생성형 인공지능 서비스의 개인정보 국외이전 및 제3자 제공 위반 사례 분석: 딥시크를 중심으로

신 민 준[1], 엄 익 채[2]
[1]전남대학교 (대학원생), [2]전남대학교 (지도교수)
jutlsgood@naver.com, iceuom@chonnam.ac.kr

Analysis of Violations in Overseas Transfer and Third-Party Provision of Personal Data by Generative AI Services: A Case Study of DeepSeek

Min-jun Shin[1], Ieck-Chae Euom[2]
[1]Chonnam National University (Graduate Student), [2]Chonnam National University (Advising Professor)

요 약

본 연구는 생성형 인공지능 서비스의 개인정보 보호법 위반 사례를 분석하기 위해, 딥시크(DeepSeek)의 웹 서비스를 대상으로 기술적 접근을 수행하였다. 클라이언트-서버 간의 네트워크 트래픽 분석을 통해, 사용자 입력 정보와 시스템 로그가 사용자의 동의 없이 외부 서버로 전송되고 있음을 확인하였다. 또한 개인정보 처리방침이 한국어로 제공되지 않았으며, 국외이전 및 제3자 제공 관련 고지와 동의 절차가 누락되어 있었다. 이와 같은 사항은 「개인정보 보호법」 제28조의8, 제30조 및 개인정보 보호 원칙에 위반되는 정황으로 판단된다. 본 연구는 딥시크 사례를 통해 생성형 인공지능 서비스에서 발생 가능한 개인정보 침해 유형을 실증적으로 제시하고, 이에 대응하기 위한 규제 개선 방향을 제안한다.

1. 서론

1.1 연구 배경과 문제 제기

최근 생성형 인공지능(Generative AI) 서비스의 확산과 함께 개인정보 국외 이전 및 제3자 제공과 관련한 법적 문제가 부각 되고 있다. 2025년 2월, 개인정보보호위원회는 개인정보보호법 위반을 이유로 딥시크(DeepSeek)의 한국 내 서비스 중단을 명령하였다. [1] 본 연구는 딥시크 사례를 통해 생성형 인공지능 서비스가 개인정보 보호법상 규제를 어떻게 위반할 수 있는지 분석하고자 한다.

1.2 연구 목적 및 방법

본 연구는 딥시크의 웹 서비스를 대상으로 실제 서비스를 직접 이용하고, 클라이언트 단말기에서 서버로 전송되는 네트워크 트래픽을 분석하는 방법으로 진행되었다. 이를 통해 개인정보 보호법상의 규정과 대조하여 법적 준수 여부를 평가하였다. 본 논문은 이러한 기술적 분석 결과를 바탕으로 딥시크의 개인정보보호법 위반 사항을 구체적으로 검토하고, 생성형 인공지능 서비스에 대한 개인정보보호 규제 방향을 제시하는 데 목적이 있다.

2. 개인정보보호법상 국외이전 및 제3자 제공 규제

본 장에서는 「개인정보 보호법」에 따라 개인정보의 국외이전에 필요한 요건과 제3자 제공 요건을 검토한다. 또한, 생성형 인공지능 서비스가 개인정보를 처리 함에 있어 발생할 수 있는 주요 법적 쟁점에 대해 분석하고, 이를 통해 향후 서비스 운영 시 유의해야 할 규제 사항을 도출하고자 한다.

2.1 국외이전 요건

「개인정보 보호법」은 개인정보의 국외이전에 대하여 엄격한 규제를 부과하고 있다. 법 제28조의8 제1항에 따르면, 개인정보를 국외로 이전하고자 하는 경우 원칙적으로 정보주체로부터 별도의 동의를 받아야 한다. 또한 개인정보를 국외로 이전 할 때에는 이전되는 개인정보의 항목, 이전 국가, 이전 시기 및 방법, 이용 목적, 보유·이용 기간 등을 정보주체에게 투명하게 고지하여야 한다. 아울러 개인정보 침해가 발생할 우려가 있는 경우, 개인정보보호위원회는 국외이전의 중지를 명령할 수 있다. [2]

