

웹 크롤링을 통한 데이터 수집행위에 대한 형사법적 문제  
검토(대법원 2022. 5. 12. 선고 2021도1533 판결)

부산지방법원 판사

장병준

## I. 머리말

### 1. 연구의 배경과 목적

#### 가. 데이터 경제의 부상

21세기에 들어 데이터는 전 세계 경제에서 핵심 자산으로서의 위치를 공고히 하고 있다. 데이터는 기업의 전략적 의사결정, 마케팅, 고객 관리, 제품 개발, 그리고 운영 효율성 증대 등 다양한 영역에서 중요한 역할을 한다. 디지털 전환이 가속화되면서 데이터는 그 자체로 하나의 경제적 가치로 평가되며, 기업들이 경쟁력을 유지하기 위해 필수적으로 활용해야 하는 자원이 되었다.

특히 빅데이터와 인공지능(AI) 기술의 발전은 방대한 양의 데이터를 수집, 분석, 처리하는 능력을 극대화시켰다. 그리하여 데이터의 수집과 분석은 이제 기업 활동의 중심으로 자리 잡았으며, 데이터 기반의 비즈니스 모델은 기업의 성공 여부를 결정짓는 중요한 요소가 되었다. 이러한 상황에서 데이터의 수집 및 활용 기술이 빠르게 발전하고 있다. 그러나 데이터 경제가 크게 성장함에 따라 데이터 수집 과정에서 발생할 수 있는 법적, 윤리적 문제가 발생하게 되었고, 그에 대한 해결방안에 대한 논의가 중요하게 되었다.

#### 나. 웹 크롤링의 개념과 역할

웹 크롤링(Web Crawling)은 데이터 수집의 대표적인 기술로, 인터넷 상의 웹사이트에서 자동으로 데이터를 수집하는 과정이다. 웹 크롤러(Web Crawler)는 일종의 자동화된 소프트웨어로, 웹 페이지를 탐색하면서 페이지 내의 정보를 체계적으로 수집하고, 이 데이터를 분석하거나 인덱싱(Indexing)하는 데 활용된다. 이러한 웹 크롤링 기술은 검색 엔진(예: 구글, 네이버), 소셜 미디어 분석, 전자상거래 플랫폼의 가격 비교 서비스, 비즈니스 인텔리전스 등 다양한 분야에서 광범위하게 사용되고 있다.

웹 크롤링의 주요 목적은 방대한 양의 웹 데이터를 효율적으로 수집하여 인사이트를 도출하는 것이다. 그러나 웹 크롤링의 확산은 다양한 법적 쟁점을 야기하고 있습니다. 데이터 수집 과정에서의 무단 접근, 저작권 침해, 개인정보 보호 위반 등이 문제가 될 수 있고, 이러한 쟁점들은 형사법적 책임이 문제가 될 수 있다. 따라서 웹 크롤링이 합법적인 데이터 수집 행위인지, 아니면 불법적 행위로서 형사법적 처벌의 대상이 되는지에 대한 법리적 논의가 필요하다.

#### 다. 대상 판결의 중요성

대법원 2022. 5. 12. 선고 2021도1533 판결(이하 ‘대상 판결’이라고 한다)은 웹 크롤링을 통한 데이터 수집행위가 형사법적 책임을 묻는 데 있어 중요한 기준을 제시한 판례이다. 대상 판결은 피고인이 경쟁사의 웹사이트에서 자동화된 크롤링 프로그램을 통해 데이터를 수집한 행위가 정보통신망 침해, 저작권 침해, 컴퓨터 등 장애 업무방해와 같은 형사범죄가 성립될 수 있는지에 대하여 판단하였다.

대상 판결을 중심으로 웹 크롤링의 데이터 수집행위가 어떠한 형사법적 책임 문제가 발생시키는지와 그 법적 쟁점에 대하여 검토하려고 한다. 이를 통해 웹 크롤링과 관련된 법적 쟁점에 대한 대법원의 판단 기준에 대하여 검토하고, 향후 법적 논의의 발전 방향에 대하여 검토하여 보려고 한다.

## 2. 연구 방법

먼저 웹 크롤링 행위가 정보통신망법, 저작권법, 형법 등에 의해 어떻게 규제되는지를 검토한다. 이를 통하여 웹 크롤링이 어떤 상황에서 형사법상 범죄로 성립할 수 있는지, 그리고 그 기준이 무엇인지에 대해 명확하게 이해할 수 있다.

그리고 대상 판결의 사실관계와 대법원의 판단에 대하여 검토하고, 유사한 법적 쟁점을 다룬 주요 우리나라 판례를 검토하고, 미국, 유럽연합(EU) 등 해외 주요 국가의 웹 크롤링 관련 법률과 판례를 비교, 검토할 것이다. 이러한 비교법적 접근은 각국의 법적 차이점을 이해하는 데 도움이 되며, 한국의 법률과의 비교를 통해 웹 크롤링 규제에 대한 국제적 시사점을 파악할 수 있을 것이다. 이를 통하여 국내 법

률의 개선 방향을 제시하고, 국제적 규제 조화의 필요성을 논의할 수 있을 것이다.

그리하여 크롤링의 데이터 수집행위가 형사법적 책임을 어떻게 구성하는지에 대한 종합적으로 검토하고, 이를 바탕으로 법적 규제의 필요성과 그 방향성에 대하여 검토하려고 한다.

## II. 대상 판결의 사실관계 및 사건의 경과

### 1. 대상 판결의 사실관계

#### 가. 사건 개요 및 당사자

피해자는 모바일 어플리케이션(이하 ‘앱’이라고만 한다)을 통해 숙박 정보를 제공하는 회사로 자사가 구축한 데이터베이스의 무단 사용으로 인해 경제적 손실과 영업상 불이익을 주장하였고, 피고인들은 숙박업체 정보 제공 서비스를 운영하는 회사의 임직원들로 피해자가 제공하는 모바일 어플리케이션의 서버에 접속하여 크롤링 방식으로 숙박업소 목록 등 피해자가 보유·관리하는 데이터베이스를 복제하였다.

피해자는 모바일 앱인 ‘바로예약 앱’(이하 ‘이 사건 앱’이라고 한다)은 ‘프론트엔드(FRONT-END) API 서버’(이하 ‘이 사건 서버’라고 한다)<sup>1)</sup>를 통하여 데이터서비스에서 정보를 받아 회원과 비회원인 이용자에게 그 정보를 제공하였는데, 피고인들을 이 사건 앱을 통하지 않고 크롤링 프로그램인 ‘패킷캡처’ 앱<sup>2)</sup>을 사용하여 바로 이 사건 서버에 접근하여 데이터를 전송받았다.

#### 나. 공소사실의 요지

---

1) 데이터 센터에서 모든 데이터를 처리할 때 과부하와 그에 따른 지연 현상을 완화하기 위하여 사용자 또는 엔드포인트(Endpoint) 단말기의 물적 위치와 인접한 곳에서 컴퓨팅을 수행하는데 이를 엣지(Edge) 컴퓨팅이라고 하고, 엣지 컴퓨터에서 최종 사용자와 가장 밀접한 계층으로 연산처리를 할 수 있는 능력이 있는 장치를 프론트엔드라고 한다(김태균, 크롤링을 통한 데이터 수집의 형사책임, 서강법률논총 제11권 제3호, 276쪽).

2) 컴퓨터 네트워크상에서 이동하는 패킷 또는 트래픽을 가로채거나 저장할 수 있는 프로그램으로 사용법도 어렵지 않고 널리 사용되고 있는 프로그램이다(김태균, 앞의 논문, 같은 쪽)

## 1) 정보통신망이용촉진및정보보호등에관한법률<sup>3)</sup>위반의 점

2016년 1월 초순경, 피고인 A는 피고인 B에게 경쟁업체의 제휴점 정보를 편리하게 수집할 수 있는 프로그램 개발을 요청하였고, 이에 따라 피고인 B는 경쟁업체의 이 사건 앱의 프로그램 소스를 분석하고, 피해자의 이 사건 서버의 모듈과 URL 주소, 그리고 해당 서버로 정보를 호출하는 명령구문들을 알아냈다.

피해자는 이 사건 서버의 URL과 명령구문들을 외부에 공개하지 않았으며, 모바일 앱 이용자들에게만 접근 권한을 부여하고 있었다. 그러나 피고인 B는 PC를 통해 이 사건 서버에 접근하여, 특정 위치를 중심으로 반경 1,000km 내의 숙박업소 정보를 모두 불러오는 기능을 탑재한 프로그램(이하 '크롤링 프로그램')을 개발하여 피고인 A에게 전달하였다.

피고인 C는 이 프로그램을 이용해 이 사건 서버에 접근, 제휴 숙박업소의 정보(업체명, 주소, 방 이름 등)를 무단으로 복제하였고, 피고인 D는 이 정보를 이메일로 보고받았다. 이후 피고인 D는 '마이룸' 객실 서비스 관련 데이터를 추가로 크롤링하기로 결정하고, 피고인 B에게 해당 기능을 프로그램에 추가하도록 지시했습니다. 이에 따라 피고인 B는 '마이룸' 객실의 판매 상태를 확인하고 1시간마다 자동 크롤링하는 기능을 추가하였다.

2016년 8월 19일부터 피고인 C는 개선된 크롤링 프로그램을 사용하여 이 사건 서버에 접근했으며, 피해자는 대량의 호출 신호를 감지하고 피고인들이 사용하던 서버의 IP 주소를 차단하였다. 이에 피고인 E는 서버의 IP 주소를 변경하여 다시 접근할 수 있도록 서버 설정을 변경하였다. 그 결과, 피고인 C는 계속해서 제휴 숙박업소의 정보를 무단으로 복제하였다.

피고인들은 공모하여 2016년 6월 1일부터 10월 3일까지 약 15,945,090회에 걸쳐 피해자의 이 사건 서버에 정당한 접근 권한 없이 침입하였다.

## 2) 저작권법위반의 점

피고인 A는 2016년 1월 초순경 피고인 B에게 경쟁업체의 데이터를 수집할

---

3) 이하 '정보통신망법'이라고만 한다.

수 있는 '크롤링 프로그램'을 개발하도록 지시하였고, 피고인 B는 이 프로그램을 개발하여 피고인 A에게 전달하였다. 이후 피고인 C는 2016년 2월 말부터 이 프로그램을 이용하여 피해자의 데이터베이스에서 제휴 숙박업소의 정보(업체명, 주소, 방 이름, 가격, 입퇴실 시간 등)를 복제하였다. 피고인 D는 이러한 복제된 정보를 피고인 C로부터 이메일로 보고받았으며, 피고인 E는 프로그램이 작동할 수 있도록 서버를 제공하였다.

결과적으로, 피고인들은 2016년 1월 11일부터 6월 27일까지 총 264회에 걸쳐 피해자의 데이터베이스를 무단으로 복제하여 데이터베이스제작자의 권리를 침해하였다.

