European Parliament

The impact of the General Data Protection Regulation (GDPR) on artificial intelligence

STUDY

Panel for the Future of Science and Technology

EPRS | European Parliamentary Research Service

Scientific Foresight Unit (STOA)

PE 641.530 – June 2020

EN

The impact of the General Data Protection Regulation (GDPR) on artificial intelligence

This study addresses the relationship between the General Data Protection Regulation (GDPR) and artificial intelligence (AI). After introducing some basic concepts of AI, it reviews the state of the art in AI technologies and focuses on the application of AI to personal data. It considers challenges and opportunities for individuals and society, and the ways in which risks can be countered and opportunities enabled through law and technology.

이 연구는 일반데이터보호규정 (GDPR)과 인공 지능 (AI)의 관계를 다룹니다. AI의 기본 개념을 소개한 후 AI 기술의 최신 기술을 검토하고 개인 데이터에 AI를적용하는 데 중점을 둡니다. 개인과 사회에 대한 도전과 기회, 그리고 법과 기술을 통해 위험을 극복할 수 있는 방법과 기회를 고려합니다.

The study then provides an analysis of how AI is regulated in the GDPR and examines the extent to which AI fits into the GDPR conceptual framework. It discusses the tensions and proximities between AI and data protection principles, such as, in particular, purpose limitation and data minimisation. It examines the legal

bases for AI applications to personal data and considers duties of information concerning AI systems, especially those involving profiling and automated decision-making. It reviews data subjects' rights, such as the rights to access, erasure, portability and object.

이 연구는 AI가 GDPR에서 어떻게 규제되는지에 대한 분석을 제공하고 AI가 GDPR 개념적 프레임 워크에 얼마나 적합한 지 조사합니다. AI와 데이터 보호 원칙 간의 긴장과 근접성, 특히 목적 제한 및 데이터 최소화와 같은 주제에 대해 설명합니다. 개인 데이터에 대한 AI 응용 프로그램의 법적 기반을 검토하고 AI 시스템, 특히 프로파일링 및 자동 의사 결정과 관련된 정보에 대한 의무를 고려합니다. 액세스 권한, 삭제, 이식성 및 개체 권한과 같은 데이터 주체의 권한을 검토합니다.

The study carries out a thorough analysis of automated decision-making, considering the extent to which automated decisions are admissible, the safeguard measures to be adopted, and whether data subjects have a right to individual explanations. It then addresses the extent to which the GDPR provides for a preventive risk-based approach, focusing on data protection by design and by default. The possibility to use AI for statistical purposes, in a way that is consistent with the GDPR, is also considered.

이 연구는 자동화된 의사 결정이 허용되는 정도, 채택할 보호 조치 및 데이터 주체가 개별 설명에 대한 권리를 가지고 있는지 여부를 고려하여 자동화된 의사 결정에 대한 철저한 분석을 수행합니다. 그런 다음 GDPR이 예방 위험 기반 접근 방식을 제공하는 정도를 다루며 설계 및 기본적으로 데이터 보호에 중점을 둡니다. GDPR과 일치하는 방식으로 통계 목적으로 AI를 사용할 가능성도 고려됩니다.

The study concludes by observing that AI can be deployed in a way that is consistent with the GDPR, but also that the GDPR does not provide sufficient guidance for controllers, and that its prescriptions need to be expanded and concretised. Some suggestions in this regard are developed.

이 연구는 AI가 GDPR과 일치하는 방식으로 배치될 수 있지만 GDPR이 컨트롤러에 충분한 지침을 제공하지 않으며 처방을 확장하고 구체화해야 한다는 것을 관찰함으로써 결론을 맺습니다. 이와 관련하여 몇 가지 제안이 개발되었습니다.

AUTHOR

The study was led by Professor Giovanni Sartor, European University Institute of Florence, at the request of the Panel for the Future of Science and Technology (STOA) and managed by the Scientific Foresight Unit, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament. It was co-authored by Professor Sartor and Dr Francesca Lagioia, European University Institute of Florence, working under his supervision.

ADMINISTRATOR RESPONSIBLE

Mihalis Kritikos, Scientific Foresight Unit (STOA)

To contact the publisher, please e-mail stoa@ep.europa.eu

LINGUISTIC VERSION

Original: EN

Manuscript completed in June 2020.

PE 641.530
ISBN: 978-92-846-6771-0
doi: 10.2861/293
QA-QA-02-20-399-EN-N

http://www.europarl.europa.eu/stoa (STOA website)
http://www.eprs.ep.parl.union.eu (intranet)
http://www.europarl.europa.eu/thinktank (internet)
http://epthinktank.eu (blog)

Executive summary

AI and big data

In the last decade, AI has gone through rapid development. It has acquired a solid scientific basis and has produced many successful applications. It provides opportunities for economic, social, and cultural development; energy sustainability; better health care; and the spread of knowledge. These opportunities are accompanied by serious risks, including unemployment, inequality, discrimination, social exclusion, surveillance, and manipulation.

지난 10년 동안 AI는 빠른 개발을 거쳤습니다. 탄탄한 과학적 기반을 획득했으며 많은 성공적인 응용 프로그램을 생산했습니다. 경제, 사회 및 문화 개발 기회를 제공합니다. 에너지 지속성; 더 나은 건강 관리; 지식의 확산. 이러한 기회에는 실업, 불평등, 차별, 사회적 배제, 감시 및 조작 등 심각한 위험이 수반됩니다.

There has been an impressive leap forward on AI since it began to

focus on the application of machine learning to mass volumes of data. Machine learning systems discover correlations between data and build corresponding models, which link possible inputs to presumably correct responses (predictions). In machine learning applications, AI systems learn to make predictions after being trained on vast sets of examples. Thus, AI has become hungry for data, and this hunger has spurred data collection, in a self-reinforcing spiral: the development of AI systems based on machine learning presupposes and fosters the creation of vast data sets, i.e., big data. The integration of AI and big data can deliver many benefits for the economic, scientific and social progress. However, it also contributes to risks for individuals and for the whole of society, such as pervasive surveillance and influence on citizens' behaviour, polarisation and fragmentation in the public sphere.

AI가 대량의 데이터에 머신러닝을 적용하는 데 초점을 맞추기 시작한 이후 AI에 대한 비약적인 발전이 있었습니다. 머신러닝 시스템은 데이터 간의 상관 관계를 발견하고 해당 모델을 구축하여 가능한 입력을 연결하여 응답을 예측할 수 있습니다 (예측). 머신러닝 애플리케이션에서 AI 시스템은 방대한 예제 세트에 대한 교육을 받은 후 예측을 학습합니다. 따라서 AI는 데이터를 필요로 하며, 이 필요는 데이터 수집을 촉진했습니다. 머신러닝을 기반으

로 하는 AI 시스템의 개발은 대규모 데이터 세트, 즉 빅 데이터의 생성을 전제하고 육성합니다. AI와 빅 데이터의 통합은 경제, 과학 및 사회 발전에 많은 이점을 제공할 수 있습니다. 그러나 그것은 또한 퍼베이시브 감시와 공공 영역에서의 시민의 행동, 양극화 및 분열에 대한 영향과 같은 개인과 사회 전체의 위험에 기여합니다.

AI and personal data

Many AI applications process personal data. On the one hand, personal data may contribute to the data sets used to train machine learning systems, namely, to build their algorithmic models. On the other hand, such models can be applied to personal data, to make inferences concerning particular individuals.

많은 AI 응용 프로그램이 개인 데이터를 처리합니다. 한편으로, 개인 데이터는 기계학습 시스템을 훈련시키는 데 사용되는 데이터 세트, 즉 알고리즘 모델을 구축하는 데 기여할 수 있습니다. 반면에, 이러한 모델은 개인 데이터에 적용되어 특정 개인에 관한 추론을 할 수 있습니다.

Thanks to AI, all kinds of personal data can be used to analyse, forecast and influence human behaviour, an opportunity that transforms such data, and the outcomes of their processing, into valuable commodities. In particular, AI enables automated decision-making even in domains that require complex choices, based on multiple factors and non -predefined criteria. In many cases, automated predictions and decisions are not only cheaper, but also more precise and impartial than human ones, as AI systems can avoid the typical fallacies of human psychology and can be subject to rigorous controls.

AI 덕분에 모든 종류의 개인 데이터를 사용하여 인간 행동을 분석하고 예측하고 영향을 미치고 그러한 데이터와 처리 결과를 가치있는 상품으로 변환할 수 있습니다. 특히 AI는 여러 요인 및 미리 정의되지 않은 기준에 따라 복잡한 선택이 필요한 도메인에서도 자동 의사 결정을 가능하게 합니다. AI 시스템은 인간 심리학의 전형적인 오류를 피할 수 있고 엄격한 통제를 받을 수 있기 때문에 많은 경우에 있어 자동 예측 및 결정은 인간보다 저렴할 뿐만 아니라 더 정확하고 공정합니다.

However, algorithmic decisions may also be mistaken or

discriminatory, reproducing human biases and introducing new ones. Even when automated assessments of individuals are fair and accurate, they are not unproblematic: they may negatively affect the individuals concerned, who are subject to pervasive surveillance, persistent evaluation, insistent influence, and possible manipulation.

그러나 알고리즘 결정은 착각이나 차별적 일 수 있으며, 인간의 편견을 재현하고 새로운 결정을 도입합니다. 개인에 대한 자동화된 평가가 공정하고 정확하더라도 문제가 되지는 않습니다. 이들은 광범위한 감시, 지속적인 평가, 지속적인 영향 및 가능한 조작의 대상이 되는 관련 개인에게 부정적인 영향을 줄 수 있습니다.

The AI-based processing of vast masses of data on individuals and their interactions has social significance: it provides opportunities for social knowledge and better governance, but it risks leading to the extremes of 'surveillance capitalism' and 'surveillance state'.

개인과 그들의 상호 작용에 대한 방대한 양의 데이터를 AI 기반으로 처리하는 것은 사회적 의미를 지니고 있습니다. 그것은 사회적 지식과 더 나은 거버넌스를 위한 기회를 제공하지만 '감시 자본주의'와 '감시 상태'의 극단을 초래할 위험이 있습니다.

A normative framework

It must be ensured that the development and deployment of AI tools takes place in a socio-technical framework – inclusive of technologies, human skills, organisational structures, and norms – where individual interests and the social good are preserved and enhanced.

AI 도구의 개발 및 배포는 기술, 인간 기술, 조직 구조 및 규범을 포함한 사회-기술적 프레임 워크에서 개인의 이익과 사회적 이익이 보존되고 향상되는 장소에서 이루어 지도록 해야 합니다.

To provide regulatory support for the creation of such a framework, ethical and legal principles are needed, together with sectorial regulations. The ethical principles include autonomy, prevention of harm, fairness and explicability; the legal ones include the rights and social values enshrined in the EU charter, in the EU treaties, as well as in national constitutions. The sectoral regulations involved include first of all data protection law, consumer protection law,

and competition law, but also other domains of the law, such as labour law, administrative law, civil liability etc.

이러한 프레임 워크를 만들기위한 규제 지원을 제공하려면 부문별 규제와 함께 윤리적 및 법적 원칙이 필요합니다. 윤리적 원칙에는 자율성, 피해 방지, 공정성 및 설명 가능성이 포함됩니다. 법적으로는 EU 헌장, EU 조약 및 국가 헌법에 명시된 권리와 사회적 가치가 포함됩니다. 관련된 부문 별 규정에는 우선 모든 데이터 보호법, 소비자 보호법 및 경쟁법 뿐만 아니라 노동법, 행정법, 민사 책임 등과 같은 법률의 다른 영역도 포함됩니다.

The pervasive impact of AI on European society is reflected in the multiplicity of the legal issues it raises.

유럽 사회에 대한 AI의 광범위한 영향은 그것이 제기하는 다양한 법적 문제에 반영됩니다.

To ensure adequate protection of citizens against the risks resulting from the misuses of AI, beside regulation and public enforcement,

the countervailing power of civil society is also needed to detect abuses, inform the public, and activate enforcement. AI-based citizen-empowering technologies can play an important role in this regard, by enabling citizens not only to protect themselves from unwanted surveillance and 'nudging', but also to detect unlawful practices, identify instances of unfair treatment, and distinguish fake and untrustworthy information.

규제 및 공공 집행 외에도 AI의 오용으로 인한 위험으로부터 시민을 적절히 보호하기 위해서는 남용을 감지하고 대중에게 알리고 집행을 활성화하는 데 시민 사회의 상반되는 힘이 필요합니다. AI 기반의 시민 역량 강화 기술은 시민들이 원치 않는 감시 및 '누설'로부터 자신을 보호할 뿐만 아니라 불법 행위를 탐지하고 부당한 대우를 식별하고 가짜와 신뢰할 수 없는 정보를 구별할 수 있게 함으로써 중요한 역할을 할 수 있습니다..

AI is compatible with the GDPR

AI is not explicitly mentioned in the GDPR, but many provisions in the GDPR are relevant to AI, and some are indeed challenged by

the new ways of processing personal data that are enabled by AI. There is indeed a tension between the traditional data protection principles – purpose limitation, data minimisation, the special treatment of 'sensitive data', the limitation on automated decisions –and the full deployment of the power of AI and big data.

AI는 GDPR에 명시적으로 언급되어 있지 않지만 GDPR의 많은 조항은 AI와 관련이 있으며 AI에 의해 가능해진 개인 데이터를 처리하는 새로운 방법으로 인해 실제로 어려움을 겪고 있습니다. 실제로 기존의 데이터 보호 원칙 (목적 제한, 데이터 최소화, '민감한 데이터'의 특수 처리, 자동화된 의사 결정의 제한) 및 AI와 빅 데이터의 완전한 배치 사이에는 긴장이 있습니다.

The latter entails the collection of vast quantities of data concerning individuals and their social relations and processing such data for purposes that were not fully determined at the time of collection. However, there are ways to interpret, apply, and develop the data protection principles that are consistent with the beneficial uses of AI and big data.

후자는 개인과 사회 관계에 관한 방대한 양의 데이터를 수집하고 수집 당시 완전히 결정되지 않은 목적으로 이러한 데이터를 처리합니다. 그러나 AI 및 빅 데이터의 유익한 사용과 일치하는 데이터 보호 원칙을 해석, 적용 및 개발하는 방법이 있습니다.

The requirement of purpose limitation can be understood in a way that is compatible with AI and big data, through a flexible application of the idea of compatibility, which allows for the reuse of personal data when this is not incompatible with the purposes for which the data were originally collected. Moreover, reuse for statistical purposes is assumed to be compatible, and thus would in general be admissible (unless it involves unacceptable risks for the data subject).

목적 제한의 요구 사항은 호환성 아이디어를 유연하게 적용하여 AI 및 빅 데이터와 호환되는 방식으로 이해될 수 있으며, 이는 개인 데이터가 원래 데이터 수집시의 목적과 호환되지 않는 경우에도 개인 데이터를 재사용할 수 있게 합니다. 또한, 통계 목적의 재사용은 호환 가능한 것으로 가정되므로 일반적으로 (데이터 주체에 허용할 수 없는 위험이 없는 한) 허용될 수 있습니다.

The principle of data minimisation can also be understood in such a way as to allow for beneficial applications of AI. Minimisation may require, in some contexts, reducing the 'personality' of the available data, rather than the amount of such data, i.e., it may require reducing, through measures such as pseudonymisation, the ease with which the data can be connected to individuals.

데이터 최소화의 원리는 AI의 유익한 응용을 가능하게 하는 방식으로도 이해될 수 있습니다. 최소화는 일부 상황에서 가용 데이터의 양보다는 가용 데이터의 '개인성'을 감소시킬 것을 요구할 수 있다. 즉, 가명 화와 같은 수단을 통해 데이터가 개인과 쉽게 연결될 수 있는 것을 감소시켜야할 수 도 있다.

The possibility of re-identification should not entail that all re-identifiable data are considered personal data to be minimised. Rather the re-identification of data subjects should be considered as creation of new personal data, which should be subject to all applicable rules. Re-identification should indeed be strictly prohibited unless all conditions for the lawful collection of personal data are met, and it should be compatible with the purposes for which the data were originally collected and subsequently

anonymised.

재식별 가능성은 최소한의 개인정보로 볼 수 있는 모든 재식별 가능한 데이터가 없어야 한다. 오히려 데이터 주체의 재식별은 새로운 개인 데이터의 생성으로 간주되어야 하며, 모든 해당 규칙이 적용되어야 합니다. 합법적인 개인 정보 수집에 대한 모든 조건이 충족되지 않는 한 재식별은 실제로 엄격히 금지되어야 하며, 데이터가 원래 수집된 후 익명이 된 목적과 호환되어야 합니다.

The information requirements established by the GDPR can be met with regard to AI-based processing, even though the complexity of AI application has to be taken into account. The information made available to data subjects should enable them to understand the purpose of each AI-based processing and its limits, even without going into unnecessary technical details.

AI 응용 프로그램의 복잡성을 고려해야 하더라도 GDPR에 의해 설정된 정보 요구 사항은 AI 기반 처리와 관련하여 충족될 수 있습니다. 데이터 주체가 이용할 수 있는 정보는 불필요한 기술적 세부 사항에 들어 가지 않아도 각 AI 기반 처리의 목적과 한계를

이해할 수 있도록 해야 합니다.

The GDPR allows for inferences based on personal data, provided that appropriate safeguards are adopted. Profiling is in principle prohibited, but there are ample exceptions (contract, law or consent). Uncertainties exist concerning the extent to which an individual explanation should be provided to the data subject. It is also uncertain to what extent reasonableness criteria may apply to automated decisions.

GDPR은적절한 보호 조치가 채택된 경우 개인 데이터를 기반으로 추론을 허용합니다. **프로파일링은 원칙적으로 금지되어 있지만 충분한 예외 (계약, 법률 또는 동의)**가 있습니다. 데이터 주제에 개별 설명을 제공해야하는 정도와 관련하여 불확실성이 존재합니다. 또한 자동화된 의사 결정에 어느 정도의 합리성 기준이 적용될 수 있는지도 불확실합니다.

The GDPR provisions on preventive measures, and in particular those concerning privacy by design and by default, do not hinder the development of AI systems, if correctly designed and

implemented, even though they may entail some additional costs. It needs to be clarified which AI applications present high risks and therefore require a preventive data protection assessment, and possibly the preventive involvement of data protection authorities.

예방 조치에 대한 GDPR 조항, 특히 설계 및 기본적으로 개인 정보 보호와 관련된 조항은 AI 시스템의 개발을 방해하지 않습니다. 어떤 AI 응용 프로그램이 위험이 높으며 **예방적 데이터 보호 평가가 필요하며 데이터 보호 기관의 예방적 개입이 필요한지 명확히 해야 합니다.**

Finally, the possibility of using personal data for statistical purposes opens opportunities for the processing of personal data in ways that do not involve the inference of new personal data. Statistical processing requires security measures that are proportionate to the risks for the data subject, and which should include at least pseudonymisation.

마지막으로, 통계 목적으로 개인 데이터를 사용할 가능성은 새로운 개인 데이터의 추론을 포함하지 않는 방식으로 개인 데이터를 처리할 수 있는 기회를 제공합니다. 통계 처리에는 데이터 주체의 위험에 비례하고 최소한 가명화를 포함해야 하는 보안 조치가 필

요합니다.

The GDPR prescriptions are often vague and open-ended

The GDPR allows for the development of AI and big data applications that successfully balance data protection and other social and economic interests, but it provides limited guidance on how to achieve this goal. It indeed abounds in vague clauses and open standards, the application of which often requires balancing competing interests. In the case of AI/big data applications, the uncertainties are aggravated by the novelty of the technologies, their complexity and the broad scope of their individual and social effects.

GDPR을 사용하면 데이터 보호 및 기타 사회적 및 경제적 이익의 균형을 맞추는 AI 및 빅 데이터 응용 프로그램을 개발할 수 있지만 이 목표를 달성하는 방법에 대한 지침은 제한적입니다. 실제로 모호한 조항과 공개 표준이 풍부하며, 그 적용에는 종종 경쟁 이해 관계의 균형이 필요합니다. AI/빅 데이터 애플리케이션의 경우, 기술의 신규성, 복잡성 및 개인 및 사회적 효과의 광범위한 범위에 의해 불확실성이 악화됩니다.

It is true that the principles of risk-prevention and accountability potentially direct the processing of personal data toward a 'positive sum' game, in which the advantages of the processing, when constrained by appropriate risk-mitigation measures, outweigh its possible disadvantages. More over these principles enable experimentation and learning, avoiding the over - and under-inclusiveness issues involved in the applications of strict rules.

위험 예방 및 책임의 원칙은 잠재적으로 개인 데이터의 처리를 ' 양의 합'게임으로 인도하는 것이 사실이며, 여기서 적절한 위험 완화 조치에 의해 제한될 때 처리의 장점이 발생 가능한 단점보다 더 큽니다. 이 원칙들보다 더 많은 것은 실험과 학습을 가능하게 하며, 엄격한 규칙의 적용과 관련된 과도하고 포괄적인 문제를 피합니다.

However, by requiring controllers to rely on such principles, the GDPR offloads the task of establishing how to manage risk and find optimal solutions onto controllers, a task that may be challenging as well as costly. The stiff penalties for non-compliance, when combined with the uncertainty on the requirements for compliance, may constitute a novel risk, which, rather than

incentivising the adoption of adequate compliance measure, may prevent small companies from engaging in new ventures.

그러나 컨트롤러가 이러한 원칙에 의존하도록 요구함으로써 GDPR은 위험 관리 방법을 설정하는 작업과 컨트롤러에 대한 최적의 솔루션을 찾는 작업, 즉 비용이 많이 들고 도전적인 작업을 오프로드합니다. 비준수에 대한 엄격한 처벌은 준수 요구 사항에 대한 불확실성과 결합될 경우적절한 준수 조치의 채택을 장려하기보다는 소기업이 새로운 벤처에 참여하는 것을 방해할 수 있는 새로운 위험을 구성할 수 있습니다.

Thus, the successful application of GDPR to AI-application depends heavily on what guidance data protection bodies and other competent authorities will provide to controllers and data subjects. Appropriate guidance would diminish the cost of legal uncertainty and would direct companies – in particular small ones that mostly need such advice – to efficient and data protection-compliant solutions.

따라서, AI 응용에 GDPR을 성공적으로 적용하는 것은 데이터 보호 기관 및 기타 유관 기관이 컨트롤러 및 데이터 주제에 제공할 지침에 크게 좌우됩니다. 적절한 지침은 법적 불확실성 비용을 감소 시키며 기업, 특히 이러한 조언이 필요한 소규모 기업을 효율

적이고 데이터 보호 준수 솔루션으로 안내합니다.

Some policy indications

The study concludes with the following indications on AI and the processing of personal data.

이 연구는 AI 및 개인 데이터 처리에 대한 다음과 같은 결론을 냅니다.

- The GDPR generally provides meaningful indications for data protection in the context of AI applications.

  GDPR은 일반적으로 AI 응용 프로그램의 맥락에서 데이터 보호에 대한 의미있는 표시를 제공합니다.

- The GDPR can be interpreted and applied in such a way that it does not substantially hinder the application of AI to personal data, and that it does not place EU companies at a disadvantage by comparison with non-European competitors.

  GDPR은 AI가 개인 데이터에적용되는 것을 실질적으로 방해하지 않으며, 유럽 이외의 경쟁사와 비교하여 EU 회사

에 불리한 점이 없도록 해석하고 적용할 수 있습니다.

- Thus, the GDPR does not require major changes in order to address AI applications.

  따라서 GDPR은 AI 응용 프로그램을 다루기 위해 큰 변화가 필요하지 않습니다.

- However, a number of AI-related data-protection issues do not have an explicit answer in the GDPR. This may lead to uncertainties and costs, and may needlessly hamper the development of AI applications.

  그러나 많은 **AI 관련 데이터 보호 문제는 GDPR에 명백한 답변이 없습니다**. 이로 인해 **불확실성과 비용이 발생할 수 있으며 AI 응용 프로그램의 개발을 불필요하게 방해할 수 있습니다**.

- Controllers and data subjects should be provided with guidance on how AI can be applied to personal data consistently with the GDPR, and on the available technologies for doing so. Such guidance can prevent costs linked to legal uncertainty, while enhancing compliance.

  컨트롤러와 데이터 주제는 AI가 GDPR과 일관되게 개인 데이터에 적용되는 방법과 사용 가능한 기술에 대한 지침

을 제공해야 합니다. 이러한 지침은 법적 불확실성과 관련된 비용을 방지하는 동시에 규정 준수를 강화할 수 있습니다.

- Providing guidance requires a multilevel approach, which involves data protection authorities, civil society, representative bodies, specialised agencies, and all stakeholders.

  지침을 제공하려면 데이터 보호 기관, 시민 사회, 대표 기관, 전문 기관 및 모든 이해 관계자가 참여하는 다단계 접근 방식이 필요합니다.

- A broad debate is needed involving not only political and administrative authorities, but also civil society and academia.

  정치 및 행정 당국 뿐만 아니라 시민 사회 및 학계와 관련된 광범위한 토론이 필요합니다.

- This debate needs to address the issues of determining what standards should apply to AI processing of personal data, particularly to ensure the acceptability, fairness and reasonability of decisions on individuals. It should also address what applications are to be barred unconditionally, and which ones may instead be admitted only under

specific circumstances and controls.

이 토론은 개인 데이터의 AI 처리에 적용할 표준을 결정하는 문제, 특히 개인의 결정에 대한 수용 가능성, 공정성 및 합리성을 보장하는 문제를 해결해야 합니다. 또한 어떤 응용 프로그램이 무조건적으로 금지되어야 하며, 특정 상황과 통제 하에서만 승인될 수 있는 응용 프로그램도 다루어야 합니다.

- Discussion of a large set of realistic examples is needed to clarify which AI applications are socially acceptable, under what circumstances and with what constraints. The debate on AI can also be an opportunity to reconsider in depth, with more precision and concreteness, some basic ideas of law and ethics, such as acceptable and practicable conceptions of fairness and non-discrimination.

  어떤 AI 응용 프로그램이 어떤 상황에서 어떤 제약 조건에서 사회적으로 수용 가능한지를 명확히 하기 **위해 현실적인 일련의 현실적인 사례에 대한 논의가 필요**합니다. 인공 지능에 대한 토론은 공정성과 차별 금지에 대한 수용 가능하고 실용 가능한 개념과 같은 법과 윤리의 기본 개념과 같은 보다 정밀하고 구체적으로 심층적으로 재고할 수 있는 기회가 될 수 있습니다.

- Political authorities, such as the European Parliament, the European Commission and the Council should provide general open-ended indications about the values at stake and ways to achieve them.

  유럽 의회, 유럽위원회, 협의회와 같은 정치 당국은 위기에 처한 가치와 이를 달성하는 방법에 대한 일반적인 개방형 표시를 제공해야 합니다.

- Data protection authorities, and in particular the Data Protection Board, should provide controllers with specific guidance on the many issues for which no precise answer can be found in the GDPR. Such guidance can often take the form of soft law instruments designed with dual legal and technical competence, as in the case of Article 29 Working Party opinions.

  데이터 보호 당국, 특히 데이터 보호위원회는 컨트롤러에 GDPR에서 정확한 답변을 찾을 수 없는 많은 문제에 대한 구체적인 지침을 제공해야 합니다. 그러한 지침은 종종 제29조 작업반 의견의 경우와 같이 이중의 법적 및 기술적 역량을 갖도록 설계된 소프트 법률 도구의 형태를 취할 수 있다.

- National Data Protection Authorities should also provide

guidance, in particular when contacted for advice by controllers, or in response to data subjects' queries.

National Data Protection Authorities는 특히 컨트롤러의 조언을 구하거나 데이터 주체의 질문에 대한 답변을 받을 때 지침을 제공해야 합니다.

- The fundamental data protection principles – especially purpose limitation and minimisation – should be interpreted in such a way that they do not exclude the use of personal data for machine learning purposes. They should not preclude the creation of training sets and the construction of algorithmic models, whenever the resulting AI systems are socially beneficial and compliant with data protection rights.

  기본 데이터 보호 원칙, 특히 목적 제한 및 최소화는 머신러닝 목적으로 개인 데이터의 사용을 배제하지 않는 방식으로 해석해야 합니다. 결과 AI 시스템이 사회적으로 이익이 되고 데이터 보호 권한을 준수할 때마다 훈련 세트 작성 및 알고리즘 모델 작성을 배제해서는 안됩니다.

- The use of personal data in a training set, for the purpose of learning general correlations and connection, should be distinguished from their use for individual profiling, which

is about making assessments about individuals.

일반적인 상관 관계 및 연결을 학습하기 위해 **훈련 세트에서 개인 데이터를 사용하는 것은 개인에 대한 평가를 하는 개인 프로파일링에 사용되는 것과 구별** 되어야 합니다.

- The inference of new personal data, as it is done in profiling, should be considered as creation of new personal data, when providing an input for making assessments and decisions. The same should apply to the re-identification of anonymous or pseudonymous data.

  프로파일링에서 수행되는 새로운 개인 데이터의 추론은 평가 및 결정을 위한 입력을 제공할 때 **새로운 개인 데이터의 생성으로 간주**되어야 합니다. 익명 또는 유사 데이터의 재식별에도 동일하게 적용됩니다.

- Guidance is needed on profiling and automated decision-making. It seems that an obligation of reasonableness – including normative and reliability aspects – should be imposed on controllers engaging in profiling, mostly, but not only when profiling is aimed at automated decision-making. Controllers should also be under an obligation to provide individual explanations, to the extent that this is

possible according to the available AI technologies, and reasonable according to costs and benefits. The explanations may be high-level, but they should still enable users to contest detrimental outcomes.

프로파일링 및 자동 의사 결정에 지침이 필요합니다. 규범 및 신뢰성 측면을 포함한 합리성의 의무는 프로파일링이 자동화된 의사 결정을 목표로 할 때 뿐만 아니라 프로파일링에 관여하는 컨트롤러에 적용 되어야하는 것으로 보입니다. 또한 컨트롤러는 이용 가능한 AI 기술에 따라 가능하고 비용과 혜택에 따라 합리적으로 개별 설명을 제공할 의무가 있어야 합니다. 설명은 수준이 높을 수 있지만 사용자는 여전히 해로운 결과에 이의를 제기할 수 있어야 합니다.

- It may be useful to establish obligations to notify data protection authorities of applications involving individualised profiling and decision-making, possibly accompanied with the right to ask for indications on compliance.

  개별 프로파일링 및 의사 결정과 관련된 응용 프로그램에 대해 데이터 보호 기관에 통지해야 할 의무를 설정하는 것이 유용할 수 있으며, 준수에 대한 표시를 요청할 권리가 있습니다.

- The content of the controller's obligation to provide information (and the corresponding rights of the data subject) about the 'logic' of an AI system need to be specified, with appropriate examples, an in relation to different technologies.

  AI 시스템의 '논리'에 대한 정보 (및 데이터 주체의 해당 권리)를 제공해야하는 **컨트롤러의 의무 내용은 적절한 기술과 관련하여 다른 기술과 관련하여 명시**해야 합니다.

- It needs to be ensured that the right to opt out of profiling and data transfers can easily be exercised, through appropriate user interfaces. The same applies to the right to be forgotten.

- 적절한 사용자 인터페이스를 통해 프로파일링 및 데이터 전송을 거부할 수 있는 권한을 쉽게 행사할 수 있어야 합니다. 잊을 권리에 동일하게 적용

- Normative and technological requirements concerning AI by design and by default need to be specified.

  설계 및 기본적으로 AI에 관한 규범적 및 기술적 요구 사항을 지정해야 합니다.

- The possibility of repurposing data for AI applications that

do not involve profiling – scientific and statistical ones – need to be broad, as long as appropriate precautions are in place preventing abuse.

과학적 및 통계적 프로파일링을 포함하지 않는 AI 응용 프로그램의 데이터 용도 변경 가능성은 학대를 방지하기 위한 적절한 예방 조치가 마련되어 있는 한 광범위 해야 합니다.

- Strong measures need to be adopted against companies and public authorities that intentionally abuse the trust of data subjects by using their data against their interests.

  의도적으로 데이터를 사용하여 데이터 주체의 신뢰를 의도적으로 남용하는 회사 및 공공 기관에 대해 강력한 조치를 취해야 합니다.

- Collective enforcement in the data protection domain should be enabled and facilitated.

  데이터 보호 영역에서의 집단 시행이 활성화되고 촉진되어야 합니다.

In conclusion, controllers engaging in AI-based processing should endorse the values of the GDPR and adopt a responsible and risk-

oriented approach. This can be done in ways that are compatible with the available technology and economic profitability (or the sustainable achievement of public interests, in the case of processing by public authorities). However, given the complexity of the matter and the gaps, vagueness and ambiguities present in the GDPR, controllers should not be left alone in this exercise.

결론적으로 AI 기반 처리에 관여하는 컨트롤러는 GDPR의 가치를 인정하고 책임감 있고 리스크 지향적인 접근 방식을 채택해야 합니다. 이것은 이용 가능한 기술 및 경제적 수익성 (또는 공공 당국에 의한 처리의 경우 지속 가능한 공공 이익 달성)과 호환되는 방식으로 수행될 수 있습니다. 그러나 GDPR에 존재하는 사안의 복잡성과 격차, 모호함 및 모호함을 감안할 때, 컨트롤러는 이 연습에서 혼자서는 안됩니다.

Institutions need to promote a broad societal debate on AI applications, and should provide high-level indications. Data protection authorities need to actively engage in a dialogue with all stakeholders, including controllers, processors, and civil society, in order to develop appropriate responses, based on shared values and effective technologies. Consistent application of data protection principles, when combined with the ability to efficiently use AI technology, can contribute to the success of AI applications,

by generating trust and preventing risks.

기관은 AI 애플리케이션에 대한 광범위한 사회적 토론을 장려해야 하며 높은 수준의 적응증을 제공해야 합니다. 데이터 보호 당국은 공유 가치와 효과적인 기술을 기반으로 적절한 대응을 개발하기 위해 컨트롤러, 프로세서 및 시민 사회를 포함한 모든 이해관계자와 대화에 적극적으로 참여해야 합니다. AI 기술을 효율적으로 사용하는 능력과 결합된 데이터 보호 원칙의 일관된 적용은 신뢰를 생성하고 위험을 예방함으로써 AI 응용 프로그램의 성공에 기여할 수 있습니다.

Table of Contents

Table of Contents

3.3.1. 제6조 (1) (a) GDPR : 동의

3.3.2. 제6조 (1) (b-e) GDPR : 필요성

3.3.3. 제6조 (1) (f) GDPR : 정당한 이익

3.3.4. 제6조 (4) GDPR : 용도 변경

3.3.5. 제9조 GDPR : AI 및 특별 범주의 데이터

3.4. AI와 투명성

3.4.1. 제13조 및 제14조 GDPR : 정보 업무

3.4.2. 자동화된 의사 결정에 대한 정보

3.5. AI 및 데이터 주체의 권리

3.5.1. 제15조 GDPR : 접근권한

3.5.2. 제17조 GDPR : 삭제권

3.5.3. 제19조 GDPR : 이동권

3.5.4. 제21조 (1) : 이의 제기권

3.5.5. 제21조 제1 항 및 제2 항 : 프로파일링 및 직접 마케팅에 반대

3.5.6. 제21조 (2). 연구 및 통계 목적의 처리에 반대

3.6. 자동화된 의사 결정

3.6.1. 제22조 (1) GDPR : 자동 결정 금지

3.6.2. 제22 (2) GDPR : 22 (1) 금지에 대한 예외

3.6.3. 제22조 (3) GDPR : 보호 조치

3.6.4. 제22조 (4) GDPR : 자동화된 의사 결정 및 민감한 데이터

Table of figures

1. Introduction

This study aims to provide a comprehensive assessment of the interactions between artificial intelligence (AI) and data protection, focusing on the 2016 EU General Data Protection Regulation (GDPR).

이 연구는 2016 EU 일반데이터보호규정 (GDPR)에 중점을 둔 **인공 지능 (AI)과 데이터 보호 간의 상호 작용에 대한 포괄적인 평가를 제공**하는 것을 목표로 합니다.

Artificial intelligence systems are populating the human and social world in multiple varieties: industrial robots in factories, service robots in houses and healthcare facilities, autonomous vehicles and unmanned aircraft in transportation, autonomous electronic agents in e-commerce and finance, autonomous weapons in the military, intelligent communicating devices embedded in every environment. AI has come to be one of the most powerful drivers of social transformation: it is changing the economy, affecting politics, and reshaping citizens' lives and interactions.

인공 지능 시스템은 공장의 산업용 로봇, 주택 및 의료 시설의 서

비스 로봇, 자율 주행 차량 및 무인 항공기 운송, 전자 상거래 및 금융의 자율적 전자 에이전트, 군대의 자율적 무기 등 다양한 종류의 인간 및 사회 세계를 채우고 있습니다. 모든 환경에 내장된 지능형 통신 장치. 인공 지능은 경제를 변화시키고 정치에 영향을 미치며 시민의 삶과 상호 작용을 변화시키는 사회 변혁의 가장 강력한 원동력 중 하나가 되었습니다.

Developing appropriate policies and regulations for AI is a priority for Europe, since AI increases opportunities and risks in ways that are of the greatest social and legal importance. AI may enhance human abilities, improve security and efficiency, and enable the universal provision of knowledge and skills. On the other hand, it may increase opportunities for control, manipulation, and discrimination; disrupt social interactions; and expose humans to harm resulting from technological failures or disregard for individual rights and social values.

AI가 사회적, 법적으로 가장 중요한 방식으로 기회와 위험을 증가시키기 때문에 AI에 대한적절한 정책과 규정을 개발하는 것이 유럽의 우선 순위입니다. AI는 인적 능력을 향상시키고 보안 및 효율성을 향상시키며 지식과 기술을 보편적으로 제공할 수 있습니다. 반면에 통제, 조작 및 차별의 기회를 증가시킬 수 있습니다. 사회적 상호 작용을 방해; 기술적 실패로 인해 인간이 피해를 입

거나 개인의 권리와 사회적 가치를 무시하는 행위.

A number of concrete ethical and legal issues have already emerged in connection with AI in several domains, such as civil liability, insurance, data protection, safety, contracts and crimes. Such issues acquire greater significance as more and more intelligent systems leave the controlled and limited environments of laboratories and factories and share the same physical and virtual spaces with humans (internet services, roads, skies, trading on the stock exchange, other markets, etc.).

민사 책임, 보험, 데이터 보호, 안전, 계약 및 범죄와 같은 여러 영역에서 AI와 관련하여 이미 많은 구체적인 윤리적 및 법적 문제가 제기되었습니다. 이러한 문제는 점점 더 지능적인 시스템이 통제된 제한된 실험실과 공장 환경을 떠나 인간 (인터넷 서비스, 도로, 하늘, 증권 거래소, 기타 시장 등)과 동일한 물리적 공간과 가상 공간을 공유함에 따라 더 큰 중요성을 얻습니다.).

Data protection is at the forefront of the relationship between AI and the law, as many AI applications involve the massive processing of personal data, including the targeting and personalised treatment of individuals on the basis of such data. This explains

why data protection has been the area of the law that has most engaged with AI, although other domains of the law are involved as well, such as consumer protection law, competition law, antidiscrimination law, and labour law.

많은 AI 응용 프로그램이 이러한 데이터를 기반으로 개인을 대상으로 하고 개인화된 치료를 포함하여 개인 데이터의 대규모 처리를 포함하기 때문에 데이터 보호는 AI와 법률 간의 관계의 최전선에 있습니다. 소비자 보호법, 경쟁법, 차별 금지법 및 노동법과 같은 법률의 다른 영역도 관련되어 있지만, 데이터 보호가 AI와 가장 관련이 있는 법의 영역 인 이유를 설명합니다.

This study will adopt an interdisciplinary perspective. Artificial intelligence technologies will be examined and assessed on the basis of most recent scientific and technological research, and their social impacts will be considered by taking account of an array of approaches, from sociology to economics and psychology. A normative perspective will be provided by works in sociology and ethics, and in particular information, computer, and machine ethics. Legal aspects will be analysed by reference to the principles and rules of European law, as well as to their application in national contexts. The report will focus on data protection and the GDPR, though it will also consider how data protection shares with other

domains of the law the task of addressing the opportunities and risks that come with AI.

이 연구는 학제 간 관점을 채택할 것입니다. 인공 지능 기술은 가장 최근의 과학 및 기술 연구를 기반으로 조사 및 평가될 것이며, 사회학에서 경제학 및 심리학에 이르기까지 다양한 접근 방식을 고려하여 사회적 영향을 고려할 것입니다. 규범적 관점은 사회학 및 윤리 분야, 특히 정보, 컴퓨터 및 기계 윤리 분야의 작업에 의해 제공됩니다. 법적 측면은 유럽 법의 원칙과 규칙은 물론 국가적 맥락에서의 적용에 따라 분석될 것입니다. 이 보고서는 데이터 보호와 GDPR에 중점을 둘 것이지만, 데이터 보호가 법의 다른 영역과 어떻게 AI와 함께 제공되는 기회와 위험을 해결하는 작업을 공유하는지 고려할 것입니다.

2. AI and personal data

This section introduces the technological and social background of the study, namely, the development of AI and its connections with the processing of personal and other data. First the concept of AI will be introduced (Section 2.1), then the parallel progress of AI and large-scale data processing will be discussed (Section 2.2), and finally, the analysis will turn to the relation between AI and the

processing of personal data (Section 2.3).

이 섹션에서는 연구의 기술 및 사회적 배경, 즉 AI 개발 및 개인 및 기타 데이터 처리와의 관련성을 소개합니다. 먼저 AI의 개념을 소개하고 (2.1 절) AI의 병렬 진행과 대규모 데이터 처리에 대해 논의하고 (2.2 절), 마지막으로 분석은 AI와 개인의 처리 사이의 관계로 전환됩니다 데이터 (섹션 2.3).

2.1. The concept and scope of AI

The concept of AI will be introduced, as well as its connections with the robotics and algorithms.

AI의 개념과 로봇 및 알고리즘과의 연결이 소개됩니다.

2.1.1. A definition of AI

The broadest definition of artificial intelligence (AI) characterises it as the attempt to build machines that 'perform functions that require intelligence when performed by people.' 1 A more

elaborate notion has been provided by the High Level Expert Group on AI (AI HLEG), set up by the EU Commission:

인공 지능 (AI)의 가장 넓은 정의는 그것을 '사람들이 수행할 때 지능이 필요한 기능을 수행하는 기계'를 구축하려는 시도로 특징 지었다. 1 EU 집행위원회가 설립한 AI HLEG (High Level Expert Group)에 의해보다 정교한 개념이 제공되었습니다.

> ***Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.2***
>
> ***인공 지능 (AI) 시스템은 복잡한 목표가 주어지면 데이터 수집을 통해 환경을 인식하고 수집된 구조적 또는 비 구조***

적 데이터를 해석하여 추론함으로써 물리적 또는 디지털 차원에서 행동하는 인간이 설계한 소프트웨어 (및 하드웨어) 시스템입니다. 이 데이터에서 파생된 지식 또는 정보 처리 및 주어진 목표를 달성하기위한 최선의 조치 결정. AI 시스템은 기호 규칙을 사용하거나 숫자 모델을 학습할 수 있으며 환경이 이전 작업의 영향을 받는 방식을 분석하여 동작을 조정할 수도 있습니다.2

This definition can be accepted with the proviso that most AI systems only perform a fraction of the activities listed in the definition: pattern recognition (e.g., recognising images of plants or animals, human faces or attitudes), language processing (e.g., understanding spoken languages, translating from one language into another, fighting spam, or answering queries), practical suggestions (e.g., recommending purchases, purveying information, performing logistic planning, or optimising industrial processes), etc. On the other hand, some systems may combine many such capacities, as in the example of self-driving vehicles or military and care robots.

대부분의 AI 시스템은 패턴 인식 (예 : 식물 또는 동물의 이미지 인식, 사람의 얼굴 또는 태도 인식), 언어 처리 (예 : 구어 이해)와 같이 대부분의 AI 시스템이 정의에 나열된 활동의 일부만 수행한

다는 조건으로 이 정의를 수용할 수 있습니다. ,한 언어에서 다른 언어로 번역, 스팸 퇴치 또는 쿼리에 응답), 실제 제안 (예 : 구매 권장, 정보 제공, 물류 계획 수행 또는 산업 프로세스 최적화) 등. 자율 주행 차량 또는 군사 및 관리 로봇의 예와 같은 용량.

The High-Level Expert Group characterises the scope of research in AI as follows:

고급 전문가 그룹은 AI의 연구 범위를 다음과 같이 특성화 합니다.

> ***As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).***
>
> ***과학적 학문으로서 AI에는 머신러닝 (딥 러닝 및 강화 학***

습이 구체적인 예임), 머신 추론 (계획, 스케줄링, 지식 표현 및 추론, 검색 및 최적화 포함)과 같은 몇 가지 접근 방식과 기법이 포함됩니다. 로봇 공학 (제어, 인식, 센서 및 액츄에이터 뿐만 아니라 다른 모든 기술을 사이버 물리 시스템에 통합).

To this definition, we could also possibly add also communication, and particularly the understanding and generation of language, as well as the domains of perception and vision.

이러한 정의에, 우리는 또한 의사 소통, 특히 언어의 이해와 생성 뿐만 아니라 인식과 비전의 영역을 추가할 수 있습니다.

1 Kurzweil (1990, 14), Russel and Norvig (2016, Section 1.1).
2 AI-HLEG (2019).

2.1.2. AI and robotics

AI constitutes the core of robotics, the discipline that aims to build 'physical agents that performs tasks by manipulating the physical world.' 3 The High-Level Expert Group describes robotics as follows

AI는 '물리적 세계를 조작하여 작업을 수행하는 물리 에이전트를 구축하는 것을 목표로 하는 학문 인 로봇 공학의 핵심을 구성합니다. 3 고급 전문가 그룹은 로봇 공학을 다음과 같이 설명합니다.

> ***Robotics can be defined as 'AI in action in the physical world' (also called embodied AI). A robot is a physical machine that has to cope with the dynamics, the uncertainties and the complexity of the physical world. Perception, reasoning, action, learning, as well as interaction capabilities with other systems are usually integrated in the control architecture of the robotic system. In addition to AI, other disciplines play a role in robot design and operation, such as mechanical engineering and control theory. Examples of robots include robotic manipulators, autonomous vehicles (e.g. cars, drones, flying taxis), humanoid robots, robotic vacuum cleaners, etc.***
>
> ***로봇 공학은 '실제 세계에서 활동중인 AI'(구체화된 AI라고도 함)로 정의할 수 있습니다. 로봇은 물리적 세계의 역학, 불확실성 및 복잡성에 대처해야하는 물리적 기계입니다. 로봇 시스템의 제어 아키텍처에는 일반적으로 인식, 추론, 행동, 학습 및 다른 시스템과의 상호 작용 기능이 통합되***

어 있습니다. AI 외에도 기계 공학 및 제어 이론과 같은 다른 분야에서도 로봇 설계 및 운영에 중요한 역할을 합니다. 로봇의 예로는 로봇 조작기, 자율 주행 차 (예 : 자동차, 드론, 비행 택시), 인간형 로봇, 로봇 식 진공 청소기 등이 있습니다.

In this report, robotics will not be separately addressed since embodied and disembodied AI systems raise similar concerns when addressed from the perspective of GDPR: in both cases personal data are collected, processed, and acted upon by intelligent system. Moreover, also software systems may have access to sensor on the physical world (e.g., cameras) or govern physical devices (e.g., doors, lights, etc.). This fact does not exclude that the specific types of interaction that may exists, or will exists, between humans and physical robots – e.g., in the medical or care domain– may require specific considerations and regulatory approaches also in the data protection domain.

이 보고서에서는 구체화 및 구현되지 않은 인공 지능 시스템이 GDPR의 관점에서 다룰 때 비슷한 문제를 제기하기 때문에 로봇 공학을 별도로 다루지 않을 것입니다. 두 경우 모두 개인 데이터가 지능 시스템에 의해 수집, 처리 및 실행됩니다. 또한 소프트웨어 시스템은 물리적 환경 (예 : 카메라)의 센서에 액세스하거나

물리적 장치 (예 : 문, 조명 등)를 제어할 수 있습니다. 이러한 사실은 인간과 물리적 로봇 사이에 존재하거나 존재하는 특정 유형의 상호 작용 (예 : 의료 또는 관리 도메인)이 데이터 보호 영역에서도 특정 고려 사항과 규제 접근 방식을 요구할 수 있다는 것을 배제하지 않습니다.

2.1.3. AI and algorithms

The term 'algorithm' is often used to refer to AI applications, e.g., through locutions such 'algorithmic decision-making.' However, the concept of an algorithm is more general that the concept of AI, since it includes any sequence of unambiguously defined instructions to execute a task, particularly but not exclusively through mathematical calculations.4 To be executed by a computer system, algorithms have to be expressed through programming languages, thus becoming machine-executable software programs.

'알고리즘'이라는 용어는 예를 들어 '알고리즘 의사 결정'과 같은 명령을 통해 AI 응용 프로그램을 지칭하는 데 종종 사용됩니다. 그러나 알고리즘의 개념은 AI의 개념보다 일반적입니다. 특히 수학 계산을 통해서만 태스크를 실행하기 위해 명확하게 정의된 명령 시퀀스가 포함되어 있기 때문입니다.4 컴퓨터 시스템에서 실행

하려면, 알고리즘은 프로그래밍 언어를 통해 표현되어야 하므로 머신 실행 가능한 소프트웨어 프로그램이 됩니다.