2.2 제3자 제공 요건

개인정보를 제3자에게 제공하는 경우에도 「개인정보 보호법」 제17조 및 제18조에 따라 엄격한 요건을 충족하여야 한다. 기본적으로 개인정보를 제3자에게 제공하려면 정보주체로부터 사전 동의를 받아야 하며, 이때 제공 목적, 제공 항목, 제공받는 자, 보유 및 이용 기간 등을 명확히 고지하여야 한다. 또한 제공 과정에서는 개인정보의 안전성 확보조치를 취하여야 하며, 제공 내역을 기록·관리하고, 목적 외 이용이나 추가 제공을 방지하여야 한다. [2]

2.3 생성형 인공지능 서비스의 법적 쟁점

생성형 인공지능 서비스는 그 특성상 대규모 데이터 수집 및 처리가 필수적이며, 이로 인해 「개인정보 보호법」과의 충돌 가능성이 빈번히 제기된다.

다음 <표 1>은 생성형 인공지능 서비스에서 논의되고 있는 주요 개인정보 보호 쟁점과 관련된 법적 고려사항을 정리한 것이다.

No	주요 쟁점	법적 고려사항 또는 위험 요인
1	정보주체 동의 없이 대규모 개인정보 수집 [3]	동의 없는 수집은 「개인정보 보호법」 제15조 및 제17조 위반 소지
2	가명처리 정보의 재식별 가능성 [4]	재식별 가능 시 가명정보가 개인정보로 간주되며, 안전조치 미흡 시 법 위반
3	제3자 또는 해외 서버로의 정보 전송 [4]	고지·동의 없는 제3자 제공은 제18조, 국외이전은 제28조의8 위반 가능성
4	암호화, 접근통제 등 기술적 보호조치 부족 [4]	기술적 보호조치 미흡은 제29조 위반이며, 유출 시 처벌 강화 요인 됨

<표 1> 생성형 인공지능 서비스의 개인정보 보호 관련 주요 쟁점 및 법적 고려사항

최근 개인정보보호위원회가 딥시크의 개인정보 국외이전 및 제3자 제공 위반을 이유로 한국 내 서비스 중단을 명령한 사례는, 생성형 인공지능 서비스가 개인정보 보호법의 규제 틀을 철저히 준수해야 함을 보여주는 대표적 사례로 평가된다. 이에 본 연구는 딥시크 사례를 중심으로 해당 법적 쟁점을 심층적으로 분석하고자 한다.

3. 딥시크 사례 분석

본 장에서는 딥시크(DeepSeek) 서비스의 「개인정보 보호법」 위반 사례에 대한 기술적 분석 결과를 제시한다. 웹(Web) 서비스에 대한 분석은 Burp Suite를 활용하여 클라이언트와 서버 간의 통신 트래픽을 수집하고, 이를 기반으로 개인정보 처리 흐름을 추적하는 방식으로 수행되었다. 이러한 분석 환경을 바탕으로, 본 장에서는 딥시크 서비스에서 이루어진 개인정보 수집, 국외이전 및 제3자 제공 과정 전반에 대해 기술적 관점에서 위반 여부를 상세히 검토하고자 한다.

3.1 딥시크 웹 서비스 분석

딥시크 웹 서비스에 대한 분석 결과, (그림 1) 및 <표 2>과 같이 클라이언트 측 스크립트인 JavaScript를 통해 사용자 시스템 로그가 약 60초 간격으로 gator.volces.com 서버로 전송되는 것을 확인할 수 있었다. (※ 이하의 분석 내용은 2025년 2월 기준의 서비스 버전을 대상으로 수행되었다.)

또한, 사용자가 특정 프롬프트를 입력할 경우, 입력된 검색 키워드가 시스템 로그에 포함되어 함께 전송되는 행위도 관찰되었다.

