

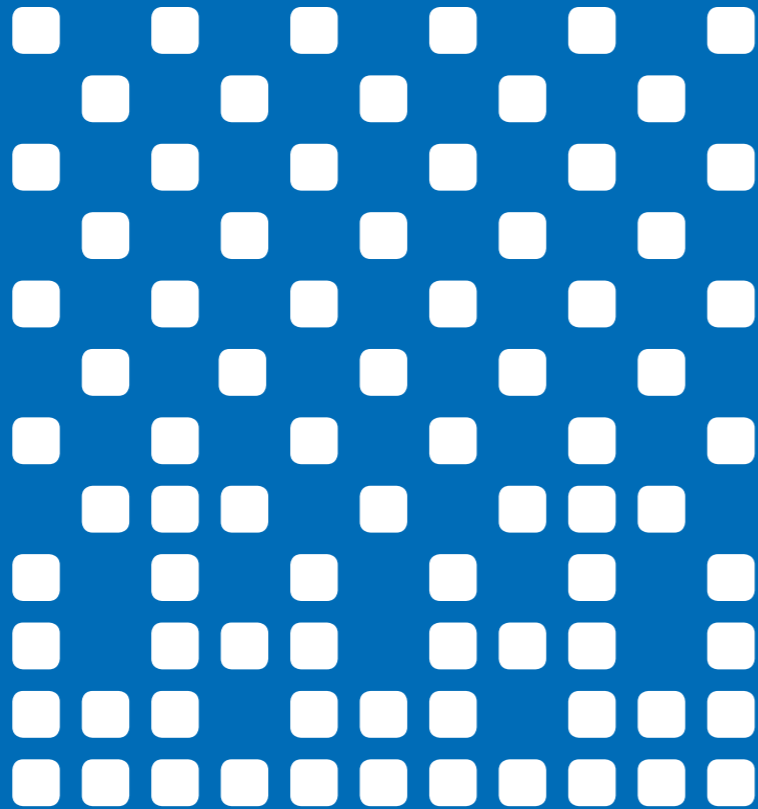


# 미디어 알고리즘과

# 민주주의

# 알고리즘의 투명성과 설명가능성 : GDPR을 중심으로

이선구(교수, 연세대학교 언더우드국제대학)



웹에서 PDF 바로 보기

## 요약

민주주의의 발전을 위해서는 언론의 안정적 운영이 중요한 의미를 가진다. 우리나라에서는 포털을 이용하여 뉴스를 접하는 인구의 비율이 월등히 높기 때문에 포털이 언론으로서 비중있는 역할을 수행한다. 따라서 포털이 사회적 신뢰를 기반으로 안정적으로 언론의 역할을 수행하는 것이 중요한데, 현재 우리 사회에서 포털에 대한 신뢰도는 다른 언론에 비해 상당히 낮은 편이다. 게다가 최근 포털들은 인공지능이라는 새로운 '편집자'를 맞이하였다. 포털의 인공지능은 기사의 추천, 배열 등 포털이 다양한 소스로부터 취합한 뉴스를 제공하는 방식을 결정한다. 인공지능은 정확성이 매우 높고 뉴스 이용자의 정보검색의 효율성을 배가시킬 수 있다는 장점이 있다. 하지만 딥러닝 등 복잡한 알고리즘으로 구성된 인공지능은 일종의 블랙박스(black box)로서 그 작동기제에 대한 불투명성이 사회적 신뢰를 얻는 데에 부정적인 영향을 가져올 수 있다. 따라서 포털이 사용하는 인공지능에 대한 신뢰를 확보하는 것은 언론의 안정성을 위하여, 나아가서는 민주주의의 발전을 위하여 사회적으로 중요한 과제이다.

인공지능에 대한 사회적 신뢰를 확보하기 위한 방안으로 인공지능의 투명성을 확보하는 것이 주로 논의되고 있다. 그런데 투명성을 제고하는 방법은 정보주체인 개인, 즉 포털 뉴스 이용자에게 인공지능에 대하여 설명을 하는 방법과 정보주체가 아닌 제3의 조직 내지는 기구가 인공지능의 작동을 관리감독을 하는 방법이 있다. 이 글에서는 둘 중 전자에 초점을 두어 인공지능에 대한 설명이 과연 투명성을 제고하여 인공지능에 대한 사회적 신뢰를 얻는 데에 기여할 수 있을지를 살펴보고자 한다.

이에 관한 논의는 2018년에 발효한 유럽연합의 일반정보보호규정(General Data Protection Regulation, 이하 'GDPR')을 중심으로 활발하게 이루어지고 있다. GDPR 제13조, 제14조, 제15조는 프로파일링을 포함한 자동화된 의사결정에 대하여 컨트롤러로 하여금 정보주체에게 '관련된 논리에 관하여 의미있는 정보'를 제공하도록 한다. 이 규정들과 전문(Recital)의 내용을 둘러싸고, 과연 GDPR이 컨트롤러에게 설명의무를 부과한 것인지 아니면 정보주체에게 설명을 요구할 권리를 인정한 것인지에 관한 논의가 분분하다. 또한 컨트롤러의 설명의무의 내용이 무엇인지에 관한 논의도 활발하다.

우선, GDPR이 컨트롤러(미디어의 맥락에서는 포털)의 설명의무를 규정한 것인지 정보주체(미디어의 맥락에서는 뉴스 이용자)의 설명을 요구할 권리(right to explanation)를 규정한 것인지에 대하여는 전문의 효력을 어떻게 보느냐에 따라 견해가 나뉘는 것이 보통이다. 본문의 규정들은 컨트롤러의 의무를 규정한 데에 반하여 전문은 정보주체에게 설명을 요구할 권리가 있다고 규정을 하였기 때문이다. 전문은 형식적으로는

법적 구속력을 가지지 않지만 유럽사법재판소는 법적 구속력을 가지는 것과 다르지 않다고 판결하였고 이와 유사한 견해를 가지는 학자들이 상당수이다. 따라서 전문에서 정한 바와 같이 설명을 요구할 권리를 가진다고 보는 견해가 있는 반면, 전문은 법적 효력이 없으며 GDPR의 입법 과정을 살펴보면 정보주체의 설명을 요구할 권리까지 인정하는 것은 입법자의 의도를 무시한 비약적 해석이라는 주장이 있다.

설명 의무의 내용에 관하여서는, GDPR과 제29조 실무그룹(Article 29 Working Party)의 가이드라인이 인공지능의 설명가능성을 전제로 하는데 반하여 현실적으로 인공지능은 상당부분 설명가능하지 않다는 지적이 다수이다. 인공지능은 설계한 사람조차 설명이 불가능한 것이 많고, 설명이 가능하다고 하더라도 설명을 제공받는 정보주체(즉, 포털 뉴스 이용자)가 과학적 전문성이 부족하여 포털이 제공하는 설명이 투명성을 확보하는데에 의미가 없다는 주장이다. 정보주체에게 의미가 있는 설명의 내용에 대하여, 인공지능의 간단한 논리에 대한 설명을 제공하는 것이 타당하다는 견해부터, 인공지능으로 인한 특정 결정에 대하여 구체적인 설명을 제공해야 한다는 견해까지 다양한 견해가 제시되고 있다.