Algorithms can be very simple, specifying, for instance, how to arrange lists of words in alphabetical order or how to find the greatest common divisor between two numbers (such as the so-called Euclidean algorithm). They can also be very complex, such as algorithms for file encryption, the compression of digital files, speech recognition, or financial forecasting. Obviously, not all algorithms involve AI, but every AI system, like any computer system, includes algorithms, some dealing with tasks that directly concern AI functions.

예를 들어, 알파벳 순서로 단어 목록을 정렬하는 방법 또는 두 숫자 사이에서 가장 큰 공약수를 찾는 방법 (예 : 소위 유클리드 알고리즘)을 지정하는 알고리즘은 매우 간단할 수 있습니다. 파일 암호화 알고리즘, 디지털 파일 압축, 음성 인식 또는 재무 예측과 같이 매우 복잡할 수도 있습니다. 분명히 모든 알고리즘에 AI가 포함되는 것은 아니지만 컴퓨터 시스템과 마찬가지로 모든 AI 시스템에는 AI 기능과 직접 관련된 작업을 처리하는 알고리즘이 포함되어 있습니다.

AI algorithms may involve different kinds of epistemic or practical reasoning (detecting patterns and shapes, applying rules, making forecasts or plans), as well different ways of learning.5 In the latter case the system can enhance itself by developing new heuristics (tentative problem-solving strategies), modifying its internal data, or even generating new algorithms. For instance, an AI system for e-commerce may apply discounts to consumers meeting certain conditions (apply rules), provide recommendations (e.g., learn and use correlations between users' features and their buying habits), optimise stock management (e.g., develop and deploy the best trading strategies).

인공 지능 알고리즘에는 다양한 종류의 전염병 또는 실제 추론 (패턴 및 형태 감지, 규칙적용, 예측 또는 계획 수립) 및 다양한 학습 방법이 포함될 수 있습니다.5 후자의 경우 새로운 휴리스틱 (잠정적 문제)을 개발하여 시스템 자체를 향상시킬 수 있습니다. - 해결 전략), 내부 데이터 수정 또는 새로운 알고리즘 생성 예를 들어, 전자 상거래를 위한 AI 시스템은 특정 조건 (규정적용)을 충족하는 소비자에게 할인을 적용하고, 추천 (예 : 사용자 기능과 구매 습관 간의 상관 관계를 배우고 사용)을 제공하고 재고 관리를 최적화합니다 (예 : 개발 및 개발) 최고의 거래 전략을 배치하십시오).

Though an AI system includes many algorithms, it can also be viewed as a single complex algorithm, combining the algorithms performing its various functions, as well as the top algorithms that orchestrate the system's functions by activating the relevant lower-level algorithms. For instance, a bot that answers queries in natural language will include an orchestrated combination of algorithms to detect sounds, capture syntactic structures, retrieve relevant knowledge, make inferences, generate answers, etc.

AI 시스템에는 많은 알고리즘이 포함되어 있지만 다양한 기능을 수행하는 알고리즘과 관련 하위 레벨 알고리즘을 활성화하여 시스템 기능을 조정하는 최상위 알고리즘을 결합하여 단일 복잡한 알고리즘으로 볼 수도 있습니다. 예를 들어, 자연 언어로 쿼리에 응답하는 봇에는 소리 감지, 구문 구조 캡처, 관련 지식 검색, 추론, 답변 생성 등을 위한 조정된 알고리즘 조합이 포함됩니다.

3 Russell and Norvig (2016).

4 Harel (2004).

5 According to Russel and Norvig (2016, 693), 'an agent is learning if it improves its performance on future tasks after making observations about the world'.

In a system that is capable of learning, the most important component will not be the learned algorithmic model, i.e., the algorithms that directly execute the tasks assigned to the system (e.g., making classifications, forecasts, or decisions) but rather the learning algorithms that modify the algorithmic model so that it better performs its function. For instance, in a classifier system that recognises images through a neural network, the crucial element is the learning algorithm (the trainer) that modifies the internal structure of the algorithmic model (the trained neural network) by changing it (by modifying its internal connections and weights) so that it correctly classifies the objects in its domain (e.g., animals, sounds, faces, attitudes, etc.).

학습할 수 있는 시스템에서 가장 중요한 요소는 학습된 알고리즘 모델, 즉 시스템에 할당된 작업 (예 : 분류, 예측 또는 결정)을 직접 실행하는 알고리즘이 아니라 학습 알고리즘입니다. 알고리즘 모델을 수정하여 기능을 더 잘 수행합니다. 예를 들어 신경망을 통해 이미지를 인식하는 분류기 시스템에서 중요한 요소는 알고리즘 모델 (훈련된 신경망)의 내부 구조를 변경하여 (내부 연결을 수정하여) 학습 알고리즘 (트레이너)입니다. 무게)) 도메인에 있는 대상 (예 : 동물, 소리, 얼굴, 태도 등)을 올바르게 분류합니다.

2.1.4. Artificial intelligence and big data

The term big data identifies vast data sets that it is difficult to manage using standard techniques, because of their special features, the so-called **three** V's: huge Volume, high Velocity and great Variety.

빅 데이터라는 용어는 특수한 기능, 이른바 3 개의 V (큰 볼륨, 높은 속도 및 다양한)로 인해 표준 기술을 사용하여 관리하기 어려운 방대한 데이터 세트를 식별합니다.

Other features associated to big data are low **Veracity** (high possibility that at least some data are inaccurate), and high Value. Such data can be created by people, but most often they are collected by machines, which capture information from the physical word (e.g., street cameras, sensors collecting climate information, devices for medical testing, etc.), or from computer-mediated activities (e.g., systems recording transactions or tracking online behaviour etc.).

빅 데이터와 관련된 다른 기능으로는 정확성 (적어도 일부 데이터가 정확하지 않을 가능성이 높음) 및 Value가 높습니다. 이러한

데이터는 사람이 만들 수 있지만 대부분 물리적인 단어 (예 : 거리 카메라, 기후 정보를 수집하는 센서, 의료 테스트 장치 등) 또는 컴퓨터 매개 활동에서 정보를 캡처하는 컴퓨터에서 수집합니다. (예 : 거래를 기록하거나 온라인 행동을 추적하는 시스템 등).

From a social and legal perspective what is most relevant in very large data sets, and which makes them 'big data' from a functional perspective, is the possibility of using such data sets for analytics, namely, for discovering correlations and making predictions, often using AI techniques, as we shall see when discussing machine learning.6 In particular, the connection with analytics and AI makes big data specifically relevant to data protection.7

사회적 및 법적 관점에서 매우 큰 데이터 세트에서 가장 관련성이 있고 기능적 관점에서 '빅 데이터'로 만드는 것은 분석을 위해, 즉 상관 관계를 발견하고 예측하기 위해 이러한 데이터 세트를 사용할 가능성입니다. 머신러닝을 논의할 때 알 수 있듯이 AI 기술을 사용합니다.6 특히 분석 및 AI와의 연결은 빅 데이터를 데이터 보호와 특별히 관련시킵니다.7

Big data can concern the non-human physical world (e.g. environmental, biological, industrial, and astronomical data), as well

as humans and their social interactions (e.g., data on social networks, health, finance, economics or transportation). Obviously, only the second kind of data is relevant to this report.

빅 데이터는 비인간 물리적 세계 (예 : 환경, 생물학적, 산업 및 천문학적 데이터)는 물론 인간과 사회적 상호 작용 (예 : 소셜 네트워크, 건강, 금융, 경제 또는 교통 관련 데이터)과 관련될 수 있습니다. 분명히 두 번째 종류의 데이터 만이 보고서와 관련이 있습니다.

2.2. AI in the new millennium

Over the last decades, AI has gone through a number of ups and downs, excessive expectations being followed by disillusion (the so-called AI winters).8 In recent years, however, there is no doubt that AI has been hugely successful. On the one hand, a solid interdisciplinary background has been constructed for AI research: the original core of computing, mathematics, and logic has been extended with models and insights from a number of other disciplines, such as statistics, economics, linguistics, neurosciences, psychology, philosophy, and law. On the other hand, an array of successful applications has been built, which have already entered

our daily lives: voice, image, and face recognition; automated translation; document analysis; question-answering; games; high-speed trading; industrial robotics; autonomous vehicles; etc.

지난 수십년 동안 AI는 여러 가지 기복을 겪었고, 과도한 기대에 뒤이어 환멸 (소위 AI winters)이 뒤따랐습니다. 한편, AI 연구를 위해 견고한 학제 간 배경이 구축되었습니다. 컴퓨팅, 수학 및 논리의 원래 핵심은 통계, 경제학, 언어학, 신경 과학, 심리학, 철학 및 법. 다른 한편으로, 우리의 일상에 이미 들어온 성공적인 응용 프로그램의 배열은 이미 음성, 이미지 및 얼굴 인식; 자동 번역; 문서 분석; 질문 답변; 계략; 고속 거래; 산업용 로봇 공학; 자율주행차; 기타

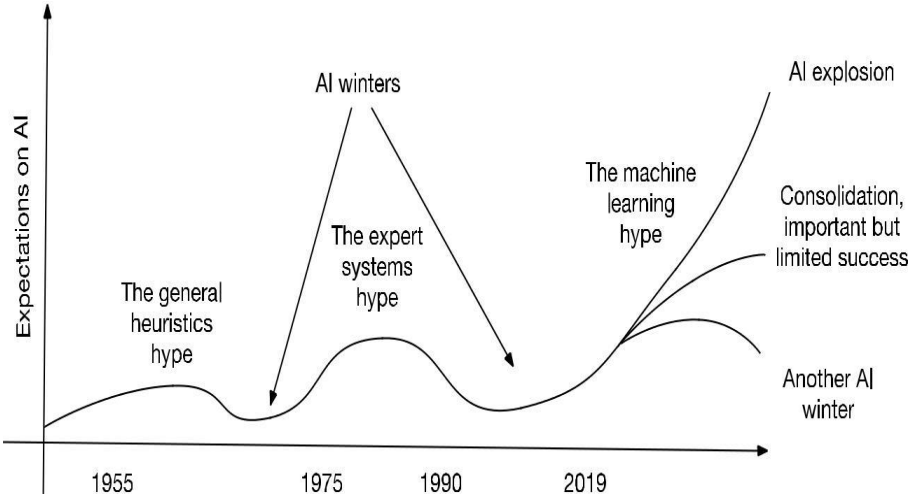

Figure 1 – Hypes and winters of AI

Based on the current successes, it is most likely that current successful applications will not only be consolidate, but will be accompanied by further growth, following probably the middle path indicated in Figure 1.

현재의 성공을 바탕으로 현재 성공한 응용 프로그램은 통합될 뿐만 아니라 그림 1에 표시된 중간 경로에 따라 추가 성장이 동반될 가능성이 높습니다.

6 See Mayer-Schoenberger and Cukier (2013, 15).
7 Hildebrandt (2014)
8 Nilsson (2010).

2.2.1. Artificial general and specific intelligence

AI research usually distinguishes two goals: 'artificial general intelligence,' also known as 'strong AI,' and 'artificial specialised intelligence,' also known as 'weak AI.'

AI 연구는 일반적으로 '강력한 AI'라고도 하는 '인공 일반 지능'과 '약한 AI'라고도 하는 '인공 특수 지능'이라는 두 가지 목표를 구분합니다.

Artificial general intelligence pursues the ambitious objective of developing computer systems that exhibit most human cognitive skills, at a human or even a superhuman level.9 Artificial specialised intelligence pursues a more modest objective, namely, the construction of systems that, at a satisfactory level, are able to engage in specific tasks requiring intelligence.

인공 일반 지능은 인간 또는 초 인간 수준에서 대부분의 인간인지 기술을 나타내는 컴퓨터 시스템을 개발하려는 야심 찬 목표를 추구합니다.9 인공 전문 지능은보다 적당한 목표, 즉 만족스러운 수준에서 시스템을 구축합니다. 지능이 필요한 특정 작업에 참여할 수 있습니다.

The future emergence of a general artificial intelligence is already raising serious concerns. A general artificial intelligence system may improve itself at an exponential speed and quickly become superhuman; through its superior intelligence it may then acquire capacities beyond human control.10 In relation to self-improving artificial intelligence, humanity may find itself in a condition of inferiority similar to that of animals in relation to humans. Some leading scientists and technologists (such as Steven Hawking, Elon Musk, and Bill Gates) have argued for the need to anticipate this existential risk,11 adopting measures meant to prevent the creation

of general artificial intelligence or to direct it towards human-friendly outcomes (e.g., by ensuring that it endorses human values and, more generally, that it adopts a benevolent attitude). Conversely, other scientists have looked favourably on the birth of an intelligence meant to overcome human capacities. In an AI system's ability to improve itself could lie the 'singularity' that will accelerate the development of science and technology, so as not only to solve current human problems (poverty, underdevelopment, etc.), but also to overcome the biological limits of human existence (illness, aging, etc.) and spread intelligence in the cosmos.12

일반적인 인공 지능의 미래 출현은 이미 심각한 문제를 제기하고 있습니다. 일반적인 인공 지능 시스템은 기하 급수적으로 향상되어 빠르게 초 인간이 될 수 있다. 우수한 지능을 통해 인간이 통제할 수 없는 능력을 습득할 수 있습니다.10 자기 개선 인공 지능과 관련하여 인류는 인간과 관련하여 동물과 유사한 열등한 상태에 있을 수 있습니다. 스티븐 호킹 (Steven Hawking), 엘론 머스크 (Elon Musk), 빌 게이츠 (Bill Gates)와 같은 일부 주요 과학자 및 기술자들은 이 존재 위험을 예상할 필요가 있다고 주장했다. (예를 들어, 그것이 인간의 가치를 지지하고 보다 일반적으로는 자비로운 태도를 취하도록 보장함으로써). 반대로, 다른 과학자들은 인간의 능력을 극복하기위한 지능의 탄생을 호의적으로 보았습니다. 인공 지능 시스템의 자체 발전 능력에서 과학과 기술의

발전을 가속화하는 '단일성'이 존재하여 현재의 인간 문제 (빈곤, 저개발 등)를 해결하고 생물학적 한계를 극복할 수 있습니다. 인간의 존재 (질병, 노화 등)와 우주에서 지능을 확산시킵니다.12

9 Bostrom (2014)

10 Bostrom (2014). This possibility was anticipated by Turing ([1951] 1966).

11 Parkin (2015).

12 See Kurzweil (2005) and Tegmark (2017).

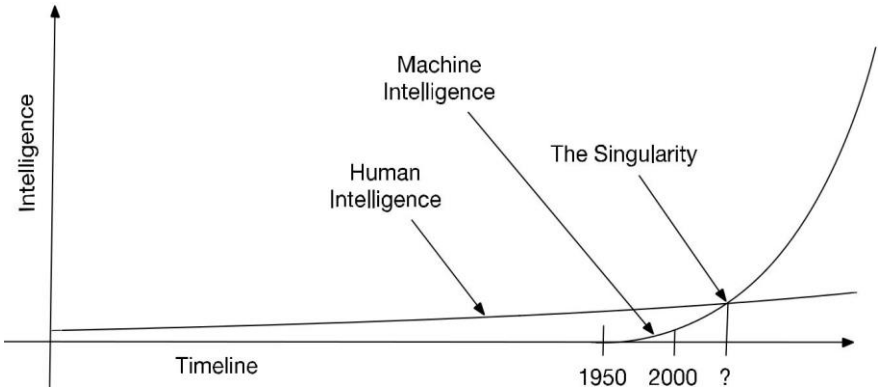

Figure 2 – General AI: The singularity

The risks related to the emergence of an 'artificial general intelligence' should not be underestimated: this is, on the contrary, a very serious problem that will pose challenges in the future. In fact, as much as scientists may disagree on whether and when 'artificial general intelligence,' will come into existence, most of

them believe that this objective will be achieved within the end of this century.13 In any case, it is too early to approach 'artificial general intelligence' at a policy level, since it lies decades ahead, and a broader experience with advanced AI is needed before we can understand both the extent and proximity of this risk, and the best ways to address it.

'인공 일반 지능'의 출현과 관련된 위험을 과소 평가해서는 안됩니다. 이는 미래에 도전이 될 매우 심각한 문제입니다. 실제로 과학자들이 '인공 일반 지능'이 언제 생길지에 대해 의견이 일치하지 않는 한, 대부분의 사람들은 이 목표가 금세기 말에 달성될 것이라고 믿고 있다. 정책 수준에서 '인공 일반 지능'에 접근하려면 수십년이 걸리기 때문에 이 위험의 범위와 근접성, 그리고 이를 해결하는 최선의 방법을 모두 이해하기 위해서는 고급 AI에 대한 광범위한 경험이 필요합니다.

Conversely, 'artificial specialised intelligence' is already with us, and is quickly transforming economic, political, and social arrangements, as well as interactions between individuals and even their private lives. The increase in economic efficiency already is reality (see Figure 2), but AI provides further opportunities: economic, social, and cultural development; energy sustainability; better health care; and the spread of knowledge. In the very recent White Paper by

the European Commission14 it is indeed affirmed that AI.

반대로, '인공 전문 지능'은 이미 우리와 함께 있으며 경제, 정치 및 사회적 준비 뿐만 아니라 개인과 개인 생활 간의 상호 작용을 빠르게 변화시키고 있습니다. 경제 효율성의 증가는 이미 현실이지만 (그림 2 참조) AI는 경제, 사회 및 문화 개발; 에너지 지속성; 더 나은 건강 관리; 지식의 확산. 유럽위원회 (European Commission)의 최근 백서 (14)에서 실제로 AI가 확인되었다.

> ***will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine.***
>
> ***건강 관리 개선 (예 : 진단의 정확성 향상, 질병 예방 개선), 농업 효율성 향상, 기후 변화 완화 및 적응에 기여, 예측 유지 보수를 통한 생산 시스템의 효율성 개선, 유럽인의 보안 강화를 통해 우리의 삶을 변화시킬 것입니다. 많은 다른 방식으로 우리는 상상하기 시작합니다.***

13 A poll among leading AI scientists can be found in Bostrom (2014).
14 White Paper 'On artificial intelligence - A European approach to excellence and trust', Brussels, 19.2.2020 COM(2020) 65 final.

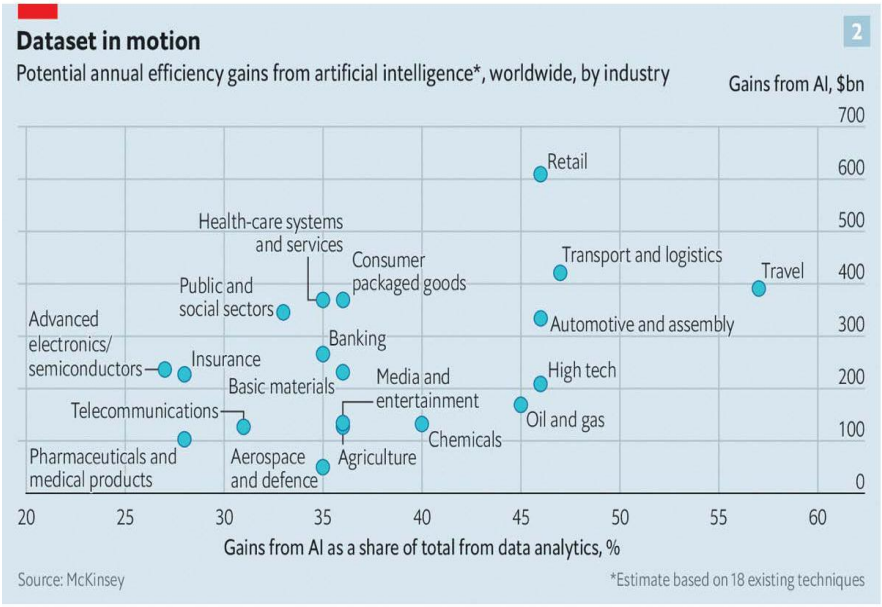

Figure 3 – Efficiency gains from AI

The opportunities offered by AI are accompanied by serious risks, including unemployment, inequality, discrimination, social exclusion, surveillance, and manipulation. It has indeed been claimed that AI should contribute to the realisation of individual

and social interests, and that it should not be 'underused, thus creating opportunity costs, nor overused and misused, thus creating risks.'15 In the just mentioned Commission's White paper, it is indeed observed that the deployment of AI

AI가 제공하는 기회에는 실업, 불평등, 차별, 사회적 배제, 감시 및 조작 등 심각한 위험이 따릅니다. 실제로 AI는 개인 및 사회적 이익을 실현하는 데 기여해야 하며, '사용률이 낮아서 기회 비용이 발생하거나 남용 및 오용되어 위험을 초래해서는 안된다'고 주장했다.15 방금 언급 한위원회의 백서에서 실제로 AI 배포가 관찰되었습니다.

> ***entails a number of potential risks, such as opaque decision-making, gender-based or other kinds of discrimination, intrusion in our private lives or being used for criminal purposes.***
>
> ***불투명한 의사 결정, 성별 기반 또는 기타 종류의 차별, 사생활 침해 또는 범죄 목적으로 사용되는 것과 같은 여러 가지 잠재적 위험이 수반됩니다.***

Because the need has been recognised to counter these risks, while

preserving scientific research and the beneficial uses of AI, a number of initiatives have been undertaken in order to design an ethical and legal framework for 'human-centred AI.' Already in 2016, the White House Office of Science and Technology Policy (OSTP), the European Parliament's Committee on Legal Affairs, and, in the UK, the House of Commons' Science and Technology Committee released their initial reports on how to prepare for the future of AI. 16 Multiple expert committees have subsequently produced reports and policy documents. Among them, the High-Level Expert Group on artificial intelligence appointed by the European Commission, the expert group on AI in Society of the Organisation for Economic Co-operation and Development (OECD), and the select committee on artificial intelligence of the United Kingdom (UK) House of Lords.17

과학적 연구와 AI의 유익한 사용을 보존하면서 이러한 위험에 대응할 필요성이 인식되어 왔기 때문에 '인간 중심 AI'에 대한 윤리적이고 합법적인 틀을 설계하기 위해 많은 이니셔티브가 수행되었습니다. 이미 2016년, 유럽 의회의 법무위원회, 백악관 과학 기술 정책국 (OSTP), 영국의 하원 과학 기술위원회 (House of Commons 'Science and Technology Committee)는 AI의 미래. 16 여러 전문가위원회가 그 후 보고서와 정책 문서를 작성했습니다. 그 중에서 유럽위원회가 임명한 인공 지능 전문가 그룹, 경제 협

력 개발기구 (OECD) 사회 인공 지능 전문가 그룹, 영국 인공 지능 선택위원회 (UK) 영주의 집 17

The Commission's White Paper affirms that two parallel policy objectives should be pursued and synergistically integrated. On the one hand research and deployment of AI should be promoted, so that the EU is competitive with the US and China.

위원회의 백서에서는 두 가지 병렬 정책 목표를 추구하고 시너지 효과적으로 통합해야 한다고 확언합니다. 한편으로 EU가 미국 및 중국과 경쟁할 수 있도록 AI의 연구 및 배포를 장려해야 합니다.

15 Floridi et al (2018, 690).
16 See Cath et al (2017).
17 For a recent review of documents on AI ethics and policy, see Jobin (2019).

The policy framework setting out measures to align efforts at European, national and regional level should aim to mobilise resources

유럽, 국가 및 지역 차원에서 노력을 조율하기위한 조치를 설정하는 정책 프레임 워크는 자원 동원을 목표로 해야 한다

to achieve an 'ecosystem of excellence' along the entire value chain, starting in research and innovation, and to create the right incentives to accelerate the adoption of solutions based on AI, including by small and medium-sized enterprises (SMEs)

연구 및 혁신을 시작으로 전체 가치 사슬에서 '우수한 생태계'를 달성하고 중소기업을 포함한 AI 기반 솔루션의 채택을 가속화하는 적절한 인센티브를 창출합니다.

On the other hand, the deployment of AI technologies should be consistent with the EU fundamental rights and social values. This requires measures to create an 'ecosystem of trust,' which should provide citizens with 'the confidence to take up AI applications' and 'companies and public organisations with the legal certainty to innovate using AI'. This ecosystem

반면 AI 기술의 배치는 EU 기본 권리 및 사회적 가치와 일치해야 합니다. 이를 위해서는 시민들에게 'AI 응용 프로그램을 채택할 수 있는 자신감'과 'AI를 사용하여 혁신할 법적 확실성을 가진 회사 및 공공 기관'을 제공하는 '신뢰 생태계'를 마련하기위한 조치가 필요합니다. 이 생태계

must ensure compliance with EU rules, including the rules protecting fundamental rights and consumers' rights, in particular for AI systems operated in the EU that pose a high risk.

기본 권리 및 소비자의 권리를 보호하는 규칙을 포함하여 특히 EU에서 운영되는 AI 시스템의 경우 위험이 높은 EU 규칙을 준수해야 합니다.

It is important to stress that the two objectives of excellence in research, innovation and implementation, and of consistency with individual rights and social values are compatible, but distinct. On the one hand the most advanced AI applications could be deployed to the detriment of citizens' rights and social values; on the other hand the effective protection of citizens' from the risks resulting from abuses AI does not provide in itself the incentives that are needed to stimulate research and innovation and promote beneficial uses. This report will argue that GDPR can contribute to address abuses of AI, and that it can be implemented in ways that do not hinder its beneficial uses. It will not address the industrial and other policies that are needed to ensure the EU

competitiveness in the AI domain.

연구, 혁신 및 실행의 우수성, 개인의 권리 및 사회적 가치와의 일관성이라는 두 가지 목표는 양립할 수 있지만 서로 다르다는 점을 강조하는 것이 중요합니다. 한편으로 가장 진보된 AI 애플리케이션은 시민의 권리와 사회적 가치를 해칠 수 있도록 배치될 수 있다. 반면에 AI 남용으로 인한 위험으로 부터 시민의 효과적인 보호는 연구와 혁신을 촉진하고 유익한 사용을 장려하는 데 필요한 인센티브를 제공하지 않습니다. 이 보고서는 GDPR이 AI 남용을 해결하는 데 기여할 수 있으며, 유익한 사용을 방해하지 않는 방식으로 구현될 수 있다고 주장합니다. AI 영역에서 EU 경쟁력을 확보하는 데 필요한 산업 및 기타 정책은 다루지 않습니다.

2.2.2. AI between logical models and machine learning

The huge success that AI has had in recent years is linked to a change in the leading paradigm in AI research and development.

AI가 최근 몇년간 큰 성공을 거둔 것은 AI 연구 개발의 주요 패러다임 변화와 관련이 있습니다.

Until a few decades ago, it was generally assumed that in order to develop an intelligent system, humans had to provide a formal representation of the relevant knowledge (usually expressed through a combination of rules and concepts), coupled with algorithms making inferences out of such knowledge. Different logical formalisms (rule languages, classical logic, modal and descriptive logics, formal argumentation, etc.) and computable models for inferential processes (deductive, defeasible, inductive, probabilistic, case-based, etc.) have been developed and applied.18

수십년 전까지 만해도 일반적으로 지능 시스템을 개발하기 위해서는 인간이 관련 지식 (일반적으로 규칙과 개념의 조합을 통해 표현됨)을 공식적으로 제시하고 알고리즘과 추론하는 알고리즘을 제공해야 한다고 가정 지식. 다른 논리 형식 (규칙 언어, 고전 논리, 모달 및 설명 논리, 형식 인수 등)과 추론적 프로세스 (연역적, 실현 가능, 유도적, 확률적, 사례 기반 등)에 대한 계산 가능한 모델이 개발되어 적용되었습니다.18

The structure for expert systems is represented in Figure 4. Note that humans appear both as users of the system and as creators of the system's knowledge base (experts, possibly helped by knowledge engineers).

전문가 시스템의 구조는 그림 4에 표시되어 있습니다. 사람은 시스템 사용자와 시스템 지식 기반 작성자 (전문가의 도움을 받을 수 있는 전문가)로 나타납니다.

18 Van Harmelen et al (2008).

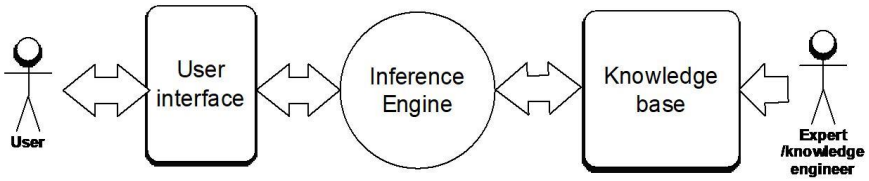

Figure 4 – Basic structure of expert systems

The theoretical results in knowledge representation and reasoning were not matched by disrupting, game-changing applications. Expert systems – i.e., computer systems including vast domain-specific knowledge bases, e.g., in medicine, law, or engineering, coupled with inferential engines – gave rise to high expectations about their ability to reason and answer users' queries.

지식 표현과 추론의 이론적 결과는 게임을 변화시키는 응용 프로그램을 방해하여 일치하지 않았습니다. 전문가 시스템 (예 : 의학,

법률 또는 엔지니어링 등의 광범위한 도메인 별 지식 기반을 포함하는 컴퓨터 시스템과 추론 엔진)은 사용자의 질문에 대한 추론 및 응답 능력에 대한 높은 기대치를 불러 일으켰습니다.

Unfortunately, such systems were often unsuccessful or only limitedly successful: they could only provide incomplete answers, were unable to address the peculiarities of individual cases, and required persistent and costly efforts to broaden and update their knowledge bases. In particular, expert-system developers had to face the so-called knowledge representation bottleneck: in order to build a successful application, the required information – including tacit and common-sense knowledge – had to be represented in advance using formalised languages. This proved to be very difficult and in many cases impractical or impossible.

불행하게도, 그러한 시스템은 종종 성공하지 못하거나 성공을 거두지 못했습니다. 불완전한 답변 만 제공할 수 있었으며, 개별 사례의 특성을 해결할 수 없었으며, 지식 기반을 넓히고 업데이트하기 위해 지속적이고 비용이 많이 드는 노력이 필요했습니다. 특히 전문가 시스템 개발자는 소위 지식 표현 병목 현상에 직면해야 했습니다. 성공적인 응용 프로그램을 구축하려면 암묵적 및 상식 지식을 포함한 필수 정보가 사전에 공식화된 언어를 사용하여 표현되어야 합니다. 이것은 매우 어렵고 많은 경우 비현실적이거나

불가능한 것으로 판명되었습니다.

In general, only in some restricted domains the logical models have led to successful application. In the legal domain, for example, logical models of great theoretical interest have been developed – dealing, for example, with arguments,19 norms, and precedents20 – and some expert systems have been successful in legal and administrative practice, in particular in dealing with tax and social security regulations. However, these studies and applications have not fundamentally transformed the legal system and the application of the law.

일반적으로 일부 제한된 도메인에서만 논리적 모델이 성공적으로 적용되었습니다. 예를 들어, 법적 영역에서 이론적 관심이 큰 논리적 모델이 개발되었으며 (예 : 논거, 19 개의 규범 및 선례 20), 일부 전문가 시스템은 특히 법적 및 행정적 관행, 특히 세금 및 사회 보장 규정. 그러나 이러한 연구와 응용은 법률 시스템과 법률의 적용을 근본적으로 변화시키지 않았습니다.

AI has made an impressive leap forward since it began to focus on the application of machine learning to mass amounts of data. This has led to a number of successful applications in many sectors –

ranging from automated translation to industrial optimisation, marketing, robotic visions, movement control, etc. – and some of these applications already have substantial economic and social impacts. In machine learning approaches, machines are provided with learning methods, rather than, or in addition to, formalised knowledge. Using such methods, they can automatically learn how to effectively accomplish their tasks by extracting/inferring relevant information from their input data. As noted, and as Alan Turing already theorised in the 1950s, a machine that is able to learn will achieve its goals in ways that are not anticipated by its creators and trainers, and in some cases without them knowing the details of its inner workings.21

AI는 대량의 데이터에 머신러닝을 적용하는 데 중점을 두기 시작하면서 인상적인 도약을 이루었습니다. 이로 인해 자동화된 번역에서 산업 최적화, 마케팅, 로봇 비전, 움직임 제어 등에 이르기까지 많은 분야에서 다수의 성공적인 응용 프로그램이 만들어졌으며 이러한 응용 프로그램 중 일부는 이미 상당한 경제적 및 사회적 영향을 미쳤습니다. 머신러닝 접근법에서, 머신에는 공식화된 지식보다는 또는 그에 더하여 학습 방법이 제공됩니다. 이러한 방법을 사용하면 입력 데이터에서 관련 정보를 추출/추론하여 작업을 효과적으로 수행하는 방법을 자동으로 배울 수 있습니다. 언급한 바와 같이 Alan Turing이 1950년대에 이미 이론화한 것처럼 학

습할 수 있는 기계는 제작자와 트레이너가 예상하지 못한 방식으로 목표를 달성할 수 있으며 경우에 따라 내부 작업의 세부 사항을 알지 못하는 경우에도 목표를 달성할 수 있습니다. 21

Even though the great success of machine learning has overshadowed the techniques for explicit and formalised knowledge representation, the latter remain highly significant. In fact, in many domains the explicit logical modelling of knowledge and reasoning can be complementary to machine learning. Logical models can explain the functioning of machine learning systems, check and govern their behaviour according to normative standards (including ethical principles and legal norms), validate their results, and develop the logical implications of such results according to conceptual knowledge and scientific theories. In the AI community the need to combine logical modelling and machine leaning is generally recognised, though different views exist on how to achieve this goal, and on the aspects to be covered by the two approaches (for a discussion on the limits of machine learning, see recently Marcus and Davis 2019).

머신러닝의 큰 성공이 명시적이고 공식화된 지식 표현을 위한 기술을 어둡게 했음에도 불구하고, 후자는 여전히 매우 중요합니다. 실제로, 많은 영역에서 지식과 추론에 대한 명시적 논리적 모델링

은 기계학습에 보완적일 수 있습니다. 논리적 모델은 기계학습 시스템의 기능을 설명하고 규범적 표준 (윤리적 원칙 및 법적 규범 포함)에 따라 행동을 점검 및 통제하고 결과를 검증하며 개념적 지식과 과학 이론에 따라 그러한 결과의 논리적 의미를 개발할 수 있습니다. AI 커뮤니티에서는 논리적 모델링과 머신 기울기를 결합해야 할 필요성이 일반적으로 인식되지만,이 목표를 달성하는 방법과 두 가지 접근 방식에 의해 다루어 질 측면 (머신러닝의 한계에 대한 논의를 위해 최근 Marcus and Davis 2019 참조).

19 Prakken, and Sartor (2015).
20 Ashley (2017).
21 Turing ([1951] 1996)

2.2.3. Approaches to learning

Three main approaches to machine learning are usually distinguished: supervised learning, reinforcement learning and unsupervised learning.

기계학습에 대한 세 가지 주요 접근 방식은 일반적으로 감독 학습, 강화 학습 및 비지도 학습입니다.

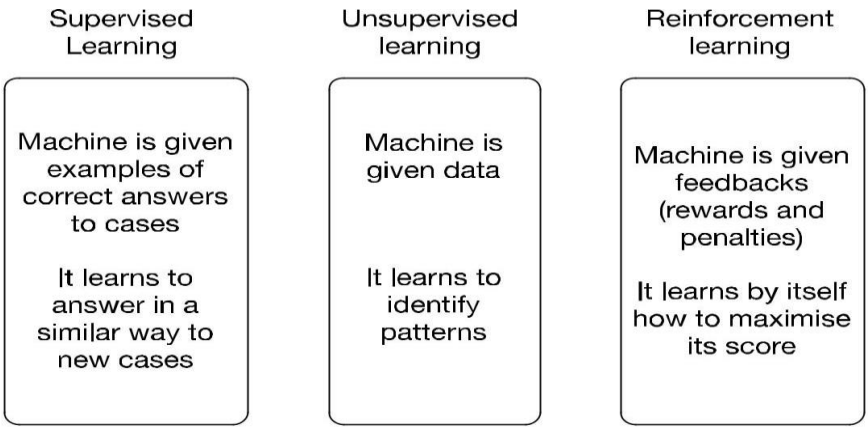

Figure 5 – Kinds of learning

Supervised learning is currently the most popular approach. In this case the machine learns through 'supervision' or 'teaching': it is given in advance a training set, i.e., a large set of (probably) correct answers to the system's task. More exactly the system is provided with a set of pairs, each linking the description of a case to the correct response for that case. Here are some examples: in systems designed to recognise objects (e.g. animals) in pictures, each picture in the training set is tagged with the name of the kind of object it contains (e.g., cat, dog, rabbit, etc.); in systems for automated translation, each (fragment of) a document in the source language is linked to its translation in the target language; in systems for personnel selection, the description of each past applicants (age, xperience, studies, etc.) is linked to whether the

application was successful (or to an indicator of the work performance for appointed candidates); in clinical decision support systems, each patient's symptoms and diagnostic tests is linked to the patient's pathologies; in recommendation systems, each consumer's features and behaviour is linked to the purchased objects; in systems for assessing loan applications, each record of a previous application is linked to whether the application was accepted (or, for successful applications, to the compliant or non-compliant behaviour of the borrower). As these examples show, the training of a system does not always require a human teacher tasked with providing correct answers to the system. In many case, the training set can be side-product of human activities (purchasing, hiring, lending, tagging, etc.), as is obtained by recording the human choices pertaining to such activities. In some cases the training set can even be gathered 'from the wild' consisting in data which is available on the open web. For instance, manually tagged images or faces, available on social networks, can be scraped and used for training automated classifiers.

지도 학습은 현재 가장 인기있는 접근법입니다. 이 경우 기계는 '감시'또는 '교육'을 통해 학습합니다. 사전에 훈련 세트, 즉 시스템 작업에 대한 (아마도) 정확한 정답 세트가 제공됩니다. 보다 정확하게 시스템에는 한 쌍의 세트가 제공되며, 각 세트는 사례에 대

한 설명을 해당 사례에 대한 올바른 응답에 연결합니다. 몇 가지 예는 다음과 같습니다. 그림에서 물체 (예 : 동물)를 인식하도록 설계된 시스템에서 훈련 세트의 각 그림에는 포함된 물체의 종류 (예 : 고양이, 개, 토끼 등)의 이름이 표시됩니다. 자동 번역을 위한 시스템에서, 소스 언어로 된 각 문서는 번역된 언어로 번역됩니다. 직원 선발 시스템에서 각 과거 지원자 (나이, 경험, 연구 등)에 대한 설명은 해당 응용 프로그램의 성공 여부 (지정된 후보자의 업무 성과 지표)와 관련이 있습니다. 임상 의사 결정 지원 시스템에서 각 환자의 증상 및 진단 검사는 환자의 병리와 관련이 있습니다. 추천 시스템에서 각 소비자의 특징과 행동은 구매한 물건과 연결되어 있습니다. 대출 신청을 평가하기위한 시스템에서, 이전 신청의 각 기록은 신청이 수락되었는지 여부, 또는 성공적인 신청을 위해 차용자의 준수 또는 비준수 행동에 연결됩니다. 이 예제에서 알 수 있듯이 시스템 교육에는 시스템에 대한 올바른 답변을 제공하는 인간 교사가 항상 필요한 것은 아닙니다. 많은 경우에, 훈련 세트는 그러한 활동에 관한 인간의 선택을 기록함으로써 얻어진 인간 활동 (구매, 고용, 대출, 태깅 등)의 부산물 일 수 있다. 경우에 따라 개방형 웹에서 사용할 수 있는 데이터로 구성된 '야생에서'훈련 세트를 수집할 수도 있습니다. 예를 들어, 소셜 네트워크에서 사용할 수 있는 수동으로 태그가 지정된 이미지 또는 얼굴을 스크랩하여 자동화된 분류기 교육에 사용할 수 있습니다.

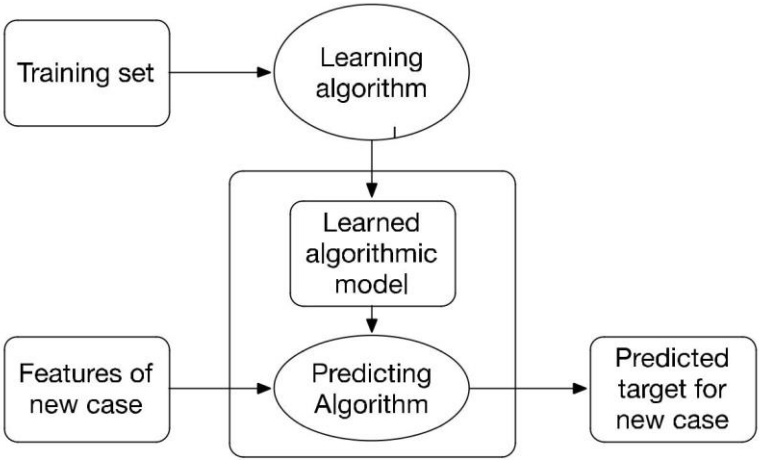

Figure 6 – Supervised learning

The learning algorithm of the system (its trainer), uses the training set to build an algorithmic model: neural network, a decision tree, a set of rules, etc. The algorithmic model is meant to capture the relevant knowledge originally embedded in the training set, namely the correlations between cases and responses. This model is then used, by a predicting algorithm, to provide hopefully correct responses to new cases, by mimicking the correlations in the training set. If the examples in the training set that come closest

to a new case (with regard to relevant features) are linked to a certain answer, the same answer will be proposed for the new case. For instance if the pictures that are most similar to a new input were tagged as cats, also the new input will also be tagged in the same way; if past applicants whose characteristic best match those of the new applicant were linked to rejection, the system will propose to reject also the new applicant; if the past workers who come closest to the new applicant performed well (or poorly), the systems will predict that also the applicant will perform likewise.

시스템의 학습 알고리즘 (트레이너)은 훈련 세트를 사용하여 신경망, 의사 결정 트리, 규칙 세트 등과 같은 알고리즘 모델을 구축합니다. 알고리즘 모델은 원래 훈련에 포함된 관련 지식을 포착하기위한 것입니다. 즉, 사례와 응답 사이의 상관 관계를 설정합니다. 그런 다음 예측 알고리즘에 의해이 모델을 사용하여 훈련 세트의 상관 관계를 모방하여 새로운 사례에 대한 올바른 응답을 제공합니다. 새로운 사례에 가장 가까운 훈련 세트의 예 (관련 기능과 관련하여)가 특정 답변에 연결되어 있으면 새로운 사례에 대해 동일한 답변이 제안됩니다. 예를 들어, 새로운 입력과 가장 유사한 그림에 고양이로 태그가 지정된 경우 새로운 입력에도 동일한 방식으로 태그가 지정됩니다. 새로운 신청자와 가장 일치하는 특성을 가진 과거 신청자가 거절과 관련이 있는 경우, 시스템은 새로운 신청자도 거부할 것을 제안합니다. 만약 새로운 지원자

에게 가장 가까이 온 과거의 노동자들이 성과가 좋았다면 (혹은 저조한), 시스템은 또한 신청자가 마찬가지로 수행할 것이라고 예측할 것입니다.

The answers by learning systems are usually called 'predictions'. However, often the context of the system's use often determines whether its proposals are be interpreted as forecasts, or rather as a suggestion to the system's user. For instance, a system's 'prediction' that a person's application for bail or parole will be accepted can be viewed by the defendant (and his or her lawyer) as a prediction of what the judge will do, and by the judge as a suggestion guiding her decision (assuming that she prefers not to depart from previous practice). The same applies to a system's prediction that a loan or a social entitlement will be granted.

학습 시스템에 의한 답변은 일반적으로 '예측'이라고 합니다. 그러나 종종 시스템 사용의 맥락에 따라 제안이 예측으로 해석되는지 또는 시스템 사용자에 대한 제안으로 해석되는지가 결정됩니다. 예를 들어, 보석이나 가석방 신청이 받아들여질 시스템의 '예측'은 피고 (및 그의 변호사)가 판사가 할 일에 대한 예측으로, 판사는 제안 안내로 볼 수 있습니다. 그녀의 결정 (이전 관행에서 떠나지 않기를 원한다고 가정). 대출이나 사회 자격이 부여될 것이라는 시스템의 예측에도 동일하게 적용됩니다.

There is also an important distinction to be drawn concerning whether the 'correct' answers in a training set are provided by the past choices by human 'experts' or rather by the factual consequences of such choices. Compare, for instance, a system whose training set consists of past loan applications linked to the corresponding lending decisions, and a system whose training set consists of successful applications linked to the outcome of the loan (repayment or non-payment). Similarly, compare a system whose training set consists of parole applications linked to judges' decisions on such application with a system whose training set consists of judicial decisions on parole applications linked to the subsequent behaviour of the applicant. In the first case, the system will learn to predict the decisions that human decision-makers (bank managers, or judges) would have made under the same circumstances. In the second case, the system will predict how a certain choice would affect the goals being pursued (preventing non-payments, preventing recidivism). In the first case the system would reproduce the virtues – accuracy, impartiality, farness – but also the vices – carelessness, partiality, unfairness – of the humans it is imitating. In the second case it would more objectively approximate the intended outcomes.

훈련 세트에서 '올바른'답변이 인간의 '전문가'에 의한 과거의 선택에 의해 제공되는지 또는 오히려 그러한 선택의 실제 결과에 의해 제공되는지에 관한 중요한 구별이 있다. 예를 들어, 교육 세트가 해당 대출 결정에 연결된 과거 대출 응용 프로그램으로 구성된 시스템과 대출 결과 (상환 또는 비 지불)에 연결된 성공적인 응용 프로그램으로 구성된 시스템을 비교하십시오. 유사하게, 훈련 세트가 그러한 신청에 대한 판사의 결정에 연결된 가석방 신청서로 구성된 시스템과 훈련 세트가 신청자의 후속 행동에 연결된 가석방 신청서에 대한 사법 결정으로 구성된 시스템과 비교하십시오. 첫 번째 경우, 시스템은 인간 의사 결정자 (은행 관리자 또는 판사)가 동일한 상황에서 내린 결정을 예측하는 방법을 배웁니다. 두 번째 경우, 시스템은 특정 선택이 추구하는 목표에 어떻게 영향을 미치는지 예측합니다 (비 지불 방지, 재범 방지). 첫 번째 경우, 시스템은 모방하는 인간의 미덕 (정확성, 공정성, 거리) 뿐만 아니라 악의 (부주의, 부분 성, 불공평)도 재현합니다. 두 번째 경우에는 의도한 결과에 더 객관적으로 접근합니다.

As a simple example of supervised learning, Figure 7, shows a (very small) training set concerning bail decisions along with the decision tree that can be learned on the basis of that training set. The decision tree captures the information in the training set through a combination of tests, to be performed sequentially. The first test

concerns whether the defendant was involved in a drug related offence. If the answer is positive, we have reached the bottom of the tree with the conclusion that bail is denied. If the answer is negative, we move to the second test, on whether the defendant used a weapon, and so on. Notice that the decision tree does not include information concerning the kind of injury, since all outcomes can be explained without reference to that information. This shows how the system's model does not merely replicate the training set; it involves generalisation: it assumes that certain combination of predictors are sufficient to determine the outcomes, other predictors being irrelevant.

지도 학습의 간단한 예로서, 그림 7은 보석(Bail) 결정에 관한 (매우 작은) 훈련 세트와 해당 훈련 세트를 기반으로 배울 수 있는 의사 결정 트리를 보여줍니다. 의사 결정 트리는 순차적으로 수행될 테스트 조합을 통해 학습 세트의 정보를 캡처합니다. 첫 번째 테스트는 피고가 마약 관련 범죄에 관여했는지 여부에 관한 것입니다. 대답이 긍정적이면 보석이 거부되었다는 결론으로 나무의 바닥에 도달했습니다. 답변이 부정적이면 피고가 무기를 사용했는지 여부 등 두 번째 테스트로 넘어갑니다. 모든 결과는 해당 정보를 참조하지 않고 설명할 수 있으므로 의사 결정 트리에는 부상 종류에 관한 정보가 포함되지 않습니다. 이것은 시스템 모델이 어떻게 훈련 세트를 복제하지 않는지를 보여줍니다. 여기에는 일반

화가 포함됩니다. 예측 변수의 특정 조합이 결과를 결정하기에 충분하고 다른 예측 변수는 관련이 없다고 가정합니다.

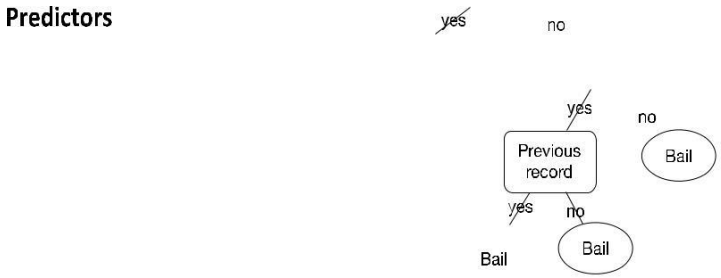

Figure 7 – Training set and decision tree for bail decisions

In this example we can distinguish the elements in Figure 6. The table in Figure 7 is the training set. The software that constructs the decision tree, is the learning algorithm. The decision tree itself, as shown in Figure 7 is the algorithmic model, which codes the logic of the human decisions in the training set. The software that processes new cases, using the decision tree, and makes predictions based on their features of such cases, is the predicting algorithm. In this example, as noted above, the decision tree reflects the attitudes of the decision-makers whose decisions are in the training set: it reproduces their virtues and biases.

