 없다. 이 점에 관해서는 일본에서도 이렇다 할 논의를 찾을 수 없는데,¹²²⁾ 우리 저작권법 전부개정안 제43조는 일본법의 문언을 별다른 고민 없이 그대로 따랐다. ‘표현의 향유’와 ‘표현된 사상 또는 감정의 향유’ 사이에 그 의미, 면책 범위, 정책적 함의 등에 있어 어떠한 차이가 있는지에 대한 검토가 법문언의 선택에 선행되는 것이 바람직할 것이다.¹²³⁾

다. 영국의 개정입법 계획 및 싱가포르의 개정 저작권법

마지막으로 비상업적 목적 제한을 제거하는 것을 포함해 TDM 면책 범위의 확장을 계획한 영국의 상황에서도 우리나라의 입법 논의에 관한 시사점을 발견할 수 있다.¹²⁴⁾ 영국이 CDPA 제29A조의 개정을 추진하는 배경에는 AI 기술의 개발·이용환경을 개선한다는 국가 비전이 놓여 있다. 스탠포드 대학이 2022년에 내놓은 통계에 의하면 지난 12년 간 학술연구 및 산업에서 AI 기술에 관한 협동연구의 양으로 매진 순위에서 영국-미국 및 영국-중국 간 협동연구가 중국-미국 간 협동연구에 이어 2위를 차지했고,¹²⁵⁾ 영국은 AI 기술에 대한 민간 투자총액과 신규 투자를 유치한 AI 기술업체의 수에 있어 미국과 중국에 이어 3위를 차지하였을 정도로,¹²⁶⁾ 영국은 AI 기술의 전 세계적인 발전, 전파를 주도하는 주요 국가로 자리매김하고 있다.¹²⁷⁾ 이처럼 학계뿐 아니라 산업계에서 높은 AI 기술 수준과 활용도를 보이고 있는 영국의 상황은, 영국이 2014년 CDPA 제29A조를 신설한 지 얼마 지나지 않아 면책 범위 확장을 내용으로 하는 재개정을 시도하고 있는 이유를 가늠케 한다. 우리 저작권법 전부개정안 제43조는 상업적 목적의 TDM을 면책 대상에 포함하는 점에서 영국의 입법 계획과 공통되는데, 이 점은 상업적 목

122) 일본 문화부의 입법검토자료(文化審議會著作権分科會, 「文化審議會著作権分科會報告書」, 2017. 4.)에서도 ‘표현의 향유’와 ‘사상 또는 감정의 향유’를 특별한 구분 없이 쓰고 있다.

123) 이에 관하여, “저작권법의 법리에 의하면 표현을 향유하지 않으면서도 사상이나 감정을 향유하는 것은 가능하다”고 하면서, 표현이 전달되지 않는 경우의 권리 제한을 의도한 입법이라면 “저작물의 표현을 향유하지 아니하는 경우”로만 규정하는 것이 간명하다는 견해가 있다. 김기정·송석현, 앞의 글(주118), 215면.

124) 다만, 영국이 이와 같은 입법계획을 실현할 수 있을지는 현재로서 단언하기 어렵다. IPO가 제출한 CDPA 개정 계획을 발표한 이후 우리나라의 문화체육관광부에 대응하는 DCMS(Department for Digital, Culture, Media and Sport)의 장관은 상원의회 앞에서 IP 보호의 중요성을 강조하면서 IPO의 입법 계획이 실현되지 않을 것이라는 견해를 표명했다. See UK House of Lords, ‘Communications and Digital Committee - Uncorrected oral evidence: A creative future’, Nov. 22, 2022, p. 18.

125) Daniel Zhang et al., ‘Artificial Intelligence Index Report 2022’, Stanford Institute for Human-Centered Artificial Intelligence, Mar. 2022, pp. 22-23.

126) *Id.*, pp. 154-157.

127) 최근 텍스트 기반 이미지 생성 모델을 오픈소스로 개방하고 기계학습 자원을 클라우드로 공유하는 등으로 AI 산업생태계 확장을 이끌고 있는 Stability AI社가 영국 런던에 본사를 둔 유니콘 스타트업이라는 점은 잘 알려진 사실이다. 김혜선, “AI그림 오픈소스로 공개하자...전세계서 ‘기발한’ 창작이 시작됐다”, 이데일리 2022. 10. 21.자 기사; 정병일, “스테빌리티AI, 클라우드 활용해 생성AI 훈련과 보급 강화”, AI타임스 2022. 12. 1.자 기사 참조.

적 이용을 개별 TDM 면책규정의 적용 범위에 포함시킬 것인지에 관한 입법정책적 판단을 실증적 근거 위에서 실시하기 위해 AI를 비롯한 우리나라의 TDM 기술 활용분야 산업 실태와 전망을 면밀히 파악할 필요가 있음을 시사한다.

