제47호 2023.06

GLOBAL TREND

세계지방자치동향

한국지방행정연구원
Korea Research Institute for Local Administration

유럽연합의 챗GPT 규제와 독일 지방행정에서의 도입 시도

개요

- 챗GPT는 LLM(Large Language Model, 거대언어모델)의 일종으로 GPT-3.5버전 출시 이후 선풍적인 인기를 끌어 이용자가 기하급수적으로 늘어나고 있으며, 2023년 4월 출시된 GPT-4버전에서는 그 성능이 대폭 개선되어옴
- 초기에는 챗GPT가 미국 변호사 자격시험, 의사 자격시험 등에서 합격 수준의 답변을 보이며 이슈화되었으며, 대부분의 나라에서 다양한 분야의 의문 사항에 답변을 주는 도구로 여겨지고 있음
- 그러나 유럽 지역에서는 이러한 열풍에 비해 개인정보 유출, 정확하지 않은 정보 제공으로 인한 비판적 시각이 다분하며, 유럽연합 차원에서 규제안을 준비 중이고, 이탈리아 정부에서는 이미 개인정보유출 등 관련 법규 위반 등으로 챗GPT 사용을 일시적으로 금지하는 등의 문제가 발생하기도 함
- 본 원고는 챗GPT에 대한 유럽 지역 AI기술 규제 및 행정도입 사례 등을 소개하고 그 시사점을 도출하고자 함

이탈리아의 챗GPT 접속 차단 조치 및 복구

- 이탈리아의 독립기관인 데이터보호감독청(GPDP: Garante per la protezione dei dati personali)은 유럽 데이터보호규정(GDPR: General Data Protection Regulation)에 근거하여 챗GPT에 대한 접속 차단 조치를 시행하였음(2023. 3. 30.)
- 발단된 사건은 챗GPT가 개인정보를 유출한 것으로, 이에 데이터보호감독청에서 별도의 조사가 시작되었음
- 공식적인 차단 사유로는 ①정보제공의 의무 위반 ②데이터 처리의 적법성 부족 ③데이터의 부정확성 ④연령 인증의 부재 등이 포함되었음
- 정보제공의 위반 사유는 챗GPT가 데이터 보호규정에 따라 개인 데이터 수집 시 통보 의무가 있으나, 이를 이행하지 않은 것이었음

- 데이터 처리의 적법성 문제는 데이터보호규정에 따라 모든 개인 데이터 처리는 명시적으로 처리되고 이것이 규정에 따라 정당화되어야 하는데, 챗GPT 운영의 기반이 되는 알고리즘 훈련 목적으로 개인 데이터를 수집하는 것은 부적절하다는 것이었음
- 데이터의 부정확성 문제 또한 데이터보호규정에 따라 가능한 한 정확하고 최신의 데이터를 제공해야 한다는 조항이 있는데, 챗GPT의 기술적 한계로 부정확한 답변을 내놓는 일명 '환각(Hallucination)증상'이나 데이터 최신성 부족이 사유였음
- 또한 13세 미만의 아동이 이용하기에는 아동의 발달적 측면에서 부정적인 영향을 줄 수 있다는 우려가 있었음
- 그러나 1개월 후인 2023년 4월 30일에 이탈리아 데이터보호감독청과 OpenAI 측의 원만한 협의와 사측의 일부 위반사항에 대한 수정조치 등 일정한 조건에 따라 차단을 해제한 바 있음

유럽연합의 데이터보호규정과 AI 규제안 준비

- 이탈리아의 차단 사례는 앞서 언급한 바와 같이 유럽연합의 강력한 데이터보호규정에 근거하고 있음
- 유럽연합의 데이터보호규정은 공공부문과 민간부문을 구분하지 않고 모든 유럽연합 국가에서 수행되는 개인정보 및 데이터 처리의 기본 규정으로 작동함
- 데이터보호규정은 사생활존중권과 개인데이터보호권을 보장하기 위해 2016년에 제정됨
- 데이터보호규정이 강력한 규제로서 작동하는 이유는 단순히 개인의 권리보장만을 위한 것이 아니며, 국제법 규범과 2차대전 이후의 유럽인권조약, IT기술의 발전으로 등장한 유럽평의회 조약 108(Council of Europe Convention 108) 등 다양한 규범을 준수하여 궁극적으로는 인간의 기본권을 보장하기 위함임
- 유럽연합 데이터보호규정의 주요한 내용은 아래의 표와 같음