이 예에서는 그림 6의 요소를 구별할 수 있습니다. 그림 7의 표는 학습 세트입니다. 의사 결정 트리를 구성하는 소프트웨어는 학습 알고리즘입니다. 그림 7에 표시된 의사 결정 트리 자체는 알고리즘 세트이며, 이는 훈련 세트에서 인간 의사 결정의 논리를 코딩합니다. 의사 결정 트리를 사용하여 새 사례를 처리하고 이러한 사례의 기능을 기반으로 예측하는 소프트웨어는 예측 알고리즘입니다. 이 예에서, 위에서 언급한 것처럼 의사 결정 트리는 의사 결정이 훈련 세트에 있는 의사 결정자의 태도를 반영합니다. 의사 결정은 미덕과 편견을 재현합니다.

For instance, according to the decision tree, the fact that the accuse concerns a drug-related offence is sufficient for bail to be denied. We may wonder whether this is a fair criterion for assessing bail requests. Note also that the decision tree (the algorithmic model) also provides answers for cases that do not fit exactly any example in the training set. For instance, no example in the training set concerns a drug -related offence with no weapon and no previous record. However, the decision tree provides an answer also for this case: there should be no bail, as this is what happens in all drug-related cases in the training set.

예를 들어, 의사 결정 트리에 따르면, 마약 관련 범죄와 관련된 고발이 보석금을 거부하기에 충분하다고 합니다. 우리는 이것이

보석금 요청을 평가하기위한 공정한 기준인지 궁금할 것입니다. 의사 결정 트리 (알고리즘 모델)는 훈련 세트의 예와 정확히 일치하지 않는 경우에 대한 답변도 제공합니다. 예를 들어, 훈련 세트의 예는 무기와 이전 기록이 없는 약물 관련 범죄와 관련이 없습니다. 그러나 의사 결정 트리는 이 경우에도 답을 제공합니다. 훈련 세트의 모든 약물 관련 사례에서 발생하는 보석이 없어야 합니다.

As another simplified example of supervised machine learning consider the training set and the rules in figure 7. In this case too, the learning algorithm, as applied to this very small set of past decisions, delivers questionable generalisation, such as the prediction that young age would always lead to a rejection of the loan applications and that middle age would always lead to acceptance.

감독된 머신러닝의 또 다른 단순화된 예로서 그림 7의 훈련 세트와 규칙을 고려하십시오. 이 경우에도, 이 아주 작은 과거 결정에 적용된 학습 알고리즘은 어린 나이 예측과 같은 의심스러운 일반화를 제공합니다. 항상 대출 신청을 거부하고 중년은 항상 수락으로 이어질 것입니다.

Usually, in order to give reliable prediction, a training set must include a vast number of examples, each described through a large set of predictors.

일반적으로, 신뢰할 수 있는 예측을 제공하기 위해, 훈련 세트는 다수의 예를 포함해야 하며, 각각은 큰 예측 변수 세트를 통해 설명된다.

Reinforcement learning is similar to supervised learning, as both involve training by way of examples. However, in the case of reinforcement learning the systems learns from the outcomes of its own action, namely, through the rewards or penalties (e.g., points gained or lost) that are linked to the outcomes of such actions. For instance, in case of a system learning how to play a game, rewards may be linked to victories and penalties to defeats; in a system learning to make investments, rewards may be linked to financial gains and penalties to losses; in a system learning to target ads effectively, rewards may be linked to users' clicks, etc.

강화 학습은 감독 학습과 유사하지만 둘 다 예를 들어 훈련을 포함합니다. 그러나 강화 학습의 경우 시스템은 자체 조치의 결과, 즉 그러한 조치의 결과와 관련된 보상 또는 처벌 (예 : 획득 또는 손실)을 통해 학습합니다. 예를 들어, 게임을 하는 방법을 배우는

시스템의 경우, 보상은 승리와 승부와 연관될 수 있습니다. 투자를 배우는 시스템 학습에서 보상은 재정적 이익과 관련이 있으며 손실에 대한 처벌과 관련될 수 있습니다. 효과적으로 광고를 타겟팅하는 시스템 학습에서 보상은 사용자의 클릭 등에 연결될 수 있습니다.

In all these cases, the system observes the outcomes of its actions, and it self-administers the corresponding rewards or penalties. Being geared towards maximising its score (its utility), the system will learn to achieve outcomes leading to rewards (victories, gains, clicks), and to prevent outcomes leading to penalties. With regard to reinforcement learning too, we can distinguish the learner (the algorithm that learns how to act successfully, based on the outcomes of previous actions by the system) and the learned model (the output of the learner, which determines the system's new actions).

이 모든 경우에 시스템은 조치의 결과를 관찰하고 해당 보상 또는 위약금을 자체 관리합니다. 점수 (유틸리티)를 극대화하기 위해 시스템은 보상 (승리, 이득, 클릭)으로 이어지는 결과를 달성하고 벌칙으로 이어지는 결과를 방지하는 방법을 배우게 됩니다. 강화 학습과 관련하여 학습자 (시스템의 이전 조치 결과에 따라 성공적으로 행동하는 방법을 학습하는 알고리즘)와 학습된 모델 (학

습자의 결과, 시스템의 새로운 조치를 결정하는 결과)을 구별할 수 있습니다.).

In unsupervised learning, finally, AI systems learn without receiving external instructions, either in advance or as feedback, about what is right or wrong. The techniques for unsupervised learning are used in particular, for clustering, i.e., for grouping the set of items that present relevant similarities or connections (e.g., documents that pertain to the same topic, people sharing relevant characteristics, or terms playing the same conceptual roles in texts). For instance, in a set of cases concerning bail or parole, we may observe that injuries are usually connected with drugs (not with weapons as expected), or that people having prior record are those who are related to weapon. These clusters might turn out to be informative to ground bail or parole policies.

비지도 학습에서 마지막으로 AI 시스템은 옳고 그른 것에 대한 사전 또는 피드백으로 외부 지시를 받지 않고 학습합니다. 비지도 학습 기술은 특히 클러스터링, 즉 관련 유사점 또는 관련성을 나타내는 항목 세트 (예 : 동일한 주제와 관련된 문서, 관련 특성을 공유하는 사람 또는 동일한 개념적 역할을 하는 용어)를 그룹화하는 데 사용됩니다. 본문에서). 예를 들어 보석이나 가석방과 관련된 일련의 사례에서 우리는 부상이 일반적으로 약물과 관련이

있거나 (예상한 무기가 아님) 사전 기록을 가진 사람들이 무기와 관련된 사람들임을 알 수 있습니다. 이 클러스터는 보석금이나 가석방 정책에 유익한 정보로 판명될 수 있습니다.

2.2.4. Neural networks and deep learning

Many techniques have been deployed in machine learning: decision trees, statistical regression, support vector machine, evolutionary algorithms, methods for reinforcement learning, etc. Recently, deep learning based on many-layered neural networks has been very successfully deployed especially, but not exclusively, where patterns have to be recognised and linked to classifications and decisions (e.g., in detecting objects in images, recognising sounds and their sources, making medical diagnosis, translating texts, choosing strategies in games, etc.). Neural networks are composed of a set of nodes, called neurons, arranged in multiple layers and connected by links.

의사 결정 트리, 통계적 회귀, 지원 벡터 머신, 진화 알고리즘, 강화 학습 방법 등과 같은 많은 기술이 머신러닝에 배포되었습니다. 최근에 여러 계층의 신경망을 기반으로 한 딥 러닝이 매우 성공적으로 배포되었지만 독점적으로는 아닙니다. 패턴을 인식하고 분

류 및 결정과 연계해야 하는 경우 (예 : 이미지의 물체 감지, 소리 및 소스 인식, 의료 진단, 텍스트 번역, 게임에서 전략 선택 등) 신경망은 뉴런이라고하는 일련의 노드로 구성되어 있으며 여러 층으로 배열되어 있으며 링크로 연결되어 있습니다.

They are so-called, since they reproduce some aspects of the human nervous system, which indeed consists of interconnected specialised cells, the biological neurons, which receive and transmit information. Neural networks were indeed developed under the assumption that artificial intelligence could be achieved by reproducing the human brain, rather than by modelling human reasoning, i.e., that artificial reasoning would naturally emerge out of an artificial brain (though we may wonder to what extent artificial neural networks and human brains really share the similar structures and processes Each neuron receives signals (numbers) from connected neurons or from the outside, and these signals are magnified or diminished as they cross incoming links, according to the weights of the latter.

그것들은 실제로 인간의 신경계의 일부 측면을 재현하기 때문에 소위, 상호 연결된 특수 세포, 생물학적 뉴런으로 구성되어 정보를 수신하고 전송합니다. 인공 지능은 인간의 추론을 모델링하는 것이 아니라 인간의 두뇌를 재생산함으로써 달성될 수 있다는 가정, 즉 인공 추론이 자연적으로 인공 두뇌에서 자연적으로 나타날 것이라는 가정하에 실제로 신경망이 개발되었다 네트워크와 인간의 뇌는 실제로 유사한 구조와 과정을 공유합니다. 각 뉴런은 연결된 뉴런 또는 외부에서 신호 (숫자)를 수신하며, 이 신호는 후자의 가중치에 따라 들어오는 링크를 통과할 때 확대되거나 축소됩니다.

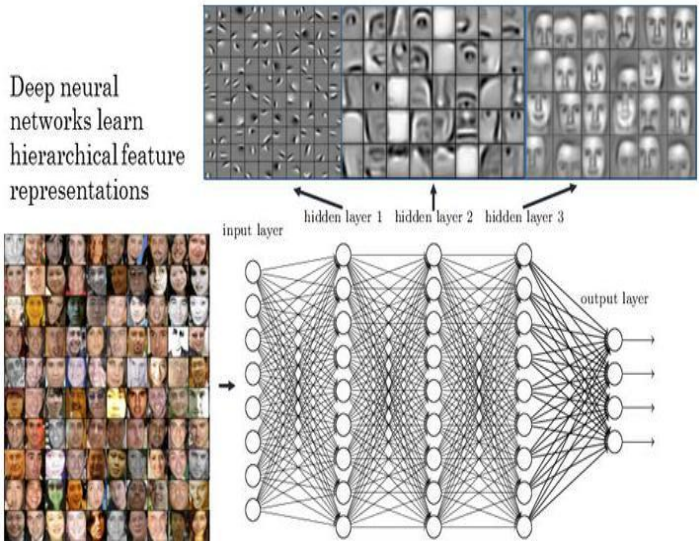

Figure 8 – Multilayered (deep) neural network for face recognition

The neuron applies some calculations to the input it receives, and if the result reaches the neuron's threshold, the neuron becomes active sending signals to the connected neurons or outside of the network. The activation starts from nodes receiving external inputs and spreads through the network. The training of the network takes place by telling the network whether its answers (its outputs) are right or wrong. If an answer by the network is wrong, the learning algorithm updates the network – i.e., it adjusts the weights of the connections – so that next time the network is presented with that input, it will give the correct answer. Figure 8 shows a simplified representation of a multi-layered neural network (real networks may have many more layers of neurons) for face recognition, where the initial layers learn very generic aspects of the images (border, colours, shapes, etc.) while higher layers engage with the elements of human faces.

뉴런은 수신한 입력에 일부 계산을 적용하고 결과가 뉴런의 임계값에 도달하면 연결된 뉴런 또는 네트워크 외부로 신호를 전송하는 활성 상태가 됩니다. 활성화는 외부 입력을 받는 노드에서 시작하여 네트워크를 통해 확산됩니다. 네트워크 교육은 네트워크의 응답 (출력)이 옳은지 아닌지를 알려주는 방식으로 진행됩니다. 네트워크의 답변이 틀린 경우 학습 알고리즘은 네트워크를 업데

이트합니다. 즉, 연결 가중치를 조정하여 다음에 네트워크에 해당 입력이 표시되면 올바른 답변을 제공합니다. 그림 8은 얼굴 인식을 위한 다층 신경망 (실제 네트워크에는 더 많은 뉴런 층이 있을 수 있음)의 단순화된 표현을 보여줍니다. 더 높은 층은 인간의 얼굴 요소와 결합합니다.

In the case of the neural network, the learning algorithm modifies the network until it achieves the desired performance level, while the outcome of the learning – algorithmic model – is the network in its final configuration.

신경망의 경우 학습 알고리즘은 원하는 성능 수준에 도달할 때까지 네트워크를 수정하는 반면 학습 결과 (알고리즘 모델)는 최종 구성의 네트워크입니다.

As previously noted, the learning algorithm is able to modify the neural network (the weights in connections and neurons) so that the network is able to provide the most appropriate answers. Under the supervised learning approach, the trained network will reproduce the behaviour in the training set; under the reinforcement learning approach, the network will adopt the behaviour that maximises its score (e.g. the reward points linked to

gains in investments or to victories in games).

앞에서 언급했듯이 학습 알고리즘은 신경망 (연결 및 뉴런의 가중치)을 수정하여 네트워크가 가장적합한 답변을 제공할 수 있습니다. 지도 학습 방식에 따라 훈련된 네트워크는 훈련 세트의 행동을 재현합니다. 강화 학습 접근법 하에서 네트워크는 점수를 극대화하는 행동 (예 : 투자 이익 또는 게임 승리와 관련된 보상 포인트)을 채택합니다.

2.2.5. Explicability (설명 가능성)

Different machine learning approaches differ in their ability to provide explanations. For instance, the outcome of a decision tree can be explained through the sequence of tests leading to that outcome. In our example, if bail is refused after testing No for Drug, Yes for Weapons and Yes for Previous record, an explanation is provided by a corresponding rule: if No Drug and Weapons and Previous Record, then No Bail.

기계학습 방식에 따라 설명 기능이 다릅니다. 예를 들어, 의사 결정 트리의 결과는 해당 결과로 이어지는 일련의 테스트를 통해 설명할 수 있습니다. 이 예에서는 약물에 대해 아니요, 무기에 대

해 예 및 이전 레코드에 대해 예를 테스트한 후 보석금이 거부된 경우 해당 규칙에 따라 설명이 제공됩니다. 약물 및 무기 및 이전 레코드가 없으면 보석금이 없습니다.

Unlike a decision tree, a neural network does not provide explanations of its outcomes. It is possible to determine how a certain output has resulted from the network's activation, and how that activation, in response to a given input, was determined by the connections between neurons (and by the weights assigned to such connections as a result of the network's training) and by the mathematical functions governing each neuron. However, this information does not show a rationale that is meaningful to humans: it does not tell us why a certain response was given. Many approaches exist to providing explanations of the behaviour of neural networks and other opaque systems (also called 'black boxes').

의사 결정 트리와 달리 신경망은 결과에 대한 설명을 제공하지 않습니다. 특정 출력이 네트워크 활성화로 인한 결과와 주어진 입력에 대한 응답으로 뉴런 간의 연결 (및 네트워크 연결의 결과로 이러한 연결에 할당된 가중치)에 의해 활성화가 결정된 방법을 결정할 수 있습니다. 훈련) 및 각 뉴런을 지배하는 수학적 기능에 의해. 그러나이 정보는 인간에게 의미가 있는 이론적 근거를 보여

주지 않습니다. 왜 특정 응답이 제공되었는지는 알려주지 않습니다. 신경망과 다른 불투명한 시스템 ('블랙 박스'라고도 함)의 동작에 대한 설명을 제공하는 많은 방법이 있습니다.

Some of these approaches look into the system to be explained, and build explanations accordingly (e.g., looking at the outcomes of the network's different layers, as in the example in Figure 8). Other approaches build explanations on the basis of the network's external behaviour: they only consider the relation between the inputs provided by the network and the outcomes it delivers, and build arguments or other explanations accordingly. However, advancements of human-understandable explanation of neural networks have so far been quite limited still.22 Unfortunately, in many domains, the systems whose functioning is less explicable provide higher performance. Thus, comparative advantages in performance and in explicability may have to be balanced, in order to determine what approach should be adopted in a machine learning system.

이러한 접근 방식 중 일부는 설명할 시스템을 조사하고 그에 따라 설명을 작성합니다 (예 : 그림 8의 예와 같이 네트워크의 다른 계층의 결과 확인). 다른 접근 방식은 네트워크의 외부 행동을 기반으로 설명을 작성합니다. 네트워크에서 제공하는 입력과 전달한

결과 간의 관계 만 고려하고 그에 따라 인수 또는 기타 설명을 작성합니다. 그러나 신경망에 대한 사람이 이해할 수 있는 설명의 발전은 지금까지 여전히 제한적이다.22 불행히도, 많은 영역에서 기능이 덜 설명 가능한 시스템은 더 높은 성능을 제공한다. 따라서 머신러닝 시스템에서 어떤 접근법을 채택해야 하는지 결정하기 위해 성능과 설명의 비교 이점이 균형을 이루어야 할 수도 있습니다.

22 Guidotti et al (2018).

The best balance also depends on the domain in which the system is used and on the importance of the interests that are affected. When public action is involved and key human interests are at stake (e.g., as in judicial decisions) explanation is paramount.

최상의 균형은 또한 시스템이 사용되는 영역과 영향을 받는 이익의 중요성에 달려 있습니다. 공공 행동이 수반되고 주요 인간의 이익이 위태로울 때 (예를 들어, 사법결정에서와 같이) 설명이 가장 중요합니다.

Even when a system can only be viewed as a black box, however,

some critical analyses of its behaviour are still possible. Through sensitivity analysis – i.e., by systematically checking whether the output changes if the value of certain input features is modified, leaving all other features unchanged – we can understand what features determine the system's output. For instance, by checking whether the prediction of a system meant to assess creditworthiness changes if we modify the place of birth or residence of the applicant, we can determine whether this input feature is relevant to the system's output. Consequently, we may wonder whether the system unduly discriminated people depending on their ethnicity or social status, which may be linked to place of birth or residence.

그러나 시스템을 블랙 박스로만 볼 수 있는 경우에도 시스템의 동작에 대한 일부 중요한 분석은 여전히 가능합니다. 민감도 분석을 통해, 즉 특정 입력 기능의 값이 수정되면 출력이 변경되는지 여부를 체계적으로 확인하여 다른 모든 기능을 변경하지 않고 시스템의 출력을 결정하는 기능을 이해할 수 있습니다. 예를 들어, 신청자의 출생지나 거주지를 수정하면 시스템의 예측이 신용도 변화를 평가할 수 있는지 여부를 확인함으로써 이 입력 기능이 시스템의 출력과 관련이 있는지 여부를 확인할 수 있습니다. 결과적으로, 우리는 시스템이 출생지 또는 거주지와 관련될 수 있는 인종 또는 사회적 지위에 따라 과도하게 차별된 사람들인지 궁금

할 수 있습니다.

2.3. AI and (personal) data

The following sections will consider the interaction between AI and big data. First, the use of big data for AI-bases predictions and assessments will be introduced. The ensuing risks and opportunities will be analysed. Then, decision-making concerning individuals will be addressed, with a focus on fairness and non-discrimination. Finally, the issues concerning profiling, influence and manipulation will be analysed, including those related to pervasive surveillance by private actors and governments.

다음 섹션에서는 AI와 빅 데이터 간의 상호 작용을 고려합니다. 먼저 AI 기반 예측 및 평가에 빅 데이터를 사용합니다. 그에 따른 위험과 기회가 분석될 것입니다. 그런 다음 공정성과 비차별에 중점을 두어 개인에 관한 의사 결정을 다룰 것입니다. 마지막으로, 개인 행위자와 정부의 광범위한 감시와 관련된 문제를 포함하여 프로파일링, 영향 및 조작에 관한 문제가 분석됩니다.

2.3.1. Data for automated predictions and assessments

To predict a certain outcome in a new case means to jump from certain known features of that case, the so-called predictors (also called independent variables, or features), to an unknown feature of that case, the target to be predicted (also called dependent variable, or label). This forecast is based of models that capture general aspects of the contexts being considered, on the basis of which it is possible to connect the values of predictors and targets. For instance a model in the medical domain may connect symptoms to diseases, a psychometric model may connect online behaviour (e.g., friends, posts and likes on a social network) to psychological attitudes; etc.

새로운 사례에서 특정 결과를 예측한다는 것은 해당 사례의 알려진 기능 (소위 예측 변수 (독립 변수 또는 기능이라고도 함))에서 해당 사례의 미지의 기능 (예상 대상)으로 점프하는 것을 의미합니다. 종속 변수 또는 레이블이라고 함). 이 예측은 예측 자와 목표의 값을 연결할 수 있는 가능성을 바탕으로 고려중인 컨텍스트의 일반적인 측면을 캡처하는 모델을 기반으로 합니다. 예를 들어, 의료 영역의 모델은 증상을 질병에 연결할 수 있고, 심리학적 모델은 온라인 행동 (예 : 소셜 네트워크의 친구, 게시물 및 좋아요)을 심리적 태도에 연결할 수 있습니다. 기타

Such models may be created by humans (who formulate the rules and concepts in the model), even when the application of the models is delegated to a machine (as in rule-based expert systems). However, as noted in Section 2.2.2, the construction (learning) of the models, and not only their application is increasingly entrusted to machines. In the machine learning approach, machines discover the probabilistic correlations between predictors and targets, and then apply these correlations to make predictions in new cases. Thanks to the combination of AI techniques, vast masses of data, and computational power, it has become possible to base automated predictions and assessments on a much larger sets of examples, taking into account a much larger set of features of each of them, so as to achieve useful level of accuracy in many domains.

이러한 모델은 모델의 적용이 기계에 위임된 경우에도 (규칙 기반 전문가 시스템에서와 같이) 사람 (모델에서 규칙 및 개념을 공식화)에 의해 생성될 수 있습니다. 그러나 2.2.2 절에 언급된 바와 같이, 모델의 구성 (학습)은 물론 그 적용 뿐만 아니라 기계에도 적용됩니다. 기계학습 접근 방식에서 기계는 예측 변수와 대상 간의 확률적 상관 관계를 발견한 다음 이러한 상관 관계를 적용하여 새로운 경우에 예측합니다. AI 기술, 방대한 양의 데이터 및 계산 능력의 조합으로 인해 훨씬 더 많은 예제 세트를 기반으로 자동화된 예측 및 평가를 기반으로 할 수 있게 되었습니다. 많은 영

역에서 유용한 수준의 정확도를 달성하기 위해

For instance, targeted advertising may be based on records linking the characteristics and behaviour of consumers (gender, age, social background, purchase history, web browsing, etc.) to their responses to ads. Similarly, the assessment of job applications may be based on records linking characteristics of previous workers (education, employment history, jobs, aptitude tests, etc.), to their work performance; the prediction of the likelihoods of recidivism by a particular offender may be based on records combining characteristics of past offenders (education, employment history, family status, criminal record, psychological tests, etc.) with data or assessments on their recidivism; the prediction of a prospective borrower's creditworthiness may be based on records linking the characteristics of past borrowers to data or assessments about their creditworthiness; the diagnosis of diseases or the suggestion of personalised medical treatments may be based on the records of

past patients, linking their characteristics and medical tests to subsequent medical conditions and treatments.

예를 들어, 타겟 광고는 소비자의 특성 및 행동 (성별, 연령, 사회적 배경, 구매 이력, 웹 브라우징 등)을 광고에 대한 응답과 연결

한 레코드를 기반으로 할 수 있습니다. 유사하게, 직무 응용 프로그램의 평가는 이전 근로자의 특성 (교육, 고용 이력, 직업,적성 검사 등)을 업무 성과와 연결하는 기록을 기반으로 할 수 있습니다. 특정 가해자에 의한 재화 가능성의 예측은 과거 가해자 (교육, 고용 이력, 가족 상태, 범죄 기록, 심리 검사 등)의 특성과 재범에 대한 데이터 또는 평가를 결합한 기록을 기반으로 할 수 있습니다. 장래 차용인의 신용도 예측은 과거 차용자의 특성을 신용도에 대한 데이터 또는 평가와 연결하는 기록에 기초할 수 있다. 질병의 진단 또는 개인화된 의학적 치료 제안은 다음의 기록에 근거할 수 있습니다.

과거 환자, 특성 및 의료 검사를 후속 의료 조건 및 치료와 연결합니다.

As a result of the need to learn by analysing vast amount of data, AI has become hungry for data, and this hunger has spurred data collection, in a self-reinforcing spiral.23 Thus, the development of AI systems based on machine learning presupposes and fosters the creation of vast data sets, i.e., big data 24.

방대한 양의 데이터를 분석하여 학습해야 할 필요성으로 인해 AI는 데이터에 대해 배고프고, 이 기아는 자체 강화 나선형으로 데이터 수집을 촉진했습니다.23 따라서 머신러닝 기반의 AI 시스템

개발 방대한 데이터 세트, 즉 빅 데이터 24 생성을 촉진합니다.

The collection of data is facilitated by the availability of electronic data as a by-product of using any kind of ICT system. Indeed, a massive digitisation has preceded most AI applications, resulting from the fact that data flows are produced in all domains where computing is deployed.9 For instance, huge amounts of data are collected every second by computers that execute economic transactions (as in e-commerce)10, by sensors monitoring and providing input to physical objects (e.g., vehicles or smart home devices), by the workflows generated by economic and governmental activities (e.g., banking, transportation, or taxation, etc.); by surveillance devices (e.g. traffic cameras, or access control systems); and systems supporting non-market activities (e.g. internet access, searching, or social networking).

데이터 수집은 모든 종류의 ICT 시스템을 사용하는 부산물로 전자 데이터의 가용성에 의해 촉진됩니다. 실제로, 디지털화는 대부분의 AI 응용 프로그램보다 우선합니다. 컴퓨팅이 배포되는 모든 도메인에서 데이터 흐름이 생성되기 때문입니다.9 예를 들어, 경제적인 트랜잭션을 실행하는 컴퓨터는 매 초마다 대량의 데이터를 수집합니다 (예 : e) 10) 경제 및 정부 활동 (예 : 은행, 운송 또는 세금 등)에 의해 생성된 워크 플로우에 의해 물리적 객체 (예 :

차량 또는 스마트 홈 장치)에 대한 모니터링 및 입력을 제공하는 센서에 의한 것; 감시 장치 (예 : 교통 카메라 또는 출입 통제 시스템) 비 시장 활동을 지원하는 시스템 (예 : 인터넷 액세스, 검색 또는 소셜 네트워킹).

In recent years, these data flows have been integrated into a global interconnected data-processing infrastructure, centred on, but not limited to, the Internet. This infrastructure constitutes a universal medium for communicating, accessing data, and delivering any kind of private and public services. It enables citizens to shop, use banking and other services, pay taxes, get government benefits and entitlements, access information and knowledge, and build social connections. Algorithms – often powered by AI – mediate citizens' access to content and services, selecting information and opportunities for them, while at the same time recording any activity. Today, this global interconnected data-processing infrastructure seems to include about 30 billion devices – computers, smart phones, industrial machines, cameras, etc. – which generate masses of electronic data (see Figure 9).

최근 몇 년 동안 이러한 데이터 흐름은 인터넷을 중심으로 한 글로벌 상호 연결된 데이터 처리 인프라에 통합되었습니다. 이 인프라는 통신, 데이터 액세스 및 모든 종류의 개인 및 공공 서비스를

제공하기위한 보편적인 매체입니다. 시민들은 쇼핑, 뱅킹 및 기타 서비스 이용, 세금 납부, 정부 혜택 및 자격 취득, 정보 및 지식 이용, 사회적 관계 구축 등을 할 수 있습니다. AI로 구동되는 알고리즘은 시민의 컨텐츠 및 서비스 액세스를 중재하고 정보 및 기회를 선택하는 동시에 활동을 기록합니다. 오늘날 전 세계적으로 상호 연결된 데이터 처리 인프라에는 컴퓨터, 스마트 폰, 산업용 기계, 카메라 등 약 300 억 개의 장치가 포함되어 있으며 대량의 전자 데이터를 생성합니다 (그림 9 참조).

23 Cristianini (2016).

24 Mayer-Schönberger and Cukier (2013).

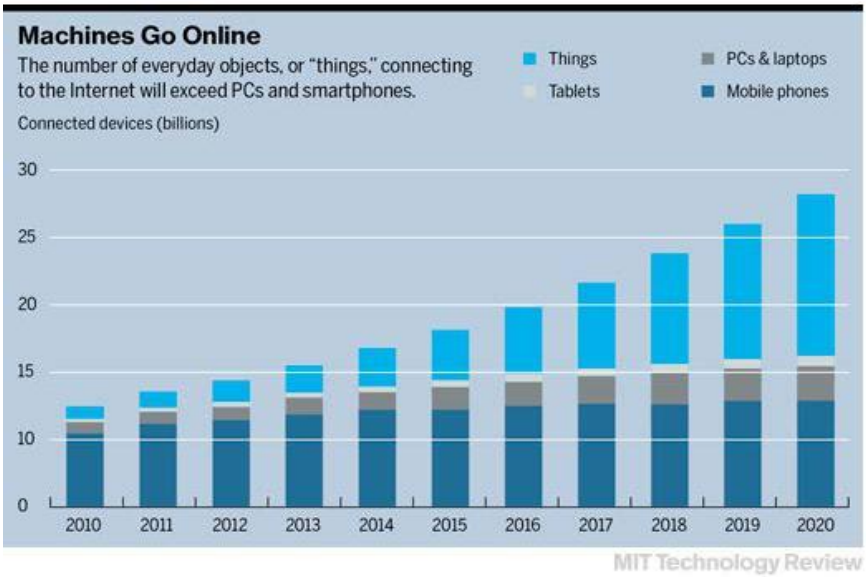

Figure 9 – Number of connected devices

Figure 10 provides a comparative overview of what takes place online every minute.

그림 10은 1 분마다 온라인에서 수행되는 작업에 대한 비교 개요를 제공합니다.

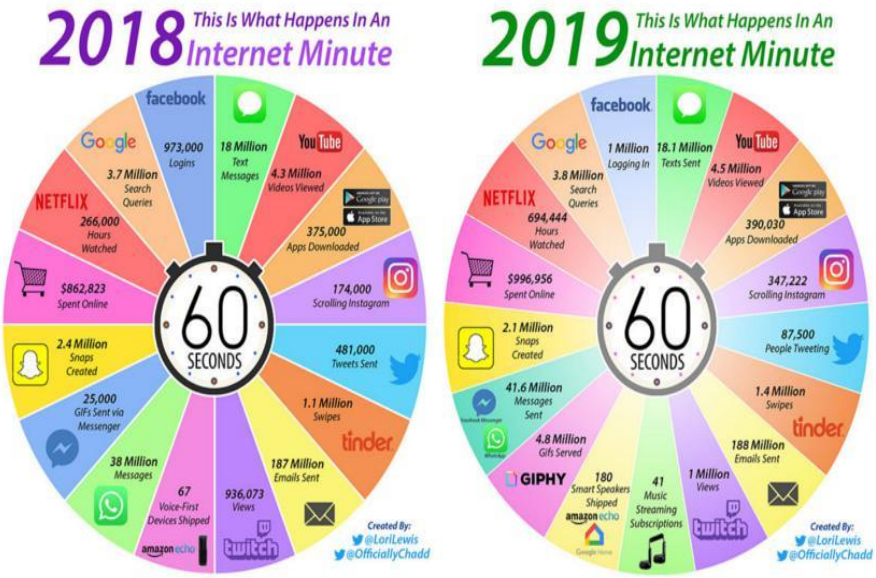

Figure 10 – Data collected in a minute of online activity worldwide

AI's hunger for data concerns any kind of information: from meteorological data, to environmental ones, to those concerning

industrial processes. Figure 4 gives an idea of the growth of data creation.

데이터에 대한 AI의 기아는 기상 데이터, 환경 데이터, 산업 프로세스 관련 정보에 이르기까지 모든 종류의 정보와 관련이 있습니다. 그림 4는 데이터 생성 증가에 대한 아이디어를 제공합니다.

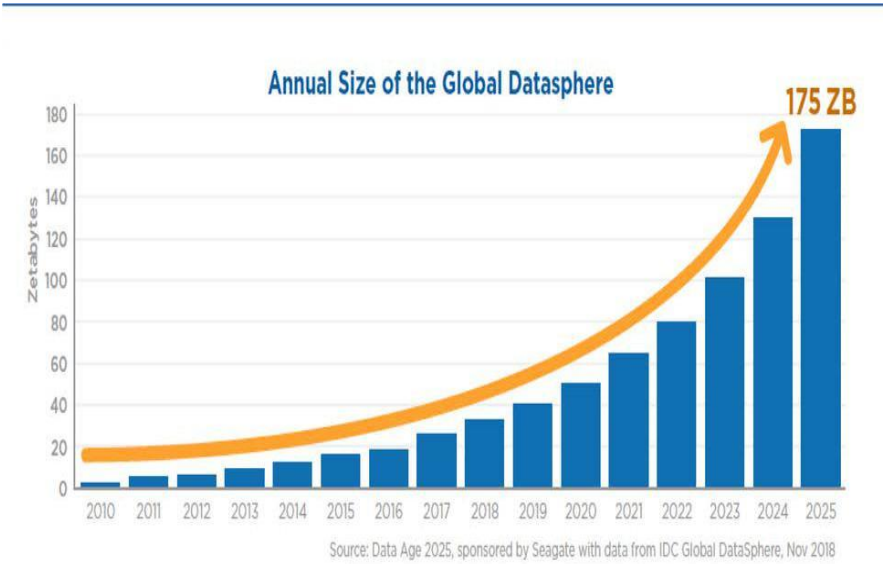

Figure 11 – Growth of global data

2.3.2. AI and big data: risks and opportunities

The integration of AI and big data technologies into the global

data-processing infrastructure can deliver a lot of benefits: better access to information; generation and distribution of knowledge across the globe; cost savings, greater productivity, and value creation; new creative and well-

paying jobs; individualised private and public services; environmentally-friendly management of utilities and logistics; novel information and consulting services; support for transparency; remedies against biases and discriminations, etc. Great advances are enabled in many domains: scientists can discover correlations, formulate hypotheses and develop evidence-based models; doctors can provide better diagnosis and personalised and targeted therapies; firms can anticipate market trends and make more efficient decisions; consumers can make more informed choices and obtain personalised services; public authorities can anticipate risks, prevent damages, optimise the management of public goods (such as the environment) and coordinate citizens' actions (e.g., the management of traffic, energy consumption, and utilities). And more good can come in the future.

AI와 빅 데이터 기술을 글로벌 데이터 처리 인프라에 통합하면 다음과 같은 많은 이점을 얻을 수 있습니다. 전세계의 지식 생성 및 배포; 비용 절감, 생산성 향상 및 가치 창출; 새로운 창의적이

고

일자리를 지불; 개별화된 개인 및 공공 서비스; 환경 친화적 유틸리티 및 물류 관리; 새로운 정보 및 컨설팅 서비스; 투명성 지원; 편견과 차별 등에 대한 구제 등 여러 분야에서 큰 발전이 가능하다. 과학자들은 상관 관계를 발견하고, 가설을 공식화하고, 증거 기반 모델을 개발할 수 있다. 의사는 더 나은 진단과 개인 맞춤형 치료법을 제공할 수 있습니다. 기업은 시장 동향을 예측하고 보다 효율적인 의사 결정을 내릴 수 있습니다. 소비자는 보다 많은 정보를 바탕으로 선택하고 개인화된 서비스를 받을 수 있습니다. 공공 기관은 위험을 예측하고 피해를 예방하며 공공 물품 관리 (예 : 환경)를 최적화하고 시민의 행동 (예 : 교통, 에너지 소비 및 유틸리티 관리)을 조정할 수 있습니다. 그리고 앞으로 더 좋은 일이 올 수 있습니다.

As has been argued by Ray Kurzweil, an inventor, futurist, and director of engineering at Google:

Google의 발명가, 미래 학자이자 엔지니어링 책임자 인 Ray Kurzweil이 주장한 바와 같이 :

Through [information] technologies we can address the

grand challenges of humanity, such as maintaining a healthy environment, providing the resources for a growing population (including energy, food, and water), overcoming disease, vastly extending human longevity, and eliminating poverty. It is only by extending ourselves with intelligent technology that we can deal with the scale of complexity needed. 25

[정보] 기술을 통해 우리는 건강한 환경을 유지하고, 증가하는 인구 (에너지, 음식 및 물 포함)를 위한 자원 제공, 질병을 극복하고, 인간의 장수를 크게 연장하고, 빈곤을 제거하는 것과 같은 인류의 큰 도전에 대처할 수 있습니다. 필요한 복잡성의 규모를 처리할 수 있는 것은 지능형 기술로 스스로를 확장하는 것입니다. 25

In some cases, AI can fully replace human activities (e.g., in driverless vehicles, cleaning robots, and certain planning and scheduling tasks in logistics). In many cases it rather complements human capacities: it enhances the human ability to know and act, it supports creativity and invention.26 Thanks to AI, it may be possible to achieve a new cooperation between humans and machines, which overcomes the classical model in which machines only performed routine and repetitive tasks. This integration was

already predicted in the early 1960s' by JK Licklider, a scientist who played a key role in the development of the Internet. He argued that in the future, cooperation between human and computer would include creative activities, i.e., 'making decisions and controlling complex situations without inflexible dependence on predetermined programs.27

경우에 따라 AI는 인간 활동을 완전히 대체할 수 있습니다 (예 : 무인 차량, 로봇 청소, 물류의 특정 계획 및 예약 작업). 많은 경우에 그것은 인간의 능력을 보완합니다 : 그것은 알고 행동하는 인간의 능력을 향상시키고, 창의성과 발명을 지원합니다.26 AI 덕분에 인간과 기계 간의 새로운 협력을 달성할 수 있습니다. 일상적인 작업과 반복적인 작업 만 수행한 시스템 이러한 통합은 1960년대 초 인터넷 개발에 중요한 역할을 한 과학자인 JK Licklider에 의해 이미 예측되었습니다. 그는 앞으로 인간과 컴퓨터 사이의 협력에는 창의적 활동, 즉 '결정된 프로그램에 의존하지 않고 결정을 내리고 복잡한 상황을 통제하는 것'이 포함된다고 주장했다.27

25 Kurzweil (2012).
26 McAfee, and Brynjolfsson (2019).

Today, it is indeed possible to integrate humans and machines in new ways that not only exploit synergies, but may also preserve

and enhance human initiative and work satisfaction.28

오늘날 시너지 효과를 활용할 뿐만 아니라 인간의 주도권과 업무 만족도를 보존하고 향상시킬 수 있는 새로운 방식으로 인간과 기계를 통합하는 것이 가능합니다.28

However, the development of AI and its convergence with big data also leads to serious risks for individuals, for groups, and for the whole of society. For one thing, AI can eliminate or devalue the jobs of those who can be replaced by machines: many risk losing the 'race against the machine',29 and therefore being excluded from or marginalised in the job market. This may lead to poverty and social exclusion, unless appropriate remedies are introduced (consider, for instance, the future impact of autonomous vehicles on taxi and truck drivers, or the impact of smart chatbots on call-centres workers).

그러나 인공 지능의 개발과 빅 데이터와의 융합은 개인, 그룹 및 사회 전체에 심각한 위험을 초래합니다. 우선 AI는 기계로 교체할 수 있는 사람들의 일자리를 제거하거나 평가 절하할 수 있다. 많은 기계가 '기계와의 경쟁'을 잃을 위험이 있다.29) 따라서 직업 시장에서 제외되거나 소외된다.적절한 구제책이 도입되지 않는 한 빈곤과 사회적 배제로 이어질 수 있다 (예를 들어, 자율 주행 자

동차가 택시 및 트럭 운전자에게 미치는 영향, 또는 스마트 챗봇이 콜센터 근로자에게 미치는 영향을 고려).

Moreover, by enabling big tech companies to make huge profits with a limited workforce, AI contributes to concentrating wealth in those who invest in such companies or provide them with high-level expertise. This trend favours economic models in which 'the winner takes all'. Within companies, monopoly positions tend to prevail, thanks to the network effect (users' preference for larger networks), coupled with economies of scale (enabled by automation) and exclusive or preferential access to data and technologies. Within workers, financial and other benefits, as well as work satisfaction, tend to accrue only to those who can engage in high -level functions that have not yet been automated. To address the adverse impact of AI, appropriate political and social strategies must ensure that everyone will benefit from AI, thanks to workers' training, human-machine interactions focused on engagement and creativity, broader access to data and technologies, wealth redistribution policies.

또한 AI는 대기업이 제한된 인력으로 막대한 수익을 올릴 수 있게함으로써 그러한 회사에 투자하거나 높은 수준의 전문 지식을 제공하는 사람들에게 부를 집중시키는 데 기여합니다. 이러한 경

향은 '승자가 모든 것을 얻는'경제 모델을 선호합니다. 회사 내에서 규모의 경제 (자동화로 가능) 및 데이터 및 기술에 대한 독점적 또는 우선적 접근과 함께 네트워크 효과 (더 큰 네트워크에 대한 사용자 선호)로 인해 독점 위치가 우세한 경향이 있습니다. 근로자 내에서 업무 만족도 뿐만 아니라 재정적 및 기타 혜택은 아직 자동화되지 않은 고급 기능을 수행할 수 있는 사람들에게만 발생하는 경향이 있습니다. AI의 악영향을 해결하려면 적절한 정치 및 사회적 전략을 통해 근로자 교육, 참여 및 창의성에 중점을 둔 인간-기계 상호 작용, 데이터 및 기술에 대한 광범위한 액세스, 자산 재배포 정책을 통해 모든 사람이 AI의 혜택을 누리도록 해야 합니다.

There is also a need to counter the new opportunities for illegal activities offered by AI and big data. In particular, AI and big data systems can fall subject to cyberattacks (designed to disable critical infrastructure, or steal or rig vast data sets, etc.), and they can even be used to commit crimes (e.g., autonomous vehicles can be used for killing or terrorist attacks, and intelligent algorithms can be used for fraud or other financial crimes).30 Even beyond the domain of outright illegal activities, the power of AI can be used to purse economic interests in ways that are harmful to individuals and society: users, consumers, and workers can be subject to

pervasive surveillance, controlled in their access to information and opportunities, manipulated in their choices.

AI와 빅 데이터가 제공하는 불법 활동에 대한 새로운 기회에 대응할 필요도 있습니다. 특히 AI 및 빅 데이터 시스템은 사이버 공격 (중요한 인프라를 비활성화하거나 방대한 데이터 세트를 도용 또는 조작하도록 설계)될 수 있으며 범죄를 저지르는 데 사용될 수도 있습니다 (예 : 자율 주행 차량은 살인 또는 테러 공격, 지능형 알고리즘은 사기 또는 기타 금융 범죄에 사용될 수 있습니다.) 30 불법 행위의 영역을 넘어서도 AI의 힘은 개인과 사회에 해로운 방식으로 경제적 이익을 추구하는 데 사용될 수 있습니다. 사용자, 소비자 및 근로자는 정보 및 기회에 대한 액세스 제어 및 선택에 따라 통제되는 광범위한 감시를 받을 수 있습니다.

Certain abuses may be incentivised by the fact that many tech companies – such as major platforms hosting user-generated content – operate in two- or many-sided markets. Their main services (search, social network management, access to content, etc.) are offered to individual consumers, but the revenue stream comes from advertisers, influencers, and opinion-makers (e.g., in political campaigns). This means not only that any information that is useful for targeted advertising will be collected and used for this purpose, but also that platforms will employ any means to capture

users,

사용자 생성 컨텐츠를 호스팅하는 주요 플랫폼과 같은 많은 기술 회사가 양면 시장에서 운영된다는 사실에 의해 특정 악용이 유발될 수 있습니다. 주요 서비스 (검색, 소셜 네트워크 관리, 콘텐츠 액세스 등)는 개별 소비자에게 제공되지만 수익원은 광고주, 영향력 있는 사람 및 의견 결정자 (예 : 정치 캠페인)에서 비롯됩니다. 이는 타겟 광고에 유용한 정보가 수집되어 이 목적으로 사용될 뿐만 아니라 플랫폼이 사용자를 사로 잡는 수단을 사용한다는 것을 의미합니다.

so that they can be exposed to ads and attempts at persuasion. This may lead not only to a massive collection of personal data about individuals, to the detriment of privacy, but also to a pervasive influence on their behaviour, to the detriment of both individual autonomy and collective interests. Additionally, profit-driven algorithms can combine in order to advance anticompetitive strategies, to the detriment not only competitors but also of consumers. AI also can contribute to polarisation and fragmentation in the public sphere,31 and to the proliferation of sensational and fake news, when used to capture users by exposing them to information they may like, or which accords with their preferences, thereby exploiting their confirmation biases.32

광고에 노출될 수 있고 설득을 시도할 수 있습니다. 이는 개인에 대한 방대한 개인 정보 수집, 개인 정보 보호에 대한 침해 뿐만 아니라 개인의 자율성과 집단적 이익 모두에 해를 끼치는 행동에 대한 광범위한 영향을 초래할 수 있습니다. 또한, 이익 중심 알고리즘은 경쟁자 뿐만 아니라 소비자에게 해를 끼치기 위해 경쟁 방지 전략을 발전시키기 위해 결합할 수 있습니다. AI는 공공 영역에서의 양극화와 단편화에 기여할 수 있다.31) 사용자가 원하는 정보에 노출되거나 선호도에 따라 정보를 노출시켜 확인 편향을 이용하여 사용자를 사로잡을 때 감각적이고 가짜 뉴스의 확산에 기여할 수 있다.32

27 Licklider (1960).
28 McAfee and Brynjolfsson (2019), Mindell (2015).
29 Brynjolfsson and McAfee (2011).
30 Bhuta et al (2015).
31 Sunstein (2007).

Just as AI can be misused by economic actors, it can also be misused by the public section. Governments have many opportunities to use AI for legitimate political and administrative purposes (e.g., efficiency, cost savings, improved services), but they may also employ it to anticipate and control citizens' behaviour in ways that restrict individual liberties and interfere with the

democratic process.

경제 행위자가 AI를 오용할 수 있는 것처럼, 공공 부문에서도 AI를 오용할 수 있습니다. 정부는 합법적인 정치 및 행정 목적 (예 : 효율성, 비용 절감, 개선된 서비스)을 위해 AI를 사용할 수 있는 많은 기회를 가지고 있지만 개인의 자유를 제한하고 민주적 절차를 방해하는 방식으로 시민의 행동을 예측하고 통제하기 위해 AI를 사용할 수도 있습니다..

2.3.3. AI in decision-making concerning individuals: fairness and discrimination (개인에 관한 의사 결정의 인공 지능 : 공정성과 차별)

The combination of AI and big data enables automated decision-making even in domains that require complex choices, based on multiple factors, and on non-predefined criteria. In recent years, wide debate has taken place on the prospects and risks of algorithmic assessments and decisions concerning individuals Some scholars have observed that in many domains automated predictions and decisions are not only cheaper, but also more precise and impartial than human ones. AI systems can avoid the typical fallacies of human psychology (overconfidence, loss

aversion, anchoring, confirmation bias, representativeness heuristics, etc.), and the widespread human inability to process statistical data,33 as well as typical human prejudice (concerning, e.g., ethnicity, gender, or social background). In many assessments and decisions – on investments, recruitment, creditworthiness, or also on judicial matters, such as bail, parole, and recidivism – algorithmic systems have often performed better, according to usual standards, than human experts.34

AI와 빅 데이터의 결합으로 여러 요인 및 사전 정의되지 않은 기준에 따라 복잡한 선택이 필요한 도메인에서도 자동 의사 결정이 가능합니다. 최근 몇년간 개인에 관한 알고리즘 평가 및 결정의 전망과 위험에 대한 광범위한 논쟁이 일어났다. 일부 학자들은 많은 영역에서 자동화된 예측과 결정이 인간의 것보다 저렴할 뿐만 아니라 더 정확하고 공평하다는 것을 관찰했다. AI 시스템은 인간 심리학의 과오 (과신, 손실 혐오, 정박, 확인 편향, 대표성 휴리스틱 스 등)와 통계 데이터를 처리할 수 없는 광범위한 인간의 무능력과 전형적인 인간 편견 (예 : 민족성, 성별 또는 사회적 배경). 투자, 채용, 신용도 또는 보석금, 가석방 및 재범과 같은 사법 문제에 대한 많은 평가와 결정에서 알고리즘 시스템은 보통 인간 표준보다 일반적인 표준에 따라 더 잘 수행되었습니다.34

Others have underscored the possibility that algorithmic decisions

may be mistaken or discriminatory. Only in rare cases will algorithms engage in explicit unlawful discrimination, so-called disparate treatment, basing their outcome on prohibited features (predictors) such as race, ethnicity or gender. More often a system's outcome will be discriminatory due to its disparate impact, i.e., since it disproportionately affects certain groups, without an acceptable rationale.

다른 사람들은 알고리즘 결정이 잘못되거나 차별적 일 가능성을 강조했다. 경우에 따라서는 알고리즘이 인종, 민족 또는 성별과 같은 금지된 지형지물 (예측자)을 기반으로 하여 명시적으로 불법적인 차별, 소위 이질적인 처리를 수행합니다. 시스템의 결과는 다른 영향으로 인해 차별적 일 것입니다. 즉, 합리적인 근거없이 특정 그룹에 불균형하게 영향을 미치기 때문입니다.

