



뉴욕지방법원, 인터넷 아카이브의 공개도서관 프로젝트는 저작권 침해책임이 있다고 판단

정진근 | 강원대학교 법학전문대학원 교수



1. 들어가는 말	2. 사실관계	3. 판결의 주요 내용	4. 관련 논의와 시사점
-----------	---------	--------------	---------------

1. 들어가는 말

최근 뉴욕지방법원에서는 세간의 관심을 일으킨 ‘출판사와 인터넷 아카이브’ 간의 분쟁에 관한 판결이 있었다. 해당 판결은 HACHETTE BOOK GROUP, INC., et al., Plaintiffs v. INTERNET ARCHIVE, et al., Defendants, 2023 WL 2623787 United States District Court, S.D. New York (2023)이다.

뉴욕지방법원 판결의 내용은 (1)인터넷 아카이브에서 도서를 이용한 것은 변형적 이용이 아니며, (2)상업적 목적을 가지고 있으므로, (3)인터넷 아카이브에서 복제하여 디지털 형식으로 배포한 목적은 공정이용을 적용하는 데 있어 부정적이라는 것이다. 또한, (4)도서라는 저작물의 성격을 고려하는 경우에도 도서를 스캔하여 대여하는 행위를 공정이용으로 볼 것 인지에 대하여 부정적이며, (5)이용한 저작물의 양과 질을 고려하여도 도서 전체를 복제하

여 이용하였으므로 공정이용을 인정하는 데 부정적일 뿐만 아니라, (6)인터넷 아카이브의 행위는 전자책 시장에서 출판사들과 경쟁관계에 있으므로, (7)결론적으로 공정이용에 해당하지 않는다는 것이다.

이러한 해당 판결의 결론은 우리가 예상하던 것과 크게 다르지 않다. 인터넷 아카이브는 비록 공공의 이익을 위한다는 철학을 토대로 하고 있기는 하나, 변형적 이용과 시장에 미치는 영향이라는 공정이용의 요건을 충족할 가능성이 크지 않으므로 공정이용에 해당하지 않는다는 판결이 예상되어 왔다.

그럼에도 불구하고, 본 판결이 세간의 관심을 일으킨 이유는 인공지능(AI)·디지털 4.0 시대에 공정이용의 범위 확장 여부에 대한 기대와 우려가 있었기 때문이다. 비록 본 판결에서는 공정이용 범위를 확장하고 있지는 아니하나, 판결문 속에는 인공지능(AI)·디지털 4.0 시대에 공정이용 여부의 판단을 위한 기준들이 제시되고 있다. 그러한 점에서 본 판결은 큰 의미를 갖는다.

2. 사실관계

원고는 Hachette Book Group 외에도 HarperCollins, Wiley, Penguin 등 미국을 대표하는 출판사들이다. 이들 출판사는 저작자들로부터 도서의 복제 및 전자책(eBook)을 포함한 아날로그/디지털 형식의 출판을 위한 배타적 권리를 획득하고 있었다.

피고인 인터넷 아카이브(이하 'IA'라 함)는 “모든 지식에 대한 보편적인 접근”을 모토로 하고 있는 비영리기관이다. 인터넷에서 얻을 수 있는 자료는 물론이거니와 도서 등을 디지털화하여 웨이백머신(Wayback Machine)이라고 불리는 웹상의 자료실을 통해 제공하고 있다. 이러한 목적에 따라 IA는 공개된 웹페이지를 이용하는 것은 물론이거니와 도서관, 박물관, 대학 등과 협력해왔다.

이번 사건은 도서관들이 전자책을 대여하는 방법과 관련이 있다. 출판사의 전자책은 ‘대행사(aggregators)’에 의해 배포되었는데, 출판사들은 ‘동시접속불허모델(one-copy, one-user model)’ 등에 의해 이익을 확보하였고, DRM은 일정한 기일이 도과하면 복제물에 대한 접근을 차단하는 방법으로 이러한 배포 모델을 가능하게 하는 기술적 수단으로 이용되었다.

