

크롤링을 통한 데이터 수집의 형사책임

- 대법원 2022. 5. 12. 선고 2021도1533 판결을 중심으로 -

김 태 균 *

- I. 사안의 개요
- II. 웹 크롤링과 관련 법적 쟁점
- III. 대법판결에 나타난 크롤링을 통한 데이터 수집의 형사책임
- IV. 자동화 도구를 사용한 정보통신망 이용
- V. 맺음말

국문초록

대법원은 숙박업체 정보 제공 서비스를 운영하는 회사의 임직원들이 경쟁자인 피해자 회사의 모바일 어플리케이션 서버에 접속하여 크롤링으로 숙박업소 목록 등 피해자 회사가 보유·관리하는 데이터베이스를 복제한 사안에서 무죄를 선고한 원심을 확정하였다. 크롤링이란 크롤러라는 정보수집 로봇을 이용하여 다른 사이트에서 필요한 정보를 수집하는 행위를 말한다. 데이터 수집 및 활용이 중요한 데이터 경제 시대에 크롤링은 이미 보편화된 기술 중 하나가 되었음에도 크롤링에 대한 형법적 평가는 거의 없었다. 대법판결은 ① 정보통신망에 대한 서비스제공자가 접근권한을 제한하고 있는지 여부는 보호조치나 이용약관 등 객관적으로 드러난 사정을 종합적으로 고려하여 판단하여야 하고, ② 데이터베이스의 상당한 부분의 복제는 양적인 측면과 질적인 측면을 모두 고려해야 한다는 내용의 새로운 법리를 설시하면서, 크롤링을 통한 데이터 수집에 있어 정보통신망침입, 데이터베이스무단복제, 컴퓨터등장애업무방해 모두 무죄라

* 서울지방법원 판사

고 실시하였다. 대상판결은 크롤링을 통한 데이터 수집의 형사책임에 관한 대법원의 최초의 판결이라는 데에 의의가 있다.

- **핵심어** 크롤링, 데이터, 데이터 수집, 정보통신망침해, 데이터베이스, 데이터베이스 복제, 업무방해

I. 사안의 개요

1. 공소사실의 요지¹⁾

피고인들은 숙박업체 정보 제공 및 예약 서비스를 제공하는 회사(이하 ‘피고인 회사’라고 한다)의 임직원들로서 경쟁 관계에 있는 피해자 회사가 운영하는 모바일 어플리케이션인 ‘바로예약’(이하 ‘이 사건 앱’이라고 한다)이나 PC용 홈페이지에 접속하여 제휴 숙박업소 목록, 주소 정보, 가격정보 등을 확인하고 있었다.

(1) 정보통신망이용촉진및정보보호등에관한법률위반(정보통신망침해등)

피고인들은 이 사건 앱의 프로그램 소스를 ‘패킷캡처’ 앱²⁾을 통하여 그 모바일 앱 프런트엔드(FRONT-END) API 서버(이하 ‘이 사건 API 서버’라고 한다)³⁾의 모

- 1) 논의의 편의성을 위하여 이 글의 주제와 관련이 있는 부분으로 한정하여 요약하였다. 이하에서 나오는 재판례의 사실관계는 모두 관련 쟁점을 파악하는 데에 필요한 한도에서 글쓴이가 요약·정리하였다.
- 2) 패킷캡처(Packet Capture, 또는 packet analyzer, packet sniffer, protocol analyzer, network analyzer라고도 한다)는 컴퓨터 네트워크상에서 이동하는 패킷 또는 트래픽을 가로채거나 저장할 수 있는 프로그램이다. 사용법도 어렵지 않고 널리 사용되고 있는 프로그램이다.
- 3) 데이터 센터에서 모든 데이터를 처리할 때 과부하와 그에 따른 지연 현상을 완화하기 위하여 사용자 또는 엔드포인트(endpoint) 단말기의 물리적 위치와 인접한 곳에서 컴퓨팅을 수행하는 것을 엣지(edge) 컴퓨팅이라고 한다(정재화, 『클라우드 컴퓨팅』, 한국방송통신대학교출판문화원, 2020년, 155쪽). 다시 말해 엣지 컴퓨팅에서는 사용자와 가장 가까운 서버에서 데이터를 처리하고 서비스를 제공하여 서비스의 지연을 방지하는 것이다. 엣지 컴퓨팅에서 최종 사용자와 가장 밀접한 계층으로 엣지 단말 기기로 연산 처리를 할 수 있는 능력이 있는 경미한 수준의 장치를 프런트 엔드(Front-End)라고 한다(정재화, 전제서, 160쪽).
API(Application Programming Interface)는 컴퓨터나 컴퓨터 프로그램 사이의 연결로서 일종의 소

둘, 해당 서버의 URL 주소 및 위도, 경도, 반경, 입실 날짜, 퇴실 날짜 등 API 서버로 정보를 호출하는 명령구문들을 알아내었다.

피고인들은 PC를 통하여 이 사건 API 서버 URL 주소에 마치 정상적인 이용자가 이 사건 앱을 이용하는 것처럼 이 사건 API 서버로 정보를 호출하는 명령구문을 입력하는 방식으로 접근하면서, 특정 위치로부터 일정 반경 내에 있는 숙박업소 정보를 모두 불러오는 기능⁴⁾을 특정 URL에 탑재한 프로그램(이하 ‘이 사건 크롤링 프로그램’이라고 한다)을 개발하였다. 피고인들은 2016. 6. 1.경부터 같은 해 10. 3.경까지 이 사건 크롤링 프로그램을 이용하여 일 1~2회 가량 이 사건 API 서버에 접근하여 제휴 숙박업소 업체명, 주소, 방 이름 등의 정보를 무단으로 복제하였다.

(2) 저작권법위반

피고인들은 위와 같은 방법으로 피해자 회사의 데이터베이스 중 제휴 숙박업소 업체명, 주소, 방 이름, 원래금액, 할인금액, 업체주소, 입실시간, 퇴실시간, 날짜와 같은 상당한 부분을 무단으로 복제하여 데이터베이스제작자의 권리를 침해하였다.

(3) 컴퓨터등장애업무방해

피고인들은 위와 같은 방법으로 정보처리장치에 부정한 명령을 입력하여 장애가 발생하게 함으로써 피해자 회사의 숙박 예약에 관한 업무를 방해하였다.

소프트웨어 인터페이스를 말한다. 사용자 인터페이스(user interface)가 컴퓨터와 인간을 연결하는 것을 의미하는 것에 비하여, API는 컴퓨터나 소프트웨어를 서로 연결하는 것으로 최종 사용자(사람)가 직접 사용하기 위하여 고안된 것이 아니며, 대신 소프트웨어에 이를 통합하고자 하는 컴퓨터 프로그래머가 사용하도록 고안된 것이다(위키백과, <https://ko.wikipedia.org/wiki/API>, 2022. 8. 3. 방문).

여기서 ‘모바일 앱 프런트엔드 API 서버’라고 함은 이 사건 앱의 사용자 인터페이스와 직접 연결된 서버를 의미한다. 이 사건 앱은 프런트엔드 API 서버가 직접 연결되어 있고, 최종적인 데이터베이스 서버까지는 다른 API 서버들이 순차적으로 연결되어 있었던 것으로 보인다.

- 4) 피고인 회사를 기준으로 반경 1,000km 이내에 있는 숙박업소 정보를 불러오도록 하였다. 이 사건 앱은 이용자의 위치로부터 7km 또는 30km의 범위 내의 숙박업소 검색만 가능하였다.

2. 사건의 경과

(1) 제1심(서울중앙지법 2019고단1777)의 판단

제1심은 공소사실 전부를 유죄로 판단하였는데, 그 주된 근거는 아래와 같다.

1) 정보통신망침입의 점의 점

① 피해자 회사는 이 사건 API의 모듈, URL 주소 및 이 사건 API 서버로 정보를 호출하는 명령구문들을 외부에 공개하지 않았고, 정상적인 이용자는 이 사건 앱을 통하지 않고서는 이 사건 API 서버에 접속할 수 없었다. ② 당시 피해자 회사의 서비스 이용약관은 일체의 가공행위를 통하여 서비스를 복제, 분해 또는 모방 기타 변형하는 행위 및 자동 접속 프로그램 등을 사용하는 등 정상적인 용법과 다른 방법으로 서비스를 이용하여 회사의 서버에 부하를 일으켜 회사의 정상적인 서비스를 방해하는 행위 등을 금지하고 있었는데, 이는 비회원에게도 적용된다. ③ 피해자 회사는 이 사건 크롤링 프로그램 이용으로 인한 대량 호출 신호를 감지하고 피고인 회사가 이용하는 아마존 웹서비스 클라우드 서버의 IP 주소를 수차례 차단하였는데, 이에 피고인들은 서버의 전원을 차단하였다가 다시 켜는 방식으로 IP 주소를 변경하여 위 차단을 회피하였다.

2) 저작권법위반의 점

① 피고인들은 6개월여 동안 264회에 걸쳐 피해자 회사의 데이터베이스를 무단 복제하였다. ② 피고인들은 경쟁에서 우위를 점하기 위해서 반복적·조직적으로 피해자 회사의 데이터베이스를 무단으로 복제하였다.

3) 컴퓨터등장애업무방해의 점

① 이 사건 앱은 자신의 위치로부터 7km 또는 30km의 범위 내의 숙박업소만 검색이 가능하였다. ② 피고인들은 이 사건 API 서버에 접속하여 전국에 있는 모든 숙박업소 정보를 요청하여 대량의 정보 호출을 발생시켰다.

(2) 원심(서울중앙지방법원 2020노611)의 판단

원심은 제1심 판결을 파기하고 공소사실 전부를 무죄로 판단하였다. 그 주된 근거는 아래와 같다.