인공지능에 대하여 정보주체에게 어느 이상의 설명을 제공하는 것이 현실적으로 곤란하거나 불가능하다면, 제3의 조직 내지는 기구를 활용하여 투명성을 제고하려는 노력을 병행하는 것이 보다 실현가능한 정책이라는 결론이다.

## 1. 서론

언론은 유권자들에게 사회적 이슈에 대하여 정확한 정보를 제공하고 다양한 의견을 전달하여 유권자들이 정치적 의사결정을 하는 데에 필수적인 역할을 한다. 현재 언론 환경은 포털의 인공지능 도입으로 인하여 획기적이고 중요한 변화를 겪고 있다. 과거에는 이용자 누구나 언론으로부터 동일한 정보를 전달받았다. 그러나 최근 포털이 인공지능을 도입하여 이용자의 뉴스 이용 행태를 분석하여 개인에게 최적화된 '맞춤형' 뉴스를 제공하기 시작함에 따라 이용자별로 조금씩 다른 정보를 제공받게 되었다. 한국에서 가장 널리 사용되는 포털인 네이버와 구글이 도입한 인공지능이 그 예이다.<sup>1</sup>

### 네이버의 AiRS

네이버는 2018년 알고리즘을 통한 뉴스 편집 기술을 도입하여(AiRS AI Recommender SYstem) 메인 화면에 게재할 기사를 선택한다. 이때 적용되는 알고리즘은 기사의 질을 판단하는 품질모델과 개인의 기사 소비 성향을 바탕으로 유사한 성향을 가진 사람들이 많이 본 기사를 추천하는 방식인 협력필터를 사용한다. 이 때 협력필터와 함께 독자가 콘텐츠를 본 순서까지 고려한 딥러닝 학습으로 정교하게 분석하여 콘텐츠를 추천하는 인공지능경망(순환신경망, Recurrent Neural Network) 기술을 사용한다.

### 구글 뉴스의 인공지능

구글 뉴스는 구글 뉴스의 설정(관심사, 출처)과 구글 검색 기록, 유튜브 등에서 이용자의 활동을 종합하여 개별화된 뉴스를 보여주는 서비스를 제공한다. 새로운 구글 뉴스는 새로운 인공지능법을 사용해 끊임없이 웹에 올라오는 정보를 실시간으로 분석하고 스토리라인 중심으로 정리한다. 이러한 접근 방식으로 구글은 이야기가 진행됨에 따라 그 이야기에 관련된 사람, 장소 및 사물을 이해하고 그 사이의 관계를 그려낼 수 있다. 또한 구글 뉴스 앱에서 "추천" 탭을 이용하면 이용가 관심을 가지고 있는 것에 대한 최신 정보를 한 곳에서 볼 수 있다. 또한 구글은 이용자를 위해 주요 헤드라인, 지역 뉴스 및 사용자의 관심사에 대한 최신 소식이 포함된 5개의 맞춤 스토리를 제공한다. 구글의 인공지능은 사용자가 앱을 더 많이 사용할수록 성능이 좋아진다. 구글은 앱에게 사용 및 접근이 가능한 제어 기능을 갖춰 이용자는 선택에 따라 특정 주제나 퍼블리셔의 스토리가 노출되는 빈도를 조절할 수 있다.

1 김민수, "구글과 네이버·카카오의 'AI 뉴스편집'은 무엇이 다를까?", 노컷뉴스, <https://www.nocutnews.co.kr/news/4968432>.

한국에서는 포털을 통해 뉴스를 제공받는 인구의 비율이 높다. 2018년 한국언론진흥재단이 발표한 바에 따르면(‘2018 언론수용자 의식조사 2018’), 응답자의 약 91%가 조사일 직전 1주일 동안 인터넷을 통하여 언론사 공식 홈페이지(기존의 신문, 방송사의 인터넷판)에 직접 접속하여 인터넷 뉴스를 이용한 적이 없다고 답했다. 모바일 뉴스의 경우도 비슷하여, 78%가 포털을 통하지 않고서는 뉴스를 접한 적이 없다고 답했다. 또한 응답자의 62%는 포털을 언론으로 여긴다고 한다. 영국의 어느 연구소가 발표한 ‘디지털뉴스 리포트(Digital News Report) 2019’는, 국민의 64%가 직접 언론사 홈페이지에서 뉴스를 접하는 스웨덴은 “직접”형 국가의 대표적인 예로, 42%가 소셜네트워크서비스(이하, SNS)로 뉴스를 접하는 칠레는 “SNS 우선”형 국가의 예로 든 반면, 75%가 포털을 통하여 뉴스를 접하는 한국은 “모아보기”형의 대표적인 예로 꼽았을 정도로 우리나라에서 포털은 중요한 언론의 기능을 담당하고 있다.<sup>2</sup>

포털 뉴스는 이용자의 정보 검색 비용을 절감한다. 여러 언론사의 뉴스를 한군데 모아놓고 주제별로 검색할 수 있게 하므로 이용자가 정보를 검색하는 데 드는 시간과 노력이 절감된다. 그리고 하나의 사건에 대하여 여러 언론사의 뉴스를 접할 수 있어 다양한 시각을 접할 수 있다. 이러한 포털 특유의 효율성에 이용자의 취향 및 뉴스 이용 패턴에 맞추어 인공지능이 제공하는 맞춤형 뉴스 서비스까지 제공됨으로서 이용자의 정보 수집 효율은 극대화된다. 가령, 대여섯 부의 신문을 한 장 한 장 들추며 자신이 관심을 가지는 주제에 대한 뉴스를 찾아 읽는 데 비하여 포털 이용자는 하나의 웹사이트에서 짧은 시간에 자신의 관심주제에 대한 뉴스를 모두 제공받아 접할 수 있다. 따라서 인공지능과 포털의 결합은 효율성의 관점에서 매우 반길 만한 일이다.

그런데 포털은 다양한 언론 매체로부터 취합한 뉴스를 편집하여 제공하는 과정에서 포털 접속자의 뉴스 이용에 일정한 영향력을 행사할 수 있다. 우리나라의 포털 뉴스 이용 행태 조사(‘2018 언론수용자 의식조사’)에 따르면, ‘포털사이트의 첫 페이지(메인화면)의 뉴스 머리기사나 사진을 보고 뉴스를 클릭해 이용한다’는 응답자의 비율은 58%에 육박했다. 또한 ‘실시간 검색 순위에 오른 인물이나 사건을 찾아서 이용한다’는 비율이 52%에 달했다. 반면, ‘내가 미리 설정한 뉴스(구독뉴스 등)를 통해 이용한다’는 비율은 13%에 그쳤다. 이러한 상황에서 인공지능이 배열 및 추천을 담당하는 새로운 ‘편집자’로서의 역할을 담당하게 되었으니, 포털의

2 Reuters Institute, Digital News Report 2019, 2019, 13면. 본 보고서에서는 검색엔진(48%)과 에그리제이터(27%)를 나누어 집계하였으나 두 언론의 조작적 정의가 명확하지 않고 한국의 상황에서는 포털로 묶어서 보아도 무방하다고 판단하여 포털이라는 하나의 언론으로 표현하였다.