As noted in Section 2.2.3, systems based on supervised learning may be trained on past human judgements and may therefore reproduce the strengths and weaknesses of the humans who made these judgements, including their propensities to error and prejudice. For example, a recruitment system trained on the past hiring decisions will learn to emulate the managers' assessment of the suitability of candidates, rather than to directly predict an applicant's performance at work. If past decisions were influenced

by prejudice, the system will reproduce the same logic.35 Prejudice baked into training sets may persist even if the inputs (the predictors) to the automated systems do not include forbidden discriminatory features, such as ethnicity or gender. This may happen whenever a correlation exists between discriminatory features and some predictors considered by the system. Assume, for instance, that a prejudiced human resources manager did not in the past hire applicants from a certain ethnic background, and that people with that background mostly live in certain neighbourhoods. A training set of decisions by that manager will teach the systems not to select people from those neighbourhoods, which would entail continuing to reject applications from the discriminated-against ethnicity.

섹션 2.2.3에 언급된 바와 같이, 지도 학습에 기반한 시스템은 과거의 인간 판단에 대해 훈련될 수 있으며, 따라서 오류와 편견에 대한 경향을 포함하여 이러한 판단을 한 인간의 강점과 약점을 재현할 수 있습니다. 예를 들어, 과거의 채용 결정에 대해 훈련된 채용 시스템은 직장에서 지원자의 성과를 직접 예측하기보다는 관리자의 후보자적합성 평가를 모방하는 방법을 배웁니다. 과거의 결정이 편견에 의해 영향을 받았다면, 시스템은 동일한 논리를 재현할 것이다.35 자동화 시스템에 대한 입력 (예측 자)에 인종이나 성별과 같은 금지된 차별적 특징이 포함되어 있지 않아도 훈련

세트에 구운 편견은 지속될 수 있다. 이것은 차별적 특징과 시스템이 고려한 일부 예측 변수 사이에 상관 관계가 존재할 때 발생할 수 있습니다. 예를 들어, 과거에 선입견이 있는 인적 자원 관리자가 특정 민족적 배경에서 지원자를 고용하지 않았으며 그 배경을 가진 사람들은 대부분 특정 지역에 살고 있다고 가정합니다. 해당 관리자의 교육 결정은 시스템에서 해당 지역의 사람들을 선택하지 않도록 가르칠 것입니다. 이는 차별적인 민족의 응용 프로그램을 계속 거부할 것입니다.

32 Pariser (2011).
33 Kahneman (2011).
34 Kahneman (2011, Ch. 21), Kleinberg et al (2019).
35 Kleinberg et all (2019).

In other cases, a training set may be biased against a certain group, since the achievement of the outcome being predicted (e.g., job performance) is approximated through a proxy that has a disparate impact on that group. Assume, for instance, that the future performance of employees (the target of interest in job hiring) is only measured by the number of hours worked in the office. This outcome criterion will lead to past hiring of women – who usually work for fewer hours than men, having to cope with heavier family burdens – being considered less successful than the hiring of men; based on this correlation (as measured on the basis of the biased

proxy), the systems will predict a poorer performance of female applicants.

다른 경우에, 훈련 세트는 특정 그룹에 대해 편향될 수 있는데, 이는 예측되는 결과의 달성 (예를 들어, 직무 수행)이 그 그룹에 다른 영향을 미치는 대리자를 통해 근사되기 때문이다. 예를 들어, 직원의 미래 성과 (직업 채용에 대한 관심 대상)는 사무실에서 근무한 시간 수로 만 측정된다고 가정하십시오. 이 결과 기준은 과거 남성 고용보다 더 많은 가족 부담에 대처해야하는 과거 여성 고용으로 이어질 것입니다. 이 상관 관계 (바이어스 프록시를 기반으로 측정된)에 기초하여, 시스템은 여성 지원자의 성과가 더 나빠질 것으로 예상합니다.

In other cases, mistakes and discriminations may pertain to the machine-learning system's biases embedded in the predictors. A system may perform unfairly, since it uses a favourable predictor (input feature) that only applies to members of a certain group (e.g., the fact of having attended a socially selective high-education institution). Unfairness may also result from taking biased human judgements as predictors (e.g., recommendation letters).

다른 경우에는 실수와 차별이 예측기에 포함된 기계학습 시스템의 편향과 관련될 수 있습니다. 시스템은 특정 그룹의 구성원에게

만적용되는 유리한 예측 자 (입력 기능)를 사용하기 때문에 불공평하게 수행될 수 있습니다 (예 : 사회적으로 선택적인 고등 교육 기관에 참석한 사실). 불공정성은 편향된 인간의 판단을 예측 자 (예 : 추천서)로 취함으로써 발생할 수 있습니다.

Finally, unfairness may derive from a data set that does reflect the statistical composition of the population. Assume for instance that in applications for bail or parole, previous criminal record plays a role, and that members of a certain groups are subject to stricter controls, so that their criminal activity is more often detected and acted upon. This would entail that members of that group will generally receive a less favourable assessment than members of other groups having behaved in the same ways.

마지막으로 불공평성은 모집단의 통계적 구성을 반영하는 데이터 세트에서 파생될 수 있습니다. 예를 들어 보석 또는 가석방 신청에서 이전 범죄 기록이 중요한 역할을 하고 특정 그룹의 구성원이 더 엄격하게 통제되어 범죄 활동이 더 자주 감지되고 행동한다고 가정합니다. 이것은 일반적으로 그 그룹의 구성원이 같은 방식으로 행동한 다른 그룹의 구성원보다 덜 유리한 평가를 받게 될 것입니다.

Members of a certain group may also suffer prejudice when that group is only represented by a very small subset of the training set, since this will reduce the accuracy of predictions for that group (e.g., consider the case of a firm that has appointed few women in the past and which uses its records of past hiring as its training set).

특정 그룹의 구성원은 해당 그룹이 훈련 세트의 매우 작은 하위 집합으로 만 표현될 경우 편견을 겪을 수 있습니다. 이렇게 하면 해당 그룹에 대한 예측의 정확도가 떨어질 수 있습니다 (예 : 여성을 거의 임명하지 않은 회사의 경우를 고려하십시오) 과거에는 과거의 고용 기록을 훈련 세트로 사용합니다).

It has also been observed that it is difficult to challenge the unfairness of automated decision-making. Challenges raised by the individuals concerned, even when justified, may be disregarded or rejected because they interfere with the system's operation, giving rise to additional costs and uncertainties. In fact, the predictions of machine-learning systems are based on statistical correlations, against which it may be difficult to argue on this basis of individual circumstances, even when exceptions would be justified. Here is the perspective of Cathy O'Neil, a machine-learning expert who has become a critic of the abuses of automation:

또한 자동화된 의사 결정의 불공평성에 도전하기가 어렵다는 것이 관찰되었습니다. 정당한 경우에도 해당 개인이 제기한 문제는 시스템 운영을 방해하여 추가 비용과 불확실성을 야기하기 때문에 무시되거나 거부될 수 있습니다. 실제로 머신러닝 시스템의 예측은 통계적 상관 관계를 기반으로 하므로 예외가 정당화될 때에도 이러한 개별 상황에 근거하여 논쟁하기가 어려울 수 있습니다. 다음은 자동화 학대를 비판하는 기계학습 전문가 인 캐시 오닐(Cathy O'Neil)의 관점입니다.

> ***An algorithm processes a slew of statistics and comes up with a probability that a certain person might be a bad hire, a risky borrower, a terrorist, or a miserable teacher. That probability is distilled into a score, which can turn someone's life upside down.***
>
> ***알고리즘은 많은 통계를 처리하고 특정 사람이 나쁜 고용인, 위험한 차용인, 테러리스트 또는 비참한 교사 일 가능성이 있습니다. 이 확률은 점수로 증류되어 누군가의 삶을 뒤집을 수 있습니다.***
>
> ***And yet when the person fights back, 'suggestive' countervailing evidence simply won't cut it. The case must be ironclad. The human victims of WMDs, we'll see***

time and again, are held to a far higher standard of evidence than the algorithms themselves.36

그러나 그 사람이 반격할 때, '추천적인'반박 증거는 단순히 그것을 잘라 내지 못할 것입니다. 사건은 반드시 철제 처리해야 합니다. 우리는 WMD의 희생자들이 알고리즘 자체보다 훨씬 더 높은 수준의 증거를 가지고 있다는 것을 알게 될 것입니다.36

These criticisms have been countered by observing that algorithmic systems, even when based on machine learning, are more controllable than human decision-makers, their faults can be identified with precision, and they can be improved and engineered to prevent unfair outcomes.

이러한 비판은 머신러닝을 기반으로 하는 알고리즘 시스템이 인간의 의사 결정자보다 제어가 용이하고 결함을 정밀하게 식별할 수 있으며 불공정한 결과를 방지하도록 개선 및 엔지니어링 될 수 있음을 관찰함으로써 반박되었습니다.

[W]ith appropriate requirements in place, the use of algorithms will make it possible to more easily examine

and interrogate the entire decision process, thereby making it far easier to know whether discrimination has occurred. By forcing a new level of specificity, the use of algorithms also highlights, and makes transparent, central trade- offs among competing values. Algorithms are not only a threat to be regulated; with the right safeguards in place, they have the potential to be a positive force for equity.37

적절한 요구 사항이 제정되면 알고리즘을 사용하면 전체 결정 프로세스를 보다 쉽게 조사하고 조사할 수 있어 차별이 발생했는지 여부를 훨씬 쉽게 알 수 있습니다. 새로운 차원의 특이성을 강요함으로써 알고리즘의 사용은 경쟁 가치들 사이에서 투명하고 중심적인 절충점을 강조하고 만듭니다. 알고리즘은 규제될 위협 일 뿐만 아니라; 올바른 보호 수단을 갖추면 자본에 긍정적인 힘이 될 가능성이 있습니다.37

36 O'Neil (2016)

In conclusion, it seems that issues that have just been presented should not lead us to exclude categorically the use of automated decision-making. The alternative to automated decision-making is

not perfect decisions but human decisions with all their flaws: a biased algorithmic system can still be fairer than an even more biased human decision-maker. In many cases, the best solution consists in integrating human and automated judgements, by enabling the affected individuals to request a human review of an automated decision as well as by favouring transparency and developing methods and technologies that enable human experts to analyse and review automated decision-making. In fact, AI systems have demonstrated an ability to successfully also act in domains traditionally entrusted the trained intuition and analysis of humans, such as medical diagnosis, financial investment, the granting of loans, etc. The future challenge will consist in finding the best combination between human and automated intelligence, taking into account the capacities and the limitations of both.

결론적으로 방금 제시된 문제로 인해 자동 의사 결정의 사용을 범주적으로 배제해서는 안됩니다. 자동화된 의사 결정의 대안은 완벽한 의사 결정이 아니라 모든 결점을 가진 인간의 의사 결정입니다. 편향된 알고리즘 시스템은 훨씬 편향된 의사 결정자보다 여전히 공정할 수 있습니다. 대부분의 경우, 최상의 솔루션은 영향을 받는 개인이 자동화된 의사 결정에 대한 인간 검토를 요청할 수 있게 하고 인간 전문가가 자동화된 의사 결정을 분석하고 검토할 수 있는 방법 및 기술을 개발함으로써 영향을 받는 개인

이 자동화된 의사 결정에 대한 인간 검토를 요청할 수 있게 함으로써 인간과 자동 판단을 통합하는 것으로 구성됩니다. -만들기. 실제로 AI 시스템은 전통적으로 의료 진단, 금융 투자, 대출 부여 등과 같이 훈련된 직관과 분석을 위임 받은 영역에서 성공적으로 행동할 수 있는 능력을 보여주었습니다. 미래의 과제는 최상의 조합을 찾는 데 있습니다. 용량과 한계 모두를 고려하여 인간 지능과 자동화 지능 사이.

2.3.4. Profiling, influence and manipulation (프로파일링, 영향 및 조작)

The use of automated assessment systems may be problematic where their performance is not worse, or even is better, than what humans would do. This is due to the fact that automation diminishes the costs of collecting information on individuals, storing this information and process it in order to evaluate individuals and make choices accordingly. Thus, automation paves the way for much more persistent and pervasive mechanisms for assessment and control.

자동화된 평가 시스템의 사용은 사람이 하는 것보다 성능이 나쁘지 않거나 더 나은 경우 문제가 될 수 있습니다. 이는 자동화로

인해 개인에 대한 정보 수집, 이 정보 저장 및 처리 비용이 줄어들 기 때문에 개인을 평가하고적절하게 선택할 수 있기 때문입니다. 따라서 자동화는 평가 및 제어를 위한 훨씬 더 지속적이고 널리 보급된 메커니즘을 위한 길을 열어줍니다.

In general, thanks to AI, all kind of personal data can be used to analyse, forecast and influence human behaviour, an opportunity that transforms them into valuable commodities. Information that was not collected or was discarded as worthless 'data exhaust' – e.g., trails of online activities –has now become a prized resource.

일반적으로 AI 덕분에 모든 종류의 개인 데이터를 사용하여 인간 행동을 분석하고 예측하고 영향을 미치며 가치 있는 상품으로 전환할 수 있습니다. 수집되지 않았거나 무가치한 '데이터 배출'(예 : 온라인 활동의 흔적)로 폐기된 정보는 이제 소중한 자원이 되었습니다.

Through AI and big data technologies – in combination with the panoply of sensor that increasingly trace any human activity – individuals can be subject to surveillance and influence in many more cases and contexts, on the basis of a broader set of personal characteristics (ranging from economic conditions to health

situation, place of residence, personal life choices and events, online and offline behaviour, etc.).

AI 및 빅 데이터 기술을 통해 (인간 활동을 점점 더 많이 추적하는 센서와 함께) 개인은 광범위한 개인 특성 (경제 상황에서 건강 상황, 거주지, 개인 생활 선택 및 행사, 온라인 및 오프라인 행동 등)을 기반으로 더 많은 사례와 상황에서 감시 및 영향을 받을 수 있습니다.

By correlating data about individuals to corresponding classifications and predictions, AI increases the potential for profiling, namely, for inferring information about individuals or groups, and adopting assessments and decisions on that basis. The term 'profile' derives from the Italian 'profilo,' from "profilare," originally meaning to draw a line, especially the contour of an object: that is precisely the idea behind profiling through data processing, which means to expand the available data of individuals of groups, so as to sketch – describe or anticipate – their traits and propensities.

AI는 개인에 대한 데이터를 해당 분류 및 예측과 연관시킴으로써 프로파일링 , 즉 개인 또는 그룹에 대한 정보를 유추하고 이를 바탕으로 평가 및 결정을 채택할 가능성을 높입니다. '프로파일'이라

는 용어는 이탈리아의 '프로필로', '프로필라어'에서 유래합니다. 원래 선, 특히 객체의 윤곽을 그리는 것을 의미합니다. 즉, 데이터 처리를 통한 프로파일링의 개념입니다. 즉, 그룹의 개인이 사용할 수 있는 데이터를 확장하여 특성과 성향을 스케치하거나 설명하거나 예측할 수 있습니다.

A profiling system establishes (predicts) that individuals having certain features F1, also have a certain likelihood of possessing certain additional features F2. For instance, assume that the system establishes (predicts) that those having a genetic patterns have the tendency to develop a higher than average chance to develop cancer, or that those having a certain education and job history or ethnicity have a certain higher-than-average likelihood to default of their debts). Then we may say that this system has profiled the group of the individuals possessing features F1: it has added to the description (the profile) of these group a new segment, namely, the likelihood of possessing the additional features F2. If the system is then given the information that a specific individual has features F1, then the system can infer that it likely that this individual also has feature F2. This may lead to the individual being treated accordingly, in a beneficial or a detrimental way. For instance, in the case in which the inferred feature of an individual is his or her

higher susceptibility to cancer, the system's indication may provide the basis for preventive therapies and tests, or rather for a raise in the insurance premium.

프로파일링 시스템은 특정 기능 F1을 가진 개인이 특정 추가 기능 F2를 가질 가능성이 있음을 확립합니다 (예측). 예를 들어, 시스템이 유전적 패턴을 가진 사람들이 평균 암 발병 가능성보다 더 높은 경향이 있거나 특정 교육 및 직업 이력 또는 민족을 가진 사람들이 채무 불이행에 대한 평균 가능성). 그런 다음이 시스템이 특징 F1을 소유한 개인 그룹을 프로파일링 했다고 말할 수 있습니다. 이 그룹의 설명 (프로파일)에 새로운 세그먼트, 즉 추가 특징 F2를 보유할 가능성을 추가했습니다. 시스템에 특정 개인에게 기능 F1이 있다는 정보가 제공되면 시스템은이 개인에게 기능 F2도 있음을 유추할 수 있습니다. 이로 인해 개인이 유익하거나 해로운 방식으로 치료될 수 있습니다. 예를 들어, 개인의 유추된 특징이 자신의 암에 대한 감수성이 높은 경우, 시스템의 적응증은 예방 요법 및 검사 또는 보험료 인상의 근거를 제공할 수 있습니다.

37 Kleinberg, Ludwig, Mullainathan, and Sunstein (2018, 113).

The information so inferred may also be conditional, that is, it may consist in the propensity to react in a certain way to given inputs. For instance, it may consist in in the propensity to respond to a therapy with improved medical condition, or in the propensity to respond to a certain kind of ad or to a certain price variation with a certain purchasing behaviour, or in the propensity to respond a certain kind of message with a change in mood or preference (e.g., relatively to political choices). When that is the case, profiling potentially leads to influence and manipulation.

그렇게 유추된 정보는 또한 조건부 일 수 있는데, 즉 주어진 입력에 대해 특정 방식으로 반응하는 경향이 있을 수 있다. 예를 들어, 그것은 의학적 상태가 개선된 요법에 반응하는 성향, 또는 특정 종류의 광고 또는 특정 구매 행동에 따른 특정 가격 변동에 반응하는 성향 또는 반응하는 성향으로 구성될 수 있다 기분이나 선호도가 변화한 특정 종류의 메시지 (예 : 정치적 선택에 따라). 이 경우 프로파일링은 잠재적으로 영향과 조작으로 이어집니다.

Assume, too, that the system connects certain values for input features (e.g., having a certain age, gender, social status, personality type, etc.) to the propensity to react to a certain message (e.g., a targeted ad) with a certain response (e.g., buying a certain product). Assume also that the system is told that a particular individual has

these values (he is a young male, working class, extrovert, etc.). Then the system would know that by administering to the individual that message, the individual can probably be induced to deliver the response.

또한 시스템이 입력 기능에 대한 특정 값 (예 : 특정 연령, 성별, 사회적 지위, 성격 유형 등)을 특정 메시지 (예 : 타겟팅된 광고)에 반응하는 경향에 연결한다고 가정합니다. 특정 응답 (예 : 특정 제품 구매). 또한 시스템은 특정 개인이 이러한 가치를 지니고 있다고 가정합니다 (청년, 노동 계급, 외향적 등). 그런 다음 시스템은 개인에게 해당 메시지를 관리함으로써 개인이 응답을 전달하도록 유도될 수 있음을 알 것입니다.

The notion of profiling just presented corresponds to this more elaborate definition:

방금 제시된 프로파일링 개념은 다음과 같은 보다 정교한 정의에 해당합니다.

> ***Profiling is a technique of (partly) automated processing of personal and/or non-personal data, aimed at producing knowledge by inferring correlations from data in the form of profiles that can subsequently be applied as a basis for decision-making. A profile is a set of***

correlated data that represents a (individual or collective) subject. Constructing profiles is the process of discovering unknown patterns between data in large data sets that can be used to create profiles. Applying profiles is the process of identifying and representing a specific individual or group as fitting a profile and of taking some form of decision based on this identification and representation.38

프로파일링은 개인 및/또는 비 개인 데이터의 (부분적으로) 자동 처리 기술로, 이후 의사 결정의 기초로 적용될 수 있는 프로파일 형식의 데이터와의 상관 관계를 유추하여 지식을 생성하는 것을 목표로 합니다. 프로파일은 (개인 또는 집단) 주제를 나타내는 상관된 데이터 세트입니다. 프로파일 구성은 프로파일을 작성하는 데 사용할 수 있는 큰 데이터 세트의 데이터 간에 알 수 없는 패턴을 발견하는 프로세스입니다. 프로파일을 적용하는 것은 특정 개인 또는 그룹을 프로파일에 적합한 것으로 식별하고 나타내는 프로세스이며 이 식별 및 표현에 따라 어떤 형태의 결정을 내립니다.38

The notion of profiling in the GDPR only covers assessments or decisions concerning individuals, based on personal data, excluding

the mere construction of group profiles:

GDPR의 프로파일링 개념은 단순한 그룹 프로파일 구성을 제외하고 개인 데이터를 기반으로 한 개인에 대한 평가 또는 결정에만 적용됩니다.

> ***'profiling'[...] consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health,personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her.***
>
> ***'프로파일링 '[...]은 자연인과 관련된 개인적 측면을 평가하는, 특히 업무, 경제 상황, 건강, 개인적 선호 또는 관심 분야에서 데이터 주체의 성과에 관한 측면을 분석하거나 예측하기 위해 모든 형태의 개인 데이터 자동 처리로 구성됩니다., 신뢰성 또는 행동, 위치 또는 움직임에 관한 법적인 영향을 미치거나 그와 비슷하게 영향을 미칩니다.***

Even when an automated assessment and decision-making system – a profile-based system – is unbiased, and meant to serve beneficial purposes, it may negatively affect the individuals concerned. Those who are subject to pervasive surveillance, persistent assessments and insistent influence come under heavy psychological pressure that affects their personal autonomy, and they are susceptible to deception, manipulation and exploitation in multiple ways.

자동 평가 및 의사 결정 시스템 (프로파일 기반 시스템)이 편견이 없고 유익한 목적을 달성하기 위해 의도된 경우에도 관련 개인에게 부정적인 영향을 줄 수 있습니다. 광범위한 감시, 지속적인 평가 및 지속적인 영향을 받는 사람들은 개인의 자율성에 영향을 미치는 심리적 압박을 받고 있으며 여러 가지 방법으로 기만, 조작 및 착취에 취약합니다.

2.3.5. The dangers of profiling: the case of Cambridge Analytica

(프로파일링의 위험 : Cambridge Analytica의 사례)

The dangers involved in profiling have emerged with clarity in the Cambridge Analytica case, concerning attempts at influencing

voting behaviour – in the United States' 2016 election and possibly also in the Brexit referendum – based of massive processing of personal data. Figure 12 shows the main steps concerning Cambridge Analytica involvement in the US elections.

프로파일링 과 관련된 위험은 개인 정보의 대규모 처리를 기반으로 한 2016년 선거와 Brexit 국민 투표에서 투표 행동에 영향을 미치는 시도와 관련하여 Cambridge Analytica 사례에서 명확하게 나타났습니다. 그림 12는 미국 선거에서의 Cambridge Analytica의 참여와 관련된 주요 단계를 보여줍니다.

38 Bosco et al (2015); see also Hildebrandt, M. (2009).

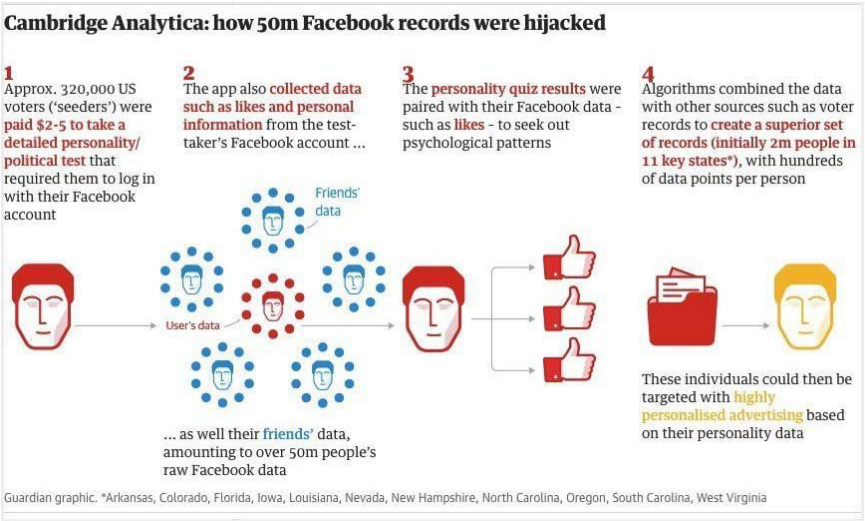

Figure 12 – The Cambridge Analytica case

First of all, people being registered as voters in the USA were invited to take a detailed personality/political test (about 120 questions), available online. The individuals taking the test would be rewarded with a small amount of money (from two to five dollars). They were told that their data would only be used for the academic research.

우선, 미국에서 유권자로 등록된 사람들은 온라인에서 이용할 수 있는 자세한 성격/정치 테스트 (약 120 개의 질문)를 하도록 초대되었습니다. 시험을 치르는 개인에게는 적은 금액의 돈이 있습니다 (2-5 달러). 그들은 자신의 데이터가 학술 연구에만 사용될 것이라고 들었습니다.

About 320 000 voters took the test. In order to be receive the reward each individual taking the test had to provide access to his or her Facebook page (step 1). This allowed the system to connect each individual's answers to the information included in his or her Facebook page.

약 320 000 명의 유권자들이 시험을 보았습니다. 보상을 받으려면 테스트를 받는 각 개인이 자신의 Facebook 페이지에 액세스할 수 있어야했습니다 (1 단계). 이를 통해 시스템은 각 개인의

답변을 자신의 Facebook 페이지에 포함된 정보에 연결할 수 있었습니다.

When accessing a test taker's page, Cambridge Analytica collected not only the Facebook page of test takers, but also the Facebook pages of their friends, between 30 and 50 million people altogether (step 2). Facebook data was also collected from other sources.

시험 응시자 페이지에 액세스할 때 Cambridge Analytica는 응시자의 Facebook 페이지 뿐만 아니라 친구의 Facebook 페이지도 3 천에서 5 천만 명 사이에 수집했습니다 (2 단계). Facebook 데이터는 다른 출처에서도 수집되었습니다.

After this data collection phase, Cambridge Analytica had at is disposition two sets of personal data to be processed (step 3): the data about the test takers, consisting in the information on their Facebook pages, paired with their answers to the questionnaire, and the data about their friends, consisting only in the information on their Facebook pages.

이 데이터 수집 단계 후, Cambridge Analytica는 처리할 두 개인 데이터 세트 (3 단계)를 처리합니다. 시험 응시자에 대한 데이터

는 Facebook 페이지의 정보로 구성되며 설문지에 대한 답변과 함께 Facebook 페이지의 정보로만 구성된 친구에 대한 데이터

Cambridge Analytica used the data about test-takers as a training set for building a model to profile their friends and other people. More precisely, the data about the test-takers constituted a vast training set, where the information on an individual's Facebook pages (likes, posts, links, etc.) provided values for predictors (features) and the answers to the questionnaire (and psychological and political attitudes expressed by such answers) provided values the targets. Thanks to its machine leaning algorithms Cambridge Analytica could use this data to build a model correlating the information in people's Facebook pages to predictions about psychology and political preferences. At this point Cambridge Analytica engaged in massive profiling, namely, in expanding the data available on the people who did not take the test (their Facebook data, and any further data that was available on them), with the predictions provided by the model. For instance, if test-takers having a certain pattern of Facebook likes and posts were classified as having a neurotic personality, the same assessment could be extended also to non-test-takers having similar patterns in their Facebook data.

Cambridge Analytica는 테스트 테이커에 대한 데이터를 친구와 다른 사람들을 프로파일링 하기위한 모델을 구축하기위한 훈련 세트로 사용했습니다. 보다 정확하게, 시험 응시자에 대한 데이터는 개인의 Facebook 페이지 (예 : 게시물, 링크 등)에 대한 정보가 예측 자 (기능) 및 설문지에 대한 답변 (심리적)에 대한 값을 제공하는 방대한 훈련 세트를 구성했습니다. 그러한 답변에 의해 표현된 정치적 태도)는 목표를 가치있게 제공했다. 기계 기울기 알고리즘 덕분에 Cambridge Analytica는이 데이터를 사용하여 사람들의 Facebook 페이지에있는 정보를 심리학과 정치 선호도에 대한 예측과 연관시키는 모델을 만들 수 있었습니다. 이 시점에서 Cambridge Analytica는 대규모 프로파일링, 즉 시험을 치르지 않은 사람들 (Facebook 데이터 및 사용 가능한 추가 데이터)에 대한 데이터를 모델에서 제공한 예측으로 확장하는 데 참여했습니다. 예를 들어, 특정 패턴의 Facebook 좋아요 및 게시물이 신경질적인 성격을 가진 것으로 분류된 테스트 타 커가 동일한 평가를 Facebook 데이터에서 유사한 패턴을 가진 비 테스트 테이커로 확장할 수도 있습니다.

Finally (stage 4), based on this personality/political profiling, potential voters who were likely to change their voting behaviour were identified (in US States in which a small change could make a difference) if prodded with appropriate messages. These voters

where targeted with personalised political ads and with other messages that could trigger the desired change in voting behaviour, possibly building upon their emotions and prejudice and without making them aware of the purpose of such messages.39

마지막으로 (4 단계),이 성격/정치적 프로파일링에 근거하여, 투표 메시지를 변경했을 경우 투표 행동을 바꿀 가능성이 있는 유권자들이 (작은 변화가 변화를 일으킬 수 있는) 미국에서 확인되었습니다. 이러한 유권자들은 개인화된 정치 광고와 투표 행동의 원하는 변화를 유발할 수 있는 다른 메시지를 목표로 했을 것이며, 그러한 감정과 편견을 기반으로 하고 그러한 메시지의 목적을 알지 못하게 할 수도 있습니다.39

2.3.6. Towards surveillance capitalism or surveillance state?

(감시 자본주의 또는 감시 국가를 향해?)

Some authors have taken a positive view of the development of systems based on the massive collection of information. They have observed that the integration of AI and big data enables increased efficiency and provides new means for managing and controlling individual and social behaviour.

When economic transactions – and more generally social interaction and individual activities– are computer-mediated, they provide for a ubiquitous and granular recording of data: computer systems can observe, verify and analyse any aspects of the activities in question.40 The recorded data can be used to construct user profiles, to personalise interactions with users (as in targeted commercial communication), to engage in experimentation (e.g., to evaluate user responses to changes in prices and messaging), to guide and control behaviour (e.g., for the purpose of economic or political persuasion). In this context, new models of economic and social interaction become possible, which are based on the possibility of observing every behaviour, and of automatically linking penalties and rewards to it. Consider for instance how online consumers trust vendors of goods and services with whom they have never had any personal contact, relying on the platform through which such goods and services are provided, and on the platform's methods for rating, scoring, selecting, and excluding. Consider too how blockchain systems – through a shared unmodifiable ledger recording all transactions – enable the creation of digital currencies, self-executing smart contracts, and digital organisations.

일부 저자는 방대한 정보 수집을 기반으로 한 시스템 개발에 대해 긍정적인 견해를 가지고 있습니다. 그들은 AI와 빅 데이터의 통합이 효율성을 높이고 개인 및 사회적 행동을 관리하고 통제할 수 있는 새로운 수단을 제공한다는 것을 관찰했습니다.

경제 거래,보다 일반적으로 사회적 상호 작용 및 개별 활동이 컴퓨터로 매개되는 경우, 데이터가 어디에나 있고 세분화된 데이터 기록을 제공합니다. 컴퓨터 시스템은 해당 활동의 모든 측면을 관찰, 확인 및 분석할 수 있습니다. 사용자 프로필을 구성하고 (대상화된 상업적 커뮤니케이션에서와 같이) 사용자와의 상호 작용을 개인화하고, 실험에 참여하고 (예 : 가격 및 메시지의 변화에 대한 사용자 응답을 평가하기 위해) 행동을 유도하고 제어하기 위해 (예 : 목적을 위해) 경제적 또는 정치적 설득의). 이러한 맥락에서, 모든 행동을 관찰하고 페널티와 보상을 자동으로 연결할 수 있는 가능성을 기반으로 하는 새로운 경제적 사회적 상호 작용 모델이 가능해집니다. 예를 들어 온라인 소비자가 개인 접촉을 하지 않은 제품 및 서비스 공급 업체를 어떻게 신뢰하는지, 해당 제품 및 서비스가 제공되는 플랫폼 및 등급, 점수, 선택 및 제외를 위한 플랫폼의 방법에 의존하는 방법을 고려하십시오. 모든 거래를 기록하는 수정 불가능한 공유 원장을 통해 블록 체인 시스템이 어떻게 디지털 통화 생성, 자체 실행 스마트 계약 및 디지털 조직을 가능하게 하는지를 고려하십시오.

According to Alex Pentland the director of the Human Dynamics Lab at the MIT Media Lab, AI and big data may enable the development of a 'social physics', i.e., a rigorous social science.41 The availability of vast masses of data and of methods and computational resources to process these data could support a social science having solid theoretical-mathematical foundations as well as operational capacities for social governance.

MIT 미디어 연구소의 인간 역학 연구소 소장 인 Alex Pentland에 따르면 AI와 빅 데이터는 '사회 물리학', 즉 엄격한 사회 과학의 개발을 가능하게 할 수 있습니다.41 방대한 양의 데이터와 방법의 가용성 이러한 자료를 처리하기위한 전산 자원은 이론적 수학적 기반이 탄탄한 사회과학 뿐만 아니라 사회 거버넌스를 위한 운영 능력을 지원할 수 있다.

By better understanding ourselves, we can potentially build a world without war or financial crashes, in which infectious disease is quickly detected and stopped, in which energy, water, and other resources are no longer wasted, and in which governments are part of the solution rather than part of the problem.

우리는 자신을 더 잘 이해함으로써 전쟁이나 재정적 충돌 없이 잠재적인 질병을 신속하게 감지하고 중단하며 에너지, 물 및 기타 자원이 더 이상 낭비되지 않고 정부가 해결책의 일부인 세상을 만들 수 있습니다 문제의 일부가 아니라

The prospect for economic and social improvement offered by AI and big data is accompanied by the risks referred to as 'surveillance capitalism' and the 'surveillance state'.

AI와 빅 데이터가 제공하는 경제 및 사회적 개선 전망에는 '감시 자본주의'와 '감시국'이라는 위험이 수반됩니다.

According to Shoshana Zuboff, surveillance capitalism is the leading economic model of the present age.42 Zuboff points out to the classic analysis by historian Karl Polanyi43 who observed that industrial capitalism also treats as commodities (products to be sold in the market) entities that are not produced for the market: human life becomes 'labour' to be bought and sold, nature becomes 'land' or 'real estate', exchange becomes 'money.' As a consequence, the dynamics of capitalism produces destructive tensions – exploitation, destruction of environment, financial crises – unless countervailing forces, such as law, politics and social

organisations (e.g., workers' and consumers' movements), intervene to counteract, moderate and mitigate excesses.

Shoshana Zuboff에 따르면, 감시 자본주의는 현재 시대의 주요 경제 모델입니다.42 Zuboff는 역사 학자 Karl Polanyi43의 고전적 분석에 따르면 산업 자본주의는 또한 시장을 위해 생산되지 않음 : 인간의 삶은 사고 파는 '노동'이 되고 자연은 '토지'또는 '부동산'이 되고 교환은 '돈'이 됩니다. 결과적으로 자본주의의 역학은 법, 정치 및 사회 단체 (예 : 노동자 및 소비자 운동)와 같은 상충하는 세력이 대응, 중간 및 초과분을 완화하십시오.

39 On the problems related to disinformation and propaganda, see Bayer et al (2019).
40 Varian (2010, 2014),
41 Pentland (2015,28),
42 Zuboff (2019), see also Cohen (2019) who prefers to speak of 'informational capitalism."
43Polanyi [1944] 2001),

According to Zuboff, the surveillance capitalism further expands commodification, extending it to human experience, which it turns into recorded and analysed behaviour, i.e., it transforms into marketable opportunities to anticipate and influence.

Zuboff에 따르면, 감시 자본주의는 상품화를 추가로 확대하여 인간 경험으로 확대하여 기록되고 분석된 행동, 즉 예상하고 영향을 미칠 수 있는 시장성 있는 기회로 전환합니다.

> ***Surveillance capitalism annexes human experience to the market dynamic so that it is reborn as behavior: the fourth 'fictional commodity.' Polanyi's first three fictional commodities – land, labor, and money – were subjected to law. Although these laws have been imperfect, the institutions of labor law, environmental law, and banking law are regulatory frameworks intended to defend society (and nature, life, and exchange) from the worst excesses of raw capitalism's destructive power. Surveillance capitalism's expropriation of human experience has faced no such impediments.44***
>
> ***감시 자본주의는 인간의 경험을 시장 역동성에 부가하여 행동으로 다시 태어나도록 합니다. 네 번째 '가상 상품': Polanyi의 첫 3 가지 허구의 상품 (토지, 노동 및 돈)은 법으로 보호되었습니다. 이러한 법은 불완전하지만 노동법, 환경법 및 은행법 제도는 최악의 초과 자본 자본의 파괴력으로부터 사회 (및 자연, 생명 및 교환)를 보호하기위한 규제 프레임 워크입니다. 인간 자본에 대한 감시 자본주의의***

수용은 그러한 장애에 직면하지 않았다.44

Zuboff observes that in the case of surveillance capitalism, raw market dynamics can lead to novel disruptive outcomes. Individuals are subject to manipulation, are deprived of control over their future and cannot develop their individuality. Social networks for collaboration are replaced by surveillance-based mechanism of incentives and disincentives.

Zuboff는 감시 자본주의의 경우 원시 시장 역학이 새로운 파괴적 결과를 초래할 수 있다고 관찰했다. 개인은 조작의 대상이 되고 미래에 대한 통제력이 박탈되어 개성을 개발할 수 없습니다. 협업을 위한 소셜 네트워크는 감시 기반의 인센티브 및 인센티브로 대체됩니다.

Consider for instance, how service platforms – such as Uber or Lyft in the ridesharing section –record the performance of workers as well the mutual reviews of workers and clients, and link multiple aspects of job performance to rewards or penalties. This new way of governing human behaviour may lead to efficient outcomes, but it affects the mental wellbeing and autonomy of the individuals concerned.45 According to Zuboff, we have not yet developed

adequate legal, political or social measures by which to check the potentially disruptive outcomes of surveillance capitalism and keep them in balance. However, she observes, the GDPR could be an important step in this direction, as a 'springboard to challenging the legitimacy of surveillance capitalism and ultimately vanquishing its instrumentarian power', towards 'society's rejection of markets based on the dispossession of human experience as a means to the prediction and control of human behavior for others' profit.'

예를 들어 승차 공유 섹션의 Uber 또는 Lyft와 같은 서비스 플랫폼이 근로자의 성과 뿐만 아니라 근로자와 고객의 상호 검토를 기록하고 작업 성과의 여러 측면을 보상 또는 위약금과 연계시키는 방법을 고려하십시오. 이 새로운 인간 행동 관리 방식은 효율적인 결과를 초래할 수 있지만 관련 개인의 정신 건강과 자율성에 영향을 미칩니다. 감시 자본주의의 결과와 균형을 유지하십시오. 그러나 그녀는 GDPR이이 방향으로 중요한 단계가 될 수 있다고 지적한다. 다른 사람의 이익을 위한 인간 행동의 예측과 통제를 의미한다. '

The need to limit the commercial use of personal data has led to new legal schemes not only in Europe, but also in California, the place where many world-leading 'surveillance capitalists' have their roots; the CCPA (California Consumer Privacy Act), which came into

effect on January 2020, provides consumers with rights to access their data and to prohibit data sales (broadly understood).

개인 데이터의 상업적 사용을 제한할 필요성으로 인해 유럽 뿐만 아니라 많은 세계적 '감시 자본가'가 뿌리를 내린 캘리포니아에서도 새로운 법적 제도가 생겨났습니다. 2020년 1 월에 발효된 CCPA (California Consumer Privacy Act)는 소비자에게 데이터에 액세스하고 데이터 판매를 금지할 수 있는 권리를 제공합니다 (광범위하게 이해됨).

At the governmental level, surveillance capitalism finds its parallel in the so-called 'surveillance state', which is characterised as follows:

정부 차원에서 감시 자본주의는 소위 '감시 상태'에서 그와 비슷한 점을 발견합니다.

> ***In the National Surveillance State, the government uses surveillance, data collection, collation, and analysis to identify problems, to head off potential threats, to govern populations, and to deliver valuable social services. The National Surveillance State is a special case of the Information State-a state that tries to identify and***

solve problems of governance through the collection, collation, analysis, and production of information.46

국가 감시국 (National Surveillance State)에서 정부는 감시, 데이터 수집, 데이터 정렬 및 분석을 사용하여 문제를 식별하고, 잠재적인 위협을 피하고, 인구를 통제하고, 귀중한 사회 서비스를 제공합니다. 국가 감시국은 정보 수집, 대조, 분석 및 정보 생산을 통해 지배 구조 문제를 식별하고 해결하려는 정보국의 특수한 사례이다.46

44 Zuboff (2019, 507).
45 Cristianini, and Scantamburlo (2019).
46 Balkin (2008, 3).

In government too, AI and big data can bring great advantages, supporting efficiency in managing public activities, coordinating citizens' behaviour, and preventing social harms. However, they may also enable new kinds of influence and control, underpinned by purposes and values that may conflict with the requirements of democratic citizenship. A paradigmatic example is that of the Chinese Social credit systems, which collects data about citizens and assigns to those citizens scores that quantify their social value and reputation. This system is based on the aggregation and

analysis of personal information. The collected data cover financial aspects (e.g., timely compliance with contractual obligations), political engagement (e.g., participation in political movements and demonstrations), involvement in civil and criminal proceedings (past and present) and social action (e.g. participation in social networks, interpersonal relationships, etc.). On the basis these data items, citizens may be assigned positive or negative points, which contribute to their social score. A citizen's overall score determines his or her access to services and social opportunities', such as universities, housing, transportation, jobs, financing, etc. The system's purported objective is to promote mutual trust, and civic virtues. One may wonder whether opportunism and conformism may be rather promoted to the detriment of individual autonomy and genuine moral and social motivations.

정부에서도 AI와 빅 데이터는 공공 활동 관리, 시민의 행동 조정 및 사회적 피해 예방에 효율성을 지원하는 큰 이점을 가져올 수 있습니다. 그러나 민주주의 시민의 요구와 상충될 수 있는 목적과 가치에 의해 뒷받침되는 새로운 종류의 영향력과 통제를 가능하게 할 수도 있습니다. 패러다임의 예는 시민에 관한 데이터를 수집하고 그들의 시민 가치에 그들의 사회적 가치와 명성을 정량화하는 중국 사회 신용 시스템의 예입니다. 이 시스템은 개인 정보의 집계 및 분석을 기반으로 합니다. 수집된 데이터는 재정적 측

면 (예 : 계약 의무 준수), 정치적 참여 (예 : 정치 운동 및 시위 참여), 민사 및 형사 소송 참여 (과거 및 현재) 및 사회적 행동 (예 : 소셜 네트워크 참여, 대인 관계 등). 이러한 데이터 항목을 기반으로 시민에게 긍정적 또는 부정적 점수가 할당되어 소셜 점수에 기여할 수 있습니다. 시민의 전체 점수는 대학, 주택, 교통, 일자리, 자금 조달 등과 같은 서비스 및 사회적 기회에 대한 액세스를 결정합니다. 이 시스템의 목표는 상호 신뢰와 시민의 미덕을 장려하는 것입니다. 기회주의와 순응주의가 개인의 자율성과 진정한 도덕적, 사회적 동기에 해를 끼칠 수 있는지 궁금할 것이다.

Thus, the perspective of an integration or symbiosis between humans and intelligent machine, while opening bright prospects, does not entail that all applications of AI should be accepted as long as they meet technological and fairness standards. It has been argued that following this approach

따라서, 인간과 지능형 기계 사이의 통합 또는 공생 관점은 밝은 전망을 열어 주지만 기술과 공정성 표준을 충족하는 한 모든 AI 적용을 수용해야 하는 것은 아닙니다. 이 접근법을 따르는 것은 논쟁의 여지가 있다

What is achieved is resignation – the normalization of

massive data capture, a one-way transfer to technology companies, and the application of automated, predictive solutions to each and every societal problem.47

대규모 데이터 캡처의 정규화, 기술 회사로의 단방향 전송, 모든 사회적 문제에 대한 자동화된 예측 솔루션의 적용 등 사직이 이루어집니다.47

Indeed, in some cases and domain AI and big data applications – even when accurate and unbiased– may have individual and social costs that outweigh their advantages. To address these cases, we need to go beyond requiring unbiasedness and fairness, and ask further questions, which may challenge the very admissibility of the AI applications at stake.

실제로, 경우에 따라서는 도메인 AI 및 빅 데이터 애플리케이션이 정확하고 편견이 없는 경우에도 장점보다 중요한 개인 및 사회적 비용이 있을 수 있습니다. 이러한 경우를 해결하려면 편견과 공정성을 요구하는 것 이상의 문제를 해결하고 추가 질문을 해야 합니다. 이로 인해 위기에 처한 AI 응용 프로그램의 허용 가능성에 도전할 수 있습니다.

Which systems really deserve to be built? Which problems most need to be tackled? Who is best placed

to build them? And who decides? We need genuine accountability mechanisms, external to companies and accessible to populations. Any A.I. system that is integrated into people's lives must be capable of contest, account, and redress to citizens and representatives of the public interest.48

어떤 시스템을 실제로 구축해야 합니까? 어떤 문제를 해결해야합니까? 누가 그들을 구축하는 것이 가장 좋은가? 그리고 누가 결정합니까? 우리는 회사 외부에서, 그리고 사람들이 이용할 수 있는 진정한 책임 메커니즘이 필요합니다. 모든 AI 사람들의 삶에 통합된 시스템은 시민과 공익의 대표자들과 경쟁하고 설명하고 구제할 수 있어야 합니다.48

Consider, for instance, systems that are able to recognise sexual orientation, or criminal tendencies from the faces of persons. Should we just ask that whether these systems provide reliable assessments, or should we rather ask whether they should be built at all. Should we 'ban them, or at least ensure they are only licensed for socially productive uses?'49 The same may concern extremely intrusive ways to monitor, analyse, punish or reword the behaviour of workers by online platforms for transportation (e.g. Uber) or

other services. Similarly, some AI-based financial application, even when inclusive, may have a negative impact on their addressees, e.g., pushing them into perpetual debt.50

예를 들어 성적 취향을 인식할 수 있는 시스템이나 사람의 얼굴에서 발생하는 범죄 경향을 고려하십시오. 이러한 시스템이 신뢰할 수 있는 평가를 제공하는지 여부를 묻거나 시스템을 전혀 구축해야 하는지 여부를 묻어 야합니다. 우리는 '금지하거나 최소한 사회적으로 생산적인 용도로만 라이센스를 받아야합니까? '49) 이와 같은 것은 온라인 플랫폼 (예 : Uber) 또는 다른 서비스. 마찬가지로 일부 AI 기반 금융 응용 프로그램은 포괄적인 경우에도 수취인에게 부정적인 영향을 미칠 수 있습니다.

2.3.7. The general problem of social sorting and differential treatment (사회 분류 및 차별 처리의 일반적인 문제)

The key aspect of AI system, of the machine learning type, is their ability to engage in differential inference: different combinations of predictor-values are correlated to different predictions. As discussed above, when the predictors concern data on individuals and their behaviour, the prediction also concerns features or attitudes of such individuals. Thus, for instance, as noted above, a

certain financial history, combined with data on residence or internet use, can lead to a prediction concerning financial reliability and possibly to a credit score.

머신러닝 유형의 AI 시스템의 주요 측면은 차등 추론에 참여할 수 있는 능력입니다. 예측 값의 서로 다른 조합은 서로 다른 예측과 관련이 있습니다. 전술한 바와 같이, 예측자가 개인 및 그들의 행동에 관한 데이터에 관한 경우, 예측은 또한 그러한 개인의 특징 또는 태도에 관한 것이다. 따라서, 예를 들어, 위에서 언급한 바와 같이, 거주 또는 인터넷 사용에 관한 데이터와 결합된 특정 재무 이력은 재무 신뢰성에 관한 예측 및 가능하게는 신용 점수로 이어질 수 있다.

47 Powles and Nissenbaum (2018).
48 Powles and Nissenbaum (2018).
49 Pasquale (2019).
50 Pasquale (2019).

A new dynamic of stereotyping and differentiation takes place. On the one hand, the individuals whose data support the same prediction, will be considered and treated in the same way. On the other hand, the individuals whose data support different predictions, will be considered and treated differently.

고정 관념과 차별화의 새로운 역학이 발생합니다. 한편으로, 데이터가 동일한 예측을 지원하는 개인은 동일한 방식으로 고려되고 처리됩니다. 반면에, 데이터가 다른 예측을 지원하는 개인은 다르게 고려되고 취급될 것입니다.

This equalisation and differentiation, depending on the domains in which it is used and on the purposes that it is meant to serve, may affect positively or negatively the individuals concerned but also broader social arrangements.

이 평등화와 차별화는 그것이 사용되는 영역과 그것이 제공하려는 목적에 따라 관련된 개인 뿐만 아니라 더 넓은 사회적 배치에도 긍정적 또는 부정적으로 영향을 줄 수 있습니다.