저작권법에 미국식 공정이용 조항인 제35조의5가 존재하는 상황에서 개별 TDM 면책규정 도입을 추진 중인 우리나라는, 공정이용 조항이 이미 마련된 저작권법제에 TDM 면책규정을 추가한 싱가포르의 상황과 비슷하다. 그러나 이와 같은 법제적 상황의 공통점은 싱가포르와 유사한 수준의 법제화를 정당화하기에 불충분하다. 결국 우리나라의 산업적·정책적 상황에 대한 진단과 법적 전통의 차이를 모두 고려하는 가운데 싱가포르의 법제를 참조하는 것이 바람직할 것이다. 싱가포르가 노동시장 내 AI 기술인력에 대한 수요의 비중이 전 세계에서 가장 높은 나라라는 점을,¹²⁸⁾ 영국이 계획한 것과 유사한 내용의 과감한 TDM 면책규정을 도입한 배경과 관련하여 참고할 수 있다.

3. 저작권법 전부개정안에 관한 단견

전술한 바와 같이 우리나라의 법제 요구 상황에 대한 기초자료가 미비하고 정책 목표가 명확하게 설정되었다고 보기 어려운 현 상황에서 저작권법 전부개정안의 개별 내용의 타당성을 논하는 것은 적절치 않다. 다만 위 개정안에는 구체적 정책 목표 설정과는 관계없이 문제로 지적할 부분이 있다. 이하에서는 재고의 필요성과 개선할 점 및 개선에 있어 고려할 사항에 관한 간략한 의견을 덧붙인다.

가. ‘적법한 접근’ 요건

‘적법한 접근(lawful access)’은 앞서 검토한 주요국 입법례 가운데 일본을 제외한 모든 국가·권역에서 면책의 기본 요건으로 요구하는 사항이다. 우리나라의 저작권법 전부개정안 제43조도 적법한 접근을 요건으로 포함한다. 이 요건에 관해서는 긍정론과 부정론이 대립하는데,¹²⁹⁾ 구체적인 당부 평가 이전에 ‘적법한 접근’의 의미와 도입 배경에 관한 이해는 반드시 필요하다. ‘적법한 접근’은 DSM 지침 및 그 바탕이 된 유럽 각국 법제에 뿌리를 둔 요건으로, DSM 지침 제정을 위한 규제영향평가서에서 언급하는 사항, 즉 ‘연구자들이 적법하게 접근 가

128) 전체 노동수요 중 AI 기술인력에 대한 수요가 차지하는 비중이 싱가포르(2.33%)는 2위인 미국(0.90%)에 비해 월등히 높다. 이 지표에 있어 우리나라는 순위권 밖으로 조사되었으나, AI 기술에 대한 민간 투자총액과 신규 투자유치 기업 수에 있어서는 2021년에 싱가포르를 앞질러(2013년부터 2021년까지의 합산 순위에서는 싱가포르보다 낮음) 빠르게 발전하는 것으로 보인다. Zhang et al., *supra* note 125, pp. 145, 154-157.

129) 긍정론에 관해서는 신창환, 앞의 글(주 41), 170면, 이규호, 앞의 글(주 112), 100면 등을, 부정론에 관해서는 앞의 주 110, 111의 문헌을 참조.

능한 콘텐츠에 대해 TDM을 수행하는 것에 관한 법적 불확실성의 제거¹³⁰⁾에서 그 정책적 배경의 힌트를 얻을 수 있다. 유럽은 오랜 학술적·문화적 전통 하에서 대형 출판사와 언론사들이 TDM 소스로 활용될 수 있는 대규모 텍스트 데이터를 축적하고 상업적으로 이를 제공하고 있다.¹³¹⁾ 이에 따라 유럽에서는 데이터에 대한 접근이 그 제공자에 의해 통제됨을 당연한 전제로 그 데이터에 접근하여 허락된 방식으로 이용할 수 있는 이용자가 TDM 활용을 위해 추가 라이선스를 받아야 하는지가 중요한 정책 관심사였다.¹³²⁾ 이 점은 DSM 지침상의 TDM 면책규정에 관한 의견들에서 빈번하게 인용된 구호가 “읽기 권한은 마이닝 권한을 포함해야 한다(the right to read should be the right to mine)”였던 점에서 확인된다.¹³³⁾ 따라서 ‘적법한 접근’을 우리 저작권법상의 TDM 면책규정의 요건으로 포함시킬 것인지는 우리 저작권법에 ‘접근권’이 배타권의 하나로 규정되어 있지 않다는 점과는 큰 관련이 없다. 그보다는 TDM 면책규정의 규율범위(coverage)의 문제라 할 수 있다. 즉 TDM의 전제가 되는 데이터 수집(예: 크롤링 등)의 위법성을 TDM의 면책에 고려할 것인지가 ‘적법한 접근’의 요건화 여부에 따라 갈리게 되는데, 유럽연합을 비롯해 ‘적법한 권한’을 요건으로 둔 법제는 전자임이 명백하고 일본은 분명하지 않다.¹³⁴⁾ 우리는 어떤 선택을 할 것인가? 이 역시 막연한 법이론적 논의가 아니라 우리나라의 현실을 고려한 정책적 판단으로써 접근할 영역이다. 우리나라의 TDM 면책규정의 요건으로 ‘적법한 권한’을 도입함에 있어서, 데이터 수집에 관한 타 법제 규율과의 조화 내지 역할 분담에 관해서도 고려할 필요가 있다. 이와 관련하여 최근 이른바 ‘야놀자 형사 및 민사 사건’¹³⁵⁾에서 정보통신망법 제48조 제1항의 “접근권한”이 쟁점화되었는데, 이를 비롯하여 저작권법상의 기술적 보호조치의 정의에 포함된 “접근” 및 부정경쟁방지법 제2조 제1호 카목의 요건으로 규정된 “접근권한”의 해석·적용에 미칠 영향 및 이들 타법제와의 체계 조화 방안에 관해서도 면밀한 고려가 필요하다.