### 3) 컴퓨터등장애업무방해의 점

피고인 D는 2016년 8월 17일 피고인 B에게 크롤링 프로그램을 개선하여 크롤링 작업을 지시하였다. 그에 따라 피고인 A, B, C, E는 각자 역할을 분담하여 2016년 8월 19일부터 1시간마다 피해자의 전국 제휴 숙박업소의 정보(업체명, 주소, 방 이름, 가격, 입퇴실 시간, 마이룸 판매 여부 등)를 확인하기 위해 피해자의 이 사건 서버에 접속했다. 이 과정에서 피고인들은 반경 1,000km 내의 모든 숙박업소 정보를 요청하는 부정한 명령을 입력하는 방법으로 대량의 정보 호출을 발생시켰다.

이로써 피고인들은 2016년 8월 27일부터 9월 17일까지 총 5회에 걸쳐 정보처리장치에 부정한 명령을 입력하여 장애를 발생시킴으로써 피해자의 숙박 예약 업무를 방해하였다.

### 나. 관련 조문

○ 구 정보통신망법(2018. 12. 24. 법률 제16021호로 개정되기 전의 것)

제2조(정의)

① 이 법에서 사용하는 용어의 뜻은 다음과 같다.

1. “정보통신망”이란 「전기통신사업법」 제2조제2호에 따른 전기통신설

비를 이용하거나 전기통신설비와 컴퓨터 및 컴퓨터의 이용기술을 활용하여 정보를 수집·가공·저장·검색·송신 또는 수신하는 정보통신체제를 말한다.

#### 제48조(정보통신망 침해행위 등의 금지)

- ① 누구든지 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망에 침입하여서는 아니 된다.

#### 제71조(벌칙)

- ① 다음 각 호의 어느 하나에 해당하는 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처한다.

9. 제48조제1항을 위반하여 정보통신망에 침입한 자

#### ○ 저작권법

#### 제2조(정의)

이 법에서 사용하는 용어의 뜻은 다음과 같다

19. “데이터베이스”는 소재를 체계적으로 배열 또는 구성한 편집물로서 개별적으로 그 소재에 접근하거나 그 소재를 검색할 수 있도록 한 것을 말한다.
20. “데이터베이스제작자”는 데이터베이스의 제작 또는 그 소재의 갱신·검증 또는 보충(이하 “갱신등”이라 한다)에 인적 또는 물적으로 상당한 투자를 한 자를 말한다.

#### 제93조(데이터베이스제작자의 권리)

- ① 데이터베이스제작자는 그의 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송(이하 이 조에서 “복제등”이라 한다)할 권리를 가진다.
- ② 데이터베이스의 개별 소재는 제1항에 따른 해당 데이터베이스의 상당한 부분으로 간주되지 아니한다. 다만, 데이터베이스의 개별 소재 또는 그 상당한 부분에 이르지 못하는 부분의 복제등이라 하더라도 반복적이거나 특정한 목적을 위하여 체계적으로 함으로써 해당 데이터베이스의 일반적

인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에는 해당 데이터베이스의 상당한 부분의 복제등으로 본다.

- ③ 이 장에 따른 보호는 데이터베이스의 구성부분이 되는 소재의 저작권 그 밖에 이 법에 따라 보호되는 권리에 영향을 미치지 아니한다.
- ④ 이 장에 따른 보호는 데이터베이스의 구성부분이 되는 소재 그 자체에는 미치지 아니한다.

#### 제136조(벌칙)

- ② 다음 각 호의 어느 하나에 해당하는 자는 3년 이하의 징역 또는 3천만원 이하의 벌금에 처하거나 이를 병과할 수 있다.

3. 제93조에 따라 보호되는 데이터베이스제작자의 권리를 복제·배포·방송 또는 전송의 방법으로 침해한 자

#### ○ 형법

#### 제314조(업무방해)

- ① 제313조의 방법 또는 위력으로써 사람의 업무를 방해한 자는 5년 이하의 징역 또는 1천500만원 이하의 벌금에 처한다.
- ② 컴퓨터등 정보처리장치 또는 전자기록등 특수매체기록을 손괴하거나 정보처리장치에 허위의 정보 또는 부정한 명령을 입력하거나 기타 방법으로 정보처리에 장애를 발생하게 하여 사람의 업무를 방해한 자도 제1항의 형과 같다.

#### 다. 사건의 경과

1) 1심 판결(서울중앙지방법원 2020. 2. 11. 선고 2019고단1777)

1심 법원은 공소사실에 대하여 모두 유죄로 인정하고, 다음과 같이 구체적으로 판단하였다.

#### 가) 정보통신망법위반의 점

피고인들이 API를 정상적인 사용 목적을 벗어나 데이터를 수집했으며, 이는 정보통신망법에서 금지하는 무단 접근에 해당한다고 판단하였다. 특히, 피고인들이

대량의 데이터를 단시간에 수집하여 피해자 회사의 시스템에 부하를 주고 서비스 운영을 방해했다고 판단하였다.

나) 저작권법위반의 점

피해자 회사가 독자적으로 수집하고 정리한 데이터베이스는 저작물로 인정되며, 피고인들이 이를 무단으로 복제하여 사용한 것은 저작권 침해에 해당한다고 판단하였다. 특히, 피고인들이 복제한 데이터의 양이 상당하고, 이를 통해 경쟁 우위를 확보하려 했다는 점을 근거로 들었다.

다) 컴퓨터등장애업무방해:

피고인들의 행위로 인해 피해자 회사의 서버에 과도한 부하가 발생하고, 서비스 이용이 느려지는 등 업무에 실질적인 지장을 초래했다고 판단하였다.

2) 항소심 판결(서울중앙지방법원 2021. 1. 31. 선고 2020노611 판결)

항소심 법원은 1심 판결을 뒤집고, 피고인들의 모든 혐의에 대해 무죄를 선고하였는데, 다음과 같은 이유를 들어 무죄 판결을 하였다.

가) 정보통신망법 위반의 점

피고인들이 사용한 API 서버는 공개된 인터페이스였고, 이를 이용하여 데이터를 수집하는 행위는 정상적인 사용 범위 내에 있다고 판단하였다. 또한, 피고인들의 행위가 피해자 회사의 시스템에 중대한 장애를 초래했다고 보기 어렵다고 판단하였다.

나) 저작권법 위반의 점

피고인들이 복제한 데이터는 공개된 정보이거나, 저작권 보호 대상이 아닌 단순한 사실이라고 판단하였다. 또한, 피고인들의 행위가 피해자 회사의 정당한 이익을 부당하게 침해했다고 보기 어렵다고 판단하였다.

다) 컴퓨터등장애업무방해의 점

피고인들의 행위가 피해자 회사의 업무에 실질적인 지장을 초래했다고 보기 어렵다고 판단하였다.

### Ⅲ. 대상판결의 판단

1. 정보통신망침입의 점에 대한 판단

가. 대법원은 ‘구 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」(2018. 12. 24. 법률 제16021호로 개정되기 전의 것, 이하 ‘구 정보통신망법’ 이라고 한다) 제48조 제1항은 누구든지 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망에 침입하는 것을 금지하고 있고, 이를 위반하여 정보통신망에 침입한 자에 대하여는 5년 이하의 징역 또는 5천만 원 이하의 벌금에 처한다(위 법 제71조 제1항 제9호). 위 규정은 이용자의 신뢰 내지 그의 이익을 보호하기 위한 규정이 아니라 정보통신망 자체의 안정성과 그 정보의 신뢰성을 보호하기 위한 것이므로, 위 규정에서 접근권한을 부여하거나 허용되는 범위를 설정하는 주체는 서비스제공자이다. 따라서 서비스제공자로부터 권한을 부여받은 이용자가 아닌 제3자가 정보통신망에 접속한 경우 그에게 접근권한이 있는지 여부는 서비스제공자가 부여한 접근권한을 기준으로 판단하여야 한다( 대법원 2005. 11. 25. 선고 2005도870 판결 등 참조). 그리고 정보통신망에 대하여 서비스제공자가 접근권한을 제한하고 있는지 여부는 보호조치나 이용약관 등 객관적으로 드러난 여러 사정을 종합적으로 고려하여 신중하게 판단하여야 한다.’ 고 법리를 설시하였다.

나. 대법원은 위 법리에 비추어 다음과 같은 이유로 피고인들의 정보통신망 침입 혐의는 인정하지 아니하였다.

1) 피고인들은 피해자 회사의 이 사건 앱과 통신하는 이 사건 서버의 URL과 명령구문을 알아내어, 자체 개발한 크롤링 프로그램을 사용해 이 사건 서버에 명령구문을 입력하여 숙박업소 정보를 수집했다.

2) 이 사건 서버의 URL과 명령구문은 피해자 회사가 공개하지는 않았지만, 패킷캡처 프로그램 등을 통해 쉽게 알아낼 수 있는 정보였다. 일반 이용자들은 회원가입 여부에 관계없이 이 사건 서버에 자유롭게 접근할 수 있었으며, 접근을 막는 별도의 보호조치는 없었다.

3) 피해자의 이용약관에서는 수집된 정보를 영리 목적으로 이용하거나 제3자에게 제공하는 것을 금지하고 있었지만, 이 약관은 정보 이용을 제한하는 내용이지, 정보 접근을 제한하는 내용은 아니었다.

4) 또한, 이용약관에서 자동접속프로그램 사용으로 인한 서버 부하 행위를 금지하고 있었으나, 이 규정은 회원가입을 하지 않은 이용자들에게는 적용되지 않았으며, 접근권한을 제한하는 것으로 볼 수 없었다.