(그림 1) gator.volces.com request 메시지

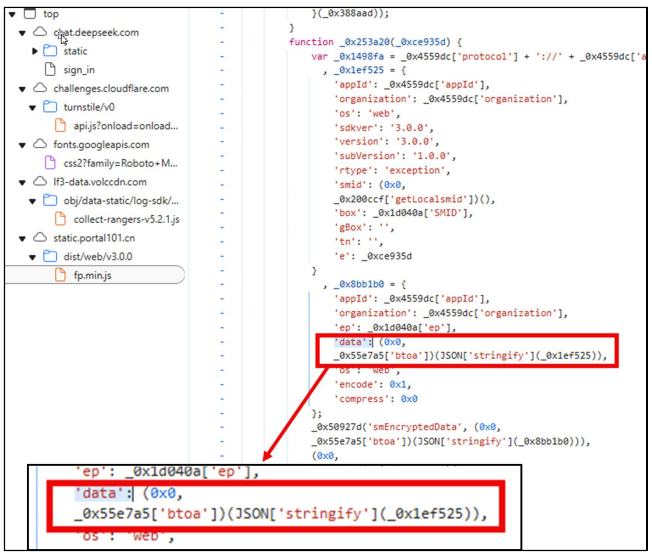
변수명	변수값	개인정보 여부
os_name	운영체제 (Windows)	결합정보
os_version	OS 버전 (10)	
device_model	기기 정보 (Windows NT 10.0)	
browser	사용 브라우저 (Microsoft Edge)	
browser_version	브라우저 버전 (132.0.0.0)	
resolution	모니터 해상도 (2560x1080)	
timezone	UTC+9 (한국 시간)	
z_offset	-32400 (초 단위 오프셋, -9시간)	
referrer	이전에 방문한 사이트	
$latest_search_keyword	사용자가 검색한 키워드 (ABCDTEST)	개인정보
origin_referrer	최초 방문한 페이지	결합정보
commit_id	버전 관리용 ID (75eaf5a9)	

<표 2> gator.volces.com request 메시지 주요 내용

다음으로, 딥시크 웹 서비스 이용 시 (그림 2)와 같이 fp-it-acc.portal101.cn 서버로 암호화된 데이터가 전송되는 것이 확인되었다.

(그림 2) fp-it-acc.portal101.cn/deviceprofile/v4 request 메시지

이 암호화된 데이터는 (그림 3)과 같이 static.portal101.cn/dist/web/v3.0.0/fp.min.js 파일을 통해 생성되는 것으로 분석되었으며,



(그림 3) static.portal101.cn/dist/web/v3.0.0/fp.min.js의 data 변수값 생성 코드

해당 파일 내 function(_0x127ee5, _0x4d7c2f, _0x35edd1) 함수에서 DES 알고리즘을 이용한 암호화 로직이 구현되어 있는 것이 확인되었다.

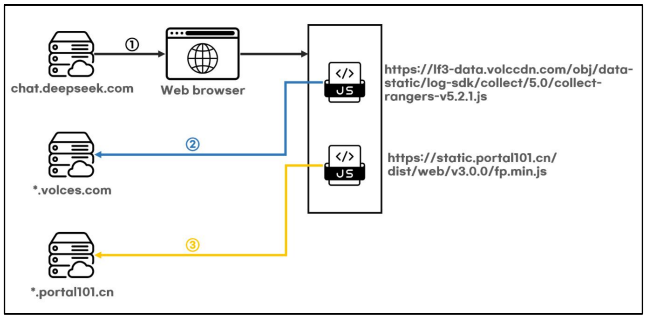
특히, 해당 함수에는 DES 암호화에 사용되는 Key 값은 <표 3>와 같이 명시되어 있으며, 이를 통해 생성된 암호화 데이터가 (그림 1)에서 제시된 바와 유사한 형태의 사용자 시스템 로그를 포함하고 있음을 확인할 수 있었다.