부연하면 야놀자 형사 사건에서는 TDM 면책 입법과 관련하여 두 가지 특기할 점이 있다. 첫째로, 대법원이 데이터베이스에 대한 접근권한을 엄격하게 해석하여 정보통신망 침입의 인정에 신중한 자세를 취한 점과 관련하여 저작권법상 보호되는 데이터베이스가 아닌 데이터에

130) European Commission, *supra* note 3, p. 111.

131) Springer, Elsevier, C.H. Beck 등 유럽에 기반한 출판사들은 대규모 학술 데이터베이스를 구축하고 라이선스를 통해 전 세계 도서관 등에 서지 및 문헌 서비스를 제공하고 있다.

132) Christopher Geiger, Giancarlo Frosio & Oleksandr Bulayenko, ‘The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects’, CEIPI Research Paper No. 2018-02, Mar. 2018, pp. 28-29.

133) Geiger et al., *supra* note 52, p. 50, n. 123.

134) 일본 저작권법 제47조의5에서는 정보해석과 그 결과 제공에 부수하는 경미한 이용에 대해 그 경미한 이용의 준비를 위해 필요한 한도에서 준비행위를 허용한다. 따라서 대량의 웹크롤링 등까지 포괄적으로 허용하는 것으로 보기 어려울 수 있다. 한편, 제30조의4에서 정보해석 그 자체가 사상·감정의 비향수적 이용에 해당한다고 하여 이를 위한 대량의 데이터 수집(복제)까지 비향수적 이용으로 볼 수 있을지는 미지수다.

135) 대법원 2022. 5. 12. 선고 2021도1533 판결; 서울고등법원 2022. 8. 25. 선고 2021나2034740 판결(확정).

대한 접근행위를 쉽게 정보통신망 침입으로 인정하게 되면 자칫 데이터 보유자에 대한 배타적 권리 창설과 같은 결과를 가져올 수 있다는 측면에서 대법원의 해석을 긍정적으로 평가하는 견해가 있다.¹³⁶⁾ 그런데 그러한 해석은 여러 논의 끝에 나온 것이고, 정보통신망 침입과 다른 맥락을 지닌 저작권법상의 접근권한에 대해서도 동일한 해석 기준이 적용되리라고 단정할 수 없다. 이 점은 TDM 면책규정에 접근권한에 관한 요건을 둘 경우 해석·적용상의 혼란에 대비하는 것이 바람직함을 시사한다. 둘째로, 이 사건에서 대법원은 데이터베이스제작자 권리 침해에 관한 요건 해석에 있어서도 준물권적 독자적 권리 부여 방식에 관한 정보 독점의 우려 및 이를 의식한 유럽연합법원의 해석례를 참고하여 매우 엄격한 관점을 취하였고, 그에 대하여 부정경쟁방지법 제2조 제1호 (파)목에 의한 보호 가능성과 부정경쟁방지법에 데이터 부정취득에 관한 규정이 최근 추가된 점을 고려할 때 타당하다는 평가가 있다.¹³⁷⁾ 만약 TDM 면책규정이 우리 저작권법에 도입되고 접근권한에 관한 내용이 포함된다면 데이터베이스제작자 권리 보호에도 일정한 영향이 미칠 수 있다. 뿐만 아니라 데이터베이스제작자 권리에 관한 현행 규정형식의 존치 여부 등에 관한 논쟁을 재점화할 가능성도 있다. 이 점도 TDM 면책규정의 설계에 있어 예상 내지는 대비하여야 할 사항 중 하나일 것이다.

나. 상업적 이용에 대한 면책 여부

유사한 태도가 TDM 면책규정의 적용 범위를 상업적 이용에까지 확장할 것인지와 관련된 정책적 판단에 있어서도 요구된다. 현재 및 장래의 우리나라 TDM 기술의 수요를 파악하고, 저작재산권 및 데이터베이스제작자 권리 제한의 정책적 정당화가 가능한 범위 설정에 이를 적극적으로 참조해야 한다. 이와 관련하여 국내에서 TDM 개별 면책규정의 범위를 학술연구 등 비상업적 이용에 국한하고 상업적 목적 이용은 저작권법 제35조의5로 다루자는 견해가 있는데,¹³⁸⁾ 공정이용 조항이 존재하는 우리 저작권법의 체계에 비추어 접근방법의 타당성은 일응 인정할 수 있다. 그러나 우리나라의 TDM 기술 개발·활용 실태에 비추어 상업적 목적 이용이 정책상 더 비중 있게 다루어져야 할 분야에 속한다면 위의 제한적 적용 범위 설정이 TDM 면책규정을 도입하려는 취지에 부합하는지 재고해 볼 필요가 있고,¹³⁹⁾ TDM 면책에 관한 규율 일부를 제35조의5에 맡기는 방안은 비상업적 이용뿐 아니라 TDM 분석에 직접 수반되지 않는 사전적 혹은 사후적 이용행위에 대해서도 고려할 수 있다는 점이 정책 형성과정에 참고가 될

136) 정현순, “크롤링에 의한 데이터 수집이 정보통신망 침입에 해당하는지 여부 - 대상판결: 대법원 2022. 5. 12. 선고 2021도1533 판결”, 『사법』 통권 제61호, 사법발전재단(2022), 380면.