표 1. 데이터보호규정의 주요 내용

조항	주요 내용
잊혀질 권리	정보 주체가 개인정보 처리를 더 이상 원하지 않거나 처리의 법적 근거가 존재하지 않으면 해당 정보를 삭제해야 함
본인 개인정보의 열람권 및 이용권	본인의 개인정보가 어떻게 처리되는지에 대한 정보를 얻을 수 있어야 하며, 명확하고 이해하기 쉬운 방식으로 접근이 가능해야 함
본인 개인정보를 이전할 권리	개인의 동의로 서비스 공급자 간 개인정보 이전을 별도의 절차 없이 가능하게 함
정보 유출을 고지받을 권리	정보 유출 사고가 발생했을 때 정보를 보유한 민간기업 또는 조직은 국가의 감독 기관에 보고하고, 유출 피해자에게도 최대한 신속하게 통보해야 함
설계 및 기본값에 의한 개인정보보호	유럽연합 지역에서 유통되는 제품 또는 서비스의 설계, 기본값 설정부터 개인정보보호에 대한 안전장치를 고려하여야 함
감독 기관	1국 1기관을 원칙으로 하며, 행정부 소속이 아닌 별도의 독립된 기관으로 GDPR에 따른 사항을 감시하여야 함
과태료	이 규정을 위반하는 경우 기업 규모에 따라 최대 2천만유로 또는 전세계 매출액의 최대 4%까지 과태료 부과가 가능함

출처: 유럽 일반데이터보호규정(gdpr.eu) 주요 내용 발췌 및 요약

- 한편, 챗GPT의 등장과 함께 유럽연합에서는 유럽 AI법(European AI Act) 제정을 가속화하고 있음
- 유럽 AI법은 챗GPT를 비롯한 다양한 AI 서비스를 포괄하는 규제 성격의 법으로 주로 기술의 위험성을 최소화하기 위해 준비 중인 법안임
- 법안의 주요 내용은 인간의 행동에 영향을 미칠 수 있는 시스템, 개인 또는 집단의 취약점을 이용하는 시스템, 생체 인식, AI를 이용한 신뢰성 평가, 범죄 예측, 무작위 안면 인식, 법의 집행과 출입국 심사 등에서 감정을 추론하는 시스템 등을 금지하는 것을 골자로 함
- 이에 더해 챗GPT를 직접적으로 겨냥한 조항도 포함되어 있는데, 대규모언어모델의 훈련에 필요한 기초 모델을 수집하는데 안전 점검과, 데이터의 적합성, 정확성 등을 확보하기 위한 위험 완화 수단을 별도로 마련하여야 한다는 조항도 있음
- 최근에는 준비 중이던 AI법 초안이 유럽연합 의회를 통과(2023. 6. 14)하여 전 세계 최초의 AI법 제정을 앞두고 있으며, 의회 측은 2023년 연말까지 AI법을 제정하겠다는 의지를 보이고 있음
- 그러나 유럽연합 회원국 간의 이해관계 조정 등 법안의 별도 심의로 인해 현시점에서는 실질적인 법 제정까지는 상당한 시간이 걸릴 것으로 예상됨