1) 정보통신망침입의 점

① 이 사건 앱이나 API의 접속에 회원 가입이나 비밀번호가 필요하고, 패킷캡처는 통상 이용되는 프로그램이다. ② 피해자 회사가 의도적으로 이 사건 API 서버의 URL을 숨기지 않았고, 이용자의 이 사건 API 서버에 대한 직접 접속을 금지하는 기술적 조치도 없었다. ③ 피고인들이 크롤링을 통하여 가져간 정보들은 피해자 회사가 자신의 숙박 예약 영업을 위하여 이용자들에게 공개한 것들이다. 크롤링이나 지역 범위 검색 명령어를 확장하지 않고도 같은 종류와 양의 정보들을 가져올 수 있다. ④ 피고인들이 사용한 API 서버의 명령구문은 이 사건 API 서버가 허용한 것이다. 피고인들이 그 검색 범위를 넓혔다는 것만으로 접근권한을 넘었다고 볼 수 없다. ⑤ 약관은 문언상 피해자 회사의 회원들에게 적용된다. ⑥ 피해자 회사가 피고인들의 반복적인 접근에 따른 대량 호출 신호를 감지하고 피고인들의 IP 주소를 수차례 차단한 것만으로 피고인들의 접근을 일률적으로 제한한 것이라고 할 수 없다.

2) 저작권법위반의 점

① 피고인들이 피해자 회사의 숙박업소 데이터베이스의 50여 항목 중 수집한 세부 항목은 ‘업체명, 주소, 지역, 방이름, 원래금액, 할인금액, 날짜, zone, 이용시간, 예의사항, 입실시간, 퇴실시간, 대실가격, 숙박가격, 타입, 카테고리’로서, 한 번에 수집한 것은 3개 내지 8개 항목이다. ② 위 8개 항목 중 숙박업소의 업체명, 업체 주소, 지역, 타입 등은 이미 상당히 알려진 것이고, 할인금액, 대실가격, 숙박가격 등은 피해자 회사가 영업을 위해서는 공개할 수밖에 없다. ③ 이와 같은 정보들은 피해자 회사가 가진 숙박업소에 대한 50여개의 데이터 항목 중 이용자에게 공개하였고, 이 사건 앱을 통해서도 확보할 수 있었다.

3) 컴퓨터등장애업무방해의 점

① 이 사건 API 서버의 목적은 주어진 명령구문에 대응하는 숙박업소 정보를 반환하는 것이다. 피고인들은 이 사건 API 서버의 명령구문에 정보를 입력하여 숙박업소의 정보를 전송받았으므로 이는 ‘허위의 정보 또는 부정한 명령의 입력’에 해당하지 않는다. ② 이 사건 앱에 접속하는 데 장애가 발생한 공소사실 기재 일자들은 순차로 토요일, 토요일, 일요일, 추석, 토요일로서 자연 이용자 증가에 따른 것이었을 수 있다.

II. 웹 크롤링과 관련 법적 쟁점

1. 웹 크롤링의 개념과 기술적 원리

웹 크롤링은 간단히 말해 웹 페이지를 통째로 다운로드 받는 프로그램이다. 웹 크롤링 프로그램인 크롤러에게 시드(seed) URL 즉 출발점이 되는 URL을 주면 크롤

러는 그 URL상의 모든 웹 페이지를 다운로드 하는데, 그 웹 페이지에서 다른 웹 페이지를 연결해 놓은 하이퍼링크들을 다시 추출하여 하이퍼링크로 연결된 웹 페이지들을 다운로드한다. 크롤러는 이러한 과정을 반복하여 데이터를 수집한다.⁵⁾ 크롤링의 등장은 웹의 등장과 거의 일치한다. Matthew Gray가 1993년 통계 목적으로 작성하여 사용했던 것이 최초의 크롤링 프로그램으로 알려져 있다.⁶⁾ 구글의 뉴스 서비스 방식은 검색로봇을 이용하여 다른 웹사이트에 게재된 뉴스 기사를 크롤링한 후 크롤링 정보를 이용하여 뉴스 기사 페이지(기사제목과 기사의 첫 단락 2~3줄 제공)를 구성하여 링크 정보를 제공해 주는 형태를 취하게 되는데, 이처럼 크롤링은 인터넷에서 정보 수집을 위해서 일반적으로 사용되는 기술이다.⁷⁾

웹 크롤링과 유사한 기술로 웹 스크래핑(web scraping)이 있다. 웹 스크래핑은 웹브라우저 화면에 표시되는 다양한 정보 중 사용자가 지정하거나 필요한 정보만 추출하여 가공, 저장하여 사용자에게 제공하는 기술임에 반하여 웹 크롤링은 특정 형태의 정보를 웹에서 자동으로 수집하여 정보를 최신으로 유지하는 소프트웨어 기술이라고 할 수 있다. 웹 스크래핑이나 웹 크롤링의 도구인 웹 크롤러(web crawler)는 조직적이고 자동화된 방법으로 월드 와이드 웹을 탐색하는 컴퓨터 프로그램으로 웹 크롤러는 앤티(ants), 자동 인덱서(automatic indexers), 봇(bots), 웜(worms), 웹 스파이더(web spider), 웹 로봇(web robot) 등으로도 불린다. 웹 스크래핑과 웹 크롤링을 엄밀하게 구분하면, 웹 스크래핑은 웹 문서를 가공하여 정보를 추출하는 과정으로서 웹 크롤링을 하지 않고도 웹 스크래핑을 할 수 있음에 반하여 웹 크롤링은 기초가 되는 시드 URL을 저장한 뒤 웹 페이지의 하이퍼링크를 인식하여 URL을 갱신하며 반복적으로 웹 링크(web link)를 찾는 과정이라고 할 수 있다.

5) Marc Najork, Web Crawler Architecture(<https://marc.najork.org/pdfs/eds2009a.pdf> 2022. 8. 5. 방문); Christopher Olston, Marc Najork, Web Crawling, Foundations and Trends® in Information Retrieval, Vol. 4, No. 3, 2010(http://infolab.stanford.edu/~olston/publications/crawling_survey.pdf 2022. 8. 5. 방문)

6) Ibid

7) 유대중, “웹검색 서비스와 ISP 책임에 관한 소고”, 장작과 권리, 2007년, 103쪽

다시 말해 웹 스크래핑은 웹으로부터 내용을 추출하는 것(pulling content from a page)이고, 웹 스크롤링은 수많은 웹 페이지에 이르기 위해 웹 링크를 따라가는 과정(following links to reach numerous pages)을 의미하는 것으로 구분할 수 있다.⁸⁾ 이를 검색엔진에서의 기능을 중심으로 보면 ‘웹 크롤링’은 검색엔진에서 가장 최신의 검색 결과를 보여주기 위해서 웹 페이지들을 돌아다니면서 새로운 페이지와 콘텐츠를 추출하는 것을 의미하고, ‘웹 스크래핑’은 대상이 된 사이트를 위해서 특별히 고안된 수단에 의하여 웹 페이지로부터 구조화된 정보(structured information)를 추출하는 과정을 말한다. 데이터 수집의 관점에서는 스크래핑이 웹 페이지의 내용 전체를 웹 코드까지 가져오는 것이라면, 웹 크롤링은 이에 더하여 웹에서 공개된 정보를 데이터화하는 것까지 포함한다고 한다.⁹⁾ 이러한 개념상의 차이는 있으나 실제에서는 웹 스크래핑이나 웹 크롤링 모두 웹에 존재하는 데이터의 수집을 위한 목적으로 사용되고 있어서, 양자를 엄밀하게 구분하지 않고 사용되고 있다.

이러한 크롤링은 인터넷상에 공개된 웹서버에 접속하여 웹 페이지 정보를 수집한다는 점에서 일반적으로 컴퓨터를 이용하여 다른 사람의 정보처리장치 또는 정보처리조직에 침입하거나 기술적인 방법으로 다른 사람의 정보처리장치가 수행하는 기능이나 전자기록에 함부로 간섭하는 일체의 행위인 해킹과는 구별된다고 본다.¹⁰⁾ 그러나 그와 같은 구별은 인터넷 기술의 원리 및 사용 목적에 따른 일반적인 구별일 뿐이고, 크롤링이 해킹의 과정 또는 방법의 하나인 정보통신망 침입이 되지 않는다고 법률적으로 밝힌 것은 대상관결이 최초이다.

2. 웹 크롤링을 통한 데이터 수집

본래 웹 크롤링은 인터넷상에서 검색 서비스를 위하여 활용되었다. 이것이 점차

8) 최대우 외, 『데이터과학입문』, 한국방송통신대학교출판문화원, 2020년, 76~77쪽

9) 김현숙, “크롤링을 이용한 공개데이터 수집·활용의 법적 쟁점에 대한 비판적 검토”, 강원법학 제 61권, 2020년, 227쪽

10) 유대중, “웹검색 서비스와 ISP 책임에 관한 소고”, 장작과 권리, 2007년, 122쪽

로 확대되어 연구 목적, 서비스 제공 또는 데이터 분석을 통해 비즈니스에 활용하기 위해서 등 다양한 목적으로 크롤링 기술을 이용하여 데이터를 수집한다. 검색 서비스 제공을 위해서는 크롤링이 필수이며, 구글, 네이버, 다음 등 일상적으로 사용하는 검색엔진은 웹 로봇(web robot)을 이용하여 정보를 수집하고 색인하여 정리하는 대표적 서비스이다.

또한 일반적인 검색사이트 외에 서로 다른 쇼핑몰의 각종 상품과 가격정보를 정리한 가격비교 사이트도 검색서비스의 다른 형태로 볼 수 있다. 이 역시 관련 사이트를 크롤링하여 비교하고 결과를 제공하는 것으로 데이터를 수집하여 제공하는 서비스 중 하나이다. 비정형데이터를 분석하여 사회적 이슈나 흐름을 알려주는 서비스들은 페이스북, 트위터, 인스타그램 등 각종 SNS의 데이터 패턴을 찾아내고 분석하기 위해서 크롤링 기술을 이용한다. 또한 경쟁회사는 벤치마킹을 위해서 상대방의 웹상 공개데이터를 수집하거나 분석하여 활용하고, 스타트업이 비즈니스를 준비할 때도 동종 업계의 데이터 수집과 분석이 필수적으로 이루어진다.

3. 웹 크롤링에 대한 법적 규제

기본적으로 크롤링의 사용이 웹에서 검색을 위한 것인 만큼 대부분의 웹사이트들은 크롤링을 허용하고 있다. 즉 검색엔진의 크롤러들이 사이트를 발견하여 크롤링함으로써 해당 웹사이트가 검색엔진에 검색될 수 있도록 하기 위한 것이다. 그런데 경쟁사의 크롤러들이 웹사이트에 들어와 정보를 다운로드해가는 상황이 발생하고 있고, 그것이 일반적인 영업 전략을 세우기 위한 시장조사 또는 자료조사의 기본적인 방법으로 인식되고 있어서 관련 사업자들 사이에 분쟁의 소재가 되고 있다.