IA는 수백만 권의 도서를 스캔하여 archive.org 또는 openlibrary.org에서 공중이 이용하도록 하는 ‘열린 도서관(Open Library)’ 사업을 전개해왔다. 제공하는 도서에는 퍼블릭 도메인(public domain)에 속한 것도 있지만, 33,000개의 출판사들로부터 출판된 3백60만 권의 저작권에 의해 보호받는 도서들이 포함되어 있었다. 그렇다고 IA가 공중이 무작위로 이들 도서들을 다운로드받도록 하지는 않았다. IA는 CDL(Controlled Digital Lending)¹⁾

을 이용하도록 했는데, CDL은 도서를 소장한 수와 대여한 수를 1:1로 유지하도록 했다.

분쟁은 2020년 COVID-19 팬데믹 상황에서 발화하였다. 팬데믹 상황이 벌어지자 IA는 국가비상도서관(NEL: National Emergency Library) 프로젝트를 시작했는데, IA는 6억 5천만 권 이상의 도서들이 대여되지 않고 있다는 이유로 도서를 소장한 수와 대여한 수를 1:1로 유지하던 원칙을 깨고 한번에 최대 1만 명의 고객이 전자책을 대여할 수 있도록 했다. 이러한 이용에 앞서 역시 저작자나 출판사들의 허락을 얻지는 않았다.

이에 본 소송이 제기되었다.

3. 판결의 주요 내용

이번 사건은 저작권 침해를 인정할 것인지가 쟁점인 사건이다.

따라서 원고들은 피고인 IA가 저작권 침해행위를 했음을 주장하였고, 피고들은 자신들의 행위가 공정이용에 해당한다고 주장하였다.

이중 원고들의 주장의 입증 및 인정은 용이하였다. 미국 저작권법 제106조는 복제, 공연, 전시, 배포 및 2차적저작물 작성에 관한 배타적 권리를 규정하고 있으며, 미국 판례는 유효한 저작권을 보유하고 있으며, 저작물에 의거하여 복제 등의 행위를 하는 경우 저작권 침해를 인정하기 때문이다.

원고의 침해주장에 대한 피고들의 공정이용 주장은 항변으로서의 기능을 하는데, 공정이용의 항변을 주장하는 자가 입증책임을 진다.

이러한 이유로 해당 판결에서의 설명은 주로 ‘공정이용의 원칙에 따른 항변이 가능한가’의 쟁점에 집중되었다. 공정이용의 원칙은 판례법에 따라 수용, 발전되어온 원칙이며, 이러한 원칙은 미국 저작권법 제107조에 성문화된 바 있다. 미국 저작권법 제107조는 공정이용의 원칙을 적용하기 위하여 저작물 이용의 목적과 성격(상업성 또는 영리성 등을 포함하여)을 고려해야 하며, 이용된 저작물의 성격, 이용된 저작물의 양과 질, 저작물 이용이 잠재적 시장에 미치는 영향을 고려해야 하는데, 이러한 4가지 요건은 사건마다 종합적으로 고려되어야 한다. 이러한 이유로 공정이용원칙은 사실(fact)과 법률(law)이 혼합된 문제로서의 특징이 있다고도 설명된다.

아래는 공정이용의 4가지 요건에 대한 뉴욕지방법원의 판단과 시사점을 정리한 것이다.

(1) 공정이용의 첫 번째 요건 - 이용의 목적

공정이용의 첫 번째 요건인 ‘이용의 목적’은 이용행위가 상업적인지 또는 비영리적이거

1) CDL이란 ‘소유한 도서의 수 대(對) 디지털방식의 대여 비율(owned-to-loaned ratio)’을 담보하는 방식의 라이선스를 말한다(<https://controleddigitallending.org/> (2023. 06. 25. visited)). 인터넷 아카이브는 CDL에 따라 도서관이 보유하고 있는 종이책의 수 범위 내로 디지털 사본을 이용자에게 제공하였고, 이용자가 다운로드한 디지털 사본은 DRM(Digital Right Management) 기술을 적용하여 14일이 지나면 접근을 차단하였다.

나 교육목적인지 등을 묻는 것인데, 최근에는 ‘변형적 이용(transformative use)’에 해당하는지가 중요한 고려사항이 되고 있다. 변형적 이용이 되기 위해서는 새로운 것이 부가됨으로써 저작물 이용이 본래 예정되었던 것과 다른 목적이나 다른 성격을 갖는 것을 의미하는데, 새로운 표현이나 의미 또는 메시지로 변경하는 것이 그 예이다. 이는 저작물의 본래 쓰임새를 넓히는 역할을 하게 된다.