## 2. 저작권법위반의 점에 대한 판단

가. 대법원은 ‘데이터베이스제작자는 그의 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송(이하 ‘복제 등’ 이라고 한다)할 권리를 가지고(저작권법 제93조 제1항), 데이터베이스의 개별 소재는 데이터베이스의 상당한 부분으로 간주되지 않지만, 개별 소재의 복제 등이라 하더라도 반복적이거나 특정한 목적을 위하여 체계적으로 함으로써 해당 데이터베이스의 통상적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에는 해당 데이터베이스의 상당한 부분의 복제 등으로 본다(저작권법 제93조 제2항). 이는 지식정보사회의 진전으로 데이터베이스에 대한 수요가 급증함에 따라 창작성의 유무를 구분하지 않고 데이터베이스를 제작하거나 그 갱신·검증 또는 보충을 위하여 상당한 투자를 한 자에 대하여는 일정기간 해당 데이터베이스의 복제 등 권리를 부여하면서도, 그로 인해 정보공유를 저해하여 정보화 사회에 역행하고 경쟁을 오히려 제한하게 되는 부정적 측면을 방지하기 위하여 단순히 데이터베이스의 개별 소재의 복제 등이나 상당한 부분에 이르지 못한 부분의 복제 등만으로는 데이터베이스제작자의 권리가 침해되지 않는다고 규정한 것이다. 데이터베이스제작자의 권리가 침해되었다고 하기 위해서는 데이터베이스제작자의 허락 없이 데이터베이스의 전부 또는 상당한 부

분의 복제 등이 되어야 하는데, 여기서 상당한 부분의 복제 등에 해당하는지를 판단할 때는 양적인 측면만이 아니라 질적인 측면도 함께 고려하여야 한다. 양적으로 상당한 부분인지 여부는 복제 등이 된 부분을 전체 데이터베이스의 규모와 비교하여 판단하여야 하며, 질적으로 상당한 부분인지 여부는 복제 등이 된 부분에 포함되어 있는 개별 소재 자체의 가치나 그 개별 소재의 생산에 들어간 투자가 아니라 데이터베이스제작자가 그 복제 등이 된 부분의 제작 또는 그 소재의 갱신·검증 또는 보충에 인적 또는 물적으로 상당한 투자를 하였는지를 기준으로 제반 사정에 비추어 판단하여야 한다. 또한 앞서 본 규정의 취지에 비추어 보면, 데이터베이스의 개별 소재 또는 상당한 부분에 이르지 못하는 부분의 반복적이거나 특정한 목적을 위한 체계적 복제 등에 의한 데이터베이스제작자의 권리 침해는 데이터베이스의 개별 소재 또는 상당하지 않은 부분에 대한 반복적이고 체계적인 복제 등으로 결국 상당한 부분의 복제 등을 한 것과 같은 결과를 발생하게 한 경우에 한하여 인정함이 타당하다.’ 고 법리를 설시하였다.

나. 대법원은 위 법리에 비추어 다음과 같은 이유로 원심(항소심)에서 피고인들의 데이터베이스제작자의 복제권 침해로 인한 저작권법 위반의 점을 무죄로 판단한 것을 수긍하였다.

1) 피고인들이 피해자의 이 사건 서버로부터 수집한 정보들은 피해자의 숙박업소 관련 데이터베이스의 일부에 해당한다.

2) 위 정보들은 이미 상당히 알려진 정보로서 그 수집에 상당한 비용이나 노력이 들었을 것으로 보이지 않거나 이미 공개되어 있어 이 사건 앱을 통해서도 확보할 수 있었던 것이고, 데이터베이스의 갱신 등에 관한 자료가 없다.

3) 이러한 피고인들의 데이터베이스 복제가 피해자의 해당 데이터베이스의 통상적인 이용과 충돌하거나 피해자의 이익을 부당하게 해치는 경우에 해당한다고 보기 어렵다.

### 3. 컴퓨터등장애업무방해위반의 점에 대한 판단

가. 대법원은 ‘형법 제314조 제2항은 컴퓨터 등 정보처리장치 또는 전자기록 등

특수매체기록을 손괴하거나 정보처리장치에 허위의 정보 또는 부정한 명령을 입력하거나 기타 방법으로 정보처리에 장애를 발생하게 하여 사람의 업무를 방해한 자를 처벌하도록 정하고 있다. 여기에서 '허위의 정보 또는 부정한 명령의 입력'이란 객관적으로 진실에 반하는 내용의 정보를 입력하거나 정보처리장치를 운영하는 본래의 목적과 상이한 명령을 입력하는 것이고, '기타 방법'이란 컴퓨터의 정보처리에 장애를 초래하는 가해수단으로 컴퓨터의 작동에 직접·간접으로 영향을 미치는 일체의 행위를 말한다. 한편 위 죄가 성립하기 위해서는 위와 같은 가해행위 결과 정보처리장치가 그 사용목적에 부합하는 기능을 하지 못하거나 사용목적과 다른 기능을 하는 등 정보처리에 장애가 현실적으로 발생하여야 한다.'고 법리를 설시하였다.

나. 대법원은 위 판시의 같은 이유로 원심의 검사가 제출한 증거들만으로 피고인들이 공모하여 정보처리장치에 부정한 명령을 입력하여 장애를 발생하게 하였다고 보기 어렵다고 보아 컴퓨터등장애업무방해위반의 점도 무죄로 판단한 것을 수긍하였다.

### Ⅲ. 웹 크롤링에 대한 일반론

#### 1. 웹 크롤링의 개념 및 목적

##### 가. 개념

인터넷에는 HTML로 작성된 텍스트를 비롯하여 이미지, 동영상 등 다양한 정보가 존재하는데, 웹 크롤링(Web crawling)이라 함은 웹 크롤러(Web crawler)라는 검색로봇인 웹페이지 수집 프로그램을 이용하여 인터넷상에 공개된 웹사이트의 웹 페이지를 순회하며 웹페이지 정보를 수집하는 행위를 말한다.<sup>4)</sup> 마치 거미줄을 치듯이 인터넷 상의 웹 페이지를 따라다니며 정보를 수집하는 자동화된 과정을 통해 방대한 양의 웹 데이터를 빠르고 효율적으로 수집하고, 이를 다양한 목적으로 활용할 수 있다.

4) 유대종, 「웹검색 서비스와 ISP책임에 관한 소고」, 창작과 권리(2007. 3.), 103쪽

## 나. 목적

웹 크롤링은 웹 페이지의 텍스트, 이미지, 링크 등 다양한 형태의 데이터를 수집하여 저장하고, 수집된 데이터를 분석하여 특정 패턴이나 트렌드를 파악하고, 의사 결정에 필요한 정보를 제공하고, 웹 크롤링을 통해 웹 페이지를 색인화하고, 사용자의 검색어와 일치하는 결과를 제공하는 검색 엔진 기능을 하기도 하고, 다양한 웹사이트의 상품 가격을 비교하여 최저가 상품을 찾거나, 시장 동향을 분석하며, 소셜 미디어 데이터를 수집하여 특정 주제에 대한 사람들의 의견이나 감정을 분석하기도 한다.<sup>5)</sup>

## 2. 다른 용어와 차이

웹 스크래핑(Web scraping)은 특정 웹 페이지에서 필요한 데이터만을 추출하는 작업으로 예를 들어, 특정 상품의 가격, 리뷰, 이미지 등 원하는 정보만을 선택적으로 추출하는 것을 특징으로 하나, 웹 크롤링은 웹 페이지 전체를 탐색하며 모든 데이터를 수집하는 광범위한 작업으로 앞서 살펴본 것과 같이 마치 거미줄을 치듯이 웹을 넓게 탐색하며 정보를 수집하는 것을 특징으로 한다.<sup>6)</sup> 그러나 실제로는 양자 모두 웹에 존재하는 데이터의 수집을 위한 목적으로 사용하므로 실질적으로 구별할 실익은 별로 없다.<sup>7)</sup>

또한 다른 사람의 정보처리장치 또는 정보처리조직에 침입하거나 기술적인 방법으로 다른 사람의 정보처리장치가 수행하는 기능이나 전자기록에 함부로 간섭하는 행위를 해킹(hacking)이라고 하는데, 공개된 웹서버를 통하여 웹페이지의 정보를 수집하는 웹 크롤링과는 차이가 있다.<sup>8)</sup>

그리고 데이터 수집방법으로 미러링(mirroring)이라는 개념이 있는데, 미러링은 해킹 등으로 하드웨어에 있는 데이터가 손실되는 것을 막기 위해 데이터를 하나 이

5) 강정희, 웹 크롤링의 데이터 수집행위와 경쟁법적 논의, 사법 제61호(2022), 사법발전재단, 469쪽

6) 장철원, 웹 크롤링 & 데이터 분석 with 파이썬, 인사이트, 2022, 82쪽

7) 김태균, 앞의 논문, 282쪽

8) 유대종, 웹검색 서비스와 ISP책임에 관한 소고, 창작과 권리 제46권, 세창출판사, 2007, 121쪽; 선종수, 크롤링을 통한 데이터 수집과 형사책임, 형사법연구 제35권 제3권(2023 가을), 316쪽

상의 장치에 중복 저장하거나 동일한 내용을 복수의 기기에 사용하게 하는 기능을 말하는데, 실제 크롤링이 문제된 사안에서는 크롤링에 통하여 미러링을 하는 경우가 있다.<sup>9)</sup>

### 3. 웹 크롤링의 주요 활용 사례<sup>10)</sup>

#### 가. 검색 엔진에서 활용

구글(Google), 네이버(Naver) 등 주요 검색 엔진은 웹 크롤링을 통해 웹 페이지를 지속적으로 업데이트하고, 사용자의 검색어와 가장 관련성이 높은 결과를 제공한다.

#### 나. e-커머스에서 활용

가격 비교 사이트, 쇼핑몰 등은 웹 크롤링을 통해 다양한 상품의 가격, 스펙, 리뷰 등을 수집하여 사용자에게 제공한다.

#### 다. 금융에서 활용

금융 시장 데이터, 뉴스 기사 등을 수집하여 시장 분석, 투자 의사 결정 등에 활용한다.

#### 라. 학술 연구에서 활용

학술 논문, 연구 데이터 등을 수집하여 연구에 필요한 자료를 확보한다.

#### 마. 마케팅 분야에서 활용

소셜 미디어, 뉴스 기사 등을 분석하여 소비자 트렌드를 파악하고, 마케팅 전략 수립에 활용한다.

#### 바. 경쟁회사 등의 문제

경쟁회사는 벤치마킹을 위해서 상대방의 웹상 공개데이터를 수집하거나 분석하여 활용하고, 스타트업이 비즈니스를 준비할 때도 종동 업계의 데이터 수집과 분석이 필수적으로 이루어진다.<sup>11)</sup> 대상판결의 사안이 이 경우에 해당한다.

---

9) 선종수, 앞의 논문, 317쪽

10) 강정희, 앞의 논문, 469쪽

11) 김태균, 앞의 논문, 283쪽

#### 4. 웹 크롤링의 기술적 발전과 한계

##### 가. 웹 크롤링의 기술적 발전

웹 크롤링은 초기에는 단순한 텍스트 데이터 수집에서 시작하여 오늘날에는 복잡하고 구조화된 데이터나 비정형 데이터를 효율적으로 수집할 수 있는 강력한 도구로 발전하였는데, 주요한 기술적 발전을 살펴보면 다음과 같다.

##### 1) 검색 엔진에서 크롤링 발전

초기의 웹 크롤러는 텍스트 기반의 웹 페이지에서 정보를 수집하는 데 한정되었으나, 구글과 같은 검색 엔진의 등장 이후 웹 크롤러는 더 복잡한 웹 페이지 구조에서도 데이터를 효과적으로 수집할 수 있게 되었다. 현대의 웹 크롤러는 수백만 개의 웹 페이지를 빠르게 분석하고 인덱싱하여 사용자에게 관련된 검색 결과를 제공하는 데 필수적인 역할을 한다.