인공지능에 대한 사회적 신뢰를 확보할 필요가 있다. 비단 인공지능의 문제가 아니더라도 이미 포털에서 제공하는 뉴스에 대한 신뢰도가 다른 뉴스 채널에 대한 신뢰도보다 낮은데, 지상파 텔레비전 방송을 매우 혹은 대체로 신뢰한다는 의견은 70%, 종합편성채널은 68%, 보도전문채널은 63%에 육박함에 반해 포털에 대한 긍정적 평가는 47%에 그쳤다. 따라서 언론의 인공지능에 대한 사회적 신뢰의 확보는 매우 중요한 정책적 과제이다.

현재까지 인공지능의 사회적 신뢰를 둘러싸고 이루어진 담론은 인공지능의 투명성을 제고하여 신뢰를 확보해야 한다는 것이다. 언론 분야도 예외는 아니다. 그런데 투명성을 확보하기 위해서는 미디어 알고리즘에 대한 설명을 제공하는 것이 필수적이므로, 투명성의 문제는 곧 인공지능의 설명가능성(explainability)의 문제로 귀착된다. 본 연구에서는 우리나라의 언론 소비 특성을 감안하여 미디어 알고리즘 중에서도 포털에서 사용하는 알고리즘의 설명가능성에 초점을 둔다.

인공지능의 설명가능성은 유럽연합의 일반정보보호규정(General Data Protection Regulation, 이하 ‘GDPR’)의 도입을 계기로 연구가 활발해진 분야이다. GDPR은 정보보호규칙(Data Protection Directive)을 대체할 목적으로 2018년 5월에 발효한 유럽연합의 개인정보보호 법령으로, 프로파일링을 포함한 자동화된 의사결정에 대하여 법적인 설명의무를 부과한 최초의 입법적 시도이다. 본 고에서는 GDPR을 둘러싼 논의를 중심으로 인공지능의 투명성과 설명의 관계, 인공지능의 설명가능성을 바탕으로 포털의 설명의무와 그 내용에 대하여 살펴본다. 마지막으로, GDPR에 따른 설명의무의 이행에 일정한 한계가 있음을 고려하여 그 대안으로 제시되는 정책들을 살펴보고자 한다.

## II. 미디어 알고리즘의 투명성과 설명(explanation), 사회적 신뢰

### 1. 포털 인공지능의 투명성

우리나라에서는 상당수의 뉴스 이용자가 포털을 통하여 뉴스를 접하기 때문에 포털의 기사 추천 방법뿐만 아니라 기사의 배열, 실시간 급상승 검색어의 순위 결정 등이 여론의 형성에 영향력을 행사할 수 있다. 포털들은 최근 인공지능을 적극적으로 도입하여 인공지능이 기사의 배열, 추천 등을 담당하도록 하였다. 그런데 인공지능의 경우 전통적인 언론 매체의 편집국에서 기사를 배열하는 방법과는 달리, 복잡하고 다소 불투명한 방식에 의하여 작동한다. 특히 딥러닝의 경우 머신러닝 알고리즘 중 하나인데 인공지능경망에 기반하여 설계된 알고리즘은 매우 높은 정확도가 그 장점이지만, 그만큼 여러 층으로 복잡하게 구성된 인공지능경망의

작동기제가 매우 복잡하고 불투명한 측면이 있다. 따라서 포털의 알고리즘에 대한 투명성을 높이는 것이 언론으로서 포털에 대한 신뢰도를 높이는 데에 관건이 될 것이다.

‘우리말샘’에 따른 투명성의 사전적 정의는 ‘일의 상황이나 성질 따위가 확실하거나 분명한 성질’이다. 따라서 알고리즘에 대하여 필요한 설명을 충분한 정도로 제공한다면 알고리즘에 대한 투명성을 제고할 수 있다. 이로써 자연스럽게 알고리즘에 의하여 발생할 수 있는 의도적 혹은 비의도적 조작 내지는 편향에 관한 의혹을 해소하고 사회적 신뢰를 구축해나갈 수 있다.<sup>3</sup>

## 2. 설명이 투명성의 제고에 가지는 의미

알고리즘의 투명성을 제고하는 방안은 크게 정보주체에 대한 정보의 제공과 정보주체가 아닌 제3자에 의한 관리감독이 있다. 정보주체는 자신의 정보가 어떻게 사용되는지를 모니터링을 하려는 유인이 있고, 알고리즘이 내린 결정에 가장 직접적인 영향을 받으므로, 정보주체에게 정보를 처리하는 알고리즘에 관한 설명을 하여 알고리즘이 차별적 요소를 가지지는 않는지 등에 관하여 경계하도록 할 수 있다.<sup>4</sup> 정보주체 외의 제3자는 정보주체에 비하여 알고리즘을 경계할 유인이 적지만 전문성을 가지고 알고리즘의 정확성, 편향성 등에 대하여 지속적인 관리 및 감독을 하여 알고리즘의 순작용이 극대화되도록 할 수 있다. GDPR은 (후술하는 바와

3 Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, 7(2), 76-99, 77면; Catherine Stupp, ‘Commission to Open Probe into Tech Companies’ Algorithms next Year’ (EurActiv.com, 8 November 2016), <http://www.euractiv.com/section/digital/news/com-mission-to-open-probe-into-tech-companies-algorithms-next-year/>; Partnership on AI, ‘Partnership on Artificial Intelligence to Benefit People and Society’ (Partnership on Artificial Intelligence to Benefit People and Society, 2016), <https://www.partnershiponai.org/>; National Science and Technology Council Committee on Technology, ‘Preparing for the Future of Artificial Intelligence’ (Executive Office of the President 2016), [https://www.whitehouse.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf); Government Office for Science, ‘Artificial Intelligence: An Overview for Policy-Makers’ (Government Office for Science 2016), <https://www.gov.uk/government/publications/artificial-intelligence-an-overview-for-policy-makers>; Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018, April). Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (p. 582). ACM.

4 Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O’Brien, D., ... & Wood, A. (2017). Accountability of AI under the law: The role of explanation. *arXiv preprint arXiv:1711.01134*.

같이) 정보주체에게 정보를 제공하는 방식을 채택하였다.