이러한 의미에서 뉴욕지방법원은 IA의 이용에는 변형적 이용이라고 할 만한 새로운 것이 없다고 판결하였다. 또한, 뉴욕지방법원은 IA의 이용이 유용성을 증가시키는 결과로 이어지지 않았다고 판단하였다. IA의 이용행위는 도서를 그대로 복제하고 디지털화한 것에 불과하므로 타당한 판결이다.

그러면서 뉴욕지방법원은 구글북스 판결을 토대로 검색목적을 위한 데이터베이스 구축은 공정이용에 해당한다고 실시한다. 새로운 것이 부가되었고, 본래 저작물 이용으로부터 예정되었던 쓰임새와는 다른 쓰임새를 창출했다는 것이다. 아울러, 뉴욕지방법원은 구글이 구축한 소위 ‘스니펫 뷰(snippet view: 전체 도서의 몇몇 줄만을 미리보기 하는 것)’는 공정이용에 해당될 수 있으나, 디지털로 변환한 저작물 전체를 공중에게 제공하였다면 저작권 침해에 해당할 수 있었다고 설명한다.

이러한 설명은 인공지능(AI) 딥러닝을 위한 목적의 데이터베이스 구축 및 저작물 이용에 대한 법원의 태도를 엿볼 수 있게 하는 대목이다. 인공지능 산업계에 긍정적인 관점에서 해석론은 인공지능 딥러닝은 인간을 위한 저작물 이용과 다른 것이면서 쓰임새를 넓힌 이용행위로 변형적 이용이 인정될 수 있기 때문이다. 또한, 딥러닝에 이용된 저작물은 인간들을 대상으로 하는 공중에게 제공된 것으로 볼 수 없으므로 공정이용에 해당될 수 있다는 설명이 가능하기 때문이다. 반면, 딥러닝을 하는 경우에도 도서가 그대로 이용되고 있다는 점에서 공정이용 인정에 부정적인 의견도 제시될 수 있다.

(2) 공정이용의 첫 번째 요건 - 이용의 성격

IA는 자신들의 도서관 프로젝트가 전적으로 비상업적(non-commercial)인 것임을 강조하였다. IA 역시 비영리기관(non-profit organization)이다.

그럼에도 불구하고 뉴욕지방법원은 이를 공정이용으로 보는 데 긍정적 근거로 보지 않았다.

오히려 뉴욕지방법원은 이와 관련하여 “영리/비영리 구분의 핵심은 사용의 유일한 동기가 금전적 이득인지 여부가 아니라 사용자가 관례적인 대가를 지불하지 않고 저작물을 이용하여 이익을 얻을 수 있는지 여부”에 있다고 설명하고 있는데, 영리 또는 비영리 여부에 대한 판단기준으로서 사용자 관점에서의 판단이 가능함을 실시하고 있음은 매우 중요한 의미가 있다.

즉, 수집하여 배포하는 자가 비영리기관인 경우에도 저작물 이용자는 그들이 지급해야 할 저작물 이용료를 내지 않음으로써 이익을 얻고 있다는 점을 고려해야 한다는 것인데, 이는 이용자가 업으로 이용하는지를 묻지 않고 상업적 목적을 인정하고 있다는 점을 유념하여야 한다.

이를 인공지능 딥러닝에 적용하면, 인공지능 사업의 영리성 여부에 관계없이 인공지능 사업자의 입장에서는 저작물 이용료를 내지 않음으로써 이익을 얻고 있으므로 상업적 목적이 인정될 가능성이 있다.

(3) 공정이용의 첫 번째 요건 - 최초판매의 원칙

최초판매의 원칙(first sale doctrine)은 미국 저작권법 제109조(a)를 근거로 하고 있으므로, 공정이용의 원칙(fair use doctrine)을 규정한 제107조의 첫 번째 요건과는 그 근거가 상이하다. 그럼에도 불구하고 뉴욕지방법원은 공정이용의 첫 번째 요건과 관련하여 최초판매의 원칙을 검토하고 있다.