독일 노르트라인-베스트팔렌(州) 교육부의 챗GPT 사용 가이드라인

- 유럽연합의 주요국인 독일에서도 챗GPT에 대한 비판적 시각이 우세하지만, 데이터보호규정 이외에 별도로 규제안을 도입할 계획은 존재하지 않음
- 다만 지방정부 차원에서 챗GPT를 비롯한 텍스트 AI 기술을 어떻게 안전하게 사용할 것인가에 대한 가이드라인을 제시하기 시작했음
- 독일 서부의 노르트라인-베스트팔렌(NRW: Nordrhein-Westfalen)주 교육부는 2023년 2월 학생들이 텍스트 AI 기술로 받는 부정적인 영향을 최소화하기 위해 텍스트 생성 AI 시스템 전반에 대한 사용 가이드라인(Umgang mit textgenenierenden KI-Systemen - Ein Handlungsleitfaden)을 제작 배포하였음
- 가이드라인은 텍스트 생성형 AI 서비스의 정의, 서비스가 잘못된 정보를 제공하는 이유에 대한 설명, 학교 수업에 사용하기 위한 지침, 이용을 위한 법적 조건, 교사가 수업에 서비스를 이용하기 위한 조건 등 전반적인 텍스트 생성형 AI 서비스의 특징과 무분별한 사용을 지양하기 위한 지침을 제공함
- 노르트라인-베스트팔렌주는 유사한 지침을 관내 대학교에도 제작·배포할 예정으로, 고등교육과 각종 연구 수행을 담당하는 대학 및 연구기관에서 AI를 올바르게 사용하도록 노력을 기울이고 있음

독일 슐레스비히-홀슈타인(州) 정부의 챗GPT 도입 시도

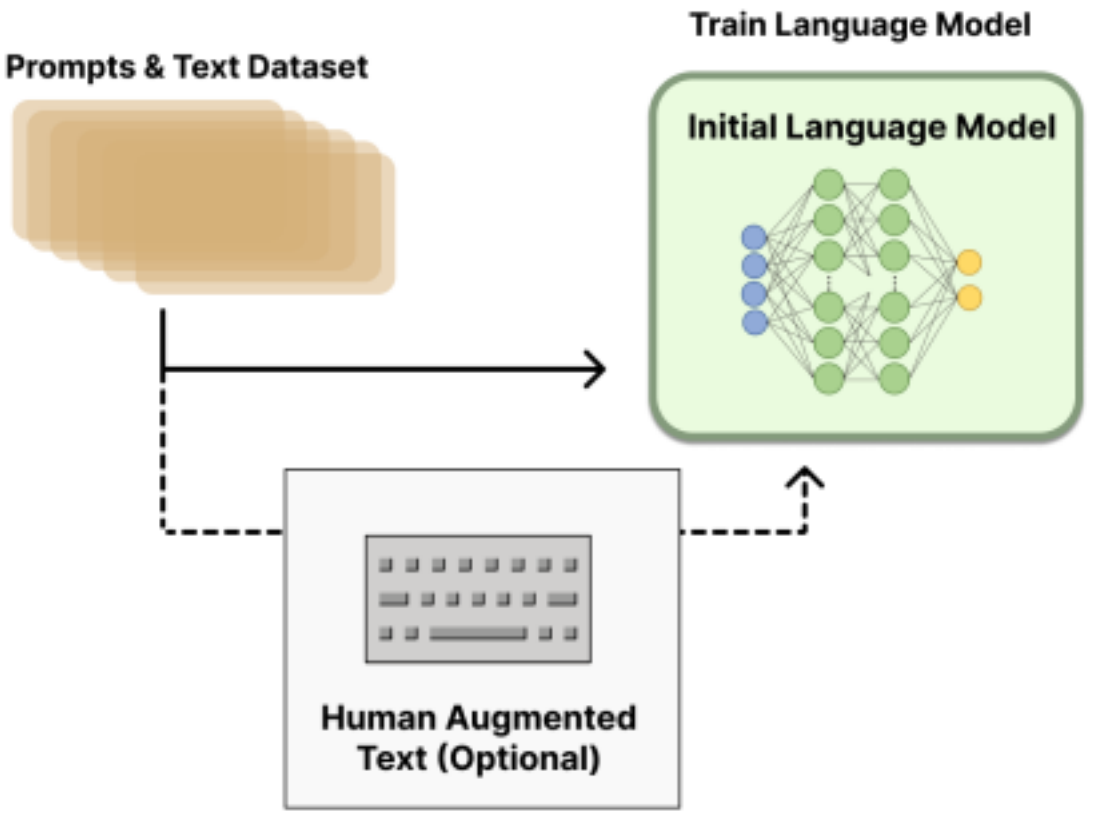
- 독일의 지방정부는 행정에 챗GPT와 같은 텍스트 생성형 AI의 도입을 적극적으로 검토하고 있음