137) 정현순, “저작권법상 데이터베이스제작자 권리 침해에 관한 연구 - 대법원 2022. 5. 12. 선고 2021도1533 판결을 중심으로 -”, 『계간 저작권』 제35권 제3호, 한국저작권위원회(2022), 190-191면.

138) 김병일 외, 앞의 글(주 42), 58면; 차상욱, 앞의 글(주 6), 47-48면; 최상필, 앞의 글(주 66), 84면 등.

139) 신창환, 앞의 글(주 41), 168면 참조.

수 있을 것이다.

다. '전송'에 대한 면책 여부

전술한 바와 같이 일본 저작권법은 비향수적 이용(제30조의4), 정보해석에 부수한 경미한 이용(제47조의5) 모두에 대해 면책되는 이용 방법에 제한을 두고 있지 않다. 반면 유럽연합 DSM 지침은 복제와 추출로 그 범위를 제한하고 있는데, '복제'는 저작물의 대표적인 이용 방법이고 '추출(extraction)'은 유럽연합 데이터베이스 지침¹⁴⁰⁾에 '재이용(re-utilization)'과 함께 등장하는 용어로서 소재의 상당한 부분의 복제를 가리킨다.¹⁴¹⁾ 우리나라 저작권법 전부개정안에서는 정보분석을 위한 비향유적 이용에 대해 복제 및 전송을 면책되는 이용행위로 규정하고 있어, 유럽연합에 비해 넓고 일본에 비해 허용되는 이용행위의 폭이 좁다. 우리나라의 TDM 면책규정안이 복제 외에 전송을 포함하는 것은 TDM 분석 과정에서 분산 서버 간에 이루어지는 데이터 전송을 포괄하고자 하는 의도로 보이나, 전송의 범위에 관해 명시적인 제한을 설정하지 않기 때문에 문언상으로는 복제물의 제3자 제공까지 허용하는 것으로 해석될 여지가 있다. "필요한 한도 안에서"라는 내적 제한을 두고 있기는 하나, 이와 같은 추상적인 한계 설정이 예측가능성의 제고를 목적으로 하는 개별 TDM 면책규정의 도입 취지에 적합한 입법인지에 대해서는 재검토가 필요하다. 이와 관련하여서는 싱가포르의 입법례가 참고가 될 수 있는데, 싱가포르 저작권법은 저작물 등에 대한 복제 외에 송신에 의한 제3자 제공을 허용하면서, 그 조건으로 '연구구결과의 검증 또는 합동연구 목적인 경우'라는 제한을 부과하고 있다. 우리 법의 입안에 있어서도 만약 '전송'을 면책 범위에 포함시켜야만 한다면 그 이유를 적절히 반영한 조건 내지 범위의 제한을 명시하는 것이 바람직할 것이다.

라. 출처명시의무의 요부

우리나라 저작권법 전부개정안 제49조는 TDM 면책규정인 제43조를 출처명시의무의 대상에서 제외하고 있다. 참고로, 유럽연합 DSM 지침은 출처명시의무를 별도로 규정하지 않고, 일본 저작권법에서 출처명시에 관해 규정하고 있는 제48조는, 제30조의4의 비향수적 이용을 출처명시의 대상에서 제외하면서, 제47조의5 제1항에 따른 부수적이고 경미한 이용에 대해서는 출처명시의 관행이 있는 때에 한해 출처명시를 하도록 규정하고 있다. 영국의 현행 CDPA 제29A조 제1항은 '충분한 출처명시'를 컴퓨터 분석 과정에서의 복제에 대한 면책의 요건으로 규정하고 있으나, 그 출처명시가 실용적인 사유(reasons of practicality) 혹은 기타 사유로 인해

140) Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

141) 정현순, 앞의 글(주 137), 171면 참조.

불가능하지 않을 것을 조건으로 한다. TDM의 성격상 대량의 데이터가 분석되므로, 서로 다른 수많은 출처를 명시하는 것은 불가능하거나 부적절할 수 있다. 따라서 저작권법 전부개정안 제49조에서 제43조에 따른 비향유적 이용에 대해 출처명시의무를 면제하는 것은 일견 적절해 보이며, 그 입법안의 기초로 삼은 일본 저작권법에서 제30조의4에 따른 이용에 대해 출처명시를 요구하지 않는 것과 동일한 맥락으로 보인다.