기술적으로는 크롤러가 웹에 접근하지 못하도록 로봇 배제 표준(robot exclusion standard), 로봇 배제 프로토콜(robot exclusion protocol)을 사용할 수 있다. 로봇 배제 표준은 웹사이트에 로봇, 즉 크롤러가 접근하는 것을 방지하기 위

한 규약으로, 일반적으로 접근 제한에 대한 설명을 robots.txt에 기술한다. 로봇이 robots.txt 파일을 읽고 접근을 중지하는 것을 목적으로 한다. 따라서, 접근 방지 설정을 하였다고 해도, 다른 사람들이 그 파일에 접근할 수 있다. robots.txt 파일은 항상 사이트의 루트 디렉토리에 위치해야 한다.¹¹⁾ 그러나 로봇 배제 표준은 권고안(Voluntary standard)에 불과하여 굳이 사이트를 크롤링하려는 모든 로봇들을 완벽하게 차단하지는 못한다.¹²⁾ 그리고 설사 그와 같이 완벽하게 크롤러들을 차단할 수 있는 기술적 방안이 마련된다고 하더라도 그것은 웹사이트를 운영하는 사업자의 입장에서 검색엔진의 검색 대상에서 자신의 웹사이트를 배제하게 되어 이용자의 접근을 막는 결과가 되기 때문에 쉽사리 적용할 수도 없을 것이다. 여기가 크롤링에 대한 법적 문제가 제기되는 지점이다.

빅데이터와 인공지능의 역할이 부상하고 있는 인더스트리 4.0¹³⁾에서 대량의 데

11) 예컨대 모든 크롤러의 접근을 금지하려면 robot.txt에 다음과 같이 입력한다.

(https://ko.wikipedia.org/wiki/%EB%A1%9C%EB%B4%87_%EB%B0%B0%EC%A0%9C_%ED%91%9C%EC%A4%80 위키백과, 2022. 8. 7. 방문)

User-agent: * Disallow: /

12) 최상진, “경쟁사의 무단 크롤링에 대한 법적 대응방안에 관한 연구”, Law & Technology 제17권 제1호, 2021년, 21쪽

13) 통상 소위 ‘4차 산업혁명’이라는 용어가 많이 사용되는데, 필자는 그러한 용어법에 동의하지 않는다. 그 이유는 다음과 같다. 첫째, 국내에는 위 용어를 2016년 다보스 포럼에서 최초로 사용된 것으로 알려져 있으나, 이는 정확하지 않다. 그 전에 이미 독일에서 ‘인더스트리 4.0(Industrie 4.0)’이라는 이름으로 국가경쟁력 발전을 위한 사업으로 진행하여 왔던 것을 다보스 포럼에서 차용된 것으로 봄이 사실에 더 가까운 관찰이다. 둘째, 위 용어는 바른 개념이라고 할 수 없다. 인더스트리 4.0은 정보기술, 인공지능 등의 발달에 적응하여 독일의 경쟁력을 살리자는 의미에 더하여 그로 인하여 노동자들의 처지가 열악해지는 것을 인식하고 그것을 막으려는 노력을 포함한다. 위 용어는 그와 같은 노동자 및 사회적 약자를 보호하지는 의미를 거의 상실한 채 사용되고 있다. 셋째, 위 용어는 사실과 가치를 의도적으로 혼동하였다. 새로운 기술의 도입과 그로 인한 사회 변화는 사실 자체일 뿐 그것이 바로 당위나 가치가 될 수 없다. 그와 같은 상황에서 우리는 얼마든지 새로운 가치와 약자를 보호하면서 사회를 구성하려는 노력을 할 수 있고, 그것이 마땅히 우리가 해야 할 당위이자 가치이다. 그런데 위 용어는 정보기술의 발전으로 인한 사회의 변화를 마치 규범으로 받아들이면서 그에 저항하고 그 속도를 늦추려는 노력 또는 거기서 오는 부작용을 제거하거나 최소화하려는 노력은 시대착오적인 반동적이라는 인식을 저변에 깔고 있다. 그러므로 나는 위 용어에 대신하여 ‘인더스트리 4.0’을 사용한다(독일의 인더스트리 4.0 중 노동 분야에서 소외되는 노동자 또는 노동을 위한 독일의 논의에 관해서는 김경래, “독일 Industrie 4.0의 특징: 노동 4.0을 중심으로”, 한독사회과학논총 제28권 제2호, 2018년, 3~26쪽 참조).

이터를 수집하고 처리·분석하는 일은 대부분의 기업활동에서 중요해졌고, 그러한 능력이 기업의 경쟁력을 좌우하고 있다.¹⁴⁾ 따라서 초기에는 거대한 양의 데이터로부터 가치를 추출해 낼 수 있을 만큼의 규모에 누가 먼저 도달하는지, 그리고 그 데이터로부터 가치를 누가 먼저 끌어내는지가 중요하다. 그러한 상황에서는 각 분야에서 빅데이터를 통해 소비자에게 효용을 주는 가치를 만들어 낸 상위 사업자가 만들어지기까지 데이터 수집에 대한 관심은 개인정보 침해에 주로 초점이 맞추어져 있었다.¹⁵⁾ 크롤링은 데이터 수집 기술이므로 크롤링 자체는 위법하다고 할 수 있을지는 회의적이다. 크롤링으로 인한 데이터 수집행위가 문제될 때 크롤링이 금지되지 않는다는 주장보다는 크롤링의 방법으로 데이터를 수집하고 이를 사용하는 행위가 문제된다.

데이터를 집적한 선두업체가 등장하기 시작하면 이제 관심은 경쟁법적 문제이다. 크롤링을 통한 데이터 수입 자체가 사업자의 사업행위 중 중요하고 기본적인 것으로 간주되고, 선두업체와 후발업체 간의 데이터 불균형이 발생하기 때문이다. 크롤링 대상 사이트 운영자 입장에서는 거래거절, 즉 접근제한 및 접근거절행위가 문제될 것이고, 크롤링하는 수집 사이트 운영자 입장에서는 크롤링 행위로 인하여 대상 사이트 운영자의 사업활동을 방해하는 경우가 문제될 것이다. 위법성 판단에 있어 전자의 경우는 경쟁제한성이 문제될 것이고, 후자의 경우는 불공정한 경쟁수단을 사용하는지 수단의 불공정성 관점에서 접근해야 할 것이다.¹⁶⁾ 크롤링과 관련하여 최근 국내 논의에서 자주 인용되고 있는 HiQ Inc. v. LinkedIn 판결¹⁷⁾ 역시 그 적용영역을 따지면 경쟁법 분야에서 발생한 것이다. 위 판결에서 법원은 구인구직 플랫폼인 링크드인(Linkedin)이 보유하고 있는 일부 회원들의 정보를 경

14) 이를 데이터 이코노미(data economy), 또는 데이터 주도 경제(data-driven economy)라고 하는데, 데이터를 기반으로 한 경제 또는 데이터가 사회와 경제를 움직이는 핵심적인 동력이 되는 새로운 시대를 의미하는 것이라고 한다(서울대 법과경제연구센터, 『데이터이코노미』, 한스미디어, 2017년, 4쪽, 245쪽)

15) 이호영, “빅데이터의 경쟁법적 함의에 관한 연구”, 법경제학연구 제15권 제3호, 2018년, 298~299쪽

16) 황태희, “웹 크롤링의 경쟁법적 고찰”, 경쟁법연구, 제43권, 2021년, 183쪽

17) HiQ Labs v. LinkedIn Corp., 31 F.4th 1180 (9th Cir, 2022).

쟁사인 하이큐 랩스(HiQ)가 크롤링하는 것을 제한하는 행위에 대하여 경쟁제한적 행위라고 판단하였다. 이처럼 데이터의 중요성이 부상하면서 크롤링이 데이터 확보 또는 수집의 방법으로 널리 활용되면서, 이는 기업 또는 사업자의 중요한 사업 활동이 되면서 경쟁법 또는 개인정보와 관련되어 문제되어 왔었다. 이미 경쟁법 부분에서는 빅데이터 또는 데이터의 수집에서 발생하는 위와 같은 문제점을 반영하여 소비자 후생의 관점에서 위법성 판단을 해야 한다는 논의가 최소한 10년 전부터 있었던 것으로 보인다.¹⁸⁾

그에 비하여 크롤링에 대한 형사법적인 관심과 논의는 별로 없었던 것으로 보인다.¹⁹⁾ 대상판결은 그런 상황에서 최초로 크롤링을 통한 데이터 수집의 형사책임을 밝힌 데에 의의가 있다.