##### 2) 머신러닝 및 자연어 처리의 도입

최근에는 머신러닝과 자연어 처리(NLP<sup>12)</sup>) 기술이 웹 크롤링에 적용되어 단순한 데이터 수집을 넘어 수집된 데이터를 분석하고 패턴을 인식하게 한다. 이러한 기술을 통해 웹 크롤러는 웹 페이지에서 단순한 텍스트 외에도 이미지를 분석하거나, 사용자 리뷰나 소셜 미디어 게시물과 같은 비정형 데이터를 처리할 수 있게 되었는데, 상품 리뷰를 크롤링하여 소비자, 이용자의 긍정적·부정적 감정 분석을 수행하거나, 소셜 미디어의 트렌드를 추적할 수 있다.

##### 3) 대규모 데이터 수집과 처리 능력

빅데이터 시대가 오면서 웹 크롤러의 대규모 데이터 수집과 처리 능력이 중요해졌다. 클라우드 컴퓨팅의 발전과 더불어 현대의 웹 크롤러는 분산 시스템을 사용하여 수백만 개의 웹 페이지에서 데이터를 수집하고, 동시에 실시간으로 처리할 수 있는 능력을 갖추게 되었다. 이로 인해 기업들은 경쟁사 분석, 시장 트렌드 추

---

12) 자연어 처리(NLP, Natural Language Processing)는 컴퓨터가 인간의 언어를 분석, 이해하고, 해석하며, 생성하는 기술이다(네이버 지식백과 참조).

적, 가격 비교와 같은 비즈니스 인텔리전스 활동을 더욱 효율적으로 수행할 수 있게 되었다.

#### 4) 동적 웹 페이지 크롤링

초기의 웹 크롤러는 주로 정적(Static) 웹 페이지를 대상으로 작업했지만, 오늘날의 웹 페이지는 자바스크립트나 AJAX<sup>13)</sup> 기술을 사용하여 동적으로 콘텐츠를 생성한다. 이에 대응하기 위해, 현대 웹 크롤러는 동적으로 생성된 콘텐츠를 수집할 수 있는 기능을 갖추고 있다. 이러한 기술은 웹 브라우저를 모방하여 자바스크립트가 실행되는 페이지에서도 필요한 데이터를 정확하게 추출할 수 있게 해준다.

#### 나. 웹 크롤링의 한계

웹 크롤링은 기술적으로 많은 발전을 이루었으나, 여전히 여러 가지 한계와 도전 과제가 존재하는데, 주요한 한계는 다음과 같다.

##### 1) 법적 규제와 윤리적 문제<sup>14)</sup>

웹 크롤링은 데이터 수집 과정에서 법적 문제를 야기할 수 있다. 예를 들어, 웹사이트의 이용 약관에 따라 크롤링이 금지되거나, 저작권법에 의해 특정 데이터를 무단으로 수집하는 것이 불법으로 간주될 수 있다. 또한, 개인정보 보호법에 의해 개인정보를 포함한 데이터 수집이 엄격하게 규제되고 있어, 웹 크롤링 기술의 적용에 법적 한계가 따른다. 윤리적으로도, 경쟁업체의 데이터를 무단으로 수집하는 행위는 공정 경쟁에 위배될 수 있다. 이러한 문제가 대상 판결 사안에서 주로 문제가 될 것이다.

##### 2) 차단 및 봇 탐지 기술<sup>15)</sup>

많은 웹사이트는 웹 크롤러로부터 서버를 보호하기 위해 다양한 차단 및 봇

---

13) (Asynchronous JavaScript and XML, 에이잭스)는 비동기적인 웹 애플리케이션의 제작을 위해 아래와 같은 조합을 이용하는 웹 개발 기법이다(위키백과 참조, <https://ko.wikipedia.org/wiki/Ajax>)

14) 김태균, 앞의 논문, 283쪽

15) 선종수, 앞의 논문, 319쪽

탐지 기술을 사용하고 있다. 예를 들어, robots.txt<sup>16)</sup> 파일을 통해 크롤링을 금지하거나, IP 차단, CAPTCHA<sup>17)</sup> 적용 등을 통해 자동화된 크롤링 시도를 막는 기술이 있다. 또한, 웹사이트는 크롤러의 활동을 탐지하기 위해 비정상적인 트래픽 패턴을 감지하여 봇 활동을 차단할 수 있다. 이는 크롤러가 목표로 하는 데이터를 수집하는 데 어려움을 초래할 수 있다.

### 3) 비정형 데이터 처리의 어려움

웹 페이지는 종종 구조화되지 않은 비정형 데이터를 포함하고 있으며, 이러한 데이터를 정확하게 수집하고 분석하는 것은 여전히 어려운 문제이다. 예를 들어, 동일한 정보가 서로 다른 웹 페이지에서 다양한 형식으로 나타날 수 있으며, 이러한 데이터의 표준화가 어렵다. 비정형 데이터를 구조화하고 의미 있는 정보를 추출하는 데는 추가적인 알고리즘이나 사전 학습된 모델이 필요하며, 이는 크롤링 작업의 복잡성을 증가시킨다.

### 4) 실시간 데이터 수집의 한계

웹 크롤러는 일반적으로 주기적으로 웹 페이지를 수집하여 데이터를 갱신하는 방식을 사용한다. 그러나 실시간 데이터를 필요로 하는 환경에서는 이러한 방식이 적합하지 않을 수 있다. 예를 들어, 주식 가격 변동이나 실시간 뉴스 업데이트와 같은 데이터는 즉각적으로 수집되고 처리되어야 하지만, 기존의 웹 크롤링 방식은 주기적 수집에 의존하기 때문에 실시간 데이터를 즉각적으로 반영하는 데 한계가 있다.

### 5) 대규모 크롤링 시 서버 부하 문제

---

16) 로봇 배제 표준(robots exclusion standard), 로봇 배제 프로토콜(robots exclusion protocol)은 웹 사이트에 로봇이 접근하는 것을 방지하기 위한 규약으로, 일반적으로 접근 제한에 대한 설명을 robots.txt에 기술한다(위키백과 참조, <https://en.wikipedia.org/wiki/Robots.txt>).

17) CAPTCHA(Completely Automated Public Turing test to tell Computers and Humans Apart, 완전 자동화된 사람과 컴퓨터 판별, 캡차)는 HIP(Human Interaction Proof) 기술의 일종으로, 어떠한 사용자가 실제 사람인지 컴퓨터 프로그램인지를 구별하기 위해 사용되는 방법이다(위키백과 참조, <https://ko.wikipedia.org/wiki/CAPTCHA>).

대규모 웹 페이지를 대상으로 하는 크롤링 작업은 대상 웹사이트에 부하를 줄 수 있다. 대량의 요청을 한 번에 처리해야 하는 웹사이트 서버는 이로 인해 과부하에 걸릴 수 있으며, 이는 서비스 제공자에게는 심각한 문제로 다가올 수 있다. 또한, 서버 부하로 인해 웹사이트의 성능 저하나 중단이 발생할 경우 크롤링 시도 자체가 차단될 수 있다.

#### 6) 데이터 정확성과 품질 문제

웹 크롤링을 통해 수집된 데이터는 웹 페이지에 의존하기 때문에 원본 데이터의 정확성과 신뢰성이 문제될 수 있다. 잘못된 정보나 오래된 데이터가 웹 페이지에 포함된 경우 크롤링 결과 역시 부정확할 수 있다. 특히, 크롤링 결과를 바탕으로 의사결정을 내리는 경우 데이터의 품질에 문제가 발생할 위험이 있다.

### 5. 웹 크롤링에 대한 법적 규제<sup>18)</sup>

웹 크롤링이 확산되면서 데이터의 수집 및 사용 과정에서 발생할 수 있는 법적 쟁점들이 중요하게 문제되고 있다. 대상 판결의 사안을 구체적으로 검토하기 전에 먼저 발생할 수 있는 법적 규제에 대하여 살펴 볼 필요가 있다. 웹 크롤링에 대한 법적 규제는 국가별, 상황별로 다르게 적용될 수 있으며, 특히 저작권 침해, 정보통신망 침입, 개인정보 보호 등 여러 법적 문제가 제기될 수 있다.

#### 가. 웹 크롤링과 저작권법

웹 크롤링을 통해 수집된 텍스트, 이미지, 데이터베이스 등은 저작권법에 의해 보호될 수 있다. 저작권법에 따라 저작물이 창작적 요소를 갖추고 있다면 저작권 보호 대상이 되고 이를 무단으로 사용하는 것은 저작권 침해 행위가 될 수 있다.

또한 많은 웹사이트가 제공하는 정보는 데이터베이스의 형태를 띠고 있는데, 저작권법에서는 데이터베이스 제작자의 권리를 보호하며, 데이터베이스의 전부 또는 상당한 부분을 무단으로 복제하거나 배포하는 행위를 저작권 침해로 간주할 수

---

18) 김태균, 앞의 논문, 283쪽

있다. 웹 크롤링을 통해 데이터베이스에서 정보를 대규모로 수집할 경우 데이터베이스 제작자의 권리를 침해할 수 있다. 이 부분이 대상 판결 사안에서 공소사실 중 하나로 문제된다.

다만 저작권법에서는 공정 이용의 원칙을 인정하고 있으며, 이는 비영리적 목적이거나 학술적 연구, 비판, 보도 목적으로 저작물을 사용하는 경우 저작권 침해가 아니라고 판단할 수 있는 기준을 제공한다. 하지만 상업적 목적의 웹 크롤링은 공정 이용으로 인정받기 쉽지 않을 것이다.

#### 나. 정보통신망 침입

정보통신망침입죄는 웹 크롤링과 관련하여 자주 논의되는 형사법적 쟁점으로 대상판결에서도 쟁점 중 하나이다. 정보통신망법은 정보통신망에 무단으로 접근하거나 접근 권한을 초과하여 침입하는 행위를 금지하고 있으며, 이를 위반할 경우 형사처벌의 대상이 된다.

#### 다. 접근 권한의 문제

많은 웹사이트는 robots.txt 파일을 통해 웹 크롤러가 접근할 수 있는 영역을 설정한다. 이 파일은 크롤러가 접근 가능한 페이지와 접근이 제한된 페이지를 명시하지만, 법적 구속력은 없다. 즉, robots.txt는 사이트 운영자의 의사 표현일 뿐 이를 위반했다고 해서 곧바로 법적 책임이 따르는 것은 아니다. 그러나 법원에서는 robots.txt를 무시하고 크롤링을 강행하는 행위를 악의적인 접근으로 간주할 수 있다.

미국에서 LinkedIn과 HiQ Labs 사건은 웹 크롤링과 관련된 정보통신망 침입 문제의 대표적인 사례이다. LinkedIn은 HiQ Labs가 자사의 공개 프로필 데이터를 크롤링하는 행위를 금지하려 했으나, 법원은 HiQ Labs가 크롤링한 데이터가 공개된 정보이기 때문에 정보통신망 침입이 아니라고 판결했다. 이는 공개된 정보에 대한 크롤링이 경우에 따라 합법적으로 인정될 수 있음을 보여준다.