이에 대한 결론은 다소 싱거운데, 최초판매의 원칙은 적법하게 취득한 복제본 등의 대여, 재판매 등에 적용되는 데 반해, IA의 복제는 권한이 없는 것으로서 적법하게 취득한 복제물이 아니므로 최초판매의 원칙을 적용할 수 없다는 것이다. 상술하면, 도서 그 자체의 구매는 적법하게 이루어졌으므로, 도서 그 자체의 재판매 또는 대여는 최초판매의 원칙에 따라 허용되는 것이나, 적법하게 구매한 도서라고 하더라도 도서의 복제를 통한 복제물의 생성은 저작권자에 의해 허락된 바가 없으므로, 복제물은 적법하게 취득한 것이 아니라는 것이다. 따라서 도서 그 자체는 최초판매 원칙의 적용을 받게 되지만, 도서의 디지털 복제물은 최초판매 원칙의 적용을 받지 않게 된다.

아쉬운 점은 뉴욕지방법원의 판결은, 적법한 저작물 취득 여부에 관해 판단함으로써 최초판매 원칙을 적용하지 않는다는 결론에 도달하고 있으므로, 해당 판결은 최초판매의 원칙이 디지털 복제물에도 적용되는지와 관련한 디지털 권리소진원칙에 관하여 명확한 입장을 보여주지 않는다는 점이다.

(4) 공정이용의 두 번째 요건 - 저작물의 성격

공정이용의 두 번째 요건인 저작물의 성격은 저작물이 사실적인지 또는 창작적인지, 출판된 것인지 또는 아닌지와 같이 저작물의 특성을 고려해야 한다는 취지의 요건이다. 즉, 저작물이 창작적인 경우라면 더 많은 창작적 표현형식을 포함하고 있을 것이므로 보호의 범위와 수준이 넓고 두터울 것인데 반해, 저작물이 사실적인 것으로 창작적인 것이 그리 많지 않다면 당연히 보호될 창작적 표현형식이 그리 많지 않게 될 것이라는 점을 반영한 것이다. 출판된 것이라면 공정이용의 가능성이 커지는 한편, 출판되지 않은 것이라면 공개

여부에 관한 저작자의 의도가 간접적으로 인정될 수 있을 것이므로 공정인정이 어렵게 될 것이다.

이번 사건에서 IA는 출판된 도서를, 창작적인 도서를 다수 포함하여 이용하고 있으므로, 깊이 논의할 필요 없이 공정이용 인정에 유리한 요건은 아닌 것으로 판단된다.

뉴욕지방법원 역시 “공정이용의 두 번째 요건은 저작권을 침해하는 검은 염소로부터 공정이용이라는 하얀 양을 분리하는 데 도움이 되지 않는다”라고 설명하고 있다.

(5) 공정이용의 세 번째 요건 - 이용된 저작물의 양과 질

공정이용의 세 번째 요건 역시 깊은 논의가 필요 없이도 판단이 가능하다.

IA가 복제하고 전자책으로 제작한 도서의 수가 33,000개의 출판사들로부터 출판된 3백60만 권의 저작권에 의해 보호받는 도서들에 이르고 있으며, 이용하는 방법 역시 도서 전체를 복제하는 방법으로 이루어지기 때문이다. 즉, IA가 이용한 도서의 양은 3백만 권을 초과하고 있으며, 질적인 관점에서도 도서 전체를 복제하고 있다. 이러한 경우 공정이용의 세 번째 요건은 출판사에게 매우 유리한 요건이 될 것이다.

뉴욕지방법원 역시 “공정이용의 세 번째 요건은 출판사들에게 매우 유리하다”는 점을 적시하고 있다.

(6) 공정이용의 네 번째 요건 - 잠재적 시장에 미치는 영향

저작물 이용이 저작자가 이익을 얻을 예정인 시장에 영향을 미치는 것은 본래 저작권법이 보호하려 했던 법익을 침해할 가능성이 높게 되므로, 본 요건 역시 공정이용 여부를 판단하는 데 있어 중요한 요건으로 평가된다. 특히, 변형적 이용인 경우 저작자가 이익을 얻을 예정인 시장에 미치는 영향은 제한적일 것이므로, 공정이용의 네 번째 요건은 공정이용의 첫 번째 요건과 관련성이 크다. 또한, 이용되는 저작물의 양과 질 역시 저작자가 이익을 얻을 예정인 시장에 미치는 영향이 클 것이므로 공정이용의 네 번째 요건은 공정이용의 세 번째 요건과도 관련성이 크다.