III. 대상판결에 나타난 크롤링을 통한 데이터 수집의 형사책임

1. 정보통신망침입의 점

(1) 정보통신망법상 정보통신망침입죄

본죄는 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망에 침입하는 행위를 처벌한다. 이는 이용자의 신뢰 내지 그의 이익을 보호하기 위한 규정 이 아니라 정보통신망 자체의 안정성과 그 정보의 신뢰성을 보호하기 위한 것이라고 할 것이므로, 위 규정에서 접근권한을 부여하거나 허용되는 범위를 설정하는 주체는 서비스제공자라 할 것이고, 따라서 서비스제공자로부터 권한을 부여받은 이용자가 아닌 제3자가 정보통신망에 접속한 경우 그에게 접근권한이 있는지 여부는 서

18) 대표적이지 가장 모범적인 것으로 강정희, 『빅데이터를 기반으로 하는 배제남용 행위의 위법성 판단기준 연구 - 소비자 선택 기준의 적용을 중심으로』, 박사학위논문, 서강대학교, 2015년 참조

19) 크롤링에 관한 법학 분야의 연구 내역에 대해서는 권세진 외, “데이터 경제 시대에 있어서 웹 크롤링의 법적 인식에 관한 연구”, 한국산업보안연구, 제11권 제3호, 2021년, 85쪽 참조

비스제공자가 부여한 접근권한을 기준으로 판단하여야 한다.²⁰⁾ 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망 침입에 밀접한 행위를 하면 실행의 착수가 있고, 그와 같이 침입이 완료되면 기수에 이르렀다고 볼 것이고 정보통신망에 침입한 후 시스템상 프로그램을 실행하거나 데이터를 송·수신하는 등의 행위까지 나아갈 것을 요구하지 않는다.²¹⁾

- 20) 대법원 2005. 11. 25. 선고 2005도870 판결, 원심은 피고인의 직속상관인 소령이 소관업무에 관한 보고를 육군본부 등에 대신 발송할 수 있도록 하는 등 업무상 필요에 의해 아이디와 비밀번호를 예하 장교와 사병들에게 공시시킴에 따라 피고인도 이를 알게 되었음을 기화로, 2회에 걸쳐 피고인의 컴퓨터로 소령의 아이디와 비밀번호를 입력하여 육군 웹 메일 등에 접속한 후 소령 명의로 대장인 1군사령관에게 “네놈이 감히 ○대위(피고인을 지칭)를 징계하려고 했던 것에 피가 거꾸로 솟구친다. 장성 하나쯤 인사 처리하는 것은 문제도 아니다.”라는 취지의 이메일을 보냄으로써 정보통신망인 소령의 육군 웹 메일 등에 침입하였다는 공소사실에 대하여, 타인의 아이디와 비밀번호를 이용하여 타인의 정보통신망에 침입하는 행위도 정당한 접근권한 없이 또는 허용된 접근권한을 초과하여 정보통신망에 침입하는 행위에 포함된다는 이유로, 위 공소사실을 유죄라고 인정하였다.

대법원은 피고인의 상고를 기각하면서, 이용자가 자신의 아이디와 비밀번호를 알려주며 사용을 승낙하여 제3자로 하여금 정보통신망을 사용하도록 한 경우라고 하더라도, 그 제3자의 사용이 이용자의 사자 내지 사실행위를 대행하는 자에 불과할 뿐 이용자의 의도에 따라 이용자의 이익을 위하여 사용되는 경우와 같이 사회통념상 이용자가 직접 사용하는 것에 불과하거나, 서비스 제공자가 이용자에게 제3자로 하여금 사용할 수 있도록 승낙하는 권한을 부여하였다고 볼 수 있거나 또는 서비스제공자에게 제3자로 하여금 사용하도록 한 사정을 고지하였다면 서비스제공자도 동의하였으리라고 추인되는 경우 등을 제외하고는, 원칙적으로 그 제3자에게는 정당한 접근권한이 없다고 봄이 상당하다고 하였다.

위 판결의 결론에 대하여 나는 약간의 의문을 갖고 있는데, 그것은 다음과 같다. 위 판결은 ‘서비스제공자에게 제3자로 하여금 사용하도록 한 사정을 고지하였다면 서비스제공자도 동의하였으리라고 추인되는 경우’를 들고 있는데, 그것이 접근권한을 판단하는 데에 과연 적절한 것인가 하는 것이다. 위 판결의 취지는 서비스제공자로부터 정당한 접근권한을 부여받아 아이디와 비밀번호까지 설정한 이용자로부터 아이디와 비밀번호를 제공받아 그것을 사용한 제3자가 이용자로부터 허락받은 범위 밖의 일을 한 경우에는 정보통신망침입이 된다는 것이다. 그런데 만약 위 사안에서 이용자인 소령이 자신의 아이디와 비밀번호를 이용하여 상급자에게 그를 모욕하는 취지의 메일을 보냈다면 그 경우에도 정보통신망침입을 인정할 것인가? 제3자가 상급자에게 위와 같은 메일을 보내는 것과 이용자가 메일을 보내는 것 사이에 어떠한 법적 차이가 있는가? 만약 대법원의 취지가 이용자가 자신의 아이디와 비밀번호를 이용하여 상급자 또는 타인을 모욕하는 메일을 보낸 경우까지 정보통신망침입을 인정하려는 것이라면, 인터넷상의 게시판 또는 포털에서 타인을 모욕하는 내용의 댓글을 게시한 경우에도 모욕죄 외에 정보통신망침입죄도 성립한다는 것인데, 이것은 불합리하게 처벌의 범위를 확장한 것이 아닌가?

한편 위 판결이 정보통신망의 범위를 부당히 확장하였다는 취지의 비판도 있는데(장윤식 외, “정보통신망법 상 정보통신망침입죄에 대한 비판적 고찰”, 경찰학연구, 제14권 제4호, 2014년, 61쪽), 그 핵심은 다음과 같다. 즉 위 판결은 ‘육군 웹메일 및 핸드오피스 계정’을 정보통신망으로 보고 있는데, 웹메일과 핸드오피스 계정은 정보통신망에 연결되어 있는 시스템에 해당할 뿐이라는 것이다.

- 21) 이창범 외, 『이론&실무 정보통신망법』, 박영사, 2021년, 266~267쪽

(2) 접근권한 유무의 판단기준 - ‘객관적으로 드러난 사정’

1) 객관적 상황

대상판결은 접근권한에 관한 보다 구체적인 기준과 그에 대한 발견을 위한 방법론적인 접근법을 제시하였다. 보호조치나 이용약관 등 객관적으로 드러난 여러 사정을 종합하여 고려하여 신중하게 판단하여야 한다는 것이다. 서비스제공자가 접근을 막기 위한 기술적 보호조치를 취하였는지 아닌지, 이용약관에 구체적으로 이용자의 사용범위나 접근권한을 밝혔는지 아닌지 등과 같은 객관적 사정을 기준으로 신중하게 파악하여야 한다는 것이다.

그렇다면 이제 접근권한을 판단하는 데에 객관적으로 외부적으로 드러난 사정을 기준으로 하되, 정보통신서비스 제공자의 주관적 의사는 배제하고 판단하여야 하는 것인가? 그렇지는 않다고 본다. 우선 후술하는 바와 같이 이용약관에 특정 행위 또는 특정 이용자의 이용을 금지하는 내용이 기재되어 있고, 보호조치가 설정되어 있다는 사정만으로 그에 기초하여 접근권한을 해석할 수는 없다고 본다. 이용약관의 내용이나 보호조치의 여부는 접근권한을 해석하는 하나의 기준이 될 뿐이지 절대적인 기준이 될 수 없다.²²⁾ 또한 이용약관이나 보호조치는 모두 서비스제공자가 정보통신망 침입 행위가 있기 전에 일방적으로 결정하는 것이어서 그 자체가 이미 서비스제공자의 주관적 의사가 반영된 것이다. 또한 정보통신망침입죄와 관련하여 대표적 판례로 언급되는 대법원 2005. 11. 25. 선고 2005도870 판결은 ‘서비스제공자에게 제3자로 하여금 사용하도록 한 사정을 고지하였다면 서비스제공자도 동의하였으리라고 추인되는 경우’라고 하여 서비스제공자의 가상적 의사를 판단기준의 하나로 들고 있기 때문이다. 위 판결은 대상판결로 명시적으로 폐기된 바 없고, 대상판결 이후에도 여전히 정보통신망침입죄에 관한 대표적 판결로서 위상을 잃지 않았

22) 그래서 대상판결의 이 부분 판시에 접근권한을 ‘종합적으로’ 고려하여 ‘신중하게’ 판단하여야 한다고 설시하였다고 본다.

다.²³⁾

그러나 대상판결이 객관적 사정을 강조한 것은 온당하다고 생각한다. 서비스제공자의 주관적 의사는 매우 불확정적이고, 이용자가 언제나 명시적·확정적으로 파악할 수 있는 것이 아니라는 점에서도 그렇다. 정보통신망의 사용이 불특정·다수의 사람들에게 일반적으로 열려있고, 인터넷이나 정보통신망의 발전이 그와 같은 개방성을 전제로 하여 왔다는 점에서도 그렇다.

2) 이용약관

대법원은 정보통신망법 제49조²⁴⁾와 관련하여 “그 전제가 되는 정보의 귀속은 정보통신서비스 제공자에 의하여 그 접근권한이 부여되거나 허용된 자가 누구인지에 따라 정해져야 할 것이고, 이는 정보통신서비스 제공자가 정한 인터넷 온라인 게임 이용약관상 계정과 비밀번호 등의 관리책임 및 그 양도나 변경의 가부, 그에 필요한 절차와 방법 및 그 준수 여부, 이용약관에 따른 의무를 이행하지 않았을 경우 행해질 수 있는 조치내용, 캐릭터 및 아이템 등 게임 정보에 관한 이용약관상 소유관계 등 여러 사정을 종합적으로 고려하여야 한다”²⁵⁾고 하여 종전부터 이용약관을 접근권한의 기준으로 하고 있었다.

그러나 이용약관을 접근권한의 판단기준으로 하는 데에는 유의하여야 할 점이 있다. 약관은 계약 당사자 중 일방의 의사에 의하여 작성되는 것이고, 그 상대방의 의사는 적극적으로 반영되지 않고 반영할 수도 없다. 또한 정보통신망서비스 제공자의 이용약관은 그것을 이용하려는 사람들에게는 일방적으로 의무를 승인 내지 동의할 것이 강제된다. 이용자가 이용약관을 승인하지 않으면 서비스 자체를 이용할 수가 없기 때문이다. 나아가 이용자는 이용약관을 승인하는 것만으로 대가 없이 무

23) 그러나 위 판결에 대한 필자의 비판적 태도는 각주 20) 참조

24) 제49조(비밀 등의 보호) 누구든지 정보통신망에 의하여 처리·보관 또는 전송되는 타인의 정보를 훼손하거나 타인의 비밀을 침해·도용 또는 누설하여서는 아니 된다.