#### 라. 개인정보 보호

개인정보 보호 역시 웹 크롤링에서 중요한 법적 이슈다. 웹 크롤러가 개인정보가 포함된 데이터를 수집할 경우 이는 개인정보 보호법에 위반될 수 있으며, 특히 유럽연합 일반 데이터 보호 규칙(GDPR, 이하 ‘GDPR’ 이라고만 한다)<sup>19)</sup>과 같은 엄격한 개인정보 보호 규제는 이에 대한 책임을 물을 수 있는데, EU 시민의 개인정보를 수집하거나 처리하는 모든 기업에 적용된다. 위 규칙은 웹 크롤러가 개인정보를 수집하는 경우 반드시 개인정보 소유자의 동의를 받아야 하며, 이를 위반할 경우 높은 벌금을 부과할 수 있다.

#### 마. 웹사이트 이용 약관

웹사이트의 이용 약관에서 웹 크롤링을 명시적으로 금지하는 경우, 크롤러가 약관을 위반하고 데이터를 수집하는 것은 법적 분쟁의 대상이 될 수 있다. 이는 특히 상업적 목적으로 웹사이트의 데이터를 대량 수집하는 경우 문제가 된다. 웹사이트의 사용자는 웹사이트의 약관에 동의함으로써 사용자는 사이트의 규정을 준수할 의무가 있으며, 이를 위반하는 행위는 계약 위반이 될 수 있다.

### IV. 웹 크롤링에 대한 비교법적 검토

#### 1. 미국

##### 가. 컴퓨터 사기 및 남용 방지법(The Computer Fraud and Abuse Act)

미국은 1996년 컴퓨터 사기 및 남용 방지법(이하 ‘CFAA’ 라고만 한다)을 연방법으로 신설하여 해킹 등 권한이 없거나 권한을 넘는 컴퓨터에 대한 접근에 대하여 처벌하고 있다.<sup>20)</sup> CFAA는 타인의 컴퓨터에 대한 고의적인 침입, 즉 해킹에 대한 대응으로 입법되었다. CFAA에서 규정하는 ‘권한 없는’ 접근이 무엇인지에 대하여 ‘코드기반 규제(Code-based restriction)<sup>21)</sup>’에 위반하여 접근한 경우에는 권

19) GDPR(General Data Protection Regulation)은 유럽연합에 속해있거나 유럽경제지역(EEA)에 속해 있는 모든 사람들의 사생활 보호와 개인정보를 보호하는 규정이다(네이버 검색결과)

20) 정현순, 크롤링에 의한 데이터 수집이 정보통신망 침입에 해당하는지 여부, 사법 제61권(2022), 사법발전재단, 352쪽

한 없는 접근에 해당한다고 보는 것에 의견이 일치하는 것으로 보이고, 가장 느슨한 방식인 ‘서비스 이용약관(terms of service)’에 위반하여 접근한 경우 권한 없는 접근에 해당하지 않는다고 해석하는 것이 일반적이라고 한다.<sup>22)</sup>

#### 나. HiQ Labs v. LinkedIn 판결<sup>23)</sup>

크롤링 프로그램을 이용하여 경쟁자의 웹 사이트로부터 사업 관련 정보를 취득하는 경우에도 CFAA 위반인지에 대하여 연방대법원은 위 판결에서 판단하였으므로, 이를 살펴보기로 한다.

##### 1) 사건의 개요

HiQ Labs는 LinkedIn 사용자들의 공개된 프로필 정보를 웹 크롤링하여 이 데이터를 기반으로 분석 서비스를 제공하던 회사다. LinkedIn은 HiQ Labs의 크롤링을 중단시키기 위해 중단 및 금지 통보(Cease and Desist)를 발송하고, 이를 위반할 경우 CFAA를 위반한 것으로 간주할 것이라고 경고했다. LinkedIn은 웹 크롤링이 자사 서버에 대한 무단 접근이라며 법적인 문제를 삼으면서 크롤링의 중단 통지 및 HiQ Labs의 IP 접속을 차단하는 조치를 취하였다. 이에 HiQ Labs가 접속차단 금지명령(Injunction)을 연방법원에 청구하였다.

##### 2) 쟁점

위 사건에서 주된 쟁점은 HiQ Labs가 공개된 LinkedIn 사용자 프로필을 크롤링하는 것이 정보통신망 침입이나 CFAA 위반으로 볼 수 있는지에 관한 것이었다. LinkedIn은 HiQ Labs의 크롤링이 이용 약관을 위반했고, 크롤링을 차단하기 위해 기

---

21) 코드기반규제란 패스워드에 의하여만 접근을 허용하거나 사용권한을 다르게 규정한 계정(account)을 부여하는 프로그램을 기반으로 접근권한을 정하여 주는 방법을 말한다(조성훈, 주거침입과 정보통신망침입: 보호법익과 ‘침입’의 해석론을 중심으로, 선진상사법률연구 80호, 법무부(2017. 10.), 144쪽)

22) 정현순, 앞의 논문, 353쪽; 조성훈, 앞의 논문, 144쪽

23) HiQ Labs, INC. v. LINKEDIN CORPORATION, 141 S. CR, 2752, 210 L. Ed, 2d902(2021)

술적 방어 수단을 사용했다고 주장하였다. 반면 HiQ Labs는 크롤링한 정보가 공개된 데이터이며, 이를 크롤링하는 것이 불법이 아니라고 주장하였다.

### 3) 법원의 판단

연방대법원은 HiQ Labs의 주장을 받아들였다. 연방대법원은 공개된 프로필 정보를 대상으로 한 크롤링은 무단 접근에 해당하지 않고, LinkedIn이 HiQ Labs의 크롤링을 방해하는 것은 HiQ Labs의 정당한 사업 활동을 침해할 수 있다고 판결하였다. 이 판결은 HiQ Labs가 공개된 정보를 크롤링하는 것이 법적으로 허용된다는 중요한 선례를 제공하였다.

#### 다. 그 후 미국 판례들의 경향

위 판결이후 미국 판례들의 경향은 누구나 이용할 수 있는 웹 사이트에 대한 웹 크롤링은 공익적 관점, 미연방 수정 헌법 제1조, 웹 크롤링과 웹 브라우징의 기술적 유사성, CFAA가 크롤링과 같은 행위를 규제하기 위하여 입법된 것은 아니라는 주장 등을 기초로 CFAA 위반인지에 대하여 엄격한 기준으로 판단하고 있다.<sup>24)</sup>

## 2. 유럽연합

유럽연합에서는 2001년에 유럽평의회(Council of Europe) 주도로 ‘유럽 사이버 범죄 협약(Convention on Cybercrime, 일명 부다페스트협약)’이 제정되어 2004. 7. 1. 효력이 발생하였고, 2020. 11. 기준으로 유럽국가를 비롯해서 미국, 일본, 캐나다 등 65개국이 가입하였으나, 우리나라는 가입하지 아니하고 있다.<sup>25)</sup>

위 협약 제2조는 정보통신망에 대한 불법접근에 대하여 다음과 같이 규정하고 있다.

각 당사국은 권한 없이 컴퓨터시스템의 일부 또는 전부에 고의적으로 접속하는 행위를 국내법상 범죄로 하는데 필요한 입법 및 그 밖의 조치를 취해야 한다. 당사

24) 구체적인 미국 판례의 경향에 대하여는 정현순, 앞의 논문, 355-364쪽 참조

25) 박희영 외 2, 사이버범죄협약 이행입법 연구, 대검찰청 연구용역 결과보고서(2015), 22쪽; 정현순, 앞의 논문, 364쪽

국은 컴퓨터 데이터를 획득할 의도로, 기타 부정한 의도로 또는 다른 컴퓨터시스템과 연결된 컴퓨터시스템과 관련하여 보호조치를 침해하는 것을 범죄로 할 수 있다.

위 규정에 따르면 ① 권한 없이 컴퓨터시스템의 일부 또는 전부에 고의적으로 접속하는 행위의 경우에 대하여만 권한 없는 접속에 대하여 규정하고 있고, 나머지 ② 컴퓨터 데이터를 획득할 의도로 보호조치를 침해하는 행위, ③ 기타 부정한 의도로 보호조치를 침해하는 행위, ④ 다른 컴퓨터시스템과 연결된 컴퓨터시스템과 관련하여 보호조치를 침해하는 행위는 보호조치 침해를 요건으로 규정하고 있다. 따라서 공중에게 자유롭게 공개되어 접근을 허용하는 컴퓨터시스템에 접근하는 행위는 불법이라거나 범죄에 해당하지 아니하고, 일상적으로 사용하는 통신프로토콜이나 프로그램에 제공되는 기본도구를 이용하는 것은 ‘권한 없이’에 해당하지 않는다고 한다.<sup>26)</sup>

## Ⅶ. 대상판결에서 웹 크롤링을 통한 데이터 수집에 대한 형사책임에 대한 검토

### 1. 형사법적 쟁점

웹 크롤링을 통한 데이터 수집행위는 인터넷 분야에서 널리 사용되고 있는데, 웹사이트 운영자와 관계에서 법적인 문제가 발생할 수 있다. 그 중 형사법적으로 문제되는 쟁점으로는 ① 웹 크롤링이 정보통신망법에서 금지하는 정보통신망침입에 해당하는지, ② 웹 크롤링을 통한 데이터 수집이 저작권법상 데이터베이스제작자의 권리를 침해하는지, ③ 웹 크롤링이 형법 제314조 제1항에서 규정한 컴퓨터등장애 업무방해에 해당하는지, ④ 웹 크롤링을 활용한 경쟁자의 데이터수집행위가 부정경쟁방지법상 위법행위에 해당하는지가 있다.

대상판결에서는 위 쟁점 중 ①, ②, ③에 대하여 공소가 제기되어 판단하고 있으므로, 각 쟁점별로 구체적으로 검토하여 보기로 한다.