Consider for instance the use of machine learning technologies to detect or anticipate health issues. When used to direct patients to therapies or preventive measures that are most suited to their particular conditions, these AI applications are certainly beneficial, and the benefits outweigh – at least when accompanied by corresponding security measures – whatever risks that may be linked to the abuse of patients' data. The benefits, moreover, concern in principle all data subjects whose data are processed for

this purpose, since each patient has an interest in a more effective and personalised treatment. Processing of health-related data may also be justified on grounds of public health (Article 9 (2)(h)), and in particular for the purpose of 'monitoring epidemics and their spread' (Recital 46). This provision has become hugely relevant in the context of the Coronavirus disease 2019 (COVID-19) epidemics. In particular a vast debate has been raised by development of applications for tracing contacts, in order to timely monitor the diffusion of the infection.51 AI is being applied in the context of the epidemics in multiple ways, e.g., to assess symptoms of individuals and to anticipate the evolution of the epidemics. Such processing should be viewed as legitimate as long as it effectively contributes to limit the diffusion and the harmfulness of the epidemics, assuming that the privacy and data protection risks are proportionate to the expected benefit, and that appropriate mitigation measures are applied.

예를 들어 머신러닝 기술을 사용하여 건강 문제를 감지하거나 예상하십시오. 환자를 특정 조건에 가장적합한 치료 또는 예방 조치로 안내하는 데 사용될 때, 이러한 AI 응용 프로그램은 확실히 유익하며, 적어도 해당 보안 조치가 수반되는 경우 그 혜택은 그 남용과 관련된 위험에 관계없이 중요합니다. 환자의 데이터. 또한, 각각의 환자는 보다 효과적이고 개인화된 치료에 관심을 갖기 때

문에, 이러한 목적을 위해 데이터를 처리하는 모든 데이터 주체는 원칙적으로 이점에 관심을 갖는다. 보건 관련 자료의 처리는 공중 보건의 근거에 따라 정당화될 수 있으며 (제9조 제2 항 (h)), 특히 '전염병 및 확산 모니터링'(Recital 46)의 목적으로 정당화될 수 있다. 이 조항은 2019년 코로나 바이러스 질병 (COVID-19) 전염병과 관련하여 크게 관련이 있습니다. 특히 감염 확산을 적시에 모니터링하기 위해 접촉 추적 응용 프로그램을 개발하여 광범위한 논쟁이 제기되었습니다.51 AI는 전염병과 관련하여 개인의 증상을 평가하기 위해 여러 가지 방법으로 적용되고 있습니다. 전염병의 진화를 예상합니다. 개인 정보 보호 및 데이터 보호 위험이 예상 이익에 비례하고 적절한 완화 조치가 적용되는 경우 전염병의 확산 및 유해성을 제한하는 데 효과적으로 기여하는 한 이러한 처리는 합법적인 것으로 간주해야 합니다.

The use of the predictions based on health data in the context of insurance deserves a much less favourable assessment. In this case there would be some gainers, namely the insured individuals getting a better deal based on their favourable heath prospects, but also some losers, namely those getting a worse deal because of their unfavourable prospects. Thus, individuals who already are disadvantaged because of their medical conditions would suffer further disadvantage, being excluded from insurance or being

subject to less favourable conditions. Insurance companies having the ability (based on the data) to distinguish the risks concerning different applicants would have a competitive advantage, being able to provide better conditions to less risky applicants, so that insurers would be pressured to collect as much personal data as possible.

보험 상황에서 건강 데이터를 기반으로 한 예측의 사용은 훨씬 덜 유리한 평가를 받을 가치가 있습니다. 이 경우, 피보험자 개인은 자신의 유리한 가망 고객 전망을 기반으로 더 나은 거래를 얻는다. 또한 일부 패자, 즉 불리한 잠재 고객으로 인해 더 나쁜 거래를 하는 사람들도 있습니다. 따라서 의료 상태로 인해 이미 불이익을 받은 개인은 보험에서 제외되거나 덜 유리한 조건에 처해질 수 있는 추가적인 불이익을 겪게 됩니다. 다른 신청자와 관련된 위험을 구별할 수 있는 능력 (데이터를 기반으로)을 보유한 보험 회사는 경쟁이 덜 유리하며 덜 위험한 신청자에게 더 나은 조건을 제공할 수 있으므로 보험 회사는 가능한 많은 개인 데이터를 수집해야 합니다.

Even less commendable would be the use of health predictions in the context of recruiting, which would involve burdening less healthy people with unemployment or with harsher work conditions. Competition between companies would also be

affected, and pressure for collecting health data would grow.

채용 상황에서 건강 예측을 사용하는 것은 칭찬 할 만한 일이 아니며, 이는 실업이 있거나 열악한 노동 조건을 가진 건강한 사람에게 부담을 줄 것입니다. 회사 간 경쟁에도 영향을 미치고 건강 데이터 수집에 대한 압력이 커질 것입니다.

Let us finally consider the domain of targeted advertising. In principle, there seems to be nothing wrong in providing consumers with ads match their interests, helping them to navigate the huge set of options that are available online. However, personalised advertising involves the massive collection of personal data, which is used in the interests of advertisers and intermediaries, possibly against the interests of data subjects. Such data provide indeed new opportunities for influence and control, they can be used to delivers deceitful, or aggressive messages, or generally messages that bypass rationality by appealing to weaknesses and emotions.

최종적으로 타겟팅 된 광고의 영역을 살펴보겠습니다. 원칙적으로 소비자에게 자신의 관심사에 맞는 광고를 제공하여 온라인에서 사용할 수 있는 다양한 옵션을 탐색하는 데 아무런 문제가 없는 것 같습니다. 그러나 개인화된 광고는 개인 정보의 대규모 수집과 관련이 있으며, 이는 데이터 주체의 이익에 반하는 광고주 및 중

개인의 이익을 위해 사용됩니다. 이러한 데이터는 실제로 영향력과 통제를 위한 새로운 기회를 제공하며, 속임수 또는 공격적인 메시지 또는 일반적으로 약점과 감정에 호소하여 합리성을 우회하는 메시지를 전달하는 데 사용될 수 있습니다.

51 See the European Data Protection Board Guidelines 04/2020 on the use of location data and contact-tracing tools in the context of the Covid-19 outbreak.

Rather than predominantly stimulating the development and exercise of conscious and deliberate reason, today's networked information flows [...] employ a radical behaviorist approach to human psychology to mobilize and reinforce patterns of motivation, cognition, and behavior that operate on automatic, near- instinctual levels and that may be manipulated instrumentally.52

오늘날의 네트워크화 된 정보 흐름은 의식적이고 의도적인 이유의 개발과 운동을 주도적으로 자극하기보다는 [...] 인간의 심리학에 근본적인 행동주의 접근 방식을 사용하여 자동, 거의 본능적인 수준에서 작동하는 동기 부여,인지 및

행동 패턴을 동원하고 강화합니다. 그것은 도구적으로 조작될 수 있습니다.52

Thus, people may be induced to purchase goods they do not need, to overspend, to engage in risky financial transactions, to indulge in their weaknesses (e.g. gambling or drug addiction). The opportunity for undue influence is emphasised by the use of psychographic techniques that enable psychological attitudes to be inferred from behaviour, and thus disclose opportunities for manipulation.53

Even outside of the domain of aggressive or misleading advertising, we may wonder what real benefits to consumers and to society may be delivered by practices such as price discrimination, namely, the policy of providing different prices and different conditions to different consumers, depending on predictions on their readiness to pay. Economist have observed that this practice may not only harm consumers but also affect the functioning of markets.

따라서 사람들은 불필요하게 금융 거래에 참여하거나 약점 (예 :

도박 또는 마약 중독)에 빠지기 위해 필요하지 않은 상품을 구매하도록 유도될 수 있습니다. 심리적 태도를 행동에서 유추하여 조작할 수 있는 기회를 공개할 수 있는 심리적 기법을 사용함으로써 과도한 영향력을 행사할 수 있는 기회가 강조됩니다.53

공격적이거나 오해의 소지가 있는 광고 영역을 벗어나도 예측에 따라 가격 차별과 같은 관행, 즉 다른 소비자에게 다른 가격과 다른 조건을 제공하는 정책과 같은 관행을 통해 소비자와 사회에 어떤 실질적인 이점이 제공되는지 궁금할 수 있습니다. 지불할 준비가 되어있는 이코노미스트는 이 관행이 소비자에게 해를 끼칠 뿐만 아니라 시장의 기능에도 영향을 미칠 수 있다는 것을 관찰했습니다.

> ***Because AI and big data enable firms to assess how much each individual values different products and is therefore willing to pay, they give these firms the power to price discriminate, to charge more to those customers who value the product more or who have fewer options. Price discrimination not only is unfair, but it also undermines the efficiency of the economy: standard economic theory is based on the absence of discriminatory pricing.54***

AI와 빅 데이터를 통해 기업은 각 개인이 서로 다른 제품을 얼마나 소중하게 생각하고 지불할 의사가 있는지 평가할 수 있기 때문에 이러한 기업은 가격을 차별화할 수 있는 힘을 제공하여 제품을 더 중요하게 생각하거나 옵션이 적은 고객에게 더 많은 비용을 청구할 수 있습니다. 가격 차별은 불공평 할 뿐만 아니라 경제의 효율성을 저해한다. 표준 경제 이론은 차별적 가격의 부재에 근거한다.54

The practice of price discrimination shows how individuals may be deprived of access to some opportunities when they are provided with personalised informational environment engineered by third parties, i.e., with informational cocoons where they are presented with data and choices that are selected by others, according to their priorities.

가격 차별 행위는 개인이 제3자가 설계한 개인화된 정보 환경, 즉 다른 사람이 선택한 데이터와 선택 사항이 제공되는 정보 고치와 함께 제공될 때 일부 기회에 대한 접근을 박탈할 수 있는 방법을 보여줍니다. 그들의 우선 순위.

Similar patterns characterise the political domain, where targeted ads and messages can enable political parties to selectively appeal

to individuals having different political preferences and psychological attitudes, without them knowing what messages are addressed to other voters, in order to direct such individuals towards the desired voting behaviour, possibly against their best judgement. In this case too, it may be wondered whether personalisation really contributes to the formation of considered political opinions, or whether it is averse to it. After the Cambridge Analytica case, some internet companies have recognised how microtargeted political advertising may negatively affect the formation of political opinion, and have consequently adopted some remedial measures. Some have refused to transmit paid political ads (Twitter), others have restricted the factors used for targeting, only allowing general features such as age, gender, or residence code, to the exclusion of other aspects, such as political affiliation or public voter records (Google).

유사한 패턴은 정치적 영역을 특징으로 하며, 타겟팅된 광고 및 메시지를 통해 정당이 원하는 투표로 향하도록 하기 위해 다른 유권자에게 어떤 메시지가 전달되는지 모른 채 다른 정치적 선호도와 심리적 태도를 가진 개인에게 선택적으로 호소할 수 있습니다. 최선의 판단에 반대되는 행동. 이 경우에도 개인화가 고려된 정치적 의견의 형성에 실제로 기여하는지 또는 그것이 반대인지에 대해 의문을 가질 수 있습니다. Cambridge Analytica 사건 이

후 일부 인터넷 회사는 소 표적 정치 광고가 정치적 견해에 부정적인 영향을 줄 수 있다는 점을 인식하고 결과적으로 일부 개선 조치를 채택했습니다. 일부는 유료 정치 광고 (Twitter)의 전송을 거부했으며 다른 일부는 연령, 성별 또는 거주지 코드와 같은 일반 기능 만 정치적 제휴 또는 공개 유권자 기록과 같은 다른 측면을 제외하도록 허용하는 데 사용되는 요소를 제한했습니다. (구글).

In conclusion we may say that AI enables new kinds of algorithmic mediated differentiations between individuals, which need to be strictly scrutinised. While in the pre-AI era differential treatments could be based on the information extracted through individual interactions (the typical job interview) and human assessments, or on few data points whose meaning was predetermined, in the AI era differential treatments can be based on vast amounts of data enabling probabilistic predictions, which may trigger algorithmically predetermined responses.

결론적으로 AI는 개인간에 새로운 종류의 알고리즘 매개 차별화를 가능하게 한다고 말할 수 있습니다. AI 이전 시대의 차등 치료는 개별 상호 작용 (일반적인 면접)과 인간 평가를 통해 추출된 정보를 기반으로 할 수 있지만, AI 시대의 차등 치료는 의미가 미리 정해진 몇 가지 데이터 포인트를 기반으로 할 수 있지만 알고

리즘적으로 미리 결정된 응답을 트리거 할 수 있는 확률적 예측을 가능하게 하는 데이터의 양.

52 Cohen 2019
53 Burr and Cristianini (2019).
54 Stiglitz (2019, 115).

The impacts of such practices can go beyond the individuals concerned, and affect important social institution, in the economical as well as in the political sphere.

이러한 관행의 영향은 경제 분야와 정치 분야에서 관련 개인을 넘어 중요한 사회 제도에 영향을 미칠 수 있습니다.

The GDPR, as we shall see in the following section, provides some constraints: the need for a legal basis for any processing of personal data, obligations concerning information and transparency, limitations on profiling and automated decision-making, requirements on anonymisation and pseudonymisation, etc. These constraints, however, need to be coupled with strong public oversight, possibly leading to the ban of socially obnoxious forms of differential treatment, or to effective measures that

prevent abuses. The decision on what forms of algorithmic differentiations to allow is a highly political one, which should be entrusted to technical authorities only under the direction of politically responsible bodies, such as in particular, parliamentary assemblies. It is a decision that concerns what society we want to live in, under what arrangement of powers and opportunities.

다음 섹션에서 볼 수 있듯이 GDPR은 몇 가지 제약 조건을 제공합니다. 개인 데이터 처리에 대한 법적 근거의 필요성, 정보 및 투명성에 대한 의무, 프로파일링 및 자동 의사 결정에 대한 제한, 익명화 및 가명 화에 대한 요구 사항, 그러나 이러한 제약은 강력한 공공 감독과 연계되어 사회적으로 불쾌한 형태의 차별적 대우를 금지하거나 남용을 방지하는 효과적인 조치를 취해야 합니다. 어떤 형태의 알고리즘 차별화가 허용되는지에 대한 결정은 매우 정치적이며, 특히 국회와 같이 정치적으로 책임있는 기구의 지시에 따라 기술 당국에 맡겨 져야합니다. 어떤 권력과 기회의 배열 하에서 우리가 살고 싶은 사회에 관한 결정입니다.

2.4. AI, legal values and norms (AI, 법적 가치 및 규범)

To promote valuable practices around the use of AI, we need to ensure that the development and deployment of AI takes place in

a sociotechnical framework (inclusive of technologies, human skills, organisational structures, and norms) where individual interests and social goods are both preserved and enhanced.

AI 사용에 관한 가치 있는 관행을 장려하기 위해서는 AI의 개발 및 배포가 기술, 인간 기술, 조직 구조 및 규범을 포함한 사회 공학적 프레임 워크에서 개인의 이익과 사회 용품이 모두 보존되도록해야 합니다. 강화되었습니다.

To provide regulatory support to the creation of such a framework, we need to focus not only on existing regulations, but also on first principles, given that the current rules may fail to provide appropriate solutions and directions to citizens, companies and enforcement authorities. First principles include fundamental rights and social values at both the ethical and the legal level.

이러한 프레임 워크 작성에 규제 지원을 제공하려면 현재 규정이 시민, 회사 및 집행 기관에 적절한 솔루션과 지침을 제공하지 못할 수 있으므로 기존 규정 뿐만 아니라 첫 번째 원칙에도 초점을 맞춰야 합니다. 첫 번째 원칙에는 윤리적 수준과 법적 차원에서 기본적인 권리와 사회적 가치가 포함됩니다.

2.4.1. The ethical framework

A high-level synthesis of the ethical framework for AI is provided for instance by the AI4People document, which describes the opportunities provided by AI and the corresponding risks as follows:15

AI에 대한 윤리적 프레임 워크의 높은 수준의 합성은 AI4People 문서에 의해 제공되며, 여기에는 AI가 제공하는 기회와 해당 위험을 다음과 같이 설명합니다.15

- enabling human self-realisation, without devaluing human abilities;

  인간의 능력을 평가하지 않고 인간의 자기 실현을 가능하게 한다.

- enhancing human agency, without removing human responsibility; and

  인간의 책임을 제거하지 않고 인간의 선택 의지를 향상시킨다. 그리고

- cultivating social cohesion, without eroding human self-

determination.

인간의 자기 결정을 침식하지 않으면서 사회적 응집력을 키울 수 있습니다.

The High-Level Expert Group on Artificial Intelligence, set up by the European Commission, recently published a set of ethics guidelines for trustworthy AI. According to the expert group, the foundation of legal, ethical and robust AI should be grounded on fundamental rights and reflect the following four ethical principles:

유럽위원회가 설립한 인공 지능에 관한 고급 전문가 그룹은 최근 신뢰할 수 있는 AI에 대한 일련의 윤리 지침을 발표했다. 전문가 그룹에 따르면, 합법적이고 윤리적이며 강력한 AI의 기초는 기본 권리에 기초해야 하며 다음 4 가지 윤리 원칙을 반영해야 합니다.

- Respect for human autonomy: humans interacting with AI must be able to keep full and effective self-determination over themselves. AI should not unjustifiably subordinate, coerce, deceive, manipulate, condition or herd humans, but should be rather designed to augment, complement and empower human cognitive, social and cultural skills.

인간 자율성 존중 : AI와 상호 작용하는 인간은 자신에 대해 완전하고 효과적인 자기 결정을 유지할 수 있어야 합니다. 인공 지능은 인간을 부당하게 종속적, 강제적, 속임수, 조작, 조건 또는 무리를 지어서는 안되며, 인간의 인지적, 사회적, 문화적 기술을 강화, 보완 및 강화하도록 설계되어야 합니다.

- Prevention of harm: the protection of human dignity as well as mental and physical integrity should be ensured. Under this principle, AI systems and the environments in which they operate must be safe and secure, they should neither cause nor exacerbate harm or otherwise adversely affect human beings.

  피해 예방 : 인간의 존엄성과 정신적, 신체적 무결성을 보호해야 합니다. 이 원칙에 따라 AI 시스템과 시스템이 작동하는 환경은 안전하고 안전해야 하며 인체에 해를 끼치거나 악화 시키거나 악영향을 미쳐서는 안됩니다.

- Fairness: it should be intended under its substantive and procedural dimension. The substantive dimension implies a commitment to: ensuring equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation. The procedural dimension entails the ability

to contest and seek effective redress against decisions made by AI systems and by the humans operating them.

공정성 : 실질적이고 절차적인 차원에서 의도되어야 합니다. 실질적인 차원은 이익과 비용의 균등하고 정당한 분배를 보장하고 개인과 그룹이 불공정한 편견, 차별 및 낙인이 없도록 보장한다는 약속을 의미합니다. 절차적 차원은 인공 지능 시스템과 시스템을 운영하는 인간의 결정에 대항하여 효과적으로 대응하고 경쟁할 수 있는 능력을 수반합니다.

- Explicability: algorithmic processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions explainable to those affected both directly and indirectly.

  설명 가능성 : 알고리즘 프로세스는 투명해야 하고, AI 시스템의 기능과 목적은 공개적으로 전달되어야 하며 의사 결정은 직간접적으로 영향을 받는 사람들에게 설명 가능해야 합니다.

According to the High-Level Expert Group, in order to implement and achieve trustworthy AI, seven requirements should be met, building on the principles mentioned above:

고급 전문가 그룹에 따르면 신뢰할 수 있는 AI를 구현하고 달성하려면 위에서 언급한 원칙을 바탕으로 7 가지 요구 사항을 충족해야 합니다.

- Human agency and oversight, including fundamental rights;

  기본권을 포함한 인간의 선택 의지와 감독;

- Technical robustness and safety, including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility;

  공격 및 보안에 대한 복원력, 대체 계획 및 일반적인 안전, 정확성, 신뢰성 및 재현성을 포함한 기술적 견고성 및 안전성;

- Privacy and data governance, including respect for privacy, quality and integrity of data, and access to data;

  개인 정보 보호, 데이터 품질 및 무결성, 데이터 액세스를 포함한 개인 정보 및 데이터 관리;

- Transparency, including traceability, explainability and communication;

  추적 성, 설명 성 및 의사 소통을 포함한 투명성;

- Diversity, non-discrimination and fairness, including the avoidance of unfair bias, accessibility and universal design, and stakeholder participation;

  불공정한 편견 회피, 접근성 및 보편적 설계, 이해 관계자 참여를 포함한 다양성, 차별 금지 및 공정성;

- Societal and environmental wellbeing, including sustainability and environmental friendliness, social impact, society and democracy;

  지속 가능성 및 환경 친 화성, 사회적 영향, 사회 및 민주주의를 포함한 사회적 및 환경적 복지;

- Accountability, including auditability, minimisation and reporting of negative impact, trade-offs and redress.

  감사, 최소화 및 부정적인 영향의 보고, 트레이드 오프 및 복구를 포함한 책임

Implementation of these requirements should occur throughout an AI system's entire life cycle as required by specific applications.

이러한 요구 사항의 구현은 특정 응용 프로그램에 필요한 AI 시스템의 전체 생명주기 동안 발생해야 합니다.

A recent comparative analysis of documents on the ethics of AI has noted a global convergence around the values of transparency, non-maleficence, responsibility, and privacy, while dignity, solidarity and responsibility are less often mentioned.55 However, substantial differences exists on how to how to balance competing requirements, i.e., on how to address cases in which some of the values just mentioned are affected, but at the same time economic, administrative, political or military advantages are also obtained.

최근 AI 윤리에 관한 문서를 비교 분석한 결과 투명성, 비 유연성, 책임 및 프라이버시의 가치에 대한 전 세계적 융합이 언급되었지만, 존엄성, 연대 및 책임은 덜 언급되었다. 경쟁 요구 사항의 균형을 맞추는 방법, 즉 방금 언급한 가치 중 일부가 영향을 받는 사례를 다루는 방법에 관한 동시에 경제적, 행정적, 정치적 또는 군사적 이점도 얻습니다.

2.4.2. Legal principles and norms (법적 원칙 및 규범)

Moving from ethics to law, AI may both promote and demote different fundamental rights and social values included in the EU

Charter and in national constitutions. AI indeed can magnify both the positive and the negative impacts of ICTs on human rights and social values.56 The rights to privacy and data protection (Articles a 7 and 8 of the Charter) are at the forefront, but other rights are also at stake: dignity (article 1), right to liberty and security, freedom of thought, conscience and religion (Article 10), freedom of expression and information (Article 11), freedom of assembly and association (Article 12), freedom of arts and science (Article 13), right to education (article 14), freedom to choose an occupation and right to engage in work (Article 15), right to equality before the law (Article 20), right to non-discrimination (article 21), equality between men and women (Article 23), rights of the child (Article 24), right to fair and just working conditions (Article 31), right to health care (article 35), right to access to services of general economic interest (Article 36),

윤리에서 법으로 전환하면서 AI는 EU 헌장과 국가 헌법에 포함된 다양한 기본 권리와 사회적 가치를 장려하고 강등할 수 있습니다. AI는 실제로 인권과 사회적 가치에 대한 ICT의 긍정적 영향과 부정적인 영향을 모두 확대할 수 있다.56 프라이버시 및 데이터 보호에 대한 권리 (헌장의 7 조와 8조)는 최전선에 있지만 다른 권리도 위태로워진다 : 존엄성 (1조), 자유와 안전에 대한 권리, 사고의 자유, 양심과 종교 (10조), 표현과 정보의 자유 (11조), 집회

와 결사의 자유 (제12조), 예술과 과학의 자유 (제13조), 교육의 권리 (제14조), 직업을 선택할 수 있는 자유와 직업에 참여할 권리 (제15조), 법 앞의 평등권 (제20조), 차별 금지의 권리 (제21조), 남녀 평등 (제23조), 아동의 권리 (제24조), 공정하고 정당한 근로 조건 (제31조), 건강 관리 권리 (제35조), 일반적인 경제적 이익을 얻는 서비스에 대한 권리 (제35조) 제36조)

55 Jobin et al (2019).
56 For a review of the impacts of ICTs on rights and values, see Sartor (2017), De Hert and Gutwirth (2009).

consumer protection (Article 38), right to good administration (Article 41), right to an effective remedy and to a fair trial (Article 47). Besides individual right also social values are at stake, such as democracy, peace, welfare, competition, social dialogue efficiency, advancement in science, art and culture, cooperation, civility, and security.

소비자 보호 (제38조), 올바른 행정권 (제41조), 효과적인 구제 및 공정한 재판 (제47조)에 대한 권리. 개인의 권리 외에도 민주주의, 평화, 복지, 경쟁, 사회적 대화 효율성, 과학, 예술 및 문화의 발전,

협력, 시민 및 보안과 같은 사회적 가치가 위태로워지고 있습니다.

Given the huge breath of its impacts on citizens' individual and social lives, AI falls under the scope of different sectorial legal regimes. These regimes include especially, though not exclusively, data protection law, consumer protection law, and competition law. As has been observed by the European Data Protection Supervisor (EDPS) in Opinion 8/18 on the legislative package 'A New Deal for Consumers,' there is synergy between the three regimes. Consumer and data protection law share the common goals of correcting imbalances of informational and market power, and, along with competition law, they contribute to ensuring that people are treated fairly. Other domains of the law are also involved in AI: labour law relative to the new forms of control over worker enabled by AI; administrative law relative to the opportunities and risk in using AI to support administrative decision-making; civil liability law relative to harm caused by AI driven systems and machines; contract law relative to the use of AI in preparing, executing and performing agreements; laws on political propaganda and elections relatively to the use of AI in political campaigns; military law on the use of AI in armed conflicts; etc.

AI가 시민의 개인 및 사회 생활에 미치는 영향이 엄청나게 높아

짐에 따라 AI는 다양한 부문 별 법률 제도의 범위에 속합니다. 이러한 제도에는 특히 데이터 보호법, 소비자 보호법 및 경쟁법이 포함됩니다. 입법 패키지 '소비자에 대한 새로운 거래'에 대한 의견 8/18의 유럽 데이터 보호 감독자 (EDPS)에 의해 관찰된 바와 같이, 세 정권 사이에는 시너지 효과가 있습니다. 소비자 및 데이터 보호법은 정보 및 시장 권력의 불균형을 수정하는 공통의 목표를 공유하며 경쟁법과 함께 사람들이 공정하게 대우받는 데 기여합니다. 법의 다른 영역들도 AI와 관련이 있다 : AI에 의해 가능해진 노동자에 대한 새로운 형태의 통제에 관한 노동법; AI를 사용하여 행정 의사 결정을 지원할 기회와 위험에 관한 행정법; AI 기반 시스템 및 기계로 인한 피해에 대한 민사 책임법; 계약 준비, 실행 및 수행에 AI 사용에 관한 계약법; 정치 캠페인에 AI를 사용하는 것에 대한 정치적 선전 및 선거에 관한 법률; 무력 충돌에서 AI 사용에 관한 군사 법; 기타

2.4.3. Some interests at stake (위기에 처한 관심사)

The significance that AI bears to different areas of the law has to do with the nature of the interest that are affected by the deployment of AI technologies. Here are some of the interests more directly and specifically involved.

AI가 법의 다른 영역에 적용되는 중요성은 AI 기술의 배치에 의해 영향을 받는 관심의 본질과 관련이 있습니다. 보다 직접적이고 구체적으로 관련된 관심사 중 일부는 다음과 같습니다.

First, there is the interest in data protection and privacy, namely, the interest in a lawful and proportionate processing of personal data subject to oversight. This is hardly compatible with an online environment where every action is tracked, and the resulting data is used to extract further information about the individuals concerned, beyond their control, and to process this information in ways that may run counter to their interests.

첫째, 데이터 보호 및 개인 정보 보호, 즉 감독 대상 개인 데이터의 합법적이고 비례적인 처리에 대한 관심이 있습니다. 이는 모든 작업을 추적하는 온라인 환경과 거의 호환되지 않으며 결과 데이터는 통제할 수 없는 관련 개인에 대한 추가 정보를 추출하고 이 정보를 자신의 이익에 반하는 방식으로 처리하는 데 사용됩니다.

The processing of personal data through AI systems may also affect citizens' interest in fair algorithmic treatment, namely, their interest in not being subject to unjustified prejudice resulting from automated processing.

AI 시스템을 통한 개인 데이터 처리는 공정한 알고리즘 처리에 대한 시민의 관심, 즉 자동화된 처리로 인한 부당한 편견을 받지 않는 것에 대한 관심에 영향을 줄 수 있습니다.

The possibility of algorithmic unfairness, as well as the need to keep the processing of personal data under control and to understand (and possibly challenge) the reasons for determinations that affect individuals, raises concern from an algorithmic transparency/explicability standpoint. Citizens want to know how and why a certain algorithmic response has been given or a decision made, so as 'to understand and hold to account the decision-making processes of AI.' 57

알고리즘 불공정의 가능성과 개인 데이터 처리를 통제하고 개인에게 영향을 미치는 결정의 이유를 이해 (및 도전)해야 할 필요성은 알고리즘 투명성/확장 성 관점에서 우려를 불러 일으킵니다. 시민들은 'AI의 의사 결정 과정을 이해하고 고려하기 위해'알고리즘 응답이 어떻게 이루어 졌는지, 왜 결정되었는지를 알고 싶어합니다. 57

Individual autonomy is affected when citizens interact with black boxes,17 whose functioning is not accessible to them, and whose

decisions remain unexplained and thus unchallengeable.58

시민들이 블랙 박스와 상호 작용할 때 개인의 자율성은 영향을 받는다.17)

As observed above, since AI systems have access to a huge amount of information about individuals and about people similar to them, they can effortlessly use this information to elicit desired behaviour for purposes that citizens may not share, possibly in violation of fiduciary expectations they have toward the organisation that is deploying the AI system in question.59 Thus, individuals have an interest in not being misled or manipulated by AI systems, but they also have an interest in being able to trust such systems, knowing that the controllers of those systems will not profit from the people's exposure (possibly resulting from personal data).

위에서 살펴본 바와 같이 AI 시스템은 개인 및 이와 유사한 사람들에 대한 방대한 양의 정보에 액세스할 수 있기 때문에 이 정보를 손쉽게 사용하여 시민이 공유할 수 없는 목적으로, 원하는 수의 기대를 위반하여 원하는 행동을 유도할 수 있습니다. 따라서 개인은 문제가 있는 AI 시스템을 배포하는 조직으로 향합니다.59 따라서 개인은 AI 시스템에 의해 오도되거나 조작되지 않는 데 관심이 있지만 해당 시스템의 컨트롤러를 알고 이러한 시스템을

신뢰할 수 있다는 데 관심이 있습니다. 사람들의 노출로 인해 이익을 얻지 못할 것입니다 (개인 데이터로 인해 발생할 수 있음).

57 Floridi et al (2018).
58 Pasquale (2015).
59 On fiduciary obligations related to the use of AI, see Balkin (2017).

Reasonable trust is needed so that individuals do to waste their limited and costly cognitive capacities in trying to fend off AI systems' attempts to mislead and manipulate them.

Finally, citizens have an indirect interest in fair algorithmic competition, i.e., in not being subject to market-power abuses resulting from exclusive control over masses of data and technologies. This is of direct concern to competitors, but the lack of competition may negatively affect consumers, too, by depriving them of valuable options and restricting their sphere of action. Moreover, the lack of competition enables the leading companies to obtain huge financial resources, which they can use to further increase their market power (e.g., by preventively buying potential competitors), or to promote their interests through influence.ng

public opinion and politics.

AI 시스템의 오도 및 조작 시도를 막기 위해 개인이 제한적이고 비용이 많이 드는 인지능력을 낭비하기 위해서는 합리적인 신뢰가 필요합니다.

마지막으로, 시민들은 공정한 알고리즘 경쟁, 즉 대량의 데이터 및 기술에 대한 독점적 통제로 인한 시장 권력 남용의 대상이 되지 않기 때문에 간접적인 관심을 가지고 있습니다. 이는 경쟁사에게 직접적인 관심사이지만 경쟁이 결여된 경우에도 귀중한 옵션을 빼앗아 행동 영역을 제한함으로써 소비자에게 부정적인 영향을 줄 수 있습니다. 또한, 경쟁이 결여되어 있는 선도 기업들은 막대한 재무 자원을 확보할 수 있으며, 이를 통해 잠재적인 경쟁업체를 사전에 구매함으로써 시장 파워를 더욱 높이거나 여론과 정치에 대한 영향을 통해 관심을 증진시킬 수 있습니다.

2.4.4. AI technologies for social and legal empowerment

(사회적 및 법적 권한 부여를 위한 AI 기술)

To ensure an effective protection of citizens' rights and to direct AI

towards individual and social goods, regulatory initiatives are an essential element. However, regulatory instruments and their implementation by public bodies may be insufficient. Indeed, AI and big data are employed in domains already characterised by a vast power imbalance, which they may contribute to accentuate. In fact, these technologies create new knowledge (analytical and forecasting abilities) and powers (control and influence capacities) and make them available to those who govern these technologies.

To ensure an adequate protection of citizens, beside regulation and public enforcement, also the countervailing power of civil society 60 is needed to detect abuses, inform the public, activate enforcement, etc. In the AI era, an effective countervailing power needs also to be supported by AI: only if citizens and their organisations are able to use AI to their advantage, can they resist, and respond to, AI-powered companies and governments.61 Moreover, active citizenship is an important value in itself, that needs to be preserved and advanced at a time in which we tend to delegate to technology (and in particular to AI) a vast amount of relevant decisions.

시민의 권리를 효과적으로 보호하고 AI를 개인 및 사회 용품으로 향하게 하려면 규제 이니셔티브가 필수 요소입니다. 그러나 규제

기관과 공공 기관의 구현은 충분하지 않을 수 있습니다. 실제로 AI와 빅 데이터는 이미 막대한 전력 불균형이 특징 인 영역에서 사용되며, 이는 강조에 기여할 수 있습니다. 실제로, 이 기술들은 새로운 지식 (분석 및 예측 능력)과 능력 (통제 능력에 영향을 미침)을 만들어 이 기술을 관리하는 사람들이 이용할 수 있게 합니다.

규제 및 공공 집행 외에도 시민의 적절한 보호를 보장하기 위해서는 남용을 감지하고 대중에게 알리고 집행을 활성화하는 등 시민 사회의 상충되는 힘이 필요합니다. AI 시대에는 효과적인 상충되는 힘도 필요합니다. AI의 지원 : 시민과 조직이 AI를 자신의 이점으로 활용할 수 있고 AI 기반 회사와 정부에 저항하고 대응할 수 있는 경우에만 가능합니다.61 또한 적극적인 시민권은 그 자체로 중요한 가치입니다. 우리가 기술 (특히 AI)에 방대한 양의 관련 결정을 위임하는 경향이 있을 때 보존되고 발전했습니다.

A few examples of citizen-empowering technologies are already with us, as in the case of ad-blocking systems as well as more traditional anti-spam software and anti-phishing techniques. Yet, there is a need to move a step forward. Services could be deployed with the goal of analysing and summarising massive amounts of

product reviews or comparing prices across a multitude of platforms. One example in this direction is offered by CLAUDETTE:62 an online system for the automatic detection of potentially unfair clauses in online contracts and in privacy policies.63 Considerable effort has also been devoted to the development of data mining techniques for detecting discrimination with the aim to build supporting tools that could identify prejudice and unfair treatments in decisions that regard consumers.64

기존의 스팸 방지 소프트웨어 및 피싱 방지 기술 뿐만 아니라 광고 차단 시스템의 경우와 같이 시민권 강화 기술의 몇 가지 예가 이미 우리와 함께 있습니다. 그러나 한 걸음 더 나아갈 필요가 있습니다. 대량의 제품 리뷰를 분석 및 요약하거나 여러 플랫폼에서 가격을 비교할 목적으로 서비스를 배포할 수 있습니다. 온라인 계약과 개인 정보 보호 정책에서 잠재적으로 불공정 조항을 자동으로 탐지하기위한 온라인 시스템 인 CLAUDETTE : 62는 이러한 방향의 한 가지 예를 제시합니다.63 목표를 통한 차별 탐지를 위한 데이터 마이닝 기술 개발에도 상당한 노력이 기울여졌습니다. 소비자를 고려한 의사 결정에서 편견과 불공정한 대우를 식별할 수 있는 지원 도구 구축 64

The growing interest in privacy and data protection has resulted in several proposals for automatically extracting, categorising and summarising information from privacy documents, and assisting users in processing and understanding their contents. Multiple AI methods to supportdata protection could be merged into integrated PDA-CDA (Privacy digital assistants/consumer digital assistants), meant to prevent excessive/unwanted/unlawful collection of personal data and well as to protect users from manipulation and fraud, provide them with awareness of fake and untrustworthy information, and facilitate their escape from 'filter bubbles' (the unwanted filtering/pushing of information).

개인 정보 보호 및 데이터 보호에 대한 관심이 높아짐에 따라 개인 정보 문서에서 정보를 자동으로 추출, 분류 및 요약하고 사용자가 내용을 처리하고 이해하는 데 도움이 되는 몇 가지 제안이 있었습니다. 데이터 보호를 지원하는 여러 가지 AI 방법을 통합 PDA-CDA (개인 정보 디지털 보조자/소비자 디지털 보조자)에 통합하여 과도한/원치 않는/불법적인 개인 데이터 수집을 방지하고 사용자를 조작 및 사기로부터 보호하고 사용자에게 제공할 수 있습니다 가짜 및 신뢰할 수 없는 정보에 대한 인식 및 '필터 버블'(원치 않는 정보 필터링/푸시)에서 벗어날 수 있습니다.

60 Galbraith (1983).
61 Lippi et al (2020).
62 https://claudette.eui.eu/
63 Contissa et al (2018), Lippi et al (2019).
64 Ruggeri, Pedreschi, and Turini (2010).

It may be worth considering how the public could support and incentivise the creation and distribution of AI tools to the benefit of data subject and citizens. Such tools would provide new opportunities for research, development, and entrepreneurship. They would contribute to reduce unfair and unlawful market behaviour and favour the development of legal and ethical business models. Finally, citizen-empowering technologies would support the involvement of civil society in monitoring and assessing the behaviour of public and private actors and of the technologies deployed by the latter, encouraging active citizenship, as a complement to the regulatory and law-enforcement activity of public bodies.

대중이 데이터 주제와 시민의 이익을 위해 AI 도구의 생성 및 배포를 지원하고 장려할 수 있는 방법을 고려해 볼 가치가 있습니다. 이러한 도구는 연구, 개발 및 기업가 정신에 새로운 기회를 제공할 것입니다. 이들은 불공정하고 불법적인 시장 행동을 줄이고 법적 및 윤리적 비즈니스 모델 개발을 선호합니다. 마지막으로

시민권 부여 기술은 공공 및 민간 행위자의 행동과 후자에 의해 전개되는 기술의 모니터링 및 평가에 시민 사회의 참여를 지원하여 시민의 규제 및 법 집행 활동에 대한 보완으로서 적극적인 시민권을 장려합니다.

3. AI in the GDPR

In this section the provisions of the GDPR are singularly analysed to determine the extent to which their application is challenged by of AI as well as the extent to which they may influence the development of AI applications.

이 섹션에서는 GDPR의 조항을 단일 분석하여 AI의 응용 프로그램이 AI에 의해 도전을 받는 정도와 AI 응용 프로그램의 개발에 영향을 미칠 수 있는 정도를 결정합니다.

3.1. AI in the conceptual framework of the GDPR

(GDPR의 개념적 틀에서의 AI)

Unlike the 1995 Data Protection Directive, the GDPR contains some terms referring to the Internet (Internet, social networks, website, links, etc.), but it does not contain the term 'artificial intelligence', nor any terms expressing related concepts, such as intelligent systems, autonomous systems, automated reasoning and inference, machine learning or even big data. This reflects the fact that the

GDPR is focussed on the challenges emerging for the Internet – which were not considered in the 1995 Data Protection Directive, but were well present at the time when GDPR was drafted – rather than on new issues pertaining to AI, which only acquired social significance in most recent years. However, as we shall see, many provisions in the GDPR are very relevant to AI.

1995년 데이터 보호 지침과 달리 GDPR에는 인터넷 (인터넷, 소셜 네트워크, 웹 사이트, 링크 등)을 가리키는 용어가 포함되어 있지만 '인공 지능'이라는 용어 나 관련 개념을 나타내는 용어는 포함되지 않습니다. 지능형 시스템, 자율 시스템, 자동화된 추론 및 추론, 기계학습 또는 빅 데이터 등이 있습니다. 이는 GDPR이 1995년 데이터 보호 지침에서 고려되지 않았지만 AI와 관련된 새로운 문제보다는 GDPR 초안 작성 당시에 존재했던 인터넷에 대한 도전에 초점을 맞추고 있다는 사실을 반영합니다. 가장 최근 몇 년 동안 사회적 중요성을 얻었습니다. 그러나 앞으로 살펴 보겠지만 GDPR의 많은 조항은 AI와 매우 관련이 있습니다.

3.1.1. Article 4(1) GDPR: Personal data (identification, identifiability, re-identification)

제4조 (1) GDPR : 개인 정보 (식별, 식별, 재식별)

The concept of personal data plays a key role in the GDPR, characterising the material scope of the regulation. The provision in the GDPR only concern personal data, to the exclusion of information that does not concerns humans (e.g., data on natural phenomena), and also to the exclusion of information that, though concerning humans does not refer to particular individuals (e.g., general medical information on human physiology or pathologies) or has been effectively anonymised so that it has lost its connection to particular individuals. Here is how personal data are defined in Article 4 (1) GDPR:

개인 정보의 개념은 GDPR에서 중요한 역할을 하며 규제의 범위를 규정합니다. GDPR의 조항은 개인 데이터, 인간과 관련이 없는 정보 (예 : 자연 현상에 대한 데이터 제외) 및 인간과 관련이 있지만 특정 개인을 언급하지 않는 정보 (예 : 인간 생리학 또는 병리에 대한 일반적인 의료 정보) 또는 특정 개인과의 연결이 끊어지도록 효과적으로 익명 처리되었습니다. 다음은 개인 정보가 제4조 (1) GDPR에 정의된 방법입니다.

> ***'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference***

to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

'개인 데이터'는 식별되거나 식별 가능한 자연인 ('데이터 주체')과 관련된 모든 정보를 의미합니다. 식별 가능한 자연인은 특히 이름, 식별 번호, 위치 데이터, 온라인 식별자와 같은 식별자 또는 물리적, 생리적, 특정 특성에 대한 하나 이상의 요소를 참조하여 직접 또는 간접적으로 식별될 수 있는 사람입니다. 그 자연인의 유전적, 정신적, 경제적, 문화적 또는 사회적 정체성;

Recital (26) addresses identifiability, namely, the conditions under which a piece of data which is not explicitly linked to a person, still counts as personal data, since the possibility exists to identify the person concerned. Identifiability depends on the availability of 'means reasonably likely to be used' for successful re-identification, which in its turn, depends on the technological and sociotechnical state of the art:

Recital (26)은 식별 가능성, 즉 사람과 명시적으로 연결되지 않은

데이터가 여전히 개인 데이터로 계산되는 조건을 다루고 있는데, 이는 해당 개인을 식별할 가능성이 있기 때문입니다. 식별 가능성은 성공적인 재식별을 위한 '합리적으로 사용될 가능성이 높은 수단'의 가용성에 달려 있으며, 이는 다시 기술 및 사회 공학적 기술 수준에 달려 있습니다.

To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.

자연인을 식별할 수 있는지 여부를 판별하려면 컨트롤러 또는 다른 사람이 자연스럽게 사람을 직접 또는 간접적으로 식별하는 데 사용하는 것과 같이 합리적으로 사용될 가능성이 있는 모든 수단을 고려해야 합니다. 자연인을 식별하기 위해 수단이 합리적으로 사용될 가능성이 있는지 확

인하려면, 비용 및 식별에 필요한 시간과 같은 모든 객관적인 요소를 고려할 때, 가용한 기술을 고려하십시오. 가공 및 기술 개발.

Through pseudonymisation, the data items that identify a person (i.e., the name) are substituted with a pseudonym, but the link between the pseudonym and the identifying data items can be retraced by using separate information (e.g., through a table linking pseudonyms and real names, or through cryptography key to decode the encrypted names). Recital (26) specifies that pseudonymised data still are personal data.

가명화를 통해 개인 (즉, 이름)을 식별하는 데이터 항목은 가명으로 대체되지만, 가명과 식별 데이터 항목 간의 연결은 별도의 정보를 사용하여 (예 : 가명과 실제를 연결하는 테이블을 통해) 되돌릴 수 있습니다. 암호화된 이름을 해독하려면 이름을 사용하거나 암호화 키를 사용하십시오. Recital (26)은 가명 화된 데이터가 여전히 개인 데이터임을 지정합니다.

Personal data which have undergone pseudonymisation, which could be attributed to natural person by the use of additional information should be considered to be

information on an identifiable natural person.

추가 정보를 사용하여 자연인에게 귀속될 수 있는 가명 화를 거친 개인 데이터는 식별 가능한 자연인에 대한 정보로 간주해야 합니다.

The connection between the personal nature of information and technological development is mentioned at Recital (9) of Regulation 2018/1807:

정보의 개인적 특성과 기술 개발 사이의 연결은 규정 2018/1807의 Recital (9)에서 언급됩니다.

If technological developments make it possible to turn anonymised data into personal data, such data are to be treated as personal data, and Regulation (EU) 2016/679 is to apply accordingly.

기술 개발로 인해 익명의 데이터를 개인 데이터로 전환할 수 있는 경우 해당 데이터는 개인 데이터로 취급되며 규정 (EU) 2016/679가 그에 따라 적용됩니다.

The concept of non-personal data is not positively defined in the EU legislation, as it includes whatever data that are not personal data as defined in the GDPR. Regulation 2018/1807,65 at Recital provides the following examples of non-personal data: aggregate and anonymised datasets used for big data analytics, data on precision farming that can help to monitor and optimise the use of pesticides and water, or data on maintenance needs for industrial machines.'

비 개인 데이터의 개념은 GDPR에 정의된 개인 데이터가 아닌 모든 데이터를 포함하므로 EU 법률에서 긍정적으로 정의되지 않습니다. Recital의 규정 2018/1807,65는 다음과 같은 비 개인 데이터의 예를 제공합니다. 빅 데이터 분석에 사용되는 집계 및 익명 데이터 세트, 농약 및 물 사용을 모니터링하고 최적화하는 데 도움이 되는 정밀 농업 데이터 또는 유지 보수 데이터 산업 기계의 필요성. '

In connection with the GDPR definition of personal data, AI raises in particular two key issues: (1) the 're-personalisation' of anonymous data, namely the re-identification of the individuals to which such data are related; (2) and the inference of further personal information from personal data that are already available.

개인 데이터의 GDPR 정의와 관련하여 AI는 특히 두 가지 주요 문제를 제기합니다. (1) 익명 데이터의 '재 개인화', 즉 해당 데이터와 관련된 개인의 재식 명; (2) 이미 사용 가능한 개인 데이터에서 추가 개인 정보를 추론합니다.

Re-identification

The first issue concerns of identifiability. AI, and more generally methods for computational statistics, increases the identifiability of apparently anonymous data, since they enable nonidentified data (including data having been anonymised or pseudonymised) to be connected to the individuals concerned

첫 번째 문제는 식별 가능성에 관한 것입니다. AI 및 보다 일반적으로 계산 통계를 위한 방법은 익명으로 식별된 데이터 (익명 또는 가명 화된 데이터 포함)를 관련 개인과 연결할 수 있기 때문에 익명의 데이터의 식별 가능성을 높입니다.

> ***[N]umerous supposedly anonymous datasets have recently been released and reidentified. In 2016, journalists reidentified politicians in an anonymized browsing history dataset of 3 million German citizens,***

uncovering their medical information and their sexual preferences. A few months before, the Australian Department of Health publicly released de-identified medical records for 10% of the population only for researchers to reidentify them 6 weeks later. Before that, studies had shown that de-identified hospital discharge data could be reidentified using basic demographic attributes and that diagnostic codes, year of birth, gender, and ethnicity could uniquely identify patients in genomic studies data. Finally, researchers were able to uniquely identify individuals in anonymized taxi trajectories in NYC27, bike sharing trips in London, subway data in Riga, and mobile phone and credit card datasets.66

수많은 익명의 데이터 세트가 최근에 릴리스 되고 재식별 되었습니다. 2016년, 언론인들은 3 백만 명의 독일 시민들의 익명의 인터넷 사용 기록 데이터 세트에서 정치인들을 재확인하여 의료 정보와 성적 취향을 밝혔습니다. 몇 개월 전 호주 보건부는 6 주 후에 연구원들이 재식별 할 수 있도록 인구의 10 %에 대한 미확인 의료 기록을 공개했습니다. 그 이전에, 연구는 기본 인구 통계 학적 속성을 사용하여 비 식별된 퇴원 데이터를 재식별 할 수 있고 진단 코드,

생년월일, 성별 및 민족성이 게놈 연구 데이터에서 환자를 고유하게 식별할 수 있음을 보여 주었다. 마지막으로, 연구원들은 NYC27의 익명 택시 궤적, 런던의 자전거 공유 여행, 리가의 지하철 데이터 및 휴대폰 및 신용 카드 데이터 세트에서 개인을 고유하게 식별할 수 있었습니다.66

The re-identification of data subjects is usually based on statistical correlations between non-identified data and personal data concerning the same individuals.

데이터 주체의 재식별은 일반적으로 동일인에 관한 비 식별 데이터와 개인 데이터 간의 통계적 상관 관계를 기반으로 합니다.

65 Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union.
66 Rocher et al (2019).