25) 대법원 2010. 7. 22. 선고 2010도63 판결

상으로 서비스를 제공받는 것이 아니다. 이용자는 서비스 이용과 관련하여 자신의 데이터를 서비스제공자에게 제공하고 있다. 이처럼 보이지 않는 교환관계가 데이터 이코노미의 기본 운영 원리가 되었다. 마지막으로 인터넷 또는 정보통신망은 개방성을 그 본질로 하고 있고 현대 사회에서는 그에 대한 접근이 사람의 기본권 또는 인권의 하나로까지 거론되고 있다. 그런 상황에서 서비스제공자인 기업의 일방적인 의사만으로 접근권한이 결정될 수는 없음이 당연하다. 가령 이용약관에서 “오페라를 좋아하는 왼손잡이는 접근을 금지한다.”라고 정하였다고 하여 실제로 그러한 사람들의 접근이 금지된다고 볼 수 없는 것과 같다.²⁶⁾

또 이용약관은 그 문언에 기초하여 매우 엄격하게 해석되어야 하고, 해석상 논란의 여지가 있을 경우 이용자의 접근권한을 확장하는 방향으로 해석되어야 한다. 이용약관에서 접근권한의 제한을 정하더라도, 그것은 접근 그 자체를 금지해야 하고, 그 접근하려는 목적을 규제할 수는 없다. 다시 말해 이용자의 접근 목적에 따라 접근을 제한하는 약관은 그 자체로 매우 불명확하고 이용자의 이용권을 부당하게 제한할 염려가 있으므로 접근권한의 판단기준으로 사용될 수 없다고 본다.

3) 보호조치

정보통신망법 제48조는 침입을 방지하기 위한 기술적 보호조치를 침해하거나 훼손할 것을 요구하지 않는다. 대법원도 같은 맥락에서 피고인들이 서비스제공자인 SK 브로드밴드로부터 정당한 접근권한을 부여받지 않고, 장애처리용 전화기를 이용하여 SK 브로드밴드 주배전반의 통신포트에 연결한 후 피고인들의 휴대폰에 전화 연결을 하는 부정한 방법으로 SK 브로드밴드의 정보통신망에 접속한 행위는 정보통신망법 제48조 제1항에서 규정하는 정당한 접근권한 없이 정보통신망에 침입하는 행위에 해당한다고 할 것이고, 시스템의 정상적인 운영을 저해함이 없이 시스템

26) Orin S. Kerr, Norms of Computer Trespass, Columbia Law Review V. 116, No 4.(<https://columbialawreview.org/content/norms-of-computer-trespass/> 2022. 8. 6. 방문)

에 접속하는 경우에는 ‘침입’에 해당하지 않는다는 피고인들의 이 부분 법리오해 주장은 이유 없다고 한 원심 판단²⁷⁾에 법리오해의 위법이 없다고 하였다.²⁸⁾

보호조치 여부도 종전부터 접근권한의 판단기준으로 검토되었다. ‘5급(행정) 공무원 공개경쟁 채용 제2차 시험’ 합격자 발표와 관련하여 담당자는 발표 전날 합격자 명단을 사이버국가고시센터 홈페이지에 올리면서, 공식 발표 시각 이전에 위 홈페이지에 게시되지 않도록 설정기간을 미리 예약하여 게시하였고, 웹브라우저 주소창에 공고 게시글의 주소를 입력해도 웹브라우저에 빈 화면이 나타나도록 소스 코드를 구성하였는데, 피고인은 위 시험에 응시한 여자친구의 합격 여부를 공식 발표 전에 미리 알기 위하여 위 사이트 게시판에 이미 게시된 ‘2016년도 외교관 후보자 선발시험 최종합격자 명단’의 첨부파일 주소를 확인하고 이를 복사하여 인터넷 주소창에 붙여넣기 한 다음 그 주소의 파일 숫자 끝 번호를 계속 변경·입력하다가 위 시험의 합격자 명단 파일의 파일 주소(URL)를 입력하게 되었고 이에 의하여 합격자 명단 파일을 다운로드한 사안에서, 제2심 법원이 “당시 사이버국가고시센터 사이트는 누구나 접속할 수 있었고, 위 합격자 명단 파일에 아무런 보호조치가 되어 있지 않았으며, 피고인은 누구나 자유롭게 이용할 수 있는 이 사건 사이트에 접속하여 우연히 알아낸 위 합격자 명단 파일의 주소를 웹브라우저 주소창에 입력하여 파일을 다운로드하였을 뿐이고, 웹브라우저 주소창에 직접 어떤 주소를 입력하는 행위가 금지되어 있다고 볼 수 없다”는 이유로 정보통신망법위반을 무죄로 판단한 것을 대법원도 그대로 인용하였다.²⁹⁾

27) 서울서부지방법원 2012. 4. 5. 선고 2012노10 판결

28) 대법원 2013. 10. 17. 선고 2012도4387 판결; KT 직원들이 전화·인터넷 등 아파트 전 세대의 통신회선이 집중되어 있는 아파트 통합통신장비실에 들어가 장애처리용 전화기를 SK 브로드밴드의 통신포트에 연결하여 개인용 휴대전화 등으로 전화를 걸어 고객의 발신번호가 표시되도록 하는 방법으로 SK 브로드밴드 가입 고객의 전화번호를 수집한 사안이다.

29) 본문의 내용은 원심(서울서부지방법원 2017. 7. 20. 선고 2017노345 판결)의 내용이다. 위 판결에 대하여 대법원은 구체적인 내용의 실시 없이 상고기각(대법원 2017. 10. 12. 선고 2017도12758 판결)

위 판결은 여러모로 흥미로운 점이 많은데, 대상판결과의 관계에서 보자면, 대상판결에서 데이터를 수집한 방법이 크롤러라는 컴퓨터 프로그램에 의한 것이라면, 위 판결은 데이터를 수집하

보호조치 역시 그것을 할 것인지 여부, 한다면 보호되는 정보와 그렇지 않은 정보의 구별, 기술적으로 서버의 계층구조에서 어느 단계에서 보호조치를 취할 것인지 등 보호조치와 관련된 모든 사항은 서비스제공자가 결정한다. 따라서 보호조치와 관련된 사정만으로 접근권한을 일의적으로 결정할 수는 없다. 위 대법원 2017. 10. 12. 선고 2017도12758 판결의 원심은 무죄 이유 중의 하나로 보호조치가 없는 사정을 들었지만, 대법원 2013. 10. 17. 선고 2012도4387 판결에서는 보호조치가 없었음에도 외부인이 접근할 수 없는 아파트 통합통신장비실에 들어가 접속한 사안에서 유죄를 인정하기도 하였다.

4) 대상판결의 태도

대상판결은 피고인들의 이 사건 크롤링 프로그램에 의한 이 사건 API 서버로의 접근이 제한되었다고 보기 어렵다고 하였다. 그 근거는 다음과 같다. ①은 보호조치에 관한 것이고, ②, ③은 이용약관에 관한 것이다. ① 이 사건 API 서버의 URL이나 명령구문은 피해자 회사가 적극적으로 공개하지는 않았지만 누구라도 간단한 기술조작이나 통상 사용되는 소위 ‘패킷캡처 프로그램’ 등을 통해 쉽게 알아낼 수 있는 정보이다. 일반 이용자들은 이 사건 앱을 통해 API 서버에 회원 가입 후 또는 회원 가입 없이 자유롭게 접근할 수 있었고, 이 사건 앱이나 API 서버로의 접근을 막는 별도의 보호조치는 없었다.³⁰⁾ ② 피해자 회사의 이 사건 앱 서비스 이용약관

기 위하여 이미 알려진 URL의 일부 숫자를 일일이 바뀌가며 입력한 것이다. 전자가 자동화 방법이라면 후자는 수기 또는 사람이 일일이 웹브라우저에 가능한 URL을 입력한 것이다(굳이 명명하자면 ‘손(hand) 크롤링’이라고 할 수 있을 것이다. 내가 이 글을 쓰게 된 동기는 다음과 같은 호기심 때문이었다. 사람이 일일이 손으로 하는 일이 무죄라면 그것을 컴퓨터 프로그램과 같은 자동화 도구를 사용하는 것도 무죄가 되어야 하는 아닌가? 만약 손으로 하는 일이 무죄이지만 그것을 자동화 도구를 사용하는 것은 유죄라면, 양자의 질적 차이는 무엇인가? 양자의 가치적·법적 평가가 달라지는 이유는 무엇인가? 자동화에도 여러 단계가 있을 터인데, 과연 자동화의 어느 지점에서 유죄와 무죄가 갈리는 것인가? 이는 크롤링뿐만 아니라 뒤에서 잠시 언급할 매크로의 형사책임과도 관련 있다. 대상판결은 위와 같은 호기심에 하나의 대답을 제시하였는데 그것은 저작권법위반에 관련된 두 번째 판시사항이다.

30) 제1심 법원도 인정하였듯이 피해자 회사는 이 사건 API 서버에 SSL(secure socket layer)을 사용하지 않았다.

에서 ‘이용자는 회사를 이용함으로써 얻은 정보를 회사의 사전 승낙 없이 복제, 송신, 출판, 배포, 방송 등 기타 방법에 의하여 영리 목적으로 이용하거나 제3자에게 이용하게 하여서는 안 된다’고 정하고 있으나, 이는 이 사건 앱 또는 API 서버로부터 취득한 정보의 이용을 제한하는 내용일 뿐, 이에 대한 접근을 제한하는 내용으로 볼 수 없다. ③ 위 이용약관에서 회원에 대하여 ‘자동접속프로그램 등을 사용하여 회사의 서버에 부하를 일으켜 회사의 정상적인 서비스를 방해하는 행위’를 금지하고 있기는 하지만, 위 약관 규정을 회원가입을 하지 않은 이용자들에게 적용할 수 있는 근거를 찾기 어렵고, 규정의 내용 또한 접근권한 자체를 제한하는 것으로 볼 수 없어 위와 같은 약관상의 규정만으로 API 서버에 대한 접근권한이 객관적으로 제한되었다고 보기 어렵다.