### 2. 정보통신망침입의 점에 대한 검토

---

26) 정현순, 앞의 논문, 365쪽

#### 가. 판례와 학설의 입장

종전 대법원은 정보통신망 침입 규정은 정보통신망 자체의 안정성과 그 정보의 신뢰성을 보호하는 것이라고 하면서 정보통신망에 대한 접근권한을 부여하거나 허용되는 범위를 설정하는 주체는 서비스제공자이고, 접근권한이 있는지 여부는 서비스제공자가 부여한 접근권한을 기준으로 판단하여야 한다고 판시하였다.<sup>27)</sup> 정당한 접근권한 없이 또는 허용된 접근권한을 넘어 정보통신망 침입에 밀접한 행위를 하면 실행의 착수가 있고, 그와 같은 침입이 완료되면 기수에 이르렀다고 볼 것이고 정보통신망에 침입한 후 시스템상 프로그램을 실행하거나 데이터를 송·수신하는 등의 행위까지 나아갈 것을 요구하지 않는다.<sup>28)</sup>

학설에는 인터넷의 특성상 정보의 접근은 자유로워야 하는 것이 원칙이므로, 서비스제공자가 접근을 제한하고 그 제한을 위반한 행위가 정보통신망 침해에 해당하려면 접근권한의 설정은 명시적으로 이루어져야 한다는 견해가 있고,<sup>29)</sup> 정보통신망에 접근한 주관적 의도로 판단하는 것이 아니고, 피고인의 웹 크롤링 행위로 인하여 정보통신망의 안정성이라는 법익이 침해되었는지 여부로 정보통신망 침해가 성립하는지를 판단하여야 한다는 견해가 있다.<sup>30)</sup> 또한 정보통신망 침입은 정보통신체계의 안정성을 보호하는 동시에 사이버공간의 재산적, 인격적 법익을 보호하는 규범이므로, 인터넷에 연결된 웹서버와 게시된 정보는 따로 접근권한을 부여할 수 없는 개방적 공간이고, 물리적, 논리적으로 망을 분리하고 코드기반 기제를 적용하는 등으로 닫힌 공간을 설정할 수 있을 뿐이어서 모든 위반행위를 처벌하는 포괄적 규정으로 전환될 위험을 피할 수 있다는 견해도 있다.<sup>31)</sup>

대사판결은 위 대법원 2005도870 판결의 법리에 따라 정보통신서비스제공자가 부여한 접근권한을 기준으로 판단하면서 서비스제공자가 접근권한을 부여하였는지를 판단하는 기준으로 서비스이용자들이 외부에서 확인할 수 없는 서비스제공자의 내심의 의사나 목적으로 판단하지 아니하고, 서비스제공자의 보호조치나 이용약관

27) 대법원 2005. 11. 25. 선고 2005도870판결

28) 김태균, 앞의 논문, 287쪽

29) 권영준, 인터넷상 정보에 대한 접근 및 취득행위의 위법성, 비교사법 14권 3호(2007.9.), 282쪽

30) 최상진, 경쟁사의 무단 크롤링에 대한 법적 대응방안에 대한 연구, Law&Technology 17권 1호, 서울대학교 기술과법센터(2021. 1.), 33-34면

31) 조성훈, 앞의 논문, 335쪽

등과 같이 객관적으로 드러나게 되는 여러 사정 등을 종합적으로 고려하여 판단하여야 한다고 판시하였다.

서비스제공자의 주관적 의사나 목적은 외부에서 파악하기 어려우므로, 그것만으로 접근권한을 판단하는 것은 법적 안정성을 해치는 것으로 보이고, 인터넷의 이용을 과도하게 위축시키는 결과를 낼 수 있어 외부에 드러난 객관적인 사정을 고려하여 판단하여야 한다는 대상판결의 법리를 타당한 것으로 보인다.<sup>32)</sup>

#### 나. 이용약관

서비스제공자는 사전에 정보통신망의 접근권한에 대하여 이용약관이라는 이름으로 계약의 내용을 설정하는 경우가 있는데, 대법원은 인터넷온라인 게임의 이용자이자 개정 명의자인 피고인이 자신의 계정을 양도한 후 그 계정을 사용 중인 양수인이 설정해 둔 비밀번호를 임의로 변경한 사안에서 정보통신서비스제공자에 의하여 그 접근권한이 부여되거나 허용된 자가 누구인지 정함에 있어서 이용약관을 고려하여야 한다는 취지로 판시하였다.<sup>33)</sup>

다만 이용약관은 서비스제공자의 일방적인 의사에 의하여 설정되는 것이기 때문에 서비스이용자의 입장이 반드시 제대로 반영되기는 어려울 것이다. 또한 서비스제공자는 서비스이용자에게 서비스를 제공하는 측면이 있으나, 서비스이용자도 서비스이용을 하면서 자신의 데이터 등을 제공하는 측면이 있어 서비스제공자의 일방적인 정함에 따라 접근권한을 판단하는 것은 비례성에 반하기도 한다. 또한 이용약관이 명확하지 않거나 실제 이용자가 제대로 이행하지 못한 상태에서 동의한 경우에도 침입이 인정된다면 과도한 형사처벌의 우려도 있다.<sup>34)</sup>

따라서 이용약관의 내용도 접근권한을 판단하는 하나의 기준으로 고려할 수 있으나, 이용약관 자체가 접근자체를 금지하는 것이 아니어서 서비스이용자가 누구든지 접근하여 정보수집이 가능한 정보통신서비스에 접근한 후 데이터 등을 가져가는 것이 정보통신서비스에 대한 침입이라고 보기는 어렵다. 아래에서 살펴보는 것처럼

32) 정현순, 앞의 논문, 369쪽

33) 대법원 2020. 7. 22. 선고 2010도63 판결

34) 선종수, 앞의 논문, 371쪽

이용약관과 함께 기술적 보호조치가 있는지도 고려하여 접근권한을 넘어서는 침입이라고 판단하여야 할 것이다.

#### 다. 기술적 보호조치

기술적 보호조치란 정보통신망의 보호를 위한 구체적인 기술적 장치를 말하는데, 대표적인 보호조치를 웹 사이트 이용시 아이디, 비밀번호를 요구하는 로그인 방식이 있다. 그런데 정보통신망 제48조 제1항 위반죄는 명시적으로 침입을 방지하기 위한 기술적 보호조치를 침해거나 훼손할 것으로 요구하지는 않는다.

이론적으로는 명시적인 규정이 없기 때문에 기술적 보호조치가 없는 경우에도 침입죄가 성립될 수 있다는 견해, 기술적 보호조치를 침해, 훼손 또는 우회하는 경우를 처벌하고자 하는 입법 취지를 고려하여 기술적 보호조치의 존재가 침입죄 성립에 요건이 되어야 한다는 견해도 있을 수 있다. 대법원은 기술적 보호조치를 구성요건으로 보고 있지는 않지만 서비스이용제공자의 접근권한을 판단하는 기준으로 검토하여야 한다고 판시하였고, 대상판결에서도 마찬가지로 법리를 판시하였다.<sup>35)</sup>

정보통신망침입죄는 해당 법규정에서 명시적으로 기술적 보호조치를 요건으로 하고 있지 아니하므로 기술적 보호조치의 존재를 전제로 보호조치를 침해, 훼손 또는 우회하여야만 위 죄가 성립된다고 보기는 어려울 것이다. 다만 기술적 보호조치가 전혀 없는 정보통신망서비스에 접근한 이용자에 대하여 서비스제공자의 이용약관이라 의사에 반한다는 이유로 위 죄를 인정한다면 역시나 과도한 형사처벌의 우려가 있다. 따라서 대상판결에서 판시한 것과 같이 기술적 보호조치를 통하여 접근권한을 부여하였거나 제한하였는지를 고려하여 접근권한을 판단하는 것이 타당하다고 볼 것이다.<sup>36)</sup>

#### 라. 검토

##### 1) 피해자가 이 사건 서버에 부여한 접근권한에 대한 검토

35) 대법원 2008. 1. 18. 선고 2007도7096 판결

36) 정현순, 앞의 논문, 375쪽;

정보통신망침입죄는 정보통신서비스제공자가 부여한 접근권한에 따라 형사 처벌되는 지가 문제가 되므로, 앞서 살펴 본 것과 같이 정보통신서비스제공자의 주관적 의사보다는 객관적인 사정을 검토하여 접근권한 범위를 판단하여야 한다. 대 상판결의 사안을 검토하여 보면, 이 사건 서버가 이 사건 앱에만 연결되어 있는 사 정은 이 사건 서버에서 이 사건 앱의 이용자와 데이터베이스 서버 사이에서 정보를 처리하는 역할 수행하기 위해서 시스템을 그러한 방식으로 구축하였다고 봄이 타당 하고, 그러한 사정만으로 피해자가 명시적, 객관적으로 이 사건 앱 이용자에게만 접근권한을 부여한 것으로 보기는 어렵다고 할 것이다.

피고인들이 이 사건 앱을 통하지 않고 이 사건 서버에 접근하였고, 이 사건 서버에 접근할 수 있는 URL 및 명령이 표시되어 있지 않은 것은 사실이나, 이 사건 앱에 이 사건 서버의 URL 등이 표시되지 않은 것은 앱의 특성상 클릭(터치)만으로 실행되므로 URL 표시 자체가 필요가 없고, URL은 인터넷 주소로 통상 숨길 필요도 없으며, 피고인들이 URL 등을 알아내기 위해 사용한 ‘패킷 캡처’라는 프로그램과 같이 일반적으로 사용하는 프로그램을 사용하는 등의 간단한 기술적 조치만으로 URL을 쉽게 파악할 수 있다는 점을 고려하면, URL이 표시되지 아니하였다는 점만 으로 피해자가 이 사건 앱 사용자들에게만 이 사건 서버에 접근할 수 있는 권한을 객관적으로 부여하였다고 보기도 어렵다고 판단된다.<sup>37)</sup>

## 2) 이용 약관 내용에 대한 검토

피해자의 이 사건 서비스 이용 약관은 그 적용 대상을 회원과 이용자로 구분 하고 있는데, 이용 약관 제19조 제6호는 ‘이용자는 회사를 이용함으로써 얻은 정보 를 회사의 사전 승낙 없이 복제, 송신, 출판, 배포, 방송 등 기타 방법에 의하여 영 리 목적으로 이용하거나 제3자에게 이용하게 하여서는 안 됩니다’ 라고 되어 있으 나, 위 약관 내용의 제한 내용은 이용자의 이용 범위의 제한이라고 보이고, 접근 자 체를 제한한 것이라고 보기는 어렵다.

또한 이용 약관 제9조 제2항 제9호는 ‘회원은 자동 접속 프로그램 등을 사용

---

37) 정현순, 앞의 논문, 376쪽

하는 등 정상적인 용법과 다른 방법으로 서비스를 이용하여 회사의 서버에 부하를 일으켜 회사의 정상적인 서비스를 방해하는 행위를 하여서는 안 됩니다.’ 라고 되어 있지만, 위 규정은 회원들에게만 적용되는 것으로 되어 있어 이용자들에게도 적용된다고 보기 어렵고, 다음 항에서 살펴보는 것과 같이 피해자가 이 사건 서버의 접근을 막기 위한 아무런 보호조치를 하지 않고 있는 점을 고려하면, 위 규정이 회원 아닌 이용자들에게 적용된다고 하더라도 이용 약관만으로 객관적으로 접근 권한을 제한한 것으로 보기는 어렵다고 판단된다.<sup>38)</sup>

### 3) 보호조치 여부에 대한 검토

피해자는 일반적으로 사용되는 기술적 보호조치나 자동수집 거부 표시 등을 전혀 하지 아니하였다. 즉 일반적으로 서버와 브라우저 사이에 주고받는 데이터를 암호화하는 기술인 SSL<sup>39)</sup>이 사용되고 있으나, 피해자는 이를 사용할 수 있었음에도 사용하지 아니하였고, 일반적으로 크롤링을 거부하기 위하여 사용하는 로봇툴<sup>40)</sup>에 따른 ‘robots.txt<sup>41)</sup>’ 도 사용하지 아니하였다.<sup>42)</sup>

앞서 살펴본 것과 같이 대상판결에서 문제된 정보통신망침해죄는 이 사건 서버에 보호조치하였는지를 구성요건으로 되어 있지는 아니하나, 피해자가 설정한 접근권한의 범위를 판단하기 위한 고려요소로 보고 있으므로, 위와 같이 피해자가 어떠한 보호조치를 하지 아니한 점도 정보통신망침해죄를 성립하는지를 판단함에 고려되어야 할 것이다.