## 3. GDPR에서 ‘공정하고 투명한 처리의 원칙’을 위한 설명

GDPR은 매우 방대한 내용의 전문(Recital)과 99개의 조항으로 이루어진 본문으로 구성되어 있다. 형식적으로 본문은 법적 구속력을 가지나, 전문은 법적 구속력을 가지지 않는다. 전문은 본문에 포함되지 않은 내용을 담기도 하고 때로는 본문의 내용을 구체화하기도 하는데, 본문의 협상 과정에서 본문에 포함되지 못한 내용이 법적 구속력을 가지지 못하는 전문에 포함된 것으로 평가된다.<sup>5</sup> 유럽사법재판소는 Case 215/99 Casa Leischhandels [1989] European Court of Justice ECR 2789 [31] 판결에서 전문이 “그 자체로 규칙이 되진 못하지만 법적 규칙을 해석하는 데에 실마리를 제공한다”고 판결하였다. 따라서 GDPR 본문의 내용이 모호할 경우 전문은 GDPR을 해석하는 데에 단서를 제시할 수 있다. 따라서 법적 구속력이 있는 본문과 비슷한 효력을 가진다고 볼 수도 있다는 해석이 다수이다.<sup>6</sup>

GDPR을 해석하고 적용하는 데에 참고할 만한 또 다른 권위 있는 문헌은 제29조 실무그룹(Article 29 Working party)의 가이드라인(이하, 가이드라인)이다. 가이드라인은 실무자들이 법을 집행할 때 참고하기 위하여 공표한 문헌이므로, 법적 구속력은 없더라도 GDPR을 해석하는 데에 유용한 자료가 된다.<sup>7</sup>

GDPR은 전문에서 ‘공정하고 투명한 정보처리(fair and transparent processing)의 원칙’을 선언하고 이 표현을 여러 번에 걸쳐서 사용하는데, 여기서 공정성은 결과의 공정성보다는 정보처리과정에서 투명성 확보를 통한 절차적 공정성을 의미하는 것으로 해석하는 것이 일반적이다.<sup>8</sup>

**공정하고 투명한 처리의 원칙에 따라 정보주체는 처리 작업의 존재 및 그 목적에 대하여 통지 받아야 한다. 컨트롤러는 개인정보가 처리되는**

5 Edwards, L., & Veale, M. (2017). Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for. *Duke L. & Tech. Rev.*, 16, 18, 50면; Kaminski, M. E. (2019). The right to explanation, explained. *Berkeley Tech. LJ*, 34, 189, 195-196, 207면.

6 Kaminski, M. E. “The right to explanation, explained”(주5), 8면; Selbst, A. D., & Powles, J. (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, 7(4), 233-242, 235면.

7 같은 글, 207면.

8 같은 글, 206면.

특정 상황 및 맥락을 참작하여 **공정하고 투명한 처리 보장에 필요한 모든 추가적인 정보를 정보주체에 제공해야** 한다. 또한 정보주체는 **프로파일링 유무와 해당 프로파일링의 결과에 대해 통지받아야** 한다. 정보주체로부터 개인정보가 수집되는 경우, 해당 정보주체는 본인이 개인정보 제공의 의무가 있는지 여부 및 해당 정보를 제공하지 않을 경우의 결과에 대해 통지받아야 한다. 그 정보는 눈에 잘 띄고 이해하기 쉬우며 가독성이 뛰어난 방식으로 예정된 처리에 대해 중요한 개략적 정보를 제공하기 위해 표준화된 아이콘과 함께 제공될 수 있다. 전자 수단을 이용하여 아이콘을 제공하는 경우에는 기계 판독이 가능해야 한다.<sup>9</sup>

전문에서는 공정하고 투명한 정보의 처리를 위해서 ① 정보주체에게 충분한 정보가 제공될 것, ② 정보주체의 이해나 권리에 발생하는 위험을 최소화할 것을 요구한다. 투명한 정보의 처리를 위하여 정보주체에게 정보의 처리에 관한 설명이 제공되어야 한다는 점에 비추어 볼 때, GDPR은 적어도 설명의무에 관한 한 개인에 대한 투명성 확보에 중점을 두고 있다고 평가할 수 있다.

하지만 실제로 설명의무를 이행하는 것, 후술하는 바와 같이, 현실적으로 매우 어려운 문제이다. 따라서 투명성을 확보하기 위한 설명의 당위성만으로는 설명이 이루어져야 한다고 결론짓기가 어렵고, 과연 얼마나 설명의무를 실현할 수 있는지에 대한 구체적인 검토가 필요하다. 이것은 인공지능이 얼마나 설명가능한지로 직결되는 문제이기도 하다. 다음 항에서는 GDPR을 둘러싼 논의를 살펴보면서 설명을 통한 투명성 제고의 가능성에 대해서 검토한다.

9 원문의 내용은 다음과 같다. The principles of fair and transparent processing require that the data subject be informed of the existence of the processing operation and its purposes. The controller should provide the data subject with any further information necessary to ensure fair and transparent processing taking into account the specific circumstances and context in which the personal data are processed. Furthermore, the data subject should be informed of the existence of profiling and the consequences of such profiling. Where the personal data are collected from the data subject, the data subject should also be informed whether he or she is obliged to provide the personal data and of the consequences, where he or she does not provide such data. That information may be provided in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner, a meaningful overview of the intended processing. Where the icons are presented electronically, they should be machine-readable.

### III. GDPR에 따른 설명의무

GDPR 본문의 설명의무에 관한 조항들은 상당히 추상적일 뿐만 아니라 전문과 차이가 있어 본문의 해석을 두고 논란이 많다. 논란의 쟁점은 크게 두 가지로 나뉘는데, GDPR이 컨트롤러가 설명을 제공할 의무를 규정한 것인지, 아니면 정보주체에게 설명을 요구할 권리를 규정한 것인지가 첫 번째 쟁점이다. 이 글의 맥락에서 컨트롤러는 포털, 정보주체는 포털의 이용자에 해당된다. 두 번째 쟁점은 설명의 내용과 정도이다. 과연 인공지능의 알고리즘은 어느 정도로 설명가능한가, 그리고 그러한 설명이 정보주체, 즉 뉴스 이용자에게 어떠한 의미를 가지는가에 관한 논의이기도 하다.