대사판결에 대해서 정보통신망침입은 인정하지 않으면서도 피해자 회사가 대량 호출 신호를 감지하고 피고인들이 사용하는 서버의 IP 주소를 차단한 이후에도 피고인들이 서버의 전원을 껐다가 다시 켜는 방식으로 IP 주소를 변경하여 이 사건 API 서버에 접속한 것은 정보통신망의 안정성을 해하는 것이므로 정보통신망침입이 인정된다는 견해도 있다.³¹⁾ 그러나 그와 같은 견해는 받아들일 수 없다. 위와 같은 피해자 회사의 IP 주소 차단으로 피해자 회사가 의도한 정보통신망의 안정성은 이미 달성된 것이다. 피고인들은 그와 같이 차단된 IP 주소를 통한 접속을 한 것이 아니고, 단순히 자신들이 이용하는 서버의 전원을 다시 켜는 뿐이다. 통상 IP 주소는 동적으로 할당(dynamic allocation)되고 정보통신망은 이것을 전제하고 설계되므로, 동적으로 할당되어 변경된 IP 주소를 통하여 접속한 행위는 정보통신망

SSL은 웹브라우저와 웹서버 간에 주고받는 데이터의 안전성과 보안성을 위하여 사용하는 통신 프로토콜이다. 구체적으로 웹브라우저가 웹서버에 SSL을 적용한 웹 페이지를 요청하면 웹서버는 웹브라우저에 공개키와 인증서를 발행하고, 웹브라우저는 증명서가 신뢰할 수 있는지 확인한 후 대칭키 방식의 암호통신을 위한 세션키를 생성한다. 웹브라우저에서 데이터는 세션키로, 세션키는 공개키로 암호화하여 암호화된 세션키와 데이터를 웹서버에 전송하면, 웹서버는 암호화된 세션키를 비밀키로 복호화하고 암호화된 데이터는 세션키로 복호화하는 방식으로 통신한다(손진곤, 길준민, 『정보통신망』, 한국방송통신대학교출판문화원, 2021년, 303쪽).

31) 최상진, “경쟁사의 무단 크롤링에 대한 법적 대응방안에 관한 연구”, Law & Technology 제17권 제1호, 2021년, 34쪽

의 안정성과는 무관하다. 또한 피해자 회사의 위 IP 주소 차단이 피고인들의 접근 권한을 제한했다고 볼 수도 없다. 앞서 본 바와 같이 접근권한의 제한은 객관적이고 명백한 사정에 따라 판단되어야 한다. 피해자 회사의 입장에서 차단된 IP 주소를 통한 접근을 금지하였다고 볼 수는 있을지언정 다른 IP 주소를 통한 접근까지 배제하였다고 볼 수 없다.³²⁾

2. 저작권법위반의 점

(1) 저작권법상 데이터베이스와 데이터베이스제작자의 권리

저작권법상 ‘데이터베이스’라 함은 소재를 체계적으로 배열 또는 구성한 편집물로서 개별적으로 그 소재에 접근하거나 그 소재를 검색할 수 있도록 한 것을 말하는데(저작권법 제2조 제19호), 데이터베이스제작자는 그의 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송(이하 ‘복제 등’이라고 한다)할 권리를 가지고(같은 법 제93조 제1항), 이에 따라 보호되는 데이터베이스제작자의 권리를 복제·배포·방송 또는 전송의 방법으로 침해한 자는 형사상 처벌의 대상이 된다(같은 법 제136조 제2항 제3호). 한편, 데이터베이스의 개별 소재는 당해 데이터베이스의 상당한 부분으로 간주되지 않고(같은 법 제93조 제2항 본문), 데이터베이스제작자에 대한 보호는 데이터베이스의 구성부분이 되는 소재 그 자체에는 미치지 않으나(같은 법 제93조 제4항), 데이터베이스의 개별 소재 또는 그 상당한 부분에 이르지 못하는 부분의 복제라 하더라도 반복적이거나 특정한 목적을 위하여 체계적으로 함으로써 해당 데이터베이스의 통상적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에는 해당 데이터베이스의 상당한 부분의 복제로 보게 되어(같은 법 제93조 제2항 단서), 같은 법에 따라 보호되는 데이터베이스제작자의 권리를 침해하는 행위로 평가될 수 있다.

32) 이 부분은 대법판결에서 언급되지 않았다.

(2) 데이터베이스제작자 권리 침해의 판단기준

데이터베이스제작자는 자신이 제작한 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송할 권리를 가지므로, 저작권법 제136조 제2항 제3호에 따라 데이터베이스제작자의 권리 침해 여부를 판단하는 데에는 그 기준은 전부 또는 상당한 부분을 복제하였는지 여부이다. ‘상당한 부분’이라는 매우 추상적인 불확정 개념의 해석이 문제된다.

1) 데이터베이스제작자의 권리 침해 여부에 관한 민사사례

① 서울고등법원 2016. 12. 15. 선고 2015나2074198 판결³³⁾

원고는 이용자들이 특정한 주제에 관한 게시물을 자유롭게 작성하여 게시하거나 이미 게시된 내용을 자유롭게 수정하는 방식으로 웹사이트(리그베다위키, rigvedawiki)를 운영하였는데 피고가 원고 웹사이트에 집적된 자료 전부를 미러링(mirroring) 방식으로 복제하여 피고 운영 웹사이트(‘엔하위키 미러’)에 게재한 사안이다. (i) 원고가 개인적으로 운영하던 상식 사전 사이트 데이터를 12,000~13,000개 항목으로 정리하고, 그 중 100여개를 선별하여 구성의 체계성, 개별 소재의 접근성, 검색 기능 등을 테스트한 후 나머지 데이터 10,000여 개를 엔하위키 게시판에 모두 업데이트하였고, 그 후에도 이용자들의 요구에 따라 목차 구조와 페이지 작성 양식 등을 만들거나 ‘최근 변경내역’ 등을 도입하여 개별 자료에의 접근성을 높인 점, (ii) 검색엔진을 변경하면서 사이트 환경에 맞추어 검색 기능 등을 추가 개발하는 등 접근 및 검색 가능성을 높여 플랫폼을 구축한 점, (iii) 원고 명의의 서버 4대를 운영하면서 약 16,000명의 가입자와 25만 개의 위키 문서가 있는 원고 사이트를 유지·관리한 점 등을 종합하여 원고의 데이터베이스제작자로서의 지위를 인정하였

33) 대법원 2017. 4. 13. 선고 2017다204315 판결(심리불속행)로 확정되었다. 일명 ‘리그베다위키’ 사건이다.

고, 피고가 원고 사이트를 미리링하는 방법으로 원고 사이트의 미리 사이트를 개설·운영하였으므로 원고의 복제권·전송권 침해가 인정된다고 하였다.

② 서울고등법원 2017. 4. 6. 선고 2016나2019365 판결³⁴⁾

피고가 경쟁업체인 원고 운영 웹사이트(‘잡코리아’)에 게시된 ‘채용정보 전부’를 원고의 동의 없이 수집 프로그램을 통해 크롤링하여 피고의 웹사이트(‘사람인’) 서버에 저장한 후 구인업체의 동의를 받는 절차를 거쳐 피고 웹사이트에 게재한 사안이다. 위 판결은 “원고 웹사이트는 여러 구인업체의 채용정보를 체계적으로 배열하여 수록함으로써 이용자가 원고 웹사이트로부터 각종 채용정보를 각 분류별로 자신이 원하는 기준에 따라 모아서 열람하거나 검색할 수 있도록 한 데이터베이스에 해당하며, 원고는 원고 웹사이트를 제작 및 소재의 갱신·검증 또는 보충을 위하여 인적 또는 물적으로 상당한 투자를 한 자로서 원고 사이트에 대한 데이터베이스제작자에 해당하며, 피고가 별도의 마케팅 비용 등의 지출 없이 피고의 영업에 이용할 목적으로 반복적, 체계적으로 원고 데이터베이스의 채용정보 부분을 복제함으로써 데이터베이스제작자인 원고의 이익을 부당하게 해쳤다”고 하였다.

③ 대 상관결의 사안과 비교

위 두 사건에서 법원은 정보의 분류, 검색 프로그램 개발, 서버의 관리 등을 투자자의 상당성 판단에서 중요한 고려요소로 보고 있으며, 피고가 원고의 정보 모두를 복제하였기 때문에 특별히 ‘상당한 복제’인지 여부를 고려할 필요는 없었던 것으로 보인다. 대 상관결의 사안은 다음과 같은 점에서 위 두 사건과 사실관계를 달리한다.³⁵⁾ (i) 웹사이트 자체가 아닌 프런드엔드(front-end)의 API 서버에의 접근이 문제가 되었는데 그 API 서버 접근에 대한 특별한 보호조치가 없었다. (ii) 크롤링을 통해 수집된 정보가 데이터베이스를 구성하는 50여개 항목 중 3 내지 8개로 양적

34) 대법원 2017. 8. 24. 선고 2017다224395 판결(심리불속행)로 확정되었다. 일명 ‘잡코리아’ 사건이다.

35) 물론 가장 큰 차이는 위 두 사건은 손해배상을 구하는 민사절차에서 이루어진 것이고, 대 상관결은 직접 크롤링에 가담한 피고인 회사의 임직원들에 대한 형사책임을 묻는 형사절차에서 이루어졌다는 점이다.

상당성을 인정하기 어려운 상황이었다. (iii) 그 정보들이 대부분 이용자들에게 공개된 것이거나 쉽게 수집할 수 있는 것이어서 질적 상당성 역시 인정하기 어려운 측면이 있었다. (iv) 약관상 크롤링 제한의 내용이 모호했고 피고인들이 약관의 적용을 받는 회원이 아니었다.

**2) 데이터베이스제작자의 권리 침해 여부에 관한 형사사례(서울남부지방법원
2021. 9. 8. 선고 2021고단588 판결, 확정)**

직업소개사업을 하는 피고인들이 수집한 정보를 판매하기 위한 목적으로, 관련 웹사이트에 접속하여 ‘인재검색 서칭 서비스’ 상품을 결제한 뒤 ATS(Advanced Target Search) 프로그램을 이용하여 ‘잡코리아’의 이력서 페이지에 약 69,000회 반복적으로 접속하여 구직자들의 사진, 이름, 연락처, 주소 일부를 제외한 나머지 정보(성별, 나이, 학력, 교육, 경력, 자격증 등)을 수집, 복제하였다. 잡코리아는 구직자정보의 경우 직무, 경력, 근무지역, 학력, 나이, 거주지역 등으로 분류하여 유형별로 ‘인재검색’이 가능하지만, 구직자정보를 모두 블라인드 처리하여 사이트 이용자가 ‘인재검색 서칭 서비스’ 상품에 가입하여야 구직자의 사진, 이름, 연락처, 주소 일부를 제외한 나머지 정보(성별, 나이, 학력, 교육, 경력, 자격증 등)를 제공한 사안에서 법원은 무단데이터복제와 정보통신망침입을 모두 유죄로 인정한 바 있다.