독일의 경우 2021년 개정 전의 저작권법 제63조는 출처명시의무를 규정하면서 복제, 배포, 공중송신을 모두 포함하고 있었으나, 2021년 개정을 통해 복제와 배포에 관한 출처명시에 관해서는 제60d조를 그 대상에서 삭제하고, 공중송신만을 출처명시의 대상으로 규정하였다. 유의할 점은 개정 전후의 제63조 모두 출처명시가 관행상 요구될 것과 가능한 경우일 것을 조건으로 규정함으로써, 출처명시가 관행적으로 요구되지 않거나 불가능한 경우를 예외로 정하고 있다는 점이다. 이 점에서 제47조의5 제1항에 대한 출처명시를 규정한 일본 저작권법과 유사한 태도이다. 앞서 언급한 바와 같이 TDM 과정에서 복제 등의 방법으로 이용되는 소재는 매우 대량이므로 저작자가 일정치 않을 수 있고 분야에 따라서는 저작자를 특정할 수 없는 경우가 많다는 점을 고려할 때 출처명시를 강제하는 것은 TDM의 특성에 반하는 측면이 있다. 다만, 전부개정안에서는 ‘전송’을 면책되는 이용행위로 포함하기 때문에 독일 입법례와 같이 TDM 분석과정의 핵심에 해당하는 복제에 대해서는 출처명시의무를 면제하더라도 전송에 대해서는 권리자의 이익 침해 정도에 비례하여 출처명시를 의무화하는 방안을 고려해 볼 수 있다. 그 경우 일본과 독일 등의 입법례를 참조하여 ‘출처명시의 관행이 있을 경우’ 및/또는 ‘출처명시가 가능한 경우’를 조건으로 명시함으로써 TDM 면책입법의 실효성을 살릴 수 있을 것이라는 입론이 가능하다. 다만, 이 경우에도 출처명시에 대해 논하기에 앞서 ‘전송’을 아무런 제한 없이 면책 범위에 포함시키는 것의 타당성 문제를 검토하는 것이 순서일 것이다.

V. 결 어

분야를 막론하고 데이터의 중요성이 날로 강조되고 있는 현실에서 TDM 기술이 인류의 지식과 편익의 증대에 크게 기여할 것임은 분명하다. 그러나 TDM은 대량의 데이터 복제를 수반하기 때문에 저작물인 데이터와 데이터베이스를 배타적 보호대상으로 삼는 저작권법과 필연적으로 긴장관계에 놓이게 된다. TDM의 유용성을 긍정하는 전제에서 우리 저작권법에 TDM 행위를 위해 특별히 마련된 권리제한조항을 도입하는 것에 관해서는 큰 이견이 없고, 그러한 공감대 위에서 저작권법 개정이 추진되고 있다.¹⁴²⁾ TDM 면책입법은 전 세계적인 추세이기 때

142) 물론 TDM 면책규정의 도입 필요성을 역설하는 주장들에 대한 비판적 견해가 전혀 없는 것은 아니다. 한

문에 우리 저작권법의 TDM 면책규정의 형태를 만들어가는 데에는 해외 입법례가 당연히 참조되어야 한다. 다만 우리나라의 정책적·법제적 상황에 가장 적합한 해결 방안을 찾기 위해서는 타국 입법의 ‘결과’만이 아니라 그 ‘과정과 배경’까지 함께 살펴보고 무엇이 좋고 무엇이 다른지를 이해하려는 자세가 필요하다.

이와 같은 견지에서 본고에서는 주요국의 최근 TDM 면책규정의 입법 동향을 그 배경과 함께 소개하고, 이를 바탕으로 우리나라의 입법 논의가 정초할 방향성에 대해 검토하였다. 필자가 저작권법 개정안 제43조의 개별 요건의 당부에 대한 평가나 구체적인 입법 방안에 관한 의견을 의식적으로 자제한 이유는 우리나라의 상황을 객관적으로 분석한 기초자료가 미비한 상태에서의 선부른 입론이 정책 형성에 도움이 되지 않는다는 판단에서였다.

각국의 입법은 저마다 고유한 특징이 있으나, 이용의 목적에 관해 비영리성, 비상업성을 요구하던 태도가 약화되는 것이 대체적인 경향임을 확인했다. 일본과 그 밖의 국가 입법례 간에는 이용 방법의 제한 유무, 비향수적 이용 요건의 유무, 적법한 권한의 요부 등의 여러 측면에서 차이가 나는데, 이처럼 서로 다른 입법례의 일부 요소를 각각 차용하는 경우에는 특히 유의할 필요가 있다. 상이한 성격의 법제를 혼합할 때 나타날 부작용 혹은 시너지에 대한 결과 예측을 바탕으로 세밀한 입법적 검토가 요구되며, 이러한 작업은 우리나라의 상황에 대한 실증적 이해에 기반하여 정책 목표를 설정하고 이를 기준으로 삼아 진행해 나가야 할 것이다. TDM 면책규정의 내용에 관한 논의에 앞서 정확한 상황 진단에 기초한 정책 목표의 설정이 선행되어야 한다는 점을 저작권법 전부개정 논의가 시작된 지 수 년이 지난 이 시점에 지적하는 것은 다소 늦은 감이 있다. 그러나 데이터가 주도하는 현대사회의 저작권법 현실에서 TDM 면책규정이 담당할 역할을 고려할 때, 공허한 논의가 가져올 시행착오를 피할 수 있다면 조금 느리더라도 한 걸음씩 다지면서 나아가는 것이 현명할 것이다. ▣

(논문접수 : 2022. 12. 19. / 심사개시 : 2023. 1. 5. / 게재확정 : 2023. 1. 20.)