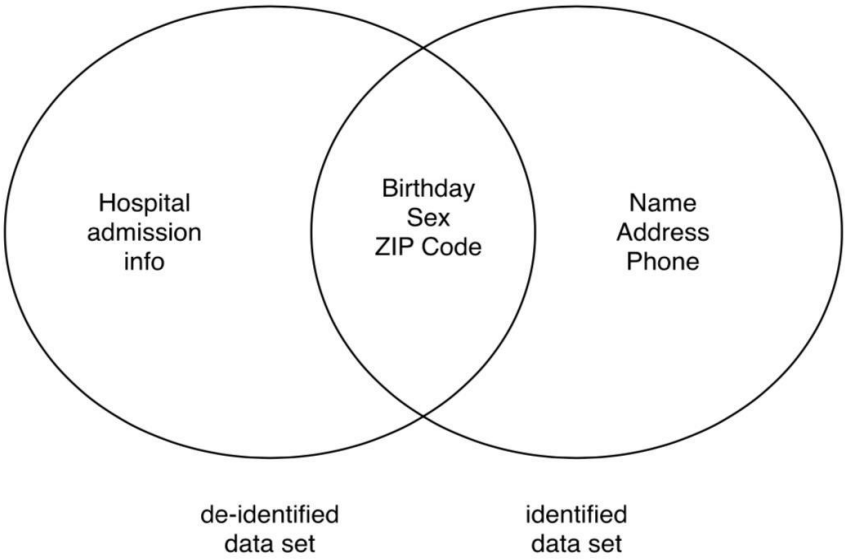

Figure 13 – The connection between identified and de-identified data

Figure 13 illustrates a connection between an identified and a de-identified data set that enabled the re-identification of the health record of the governor of Massachusetts. This result was obtained by searching for de-identified data that matched the Governor's date of birth, ZIP code and gender.67 Another classic example is provided the Netflix price database case, in which anonymised movie ratings could be re-identified by linking them to non-

anonymous ratings in IMDb (Internet Movie Database). In fact, knowing only two non-anonymous reviews by an IMDb user, it was possible to identify the reviews by the same user in the anonymous database. Similarly, it has been shown that an anonymous user of an online service can be re-identified by that service, if the service knows that the user has installed four apps on his or her device, and the service has access to the whole list of apps installed by each user.68

그림 13은 매사추세츠 주지사의 건강 기록을 재식별 할 수 있는 식별된 데이터 세트와 비 식별된 데이터 세트 간의 연결을 보여줍니다. 이 결과는 주지사의 생년월일, 우편 번호 및 성별과 일치하는 비 식별된 데이터를 검색하여 얻은 것입니다.67 또 다른 고전적인 예는 Netflix 가격 데이터베이스 사례입니다. 익명의 영화 등급을 다음과 연결하여 다시 식별할 수 있습니다. IMDb (인터넷 영화 데이터베이스)의 비 익명 등급 실제로 IMDb 사용자가 익명이 아닌 두 개의 리뷰 만 알면 익명 데이터베이스에서 동일한 사용자가 리뷰를 식별할 수 있었습니다. 마찬가지로 서비스에서 사용자가 자신의 기기에 4 개의 앱을 설치했으며 서비스가 전체 목록에 액세스할 수 있음을 알고 있는 경우 온라인 서비스의 익명 사용자를 해당 서비스에서 다시 식별할 수 있음을 보여줍니다 각 사용자가 설치한 앱

Re-identification can be viewed as a specific kind of inference of personal data: through re-identification. A personal identifier is associated to previously non-identified data items, which, as a consequence, become personal data. Note that for an item to be linked to a person, it is not necessary that the data subject be identified with absolute certainty; a degree of probability may be sufficient to enable a differential treatment of the same individual (e.g., the sending of targeted advertising).

재식별은 재식별을 통한 개인 데이터의 특정 유추로 볼 수 있습니다. 개인 식별자는 이전에 식별되지 않은 데이터 항목과 연관되어 결과적으로 개인 데이터가 됩니다. 항목이 사람과 연결되려면 데이터 주제를 절대 확실하게 식별할 필요는 없습니다. 동일한 개인 (예를 들어, 타겟 광고의 전송)의 차등 처리를 가능하게 하기 위해 확률의 정도가 충분할 수 있다.

Thanks to AI and big data the identifiability of the data subjects has vastly increased. The personal nature of a data idem no longer is a feature of that item separately considered. It has rather become a contextual feature. As shown above, an apparently anonymous data item becomes personal in the context of further personal data that enable re-identification. For instance, the identifiability of the Netflix movie reviewers supervened on the availability of their

named reviews on IMDb. As it has been argued, 'in any "reasonable" setting there is a piece of information that is in itself innocent, yet in conjunction with even a modified (noisy) version of the data yields a privacy breach.' 69

AI와 빅 데이터 덕분에 데이터 주체의 식별 가능성이 크게 증가했습니다. 데이터의 개인 특성은 더 이상 개별적으로 고려되는 해당 항목의 기능이 아닙니다. 오히려 상황에 맞는 기능이 되었습니다. 위에서 보여 지듯이, 명백히 익명 인 데이터 아이템은 재식별을 가능하게 하는 추가 개인 데이터의 맥락에서 개인화된다. 예를 들어, Netflix 영화 리뷰어의 식별 가능성은 IMDb에서 명명된 리뷰의 가용성을 감독했습니다. 논쟁의 여지가 있지만, "모든"합리적인 "환경에는 그 자체로 결백한 정보가 있지만 수정된 (잡음) 버전의 데이터 와도 프라이버시 침해가 발생합니다. ' 69

67 Sweeney (2000).
68 Achara et al (2015)
69 Dwork and Naor (2010, 93).

This possibility can be addressed in two ways, neither of which is fail-proof. The first consists in ensuring that data is de-identified in ways that make it more difficult to re-identify the data subject; the second consists in implementing security processes and measures

for the release of data that contribute to this outcome.70

이 가능성은 두 가지 방식으로 해결될 수 있으며, 두 가지 모두 실패 방지 방법이 아닙니다. 첫 번째는 데이터 주체를 다시 식별하기 어려운 방식으로 데이터를 비 식별 화하는 것입니다. 두 번째는 보안 프로세스를 구현하고 이 결과에 기여하는 데이터 릴리스를 위한 조치로 구성됩니다.70

Inferred personal data

As noted above, AI systems may infer new information about data subjects, by applying algorithmic models to their personal data. The key issue, from a data protection perspective, is whether the inferred information should be considered as new personal data, distinct from the data from which it has been inferred. Assume for instance, that an individual's sexual orientation is inferred from his or her facial features or that an individual's personality type is inferred from his or her online activity. Is the inferred sexual orientation or personality type a new item of personal data? Even when the inference only is probabilistic? If the inferred information counts as new personal data, then automated inferences would trigger all the consequences that the processing of personal data

entails according to the GDPR: the need of a legal basis, the conditions for processing sensitive data, the data subject's rights, etc.

위에서 언급한 것처럼 AI 시스템은 알고리즘 모델을 개인 데이터에 적용하여 데이터 주체에 대한 새로운 정보를 유추할 수 있습니다. 데이터 보호 관점에서 중요한 문제는 추론된 정보가 추론된 데이터와는 다른 새로운 개인 데이터로 간주되어야 하는지 여부입니다. 예를 들어, 개인의 성적 취향은 자신의 얼굴 특징에서 추론되거나 개인의 성격 유형이 온라인 활동에서 추론된다고 가정하십시오. 유추된 성적 취향 또는 성격 유형이 새로운 개인 정보 항목입니까? 추론 만이 확률론 일 때에도? 추론된 정보가 새로운 개인 데이터로 계산되면 자동화된 추론은 개인 정보 처리가 GDPR에 따라 수반되는 모든 결과를 촉발합니다 : 법적 근거의 필요성, 민감한 데이터 처리 조건, 데이터 주체의 권리 등.

Some clues on the legal status of automatically inferred information can be obtained by considering the status of information inferred by humans. There is uncertainty about whether assertions concerning individuals, resulting from human inferences and reasoning may be regarded as personal data. This issue has been examined by the ECJ in Joint Cases C-141 and 372/12, where it was denied that the legal analysis, by the

competent officer, on an application for a residence permit could be deemed personal data.71 According to the ECJ and the Advocate General, only the data on which the analysis was based (the input data about the applicant) as well as the final conclusion of the analysis (the holding that the application was to be denied) were to be regarded as personal data. This qualification did not apply to the intermediate steps (the intermediate conclusions in the argument chain) leading to the final conclusion.

인간이 유추한 정보의 상태를 고려하여 자동으로 유추된 정보의 법적 지위에 대한 단서를 얻을 수 있습니다. 인간의 추론과 추론으로 인해 개인에 관한 주장이 개인 데이터로 간주될 수 있는지에 대한 불확실성이 있다. 이 문제는 ECJ가 공동 사례 C-141 및 372/12에서 조사한 결과, 유자격 공무원에 의한 거주 허가 신청에 대한 법적 분석은 개인 데이터로 간주될 수 없다고 거부되었습니다. ECJ와 대변인은 분석의 근거가 된 데이터 (신청자에 대한 입력 데이터)와 분석의 최종 결론 (신청이 거부된 보유) 만 개인 데이터로 간주했습니다. 이 자격은 최종 결론으로 이어지는 중간 단계 (인수 체인의 중간 결론)에는 적용되지 않았습니다.

In the subsequent decision on Case C-434/16,72 concerning a candidate's request to exercise data protection rights relative to an exam script and the examiners' comments, the ECJ apparently

departed from the principle stated in Joint Cases C-141 and 372/12, arguing that the examiner's comments, too, were personal data. However, the Court held that data protection rights, and in particular the right to rectification, should be understood in connection with the purpose of the data at issue. Thus, according to the Court, the right to rectification does not include a right to correct a candidate's answers or the examiner's comments (unless they were incorrectly recorded). In fact,

according to the ECJ, data protection law is not intended to ensure the accuracy of decision-making processes or good administrative practices. Thus, an examinee has the right to access both to the exam data (the exam responses) and the reasoning based on such data (the comments), but he or she does not have a right to correct the examiners' inferences (the reasoning) or the final result.

시험 스크립트 및 심사관의 의견과 관련하여 데이터 보호 권한을 행사하기위한 후보자의 요청과 관련하여 사례 C-434/16,72에 대한 후속 결정에서 ECJ는 분명히 공동 사례 C-141 및 372 /에 명시된 원칙에서 벗어났습니다. 12, 심사관의 의견도 개인 데이터라고 주장했다. 그러나 법원은 문제의 데이터의 목적과 관련하여 데이터 보호 권한, 특히 수정 권한이 이해되어야 한다고 판결했습니다. 따라서, 법원에 따르면, 정정할 권리에는 응시자의 답변 또는

심사관의 의견을 정정할 권리가 포함되지 않습니다 (잘못 기록되지 않은 한). 사실로,

The view that inferred data are personal data was endorsed by the Article 29 WP, being implied in particular by the broad concept of personal data adopted in Opinion 4/2007. 73 This broad concept of personal data is presupposed by the Article 29 WP's statement, that in case of automated inference (profiling) data subjects have the right to access both the input data and the (final or intermediate) conclusions automatically inferred from such data.74

추론된 데이터가 개인 데이터라는 견해는 제29조 WP에 의해 승인되었으며, 특히 의견 4/2007에 채택된 광범위한 개인 데이터 개념에 의해 암시됩니다. 73 개인 정보의 이 넓은 개념은 제29조 WP의 진술에 의해 가정되며, 자동 추론 (프로파일링) 데이터의 경우 주체는 입력 데이터와 그러한 데이터에서 자동으로 추론된 (최종 또는 중간) 결론에 액세스할 권리가 있다. 74

70 Rubinstein and Harzog (2016).

71 Joint cases c-141 and 372/12. See Joined Cases C–141 & 372/12, YS, M and S v. Minister voor Immigratie, Integratie en Asiel, 2014 E.C.R. I-2081, ¶ 48.

72 Case C-434/16, Peter Nowak v. Data Protection Commissioner, 34.

73 Opinion 4/2007

74 Opinion 216/679, adopted on 3 October 2017, revised in 6 February 2018.

3.1.2. Article 4(2) GDPR: Profiling

The definition of profiling, while not using explicitly referring to AI, addresses processing that is today is typically accomplished using AI technologies. This processing consists in using the data concerning person to infer information on further aspects of that person:

프로파일링의 정의는 AI를 명시적으로 참조하지는 않지만 오늘날 처리되는 주소를 처리하는 것은 일반적으로 AI 기술을 사용하여 수행됩니다. 이 처리는 개인 관련 데이터를 사용하여 해당 개인의 추가 측면에 대한 정보를 유추합니다.

> ***'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences,***

interests, reliability, behaviour, location or movements;

'프로파일링 '이란 자연인과 관련된 특정 개인적 측면을 평가하기 위해, 특히 직장에서 자연인의 성과, 경제적 상황, 건강, 개인적 취향, 관심사, 신뢰성, 행동, 위치 또는 움직임;

According to the Article 29 WP,75 profiling aims at classifying persons into categories of groups sharing the features being inferred:

제29조 WP, 75에 따르면 프로파일링은 추론되는 기능을 공유하는 그룹을 사람을 범주로 분류하는 것을 목표로 한다.

'broadly speaking, profiling means gathering information about an individual (or group of individuals) and evaluating their characteristics or behaviour patterns in order to place them into a certain category or group, in particular to analyse and/or make predictions about, for example, their:

ability to perform a task;

interests; or

likely behaviour.'

'폭넓게 말하면, 프로파일링 이란 개인 (또는 개인 그룹)에 대한 정보를 수집하고 특정 범주 또는 그룹에 배치하기 위해 특히 예를 들어 분석 및/또는 예측하기 위해 특성 또는 행동 패턴을 평가하는 것을 의미합니다. 그들의:

작업 수행 능력;

관심;

취향 '

AI and profiling

AI and big data, in combination with the availability of extensive computer resources, have vastly increased the opportunities for profiling. Indeed, machine learning-based approaches, as described in the previous sections, are often meant to provide inferences – classifications, predictions or decisions – when applied to data concerning individuals.

AI와 빅 데이터는 광범위한 컴퓨터 리소스의 가용성과 함께 프로

파일링 기회를 크게 늘렸습니다. 실제로 이전 섹션에서 설명한 기계학습 기반 접근 방식은 종종 개인 관련 데이터에 적용될 때 분류, 예측 또는 결정과 같은 추론을 제공하기위한 것입니다..

Assume that a classifier has trained on a vast set of past examples, which link certain features of individuals (the predictors), to another feature of the same individuals (the target). Through the training, the system has learned an algorithmic model can be applied to new cases: if the model is given predictors-values concerning a new individual, it infers a corresponding target value for that individual, i.e., a new data item concerning him or her.

분류자가 개인의 특정 기능 (예측 자)과 동일한 개인의 다른 기능 (목표)을 연결하는 방대한 과거 예제에 대해 학습했다고 가정합니다. 훈련을 통해 시스템은 알고리즘 모델이 새로운 사례에 적용될 수 있음을 배웠습니다. 모델에 새로운 개인에 대한 예측 자 값이 제공되면 해당 개인에 대한 해당 목표 값, 즉 그에 대한 새로운 데이터 항목 또는 그녀.

For instance, the likelihood of heart disease of applicants for insurance may be predicted on the basis of their health records, but also on the basis of their habits (on eating, physical exercise,

etc.) or social conditions; the creditworthiness of loan applicants may be predicted on the basis of their financial history but also on the basis of their online activity and social condition; the likelihood that convicted persons may reoffend may be predicted on the basis their criminal history, but also possibly their character (as identified by personality test) and personal background. These predictions may trigger automated determinations concerning, respectively, the price of a health insurance, the granting of a loan, or the release on parole.

예를 들어, 보험 신청자들의 심장병 가능성은 그들의 건강 기록뿐만 아니라 그들의 습관 (식사, 운동 등) 또는 사회적 조건에 기초하여 예측될 수 있습니다; 대출 신청자의 신용도는 재무 기록뿐만 아니라 온라인 활동 및 사회적 조건을 기반으로 예측될 수 있습니다. 유죄 판결을 받은 사람이 다시 불쾌감을 줄 가능성은 자신의 범죄 기록뿐만 아니라 성격 (성격 테스트에 의해 식별된) 및 개인적 배경에 기초하여 예측될 수 있습니다. 이러한 예측은 각각 건강 보험 가격, 대출 승인 또는 가석방 해제에 관한 자동 결정을 유발할 수 있습니다.

A learned correlation may also concern a person's propensity to respond in certain ways to certain stimuli. This would enable the transition from prediction to behaviour modification (both

legitimate influence and illegal or unethical manipulation). Assume, for instance that a system learns a correlation between certain features and activities (purchases, likes, etc.) of a person and his or her profile as a specific type of consumer, and that the system has also learned (or has been told) that this kind of consumer is interested in certain products and is likely to respond to certain kinds of ads. Consequently, a person who has these features and has engaged in such activities may be sent the messages that are most likely to trigger the desired purchasing behaviour. The same model can be extended to politics, with regard to messages that may trigger desired voting behaviour.

학습된 상관 관계는 특정 자극에 특정 방식으로 반응하는 사람의 성향과 관련될 수도 있습니다. 이를 통해 예측에서 행동 수정 (적법한 영향 및 불법적 또는 비 윤리적 조작)으로 전환할 수 있습니다. 예를 들어, 시스템이 특정 유형의 소비자로서 특정 기능과 활동 (구매, 좋아요 등)과 개인의 프로파일 사이의 상관 관계를 학습하고 시스템이 또한 학습했거나 들었다고 가정합니다.) 이러한 종류의 소비자는 특정 제품에 관심이 있으며 특정 종류의 광고에 반응할 가능성이 높습니다. 결과적으로, 이러한 기능을 가지고 있고 그러한 활동에 참여한 사람에게는 원하는 구매 행동을 유발할 가능성이 가장 높은 메시지가 전송될 수 있습니다. 원하는 투표 동작을 유발할 수 있는 메시지와 관련하여 동일한 모델을

정치로 확장할 수 있습니다.

75 Opinion 216/679, adopted on 3 October 2017, revised in 6 February 2018.

Inferences as personal data

As noted above, the data inferred through profiling should be considered personal data. In this connection, we need to distinguish the general correlations that are captured by the learned algorithmic model, and the results of applying that model to the description of a particular individual. Consider for instance a machine learning system that has learned a model (e.g., a neural network or a decision tree) from a training set consisting of previous loan applications and outcomes.

위에서 언급했듯이 프로파일링을 통해 추론된 데이터는 개인 데이터로 간주되어야 합니다. 이와 관련하여, 우리는 학습된 알고리즘 모델에 의해 포착되는 일반적인 상관 관계와 해당 모델을 특정 개인의 설명에적용한 결과를 구별해야 합니다. 예를 들어 이전 대출 신청 및 결과로 구성된 훈련 세트에서 모델 (예 : 신경망 또는 의사 결정 트리)을 학습한 기계학습 시스템을 고려하십시오.

In this example, the system's training set consists of personal data: e.g., for each borrower, his name, the data collected on him or her – age, economic condition, education, job, etc. – and the information on whether he or she defaulted on the loan. The learned algorithmic model no longer contains personal data, since it links any possible combinations of possible input values (predictors) to a corresponding likelihood of default (target). The correlations embedded in the algorithmic model are not personal data, since they apply to all individuals sharing similar characteristics. We can possibly view them as group data, concerning the set of such individuals (e.g., those who are assigned a higher likelihood of default, since they have a low revenue, live in a poor neighbourhood, etc.).

이 예에서 시스템의 교육 세트는 개인 데이터 (예 : 각 차용자, 이름, 수집된 데이터 (연령, 경제 상태, 교육, 직업 등) 및 개인 정보)로 구성됩니다. 대출에 불이행. 학습된 알고리즘 모델은 가능한 입력 값 (예측 자)의 가능한 조합을 해당 기본값 (목표)에 연결하기 때문에 더 이상 개인 데이터를 포함하지 않습니다. 알고리즘 모델에 포함된 상관 관계는 유사한 특성을 공유하는 모든 개인에게 적용되므로 개인 데이터가 아닙니다. 그러한 개인들 (예를 들어, 수입이 적고 이웃이 열악하기 때문에 채무 불이행 가능성이 높은 사람들)에 관한 그룹 데이터로 볼 수 있습니다.

Assume that the algorithmic model is then applied to the input data consisting in the description of a new applicant, in order to determine that applicant's risk of default. In this case both the description of the applicant and the default risk attributed to him or her by the model represent personal data, the first being collected data, and the second inferred data.

그런 다음 신청자가 불이행의 위험을 판단하기 위해 알고리즘 모델이 새로운 신청자의 설명으로 구성된 입력 데이터에 적용된다고 가정하십시오. 이 경우, 신청자에 대한 설명과 모델에 의해 귀속된 기본 위험은 모두 개인 데이터, 첫 번째 수집 데이터 및 두 번째 유추 데이터를 나타냅니다.

Rights over inferences

Since inferred data concerning individuals also are personal data under the GDPR – at least when they are used to derive conclusions that are or may be acted upon – data protection rights should in principle also apply, though concurrent remedies and interests have to be taken into account. As noted above, according to the Article 29 Working Party, in the case of automated inferences (profiling) data subjects have a right to access both the personal data used

as input for the inference, and the personal data obtained as (final or intermediate) inferred output. On the contrary, the right to rectification only applies to a limited extend. When the data are processed by a public authority, it should be considered whether review procedures already exist which provide for access and control. In the case of processing by private controllers, the right to rectify the data should be balanced with the respect for autonomy of private assessments and decisions.76

개인에 관한 유추된 데이터는 GDPR에 따른 개인 데이터이기 때문에 (적어도 그들이 결론을 내릴 때 또는 그에 따라 결정될 수 있는 경우) 원칙적으로 데이터 보호 권리가 적용되어야 하지만, 동시적인 구제 조치와 관심사를 고려해야 합니다.. 위에서 언급한 바와 같이, 제29조 작업반에 따르면, 자동화된 추론 (프로파일링) 데이터의 경우, 주체는 추론을 위한 입력으로 사용된 개인 데이터와 (최종 또는 중간)으로 얻은 개인 데이터 모두에 액세스할 권리가 있습니다. 유추된 출력. 반대로, 정류 권리는 제한된 범위에만 적용됩니다. 공공 기관이 데이터를 처리할 때는 액세스 및 제어를 제공하는 검토 절차가 이미 존재하는지 여부를 고려해야 합니다. 개인 컨트롤러에 의한 처리의 경우, 데이터를 수정할 권리는 개인 평가 및 결정의 자율성을 존중해야 합니다.76

According to the Article 29 Working Party data subjects have a

right to rectification of inferred information not only when the inferred information is 'verifiable' (its correctness can be objectively determined), but also when it is the outcome of unverifiable or probabilistic inferences (e.g., the likelihood of developing heart disease in the future). In the latter case, rectification may be needed not only when the statistical inference was mistaken, but also when the data subject provides specific additional data that support a different, more specific, statistical conclusion. This is linked to the fact that statistical inferences concerning a class may not apply to subclasses of it: it may be the case that students from university A usually have lower skills that students from university B, but this does not apply to the A students having top marks. Accordingly, a top student from university A should have the right to contest the inference that put him or her at a disadvantage relative to an average student from B.

제 29조 작업반 자료에 따르면 주체는 유추된 정보가 '확인 가능' 할 때 (정확성이 객관적으로 결정될 수 있을 때)뿐만 아니라 그것이 검증 불가능하거나 확률론적 추론 (예 : , 장래에 심장병이 발생할 가능성). 후자의 경우, 통계적 추론이 잘못되었을때 뿐만 아니라 데이터 주체가 다른 보다 구체적인 통계적 결론을 뒷받침하는 특정 추가 데이터를 제공할 때도 정류가 필요할 수 있다. 이것은 클래스에 관한 통계적 추론이 클래스의 서브 클래스에 적용되

지 않을 수 있다는 사실과 관련이 있습니다. A 대학의 학생들은 일반적으로 B 대학의 학생들보다 낮은 기술을 가지고 있을 수 있지만, A 학생들에게는 적용되지 않습니다. 최고 인증. 따라서 대학 A의 최고 학생은 B의 평균 학생과 비교하여 자신에게 불리한 추론에 이의를 제기할 권리가 있어야 합니다.

76 Wachter and Mittelstadt (2019).

Legal scholars have argued that data subjects should be granted a general right to 'reasonable inference' namely, the right that any assessment of decision affecting them is obtained through automated inferences that are reasonable, respecting both ethical and epistemic standards.

법률 학자들은 데이터 주체에게 '합리적인 추론'에 대한 일반적인 권리, 즉 그에 영향을 미치는 결정에 대한 평가가 윤리적 및 비판적 표준을 존중하는 합리적인 추론을 통해 얻을 수 있는 권리를 부여해야 한다고 주장했다.

Accordingly, data subject should be entitled to challenge the inferences (e.g. credit scores) made by an AI system, and not only

the decisions based on such inferences (e.g., the granting of loans). It has been argued that for an inference to be reasonable it should satisfy the following criteria:77

따라서 데이터 주체는 그러한 추론에 근거한 결정 (예 : 대출 부여)뿐만 아니라 AI 시스템에 의해 만들어진 추론 (예 : 신용 점수)에 이의를 제기할 자격이 있어야 합니다. 추론이 합리적이려면 다음 기준을 만족해야 한다고 주장했다.77

(a) Acceptability: the input data (the predictors) for the inference should be normatively acceptable as a basis for inferences concerning individuals (e.g., to the exclusion of prohibited features, such as sexual orientation);

수용성 : 추론에 대한 입력 데이터 (예측 자)는 개인에 관한 추론의 근거로서 규범적으로 수용 가능해야 한다 (예 : 성적 취향과 같은 금지된 특징의 배제).

(b) Relevance: the inferred information (the target) should be relevant to the purpose of the decision and normatively acceptable in that connection (e.g., ethnicity should not be inferred for the purpose of giving a loan).

관련성 : 유추된 정보 (목표)는 결정의 목적과 관련이 있

어야 하고 그와 관련하여 규범적으로 수용 가능해야 합니다 (예 : 대출을 목적으로 민족성을 유추해서는 안 됨).

(c) Reliability: both input data, including the training set, and the methods to process them should be accurate and statistically reliable (see Section 2.3.3).

신뢰성 : 훈련 세트를 포함한 입력 데이터와 이를 처리하는 방법은 정확하고 통계적으로 신뢰할 수 있어야한다 (섹션 2.3.3 참조).

Controllers, conversely, should be prohibited to base their assessment or decisions on unreasonable inferences, and they should also have the obligation to demonstrate the reasonableness of their inferences.

반대로, 관리자는 자신의 평가 나 결정을 불합리한 추론에 근거하여 금지해야 하며 추론의 합리성을 입증할 의무도 있어야 합니다.

The idea the unreasonable automated inference should be prohibited only applies to inferences meant to lead to assessments and decisions affecting the data subject. They should not apply to inquiries that are motivated by merely cognitive purposes, such as

those pertaining to scientific research.

불합리한 자동 추론이 금지되어야 한다는 생각은 데이터 주제에 영향을 미치는 평가 및 결정으로 이어지는 추론에만 적용됩니다. 과학적 연구와 같은 인지적 목적에 의해 동기가 부여된 문의에는 적용되지 않아야 합니다.

3.1.3. Article 4(11) GDPR: Consent

Consent according to Article 4(11) GDPR should be freely given, specific, informed and unambiguous, and be expressed through a clear affirmative action:

제 4조 (11)에 따라 동의 GDPR은 자유롭게 제공되고, 구체적이며, 정보가 있고 모호하지 않아야 하며, 명확한 긍정적 조치를 통해 표현되어야 합니다.

> ***'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to***

the processing of personal data relating to him or her;

데이터 주체의 '동의'는 진술 또는 명확한 긍정 조치에 의해 데이터 주체의 희망에 대한 자유롭고, 구체적이며, 정보가 있고 모호하지 않은 표시를 의미하며, 그 또는 그녀와 관련된 개인 정보 처리에 대한 동의를 나타냅니다.;

This definition is complemented by Recital (32) which specifies that consent should be granular, i.e., it should be given for all the purposes of the processing.

이 정의는 Recital (32)에 의해 보완되는데, 이는 동의가 세분화되어야 한다는 것, 즉 처리의 모든 목적을 위해 제공되어야 함을 지정합니다.

Consent should cover all processing activities carried out for the same purpose or purposes. When the processing has multiple purposes, consent should be given for all of them.

동의는 동일한 목적 또는 목적으로 수행된 모든 처리 활동을 포함해야 합니다. 처리에 여러 목적이 있는 경우 모든 목적에 동의해야 합니다.

Consent plays a key role in the traditional understanding of data protection, based indeed on the 'notice and consent' model, according to which data protection is aimed at protecting a right to 'informational self-determination.' This right is indeed exercised by consenting or refusing to content to the processing of one's data, after having been given adequate notice. Against this approach two main criticism have been raised.78

동의는 실제로 '통지 및 동의'모델을 기반으로 데이터 보호에 대한 전통적인 이해에서 중요한 역할을 하며, 정보 보호는 '정보적 자기 결정'에 대한 권리를 보호하는 데 목적이 있습니다. 이 권리는 실제로적절한 통지를 받은 후 데이터 처리에 대한 내용에 동의하거나 거부함으로써 행사됩니다. 이 접근법에 반대하여 두 가지 주요 비판이 제기되었습니다.78

The first criticism it that consent is most often meaningless: usually is not based on real knowledge of the processing at stake, nor on a real opportunity to choose. On the one hand, today's processing of personal data is so complex that most data subjects to do not have the skills to understand them and anticipate the involved risks.

동의한다는 첫 번째 비판은 가장 의미가 없는 경우가 많다. 보통

위기에 처한 처리에 대한 실제 지식이나 선택의 여지가 없다는 사실에 근거하지 않는다. 한편으로, 오늘날의 개인 데이터 처리는 너무 복잡하여 대부분의 데이터 주체는 이를 이해하고 관련 위험을 예측할 수 있는 기술이 없습니다.

77 Wachter and Mittelstadt (2019).
78 See Cate et all (2014).

Moreover, even if data subjects possessed such skills, still they would not have the time and energy to go through the details of each privacy policy. On the other hand, a refusal to consent may imply the impossibility to use (or limitation in the use of) services that are important or even necessary to the data subjects.

더욱이, 데이터 주체가 그러한 기술을 보유하고 있더라도 여전히 각 개인 정보 보호 정책의 세부 사항을 살펴볼 시간과 에너지가 없을 것입니다. 반면에 동의 거부는 데이터 주체에게 중요하거나 필요한 서비스를 사용할 수 없거나 사용이 제한될 수 있음을 의미할 수 있습니다.

The second criticism is that consent, when targeted on specific purposes, does not include (and therefore precludes, when

considered a necessary basis of the processing) future, often unknown, uses of the data, even when such uses are socially beneficial. Thus, the requirement of consent can 'interfere with future benefits and hinder valuable new discoveries', as exemplified in 'myriad examples', including 'examining health records and lab results for medical research, analysing billions of Internet search records to map flu outbreaks and identify dangerous drug interactions, searching financial records to detect and prevent money laundering, and tracking vehicles and pedestrians to aid in infrastructure planning.'79

두 번째 비판은 동의가 특정 목적을 대상으로 할 때 그러한 용도가 사회적으로 유익한 경우에도 데이터의 미래, 종종 알려지지 않은 미래의 사용을 포함하지 않기 때문이다. 따라서 동의 요구 사항은 '수많은 인터넷 검색 기록을 분석하여 독감 발생 및 위험한 약물 상호 작용을 식별하고, 자금 세탁을 감지하고 방지하기 위해 재무 기록을 검색하며, 인프라 계획을 지원하기 위해 차량과 보행자를 추적합니다.'79

These criticisms of consent have been countered by observing that it is possible to implement the principles of consent and purpose limitation in ways that are both meaningful to the data subject and consistent with allowing for future beneficial uses of the data.80

이러한 동의에 대한 비판은 데이터 주체에게 의미가 있고 미래의 유익한 데이터 사용을 허용하는 방식으로 동의 및 목적 제한 원칙을 구현할 수 있음을 관찰함으로써 해결되었습니다.80

Firstly, it has been argued that notices should focus on most important issue, and that they should be user-friendly and direct. In particular, simple and clear information should be given on how to opt-in or opt-out relative to critical processing, such as those involving the tracking of users or the transmission of data to third parties. An interesting example is provided by the new California Data Privacy Act, which requires companies to include in their website a link with the words 'do not sell my data' (or a corresponding logo-button) to enable users to exclude transmission of their data to third parties. Further opt-out or opt-in buttons could be presented to all users, to provide ways to express their preferences relatively to tracking, profiling, etc.

첫째, 통지는 가장 중요한 문제에 중점을 두어야 하며, 사용자에게 친숙하고 직접적이어야 한다고 주장했습니다. 특히, 사용자 추적 또는 제3 자에게 데이터 전송과 같은 중요한 처리와 관련하여 옵트인 또는 옵트아웃하는 방법에 대한 간단하고 명확한 정보를 제공해야 합니다. 흥미로운 캘리포니아 데이터 개인 정보 보호법(California Data Privacy Act)에 의해 흥미로운 예가 제시되는데, 이

는 회사가 웹 사이트에 '내 데이터를 판매하지 않음'(또는 해당 로고 버튼)이라는 단어의 링크를 포함시켜 사용자가 데이터 전송을 배제할 수 있도록 합니다. 제3 자에게. 추적, 프로파일링 등에 상대적으로 선호도를 표현하는 방법을 제공하기 위해 추가 옵트 아웃 또는 옵트 인 버튼을 모든 사용자에게 제공할 수 있습니다.

Secondly, the GDPR allows that the data that were collected for certain purposes are processed for further purposes, as long as the latter purposes are compatible with the original ones (see Section 3.3.4).

둘째, GDPR은 특정 목적을 위해 수집된 데이터가 후자의 목적이 원래의 목적과 호환되는 한 추가 목적을 위해 처리되도록 허용합니다 (섹션 3.3.4 참조).

In conclusion, it seems that, as we shall see in the following, the concepts of consent and purpose limitation can be interpreted in ways that are consistent with both the protection of the data subject and the need of enabling beneficial uses of AI. However, AI and big data raise three key issues concerning consent: specificity, granularity, and freedom.

결론적으로, 다음에서 볼 수 있듯이 동의 및 목적 제한의 개념은 데이터 주체의 보호와 AI의 유익한 사용의 필요성과 일치하는 방식으로 해석될 수 있는 것으로 보인다. 그러나 AI와 빅 데이터는 동의와 관련된 세 가지 주요 문제인 특이성, 세분성 및 자유를 유발합니다.

Specificity (특성)

The first issue pertains to the specificity of consent: does consent to the processing for a certain purpose also cover further AI-based processing, typically for data analytics and profiling? – e.g., can data on sales be used to analyse consumer preferences and send targeted advertising? This seems to be ruled out, since consent needs to be specific, so that it cannot extend beyond what is explicitly indicated. However, the fact that the data subject has only consented to processing for a certain purpose (e.g., client management) does not necessarily rule out that the data can be processed for further legitimate purpose (e.g., business analytics): the further processing is permissible when it is covered by a legal basis, and it is not incompatible with the purpose for which the data were collected.

첫 번째 문제는 동의의 특수성과 관련이 있습니다. 특정 목적의

처리에 대한 동의도 일반적으로 데이터 분석 및 프로파일링을 위한 추가 AI 기반 처리를 포함합니까? – 예를 들어 판매 데이터를 사용하여 소비자 선호도를 분석하고 대상 광고를 보낼 수 있습니까? 동의는 구체적이어야 하기 때문에 명시적으로 지시된 것 이상으로 확장될 수 없기 때문에 이것은 배제된 것으로 보인다. 그러나 데이터 주체가 특정 목적 (예 : 고객 관리)을 위한 처리에만 동의했다는 사실이 데이터가 추가 합법적인 목적 (예 : 비즈니스 분석)을 위해 처리될 수 있다는 것을 반드시 배제하지는 않습니다. 추가 처리가 허용됩니다. 법적 근거가 있고 데이터 수집 목적과 호환되지 않는 경우

The requirement of specificity is attenuated for scientific research as stated in Recital (33), which allows consent to be given not only for specific research projects, but also for areas of scientific research.

Recital (33)에 명시된 바와 같이 과학 연구에 대한 특이성 요구 사항이 완화되어 특정 연구 프로젝트 뿐만 아니라 과학 연구 분야에도 동의할 수 있습니다.

79 Cate et al (2014, 9).
80 Cavoukian (2015), Calo (2012).

It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.

데이터 수집 시 과학적 연구 목적으로 개인 데이터 처리의 목적을 완전히 식별할 수 없는 경우가 종종 있습니다. 그러므로 과학 연구에 대한 인정된 윤리적 기준에 따라 데이터 과목은 과학 연구의 특정 영역에 동의할 수 있어야한다. 데이터 주체는 의도한 목적에 따라 허용되는 범위 내에서 특정 연구 영역이나 연구 프로젝트의 일부에만 동의할 수 있는 기회를 가져야합니다.

Granularity (세분성, the scale or level of detail present in a set of data or other phenomenon.)

The second issue pertains to the granularity of consent. For instance, is a general consent to any kind of analytics and profiling sufficient to authorise the AI-based sending of targeted commercial

or political advertising? Recital (43) addresses granularity as follows:

두 번째 문제는 동의의 세분성에 관한 것입니다. 예를 들어 AI를 기반으로 한 타깃 광고 또는 정치 광고의 전송을 승인하기에 충분한 모든 종류의 분석 및 프로파일링에 대한 일반적인 동의가 있습니까? Recital (43)은 다음과 같이 세분성을 처리합니다.

> ***Consent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case.***
>
> ***개별 사례에 적합하더라도 다른 개인 데이터 처리 작업에 대해 별도의 동의가 허용되지 않으면 동의가 자유롭게 제공되지 않는 것으로 간주됩니다.***

This has two implications for AI application. First it seems that the data subject should not be required to jointly consent to essentially different kinds of AI-based processing (e.g., to economic and political ads). Second, the use of a service should not in principle be dependent on an agreement to be subject to profiling practices. Consent to profiling must be separate from access to the service.81

이것은 AI 응용 프로그램에 두 가지 의미가 있습니다. 먼저 데이터 주체가 본질적으로 다른 종류의 AI 기반 처리 (예 : 경제 및 정치 광고)에 공동으로 동의할 필요는 없는 것 같습니다. 둘째, 서비스의 사용이 원칙적으로 프로파일링 관례에 따른 계약에 의존해서는 안됩니다. 프로파일링 동의는 서비스 액세스와 분리되어야 합니다.81

Freedom

The third issue pertains to the freedom of consent: can consent to profiling be considered freely given? This issue is addressed in Recital (42), which excludes the freedom of consent when 'the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment.' According to Recital (43), consent is not free under situations of 'clear imbalance:'

세 번째 문제는 동의의 자유와 관련이 있습니다. 프로파일링에 대한 동의가 자유롭게 주어질 수 있습니까? 이 문제는 Recital (42)에서 다루어지며, '데이터 주체가 진실하거나 자유로운 선택이 없거나 손해없이 동의를 거부하거나 철회할 수 없는 경우'동의의 자유를 배제합니다. Recital (43)에 따르면, '명확한 불균형'상황에서는 동의가 자유롭지 않다 :

In order to ensure that consent is freely given, consent should not provide a valid legal ground for the processing of personal data in a specific case where there is a clear imbalance between the data subject and the controller.

동의가 자유롭게 이루어 지려면, 데이터 주체와 관리자 사이에 불균형이 분명한 경우에 개인 정보 처리에 대한 유효한 법적 근거를 제공해서는 안됩니다.

Situations of imbalance are prevalent in the typical contexts in which AI and data analytics are applied to personal data. Such situations exist in the private sector, especially when a party enjoys market dominance (as is the case for leading platforms), or a position of private power (as is the case for employers relative to their employees). They also exist between public authorities and the individuals who are subject to the powers by such authorities. In all these cases, consent cannot provide a sufficient legal basis, unless it can be shown that there are no risks of 'deception, intimidation, coercion or significant negative consequence if [the data subject] does not consent.'82

AI와 데이터 분석이 개인 데이터에 적용되는 일반적인 상황에서 불균형 상황이 널리 퍼져 있습니다. 이러한 상황은 민간 부문, 특히 당사자가 시장 지배력 (선도적 플랫폼의 경우) 또는 개인 권력의 위치 (직원에 대한 고용주의 경우)를 누릴 때 존재합니다. 그들은 또한 공공 당국과 그러한 당국의 권력을 받는 개인들 사이에도 존재합니다. 이러한 모든 경우에, '데이터 주체'가 동의하지 않는 경우 '기만, 협박, 강요 또는 중대한 부정적인 결과의 위험이 없음을 나타내지 않는 한 동의는 충분한 법적 근거를 제공할 수 없습니다.'82

Finally, consent should be invalid when refusal or withdrawal of consent is linked to a detriment that is unrelated to the availability of the personal data for which consent was refused (e.g., a patients are told that in order to obtain a medical treatment they must consent that their medical data are used for purposes that are not needed for that treatment). This also applies to cases in which consent is required by the provider of a service, even though the processing is not necessary for performing the service.

마지막으로, 동의 거부 또는 철회가 동의가 거부된 개인 데이터의 가용성과 관련이 없는 손해와 관련이 있는 경우 동의가 유효하지 않아야 합니다 (예 : 환자는 의학적 치료를 받기 위해서는 반드시 동의해야 함 그들의 의료 데이터는 그 치료에 필요하지 않은 목

적으로 사용된다). 이는 서비스 수행에 처리가 필요하지 않더라도 서비스 제공 업체가 동의해야 하는 경우에도 적용됩니다.

> ***if the performance of a contract, including the provision of a service, is dependent on the consent despite such consent not being necessary for such performance.***
>
> ***서비스 제공을 포함하여 계약의 이행이 그러한 이행에 필요하지 않은 동의에도 불구하고 동의에 의존하는 경우***

This typically is the case when the closing of a contract for a service is conditioned on the user's consent to being profiled, the profiling not being needed to provide the service to the individual user.

이는 일반적으로 서비스 계약의 종결이 프로파일링에 대한 사용자의 동의에 따라 결정되는 경우이며, 프로파일링은 개별 사용자에게 서비스를 제공하는 데 필요하지 않습니다.

81 Article 29 Working Party Guidelines on consent under Regulation 2016/679. Wp259

82 Article 29 Working Party Guidelines on consent under Regulation 2016/679. Wp259, 7

3.2. AI and the data protection principles

As many authors have observed, AI and big data challenge key data protection principles. In this section, we shall consider each principle separately, so as to determine the extent to which it may constrain intelligent processing.

많은 저자들이 관찰한 것처럼 AI와 빅 데이터는 주요 데이터 보호 원칙에 도전합니다. 이 섹션에서는 지능형 처리를 제한할 수 있는 정도를 결정하기 위해 각 원칙을 개별적으로 고려합니다.

3.2.1. Article 5(1)(a) GDPR: Fairness, transparency

제5조 (1) (a) GDPR : 공정성, 투명성

Article 5(1)(a) requires that personal data should be processed 'lawfully, fairly and in a transparent manner in relation to the data subject.'

제 5 (1) (a) 조는 개인 정보를 '정보 주체와 관련하여 합법적이고

공정하며 투명한 방식으로 처리'할 것을 요구하고 있습니다.

Transparency

The idea of transparency is specified in Recital 58, which focuses on conciseness, accessibility and understandability.

투명성에 대한 아이디어는 간결성, 접근성 및 이해에 중점을 둔 Recital 58에 명시되어 있습니다.

> ***The principle of transparency requires that any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used.***
>
> ***투명성 원칙은 일반인이나 데이터 주체에게 전달되는 모든 정보가 간결하고 쉽게 접근 가능하고 이해하기 쉬워야 하며 명확하고 평범한 언어와 추가로적절한 경우 시각화가 사용되어야 합니다.***

As we shall clarify in what follows, this idea is related, but distinct, from the idea of transparent and explainable AI. In fact, the latter idea involves building a 'scientific' model of the functioning of an AI system, rather than providing sufficient information to lay people, relatively to issues that are relevant to them.

다음에 나오는 내용을 명확하게 설명할 때이 아이디어는 투명하고 설명 가능한 AI 아이디어와 관련이 있지만 별개입니다. 실제로 후자의 아이디어는 AI 시스템의 기능에 대한 '과학적인'모델을 구축하는 것인데, 관련 문제와 관련하여 사람들을 배치하기에 충분한 정보를 제공하는 것이 아닙니다.

Informational fairness

Two different concepts of fairness can be distinguished in the GDPR. The first, which we may call 'information fairness' is strictly connected to the idea of transparency It requires that data subjects are not deceived or misled concerning the processing of their data, as is explicated in Recital (60):

GDPR에서 두 가지 다른 공정성 개념을 구별할 수 있습니다. 우리가 '정보 공정성'이라고 부르는 첫 번째는 투명성이라는 개념과

밀접한 관련이 있습니다. Recital (60)에 설명 된 대로 데이터 주제에 대한 데이터 처리와 관련하여 데이터 주체를 속이거나 오도하지 않아야 합니다.

> ***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.***
>
> ***공정하고 투명한 처리의 원칙은 데이터 주체에게 처리 작업의 존재 및 그 목적에 대한 정보를 제공해야 합니다. 컨트롤러는 개인 정보가 처리되는 특정 상황과 상황을 고려하여 공정하고 투명한 처리를 보장하는 데 필요한 추가 정보를 데이터 주체에게 제공해야 합니다.***

The same recital explicitly requires that information is provided on profiling:

동일한 리사이틀은 정보가 프로파일링에 제공되도록 명시적으로

요구합니다.

> ***Furthermore, the data subject should be informed of the existence of profiling and the consequences of such profiling.***
>
> ***또한, 데이터 주체는 프로파일링의 존재 및 프로파일링의 결과에 대해 통보 받아야한다.***

Informational fairness is also linked to accountability, since it presumes that the information to be provided makes it possible to check for compliance. Informational fairness raises specific issues in connection with AI and big data, because of the complexity of the processing involved in AI-applications, the uncertainty of its outcome, and the multiplicity of its purposes. The new dimension of the principle pertains to the explicability of automated decisions, an idea that is explicitly affirmed in the GDPR, as we shall see in the following section. Arguably, the idea of transparency as explicability can be extended to automated inferences, even when a specific decision has not yet been adopted.

정보의 공정성은 책임 성과 연결되어 있습니다. 제공되는 정보가

준수 여부를 확인할 수 있다고 가정하기 때문입니다. 인공 지능 응용 프로그램과 관련된 처리의 복잡성, 결과의 불확실성 및 그 목적의 다양성으로 인해 정보 공평성은 AI 및 빅 데이터와 관련하여 특정 문제를 제기합니다. 이 원칙의 새로운 차원은 다음 섹션에서 볼 수 있듯이 GDPR에서 명시적으로 확인된 아이디어 인 자동화된 결정의 설명 가능성과 관련이 있습니다. 아마도 특정 결정이 아직 채택되지 않은 경우에도 설명 가능성으로서의 투명성이라는 개념은 자동화된 추론으로 확장될 수 있습니다.

A specific aspect of transparency in the context of machine learning concerns access to data, in particular to the system's training set. Access to data may be needed to identify possible causes of unfairness resulting from inadequate or biased data or training algorithm. This is particularly important when the learned algorithmic model is opaque, so that possible flaws cannot be detected though its inspection.

머신러닝 맥락에서 투명성의 특정 측면은 데이터, 특히 시스템의 훈련 세트에 대한 액세스에 관한 것이다. 부적절하거나 편향된 데이터 또는 교육 알고리즘으로 인해 발생할 수 있는 불공정의 원인을 식별하기 위해 데이터에 액세스해야 할 수 있습니다. 이것은 학습된 알고리즘 모델이 불투명할 때 특히 중요하므로 검사를 통해 가능한 결함을 감지할 수 없습니다.

Substantive fairness 실질적인 공정성

Recital (71) points to a different dimension of fairness, i.e. what we may call substantive fairness, which concerns the fairness of the content of an automated inference or decision, under a combination of criteria, which may be summarised by referring to the aforementioned standards of acceptability, relevance and reliability (see Section3.1.2):

결정론 (71)은 다른 차원의 공정성, 즉 자동화된 추론 또는 결정의 내용의 공정성과 관련하여 언급된 표준을 참조하여 요약할 수 있는 기준의 조합에 따라 실질적인 공정성이라고 부를 수 있는 것 수용성, 관련성 및 신뢰성 (섹션 3.1.2 참조) :

In order to ensure fair and transparent processing in respect of the data subject, taking into account the specific circumstances and context in which the personal data are processed, the controller should use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures

appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect.

개인 정보가 처리되는 특정 상황과 상황을 고려하여 데이터 주체와 관련하여 공정하고 투명한 처리를 보장하기 위해, 관리자는 프로파일링을 위해적절한 수학적 또는 통계적 절차를 사용하고 적절한 기술적 및 조직적 조치를 이행해야 합니다 특히 개인 정보의 부정확성을 초래하는 요소가 수정되고 오류의 위험이 최소화되도록 보장하기 위해, 데이터 주체의 이익과 권리에 관련된 잠재적 위험을 고려하여 개인 정보를 보호합니다. 특히 인종 또는 민족, 정치적 견해, 종교 또는 신념, 노동 조합 가입, 유전자 또는 건강 상태 또는 성적 취향에 기초한 자연인에 대한 차별적 영향을 방지하거나 그러한 영향을 미치는 조치를 초래합니다.