- 독일 북서부의 슐레스비히-홀슈타인(SH: Schleswig-Holstein)주는 독일 지방정부 최초로 행정에 챗GPT를 도입하겠다고 선언한 바 있음
- 2023년 4월 연방 디지털부 장관 디르크 슈뢰터(Dirk Schrödter)는 온라인 매체 하이제(Heise)와의 인터뷰에서 “챗GPT와 같은 언어모델이 행정에 혁신을 가져올 것”, “행정의 미래는 자동화, 알고리즘화, 클라우드화, 데이터 기반이 될 것”이라며 이를 지지하는 발언을 하기도 하였음
- 그러나 거대언어모델의 기술적 한계로 행정업무 및 서비스 전반에 챗GPT를 도입하기에는 아직 무리가 있으며, 슐레스비히-홀슈타인주에서는 초기 단계로 우리나라와 유사하게 연설 준비, 업무처리의 설명 또는 요약 등의 업무에 기술을 사용하고자 함
- 또한 디지털부 장관이 언급한 행정의 자동화에 도달하기 위한 과정에서 유럽 데이터보호규정의 규제가 행정의 챗GPT 도입에 큰 걸림돌이 되고 있는 실정임
- 슐레스비히-홀슈타인주 데이터보호담당관은 “챗GPT의 행정도입을 위해 위험성 및 데이터보호규정 및 독일의 정보통신이용법 등의 관련 규정을 준수하는지 검토가 우선이며, 현재로서는 불가능하다”는 의견을 밝혔음[1)]
- 가장 큰 문제는 행정의 내부 데이터를 타국의 회사인 OpenAI에 제공하여야 한다는 점인데, 여기에는 행정의 비공개 정보, 시민의 개인정보 등 민감한 정보를 처리하기 위한 문제점 또한 존재하며, 아직 이러한 사항과 관련된 법규는 독일에 없는 상황임
- 연방 디지털부 장관의 발언이나 슐레스비히-홀슈타인주의 선언은 다소 정치적인 수사로 볼 수도 있으나, 초기 도입의 범위를 정한다던가, 관련 법규에 대한 검토를 시작했다는 점에서 긍정적으로 볼 수 있는 측면이 있음