예로, 위의 주장들에 대해 “TDM을 지상과제로 설정하고 공정이용으로 해결되지 않는다면 법을 고쳐서라도 문제를 해결하고자 하는 시도”라고 평가하면서, “TDM 면책은 빅데이터를 구축한 빅테크 기업들에게 백지위임을 주는 것과 같고, 공정이용의 기능인 부의 재분배를 역으로 작동시킨다”고 비판하는 견해가 있다. 남형두, “신발을 위한 영화 - 플랫폼 사업자의 콘텐츠 확보에 따른 법적 문제 -”, 『경제규제와 법』 제11권 제2호, 서울대학교 법학연구소(2018), 308-311면 참조.

참 고 문 헌

1. 국내 문헌

<논문/보고서>

- 고수윤, “디지털저작물에 있어 일본의 유연한 권리제한규정(柔軟な權利制限規定)의 시사점 - 비향수이용 규정(일본 저작권법 제30조의4)을 중심으로 -”, 「계간 저작권」 제32권 제3호, 한국저작권위원회(2019).
- 김경숙, “TDM 관련 저작권법 개정안의 비판적 고찰 - 기술적보호조치 관점에서의 비교법적 검토 -”, 「경영법률」 제31집 제3호, 한국경영법률학회(2021).
- 김기정·송석현, “딥러닝과 데이터마이닝에 있어서의 데이터베이스권의 범위와 제한에 관한 고찰”, 「정보법학」 제25권 제2호, 한국정보법학회(2021).
- 김병일·신현철·안창원, “빅데이터 분석과 데이터 마이닝을 위한 저작권 제한”, 「계간 저작권」 제30권 제1호, 한국저작권위원회(2017).
- 김용주, “텍스트 및 데이터마이닝을 위한 저작권법 개정 방향”, 「법학연구」 제61권 제2호, 부산대학교 법학연구소(2020).
- 김창화, “데이터 마이닝과 저작권 면책의 범위 및 한계”, 「계간 저작권」 제34권 제2호, 한국저작권위원회(2021).
- 김혜창 외 3명, “인공지능 및 데이터베이스 등 기술변화에 따른 저작권법 개선 방안 연구”, 한국저작권위원회 연구보고서, 2021. 12.
- 남형두, “신발을 위한 영화 - 플랫폼 사업자의 콘텐츠 확보에 따른 법적 문제 -”, 「경제규제와 법」 제11권 제2호, 서울대학교 법학연구소(2018).
- 박성호, “저작권법 전면 개정안 주요내용의 비판적 고찰”, 「정보법학」 제24권 제3호, 한국정보법학회(2020).
- _____, “텍스트 및 데이터 마이닝을 목적으로 하는 타인의 저작물의 수집·이용과 저작재산권의 제한 - 인공지능의 빅데이터 활용을 중심으로 -”, 「인권과 정의」 제494호, 대한변호사협회(2020).
- 신창환, “인공지능 시대의 저작물 대량 디지털화에 관한 소고”, 「성균관법학」 제31권 제1호, 성균관대학교 법학연구원(2019).
- 신현철, “제4차 산업혁명시대의 일본 저작권법 개정과 시사점 - ‘유연성 있는 권리제한규정’을 중심으로 -”, 「법학논총」 제26집 제3호, 조선대학교 법학연구원(2019).
- 오승중, “데이터마이닝 및 텍스트마이닝과 저작재산권의 제한”, 「홍익법학」 제20권 제2호,

홍익대학교 법학연구소(2019).

유현우, “인공지능 시대 대응을 위한 「저작권법」 전부개정안 검토 - 정보분석을 위한 복제·전송 허용 규정을 중심으로”, 「AIIA」 제44호, 지능정보산업협회(2021. 3.).

이규호, “인공지능 학습용 데이터세트에 대한 저작권법과 부정경쟁방지법상 보호와 그 한계”, 「인권과 정의」 제494호, 대한변호사협회(2020).

이상용, “데이터의 비계약적 이용 - 데이터 마이닝을 위한 저작권 제한을 중심으로 -”, 「강원법학」 제65권, 강원대학교 비교법학연구소(2021).

전웅준, “인공지능 관련 저작권 침해에 관한 시론”, 「경영법률」 제31집 제4호, 한국경영법률학회(2021).

정진근 외 8명, “신기술에 대응하는 저작권 보호방안 연구”, 한국저작권보호원 연구보고서, 2021. 12.

정현순, “저작권법상 데이터베이스제작자 권리 침해에 관한 연구 - 대법원 2022. 5. 12. 선고 2021도1533 판결을 중심으로 -”, 「계간 저작권」 제35권 제3호, 한국저작권위원회(2022).

_____, “크롤링에 의한 데이터 수집이 정보통신망 침입에 해당하는지 여부 - 대상판결: 대법원 2022. 5. 12. 선고 2021도1533 판결”, 「사법」 통권 제61호, 사법발전재단(2022).

차상욱, “저작권법상 인공지능 학습용 데이터셋의 보호와 쟁점 - 텍스트·데이터마이닝(TDM) 면책규정을 중심으로 -”, 「경영법률」 제31집 제1호, 한국경영법률학회(2021).