3.2.2. Article 5(1)(b) GDPR: Purpose limitation

제5 (1) (b) GDPR : 목적 제한

Article 5(1)(b) sets forth the principle of purpose limitation, according to which personal data should be

제 5조 (1) (b)는 개인 정보에 따라 목적 제한의 원칙을 명시하고 있다.

collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation')

명시적이고 명시적이며 합법적인 목적으로 수집되고 해당 목적과 양립할 수 없는 방식으로 추가 처리되지 않은 경우

공공 이익, 과학 또는 역사적 연구 목적 또는 통계적 목적으로 아카이빙 목적을 위한 추가 처리는 제89 (1) 조에 따라 초기 목적과 양립할 수 없는 것으로 간주되지 않아야 한다 ('목적 제한')

The concept of a purpose also figures in Article 6, which establishes a link between the purpose of processing operations and their legal basis. The notion of a purpose is explicitly mentioned in Article 6 only in relation to the first legal basis, namely, consent, which should be given 'for one or more specific purposes', and for the last legal basis, namely 'the purposes of the legitimate interests pursued by the controller or by a third party'. However, the need for legitimate purpose is implicit in the other legal bases, which consist in the necessity of the processing for performing a contract, complying with a legal obligation, protecting vital interests, performing a task in the public interest or exercising a legitimate authority. Finally, the notion of a purpose also comes up in Articles 13(1)(c) and 14(1)(c), requiring controllers to provide information concerning 'the purposes of the processing for which the personal data are intended as well as the legal basis for the processing.'

목적의 개념은 또한 제6 조에 나와 있으며, 이는 처리 작업의 목적과 법적 근거 사이의 연결 고리를 설정합니다. 목적의 개념은

첫 번째 법적 근거, 즉 '하나 이상의 특정 목적을 위해', 그리고 마지막 법적 근거를 위해, 즉 ' 관리자 또는 제3자가 추구하는 합법적인 이익. 그러나 합법적인 목적의 필요성은 계약 수행, 법적 의무 준수, 중요한 이익 보호, 공익 활동을 수행하거나 합법적인 권한을 행사하는 처리의 필요성으로 구성된 다른 법적 근거에 내재되어 있습니다.. 마지막으로, 목적의 개념은 제13 (1) (c)조 및 제14 (1) (c) 조에도 나타나며, 개인 정보 처리 계획의 목적 뿐만 아니라 개인 정보의 처리 목적에 관한 정보를 관리자에게 제공하도록 요구합니다. 처리의 법적 근거. '

AI and repurposing AI와 용도 변경

A tension exists between the use of AI and big data technologies and the purpose limitation requirement. These technologies enable the useful reuse of personal data for new purposes that are different from those for which the data were originally collected. For instance, data collected for the purpose of contract management can be processed to learn consumers' preferences and send targeted advertising; 'likes' that are meant to express and communicate one's opinion may be used to detect psychological attitudes, political or commercial preferences, etc.

AI와 빅 데이터 기술의 사용과 목적 제한 요건 사이에는 긴장이 존재합니다. 이러한 기술을 사용하면 데이터가 원래 수집된 것과 다른 새로운 목적으로 개인 데이터를 유용하게 재사용할 수 있습니다. 예를 들어, 계약 관리 목적으로 수집된 데이터는 소비자의 선호도를 배우고 대상 광고를 보내기 위해 처리될 수 있습니다. 자신의 의견을 표현하고 전달하기위한 '좋아요'는 심리적 태도, 정치적 또는 상업적 선호 등을 탐지하는 데 사용될 수 있습니다.

To establish whether the repurposing of data is legitimate, we need to determine whether a new purpose is 'compatible' or 'not incompatible' with the purpose for which the data were originally collected. According to the Article 29 WP, the relevant criteria are (a) the distance between the new purpose and the original purpose, (b) the alignment of the new purpose with the data subjects' expectations, the nature of the data and their impact on the data subjects' interests, and (c) the safeguards adopted by the controller to ensure fair processing and prevent undue impacts.83

데이터의 용도 변경이 합법적인지 여부를 확인하기 위해, 새로운 목적이 데이터가 원래 수집된 목적과 '호환'되는지 '호환되지 않는'지 여부를 결정해야 합니다. 제29조 WP에 따르면 관련 기준은 (a) 새로운 목적과 원래의 목적 사이의 거리, (b) 새로운 목적과 데이터 주체의 기대, 데이터의 성격 및 그 영향에 대한 정렬 데이

터 주체의 이익 및 (c) 공정한 처리를 보장하고 과도한 영향을 방지하기 위해 컨트롤러가 채택한 보호 조치.83

Though all these criteria are relevant to the issue of compatibility, they do not provide a definite answer to the typical issues pertaining to the reuse of personal data in AI applications. To what extent can the repurposing of personal data for analytics and AI be compatible with the purpose of the original collection? Should the data subjects be informed that their data is being repurposed? To address such issues, we need to distinguish what is at stake in the inclusion of a person's data in a training set from the application of a trained model to a particular individual.

이러한 모든 기준은 호환성 문제와 관련이 있지만 AI 응용 프로그램에서 개인 데이터의 재사용과 관련된 일반적인 문제에 대한 명확한 답변을 제공하지는 않습니다. 분석 및 AI를위한 개인 데이터의 용도 변경은 원본 수집의 목적과 어느 정도 호환될 수 있습니까? 데이터 주체에게 데이터의 용도가 변경되었다는 알림을 받아야합니까? 이러한 문제를 해결하기 위해 훈련된 모델을 특정 개인에게 적용하는 것과 훈련 세트에 개인 데이터를 포함시키는 데 있어 중요한 사항을 구별해야 합니다.

Personal data in a training set

In general, the inclusion of a person's data in a training set is not going to affect to a large extent that particular person, since the record concerning a single individual is unlikely to a make a difference in a model that is based in a vast set of such records. However, the inclusion of a single record exposes the data subject to risks concerning the possible misuse of his or her data, unless the information concerning that person is anonymised or deleted once the model is constructed.

일반적으로 훈련 세트에 개인의 데이터를 포함시키는 것은 특정 개인에 크게 영향을 미치지 않을 것입니다. 단일 개인에 관한 기록이 방대한 세트에 기반을 둔 모델에서 차이를 만들지 않기 때문입니다. 그러한 기록의. 그러나 단일 레코드를 포함하면 모델을 구성한 후에 해당 개인에 대한 정보가 익명화되거나 삭제되지 않는 한 데이터의 오용 가능성과 관련된 위험에 노출됩니다.

Moreover, when considered together with the data provided by similar individuals, the data concerning a person, once included in the training set, contribute to enabling the system's inference concerning a group of people, i.e., the group of all the individuals

who share the similarities supporting the inference. Therefore, we may say that the set of all such records affects the common interest of the group in which that person is included. Consider for instance the use of a patient's genetic data to train a model that is then used to diagnose present diseases, or to determine their propensity to develop a disease in the future. The inclusion of a patient's data in a training set will contribute little to the model's predictive power, and it will not specifically affect the patient (unless his or her data are misused). However, the inclusion of the patient's data, alongside with the data of other similar patients, may create a risk for the group of all the patients who might be affected by predictions based on such data. For instance, assume that the trained model links certain predictors to a high probability of a future health issue. Patients who share such predictors, when their data is fed to the model, may either find themselves at an advantage (prevention based on predictive medicine) or at a disadvantage (e.g., discrimination in recruitment or insurance) depending on the how the prediction is used. The risks for the group increase if the predictive model is made available to third parties, which may use it in ways that the data subjects did not anticipate when providing their data.

더욱이, 유사한 개인에 의해 제공된 데이터와 함께 고려될 때, 일

단 훈련 세트에 포함된 개인에 관한 데이터는 한 그룹의 사람들, 즉 유사성을 공유하는 모든 개인의 그룹에 관한 시스템의 추론을 가능하게 하는 데 기여한다 추론을 지원합니다. 그러므로, 우리는 그러한 모든 기록이 그 사람이 포함된 집단의 공통 관심사에 영향을 미친다고 말할 수 있습니다. 예를 들어 환자의 유전자 데이터를 사용하여 현재 질병을 진단하는 데 사용되는 모델을 훈련시키거나 미래에 질병을 발병시키는 경향을 결정하는 것을 고려하십시오. 훈련 세트에 환자 데이터를 포함시키는 것은 모델의 예측력에 거의 영향을 미치지 않으며, 환자의 데이터가 잘못 사용되지 않는 한 환자에게 특별히 영향을 미치지 않습니다. 그러나 다른 유사한 환자의 데이터와 함께 환자의 데이터를 포함하면 그러한 데이터를 기반으로 한 예측에 영향을 받을 수 있는 모든 환자 그룹에 위험을 초래할 수 있습니다. 예를 들어, 훈련된 모델이 특정 예측 변수를 미래 건강 문제의 가능성이 높은 것으로 연결한다고 가정합니다. 데이터를 모델에 제공할 때 이러한 예측 변수를 공유하는 환자는 예측 사용 방법에 따라 장점 (예측 의학에 근거한 예방) 또는 단점 (예 : 모집 또는 보험 차별)에 처할 수 있습니다.. 예측 모델을 제3 자에게 제공할 경우 그룹에 대한 위험이 증가하여 데이터 주체가 데이터를 제공할 때 예상하지 못한 방식으로 이를 사용할 수 있습니다.

Personal data for individualised inferences

개별 추론을 위한 개인 데이터

While, as just noted, the inclusion of a person's data in a training set does not lead to significant impacts on that person, an individual is directly affected when his or her personal data are used as input in the algorithmic model that has been created on the basis of that training set, in order to make inferences concerning that individual. Consider, for instance, the case in which someone's medical data are entered into a model to make a medical diagnosis or to determine that person's prospective health condition. In such a case, we are clearly in the domain of profiling, since the input data (the predictors) concerning an individual are used to infer further personal data concerning him or her.

방금 언급했듯이 훈련 세트에 개인 데이터를 포함시키는 것이 그 개인에게 큰 영향을 미치지는 않지만 개인은 개인 데이터가 생성된 알고리즘 모델에서 입력으로 사용될 때 직접 영향을 받습니다. 그 개인에 관한 추론을 하기 위해 그 훈련 세트의 기초. 예를 들어, 누군가의 의료 데이터가 의료 진단을 내리거나 그 사람의 건강 상태를 판단하기 위해 모델에 입력된 경우를 고려하십시오. 이 경우 개인에 관한 입력 데이터 (예측 자)가 개인에 관한 추가 개인 데이터를 유추하는 데 사용되므로 프로파일링 영역에 있습니다.

Let us now consider how the criteria for non-incompatibility established by the Article 29 WP apply on the one hand to the inclusion of personal data in a training set, and on the other hand to the use of personal data as input to profiling algorithms.

제 29조 WP에 의해 확립된 비 호환성에 대한 기준이 한편으로는 훈련 세트에 개인 데이터를 포함시키는 한편, 다른 한편으로는 프로파일링 알고리즘에 대한 입력으로서 개인 데이터를 이용하는 것에 어떻게 적용되는지 살펴 보자.

83 Opinion 03/2013 on purpose limitation.

With regard to the use of a person's data in a training set, it seems that since the person is not directly affected by the use of her personal data, the distance between the new purpose and the original purpose should not be a primary concern, nor should be the data subject's expectations. However, we need to consider the risk that the data are misused, against the interest of the data subject (the risk is particularly serious for data on health or other sensitive conditions), as well as the possibility of mitigating this risk through anonymisation or pseudonymisation. Adequate security

measures also are the key precondition for the legitimate use of personal data in a training set.

훈련 세트에서 개인의 데이터 사용과 관련하여, 개인 데이터의 사용에 의해 개인이 직접 영향을 받지 않기 때문에, 새로운 목적과 원래 목적 사이의 거리가 주요 관심사가 되어서는 안되며 데이터 주체의 기대가 되어야 합니다. 그러나 데이터 주체의 이익에 대한 데이터 오용의 위험성 (건강 또는 기타 민감한 조건에 대한 데이터의 위험이 특히 심각함)과 익명화 또는 가명 화를 통해 이 위험을 완화할 수 있는 가능성을 고려해야 합니다.. 훈련 세트에서 개인 데이터를 합법적으로 사용하기위한 적절한 전제 조건도 적절한 보안 조치입니다.

Different considerations pertain to the use of a personal data as input to algorithmic models that provide inferences concerning the data subject. This case clearly falls within the domain of profiling as the inference directly affects the individuals concerned. Therefore, the criteria indicated by the Article 29 WP have to be rigorously applied.

데이터 주체와 관련된 추론을 제공하는 알고리즘 모델에 대한 입력으로 개인 데이터를 사용하는 것과는 다른 고려 사항이 있습니다. 추론이 관련 개인에게 직접 영향을 미치기 때문에 이 사례는

프로파일링 영역에 분명히 속한다. 그러므로 제29조 WP에 의해 표시된 기준은 엄격하게 적용되어야 한다.

Obviously, the two uses of personal data may be connected in practice: personal data (for instance data outlining an individual's clinical history, or the history of his or her online purchases) can be processed to learn an algorithmic model, but they can also be used as inputs for the same or other algorithmic models (e.g., to predict additional health issues, or further purchases).

개인 데이터의 두 가지 용도는 실제로 연결될 수 있습니다. 개인 데이터 (예 : 개인의 임상 기록 또는 온라인 구매 기록을 나타내는 데이터)는 알고리즘 모델을 배우기 위해 처리될 수 있지만, 동일하거나 다른 알고리즘 모델의 입력으로 사용됩니다 (예 : 추가 건강 문제 또는 추가 구매 예측).

3.2.3. Article 5(1)(c) GDPR: Data minimisation

제5조 (1) (c) GDPR : 데이터 최소화

Article 5(1)(c) states the principle of data minimisation, according

to which personal data should be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.' The principle of minimisation is also contained in Recital 78, requiring the 'minimisation of personal data' as an organisational measure for data protection by design and by default.

제 5 (1) (c) 조는 개인 정보가 '처리 목적과 관련하여 필요한 것으로 적절하고, 적절하며, 제한되어야 한다'는 정보 최소화 원칙을 명시하고있다. 최소화의 원칙은 Recital 78에도 포함되어 있으며, 설계 및 기본적으로 데이터 보호를 위한 조직적 조치로 '개인 데이터 최소화'가 필요합니다.

There is a tension between the principle of minimisation and the very idea of big data and data analytics, which involves using AI and statistical methods to discover new unexpected correlations in vast datasets. This tension may be reduced by the following considerations.

최소화 원칙과 빅 데이터 및 데이터 분석이라는 아이디어 사이에는 긴장이 있습니다. AI와 통계 방법을 사용하여 방대한 데이터 세트에서 예기치 않은 새로운 상관 관계를 발견합니다. 이 장력은 다음 고려 사항에 의해 감소될 수 있습니다.

First, the idea of minimisation should be linked to an idea of proportionality. Minimisation does not exclude the inclusion of additional personal data in a processing, as long as the addition of such data provides a benefit, relatively to the purposes of the processing that outweigh the additional risks for the data subjects. Even the utility of future processing may justify retaining the data, as long as adequate security measures are in place. In particular, pseudonymisation, in combination with other security measures, may contribute to limit risks and increase therefore the compatibility of retention with minimisation.

첫째, 최소화의 아이디어는 비례의 아이디어와 연결되어야 합니다. 데이터 주체에 대한 추가 위험을 능가하는 처리의 목적에 비해, 그러한 데이터의 추가가 이익을 제공하는 한, 처리에 추가 개인 데이터를 포함시키는 것을 최소화하는 것은 아니다. 향후 처리의 유용성조차도 적절한 보안 조치가 마련되어 있는 한 데이터 유지가 정당화될 수 있습니다. 특히 가명 화는 다른 보안 조치와 함께 위험을 제한하고 최소화와의 보존 호환성을 높이는 데 기여할 수 있습니다.

Second, the processing of personal data for merely statistical

purposes may be subject to looser minimisation requirements. In such a case the data subjects' information is considered only as an input to a training set (or a statistical database) and is not used for predictions or decisions concerning individuals. This is stated in Recital (162) which links statistical processing to the objective of producing statistical surveys or results:

둘째, 단지 통계적 목적으로 개인 데이터를 처리하는 것은 최소화 요구 사항이 느슨해 질 수 있습니다. 이 경우 데이터 주체의 정보는 훈련 세트 (또는 통계 데이터베이스)에 대한 입력으로 만 간주되며 개인에 관한 예측 또는 결정에는 사용되지 않습니다. 이는 통계 처리 또는 통계 조사 또는 결과 생성의 목표와 연결되는 Recital (162)에 명시되어 있습니다.

> ***Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. Those statistical results may further be used for different purposes, including a scientific research purpose.***
>
> ***통계 목적은 통계 조사 또는 통계 결과 생성에 필요한 개인 데이터 수집 및 처리 작업을 의미합니다. 이러한 통계 결과는 과학적 연구 목적을 포함하여 다른 목적으로 추가***

로 사용될 수 있습니다.

Thus, the processing of personal data for statistical purposes should not deliver personal data as its final result. In particular, the personal data processed for statistical purpose should not be used for adopting decisions on individuals.

따라서 통계 목적으로 개인 데이터를 처리할 때 개인 데이터를 최종 결과로 제공해서는 안됩니다. 특히, 통계 목적으로 처리된 개인 데이터는 개인의 결정을 채택하는 데 사용되어서는 안됩니다.

The statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person.

통계적 목적은 통계적 목적에 대한 처리 결과가 개인 데이터가 아니라 집계 데이터이며 이 결과 또는 개인 데이터가 특정 자연인에 대한 측정 또는 결정을 지원하는 데 사용되지 않음을 의미합니다

Since the data subject is not individually affected by statistical processing, the proportionality assessment, as far as data protection is concerned, concerns the comparison between the (legitimate) interest in obtaining the statistical results, and the risks of the data being misused for non-statistical purposes.

데이터 주체는 통계적 처리에 의해 개별적으로 영향을 받지 않기 때문에, 데이터 보호와 관련하여 비례 평가는 통계적 결과를 얻는 데 대한 (정상적인) 관심과 비 통계적 목적으로 데이터가 오용될 위험 사이의 비교에 관한 것입니다.

It is true that the results of statistical processing can affect the collective interests of the data subjects who share the factors that are correlated to certain inferences (e.g., the individuals whose live style and activities are correlated to certain pathologies, certain psychological attitudes, or certain market preferences or political views). The availability of this correlation exposes all members of the group – as soon as their membership in the group is known – to such inferences. However, as long as the correlation is not meant to be applied to particular individuals, on the basis of data concerning such individual (data determining its belonging to the

group) statistical processing remains outside of data protection. On the contrary, the information used to ascribe a person to a group and the person's ascription to that group are personal data, and so are the consequentially inferred data concerning that person. This idea is expressed in at footnote 5 in the 2017 Council of Europe Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data

통계적 처리 결과가 특정 추론과 관련된 요소를 공유하는 데이터 주체의 집단적 이해에 영향을 미칠 수 있다는 것은 사실입니다 (예 : 라이브 스타일과 활동이 특정 병리, 특정 심리적 태도 또는 특정과 관련이 있는 개인) 시장 선호도 또는 정치적 견해). 이 상관 관계의 가용성은 그룹의 구성원이 알려진 즉시 그룹의 모든 구성원을 그러한 추론에 노출시킵니다. 그러나, 상관이 특정 개인에게 적용되도록 의도되지 않는 한, 그러한 개인에 관한 데이터 (그룹에 속하는 데이터를 결정하는 데이터)에 기초하여 통계 처리는 데이터 보호 외부에 남아있다. 반대로, 개인을 그룹에 등록하는 데 사용된 정보와 해당 그룹에 대한 개인의 설명은 개인 데이터이므로 해당 개인에 대한 결과적으로 추론된 데이터도 마찬가지입니다. 이 아이디어는 빅 데이터 세계에서 개인 데이터 처리와 관련하여 개인 보호에 관한 2017 유럽위원회 지침의 각주 5에 표시되어 있습니다.

personal data are also any information used to single out people from data sets, to take decisions affecting them on the basis of group profiling information.

개인 데이터는 또한 그룹 프로파일링 정보를 기반으로 데이터 세트에 영향을 미치는 결정을 내리기 위해 데이터 세트에서 사람들을 발굴하는 데 사용되는 모든 정보입니다.

Thus, neither in the GDPR nor in the in Guidelines can we yet find an explicit endorsement of group privacy as an aspect of data protection. On the contrary, the need to take into account group privacy has been advocated by many scholars.84 However, as we shall see in the following, a preventive risk-management approach can contribute to the protection of group privacy also in the context of GDPR.

따라서 GDPR이나 가이드 라인에서 데이터 보호의 측면으로서 그룹 프라이버시를 명시적으로 보증할 수는 없습니다. 반대로 많은 학자들은 단체 프라이버시를 고려해야 한다는 주장을 옹호 해왔다.84 그러나 다음과 같이 예방적 리스크 관리 접근법은 GDPR의 맥락에서도 단체 프라이버시 보호에 기여할 수 있다..

3.2.4. Article 5(1)(d) GDPR: Accuracy

제5조 (1) (d) GDPR : 정확성

The principle of accuracy is stated in Article 5(1) GDPR that requires data to be 'accurate and, where necessary, kept up to date,' and that initiatives are taken to address inaccuracies:

정확성의 원칙은 데이터가 '정확하고 필요할 경우 최신 상태로 유지'되어야 하고 부정확성을 해결하기위한 조치가 취해 지도록 요구하는 GDPR 5 (1) GDPR에 명시되어 있습니다.

> ***every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.***
>
> ***처리 목적과 관련하여 부정확한 개인 데이터가 지체없이 삭제되거나 수정되도록 모든 합리적인 조치를 취해야 합니다.***

This principle also applies to personal data that are used as an

input for AI system, particularly when personal data are used to make inferences or decisions about data subjects. Inaccurate data may expose data subjects to harm, whenever they are considered and treated in ways that do not fit their identity.

이 원칙은 AI 시스템에 대한 입력으로 사용되는 개인 데이터에도 적용됩니다. 특히 개인 데이터를 사용하여 데이터 주제에 대해 추론하거나 결정하는 경우에 사용됩니다. 부정확한 데이터는 신원에 맞지 않는 방식으로 간주되고 취급될 때마다 데이터 주체를 해칠 수 있습니다.

With regard to machine learning systems, we need to distinguish whether personal data are used only in a training set, to learn general statistical correlations, or rather as input to a profiling algorithm. Obviously, once that the data are available for the training set, the temptation to use the same data to make also individualised inferences will be very strong. Anonymisation, or pseudonymisation, with strong security measures can contribute to reducing the risk

머신러닝 시스템과 관련하여 개인 데이터가 훈련 세트에서만 사용되는지, 일반적인 통계적 상관 관계를 배우기 위해 또는 프로파일링 알고리즘에 대한 입력으로 사용되는지 구별해야 합니다. 교

육 세트에 데이터를 사용할 수 있게 되면 동일한 데이터를 사용하여 개별화된 추론을 수행하려는 유혹이 커질 것입니다. 강력한 보안 조치를 사용하는 익명화 또는 가명화는 위험을 줄이는 데 기여할 수 있습니다.

3.2.5. Article 5(1)(e) GDPR: Storage limitation

제5조 (1) (e) GDPR : 보관 제한

The principle of storage limitation is stated in GDPR at Article 5(1)(e), which prohibits to keep personal data when they are no longer needed for the purposes of the processing.

저장 제한의 원칙은 GDPR에 5 (1) (e) 항에 명시되어 있으며, 처리 목적으로 더 이상 필요하지 않은 개인 데이터를 유지하는 것을 금지합니다.

84 On the Guidelines, see Mantelero (2017).

[Personal data should be] kept in a form which permits identification of data subjects for no longer than is

necessary for the purposes for which the personal data are processed.

[개인 데이터는] 개인 데이터가 처리되는 목적에 필요한만큼 더 이상 데이터 주체를 식별할 수 있는 형태로 유지되어야 합니다.

Longer storage is however allowed for archiving, research, or statistical purposes.

그러나 보관, 연구 또는 통계 목적으로 더 긴 저장 공간이 허용됩니다.

[P]ersonal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');

[P] 개인 정보는 적절한 이행에 따라 제89 조제 1 항에 따라 공공의 이익, 과학 또는 역사적 연구 목적 또는 통계적 목적으로 만 보관 목적으로 만 처리되는 한 더 오랜 기간 동안 저장될 수 있다. 데이터 주체의 권리와 자유를 보호하기 위해이 규정에서 요구하는 기술적 및 조직적 조치 ('저장소 제한');

There is undoubtable tension between the AI-based processing of large sets of personal data and the principle of storage limitation. This tension can be limited to the extent that the data are used for statistical purposes, and appropriate measures are adopted at national level, as discussed above in 3.2.3.

대규모 개인 데이터 세트의 AI 기반 처리와 스토리지 제한 원칙 사이에는 의심할 여지가 없는 긴장이 있습니다. 이 장력은 데이터가 통계 목적으로 사용되는 정도까지 제한될 수 있으며, 3.2.3에서 논의된 바와 같이 국가 차원에서 적절한 조치가 채택된다.

3.3. AI and legal bases (AI 및 법적 근거)

Article 6 GDPR states that all processing of personal data requires

a legal basis. This idea was first introduced in the 1995 Data Protection Directive, and was subsequently constitutionalised in Article 8 of the European Charter of Fundamental Rights, according to which personal data 'must be processed [...] on the basis of the consent of the person concerned or some other legitimate basis laid down by law.'

GDPR 제6 조에 따르면 모든 개인 정보 처리에는 법적 근거가 필요합니다. 이 아이디어는 1995년 데이터 보호 지침에서 처음 소개되었으며, 유럽 기본권 헌장 제8 조에서 헌법 화되었으며, 이에 따라 개인 또는 개인의 동의에 따라 개인 정보가 처리되어야 한다. 법에 의해 규정된 다른 합법적인 근거. '

The processing of personal data in the context of AI application raises some issues relating to the existence of a valid legal basis. To determine when a legal basis may support AI-based processing, we need to separately consider the legal bases set forth in Article 6 GDPR, which states that the processing of personal data only is lawful under the following conditions: (a) consent of the data subject, or necessity (b) for performing or entering into a contract, (c) for complying with a legal obligation, (d) for protecting vital interests (e) for performing a task in the public interest or in the exercise of public authority, or (f) for a legitimate interest.

AI 응용 프로그램의 맥락에서 개인 데이터를 처리하면 유효한 법적 근거의 존재와 관련된 몇 가지 문제가 발생합니다. 법적 근거가 AI 기반 처리를 지원할 수 있는 시점을 결정하기 위해, 당사는 제6조 GDPR에 명시된 법적 근거를 별도로 고려해야 합니다. 이 기준은 개인 데이터 처리가 다음 조건 하에서만 합법적이라고 명시합니다. 데이터 주체, 또는 (b) 계약을 수행하거나 계약을 체결하기 위해 필요, (c) 법적 의무를 준수하기 위해, (d) 중요한 이익을 보호하기 위해 (e) 공익 또는 공공 운동으로 업무를 수행하기 위해 권한, 또는 (f) 정당한 이익을 위해.

3.3.1. Article 6(1)(a) GDPR: Consent(제6조 (1) (a) GDPR : 동의)

A data subject's consent to the processing of his or her personal data by an AI system can have two possibly concurring objects: including such data in a training set, or providing them to an algorithmic model meant to deliver individualised responses. Usually, the data subject's consent covers both. As noted in Section 3.1.3, consent has to be specific, granular and free. It is not easy for all these conditions to be satisfied with regard to the AI-based processing of personal data. Thus,

AI 시스템에 의한 개인 데이터 처리에 대한 데이터 주체의 동의

에는 두 가지 상충되는 개체가 있을 수 있습니다. 트레이닝 세트에 이러한 데이터를 포함시키거나 개별화된 응답을 전달하기위한 알고리즘 모델에 제공하는 것. 일반적으로 데이터 주체의 동의는 두 가지 모두를 다룹니다. 섹션 3.1.3에 명시된 바와 같이, 동의는 구체적이고 세분화되고 자유로워야 합니다. AI 기반의 개인 데이터 처리와 관련하여 이러한 모든 조건을 만족시키는 것은 쉽지 않습니다. 그러므로,

this processing usually needs to rely alternatively or additionally on other legal bases.

이 처리는 일반적으로 다른 법적 기반에 대안적으로 또는 추가로 의존해야 합니다.

The processing of personal data for scientific or statistical purposes may be based on the social significance of such purposes (Article 6(1)(f)), beside the endorsement of such purposes by the data subject. Consent to individual profiling may concur with the necessity or usefulness of such processing for the purposes indicated in the subsequent items of Article 6.

과학적 또는 통계적 목적을 위한 개인 데이터의 처리는 데이터

주체에 의한 그러한 목적의 승인 외에, 그러한 목적의 사회적 중요성에 근거할 수 있습니다 (제6 (1) (f)). 개별 프로파일링에 대한 동의는 제6 조의 후속 항목에 명시된 목적을 위해 그러한 처리의 필요성 또는 유용성과 일치할 수 있습니다.

3.3.2. Article 6(1)(b-e) GDPR: Necessity 필요성

The legal bases from (b) to (e) can be treated together here since they all involve establishing the necessity of the processing for a certain aim: (b) performing or entering (at the request of the data subject) into a contract, (c) for complying with a legal obligation, (d) protecting vital interests (e) performing a task in the public interest or in the exercise of public authority. Thus, such legal bases do not apply to the AI-based processing that is subsequent to or independent of such aims in the specific case at hand.

(b)에서 (e)까지의 법적 근거는 모두 특정 목표에 대한 처리의 필요성을 확립하는 것을 포함하므로 여기에서 함께 취급될 수 있습니다. (b) (데이터 주체의 요청에 따라) 계약에 이행 또는 입력, (c) 법적 의무를 준수하기 위해, (d) 중요한 이익을 보호하기 위해 (e) 공익 또는 공공 기관의 행사를 수행하는 행위 따라서, 그러한 법적 근거는 특정 상황에서 그러한 목표에 뒤 따르거나 독립적인

AI 기반 처리에는 적용되지 않습니다.

For instance, the necessity of using personal data for performing or entering a particular contract does not cover the subsequent use of such data for purposes of business analytics. Similarly, this legal basis does not cover the subsequent use of contract data as input to a predictive-decisional model concerning the data subject, even when the data are used for offering a different contract to the same person. Assume, for instance that the data subject's health data are necessary for performing an insurance contract with the data subject. This necessity would not cover to the use of the same data for offering a new contract to the same data subject, unless the data subject has requested to be considered for a new contract, i.e., unless the data are necessary 'in order to take steps at the request of the data subject prior to entering into a contract' (Article 6(b)).

예를 들어, 특정 계약을 수행하거나 입력하기 위해 개인 데이터를 사용해야하는 것은 비즈니스 분석 목적으로 이러한 데이터의 후속 사용을 다루지 않습니다. 마찬가지로, 이 법적 근거는 데이터가 동일한 사람에게 다른 계약을 제공하는 데 사용되는 경우에도 계약 데이터를 데이터 주체에 관한 예측 결정 모델에 대한 입력으로 후속 사용을 다루지 않습니다. 예를 들어, 데이터 주체와 건

강 보험 계약을 수행하려면 데이터 주체의 건강 데이터가 필요하다고 가정하십시오. 데이터 주체가 새로운 계약에 대해 고려하도록 요청하지 않은 경우, 즉 데이터가 '단계를 수행하기 위해 필요하지 않은 경우'가 아니라면, 동일한 데이터를 동일한 데이터 주체에 새로운 계약을 제공하기 위해 동일한 데이터를 사용하는 데에는 이러한 필요성이 포함되지 않습니다. 계약을 체결하기 전에 데이터 주체의 요청에 따라 '(제6조 (b)).

3.3.3. Article 6(1)(f) GDPR: Legitimate interest 정당한 관심

Article 6(1)(f) provides a general legal basis to the processing of personal data, namely, the necessity of the processing

제 6 (1) (f) 조는 개인 정보 처리, 즉 처리의 필요성에 대한 일반적인 법적 근거를 제공합니다.

> ***for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data,***

개인 정보 보호가 필요한 데이터 주체의 이익 또는 기본 권리 및 자유에 의해 우선권을 행사하는 경우를 제외하고는, 컨트롤러 또는 제3자가 추구하는 합법적 이익의 목적으로,

We may wonder to what extent Article 6(1)(f) may apply to the AI-processing of personal data.85 We have to distinguish the use of personal data in a training set to build/learn an algorithmic model, and their use as an input to a given algorithmic model. In the first case, as long as strong security measures are adopted it – which usually should involve pseudonymisation of the data, and their anonymisation as soon as the model has been completed – seems that the data subject's interests are not severely affected. If the controller is pursuing an interest that is permissible under the law (including an economic interests), it seems that the standard set forth in Article 6(1)(f) could be met.

개인 데이터의 AI 처리에 6 (1) (f) 조가 어느 정도 적용될 수 있는지 궁금할 것입니다.85 알고리즘 모델을 구축/학습하기 위해 훈련 세트에서 개인 데이터의 사용과 그 사용을 구별해야 합니다. 주어진 알고리즘 모델에 대한 입력으로. 첫 번째 경우, 강력한 보안 조치가 채택되는 한 (보통 데이터 가명 화 및 모델이 완료되자마자 익명화가 포함되어야 함) 데이터 주체의 관심사는 크게 영

향을 받지 않는 것으로 보입니다. 지배인이 법에 따라 허용되는 이익 (경제적 이익 포함)을 추구하는 경우, 6 (1) (f) 조에 명시된 표준이 충족될 수 있는 것 같습니다.

The situation is much different when the data subjects' data are used in an algorithmic model, to derive conclusions concerning the data subject. Under such a case, the interest of the data subject should be given priority, according to his or her assessment. Thus, the data subject should be asked for his or her consent and have the opportunity to opt out.

데이터 주체의 데이터가 알고리즘 모델에서 사용되어 데이터 주체에 관한 결론을 도출하는 상황은 매우 다릅니다. 그러한 경우, 평가에 따라 데이터 주체의 관심이 우선시 되어야 한다. 따라서 데이터 주체는 동의를 구하고 선택 해제할 수 있는 기회를 가져야합니다.

The legitimate interest test may be important to address the admissibility of those applications that may seriously affect individuals and society, even when they are technologically sound and non-discriminatory.

합법적인 관심사 테스트는 기술적으로 건전하고 비 차별적 일지라도 개인과 사회에 심각한 영향을 줄 수 있는 응용 프로그램의 허용 가능성을 해결하는 데 중요할 수 있습니다.

When an application provides benefits that are outweighed by the disadvantages imposed on the data subjects, we should conclude that the application fails to have a basis according to Article 6(1)(f). This may be the case, as noted above, for systems meant to detect individuals' attitudes from faces, or also to assess workers' performance based on pervasive surveillance, or to detect and influence political views, etc. In all such instances, given the difference in knowledge and power and lack of adequate information, consent by the data subject would not meet the requirement of freedom and information in the GDPR, and thus could not provide an alternative legal basis. Thus, the processing should be considered to be unlawful.

신청서가 데이터 주체에 부과된 단점보다 중요한 이점을 제공할 때, 우리는 신청서가 제6 (1) (f) 조에 따라 근거를 가지고 있지 않다는 결론을 내려야합니다. 위에서 언급한 바와 같이, 이는 얼굴로부터 개인의 태도를 감지하거나, 광범위한 감시에 기초하여 근로자의 성과를 평가하거나 정치적 견해 등을 감지하고 영향을 미치는 시스템에 해당될 수 있습니다. 지식과 힘의 차이와 적절한

정보의 부족, 데이터 주체의 동의는 GDPR의 자유와 정보의 요구 사항을 충족시키지 못하므로 대체 법적 근거를 제공할 수 없습니다. 따라서 처리는 불법으로 간주되어야 합니다.

A limitation of the scope of Article 6(1)(f) may consist in the fact that it seems to adopt individualistic perspective, as it only requires a balance between the interests of controllers and on data subject, without taking into accounts broader interests, pertaining to groups or even to society as a whole. However, this limitation of the scope of the balancing test according to Article 6(1)(f) may have a reason, since the assessment of the social merit of a processing operation, and the decision to outlaw it based on this assessment, should be adopted on the basis of on a wide debate, and according to the determination or at least to the directions, of politically responsible bodies.

제 6 (1) (f) 조의 범위의 제한은 개인의 관점을 채택하는 것처럼 보일 수 있는데, 이는 광범위한 이해 관계를 고려하지 않고 통제관의 이해와 데이터 주제에 대한 균형 만 필요하기 때문입니다. 그룹이나 사회 전체에 관한 것입니다. 그러나, 제6 (1) (f) 조에 따른 균형 시험 범위의 이러한 제한은 처리 작업의 사회적 가치 평가 및이 평가에 근거하여 이를 금지하기로 결정한 이유가 있을 수 있다. 정치적으로 책임있는 기구의 결정에 따라 또는 적어도

지시에 따라 폭 넓은 토론을 바탕으로 채택되어야 한다.

85 On legitimate interest, see Kamara and De Hert (2019).

3.3.4. Article 6(4) GDPR: Repurposing 용도 변경

A key issue concerning AI applications pertains to repurposing of personal data. This is an issue on which the provision of the GDPR are unclear. The general idea is stated Article 5(1)(b) as an articulation of the principle of purpose limitation. Personal data shall be 'not further processed in a manner that is incompatible' with the original purposes. The prohibition of repurposing is also affirmed Recital 50, according to which the further processing of personal data for new purposes is only allowed when it is compatible with the original purposes:

AI 응용 프로그램과 관련된 주요 문제는 개인 데이터의 용도 변경과 관련이 있습니다. 이것은 GDPR의 조항이 명확하지 않은 문제입니다. 일반적인 아이디어는 목적 제한의 원칙을 나타내는 것으로서 5 (1) (b)에 명시되어 있습니다. 개인 정보는 원래 목적과 호환되지 않는 방식으로 더 이상 처리되지 않아야 합니다. 용도

변경 금지는 Recital 50에서도 확인되며, 새로운 목적을 위한 개인 데이터의 추가 처리는 원래 목적과 호환되는 경우에만 허용됩니다.

> ***The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected.***
>
> ***개인 정보가 처음 수집된 목적 이외의 목적으로 개인 데이터를 처리하는 것은 처리가 개인 데이터가 처음 수집된 목적과 호환되는 경우에만 허용되어야 합니다.***

Compatibility is however presumed, according to 5(1)(b) when the further processing is meant to serve purposes pertaining to archiving, scientific or historical research or statistics:

그러나 추가 처리가 아카이빙, 과학 또는 역사적 연구 또는 통계와 관련된 목적을 위해 사용되는 경우 5 (1) (b)에 따르면 호환성은 추정됩니다.

further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes

공공 이익, 과학 또는 역사적 연구 목적 또는 통계적 목적으로 아카이빙 목적을 위한 추가 처리는 제89 (1) 조에 따라 초기 목적과 양립할 수 없는 것으로 간주된다

Compatibility is also presumed when the new processing is based on a law, for reasons of public interest:

새로운 처리가 법에 근거한 경우 공익성을 이유로 호환성도 추정됩니다.

If the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, Union or Member State law may determine and specify the tasks and purposes for which the further processing should be regarded as compatible and lawful.

공공의 이익을 위해 또는 관리자에 대한 공식 권한을 행사하여 수행된 업무의 수행에 처리가 필요한 경우, 연합 또는 회원국 법률은 추가 처리를 고려해야할 업무 및 목적을 결정할 수 있습니다. 양립하고 합법적입니다.

Article 6(4) specifies that the law allowing for repurposing 'constitutes a necessary and proportionate measure in a democratic society' and that compatibility is established (or substituted) by the data subject's consent. It also spells out possible factors to be taken into account to determine compatibility:

제 6 (4) 조는 '민주주의 사회에서 필요한 헌법적 헌법을 재구성할 수 있는 법'을 규정하고, 데이터 주체의 동의에 의해 양립성을 확립 (또는 대체)한다고 명시하고있다. 또한 호환성을 결정하기 위해 고려해야할 가능한 요소를 설명합니다.

Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent or on a Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard the objectives referred to in Article 23(1), the

controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia:

개인 정보를 수집한 목적 이외의 목적으로 처리하는 것이 데이터 주체의 동의 또는 민주주의 사회에서 언급된 목표를 보호하기 위해 필요하고 비례적인 조치를 구성하는 연합 또는 회원국 법률에 근거하지 않는 경우 제23 (1) 조에서, 개인 정보 처리자는 다른 목적에 대한 처리가 개인 정보의 최초 수집 목적에 적합한 지 여부를 확인하기 위해 특히 다음 사항을 고려해야 한다.

- *any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;*

  개인 정보 수집 목적과 추가 처리 목적 간의 관계;

- *the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;*

  특히 데이터 주체와 컨트롤러 사이의 관계와 관련하

여 개인 데이터가 수집된 상황;

- ***the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10;***

  개인 정보의 특성, 특히 제9 조에 따라 특별 범주의 개인 정보가 처리되는지 또는 제10 조에 따라 범죄 유죄 및 범죄와 관련된 개인 정보가 처리되는지 여부

- ***the possible consequences of the intended further processing for data subjects;***

  데이터 주체에 대한 추가 처리의 가능한 결과;

- ***the existence of appropriate safeguards, which may include encryption or pseudonymisation.***

  암호화 또는 가명화를 포함할 수 있는 적절한 보호 수단의 존재.

The issues of the admissibility of processing personal data for new and different purposes has become crucial in the era of AI and big

data, when vast and diverse masses of data are available and artificial intelligence or statistical methods are then deployed to discover correlations and identify possible causal links. As noted above this may lead to the discovery of unexpected connections based on the combination of disparate sets of data (e.g., connections between lifestyle preferences in social networks and health conditions, between consumer behaviour and market trends, between internet queries and the spread of diseases, between internet likes and political preferences, etc.). The results of these analyses (e.g., correlations discovered between consumers' data and their preferences, spending capacities and purchasing propensities, etc.) can then be used to assess or influence individual behaviour (e.g., by sending targeted advertisements).

AI 및 빅 데이터 시대에 광범위하고 다양한 대량의 데이터를 사용할 수 있고 인공 지능 또는 통계적 방법을 사용하여 상관 관계를 발견하고 가능한지를 식별할 때 새롭고 다양한 목적으로 개인 데이터를 처리할 수 있는 문제가 중요해졌습니다. 인과 관계 링크. 위에서 언급한 바와 같이, 이것은 서로 다른 데이터 세트의 조합 (예를 들어, 소셜 네트워크의 라이프 스타일 선호와 건강 상태, 소비자 행동과 시장 동향, 인터넷 쿼리와 질병의 확산 사이의 연결) 을 기반으로 예기치 않은 연결을 발견할 수 있습니다. , 인터넷 선호도와 정치적 선호도 등). 이러한 분석 결과 (예 : 소비자 데이터

와 선호도 사이에서 발견된 상관 관계, 지출 용량 및 구매 성향 등)를 사용하여 개별 행동을 평가하거나 영향을 줄 수 있습니다 (예 : 대상 광고를 전송).

Repurposing is key in the domain of big data and AI, since the construction of big data sets often involves merging data that had been separately collected for different purposes, and processing such data to address issues that were not contemplated at the time of collection. A key issue for the future of the GDPR pertains to the extent to which the compatibility test will enable us to draw a sensible distinction between admissible and inadmissible reuses of the data for the purposes of analytics.

용도 변경은 빅 데이터 및 AI 영역의 핵심입니다. 빅 데이터 세트의 구성에는 종종 다른 목적으로 별도로 수집된 데이터를 병합하고 수집 시 고려되지 않은 문제를 해결하기 위해 이러한 데이터를 처리하는 과정이 포함되기 때문입니다. GDPR의 미래에 대한 주요 문제는 호환성 테스트를 통해 분석 목적으로 데이터의 허용 및 허용 불가능 재사용 사이의 현명한 차이를 도출할 수 있는 정도와 관련이 있습니다.

Recital (50) does not help us much in addressing this issue, since

it seems to indicate that no legal basis is required for compatible repurposing: 'where the processing is compatible with the purposes for which the personal data were initially collected [...] no legal basis separate from that which allowed the collection of the personal data is required.' Moreover, Recital (50) seems to presume that all processing for statistical purposes is admissible, by affirming that 'further processing for ... statistical purposes should be considered to be compatible lawful processing operations.' This presumption has been limited by the Article 29 WP, who has argued that compatibility must be checked also in the case of statistical processing.

Recital (50)은이 문제를 해결하는 데 큰 도움이 되지 않습니다. 이는 호환 가능한 용도 변경에 법적 근거가 필요하지 않음을 나타냅니다. '처리가 개인 데이터가 처음 수집된 목적과 호환되는 경우 [...] no 개인 정보 수집이 허용된 것과는 별도의 법적 근거가 요구됩니다. ' 더욱이 Recital (50)은 '통계적 목적을 위한 추가 처리는 합법적인 합법적 처리 작업으로 간주되어야 한다'는 것을 확인함으로써 통계적 목적을 위한 모든 처리가 허용되는 것으로 가정한 것으로 보인다. 이 추정은 제29조 WP에 의해 제한되었으며, 통계 처리의 경우에도 호환성을 확인해야 한다고 주장했다.

In conclusion, it seems that two requirements are needed for

repurposing to be permissible: (a) the new processing must be compatible with the purpose for which the data were collected, and (b) the new processing must have a legal basis (that may be, but is not necessarily, the same of the original processing). Following Recital (50) it seems that statistical processing should be presumed to be compatible, unless reasons for incompatibility appear to exist.

결론적으로, 용도 변경이 허용 되려면 두 가지 요구 사항이 필요한 것 같습니다 : (a) 새로운 처리는 데이터가 수집된 목적과 호환되어야 하며, (b) 새로운 처리는 법적 근거를 가져야합니다 (반드시 원래 처리와 동일해야 하는 것은 아닙니다. Recital (50)에 따르면, 비 호환성 이유가 존재하지 않는 한 통계 처리는 호환되는 것으로 가정해야 합니다.

By applying these criteria to the AI-based reuse of data, we must distinguish whether the data are reused for statistical purposes or rather for profiling. Reuse for a merely statistical purpose should in general be acceptable since it does affect individually the data subject, and thus it should be compatible with the original processing. If the statistical processing is directed towards a permissible goal, such as security or market research, it can also rely on the legal basis of Article 6(1)(f), i.e., on its necessity for

achieving purposes pertaining to legitimate interests.

이러한 기준을 AI 기반의 데이터 재사용에 적용함으로써 데이터를 통계 목적으로 재사용할지 프로파일링에 재사용 할지를 구분해야 합니다. 단순한 통계적 목적의 재사용은 일반적으로 데이터 주제에 개별적으로 영향을 미치므로 원래 처리와 호환 가능해야 하기 때문에 일반적으로 수용 가능해야 합니다. 통계 처리가 안보 또는 시장 조사와 같은 허용 가능한 목표를 향한 경우, 이는 또한 6조 1 항 (f)의 법적 근거, 즉 정당한 이해와 관련된 목적을 달성하기위한 필요성에 의존할 수 있습니다.

Different would be the case for profiling. In such a case, the compatibility assessment is much more uncertain. It should lead to a negative outcome whenever AI-based predictions or decisions may affect the data subject in a way that negatively reverberates on the original purpose of the processing. Consider, for instance, the case in which a person's data collected for medical purpose are inputted to an algorithmic model that determines an insurance price for that person.

프로파일링의 경우가 다릅니다. 이 경우 호환성 평가가 훨씬 불확실합니다. AI 기반 예측 또는 결정이 처리의 원래 목적에 부정적인 영향을 미치는 방식으로 데이터 주제에 영향을 줄 수 있을 때

마다 부정적인 결과로 이어져야 합니다. 예를 들어, 의료 목적으로 수집된 개인의 데이터가 해당 개인의 보험료를 결정하는 알고리즘 모델에 입력되는 경우를 고려하십시오.

It has been argued that the possibility to repurpose personal data for statistical processing is very important for European economy, since European companies need to extract information on markets and social trends – as US and Asian companies do – in order to be competitive.86 The use of personal data for merely statistical purposes should enable companies to obtain the information they need without interfering with the data subjects rights. In fact, as we noted above, according to Recital (162) the processing remains statistical only as long as the result the processing

유럽 기업들은 경쟁력을 유지하기 위해 미국 및 아시아 기업과 마찬가지로 시장 및 사회적 동향에 대한 정보를 추출해야 하므로 통계 처리를 위해 개인 데이터의 용도 변경 가능성이 유럽 경제에 매우 중요하다고 주장되었습니다. 통계적 목적으로 개인 정보를 수집하는 것은 회사가 데이터 주체 권리를 방해하지 않고 필요한 정보를 얻을 수 있도록 해야 합니다. 실제로, 위에서 언급한 바와 같이, Recital (162)에 따르면 처리 결과는 처리 결과만큼만 통계적으로 유지됩니다.

is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person.

개인 데이터가 아니라 집계 데이터이며 이 결과 또는 개인 데이터는 특정 자연인에 대한 측정 또는 결정을 지원하는 데 사용되지 않습니다.

86 On statistical uses and big data, see Mayer-Schonberger and Padova 2016)

3.3.5. Article 9 GDPR: AI and special categories of data (AI 및 특수 데이터 범주)

Article 9 GDPR addresses the so-called sensitive data, namely those personal data whose processing may affect to a larger extent the data subjects, exposing them to severe risks. In this regard AI presents some specific challenges.