### 4) 소결

피해자의 이 사건 서버는 누구나 접근이 가능하고, 이에 대한 접근을 막는

38) 정현순, 앞의 논문, 378쪽; 김태균, 앞의 논문, 293쪽.

39) SSL(Secure Sockets Layer)은 인터넷 프로토콜(Internet protocol)이 보안면에서 기밀성을 유지하지 못한다는 문제를 극복하기 위해 개발되었다. 현재 전세계에서 사용되는 인터넷 상거래시 요구되는 개인 정보와 신용카드 정보의 보안 유지에 가장 많이 사용되고 있는 프로토콜이다. 최종 사용자와 가맹점간의 지불 정보 보안에 관한 프로토콜이라고 할 수 있다[네이버 지식백과], (두산백과 두피디아, 두산백과).

40) ‘로봇배제 표준규약’이라고도 한다.

41) 크롤링을 거부하는 경우 ‘disallow’라고 기재하기도 한다.

42) 정현순, 앞의 논문, 377쪽.

기술적 보호조치나 자동수집제한표시도 없으며, 이용약관에는 특정한 방식의 이용만 제한하고 있으므로, 이 사건 서버에 대한 접근권한이 객관적 또는 명시적으로 제한되었다고 보기 어렵다. 따라서 피고인들이 통상적으로 이용되는 크롤링 프로그램을 이용하여 공개된 정보를 수집하기 위해 이 사건 서버에 접근한 행위를 피해자의 정보통신망을 침입하였다고 인정할 수 없으므로, 대상판결의 결론은 타당하다고 판단된다.

### 3. 저작권법위반의 점에 대한 검토

#### 가. 저작권법에 따른 데이터베이스와 데이터제작자의 권리

앞서 본 것과 같이 저작권법 제2조 제19호에서 데이터베이스의 정의를 규정하면서 데이터베이스가 되기 위해서는 단순히 소재를 모아놓은 것만으로는 부족하고, 소재를 체계적으로 배열 또는 구성하여야 한다. 다만 데이터베이스는 정형데이터(structured data)뿐만 아니라 비정형데이터(unstructured data)도 포함되는데<sup>43)</sup>, 비정형데이터인 경우 구조적으로 체계성이 있어야 보호대상이 될 것이다.<sup>44)</sup>

저작권법 제2조 제20호에 따라 데이터베이스제작자는 데이터베이스의 제작 또는 그 소재의 갱신·검증 또는 보충에 인적 또는 물적으로 상당한 투자를 한 자라고 규정하고 있고, 데이터베이스제작자는 그 데이터베이스의 전부 또는 상당한 부분을 복제·배포·방송 또는 전송할 권리를 가진다(저작권법 제93조 제1항). 다만 데이터베이스의 개별 소재 또는 그 상당한 부분에 이르지 못하는 부분의 복제 등에 대하여는 원칙적으로 데이터베이스제작자의 권리를 침해한 것은 아니나, 반복적이거나 특정한 목적을 위하여 체계적으로 함으로써 해당 데이터베이스의 일반적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에는 데이터베이스의 상당한 부분의 복제 등으로 보아 데이터베이스제작자의 권리를 침해한 것으로 본다(저작권법 제93조 제2항).

43) '정형데이터'는 미리 정해 놓은 형식과 구조에 따라 사용자가 쉽게 이행하고 시스템에 쉽게 적용될 수 있도록 잘 알려진 포맷이나 구조방법에 따라 저장되도록 구성된 데이터를 말하고, '비정형데이터'는 정의된 구조가 없이 정형화되지 않는 데이터를 말한다. '비정형데이터'의 예로는 오디오 파일, 동영상 파일, 사진, 보고서, 메일 본문 등이 있다.

44) 선종수, 앞의 논문, 329쪽

#### 나. 데이터베이스제작자 권리침해의 판단기준

저작권법의 규정에 따르면 어떠한 행위가 데이터제작자의 권리 침해를 하였는지를 판단하는 경우 먼저 데이터베이스 전부 또는 상당한 부분을 복제하였는지를 기준으로 검토하여야 한다. 그런데 ‘전부 또는 상당한 부분’이라는 개념은 불확정성을 가지고 있기 때문에 구체적으로 어떻게 판단하여야 할지는 법원에게 그 해석이 맡겨져 있다고 할 것이다.

또한 상당한 부분에 이르지 아니한 부분의 복제 등이라면 반복적이거나 특정한 목적을 위하여 체계적으로 복제 등을 하여 데이터베이스의 일반적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에 해당하는지를 검토하여야 한다.

#### 다. 상당한 부분에 대한 판단기준

데이터베이스제작자 권리침해의 판단기준으로 상당한 부분을 검토하는 경우 데이터베이스제작자의 권리성립요건인 투자의 상당성과 밀접하게 관련되어 권리 보호와 공정한 이용 사이의 균형 작용을 하게 된다.<sup>45)</sup>

데이터베이스제작자의 관점에서 침해행위 자체를 기준으로 판단하여야 하고, 추출되거나 재이용된 분량을 전체 데이터베이스 규모에 비추어 판단하여야 한다는 견해가 있고,<sup>46)</sup> 양적인 상당성만 아니라 질적인 상당성도 고려하여 데이터베이스 내용 중 핵심적인 정보를 지니고 있는지 여부, 가장 전략적이고 최신의 정보인지 여부, 데이터베이스 제작과정에서 문제된 부분을 수집·검증·표현하기 위하여 한 투자의 정도 등을 고려하여 판단하여야 한다는 견해가 있다.<sup>47)</sup>

종전에 사실심 법원은 데이터베이스 전부를 복제한 사안을 제외한 상당부분이 생략된 사안에서 정보의 수집이용을 중요하게 고려하면서 복제한 데이터의 양, 즉

45) 박익환, 데이터베이스 제작자의 법적 보호:현행 저작권법의 해석을 중심으로, 한국지식재산학회, 산업재산권 제15호, 269쪽

46) 한지영, 저작권법상 데이터베이스관련 규정의 해석에 관한 고찰, 대한변호사협회, 인권과 정의, 2006년 3월 제355호, 208쪽

47) 엄호준, 데이터베이스의 보호, 저작권법 누구를 위한 법인가, 서울대학교 기술과 법센터, 2006년, 31쪽

양적 상당성을 중요한 고려사항으로 보면서도 질적 상당성도 주요한 판단요소로 고려하여야 한다고 하였다.<sup>48)</sup>

유럽연합의 EU 데이터베이스 지침 제7조 제1호에서 그 보호 대상으로 ‘회원국은 데이터베이스의 내용의 취득, 검증 또는 표현에 대하여 양적으로 그리고/또는 질적으로 상당한 투자를 하였음을 나타내는 데이터베이스의 제작자에게 데이터베이스 내용의 전부 또는 약정으로 그리고/또는 질적으로 상당한 것으로 평가되는 부분의 추출 그리고/또는 재이용을 금지할 권리를 부여하여야 한다.’고 규정하고 있는데, 결국 우리나라 학설이나 사실심 판례와 같이 상당한 부분의 판단시 ‘양적 상당성’과 ‘질적 상당성’을 종합적으로 고려하여 판단하여야 한다는 것이다.<sup>49)</sup>

#### 라. 반복적·체계적 복제 등의 판단기준

‘반복적·체계적’이라는 의미는 어떤 계획에 따라 반복적으로 이루어지는 것을 의미하며, 임의적이거나 우발적인 이용행위를 배제하는 것을 의미하고, ‘통상적인 이용과 충돌’한다는 것은 데이터베이스의 개별 소재를 상당한 정도 복제 등의 방법으로 추출한 자가 이를 이용하여 제작한 상품·서비스가 시장에서 자신의 데이터베이스와 경쟁하는 상황에 놓인다거나 자신의 현재적이거나 잠재적 시장에 영향을 미칠 정도로 복제 등이 되는 경우를 의미한다.<sup>50)</sup> ‘이익을 부당하게 해치는 경우’는 제작자가 통상적으로 이용허락을 해서 얻을 수 있는 이익을 부당하게 상실하는 것을 의미한다고 한다.<sup>51)</sup>

종전 사실심 법원은 웹 크롤링 방법으로 데이터 등 정보를 수집한 경우 반복적·체계적 복제로 인정하였고, 쌍방이 경쟁관계에 있는 것을 전제로 권리 침해를 인정하였으며, 이용자들의 광고 노출 감소, 무료 사이트를 홍보하여 고객 유인, 판매수익 또는 광고 수익 감소 등의 사정을 들어 해당 데이터베이스의 통상적인 이용과 충돌하거나 데이터베이스제작자의 이익을 부당하게 해치는 경우에 해당된다고

48) 서울고등법원 2016. 5. 12. 선고 2015나20004441 판결(대법원에서 심리불속행기각되어 확정됨, 대법원 2016다223227 판결)

49) 박성호, 저작권법(제2판), 박영사(2017), 419쪽; 한지영, 데이터베이스의 법적 보호에 관한 연구, 서울대학교 법학박사학위논문(2005), 40쪽

50) 이해완, 저작권법(제4판), 박영사(2019), 1016쪽

51) 오승종, 저작권법(제5판), 박영사(2020), 1068쪽

인정하였다.<sup>52)</sup>

유럽연합의 EU 데이터베이스 지침 제7조 제5호는 ‘데이터베이스의 통상적인 이용과 충돌하거나 데이터베이스 제작자의 합법적인 이익을 부당하게 해치는 행위를 수반하는 데이터베이스 내용의 상당하지 않는 부분에 대한 반복적이고 체계적인 추출과 재이용은 허용되어서는 아니 된다.’고 규정하고 있어 우리 저작권법상의 관련 규정과 유사하여 그 논의를 참고할 필요가 있다고 보이는데, 유럽사법재판소도 위 조항의 목적이 사용자의 반복적 행위를 통해 데이터베이스의 전부 또는 상당한 부분을 재구성하는 것을 방지하기 위한 것이라고 보았고, 상당하지 아니한 부분을 반복적으로 재이용하는 경우는 여기에 해당하지 않는다고 보았다.<sup>53)</sup>

#### 마. 검토

피고인들이 크롤링 프로그램을 통하여 264회에 걸쳐 피해자의 데이터를 복제하였는데, 피고인들이 1회 접속시 3-9개 항목에 관한 정보를 수집하였을 뿐이어서 전체 데이터베이스가 50여 개의 항목으로 구성되어 있다는 점에서 1회 복제만으로 저작권법 제93조 제1항에 정한 전부 또는 상당한 부분의 복제에 해당하지는 아니한다. 그러므로 저작권법 제93조 제2항에 따라 피고인들이 반복적·체계적으로 복제한 데이터 전체가 ‘양적 상당성’과 ‘질적 상당성’을 종합적으로 고려하여 상당한 복제에 해당하는지를 검토하여야 할 것이다.