#### 1. 설명의무인가, 설명을 요구할 수 있는 권리인가?

##### 1) GDPR의 내용

GDPR이 컨트롤러의 설명의무를 규정하고 있다는 데에는 별다른 이견이 없지만, 과연 정보주체에게 얼마나 구체적이고 적극적인 권리를 인정한 것인지에 대해서는 논쟁이 끊이지 않고 있다. 우선 본문의 경우 분명히 컨트롤러가 정보주체에게 일정한 설명을 할 의무를 부과한다. 제13조, 제14조, 제15조는 공통적으로 정보가 수집, 처리되는 과정에서 컨트롤러는 정보주체에게 ① 프로파일링을 포함한 자동화된 의사결정이 이루어지고 있다는 점, ② 관련된 논리(logic)에 관한 의미 있는(meaningful) 정보와 ③ 정보주체의 정보를 처리하는 것의 의의와 그에 따라 예상되는 결과를 설명해주어야 한다고 정한다. GDPR 제4조는 프로파일링은 ‘특히 자연인의 업무 성과, 경제적 상황, 건강, 개인적 선호, 관심사, 신뢰도, 행태, 위치 또는 이동에 관한 측면을 분석하거나 예측하기 위해 행해지는 경우로서, 자연인에 관련한 개인적인 특정 측면을 평가하기 위해 개인정보를 사용하여 이루어지는 모든 형태의 자동화된 개인정보의 처리’라고 정의한다. 따라서 개개인에게 맞춤형 뉴스를 제공하기 위하여 뉴스 이용 패턴, 선호도 등의 정보를 처리하는 것은 프로파일링에 해당한다. 그런데 전문에서는 여기서 한걸음 더 나아가 정보주체에게 설명을 ‘요구할 권리’가 있다고 규정한다. 즉, 전문 (71)는 “어떠한 경우에도, 정보주체에게는.... 이러한 (알고리즘에 기초한) 평가에 따른 결정이 있을 후 그 결정에 대한 설명을 요구할 권리, 그리고 결정에 이의를 제기할(challenge) 권리가 있다”고 정한다. 그리고 가이드라인은 이 ‘설명을 요구할 권리’라는 표현을 무려 세 번이나 인용한다.<sup>10</sup>

10 Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01), [https://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=612053](https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053).

### 전문(71)

전문 정보주체는 온라인 신용신청의 자동 거절이나 인적개입 없이 이루어지는 전자채용 관행 등 자동화 처리에만 근거하여 본인에 관한 개인적인 면을 평가하고 본인에게 법적인 영향 또는 유사하게 중대한 영향을 초래하는 조치를 포함할 수 있는 결정에 따르지 않을 권리를 가진다. 그 처리에는 개인에 관한 개인적인 측면을 평가하는 모든 형태의 개인정보 자동화 처리로 구성된 '프로파일링'이 포함되고, 특히 정보주체의 업무능력, 경제적 상황, 건강, 개인의 선호나 관심사, 신뢰성 또는 행동, 위치나 움직임에 관한 측면을 분석 또는 예측하며, 정보주체에게 법적인 영향이나 이에 상응하는 중대한 영향을 미치는 경우 그러하다. 그러나 프로파일링 등 그러한 처리에 근거한 의사결정은 컨트롤러가 적용받는 유럽연합 또는 회원국 법률에서 명시적으로 승인하는 경우 허용되어야 하며, 유럽연합 산하기구 또는 회원국 감독기구의 규정, 기준 및 권고에 따라 실시되는 사기 및 탈세의 감시·예방 목적으로나 컨트롤러가 제공하는 서비스의 보안 및 신뢰성을 보장하기 위해, 또는 정보주체와 컨트롤러 간의 계약 체결이나 이행에 필요하거나 정보주체가 명시적인 동의를 제공하였을 때 등이 이에 해당한다. 어떠한 경우에도, 그러한 처리는 정보주체에게 제공되는 특정 정보, 인적개입을 획득할 권리, 견해를 표현할 권리, 상기 평가 후 내려진 결정에 대한 설명을 얻을 권리, 해당 결정에 이의를 제기할 권리 등 적절한 안전장치를 적용받아야 한다. 그 같은 조치에 아동은 관여되지 않는다.

### 제13조 개인정보가 정보주체로부터 수집되는 경우 제공되는 정보

1. 정보주체에 관련된 개인정보를 정보주체로부터 수집하는 경우, 컨트롤러는 개인정보를 취득할 당시 정보주체에게 다음 각 호의 정보 일체를 제공해야 한다.  
...
  2. 제1항의 정보와 함께, 컨트롤러는 개인정보가 입수될 때 공정하고 투명한 처리를 보장하는 데 필요한, 다음 각 호의 추가 정보를 정보주체에 제공해야 한다.  
...
- (f) 제22조(1) 및 (4)에 규정된 프로파일링 등, 자동화된 의사결정의 유무. 최소한 이 경우, 관련 논리에 관한 유의미한 정보와 그 같은 처리가 정보주체에 미치는 중대성 및 예상되는 결과

### 제14조 개인정보가 정보주체로부터 수집되지 않은 경우 제공되는 정보

1. 개인정보가 정보주체로부터 수집되지 않은 경우, 컨트롤러는 다음 각 호의 정보를 정보주체에게 제공해야 한다.  
...
  2. 제1항의 정보와 함께, 컨트롤러는 정보주체와 관련한 공정하고 투명한 처리를 보장하는데 필요한, 다음 각 호의 정보를 정보주체에 제공해야 한다.  
...
- (g) 제22조(1) 및 (4)에 규정된 프로파일링 등, 자동화된 의사결정의 유무. 최소한 이 경우, 관련 논리에 관한 유의미한 정보와 그 같은 처리가 정보주체에 미치는 중대성 및 예상되는 결과
3. 컨트롤러는 제1항 및 제2항에 명시된 정보를 다음 각 호와 같이 제공해야 한다.  
...

### 제15조 정보주체의 열람권

1. 정보주체는 본인에 관련된 개인정보가 처리되고 있는지 여부에 관련해 컨트롤러로부터 확답을 얻을 권리를 가지며, 이 경우, 개인정보 및 다음 각 호의 정보에 대한 열람권을 가진다.  
...
- (h) 제22조(1) 및 (4)에 규정된 프로파일링 등 자동화된 의사결정의 유무. 최소한 이 경우, 관련 논리에 관한 유의미한 정보와 그 같은 처리가 정보주체에 미치는 중대성 및 예상되는 결과

### 제22조 프로파일링 등 자동화된 개별 의사결정

1. 정보주체는 프로파일링 등, 본인에 관한 법적 효력을 초래하거나 이와 유사하게 본인에게 중대한 영향을 미치는 자동화된 처리에만 의존하는 결정의 적용을 받지 않을 권리를 가진다.  
...

이러한 전문과 본문, 가이드라인의 내용을 두고 과연 GDPR이 컨트롤러, 즉 포털운영자에게 인공지능에 관하여 일정한 설명을 할 의무를 부과한 것인지, 아니면 포털의 이용자 개개인에게 '설명을 요구할 권리(right to explanation)'를 부여한 것인가에 대한 논의가 있다.

