

1. 딥시크(DeepSeek)의 정의와 역할

- **정의**
딥시크(DeepSeek)는 대규모 인공지능 모델 학습과 자연어 처리(NLP)에 특화된 AI 플랫폼 또는 모델을 의미하며, 방대한 데이터셋을 기반으로 텍스트 분석, 생성, 예측 등을 수행한다.
- **역할**
 - 대규모 데이터 분석 및 의미 추출
 - 자연어 질의응답, 요약, 번역 등 언어 이해 및 생성
 - 기업·기관의 데이터 기반 의사결정 지원
 - 다양한 AI 서비스(챗봇, 검색, 추천 시스템 등)의 핵심 엔진 역할

2. 딥시크 관련 개인정보보호 문제점

1. 데이터 수집 단계
 - 비식별화가 충분하지 않은 상태에서 개인정보가 포함될 수 있음
 - 사용자의 동의 절차가 불분명하거나 형식적일 가능성
- 데이터 저장 단계
 - 암호화 미흡으로 인한 데이터 유출 위험
 - 접근권한 관리 부실로 인한 내부자 유출 가능성
- 데이터 학습 및 활용 단계
 - 학습 데이터에 포함된 개인정보가 모델 출력에 재노출될 가능성(모델 반출 공격, Membership Inference Attack)
 - 데이터 오염(데이터 포이즈닝)에 따른 잘못된 모델 동작
- 제3자 제공
 - 외부 API, 클라우드 환경에서의 전송 시 암호화 미흡
 - 위탁업체와의 계약서 내 개인정보보호 조항 미흡

3. 보안 위협 대응 방안

1. 수집·저장 단계
 - 개인정보 최소 수집 원칙 적용
 - 저장 시 AES-256 등 강력한 암호화 사용
 - 접근제어(권한분리, MFA) 철저히 시행
- 학습·처리 단계
 - 데이터 비식별화(마스킹, 가명처리) 후 학습
 - Differential Privacy, Federated Learning 등 개인정보 보호 학습 기법 적용
 - 데이터 포이즈닝 방지 위한 데이터 무결성 검증
- 전송·공유 단계
 - TLS 1.3 등 안전한 통신 프로토콜 사용
 - 제3자와의 계약 시 개인정보보호 책임 범위 명확화
- 운영·모니터링 단계
 - 로그·모델 모니터링을 통한 비정상 접근 탐지
 - 보안 사고 대응 프로세스 수립 및 모의훈련 수행