이러한 점을 고려할 때, 잠재적 시장에 미치는 영향은 공정이용을 부정하는 출판사들에게 유리하게 작동할 가능성이 커진다. 공정이용의 첫 번째 요건과 세 번째 요건 모두 공정이용을 부정하는 출판사들에게 유리하게 해석될 수 있음은 앞서 기술한 바와 같다.

뉴욕지방법원은 IA의 열린 도서관 프로젝트 등의 이용으로 동일한 도서를 연이어 구매해야 할 필요성을 감소시키고 있다고 설명하고 있는데, 특히 전자책과 같이 구현 형식이 다르다는 이유로 도서관이나 사용자가 도서를 구매할 필요성을 약화시키고 있음을 강조하였다.

4. 관련 논의와 시사점

이번 판결은 디지털 4.0시대와 인공지능 시대를 맞이하여 미국 법원이 공정이용의 원칙에 관한 인정 범위를 넓힐 것인지에 세간의 관심이 모아졌다. 피고인 IA는 지식에의 보편적 접근이라는 인류사회의 보편적 원칙에 근거하고 있고, COVID-19 팬데믹이라는 특수한 상황에서 디지털 시대가 속성되어 가고 있는 상황이었기 때문이다. 어떻게 보면, 도서관에 보관되어 있는 출력된 저작물의 숫자만큼만 전자책으로 대여하면서 저작권자의 저작권 행사에 대비하던 IA가 COVID-19 팬데믹을 계기로 공정이용의 범위를 확장하기 위한 시험적인 행위를 하였다고 이해할 수도 있다.

그러나 뉴욕지방법원은 철저히 지금까지 축적된 공정이용 원칙을 원용한 판례들을 토대로 이번 사건을 판단하였다. 변형적 이용, 이용행위의 상업성, 이용된 저작물의 성격, 이용된 저작물의 양과 질, 이용행위로 인해 영향받는 잠재적 시장에 관한 판단 모두 선행 판례를 태도로 하고 있다. 아울러, 기존 판례 중 구글 판결, 소니 판결이 이번 사건의 해당 판결과 다른 점을 설명하면서, 기존 판례의 사건이 이번 사건과 다르다는 점을 명확히 설명하고 있다.

이를 요약하면, 도서를 디지털화하여 목록을 정리하거나 도서를 찾을 수 있도록 검색 정보를 제공하는 행위는 공정이용에 해당하지만, 도서 그 자체를 이용하도록 하는 행위는 공정이용으로 볼 수 없다는 것이다. 영리성을 갖는지 여부는 저작물을 이용함으로써 저작물 구매를 대체할 수 있는 자가 존재하는 경우 인정될 수 있음을 분명히 하고 있다. 또한, 디지털화하여 유통하는 행위는 디지털화된 대상이 유체물이건 무체물인 디지털 파일이건 상관없이 적법한 권한을 갖는 자에 의한 복제가 아니므로, 최초판매의 원칙 또는 권리소진의 원칙이 적용될 여지가 없음을 확인하고 있다.

이러한 점을 고려할 때, 인공지능 딥러닝 또는 데이터 분석을 위한 웹 크롤링 등의 경우에도 현재의 공정이용의 판단기준을 토대로 저작권 침해 여부를 판단하여야 한다. 공정이용원칙의 적용은 사실과 법률의 혼합적이면서 종합적인 고려가 필요한 영역이므로, 개별 사건 별로 결과가 상이할 수 있다. 생각건대, 저작물 이용의 주체는 인간이며, 저작물 이용을 위한 공중 역시 인간을 전제하고 있으므로, 디지털화된 저작물이 TDM(Text and Data Mining)을 넘어 인간의 표현적 향유가 존재하지 않는 한 그러한 저작물 이용은 공정이용으로 볼 가능성이 클 것으로 생각된다. 이번 판결은 공정이용과 저작권 침해의 경계를 분명히 하는 판례로서 그 의미가 있다.