2) GDPR이 정보주체에게 설명을 요구할 권리를 인정한다고 보는 견해 이 입장에서는 설명을 요구할 권리를 명확하게 규정한 전문을 고려하여 GDPR을 해석하여야 한다고 주장한다.<sup>11</sup> 대부분 전문은 비록 법적 구속력은 없더라도 규범적 효력의 측면에서 법적 구속력이 있는 본문과 별다른 차이가 없는 점(상술한 유럽사법재판소의 판결 참조), 가이드라인에서도 같은 견해를 반복적으로 강조하고 있다는 점을 근거로 든다. 한편, GDPR의 제13조와 제14조에 근거하여 설명을 요구할 권리가 있다고 주장하는 입장도 있다.<sup>12</sup> 위 조항들이 명확하게 권리라는 표현을 사용하지는 않았지만 법의 전체적인 맥락과 GDPR이 자동화된 의사결정에 대하여 적용하는 규제들을 고려하면 설명을 요구할 권리를 인정하는 것이 타당하다는 것이다. 이에 부연하여 실제로 설명을 하기 어렵다는 기술적인 한계를 이유로 설명을 요구할 권리를 부정할 수는 없고, 기술적 한계로 인하여 의무를 이행할 수 없다면, 기술 발전을 통해 이를 극복하거나 법률을 개정해 해결해야 한다고 주장하기도 한다.<sup>13</sup>

3) GDPR이 컨트롤러에게 설명을 제공할 의무를 부과할 뿐이라고 보는 견해  
 GDPR이 정보주체에게 설명을 요구할 권리를 부여하지 않고 있다는 견해는, 입법자들이 제22조에 '자동화된 의사결정'에 관한 규정을 마련하면서도 설명을 요구할 권리를 명시하지 않은 이유는 설명을 요구할 권리를 법적으로 보장하지 않으려는 의도에서 비롯된 것이라고 주장한다.<sup>14</sup> 이들은 GDPR의 입법 과정에서 설명을 요구할 권리에 대한 문구를 본문에 넣을 것인가를 둘러싸고 유럽의회, 유럽이사회, 유럽연합 집행위원회 사이에서 벌어진 논의들을 고려하면, 결과적으로 본문에서 이러한 권리를 규정하지 않은 데에서 입법자들의 의도를 읽을 수 있다고 주장한다.<sup>15</sup> 입법자들이 자동화된 의사결정과 프로파일링에 더 강력한 보호장치를 두고자 했음(제22조)에도 결국 설명을 요구할 권리를 본문에서 제외한

11 Kaminski, M. E. "The right to explanation, explained"(주5), 9면; Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, 38(3), 50-57, 50-57면.  
 12 같은 글, 50-57면; Selbst, A. D., & Powles, J. "Meaningful information and the right to explanation"(주6), 234-235면.  
 13 Goodman, B., & Flaxman, S. "European Union regulations on algorithmic decision-making and a "right to explanation""(주11); Selbst, A. D., & Powles, J. "Meaningful information and the right to explanation"(주6), 234-235면..  
 14 Wachter, S., Mittelstadt, B., & Floridi, L. "Why a right to explanation of automated decision-making does not exist in the general data protection regulation"(주3), 81면.  
 15 같은 글, 78면.

데에는 이유가 있다는 것이다.<sup>16</sup> 또한 법문이 '권리'에 관한 명시적 조항을 두고 있지 않은데도 정보주체에게 권리를 인정하는 것은 지나친 논리적 비약이라는 주장도 있다.<sup>17</sup> 이러한 주장에 따르면 컨트롤러가 정보를 제공할 의무는 있지만, (그리고 이용자가 그에 따른 반사적 이익으로 일정한 정보를 제공받게 되지만) 이용자 개개인이 특정 정보를 요구할 권리는 없다.

## 2. 설명의 내용

### 1) 인공지능의 설명가능성

GDPR은 자동화된 의사결정이 있는 경우 정보주체에게 "관련된 논리에 관한 의미 있는 정보(meaningful information about the logic involved)"를 제공할 것을 요구한다. 인공지능이 블랙박스(black box) 즉, 기능은 알 수 있지만 작동 원리를 이해할 수 없는 복잡한 기계 장치로서는 사회적 신뢰를 얻기 어렵기 때문에 이용자에게 그 작동 원리에 대하여 어느 정도의 설명을 제공하여 인공지능의 의사결정의 투명성을 제고하기 위함이다.

하지만 과연 이러한 설명의무가 실제로 현실에서 이행될 수 있을지, 또 투명성을 제고하는 데에 도움이 될지에 관하여는 의견이 분분하다. 실제로 인공지능의 설명가능성에 대한 논의를 활성화시킨 미국 국방위고등연구계획국(DARPA)의 보고서는 인공지능과 관련하여 두 가지 목표를 제시하였다. 이 때, 인공지능을 설명가능하게 하는 것과 동일한 선상에서 제시된 다른 목표가 바로 설명가능한 인공지능을 생성하는 것이 그것이다.<sup>18</sup> 인공지능을 설명하는 것은 쉬운 일이 아니다. 설명가능성을 중심으로 기존 인공지능의 발전 동향을 정리한 연구들을 보면, 인공지능에는 크게 인간이 그 작동기제를 충분히 이해하고 설명할 수 있는 인공지능과 인간이 인공지능의 의사결정 결과를 놓고 그에 대한 해석을 제공하거나 일정한 설명을 제공할 수 있는 정도의 인공지능으로 분류를 할 정도로, 인공지능에 대하여 완전한 설명을 제공할 수 없는 경우가 상당수이다.<sup>19</sup>

16 같은 글, 81면.

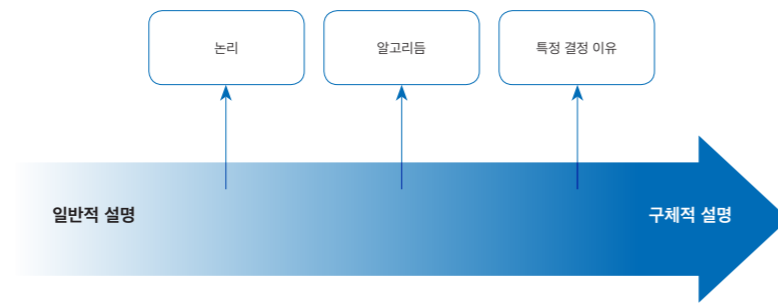
17 같은 글, 82-84면.

18 Defense Advanced Research Projects Agency(DARPA), Explainable Artificial Intelligence(XAI), Dr. Matt Turek, <https://www.darpa.mil/program/explainable-artificial-intelligence>.

19 Biran, O., & Cotton, C. (2017, August). Explanation and justification in machine learning: A survey. In *IJCAI-17 workshop on explainable AI (XAI)* (Vol. 8, p. 1).



인공지능은 그것을 생성한 사람들조차 그것에 대하여 다 알지 못하는 경우가 적지 않다. 머신러닝 중에서도 딥러닝은 특히 그것을 설계한 사람도 완전히 이해하기 어려운 경우가 많다. 따라서 인공지능에 대하여 정보주체에게 제공될 수 있는 설명의 내용이 무엇인지를 두고, 완전한 설명과 차별화하여 결정의 ‘정당한 이유(justification)’, ‘해석(interpretation)’, ‘인식(intelligibility)’을 제시하는 입장도 있다.<sup>20</sup> 가령, 정당한 이유는 의사 결정 과정은 생략한 채 결과의 타당성을 설명하는 것이고, 해석은 인간에 의해서 이해가 가능한 만큼의 설명을 말하며 인식은 인공지능이 어떠한 정보를 사용하고, 어떻게 그 정보들을 수집하였고, 그 정보를 이용하여 무엇을 할 것인지를 설명해주는 것을 말한다고 설명하는 논문도 있다.<sup>21</sup> 하지만 아직까지는 정당한 이유, 해석, 인식 등 용어의 정의가 명확하지 않고 용어의 사용도 일관되지 않다.<sup>22</sup> 따라서 용어와 무관하게, 인공지능에 대하여 제공될 수 있는 설명의 스펙트럼을 정리하자면, 결정의 논리, 알고리즘, 특정한 결정의 이유로 나눌 수 있다.