거대언어모델(LLM)의 기술적 한계

- 행정에 챗GPT와 같은 거대언어모델을 도입하기 위해서는 투명성이 우선시 되어야 함
- 그러나 거대언어모델, 즉 머신러닝(Machine Learning)의 특징을 살펴보면 정보처리의 투명성 확보가 어려운 단점이 있음

1) https://www.heise.de/news/Bundesweite-Vorreiterrolle-Wie-Schleswig-Holstein-챗GPT-nutzen-will-8991494.html

그림 1. 언어모델의 기본적인 훈련 과정



- 언어모델의 기초 생성 과정을 살펴보면 위의 그림과 같이 수많은 텍스트 자료가 언어모델에 투입(파란색 동그라미)되고 알고리즘을 거쳐 텍스트를 생성(노란색 동그라미)하게 됨
- 언어모델은 정확성을 위해 인간에 의해 임의로 조정하는 과정이 필수적이며, 언어모델이 훈련을 거치는 동안의 과정(연두색 동그라미)이 어떻게 작동하는지 현재까지의 머신러닝(Machine Learning) 연구에 의하면 “거대한 양의 텍스트를 모델이 학습하고 처리하는 중간 과정에서 무슨 일이 일어나는지 모른다”가 정설로 통용되고 있음
- 인간이 알고리즘을 개발하기는 했지만, 아직 학습과정을 알 수 없는 단계이고, 학습은 인간이 아닌 모델(컴퓨터)이 수행하며, 거기에 인간의 미세조정이 투입된다는 점에서 완벽한 투명성이 확보되기 어려움
- 또 다른 기술적 한계점은 거대언어모델의 가장 큰 단점인 잘못된 답변의 생성, 일명 ‘환각(Hallucination)’ 증상임
- 언어모델의 환각 증상은 기초 훈련자료 수집단계에서 방대한 자료를 투입할 때 잘못된 정보 또한 투입되기 때문에 발생하는 현상임
- 챗GPT 또한 GPT-4 버전을 출시하여 환각 증상을 많이 개선시켰으나, 고의로 환각 증상을 유도하는 질문을 하면 여전히 부정확한 정보를 생성하는 문제점이 있음

시사점

- AI의 발전과 챗GPT와 같은 정교한 거대언어모델의 등장은 우리의 일상뿐만 아니라 행정 분야를 비롯하여 시민들에 대한 다양한 공공서비스 제공에 영향을 미치고 있음
- 하지만 전 세계적인 열풍과는 달리 유럽 지역, 특히 유럽연합 차원에서는 AI 기술의 위험성을 일찍이 인지하고 기존의 데이터보호규정을 이용해 선제적인 규제를 실시하기도 하였으며, 보다 안전한 기술 사용을 위한 AI 법 제정을 준비하는 등 신기술에 대한 규제를 적극적으로 검토하고 있음
- 또한 시민의 삶과 밀접하게 닿아 있는 행정 분야에서도 AI 기술의 도입을 적극적으로 검토하고 있음
- 다만, 시민의 삶과 밀접한 관계를 갖고 있고, 그 영향이 큰 행정 분야의 특성상 새로운 기술 도입을 통한 혁신도 중요하지만 보다 보수적인 관점에서 기술의 안전성을 확보하고 위험 요소를 제거하는 것이 우선시 되어야 함
- 특히 민원서비스 제공 등의 행정서비스는 정확성과 적시성이 중요한데, 앞서 언급한 잘못된 정보를 제공하는 환각 증상과 같은 기술적 한계점들이 극복되지 못한다면 책임 있는 행정의 자동화가 아닌 현재의 평범한 챗봇 수준과 크게 다르지 않은 서비스가 될 가능성도 있음
- 2차 산업혁명을 통한 산업의 비약적인 생산성 발전은 환경오염에 대한 규제를 낳았고, 3차 산업혁명으로 불리는 IT 기술의 발전과 보급은 개인정보보호라는 규제를 만들어 내기도 했던 것처럼, 패러다임 변화에 따른 기술의 안전한 이용과 위험성을 최소화하기 위해서는 규제가 뒤따를 필요가 있음
- 우리나라에서도 AI 법안(인공지능산업 육성 및 신뢰기반 조성 등에 관한 법률안)이 법안소위를 통과(2023.2.14.)하여 법사위와 본회의 통과를 남겨두고 있음
- 그러나, 현재 계류 중인 우리의 법안은 산업 육성에 초점이 맞추어져 있고, 규제의 성격은 매우 약함
- 특히 우선허용·사후규제라는 조항으로 AI 기술을 악의적으로 이용할 수 있는 여지를 남겨두기도 하였으며, 법률 위반에 따른 처벌 규정 또한 미약한 수준임
- 따라서, 앞으로 우리의 법안 심사 과정에서 새로운 기술을 어떻게 받아들이고 산업을 육성해 나아갈 것인가, 기술 도입과 이용을 어떻게 할 것인가에 주요한 초점을 맞추기보다는 안전한 기술 이용을 위한 최소한의 규제 조항을 고민해 나가야 할 단계로 판단됨

참고자료

1) Lambert et al. (2022) Illustrating Reinforcement Learning from Human Feedback(RLHF)

장인성 통신원
drong85@naver.com
독일 아헨공과대학교 (RWTH Aachen University)