(3) 대상판결의 태도

대상판결에 따르면 데이터베이스제작자의 권리 침해가 인정되는 경우는 다음과 같다. ① 데이터베이스의 전부 또는 상당한 부분의 복제가 있는 경우인데, 여기서 ‘전부 복제’는 판단의 어려움이 없다. 문제는 ‘상당한 부분’의 복제는 어떠한 정도에 이르러야 하는지 인데, 전체 데이터베이스와 복제된 부분의 규모를 비교하는 양적인 측면과 데이터베이스제작자가 그 복제된 부분의 제작 또는 그 소재의 갱신·검증 또는 보충에 상당한 투자를 하였는지를 확인하는 질적인 측면을 모두 고려하

여야 한다. ③ 상당한 부분의 복제에 이르지 못하였어도 반복적이거나 특정한 목적을 위하여 체계적으로 복제하여 데이터베이스의 통상적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에는 ‘상당한 부분의 복제’에 이르렀다고 간주되나, 이때에도 데이터베이스의 개별 소재 또는 상당하지 않은 부분에 대한 반복적이고 체계적인 복제로 결국 상당한 부분의 복제 등을 한 것과 같은 결과를 발생하게 한 경우에 한하여 인정하여야 한다. 아래에서 보는 바와 같이 대상판결은 데이터베이스제작자의 권리 침해를 인정하지 않은 원심을 그대로 인용하였으므로 위 ③과 같이 상당한 부분의 복제 이르지 않은 경우에도 결국 상당한 부분의 복제 등을 한 것과 같은 결과를 발생한 경우가 구체적으로 어떤 경우가 될 것인가는 후속 판결들을 기다려보는 수밖에 없다.

3. 컴퓨터등장애업무방해의 점

(1) 형법 제314조 제2항의 컴퓨터등장애업무방해죄

대상판결에서 문제되는 것은 형법 제314조 제2항 후단이다. ‘허위의 정보’는 객관적으로 진실에 반하는 내용의 정보를 의미하고, ‘부정한 명령’은 객관적으로 정당하지 않은 명령이나 사무처리 과정에서 주어서는 안 되는 명령을 의미하는데, 구체적으로 관리자가 정보처리장치를 운영하는 본래의 목적과 상이한 명령을 입력하는 것을 의미한다.³⁶⁾ 또한 ‘정보처리에 장애’를 발생하게 하여 ‘업무를 방해’하여야 하는데, 정보처리의 장애란 가해행위 결과 정보처리장치가 그 사용목적에 부합하는 기능을 하지 못하거나 사용목적과 다른 기능을 하는 것을 말한다.³⁷⁾

36) 대법원 2006. 3. 10. 선고 2005도382 판결, 대법원 2009. 4. 9. 선고 2008도11978 판결 등 참조

37) 대법원 2009. 4. 9. 선고 2008도11978 판결

(2) 대상판결의 태도³⁸⁾

공소사실은 피고인들이 이 사건 API 서버에 접속하여 위도, 경도, 반경에 관하여 피고인 회사의 위치를 기준으로 반경 1,000km로 하여 그 안에 있는 모든 숙박업소 정보를 요청하는 것이 부정한 명령을 입력한 것이라고 한다. 결국 2가지가 문제되는데, 첫째는 앱을 통하지 않고 PC를 통하여 접속하였다는 것과 반경을 피해자 회사가 설정해 놓은 범위를 넘어 1,000km로 하였다는 것에 대한 형법적 평가이다.

전술한 바와 같이 이 사건 API 서버에 이 사건 크롤링 프로그램을 통하여 명령구문을 입력한 것이 명령입력 권한이 없는 행위라고 할 수 없다고 본다. 왜냐하면 피해자 회사는 이 사건 API 서버에 명령 입력의 범위를 제한하지 아니하였기 때문이다. 이 사건 앱에서만 검색범위의 제한이 일부 설정되어 있었다고 하더라도 이 사건 API 서버는 기본적으로 주어진 명령구문에 대응하는 숙박업소 정보를 반환하는 것에 있으므로 반경을 넓혀서 검색한 것만으로 그것이 부정하다고 할 수 없다. 물론 피해자 회사가 검색범위를 3km로 제한해 놓은 의도를 대상판결 및 관련 하급심 판결의 내용만으로는 정확히 알 수 없다. 그와 같은 제한이 이 사건 앱의 접속시간과 접속 횟수를 늘리기 위한 피해자 회사의 경영적 판단일 수도 있고, 단순히 이용자에게 그의 위치를 기준으로 적절한 정보를 제공해 주기 위한 이용자 편의를 위한 것일 수도 있고, 아니면 그 밖의 다른 이유가 개재되어 있었을 수도 있다. 그러나 그와 같은 피해자 회사의 일방적 의사에 따라 설정된 검색범위의 제한을 우회하는 이용을 부정한 것이라고 한다면, 정보통신망 관리자의 주관적 의사에 의하여 형사처벌이 불합리하게 확대되는 결과가 될 것이다.

따라서 ‘부정한 명령’의 판단도 정보통신망 침입의 접근권한과 같이 객관적 사정을 중심으로 판단해야 한다. 본죄의 성립을 인정한 주요 선례들³⁹⁾을 보아도, 권한

38) 이 부분 대해서 대상판결은 기존의 법리를 기초로 원심의 판단이 위법하지 않다고 하였을 뿐, 새로운 법리를 실시하지 않았다.

39) 대법원 2009. 4. 9. 선고 2008도11978 판결, 대법원 2013. 3. 28. 선고 2010도14607 판결, 대법원 2020. 2. 13. 선고 2019도12194 판결 등

없는 자가 아이디와 비밀번호를 무단으로 변경하거나, 허위의 클릭정보를 전송하거나, 권한 없는 자가 함부로 컴퓨터에 비밀번호를 설정한 행위 등으로 명백하게 허위의 정보를 전송하거나 정보처리장치의 관리 권한 없으면서도 아이디와 비밀번호를 무단으로 입력하거나 비밀번호를 설정한 사안들이어서, 그와 같은 것은 객관적 사정에 비추어 보아도 부정한 것이라고 볼 수 있다.

IV. 자동화 도구를 사용한 정보통신망 이용

1. 매크로(macro)의 사용에 대한 형사책임

대상관결과 관련된 법적 논의의 근본문제는 정보통신망의 이용자가 자동화 알고리즘 또는 자동화 프로그램과 같은 ‘자동화 도구(automatic tool)’⁴⁰⁾를 통하여 접근하는 것에 대한 법적 평가를 어떻게 하는가이다. 다시 말해 사람이 손으로 직접 키보드나 마우스를 조작하여 이용하면 별다른 문제가 되지 않는 정보통신망의 이용이 자동화 도구를 사용하여 짧은 시간에 수회에 걸쳐 반복적으로 이루어진 경우에는 무엇이 어떻게 달라지는가? 양자의 질적 차이는 무엇인가? 양자의 법적 평가는 달라져야 한다면 그 이유는 무엇이고, 그렇게 평가가 달라지는 구체적인 지점은 어디인가? 이는 특히 인터넷 4.0에서는 무수히 많은 자동화 알고리즘이 개발되고 그것의 사용이 용이해지고 보편화되고 있어서 이와 같은 문제가 계속 발생할 것이다. 그에 대한 선부른 해답을 제시하는 것은 이 글의 의도가 아니다. 그 해결의 실마리, 도움이 될 만한 기존의 논의들 또는 방법론에 대한 아이디어를 제시하고자 할 뿐이다.

대상관결에서 다른 크롤링 외의 다른 자동화 도구인 매크로의 사용에 대한 형사 책임을 살펴본다. 매크로는 컴퓨터 환경에서 순차적으로 수행되어야 할 여러 개의

40) 이는 필자의 용어법이다.

명령어들이 반복되어 사용될 때, 이 명령어들을 묶어서 하나의 키 입력 동작으로 만드는 것을 말한다. 이를 상술하자면 일련의 반복된 작업을 수행하기 위해 조작자가 매번 개개의 명령어를 입력하지 않고, 순차적으로 수행될 여러 개의 명령어들을 하나의 새로운 명령어로 정의하여 이를 키보드의 특정 키 또는 마우스 클릭으로 설정해두는 것을 매크로 명령어라고 하고, 이렇게 한 번의 키 입력으로 미리 기록해 놓은 여러 단계의 명령군을 자동으로 실행하도록 하는 프로그램을 매크로 프로그램이라고 한다.⁴¹⁾

재판례를 보면, 특정 회사가 제공하는 게임사이트에서 정상적인 포커게임을 하고 있는 것처럼 가장하면서 통상적인 업무처리 과정에서 적발해 내기 어려운 사설 프로그램⁴²⁾을 이용하여 약관상 양도가 금지된 포커머니를 약속된 상대방에게 이전해 준 사안에서, 대법원은 위 프로그램이 정보통신망법 제48조 제2항에서 정한 ‘악성 프로그램’에 해당하지 않고, 위 프로그램에 입력된 명령이 형법 제314조 제2항의 ‘부정한 명령의 입력’에 해당하지 않으며, 위 프로그램의 사용으로 인하여 엔에이치엔 주식회사가 운영하는 게임서버에 장애가 발생하였다고 볼 수 없다는 취지로 판단한 다음, 피고인들에 대한 공소사실 중 정보통신망법위반(정보통신망침해등)의 점 및 컴퓨터등장애업무방해의 점을 모두 무죄로 본 원심판결이 정당하다고 하였으나, 업무방해죄의 성립은 인정하였다.⁴³⁾ 또 피고인들이 자동접속프로그램을 이용하여 특정 홈페이지에 접속한 것 이외에 로그인 절차를 요구하는 개인정보 등에 권한 없이 접근한 것이 아니므로 위 홈페이지를 침입하였다고 볼 수 없다는 원심판결을 인용하였다.⁴⁴⁾

41) 최란, “매크로 프로그램을 사용하여 티켓 등 이용권을 예매하는 행위의 가벌성 검토”, 과학기술법연구 제24집 제3호, 한남대학교 과학기술법연구원, 2019년, 245~246쪽

42) 포커머니 거래용 자동 프로그램으로 속칭 ‘수혈프로그램’으로 불렸다. 피고인들은 포커머니를 판매하기 위하여 매수자들에게 고의로 게임에서 패하도록 하는 위 프로그램을 사용하였다.