20 Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018, April). "Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda"(주3), 582면; Biran, O., & Cotton, C. "Explanation and justification in machine learning: A survey"(주19); Burt, A., Leong, B., Shirrell, S., & Wang, X. G. (2018). Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models. In Future of Privacy Forum. Biran, O., & Cotton, C. (2017, August). Explanation and justification in machine learning: A survey. In IJCAI-17 workshop on explainable AI (XAI) (Vol. 8, p. 1).

21 Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018, April). "Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda"(주3), 582면; Biran, O., & Cotton, C. "Explanation and justification in machine learning: A survey"(주19).

22 Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., ... & Wood, A. (2017). Accountability of AI under the law: The role of explanation. arXiv preprint arXiv:1711.01134.

## 2) 인공지능에 대한 설명의 내용

### (1) 결정의 논리

인공지능이 내린 특정한 결정의 이유가 아닌 인공지능의 일반적인 정보처리절차에 대하여 알려주는 방법이 있다. 가장 일반적으로 제공될 수 있는 설명은 정보처리의 논리를 알려주는 것이다.<sup>23</sup> 이러한 설명에는 알고리즘에 대한 전문적 지식을 필요로 하지 않는다. 법에서 정보주체에게 제공되는 정보를 ‘의미있는 정보’로 한정된 것을 고려하여, 일반인인 정보주체가 이해하기 힘든 전문적인 알고리즘보다는 결정에 사용된 기준, 분류체계, 의사결정모형(decision tree)의 논리 정도를 알려주는 것으로 족하다는 입장이다.

### (2) 알고리즘

결정의 논리에서 한걸음 더 나아가 알고리즘의 구조에 대해 설명할 것을 주장하는 학자도 있다.<sup>24</sup> 어떠한 결정을 내리기 위해 거친 공식화된 일련의 절차나 방법을 알려주는 것이다.

### (3) 특정 의사결정의 이유

일반적인 설명에 그치지 않고 특정 결정에 대하여 설명을 제공하는 것도 가능하다. 결정의 논리, 처리의 결과뿐만 아니라 구체적인 결과에 대한 결정요소 등을 설명하는 것이다. 대표적으로, 가이드라인에서는 다음과 같은 정보가 제공될 것을 제안하는데<sup>25</sup> 의사결정에 어떤 종류의 정보가 사용되었는지, 왜 이런 종류의 정보가 유의미하다고 판단했는지, 분석에 포함되는 모든 통계(statistics)를 포함하여 어떻게 프로파일을 형성하였는지, 왜 이 프로파일이 자동화된 의사결정에서 의미가 있는지, 이러한 프로파일이 정보주체에 대한 의사결정에서 어떻게 사용되었는지가 포함된다.

가이드라인에 의하면, 설명의무는 개인이 자동화된 의사결정을 투명하게 이해함으로써 자신의 정보가 불합리하게 사용되지 않게 할 권리를 행사할 수 있게 하는 데에 의의가 있다. 따라서 개인이 그러한 권리를 행사하기 위해 필요한 모든 정보를 제공해야 한다.<sup>26</sup> 가령, GDPR이

23 Wachter, S., Mittelstadt, B., & Floridi, L., "Why a right to explanation of automated decision-making does not exist in the general data protection regulation"(주3), 90면.

24 Selbst, A. D., & Powles, J. "Meaningful information and the right to explanation"(주6), 236-237면.

25 가이드라인(주10), 31면.

26 가이드라인(주10), 13, 27면. Kaminski, M. E. "The right to explanation, explained"(주5), 208면.

정보주체에게 차별받지 않을 권리를 보장하기 때문에 정보주체는 그 권리 행사를 위해 필요한 정보를 모두 제공받아야 한다. 이것은 정보주체의 어떠한 정보가 어떠한 비중을 두고 사용되었으며 결과에 어떻게 영향을 미쳤는지를 포함한다. 이 입장을 지지하는 자들은 이러한 구체적인 정보는 난해한 것이겠지만, 컨트롤러와 정보주체 사이의 소통은 모두 평이하고 명확한 언어를 사용하여야 한다는 GDPR 제12조의 취지에 쫓아, 컨트롤러는 정보주체가 이해할 수 있게 설명하는 것이 중요하다고 강조한다.<sup>27</sup>

이러한 주장에 대하여, 현실적으로 불가능한 제안이며 일반인의 전문적 지식의 한계가 설명의 내용에서 고려되어야 한다고 반론이 보편적이다. 과연 기술에 관하여 특별한 지식이 없는 뉴스 이용자에게 알고리즘의 구조에 대한 설명은 '의미있는' 설명인가? 그러한 설명을 통하여 궁극적으로 뉴스 이용자는 그 알고리즘을 신뢰할 수 있을까? 구체적인 결과의 결정 요소에 대한 설명은 어떨까? 가령, 기사배열에 과거에 읽은 기사들, 인터넷 사용 패턴 등이 고려되었다는 점을 알려준다면 어느 정도 의미가 있겠지만, 각각의 요소들이 어떤 비중으로 고려되었고 어떠한 수학적 혹은 통계적 메커니즘을 통하여 결론에 도달하였다는 점을 알려준다면, 뉴스 이용자에게 그 정보는 어떠한 의미가 있을까? 일반인인 정보주체가 과학지식이 부족하다는 점을 근거로 설명의 내용이 평이한 내용에 그치는 것이 타당하다는 견해가 다수이다(물론 학자에 따라 내용의 범위에 대한 입장의 차이는 존재한다).

27 GDPR 제12조 정보주체의 권리 행사를 위한 투명한 정보, 통지 및 형식 1. 컨트롤러는 처리와 관련한 제13조 및 제14조에 명시된 일체의 정보, 제15조부터 제22조까지의 조문 및 제34조에 규정된 일체의 통지를 정확하고, 투명하며, 이해하기 쉬운 형식으로 명확하고 평이한 언어를 사용하여 정보주체에게 제공하기 위한 적절한 조치를 취해야 하고, 특히 아동을 특정 대상으로 할 때 더욱 그러해야 한다. 해당 정보는 서면이나 적절한 경우, 전자수단 등 기타 수단을 이용하여 제공되어야 한다. 정보주체가 요청하는 경우, 다른 수단을 통해 정보주체의 신원이 입증되면, 해당 정보는 구두로 제공될 수 있다. Kaminski, M. E. "The right to explanation, explained"(주5), 210면.