인자값	Key 값	난독화 변수명	인자값 내용(추정)
appId	q8zb1bs1	vw	애플리케이션 ID
box	wncbzi1u	fu	추정 불가
canvas	qy1d6fmu	mi	브라우저 정보
clientSize	wqmvv9z8	vq	브라우저 창 크기
organization	ht432iov	tj	사용자 국가
os	i7g8jtql	vh	운영체제 정보
platform	mcbmtg5y	ye	플랫폼 정보
plugins	j916a142	vc	추정 불가
pmf	d35cb8mv	jw	추정 불가
protocol	-	protocol	현재 페이지 프로토콜
referer	rf149ntc	hv	직전 방문 페이지
res	rbtd2cl6	pi	디스플레이 해상도
sdkver	rw74ssux	xh	SDK 버전
status	4xx5u0ww	uu	추정 불가
subVersion	9wg0vhb9	xi	서브 버전 정보
svm	mqcala0h	nu	
time	8e224y3f	gw	현재 시간
timezone	3w9lg8pr	oi	시간대 정보
rtype	rx7ob4gc	na	요청 유형
tn	ilnts67v	nb	추정 불가
trees	al8x9zt7	hc	추정 불가
ua	z5jc9qyp	wr	추정 불가
url	6tbl1wpw	no	현재 접속한 URL
vpw	xzvnu5jt	st	뷰포트 너비

<표 3> function(_0x127ee5, _0x4d7c2f, _0x35edd1) 함수 분석 정보(추정)

또한, function(_0x127ee5, _0x4d7c2f, _0x35edd1) 함수의 구조와 <표 3>에 제시된 Key 값이 별도의 난독화나 보호 조치 없이 노출되어 있는 점에 비추어볼 때, 해당 암호화는 정보의 기밀성 확보보다는 사용자가 프록시 도구 등을 활용하여 데이터를 임의로 변조하는 행위를 방지하기 위한 용도로 구현된 것으로 판단된다.

위와 같은 분석 결과를 종합해보면, (그림 4)와 같은 구조로 이용자의 로그를 전달하며, <표 4>과 같이 딥시크 자체 서버 외에도 최소 두 개 이상의 외부 서버에 사용자 정보가 전송되고 있는 사실을 확인할 수 있었다.



(그림 4) 딥시크 웹 서비스의 사용자 로그 수집 및 외부 서버 전송 구조도

URL	IP주소	국가	owner Name
www.deepseek.com	104.18.26.90	USA	cloudflare, Inc.
chat.deepseek.com	104.18.27.90	USA	cloudflare, Inc.
csp-reporting.cloudflare.com	104.18.20.157	USA	cloudflare, Inc.
nexus-websocket-a.intercom.io	35.174.127.31	USA	Amazon Technologies Inc.
*.volces.com ※ ByteDance 사의 인과관계 추론 SaaS 서비스	47.246.29.251	USA	Alibaba Cloud LLC
*.portal101.cn ※ 수메이 테크놀로지 사의 SaaS기반 AI위협 탐지 플랫폼	43.155.149.X	Korea	Tencen ※ 텐센트 클라우드 한국리전 추정

<표 4> 딥시크 서비스를 통한 통신서버 정리

3.2 법 쟁점 사항 평가

딥시크(DeepSeek)는 2025년 2월 당시 서비스 약관 및 개인정보 처리방침에 서비스 제공 지역을 명시하지 않았으나, 국내 앱 마켓을 통한 공식 출시(2025년 2월 기준, 한국 앱 마켓 다운로드 순위 1위), 한국어 기반의 서비스 제공, 한국어 데이터 학습 등의 정황을 종합할 때 「개인정보 보호법」의 적용 대상에 해당한다고 판단할 수 있다.

딥시크는 중국어, 영어, 일본어 버전의 개인정보 처리방침은 제공하고 있었으나, 한국어 처리방침은 별도로 수립하거나 공개하지 않았다.

또한 당시 공개된 처리방침에는 개인정보의 파기 절차 및 방법, 개인정보 보호책임자의 성명과 연락처 등 「개인정보 보호법」 제30조 제1항에서 규정한 주요 항목들이 누락되어 있었다.