최상필, “빅데이터의 분석과 활용을 위한 TDM의 저작권적 쟁점과 입법론”, 「계간 저작권」 제35권 제1호, 한국저작권위원회(2022).

홍승기, “데이터마이닝 면책 입법 방향에 대한 의문”, 「경영법률」 제32집 제4호, 한국경영법률학회(2022).

<기타 자료>

문화체육관광부, “창작자 권리보호는 강하게, 저작물 이용은 손쉽게 - 문체부, ‘추가보상 청구권’, ‘확대된 집중관리’ 및 인공지능 개발·활용 촉진 등 담은 「저작권법」 전부개정 추진 -”, 문화체육관광부 보도자료, 2020. 7. 1.

안효길, “빅데이터 활용과 인공지능 개발을 위한 Text and Data Mining”, 법률신문 2018. 12. 24.자 기사.

2. 해외 문헌

<단행본>

- Han, Jiawei, Jian Pei & Hanghang Tong, *Data Mining: Concepts and Techniques* (4th ed.), Morgan Kaufmann, 2022.
- McCue, Colleen, *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis*, Butterworth-Heinemann, 2007.
- Wandtke, Artur-Axel & Winfried Bullinger, *Praxiskommentar Urheberrecht* (6th ed.), 2022
- 小倉秀夫·金井重彦(編), 「著作権法コンメンタルⅡ(改訂版)」, 第一法規, 2020.
- 中山信弘, 「著作権法(第3版)」, 有斐閣, 2020.

<논문/보고서>

- Ang, Darren, 'The Web Scraper's World of Copyright Exceptions and Contractual Overrides', *SLR Juris Illuminae*, vol. 13, 2021.
- Bunge, John A. & Dean H. Judson, 'Data Mining', in *Encyclopedia of Social Measurement* (Kimberly Kempf-Leonard ed.), Elsevier, 2005.
- Carroll, Michael W., 'Copyright and the Progress of Science: Why Text and Data Mining Is Lawful', *UC Davis Law Review*, vol. 53, 2019.
- Ducato, Rossana & Alain Strowel, 'Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out', CRIDES Working Paper Series No. 1/2021, Feb. 2021.
- Dusollier, Séverine, 'The 2019 Directive on Copyright in the Digital Single Market: Some progress, a few bad choices, and overall a failed ambition', *Common Market Law Review*, vol. 57(4), 2020.
- Ferri, Federico, 'The dark side(s) of the EU Directive on copyright and related rights in the Digital Single Market', *China-EU Law Journal*, vol. 7, 2021.
- Filippov, Sergey & Paul Hofheinz, 'Text and Data Mining for Research and Innovation', *The Lisbon Council*, iss. 20/2016, 2016.
- Flynn, Sean et al., 'Comment in response to the WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI)', WIPO/IP/AI/2/GE/201, available at:

- <https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_global_expert_network_on_copyright_user_rights.pdf>.
- Geiger, Christopher, 'The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Fundamental Right to Research and the Failed Text and Data Mining Limitations in the CSDM Directive', PIJIP/TLS Research Paper Series 2021, No. 66, Apr. 2021.
- _____, Giancarlo Frosio & Oleksandr Bulayenko, 'Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU', in *Propiedad Intelectual y Mercado Único Digital Europeo* (C. Saiz Garcia & R. Evangelio Llorca eds., 2019).
- _____, _____ & _____, 'The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects', CEIPI Research Paper No. 2018-02, Mar. 2018.
- Gerrish, Charlotte & Anders Molander Skavlan, 'European Copyright Law and the Text and Data Mining Exceptions and Limitations', *Stockholm Intellectual Property Law Review*, vol. 2(2), 2019.
- Gilotte, Jessica L., 'Copyright Infringement in AI-Generated Artworks', *U.C. Davis Law Review*, vol. 53, 2020.
- Kop, Mauritz, 'Machine Learning & EU Data Sharing Practices', *Stanford-Vienna Transatlantic Antitrust and IPR Developments*, iss. 1/2020, Mar. 2020.
- Lemley, Mark A. & Bryan Casey, 'Fair Learning', *Texas Law Review*, vol. 99, 2021.
- Margoni, Thomas, 'Text and Data Mining in Intellectual Property Law: Towards an Autonomous Classification of Computational Legal Methods', CREATE Working Paper 2020/1.
- _____, & Martin Kretschmer, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology', *GRUR International*, vol. 71(8), 2022.
- Sag, Matthew, 'The New Legal Landscape for Text Mining and Machine Learning', *Journal of the Copyright Society of the USA*, vol. 66, 2019.
- Salloum, Said A., Mostafa Al-Emran & Khaled Shaalan, 'Mining Social Media Text: Extracting Knowledge from Facebook', *International Journal of Computing and Digital Systems*, vol. 6(2), 2017.
- Senftleben, Martin R.F., 'Study on EU Copyright and Related Rights and Access to and Reuse of Data', European Commission, Mar. 2022.