#### IV. 결어

미디어 알고리즘에 대한 투명성을 확보하는 것은 중요한 사회적 과제이지만, 설명의무를 통하여 이 과제를 해결하는 것은 간단하지 않으며 상당히 유연하고 기능적 접근이 필요하다. 현재 설명가능하지 않은 인공지능들이 다수라는 사실을 고려하고, 설명을 제공받더라도 이를 이해하기 어려워하는 정보주체의 전문성 부족을 시인하지 않을 수 없다. 그리고 이러한 설명의무를 통하여 성취하고자 하는 투명성의 구체적인 목표가 무엇인지도 정하여야 한다. 가령, 사용하고 있는 알고리즘에 대한 일반적인 불안감을 없애는 것인지, 사용자에게 사용 유무에 대한 결정권을 주고자 함인지, 아니면 알고리즘으로 인하여 발생할 수 있는 권리의 침해에 대응하여 사용자가 어떤 대응을 할 수 있게 함인지를 정하여야 한다. 실제로 추천 서비스에 대하여 제공되는 설명에 대한 만족도를 조사한 연구결과들을 보면, 이용자들은 간단하고 직관적으로 이 추천이 적당한지에 관한 설명을 들었을 때 가장 만족도가 높았다. 그리고 이용자들은 이 설명을 통하여 추천 서비스의 이용 여부를 결정하였다. 설명의무의 정책적 목표인 '투명성'을 어떻게 정의하느냐에 따라 설명의무의 내용도 달라질 수 있다.

알고리즘이 현재 우리에게 주는 실용성, 효율성에 대하여 충분히 고려하되, 알고리즘의 부정적인 모습을 최소화하기 위한 노력이 필요하다. 만약 개인에 대한 설명이 투명성의 제고에 일정한 한계를 가진다면, 다른 시각의 정책에도 관심을 둘 필요가 있다. 실제로 많은 학자들과 가이드라인은 정보주체에게 설명을 하는 것 외에도 투명성을 높일 방안을 모색해왔다. 설명가능한 모델을 개발하는 것, 조직 차원의 관리감독기구를 설치하는 것이나 내부감사(auditing), 품질 보증(quality assurance), 제3자에 의한 관리감독을 철저히 하는 것이 그 예가 될 수 있겠다. 정보주체에 대한 설명이 투명성의 제고에 가지는 현실적 한계를 고려할 때, 아마도 정보주체에 대한 설명과 더불어 이러한 조직적 차원의 노력을 균형있게 병행하는 것이 인공지능이라는 포털 언론의 새로운 편집자가 등장한 지금 우리 앞에 놓인 과제가 아닐까 생각된다.

국내 문헌

김민수, "구글과 네이버·카카오의 'AI 뉴스편집'은 무엇이 다를까?", 노컷뉴스, <https://www.nocutnews.co.kr/news/4968432>

스튜어드 러셀, 피터 노빅, 「인공지능 현대적 접근방식 1」(제3판), 제이펍, 2009

스튜어드 러셀, 피터 노빅, 「인공지능 현대적 접근방식 2」(제3판), 제이펍, 2009

이기범, "네이버, "내년부터 설명가능한 뉴스 추천 서비스 도입", <http://www.bloter.net/archives/326877>

이석구, 네이버, 뉴스편집 영역 AI 추천 '에어스'로 완전 대체한다, 뉴시안, <http://www.newsian.co.kr/news/articleView.html?idxno=34912>

해외 문헌

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018, April). Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In Proceedings of the 2018 CHI conference on human factors in computing systems (p. 582). ACM

Biran, O., & Cotton, C. (2017, August). Explanation and justification in machine learning: A survey. In IJCAI-17 workshop on explainable AI (XAI) (Vol. 8, p. 1).

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 2053951715622512

Burt, A., Leong, B., Shirrell, S., & Wang, X. G. (2018). Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models. In Future of Privacy Forum

Catherine Stupp, 'Commission to Open Probe into Tech Companies' Algorithms next Year' (EurActiv.com, 8 November 2016), <https://www.euractiv.com/section/digital/news/commission-to-open-probe-into-tech-companies-algorithms-next-year/>

Defense Advanced Research Projects Agency(DARPA), Explainable Artificial Intelligence(XAI), Dr. Matt Turek, <https://www.darpa.mil/program/explainable-artificial-intelligence>

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., ... & Wood, A. (2017). Accountability of AI under the law: The role of explanation. arXiv preprint arXiv:1711.01134

Edwards, L., & Veale, M. (2017). Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for. Duke L. & Tech. Rev., 16, 18

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". AI Magazine, 38(3), 50-57

Government Office for Science, 'Artificial Intelligence: An Overview for Policy-Makers' (Government Office for Science 2016), <https://www.gov.uk/government/publications/artificial-intelligence-an-overview-for-policy-makers>

Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01), [https://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=612053](https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053)

Kaminski, M. E. (2019). The right to explanation, explained. Berkeley Tech. LJ, 34, 189

National Science and Technology Council Committee on Technology, 'Preparing for the Future of Artificial Intelligence' (Executive Office of the President 2016), [https://www.whitehouse.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparin\\_g\\_for\\_the\\_future\\_of\\_ai.pdf](https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparin_g_for_the_future_of_ai.pdf)

Partnership on AI, 'Partnership on Artificial Intelligence to Benefit People and Society' (Partnership on Artificial Intelligence to Benefit People and Society, 2016), <https://www.partnershiponai.org/>

Reuters Institute. (2019). Digital News Report 2019

Selbst, A. D., & Powles, J. (2017). Meaningful information and the right to explanation. International Data Privacy Law, 7(4), 233-242

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. International Data Privacy Law, 7(2), 76-99

## 서울대학교 인공지능정책 이니셔티브 안내

서울대학교 인공지능정책 이니셔티브는 인공지능과 관련된 다양한 사회경제적, 법적, 정책적 이슈들을 연구하고 논의하기 위해 시작된 서울대학교 법과경제연구센터의 프로그램입니다. ‘소셜랩(Social Lab)’ 개념을 지향하여, 여러 배경과 관심을 가진 분들 사이의 협업과 지속적인 대화를 추구합니다. 서울대학교 법학전문대학원의 교수와 임용 교수가 함께 이끌고 있습니다.

### 1. 발간물 안내

서울대학교 인공지능정책 이니셔티브의 주요 발간물은 이슈페이퍼와 워킹페이퍼가 있고, 비정기적으로 발간되는 단행본 및 학술행사 자료집 등이 있습니다. 이슈페이퍼와 워킹페이퍼 등의 자료들은 홈페이지를 통해 다운로드 받으실 수 있습니다.

### 2. 행사 안내

서울대학교 인공지능정책 이니셔티브의 주요 행사는 이슈페이퍼를 발표하고 논의하는 행사(상반기 및 하반기 각 1회) 그리고 국내외 연구자들을 초빙하여 진행하는 대규모 국제학술대회(연 1회) 등이 있습니다. 그 이외에 비정기적으로 진행하는 행사들도 있습니다.

### 3. 이슈페이퍼 2019-2

이번 이슈페이퍼는 서울대학교 인공지능정책 이니셔티브의 두 번째 이슈페이퍼로, 2019.11.07. D2 Startup Factory에서 열린 행사에 맞춰 준비되었습니다.