이와 함께, 딥시크는 이용자가 프롬프트에 입력한 정보를 중국 소재의 volces.com 서버로 전송하면서도 이에 대한 사전 동의를 받지 않았고, 관련 내용 역시 처리방침에 공개하지 않았다.

뿐만 아니라, volces.com을 포함한 중국 및 미국 내 다수의 외부 서버로 이용자 정보가 전송되고 있음에도 불구하고, 해당 사실을 명시적으로 고지하지 않았으며, 이에 대한 동의도 확보하지 않은 것으로 확인되었다.

4. 시사점 및 결론

본 연구에서는 생성형 인공지능 서비스인 딥시크(DeepSeek)를 사례로, 웹 서비스를 대상으로 한 기술적 분석을 통해 개인정보 수집·이용·전송 방식의 법적 쟁점을 검토하였다. 특히 「개인정보 보호법」상 국외이전 및 제3자 제공 요건을 중심으로 위반 가능성을 분석한 결과, 다수의 개인정보 관련 법령 위반 정황이 확인되었다.

먼저, 딥시크는 한국어 기반으로 서비스를 제공하고 있음에도 불구하고, 한국어 처리방침을 제공하지 않거나, 법령에서 요구하는 필수 고지 항목들을 누락한 점에서 이용자 알 권리를 침해하였다. 또한, 수집된 개인정보를 중국 및 미국 등 해외 서버에 전송하면서도 정보주체의 동의를 받지 않았고, 해당 사실을 공개하지 않아 국외이전 관련 법적 요건을 충족하지 못한 것으로 판단된다. 이러한 행위는 개인정보의 국외이전(제28조의8) 및 처리방침 수립·공개 의무(제30조 제1항) 위반 소지가 있으며, 개인정보 보호 원칙(제3조)에도 저촉된다.

이러한 사례는 생성형 인공지능 서비스가 기존의 웹 기반 서비스보다 훨씬 더 복잡하고 고도화된 데이터 수집 및 활용 방식을 사용함에 따라, 기존 개인정보보호 체계로는 충분히 대응하기 어려운 지점이 존재함을 시사한다. 이에 따라 다음과 같은 정책적·제도적 시사점을 제안한다.

첫째, 생성형 인공지능 서비스에 특화된 개인정보 보호 가이드라인 또는 기술 기준을 마련할 필요가 있다. 예를 들어, 사용자 입력 데이터의 자동 수집 행위에 대한 명확한 법적 기준 및 고지 방식이 필요하다.

둘째, 국외이전 대상 서버의 위치·용도·수탁자 정보를 명확하게 기술하도록 하는 처리방침 작성 기준의 강화가 요구된다. 특히 SaaS 기반 API 연동 및 추적형 스크립트를 사용하는 경우에도 국외이전 여부를 명시적으로 판단할 수 있는 해석기준이 필요하다.

셋째, 이용자의 이해 가능성과 접근성을 높이기 위한 다국어 처리방침 제공 기준이 마련되어야 한다. 단순히 영문 처리방침을 제공하는 수준을 넘어서, 서비스 제공 국가의 언어로 된 정식 처리방침 수립·공개가 필수적으로 요구되어야 한다.

결론적으로, 생성형 인공지능 서비스는 새로운 데이터 활용의 가능성을 제시하는 동시에, 기존 개인정보 보호 체계의 한계를 시험하고 있다. 딥시크 사례는 이러한 변화를 보여주는 대표적 사례로서, 향후 AI 서비스의 확산에 따라 개인정보 보호정책의 정교화와 국제적 조화가 더욱 중요해질 것이다.

참고문헌

[1] 개인정보보호위원회. "딥시크 앱 국내 서비스, 잠정 중단 후 개선·보완키로", 보도참고자료: 2025.02.17

[2] 개인정보보호위원회. “「개인정보 보호법」” 법률 제19234호(시행 2025.03.13)

[3] 안정민 (2024). "AI 가 불러온 글로벌 데이터 규제의 변화." 법학논집 29(1): 331-363.

[4] 개인정보보호위원회. (2024). “가명정보 처리 가이드라인” 발간등록번호: 11-1790365-000029-01