43) 대법원 2009. 10. 15. 선고 2007도9334 판결

44) 대법원 2013. 3. 14. 선고 2010도410 판결

특히 자동 회원가입, 자동 방문 및 이웃신청 등의 기능을 이용하여 네이버 카페나 블로그 등에 자동적으로 게시 글과 댓글을 등록하고 쪽지와 초대장을 발송하는 등의 작업을 반복 수행하는 프로그램에 대해서, 대법원은 다음과 같은 이유로 그것이 정보통신망법 제48조 제2항의 악성프로그램에 해당하지 않는다고 하였다.⁴⁵⁾ ① 위 프로그램은 일반 사용자가 통상적으로 작업하는 것보다 빠른 속도로 작업하기 위하여 자동적으로 댓글의 등록이나 쪽지의 발송 등의 작업을 반복 수행할 뿐이고, 기본적으로 일반 사용자가 직접 작업하는 것과 동일한 경로와 방법으로 위와 같은 작업을 수행한다. ② 위 프로그램 중 일부는 프록시 서버를 이용하여 네이버 등에 간접적으로 접속할 수 있도록 함으로써 네이버 등의 정보통신시스템 등이 IP를 차단하는 것을 회피할 수 있도록 설계되어 있으나 이는 네이버 등의 정보통신시스템 등을 훼손·멸실·변경·위조하는 등 그 기능을 물리적으로 수행하지 못하게 하는 방법으로 IP 차단을 방해하는 것이 아닌, 위 정보통신시스템 등이 예정한 대로 작동하는 범위 내에서 IP 차단 사유에 해당하지 않고 통과할 수 있도록 도와주는 것에 불과하다.

사례가 많지 않아 일반화하기는 어렵지만, 개략적으로나마 법원의 입장은 매크로 프로그램 자체는 악성프로그램으로 볼 수 없고, 특별한 사정이 없는 이상 매크로 프로그램을 사용하였다고 하여 그것이 정보통신망침입이 될 수도 없지만, 경우에 따라 형법상 업무방해죄가 성립할 수는 있다는 정도로 정리할 수 있다.

2. 자동화 도구의 사용에 대한 형법해석

우선 기술중립적인 해석이 필요하다. 자동화 도구는 말 그대로 도구일 뿐이고, 법률의 규정과는 무관하게 다양한 모습으로 새롭게 개발되고 되고 있다. 도구 또는 기술 그 자체는 법률적 책임과 무관하다. 문제가 되는 것은 새로운 도구의 사용이

45) 대법원 2019. 12. 12. 선고 2017도16520 판결

기존의 법률 체계 또는 그것이 보호하려는 이익 또는 법익에 어떠한 영향을 미치는지 가 될 것이다. 그와 같은 문제점은 자연스럽게 처벌규정의 보호법익과 연결된다. 오늘날 우리의 보호법익론은 헌법상 과잉금지원칙과 그것을 구체적으로 드러내는 헌법재판의 활성화로 인하여 실질적 범죄를 규정하는 기능을 상실하였다고 봐도 과언이 아니다.⁴⁶⁾ 그러나 보호법익을 세밀하게 밝히는 것은 해당 구성요건의 해석과 그를 통한 형사처벌의 범위를 명확히 하는데 도움을 줄 수 있을 것이다. 또한 그렇게 세밀화된 법익론은 새로운 기술 또는 새로운 행위태양에 대해서도 탄력적으로 대응할 수 있을 것이다. 새로운 사태에 대해서 반드시 새로운 입법과 법리가 필요한 것은 아니라고 본다. 유추할 수 있는 기존의 법리를 차용하는 것은 불필요한 입법과 과잉 법리를 저지할 것이다. 그와 같은 유추의 근거이자 제한원리가 보호법익이 될 수 있다(예컨대 정보통신망침해는 종래 주거침입에 대한 법리를 차용할 수 있다고 본다. 그러나 양죄의 보호법익의 성격이 상이하므로, 전자는 사회적 법익의 성격이 짙고, 후자는 개인적 법익에 관한 것이므로 그와 같은 차이는 유추의 제한원리가 될 것이다). 마지막으로 형법의 겸양성이다. 처벌의 필요성과 현행법상 처벌이 가능한지는 전혀 다른 별개의 문제이다. 특정 행위에 대하여 제재가 필요하다고 하더라도 굳이 형사법이 그것을 온전히 담당할 필요는 없다. 형사법의 법리가 끝나는 곳에서 민사법의 논의가 시작될 수도 있을 것이다.⁴⁷⁾

V. 맺음말

요컨대 대상판결은 크롤링을 통한 데이터 수집의 형사책임에 대하여 최초로 판

46) 이에 대해서는 이진국, “형사입법에서 법익개념의 체계비판적 기능”, 동아법학, 동아대학교출판부, 2005년, 103~104쪽

47) 실제로 대상판결의 사안에 대한 민사판결에서는 피고인 회사의 손해배상책임이 인정되었다(서울중앙지방법원 2018가합508729 판결. 위 판결은 피고인 회사에 대해서 부정경쟁방지법 제2조 제1호 (가)목에 의한 손해배상책임을 인정하였다. 이에 대한 항소심인 서울고등법원 2021나2034740 판결도 피고인 회사의 손해배상책임에 관한 제1심 판결과 같은 취지로 판단하였다. 위 고등법원 판결은 상소되지 않아 그대로 확정되었다).

시하였다는 데에 그 의의가 있다. 대상판결은 해당 사안에서 크롤링을 통한 데이터 수집이 정보통신망의 침입이 되지 않고, 저작권법상 데이터베이스제작자의 권리를 침해하지도 않았으며, 컴퓨터등장애업무방해도 성립하지 않는다고 하면서 정보통신망법과 저작권법의 새로운 법리를 밝혔다. 다만 기술 자체에 대하여 법률적 가치 판단을 할 수는 없는 것이므로 대상판결의 취지가 크롤링 기술에 대해 형사책임이 없다는 것으로 읽혀서는 안 된다. 또한 향후에 크롤링을 통한 데이터 수집이 문제되는 다른 사안에서는 보호조치 여부 및 그 정도, 이용약관의 내용과 적용 범위, 원본 데이터베이스와 비교한 수집한 정보의 양과 서비스제공자가 그 관리에 쏟은 인적·물적 자원의 정도, 정보통신망에 접속한 방법 등에 따라서 형사책임에 관한 결론이 대상판결과 얼마든지 달라질 수 있을 것이다.

■ 참고문헌

- 강정희, 『빅데이터를 기반으로 하는 배제남용 행위의 위법성 판단기준 연구 - 소비자 선택 기준의 적용을 중심으로』, 박사학위논문, 서강대학교, 2015년
- 권세진 외, “데이터 경제 시대에 있어서 웹 크롤링의 법적 인식에 관한 연구”, 한국산업보안연구, 제11권 제3호, 2021년
- 길준민, 손진곤, 『정보통신망』, 한국방송통신대학교출판문화원, 2021년
- 김경래, “독일 Industrie 4.0의 특징: 노동 4.0을 중심으로”, 한독사회과학논총 제28권 제2호, 2018년
- 김현숙, “크롤링을 이용한 공개데이터 수집·활용의 법적 쟁점에 대한 비판적 검토”, 강원법학 제61권, 2020년
- 서울대 법과경제연구센터, 『데이터이코노미』, 한스미디어, 2017년
- 유대중, “웹검색 서비스와 ISP 책임에 관한 소고”, 장작과 권리, 2007년
- 이진국, “형사입법에서 법익개념의 체계비판적 기능”, 동아법학, 동아대학교출판부, 2005년
- 이창범 외, 『이론&실무 정보통신망법』, 박영사, 2021년
- 이호영, “빅데이터의 경쟁법적 함의에 관한 연구”, 법경제학연구 제15권 제3호, 2018년
- 장윤식 외, “정보통신망법 상 정보통신망침입죄에 대한 비판적 고찰”, 경찰학연구, 제14권 제4호, 2014년
- 정재화, 『클라우드 컴퓨팅』, 한국방송통신대학교출판문화원, 2020년
- 최대우 외, 『데이터과학입문』, 한국방송통신대학교출판문화원, 2020년
- 최란, “매크로 프로그램을 사용하여 티켓 등 이용권을 예약하는 행위의 가벌성 검토”, 과학기술법연구 제24집 제3호, 한남대학교 과학기술법연구원, 2019년
- 최상진, “경쟁사의 무단 크롤링에 대한 법적 대응방안에 관한 연구”, Law & Technology 제17권 제1호, 2021년
- 황태희, “웹 크롤링의 경쟁법적 고찰”, 경쟁법연구, 제43권, 2021년
- Christopher Olston, Marc Najork, 『Web Crawling』, Foundations and Trends® in Information Retrieval, Vol. 4, No. 3, 2010
(http://infolab.stanford.edu/~olston/publications/crawling_survey.pdf 2022. 8. 5. 방문)
- Marc Najork, “Web Crawler Architecture”
(<https://marc.najork.org/pdfs/eds2009a.pdf> 2022. 8. 5. 방문)
- Orin S. Kerr, “Norms of Computer Trespass”, Columbia Law Review V. 116, No 4.
(<https://columbialawreview.org/content/norms-of-computer-trespass/> 2022. 8. 6. 방문)

■ Abstract

Web Crawling Data Collection and Criminal responsibility

Kim Taekyun *

The Case has found, in the case where the defendants, the employees of a company running a accommodation information offering service accessed the competitor's mobile application server and copied the database such as a list of accommodation through their crawling computer program, that the defendants are not guilty. Crawling(web crawling) means an act of collecting data from websites by using data-collecting robots called crawlers. Even though there comes a data economy where data collection is more important than ever and crawling technology is already one of the wide-used technic, there has hardly been studies about the criminal responsibility of data collecting through crawling. The Case showed new legal principles, ① whether the service provider has set any limitation on access right to network should be determined by object things such as protective actions or terms and conditions of use, ② significant copy of database should be determined in the aspects of quantity and quality both. And The Case sentenced not guilty to the charges(invasion of network, piracy of database and obstruction of business). The Case was the first judgement of the Supreme Court on crawling data collection, so it has a great signification.

• **Key Words** Crawling , Data, Data collection, Intrusion on an information and communications network, Database, Reproduction of database, Interference with business

* Seoul Central District Court, Judge