- Sobel, Benjamin L.W., 'Artificial Intelligence's Fair Use Crisis', *Columbia Journal of Law & the Arts*, vol. 41, 2017.
- Ueno, Tatsuhiro, 'The Flexible Copyright Exception for 'Non-Enjoyment' Purposes - Recent Amendment in Japan and Its Implication', *GRUR International*, vol. 70(2), 2021
- Underwood, Ted, David Bamman & Sabrina Lee, 'The Transformation of Gender in English-Language Fiction', *Journal of Cultural Analytics*, vol. 3(2), 2018.
- Zhang, Daniel et al., 'Artificial Intelligence Index Report 2022', Stanford Institute for Human-Centered Artificial Intelligence, Mar. 2022.

<기타 자료>

- AFP, "We can be proud: France becomes first country to adopt EU copyright reform", *THE LOCAL*, Jul. 24, 2019, available at:
<<https://www.thelocal.fr/20190724/french-parliament-adopts-eu-copyright-reform/>>.
- European Commission, 'Impact Assessment on the Modernisation of EU Copyright Rules', Commission Staff Working Document SWD(2016) 301 final Part 1/3, Sep. 2016.
- _____, 'Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market', COM/2016/0593 final - 2016/0280 (COD), Explanatory Memorandum, Sep. 14, 2016.
- Ministry of Law Singapore, 'Singapore Copyright Review - Enhancing Creators' Rights and Users' Access to Copyrighted Works', Press Releases, Jan. 17, 2019, available at:
<<https://www.mlaw.gov.sg/news/press-releases/singapore-copyright-review-report-2019>>.
- Trapova, Alina & João Pedro Quintais, 'The UK government moves forward with a text and data mining exception for all purposes', *Kluwer Copyright Blog*, Aug. 24, 2022, available at:
<<http://copyrightblog.kluweriplaw.com/2022/08/24/the-uk-government-moves-forward-with-a-text-and-data-mining-exception-for-all-purposes/>>.
- UK House of Lords, 'Communications and Digital Committee - Uncorrected oral evidence: A creative future', Nov. 22, 2022.

UK IPO, 'Consultation outcome - Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation', Jun. 28, 2022, available at:

<<https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copy-right-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation>>.

_____, 'Consultation stage impact assessment on Artificial Intelligence and Intellectual Property', RPC-BEIS-IPO-5101(1), Oct. 29, 2021.

_____, 'Exceptions to copyright: Research', Oct. 2014.

_____, "Research and analysis: Intellectual property and investment in Artificial Intelligence", Jul. 4, 2022, available at:

<<https://www.gov.uk/government/publications/intellectual-property-and-investment-in-artificial-intelligence/intellectual-property-and-investment-in-artificial-intelligence#appendix>>.

文化審議會著作權分科會, 「文化審議會著作權分科會報告書」, 2017. 4.

Abstract

A Study on Introducing a Specific Copyright Exception Clause to Text and Data Mining

Ryu, Shi-Won

Modern economy relies heavily on the important role of data. As such, text and data mining (TDM) technology is expected to contribute to the improvement of human knowledge and social benefits. However, since TDM is generally accompanied by reproduction of large volume of data and such data might belong to the subject matter of copyright law, i.e., copyrighted works and databases, TDM activities are inevitably put under tension with the copyright systems. In order to cope with the social needs of TDM technology and its possible collision with copyright systems, there arises a discussion in most countries regarding the amendments to the copyright law introducing a specific exception clause to TDM activities. It is necessary and adequate to refer to TDM exception clauses in other jurisdictions in the discussion of domestic legislation. However, in order to identify the best solution in our own situations, consideration should be given to not only the outcome but also the process and background of other countries' legislative reform.

In such a vein, this study reviewed the recent trends of several major countries' legislative outcomes or proposals related to copyrights exception to TDM activities along with their respective background of the legislation. A few of the subjects reviewed are as follows: While UK's CDPA Section 29A is referenced as one of the earliest individual TDM exception clause and the model of exception requiring non-commercial purposes, UK IPO announced its plan to amend the current legislation so as to extend its coverage to TDM activities with commercial purposes; Japan's amended Copyright Act Section 30-1 characterized with non-enjoyment use has been introduced as a flexible exception clause chosen as an alternative to general exception resembling US fair use clause, thus should be referred to critically in our condition where introduction of individual TDM exception is discussed in order to improve predictability under the domestic copyright law's fair use

regime; Singapore's recently amended copyright law, which employs an individual TDM exception clause having large coverage including commercial use and such clause was introduced even where a fair use clause existed, bears important implication for our copyright law amendment due to the resemblance in terms of legislative background and policy conditions.

It was emphasized in this study that the domestic legislative discussions on TDM exception for copyright should put its basis on empirical understanding of Korea's current situation and the policy goals established based thereupon. The detailed review of the Section 43 of the domestic copyright act draft was intentionally avoided under such consideration that hasted argumentation without base material for the diagnosis of our own conditions would lead to unwanted outcome. When adopting different elements from different legislative approaches, careful consideration should be given to the synergy or side effects of such combinatorial adoption. The discussion of introducing TDM exception in our copyright law is of early stage. Hopefully, this study serves to reexamining of the the desirable directions to which legislative discussion is to be guided.

Key Words : Text and Data Mining, TDM, Copyright Exception, Fair Use, Non-enjoyment Use, Information Analysis, Data Analysis, Computational Analysis, Artificial Intelligence