먼저 ‘양적 상당성’이 있는지에 관하여 살펴보면, 양적 상당성은 전체 데이터베이스의 규모를 고려하여 판단하여야 하는데, 항소심법원은 피고인들이 복제한 정보의 양은 전체 데이터베이스의 일부에 불과한 점, 전체 데이터베이스에서 차지하는 비율이 높다고 보기 어려운 점을 고려하면 양적 상당성이 있다고 단정하기 어렵다고 판단하였는데, 그러한 사정을 고려하면 위 결론은 타당하다고 할 것이다.<sup>54)</sup>

다음으로 ‘질적 상당성’이 있는지에 관하여 살펴보면, 데이터베이스제작자

52) 서울고등법원 2021. 12. 9. 선고 2020나2042706 판결(확정); 서울중앙지방법원 2020. 7. 9. 선고 2018가합528464 판결(확정); 서울중앙지방법원 2018. 6. 27. 선고 2017가합556332 판결; 서울고등법원 2010. 6. 9. 선고 2009나96306 판결(확정)

53) 한지영, 앞의 논문, 서울대학교 법학박사학위논문(2005), 129쪽

54) 선종수, 앞의 논문, 333쪽

의 권리 보호는 데이터베이스의 제작 등에 대한 투자를 보호하기 위한 것이므로 개별 소재의 가치나 개별 소재 생산에 투입된 투자가 아닌 데이터베이스 자체 투자의 정도를 고려하여 피고인들이 복제한 데이터베이스 부분이 질적인 상당성이 있는지를 판단하여야 한다. 그런데, 피고인들이 복제한 정보는 숙박 관련 정보로 공개된 것으로 피해자도 공개된 사이트를 접속하여 취득하였거나 정보 개별 소재의 생산과 구별되는 소재들의 수집, 검증 또는 표현에 상당한 투자를 하였다고 볼 만한 주장이나 자료가 없어 피고인들이 수집한 데이터가 피해자가 질적으로 상당히 투자한 부분이라고 보기 어렵다고 할 것이다. 또한 피고인들의 데이터 복제행위로 인하여 직접적으로 고객 유인 효과를 발생하게 하거나 이용허락을 통한 이익을 상실하게 하는 등으로 피해자의 데이터베이스의 통상적인 이용과 충돌하거나 피해자의 이익을 부당하게 해치는 경우에 해당한다고 보기도 어렵다.<sup>55)</sup>

따라서 피고인들의 행위는 피해자의 데이터베이스 제작자의 권리를 침해하였다고 인정하기 어렵다고 할 것이고, 이와 같은 결론의 항소심법원의 판단을 수긍한 대법원의 입장은 타당하다고 할 것이다.

#### 4. 컴퓨터등장애업무방해의 점에 대한 검토

##### 가. 일반론

대법원에서 문제되는 부분은 형법 제314조 제2항 후단으로 구성요건으로 ‘허위의 정보 또는 부정한 명령을 입력하거나 기타 방법으로’와 ‘정보처리에 장애를 발생하게 하여 업무를 방해’한 것이므로, 각 요건에 대한 검토가 필요하다.

##### 나. 허위의 정보 또는 부정한 명령의 입력

‘허위의 정보 입력’은 사실관계에 일치하지 않는 자료를 입력한 것을 말하고, ‘부정한 명령 입력’은 프로그램을 조작하는 경우로 프로그램을 구성하는 개개의 명령을 부정하게 변경, 삭제, 추가하거나 프로그램을 변경하는 것을 의미한다.<sup>56)</sup>

55) 선종수, 앞의 논문, 334쪽

56) 대법원 2006. 3. 10. 선고 2005도382판결; 대법원 2009. 4. 9. 선고 200811978 판결 등 참조; 김

다. 정보처리에 장애를 발생하게 하여 업무를 방해

‘정보처리의 장애’란 가해행위 결과 정보처리장치가 그 사용목적에 부합하는 기능을 하지 못하거나 사용목적과 다른 기능을 하는 것을 의미하고, 정보처리의 장애가 현실적으로 발생할 것을 필요로 하지만, 장애가 발생하여 업무방해의 결과를 초래할 위험이 발생한 이상, 업무방해의 결과가 실제로 발생하지 않더라도 업무방해죄가 성립한다고 본다.<sup>57)</sup>

라. 검토

1) 피고인들이 부정한 명령을 입력한 것인지에 대한 판단

공소사실에서 본 것과 같이 피고인들이 피해자의 이 사건 앱을 통하지 않고 이 사건 서버에 직접 명령을 입력한 행위와 이 사건 앱을 통한 숙박업소 검색 범위(현 위치 기준 7km ~ 30km)를 벗어나 크롤링 프로그램을 통하여 1,000km 이내 숙박업소 정보를 검색한 행위가 권한 없는 행위로 ‘부정한 명령’을 입력한 행위로 볼 수 있는지를 검토하여야 한다.

피해자는 이 사건 서버에 명령 입력 범위를 제한하지 아니한 채 이 사건 앱에서만 검색범위의 일부 제한만 설정하였을 뿐이고, 이 사건 서버는 명령 입력 범위에 제한 없이 기본적으로 주어진 명령에 대응하여 숙박업소 정보를 제공하는 것을 목적으로 하고 있었으므로, 피고인들이 입력한 명령문이 이 사건 앱에서 설정한 제한 범위를 확대하여 검색하는 것이라고 하여 동일한 제한이 없는 이 사건 서버에 본래 목적과 상이한 부정한 명령을 입력하는 것이라고 보기 어렵다.<sup>58)</sup> 즉 웹 크롤링이 검색 및 정보 수집의 효율성을 위하여 활용되고 있는 점을 고려하면 위와 같은 제한이 없는 이 사건 서버에 이 사건 앱이 설정한 검색 범위를 확장하여 입력한 것을 부정한 명령이라고 볼 수 없다고 할 것이다.

총전에 컴퓨터등장애업무방해죄가 인정된 대법원 판례들을 살펴보면, 권한

---

태균, 앞의 논문, 298쪽

57) 대법원 2009. 4. 9. 선고 2008도11978 판결

58) 김태균, 앞의 논문, 299쪽

없는 자가 아이디와 비밀번호를 무단으로 변경한 사안, 허위의 클릭정보를 전송한 사안, 권한 없는 자가 임의로 컴퓨터에 비밀번호를 설정한 사안으로 권한이 없는 자가 명백하게 허위의 정보를 전송하거나 아이디, 비밀번호를 무단으로 입력하거나 설정한 경우로 정보통신망 침입의 접근권한이 없으면서 행위한 것으로 객관적인 사정에 비추어 부정한 것으로 충분히 볼 수 있다.<sup>59)</sup>

## 2) 정보처리 장애의 현실적 발생하였는지에 대한 판단

검사는 피고인들의 대량 정보 수집을 위한 명령 입력으로 이 사건 서버가 본래의 목적과 달리 정보 탈취 목적으로 사용되었으므로, 그러한 작동 자체로 이 사건 서버의 정보처리에 장애가 현실적으로 발생하였고, 최대 초당 90회 정도의 접속이 이루어져 이 사건 서버의 트래픽이 상당히 증가하는 등으로 정보처리에 장애가 발생하였다고 주장하였다.

그러나 피고인들은 이 사건 서버의 목적인 숙박업소의 정보 수집을 위하여 검색을 하면서 웹 크롤링 프로그램을 사용하여 명령을 입력한 것은 이 사건 서버의 목적에 따라 기능하였다고 볼 수 있고, 피고인들의 명령 입력으로 이 사건 서버의 트래픽이 증가하였다고 하더라도 이 사건 서버가 사용목적에 부합하는 기능을 하지 못하거나 다른 기능을 하는 등으로 정보처리에 장애가 발생한 것이 아니라면, 단순 트래픽 증가만으로 정보처리에 장애가 현실적으로 발생하지 아니하였다고 보기 어렵다고 판단된다.

## 3) 소결

따라서 컴퓨터등장애업무방해하였다는 점을 인정할 수 없다고 할 것이고, 항소심법원의 판단을 그대로 인정한 대상판결의 결론은 타당하다고 판단된다.

## VII. 결론

---

59) 김태균, 앞의 논문 300쪽

대상판결은 웹 크롤링을 이용한 데이터 수집행위에 대한 형사책임에 대한 명시적인 판단을 한 최초의 사례로서 선례적인 의미가 있다. 대상판결에서는 쟁점이 된 정보통신망침해, 저작권법상 데이터베이스저작권침해, 형법상 컴퓨터등장애업무방해침해에 대하여 항소심법원의 결론과 같이 모두 무죄를 확정하였는데, 웹 크롤링을 이용한 데이터 활용의 중요성이 증가함에 따라, 데이터의 저작권 보호와 데이터 활용이라는 양쪽의 이익을 조화롭게 보호하여야 할 필요성이 있는 상황에서 대상판결의 결론은 데이터의 보호의 필요성만으로 웹 크롤링과 관련된 형사법적 불확실성이 어느 정도 해소되었다고 할 것이다.

다만 모든 웹 크롤링을 통한 데이터 수집행위가 형사책임이 없다는 것은 아니고, 앞서 본 여러 사정을 고려하여 형사책임의 결론이 판단되어야 할 것이므로, 구체적인 사안의 사실관계에 따라 형사책임이 인정되는 경우도 있을 것이다. 그러나 헌법상 과잉금지원칙과 형사처벌의 보충성, 불확정성의 해소라는 면을 고려하면, 위와 같은 문제에 있어서 형사법적인 해결방안보다는 우선 민사법적인 해결방법이 있는지를 검토하는 것이 바람직하다고 보인다.<sup>60)</sup>

앞으로도 인공지능, 빅데이터 등 새로운 기술과 결합된 웹 크롤링 기술의 등장에 따른 다양한 법적인 문제가 발생할 것이다. 따라서 저작권자 등 권리의 보호와 충돌하는 경우가 많을 것이므로, 데이터 활용을 위한 산업 발달의 측면과 데이터 보호라는 측면을 조화롭게 고려한다면 형사책임이 인정되는지 문제뿐만 아니라 다른 법적 해결방법이 있는지, 더 나아가 입법적인 해결방안에 대하여도 고민하고 연구하여 볼 필요성이 있다고 할 것이다.

---

60) 대상판결의 사안에 대한 민사판결에서 피고인 회사의 부정경쟁방지법위반으로 불법행위로 인한 손해 배상책임이 인정되었다(서울고등법원 2021나203474 판결, 상고하지 아니하여 확정됨).