제9조 GDPR은 소위 민감한 데이터, 즉 데이터 처리가 데이터 주

체에 더 큰 영향을 미쳐 심각한 위험에 노출될 수 있는 개인 데이터를 다룹니다. 이와 관련하여 AI는 몇 가지 특정 과제를 제시합니다.

The first challenge is connected to re-identifiability. As noted in Section 3.1.1, thanks to AI and big data, pieces of data that apparently are unidentified, not being linked to a specific individual, may be re-identified, and reconnected to the individuals concerned. The re-identification of sensitive data may have serious consequences for the data subject. Consider for instance the case in which de-identified medical records that have been made accessible to the public are re-identified at a later stage, so that the public comes to know the medical conditions of the individuals concerned.

첫 번째 과제는 재식별성과 관련이 있습니다. 3.1.1 절에 언급된 바와 같이, AI 및 빅 데이터 덕분에, 특정 개인과 연결되지 않은 것으로 확인되지 않은 것으로 보이는 데이터는 재식별되고 관련 개인과 다시 연결될 수 있다. 민감한 데이터의 재식별은 데이터 주제에 심각한 결과를 초래할 수 있습니다. 예를 들어, 대중이 접근할 수 있는 비 식별 의료 기록이 이후 단계에서 재식별되어 대중이 관련 개인의 의학적 상태를 알게되는 경우를 고려하십시오.

The second challenge is connected to inference. Thanks to AI and big data, it may be possible to link observable behaviour and known features of individuals – online activity, purchases, likes, movements – to non-observable sensitive data on them such as their psychological attitudes, their health condition their sexual orientation, or their political preferences. Such inferences may expose the concerned individuals to discrimination or manipulation.

두 번째 과제는 추론과 관련이 있습니다. AI와 빅 데이터 덕분에 관찰 가능한 행동과 개인의 알려진 활동 (온라인 활동, 구매, 좋아하는 것, 움직임 등)을 심리적 태도, 건강 상태, 성적 취향 등 관찰할 수 없는 민감한 데이터에 연결할 수 있습니다. 또는 정치적 선호도. 이러한 추론은 관련 개인을 차별 또는 조작에 노출시킬 수 있습니다.

3.4. AI and transparency AI와 투명성

The complexity of AI-based processing, and the fact that such processing cannot be completely anticipated, especially when

based on machine learning, makes it particularly difficult to ensure transparency. The issue of transparency can come up at two points in time, when a data subject's information is inputted in an information system that includes AI algorithms (ex-ante transparency), or after the system's algorithmic model has been applied to the data subject, to deliver specific outcomes concerning his or her (ex-post transparency).

AI 기반 처리의 복잡성과 이러한 처리를 특히 기계학습을 기반으로할 때 완전히 예측할 수 없다는 사실은 특히 투명성을 보장하기 어렵습니다. 투명성 문제는 AI 알고리즘을 포함하는 정보 시스템 (예 : 투명도)에 데이터 주체의 정보가 입력되거나 시스템 알고리즘 모델이 데이터 주체에 적용된 후 두 시점에 발생할 수 있습니다. 자신에 관한 구체적인 결과를 제공합니다 (사후 투명성).

3.4.1. Articles 13 and 14 GDPR: Information duties 정보 업무

Transparency at the stage in which personal data are collected or repurposed is addressed in Articles 13 and 14 GDPR, which require that the data subject be informed about

개인 정보 수집 또는 용도 변경 단계의 투명성은 제13조 및 제14

조 GDPR에 명시되어 있으며, 이는 데이터 주체에 대해 정보를 제공해야 합니다.

> ***the purposes of the processing for which the personal data are intended as well as the legal basis for the processing.***
>
> ***개인 정보가 의도된 처리의 목적 및 처리의 법적 근거.***

Information must also be provided about 'the legitimate interests pursued by the controller or by a third party' where the processing is based on legitimate interest (Article 6(1)(f)). When the data are processed for purposes that could not be foreseen at the time the data were collected – as it is often the case with machine learning applications– the information has to be provided before the new processing, as specified in Article 13(3) and 14(4):

처리가 정당한 이익에 근거한 '관리자 또는 제3자가 추구하는 정당한 이익'에 관한 정보도 제공해야 합니다 (제6조 (1) (f)). 머신러닝 응용 프로그램의 경우와 같이 데이터를 수집할 때 예측할 수 없는 목적으로 데이터를 처리하는 경우, 제13 (3) 조에 명시된대로 새로운 처리 전에 정보를 제공해야 합니다. 14 (4) :

Where the controller intends to further process the personal data for a purpose other than that for which the personal data were collected, the controller shall provide the data subject prior to that further processing with information on that other purpose and with any relevant further information

개인 정보 수집자가 개인 정보를 수집한 목적 이외의 다른 목적으로 개인 정보를 추가로 처리하려는 경우, 해당 개인 정보는 해당 추가 처리에 앞서 해당 다른 목적에 대한 정보 및 관련 추가 정보를 제공해야 합니다.

The obligation to inform the data subject is waved when compliance is impossible, requires a disproportionate effort or impairs the achievement of the objective of the processing (Article 14(5)(b)):

준수가 불가능하거나 불균형한 노력이 필요하거나 처리 목표 달성을 손상시키는 경우, 데이터 주체에게 정보를 제공할 의무가 흔들린다 (제14조 (5) (b)).

[The obligation to provide information to the data subject does not apply when] the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation referred to in paragraph 1 of this Article is likely to render impossible or seriously impair the achievement of the objectives of that processing. In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available.

[정보 주체에 정보를 제공할 의무는 적용되지 않음] 그러한 정보의 제공이 불가능하거나 불공평한 노력, 특히 공익, 과학 또는 역사적 연구 목적 또는 통계적 목적으로 보관 목적으로 처리하는 경우 제89 (1) 조에 언급된 조건과 보호 조치에 따라 또는 이조 제1 항에 언급된 의무가 해당 처리의 목표 달성을 불가능하게 하거나 심각하게 손상시킬 수 있는 한. 그러한 경우, 관리자는 정보를 공개적으로 제공하는 것을 포함하여 데이터 주체의 권리와 자유 및 정당

한 이익을 보호하기위한 적절한 조치를 취해야 합니다.

This limitation only applies when the data have not been collected from the data subject. It is hard to understand why this is the case. In fact, the reasons that justify an exception to the information obligation when the data were not obtained from the data subject, should also justify the same exception when the data were collected from him or her.

이 제한 사항은 데이터 주제에서 데이터를 수집하지 않은 경우에만 적용됩니다. 왜 그런지 이해하기 어렵습니다. 실제로, 데이터 주체로부터 데이터를 얻지 않았을 때 정보 의무에 대한 예외를 정당화하는 이유는 데이터가 데이터 수집 시 동일한 예외를 정당화해야 합니다.

3.4.2. Information on automated decision-making

자동화된 의사 결정에 대한 정보

Article 13(2)(f) and 14(2)(g) GDPR address a key aspect of AI applications, i.e. automated decision-making. The controller has the

obligation to provide:

제 13 (2) (f) 및 14 (2) (g) GDPR은 AI 응용의 주요 측면, 즉 자동화된 의사 결정을 다루고 있습니다. 컨트롤러는 다음을 제공할 의무가 있습니다.

- ***information on 'the existence of automated decision-making, including profiling, referred to in Article 22(1)' and***

  '제 22 조제 1 항에 언급된 프로파일링을 포함한 자동화된 의사 결정의 존재'에 관한 정보

- ***'at least in those cases meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.'***

  '적어도 이러한 경우에 관련된 논리에 대한 의미있는 정보는 물론 데이터 주체에 대한 그러한 처리의 의미 및 예상 결과도 있습니다.'

This provision has been at the centre of a vast debate in the

research community, where this legal requirement has been related to the more general, and indeed fundamental issue of explaining AI systems and their outcomes. Indeed, according to the AI4People document,87 explainability (or explicability) is indeed one of the principles that should inspire the development of AI, along with beneficence, non-maleficence, autonomy and justice. In the current discussion on explainability different perspectives have been put forward.

이 조항은 리서치 커뮤니티에서 광범위한 토론의 중심에 있으며, 이 법적 요구 사항은 AI 시스템과 그 결과를 설명하는 보다 일반적이고 근본적인 문제와 관련이 있습니다. 실제로, AI4People 문서에 따르면, 87 설명 가능성 (또는 설명 가능성)은 실제로 AI의 발전을 촉진해야 하는 원칙 중 하나이며, 수혜, 비 악성, 자율성 및 정의입니다. 설명 가능성에 대한 현재 논의에서 다른 관점이 제시되었습니다.

Computer scientists have focused on the technological possibility of providing understandable models of opaque AI systems (and, in particular, of deep neural networks), i.e., model of the functioning of such systems that can be mastered by human experts. For instance, the following kinds of explanations are at the core of current research on explainable AI:88

컴퓨터 과학자들은 이해하기 쉬운 불투명 AI 시스템 (특히 심층 신경망) 모델, 즉 인간 전문가가 마스터할 수 있는 시스템의 기능 모델을 제공할 수 있는 기술적 가능성에 중점을 두었습니다. 예를 들어, 다음과 같은 종류의 설명은 설명 가능한 AI에 대한 현재 연구의 핵심입니다.

- ***Model explanation, i.e., the global explanation of an opaque AI system through an interpretable and transparent model that fully captures the logic of the opaque system. This would be obtained for instance, if a decision tree or a set of rules was provided, whose activation exactly (or almost exactly) reproduces the functioning of a neural network.***

  모델 설명, 즉 불투명 시스템의 논리를 완전히 포착하는 해석 가능하고 투명한 모델을 통한 불투명 AI 시스템의 글로벌 설명. 예를 들어, 의사 결정 트리 또는 일련의 규칙이 제공된 경우 활성화는 신경망의 기능을 정확하게 (또는 거의 정확하게) 재현합니다.

- ***Model inspection, i.e., a representation that makes it possible to understanding of some specific***

properties of an opaque model or of its predictions. It may concern the patterns of activation in the system's neural networks, or the system's sensitivity to changes in its input factors (e.g. how a change in the applicant's revenue or age makes a difference in the grant of a loan application).

모델 검사, 즉 불투명 모델의 특정 특성이나 예측에 대한 이해를 가능하게 하는 표현. 시스템의 신경망에서의 활성화 패턴 또는 입력 요소의 변화에 대한 시스템의 민감성 (예 : 신청자의 수입 또는 연령의 변화가 대출 신청의 부여에 차이를 만드는 방법)에 관한 것일 수 있습니다.

- *Outcome explanation, i.e., an account of the outcome of an opaque AI in a particular instance. For instance, a special decision concerning an individual can be explained by listing the choices that lead to that conclusions in a decision tree (e.g., the loan was denied because of the applicant's income fell below a certain threshold, his age above a certain threshold, and he did not have enough ownership interest in any real estate available as collateral).*

결과 설명, 즉 특정 인스턴스에서 불투명한 AI의 결과에 대한 설명. 예를 들어, 개인에 관한 특별 결정은 결정 트리에 해당 결론으로 이어지는 선택 사항을 나열하여 설명할 수 있습니다 (예 : 신청자의 소득이 특정 임계 값 아래로 떨어졌고, 연령이 특정 임계 값을 초과하여 대출이 거부된 경우, 또한 담보로 제공되는 부동산에 대한 소유권이 충분하지 않았습니다.

Floridi et al (2018).
Guidotti et al (2019).

The explanatory techniques and models developed within computer science are intended for technological experts and assume ample access to the system being explained.

컴퓨터 과학 내에서 개발된 설명 기술 및 모델은 기술 전문가를 위한 것이며 설명할 시스템에 대한 충분한 액세스를 가정합니다.

Social scientists, on the contrary have focused on the objective of making explanations accessible to lay people, thus addressing the

communicative and dialectical dimensions of explanations. For instance, it has been argued that the following approaches are needed.89

반대로, 사회 과학자들은 사람들이 평신도를 설명할 수 있도록 하는 데 중점을 두어 설명의 의사 소통적이고 변증법적인 차원을 다루었 다. 예를 들어 다음과 같은 접근 방식이 필요하다고 주장했습니다.89

- ***Contrastive explanation: specifying what input values made a difference, determining the adoption of a certain decision (e.g., refusing a loan) rather than possible alternatives (granting the loan);***

  대립적 설명 : 어떤 대안이 아닌 (대출 거부) 특정 결정의 채택 (예 : 대출 거부)을 결정하고 어떤 입력 값이 차이를 만들었는지 명시;

- ***Selective explanation: focusing on those factors that are most relevant according to human judgement;***

  선택적 설명 : 인간의 판단에 따라 가장 관련성이 높은 요소에 중점을 둡니다.

- ***Causal explanation: focusing on causes, rather than***

on merely statistical correlations (e.g., a refusal of a loan can be causally explained by the financial situation of the applicant, not by the kind of Facebook activity that is common for unreliable borrowers);

인과적 설명 : 단순한 통계적 상관 관계보다는 원인에 초점을 맞추는 것 (예 : 대출 거부는 신뢰할 수 없는 차용자에게 공통적인 Facebook 활동의 종류가 아니라 신청자의 재무 상황에 의해 인과적으로 설명될 수 있음);

- ***Social explanation: adopting an interactive and conversational approach in which information is tailored according to the recipient's beliefs and comprehension capacities.***

  사회적 설명 : 정보가 수신자의 신념과 이해력에 따라 조정되는 대화식 대화 방식을 채택합니다.

While the latter suggestions are useful for the ex-post explanation of specific decisions by a system, they cannot be easily applied ex-ante, at the time of data collection (or repurposing). At that time – i.e., before the user's data are inputted either in the training

algorithm, or in the prediction algorithm (using the algorithmic model) – what can be provided to the user is just an indication on the system's general functioning. At this stage, the user should ideally be provided with the following information:

후자의 제안은 시스템에 의한 특정 결정에 대한 사후 설명에 유용하지만, 데이터 수집 (또는 용도 변경)시 사전에 쉽게 적용할 수는 없습니다. 그 당시, 즉 사용자의 데이터가 훈련 알고리즘 또는 예측 알고리즘 (알고리즘 모델 사용)에 입력되기 전에 사용자에게 제공될 수 있는 것은 시스템의 일반적인 기능에 대한 표시일 뿐입니다. 이 단계에서 사용자에게는 이상적으로 다음 정보가 제공되어야 합니다.

- ***The input data that the system takes into consideration (e.g., for a loan application, the applicant's income, gender, assets, job, etc.), and whether different data items are favouring or rather disfavouring the outcome that the applicant hopes for;***

  시스템이 고려하는 입력 데이터 (예 : 대출 신청, 신청자의 소득, 성별, 자산, 직업 등) 및 다른 데이터 항목이 신청자가 원하는 결과를 선호하는지 또는 선호

하지 않는지;

- ***The target values that the system is meant to compute (e.g., a level of creditworthiness, and possibly the threshold to be reached in order for the loan to be approved);***

  시스템이 계산해야 하는 목표값 (예 : 신용도 수준 및 대출이 승인되기 위해 임계 값에 도달할 수 있음)

- ***The envisaged consequence of the automated assessment/decision (e.g., the approval or denial of the loan application).***

  자동 평가/결정의 예상 결과 (예 : 대출 신청의 승인 또는 거부).

It may also be useful to specify what are the overall purposes that the system is aimed to achieve. In the current practice the information that is provided about AI applications is quite scanty, even when profiling is involved. For example, Airbnb explains its profiling practice by asserting that it will:

시스템이 달성하려는 전체적인 목적을 지정하는 것이 유용할 수도 있습니다. 현재의 실무에서 AI 응용 프로그램에 대해 제공되는

정보는 프로파일링 이 관련되어 있어도 상당히 빈약합니다. 예를 들어, 에어비앤비는 다음과 같이 주장하면서 프로파일링 실습을 설명합니다.

> ***conduct profiling on your characteristics and preferences (based on the information you provide to us, your interactions with the Airbnb Platform, information obtained from third parties, and your search and booking history) to send you promotional messages, marketing, advertising and other information that we think may be of interest to you.***
>
> ***귀하가 제공하는 정보, Airbnb 플랫폼과의 상호 작용, 제3자로부터 얻은 정보 및 검색 및 예약 기록에 근거하여 귀하의 특성 및 선호 사항에 대한 프로파일링을 수행하여 귀하에게 판촉 메시지, 마케팅, 광고 및 기타 정보를 보냅니다. 우리는 당신에게 관심이 있을 것이라고 생각합니다.***

89 Miller (2019). Mittelstadt and Wachter (2019).

The data subject would benefit from more precise and relevant information, especially when important decisions are at stake. In

particular, with regard to complex AI systems, the possibility of providing modular information should be explored, i.e., providing bullet points that laypeople can understand, with links to access more detailed information possibly covering technical aspects.

데이터 주제는 특히 중요한 결정이 내려 질 때보다 정확하고 관련성있는 정보를 활용할 수 있습니다. 특히, 복잡한 AI 시스템과 관련하여 모듈 식 정보를 제공할 수 있는 가능성, 즉 평신도가 이해할 수 있는 총알 점을 제공하고 기술적인 측면을 포괄하는 보다 자세한 정보에 액세스할 수 있는 링크를 제공해야 합니다.

However, it is unlikely that the information that is provided to the general public will be sufficient to gain an understanding that is sufficient for identifying potential problems, dysfunctions, unfairness. This would assume access to the algorithmic model, or at least the possibly of subjecting it to extensive testing, and in the case of machine learning approaches, access to the system's training set.

그러나 일반 대중에게 제공되는 정보가 잠재적 문제, 기능 장애, 불공평을 식별하기에 충분한 이해를 얻기에 충분하지는 않습니다. 이것은 알고리즘 모델에 대한 접근, 또는 적어도 광범위한 모델에 적용되고 머신러닝 접근의 경우 시스템의 훈련 세트에 대한 접근

을 가정할 수 있습니다.

It has been argued that it would important to enable citizen to engage in 'black box tinkering', i.e., on a limited reverse-engineering exercise that consists in submitting test cases to a system and analysing the system's responses to detect faults and biases.90 This approach, which involves a distributed and non-systematic attempt at sensitivity analysis, has the advantage of democratising controls but is likely to have a limited success given the complexity of AI applications and the limitations on access to them.

시민이 '블랙 박스 땜질 (black box tinkering)'에 참여할 수 있게 하는 것이 중요하다고 주장했다. 즉, 테스트 케이스를 시스템에 제출하고 결함과 편견을 탐지하기 위해 시스템의 응답을 분석하는 제한된 리버스 엔지니어링 운동에 참여할 수 있다고 주장했다.90 감도 분석을 위한 분산적이고 비 체계적인 시도가 포함 된 이 접근 방식은 제어 민주화의 이점을 가지지 만 AI 응용 프로그램의 복잡성과 액세스에 대한 제한으로 인해 성공이 제한될 수 있습니다.

3.5. AI and data subjects' rights AI 및 데이터 주체의 권리

AI is relevant to distinct data protection rights. The GDPR expressly refers to profiling and automated decision-making in connection with the rights to access and the right to object, but AI also raises specific issues relative to other rights such as in particular, the rights to erasure and portability.

AI는 고유한 데이터 보호 권한과 관련이 있습니다. GDPR은 접근 권한 및 이의 제기 권리와 관련하여 프로파일링 및 자동화된 의사 결정을 명시적으로 언급하지만 AI는 특히 삭제 및 이식 권한과 같은 다른 권한과 관련하여 특정 문제를 제기합니다.

3.5.1. Article 15 GDPR: The right to access 접근할 권리

A key aspect of transparency (and consequently of accountability) consist in the data subjects' rights to access information about the processing of their data. Data subjects, according to Article 15 GDPR, have

투명성 (및 그에 따른 책임)의 주요 측면은 데이터 처리에 관한 정보에 액세스할 수 있는 데이터 주체의 권리로 구성됩니다. 제15조 GDPR에 따르면 데이터 주체는

the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and [...] information' [about their processing].

개인 정보가 처리되고 있는지의 여부, 그리고 그 경우에 개인 정보에 대한 액세스 및 [...] 처리 정보에 대한 정보를 컨트롤러로부터 확인을 받을 권리.

Article 15(1)(f) specifically addresses automated decision-making, requiring the controller to provide, when requested by the data subject, the same information that should have been provided before starting the processing according to 13(2)(f) and 14(2)(g). The mandatory information concerns

제 15 (1) (f) 조는 구체적으로 자동화된 의사 결정을 다루고 있으며, 데이터 주체가 요청할 경우 13 (2) (f)에 따라 처리를 시작하기 전에 제공해야하는 동일한 정보를 컨트롤러가 제공하도록 컨트롤러에 요구합니다. 14 (2) (g). 필수 정보 문제

the existence of automated decision-making' and

'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

자동화된 의사 결정의 존재 '및'관련 논리에 대한 의미있는 정보, 데이터 주체에 대한 그러한 처리의 의미 및 예상 결과.

The right to access information is also addressed in Recital 63. The recital first states that the right of access includes the data subject's right to know

정보에 접근할 수 있는 권리는 Recital 63에서도 다루고 있습니다. 리사이틀은 우선 접근할 수 있는 권리에는 데이터 주체가 알아야 할 권리가 포함되어 있다고 명시하고 있습니다

where possible [...] the logic involved in any automatic personal data processing and, at least when based on profiling, the consequences of such processing.

가능한 경우 자동 개인 데이터 처리와 관련된 논리, 그리고 적어도 프로파일링을 기반으로 할 때 그러한 처리의 결과.

90 Perel and Elkin-Koren (2017).

The scope of the right to access, or the ways of implementing it are limited by the requirement that but it

액세스 권한의 범위 또는 구현 방법은 다음과 같은 요구 사항에 의해 제한됩니다.

> ***should not adversely affect the rights or freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software.***
>
> ***영업 비밀 또는 지적 재산권, 특히 소프트웨어를 보호하는 저작권을 포함하여 타인의 권리 또는 자유에 부정적인 영향을 미치지 않아야 합니다.***

This limitation, however, should not entail a complete denial of the right to information:

그러나이 제한으로 인해 정보에 대한 권리가 완전히 거부되는 것은 아닙니다.

[T]he result of these considerations should not be a refusal to provide all information to the data subject. Where the controller processes a large quantity of information concerning the data subject, the controller should be able to request that, before the information is delivered, the data subject specify the information or processing activities to which the request relates'

그러한 고려 사항의 결과는 모든 정보를 데이터 주체에게 제공하는 것을 거부해서는 안된다. 컨트롤러가 데이터 주체에 관한 대량의 정보를 처리하는 경우, 컨트롤러는 정보가 전달되기 전에 데이터 주체가 요청과 관련된 정보 또는 처리 활동을 지정하도록 요청할 수 있어야 합니다. '

There has been a wide discussion on whether Article 15 should be read as granting data subjects the right to obtain an individualised explanation of automated assessments and decisions.91

자동화된 평가 및 결정에 대한 개별화된 설명을 얻을 수 있는 권리를 데이터 주체에게 부여하는 것으로 제15 조를 읽어야 하는지에 대한 광범위한 논의가 있었다.91

Unfortunately, the formulation of Article 15 is very ambiguous, and that ambiguity is reflected in Recital 63. In particular it is not specified whether the obligation to provide information on the 'logic involved' only concerns providing general information on the methods adopted in the system, or rather specific information on how these methods where applied to the data subject (i.e., an individual explanation, as we shall see in Section 3.6.5).

불행하게도, 제15 조의 공식은 매우 모호하며, 모호성은 Recital 63에 반영된다. 특히 '논리 관련'에 관한 정보를 제공할 의무가 시스템에 채택된 방법에 대한 일반적인 정보를 제공하는 것에 만 관련되는지 여부는 명시되어 있지 않다 또는 데이터 주체에 이러한 방법이 어떻게 적용되는지에 대한 구체적인 정보 (즉, 섹션 3.6.5에서 볼 수 있는 개별 설명).

3.5.2. Article 17 GDPR: The right to erasure

The right to erasure (or to be forgotten) consists in the data subjects' right to 'obtain from the controller the erasure of personal data concerning him or her without undue delay ', when the

conditions for lawful processing no longer obtain (such conditions are forth in Article 17 (1)). An issue may concern whether even inferred personal data or also inferred group data (such as a trained algorithmic model) should be deleted as a consequence of the obligation to erase the collected personal data that have enabled such inferences to be drawn. The answer seems positive in the first case and negative in the second, since the data that are embedded in an algorithmic model are no longer personal. However, erasing the data used for constructing an algorithmic model, may make it difficult or impossible to demonstrate the correctness of that model.

합법적인 처리 조건이 더 이상 얻지 않을 때 (지우지 않고 지체할 권한) 데이터 주체는 '지체없이 지체없이 개인 정보를 지워야' 하는 데이터 주체의 권리로 구성됩니다. 제17 조제 1 항에 따른다. 이러한 추론을 도출할 수 있는 수집된 개인 데이터를 지워야 하는 의무의 결과로 유추된 개인 데이터 또는 유추된 그룹 데이터 (예 : 훈련된 알고리즘 모델)도 삭제해야 하는지에 대한 문제가 발생할 수 있습니다. 알고리즘 모델에 포함된 데이터가 더 이상 개인 정보가 아니기 때문에 첫 번째 경우에는 긍정적인 것으로 보이며 두 번째 경우에는 부정적인 것으로 보입니다. 그러나 알고리즘 모델을 구성하는 데 사용된 데이터를 지우면 해당 모델의 정확성을 입증하기가 어렵거나 불가능할 수 있습니다.

3.5.3. Article 19 GDPR: The right to portability

The data subject has the 'right to receive the personal data concerning him or her, which he or she has provided to a controller in a structured, commonly used and machine-readable format' and 'to transfer the data to other controller'. This right only applies when the processing is based on consent. Thus, the right to portability has a smaller scope that the right to access, which applies to all processing personal data, regardless of the applicable legal basis.

데이터 주체는 '자신에 관한 개인 데이터를 수신할 권리가 있으며, 이 데이터는 구조적이고 일반적으로 사용되며 기계가 읽을 수 있는 형식으로 컨트롤러에 제공한 데이터와 다른 컨트롤러로 데이터를 전송합니다. 이 권한은 처리가 동의에 기반한 경우에만 적용됩니다. 따라서, 이식성에 대한 권리는 접근할 수 있는 권리의 범위가 더 작으며, 이는 적용 가능한 법적 근거에 관계없이 모든 처리 개인 데이터에 적용됩니다.

It is not easy to determine the scope of this right with regard to AI-based processing. First, it needs to be determined whether the data 'provided' by the data subject only concern the data entered

by the data subject (e.g., keying his or her particulars) or also the data collected by the system when tracking the data subject's activity. Second, it is to be determined whether the right also concerns the data inferred from the collected data about the data subject. A clarification would be useful in this regard.

AI 기반 처리와 관련하여 이 권한의 범위를 결정하는 것은 쉽지 않습니다. 먼저, 데이터 주체에 의해 '제공된'데이터가 데이터 주체에 의해 입력된 데이터 (예를 들어, 자신의 특정 사항을 키잉) 또는 데이터 주체의 활동을 추적할 때 시스템에 의해 수집된 데이터에만 관련되는지 여부를 결정해야 한다. 둘째, 권리가 또한 데이터 주체에 관한 수집된 데이터로부터 추론된 데이터에 관한 것인지의 여부가 결정되어야 한다. 이와 관련하여 설명이 유용할 것입니다.

3.5.4. Article 21 (1): The right to object 이의 제기 권리

The right to object enables data subjects to request (and obtain) that the processing of their data be terminated. This right can be exercised under the following conditions:

이의를 제기할 수 있는 권리는 데이터 주체가 자신의 데이터 처

리를 종료하도록 요청 (및 획득)할 수 있게 합니다. 이 권한은 다음 조건에서 행사할 수 있습니다.

91 Wachter et al (2016), Edwards and Veale (2019).

1. The data subject has grounds relating to his or her particular situation that support the request.

   데이터 주체는 요청을 지원하는 특정 상황과 관련된 근거를 가지고 있습니다.

2. The processing is based on the legal basis of Article 6 (3)(e), i.e. necessity of the processing for performing a public task in the public interest or for the exercise of legitimate authority, or on the legal basis of Article 6 (3)(f), i.e., necessity of the processing for purposes of the legitimate interests pursued by the controller or by a third party.

   처리는 제6조 (3) (e)의 법적 근거, 즉 공공의 이익을 위해 공공 업무를 수행하거나 합법적인 권한을 행사하기위한 처리의 필요성 또는 제6조 (3)의 법적 근거를 기초로 한다.) (f), 즉 지배인이나 제3자가 추구하는 정당한 이익을 목적으로 하는 처리의 필요성.

3. The controller fails to demonstrate compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject.

   컨트롤러는 데이터 주체의 이익, 권리 및 자유를 무시하는 처리에 대한 강력한 근거를 입증하지 못합니다.

If all these conditions are satisfied, the controller has the obligation to terminate the processing.

이러한 모든 조건이 충족되면 컨트롤러는 처리를 종료할 의무가 있습니다.

The right to object is particularly significant with regard to profiling, since it seems that only in very special cases the controller may have overriding compelling legitimate grounds for continuing to profile a data subject which objects to the profiling on personal grounds.

이의 제기 권은 프로파일링 과 관련하여 특히 중요합니다. 매우 특별한 경우에만 컨트롤러가 개인적인 근거로 프로파일링에 반대하는 데이터 주제를 계속 프로파일링 하기위한 강력한 합법적인 근거를 무시할 수 있는 것 같습니다.

The right to object does not apply to a processing that is based on the data subject's consent, since in this case the data subject can impede the continuation of the processing just by withdrawing consent (according to Article 7 (3) GDPR).

데이터 주체의 동의에 근거한 처리에는 이의를 제기할 수 있는 권리가 적용되지 않습니다. 이 경우, 데이터 주체는 동의를 철회하는 것만으로 처리의 계속을 방해할 수 있기 때문입니다 (제7조 (3) GDPR에 따름).

The GDPR, in regulating the right to object, explicitly refers to profiling, and introduces special norms concerning direct marketing and statistical processings. Such provisions are relevant to AI, given that profiling and statistics are indeed key applications of AI to personal data.

이의 제기권을 규제함에 있어 GDPR은 명시적으로 프로파일링을 말하며 직접 마케팅 및 통계 처리에 관한 특별 규범을 도입합니다. 프로파일링 및 통계가 실제로 AI를 개인 데이터에 적용하는 주요 응용 프로그램이라는 점을 감안할 때 이러한 조항은 AI와 관련이 있습니다.

3.5.5. Article 21 (1) and (2): Objecting to profiling and direct marketing 프로파일링 및 다이렉트 마케팅 반대

Article 21 (1) specifies that the right to object also applies to profiling:

제 21조 제1 항은 이의 제기 권리가 프로파일링에도 적용됨을 명시하고있다.

> ***The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions.***
>
> ***데이터 주체는 자신의 특정 상황과 관련하여 언제든지 제 6조 (1)의 (e) 또는 (f)에 근거한 개인 데이터의 처리에 대해 이의를 제기할 권리가 있습니다. 해당 규정에 따른 프로파일링을 포함합니다.***

Profiling in the context of direct marketing is addressed in Article 21 (2), which recognises an unconditioned right to object:

직접 마케팅의 맥락에서 프로파일링은 제21조 제2 항에 언급되어 있으며, 이는 조건에 반대할 권리를 인정한다.

> ***Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing.***
>
> ***개인 데이터가 직접 마케팅 목적으로 처리되는 경우, 데이터 주체는 언제든지 해당 마케팅에 대한 개인 데이터의 처리에 대해 이의를 제기할 권리가 있으며, 여기에는 해당 직접 마케팅과 관련된 정도까지의 프로파일링 이 포함됩니다.***

This means that the data subject does not need to invoke specific grounds when objecting to processing for direct marketing purposes, and that such purposes cannot be 'compelling legitimate grounds for the processing which override the interests, rights and

freedoms of the data subject'.

이는 데이터 주체가 직접 마케팅 목적으로 처리에 반대할 때 특정 근거를 요구할 필요가 없으며, 그러한 목적이 '데이터 주체의 이익, 권리 및 자유를 무시하는 처리에 대한 정당한 근거를 강구할 수 없음'을 의미합니다.

Given the importance of profiling for marketing purposes, the unconditional right to object to such processing is particularly significant for the self-protection of data subjects. Controllers should be required to provide easy, intuitive and standardised ways to facilitate the exercise of this right.

마케팅 목적으로 프로파일링의 중요성을 고려할 때, 이러한 처리에 반대하는 무조건적인 권리는 데이터 주체의 자체 보호에 특히 중요합니다. 컨트롤러는 이 권한을 쉽게 발휘할 수 있는 쉽고 직관적이며 표준화된 방법을 제공해야 합니다.

3.5.6. Article 21 (2). Objecting to processing for research and statistical purposes 연구 및 통계 목적의 처리에 반대

The right to object also applies to processing for scientific or historical research purposes and for statistical purposes. In such cases, the objection concerns the inclusion of the data subject information in the input data for the processings at stake (as the result of research and statistics cannot consist in personal data).

반대할 권리는 과학적 또는 역사적 연구 목적 및 통계적 목적을 위한 처리에도 적용됩니다. 그러한 경우, 이의 제기는 처리중인 처리를 위해 입력 데이터에 데이터 주제 정보를 포함시키는 것과 관련이 있습니다 (연구 및 통계 결과는 개인 데이터로 구성될 수 없음).

The right to object does not apply when the processing is carried out for reasons of public interest (it therefore applies, a contrario, when the processing is aimed at private commercial purposes):

처리가 공익상의 이유로 수행되는 경우 이의 제기 권리가 적용되지 않습니다 (따라서 처리가 개인의 상업적 목적을 목표로 하는 경우에는 반대).

> ***Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the***

right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

개인 정보가 제89 조제 1 항에 따라 과학적 또는 역사적 연구 목적 또는 통계적 목적으로 처리되는 경우, 데이터 주체는 자신의 특정한 상황과 관련하여 그 또는 그에 관한 개인 정보의 처리에 반대할 권리를 갖습니다. 그녀는 공공의 이익을 위해 수행된 작업 수행에 처리가 필요하지 않은 한.

further limitation is introduced by Article 17(3)(d), which limits the right to erasure when its exercise would make it impossible or would seriously undercut the ability to achieve the objectives of the processing for archiving, research or statistical purposes. This limitation would probably find limited application to big data, since the exclusion of a single records from the processing would likely have little impact on the system's training or, at any rate, on the definition of its algorithmic model.

추가 제한은 제17 (3) (d) 조에 의해 도입되며, 이로 인해 운동으로 인해 보관, 연구 또는 통계 목적으로 처리의 목표를 달성할 수 있

는 능력을 심각하게 약화시킬 수 있는 권리를 제한할 수 있습니다. 처리에서 단일 레코드를 제외하면 시스템 교육에 영향을 미치지 않거나 알고리즘 모델의 정의에 영향을 미치지 않기 때문에 이 제한은 아마도 빅 데이터에 대한적용이 제한적일 것입니다.

3.6. Automated decision-making

Article 22, which deals with automated decision-making, is most relevant to AI. As we shall see in what follows, this provision combines a general prohibition on automated decision-making, with broad exceptions.

자동화된 의사 결정을 다루는 제22 조는 AI와 가장 관련이 있습니다. 다음 내용에서 알 수 있듯이, 이 조항은 자동화된 의사 결정에 대한 일반적인 금지와 광범위한 예외를 결합합니다.

3.6.1. Article 22(1) GDPR: The prohibition of automated decisions

The first paragraph of Article 22 provides for a general right not to be subject to completely automated decisions significantly

affecting the data subject:

제 22 조의 첫 번째 단락은 데이터 주제에 중대한 영향을 미치는 완전히 자동화된 결정의 대상이 되지 않을 일반적인 권리를 규정합니다.

> ***The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.***
>
> ***데이터 주체는 프로파일링을 포함하여 자동화된 처리만을 기반으로 한 결정에 종속되지 않을 권리가 있으며, 이는 프로파일링을 포함하여 법적 영향을 미치거나 마찬가지로 자신에게도 큰 영향을 미칩니다.***

Even though this provision refers to a right, it does not provide for a right to object to automated decision-making, namely, it does not assume that automated decision-making is in general permissible as long as the data subject does not object to it. It rather introduces a prohibition upon controllers: automated

decisions affecting data subjects are prohibited, unless they fit in one of the exceptions provided in paragraph 2.92 According to the Article 29 Working Party:

이 조항은 권리를 언급하지만, 자동화된 의사 결정에 반대할 권리를 제공하지 않습니다. 즉, 데이터 주체가 반대하지 않는 한 자동화된 의사 결정이 일반적으로 허용되는 것으로 가정하지 않습니다. 제29조 작업반에 따르면, 2.92 항에 규정된 예외 중 하나에 해당하지 않는 한, 데이터 주체에 영향을 미치는 자동 결정은 금지된다.

> ***as a rule, there is a general prohibition on fully automated individual decision-making, including profiling that has a legal or similarly significant effect.93***
>
> ***일반적으로 법적 또는 유사하게 중요한 영향을 미치는 프로파일링을 포함하여 완전히 자동화된 개별 의사 결정에는 일반적으로 금지됩니다.93***

For the application of the prohibition established by Article 22(1), four conditions are needed: a decision must be taken, (2) it must be solely based on automated processing, (3) it must include

profiling, (4) it must have legal or anyway significant effect.

제 22 (1) 조에 의해 금지된 금지를 적용하기 위해서는 4 가지 조건이 필요하다 : 결정이 내려져야 한다. (2) 자동화된 처리만을 기반으로 해야 한다. (3) 프로파일링을 포함해야 한다. 법적 또는 어쨌든 중대한 영향을 미칩니다.

The first condition requires that a stance be taken toward a person, and that this stance is likely to be acted upon (as when assigning a credit score).

첫 번째 조건은 사람을 향한 자세를 취해야 하며, 이 자세는 신용 점수를 지정할 때와 같이 행동해야 합니다.

The second condition requires that humans do not exercise any real influence on the outcome of a decision-making process, even though the final decision is formally ascribed to a person. This condition is not satisfied when the system is only used as a decision-support tool for human beings, who are responsible for the decision, deliberate on the merit of each case, and autonomously decide whether to accept or reject the system's suggestions.94

두 번째 조건은 최종 결정이 공식적으로 사람에게 속하더라도 인간이 의사 결정 과정의 결과에 실질적인 영향을 미치지 않도록 요구합니다. 시스템이 의사 결정을 담당하고 각 사례의 장점을 고의적으로 고려하고 시스템 제안을 수락할지 거부할지 자율적으로 결정하는 사람을 위한 의사 결정 지원 도구로만 사용되는 경우에는 이 조건이 충족되지 않습니다.94

92 Mendoza and Bygrave (2017).
93 Article 29, WP251/2017 last revised 2018, 19.

The third condition requires that the automated processing determining the decision includes profiling. A different interpretation could be suggested by the comma that separates 'processing' and 'including profiling' in Article 22(1), which seems to indicate that profiling only is an optional component of the kind of automated decisions that are in principle prohibited by Article 22(1). However, the first interpretation (the necessity of profiling) is confirmed by Recital (71), according to which the processing at stake in the regulation of automated decision must include profiling:

세 번째 조건은 결정을 결정하는 자동화된 처리에 프로파일링 이 포함되어야 합니다. 제22 (1) 조에서 '프로세싱'과 '프로파일링 포

함'을 구분하는 다른 해석이 쉼표로 제안될 수 있는데, 이는 프로파일링 이 기본적으로 조항에 의해 금지되는 자동화된 결정의 선택적 구성 요소임을 나타냅니다. 22 (1). 그러나 첫 번째 해석 (프로파일링의 필요성)은 Recital (71)에 의해 확인되며, 자동 결정의 규제에 처한 처리에는 프로파일링 이 포함되어야 합니다.

> ***Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements.***
>
> ***이러한 처리에는 자연인과 관련된 개인의 측면을 평가하는, 특히 업무, 경제 상황, 건강, 개인의 선호도에 대한 데이터 주체의 성능에 관한 측면을 분석하거나 예측하기위한 모든 형태의 개인 데이터 자동 처리로 구성된 '프로파일링 '이 포함됩니다. 관심사, 신뢰성 또는 행동, 위치 또는 움직임.***

The fourth condition requires that the decision

네 번째 조건은 결정을 요구합니다

> ***produces legal effects concerning [the data subject] or similarly significantly affects him or her.***
>
> ***[데이터 주체]와 관련하여 법적 효력을 발생시키거나 유사하게 그 또는 그녀에게 상당한 영향을 미칩니다.***

Recital (71) mentions the following examples of decision having significant effects: the 'automatic refusal of an online credit application or e-recruiting practices'.95 It has been argued that such effects cannot be merely emotional, and that usually they are not caused by targeted advertising, unless 'advertising involves blatantly unfair discrimination in the form of web-lining and the discrimination has non-trivial economic consequences (e.g., the data subject must pay a substantially higher price for goods or services than other persons).' 96

Recital (71)은 다음과 같은 중대한 영향을 미치는 결정의 예를 언급합니다 : '온라인 신용 신청 또는 전자 모집 관행의 자동 거부'. 그러한 영향은 단순히 감정적 일 수 없으며 일반적으로 발생하지 않는다고 주장되었습니다. '광고가 웹 라이닝의 형태로 뻔뻔한 차

별을 수반하고 차별이 사소한 경제적 결과를 초래하지 않는 한 (예를 들어, 데이터 주체는 다른 사람보다 상품이나 서비스에 대해 실질적으로 더 높은 가격을 지불해야 한다 ")를 제외하고는 표적 광고에 의해. 96

Many decisions made today by AI systems fall under the scope of Article 21(1), as AI algorithms are increasingly deployed in recruitment, lending, access to insurance, health services, social security, education, etc. The use of AI makes it more likely that a decision will be based 'solely' on automated processing. This is due to the fact that humans may not have access to all the information that is used by AI systems, and may not have the ability to analyse and review the way in which this information is used. It may be impossible, or it may take an excessive effort to carry out an effective review – unless the system has been effectively engineered for transparency, which in some cases may be beyond the state of the art. Thus, especially when a large-scale opaque system is deployed, humans are likely to merely execute the automated suggestions by AI, even when they are formally in charge. Moreover, human intervention may be prevented by the costs-and-incentives structure in place: humans are likely not to substantially review automated decision, when the cost of

engaging in the review – from an individual or an institutional perspective– exceeds the significance of the decision (according to the decision-maker's perspective).

AI 알고리즘이 채용, 대출, 보험, 의료 서비스, 사회 보장, 교육 등에 점점 더 많이 보급됨에 따라 AI 시스템에 의해 오늘날 내려진 많은 결정은 제21 (1) 조의 범위에 속합니다. 결정은 자동화된 처리에만 '전적으로'기초할 것입니다. 이는 인간이 AI 시스템에서 사용하는 모든 정보에 액세스할 수 없고 이 정보가 사용되는 방식을 분석하고 검토할 수 있는 능력이 없기 때문입니다. 시스템이 투명성을 위해 효과적으로 설계되지 않은 경우가 아니면 불가능할 수도 있고 효과적인 검토를 수행하기 위해 과도한 노력이 필요할 수도 있습니다. 따라서, 특히 대규모 불투명 시스템이 배치될 때, 인간은 공식적으로 책임을 지고 있을 때조차도 AI에 의한 자동화된 제안을 실행하기 만합니다. 또한, 비용 및 인센티브 구조를 통해 인적 개입을 방지할 수 있습니다. 개인 또는 제도적 관점에서 검토에 참여하는 비용이 시스템의 중요성을 초과할 경우 자동 결정을 실질적으로 검토하지 않을 가능성이 높습니다. 결정 (의사 결정자의 관점에 따라).

3.6.2. Article 22(2) GDPR: Exceptions to the prohibition of 22(1)

22 (1) 금지에 대한 예외

Paragraph 2 of Article 22 provides for three broad exceptions to Paragraph 1. It states that the prohibition on automated decision-making does not apply when the processing upon which the decision is based

제22조 제2항은 제1 항에 대해 세 가지 광범위한 예외를 규정하고있다. 결정에 근거한 처리가 자동화된 의사 결정에 대한 금지는 적용되지 않는다고 명시되어 있다.

94 Article 29, WP251/2017 last revised 2018, 21-22.
95 For an analysis of legal effects and of similarly relevant effects, see Article 29, WP251/2017 last revised 2018,
96 Medoza and Bygrave (2017, 89).

(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;

데이터 주체와 데이터 컨트롤러 간의 계약 체결 또는 이행에 필요하다.

(b) is authorised by Union or Member State law to which the controller is subject, and which also

lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

개인 정보 처리자의 권리와 자유 및 정당한 이익을 보호하기위한 적절한 조치를 규정하고 있으며, 또한 관할권이 적용되는 연합 또는 회원국 법률에 의해 승인됩니다. 또는

(c) is based on the data subject's explicit consent.

데이터 주체의 명시적 동의에 근거합니다.

Based on the broad exception of item (a), automated decision-making is enabled in key areas such as recruitment and lending. However, for the exception to apply, decisions based solely on automated processing must be 'necessary.' Such necessity may depend on the high number of cases to be examined (e.g., a very high number of applications to a job). The necessity of using AI in decision-making may also be connected to AI capacities to outperform human judgement. In this connection we may wonder whether human involvement will still contribute to a stronger protection of data subjects, or whether the better performance of machines – even with regard to the political and legal values at

stake, e.g., ensuring 'fair equality of opportunity' for all applicants to position97 – will make human intervention redundant or dysfunctional. Outside of the domain of contract and legal authorisation, consent may provide a basis for automated decision-making according to Article 22(2)(c). However, the conditions for valid consent not always obtain, even in cases when automated decision-making seems appropriate. Consider for instance the case in which an NGO uses an automated method for classifying (profiling) applicants to determine their need and consequently allocate certain benefits to them. In such a case, it is very doubtful that an applicant's consent may be viewed as free (as not consenting would entail being excluded from the benefit), but the system seems socially acceptable and beneficial even so.

항목 (a)의 광범위한 예외를 기반으로 채용 및 대출과 같은 주요 영역에서 자동 의사 결정이 가능합니다. 그러나 예외가 적용 되려면 자동화된 처리만을 기반으로 하는 결정이 '필요'해야 합니다. 이러한 필요성은 조사될 많은 사례 (예를 들어, 업무에 대한 매우 많은 애플리케이션)에 의존할 수 있다. 의사 결정에 AI를 사용해야할 필요성은 AI 판단력과 연결되어 인간의 판단력을 능가할 수 있습니다. 이와 관련하여 우리는 인간의 참여가 여전히 데이터 주체를 더욱 강력하게 보호할 것인지, 또는 기계의 성능을 향상시키는 데 도움이 될지 궁금할 수 있습니다. 위치 97에 대한 모든 지

원자 – 사람의 개입이 불필요하거나 역기능을 하게 됩니다. 계약 및 법적 승인 영역을 벗어나서, 동의는 제22 (2) (c) 조에 따라 자동화된 의사 결정의 기초를 제공할 수 있습니다. 그러나 자동화된 의사 결정이적절 해 보이는 경우에도 유효한 동의 조건이 항상 확보되는 것은 아닙니다. 예를 들어 NGO가 신청자를 분류 (프로파일링)하기 위해 자동화된 방법을 사용하여 자신의 필요를 결정하고 결과적으로 그들에게 특정 혜택을 할당하는 경우를 고려하십시오. 그러한 경우, 신청자의 동의가 자유로 간주될 수 있다는 것은 의심의 여지가 있지만 (동의하지 않는 것이 이익에서 제외될 수 있기 때문에) 시스템은 사회적으로 수용 가능하고 유익한 것처럼 보입니다.

3.6.3. Article 22(3) GDPR: Safeguard measures

In the cases under Article 22(2)(a) and (c) – i.e. when the automated decision is necessary to contract or explicitly consented – Article 22(3) requires suitable safeguard measures:

제 22 (2) (a) 및 (c) 조의 경우 – 즉, 자동 결정이 계약 또는 명시적으로 동의해야 하는 경우 – 제22 (3) 조에는 적절한 보호 조치가 필요합니다.

the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

데이터 관리자는 데이터 주체의 권리와 자유 및 적법한 이익을 보호하기 위해 최소한 컨트롤러의 일부에 대한 사람의 개입을 얻거나 자신의 관점을 표현하고 결정에 이의를 제기할 수 있는 적절한 조치를 이행해야 합니다.

According to Article 29 Working Party, some of these measures concern risk reduction, Examples are quality assurance checks, algorithmic auditing, data minimisation, and anonymisation or pseudonymisation, and certification mechanisms.98 Such measures should ensure that the requirements set forth in Recital (71) – concerning acceptability, accuracy and reliability – are respected

제 29조 작업반에 따르면, 이러한 조치 중 일부는 위험 감소에 관한 것이며, 예는 품질 보증 점검, 알고리즘 감사, 데이터 최소화, 익명화 또는 가명 화 및 인증 메커니즘이다.98 이러한 조치는 Recital (71)) – 수용성, 정확성 및 신뢰성에 관한 – 존중

the controller should use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect

관리자는 프로파일링을 위해적절한 수학적 또는 통계적 절차를 사용해야 하며, 특히 개인 데이터의 부정확성을 초래하는 요소가 수정되고 오류의 위험이 최소화되고 개인 데이터를 안전한 방식으로 확보할 수 있도록 적절한 기술적 및 조직적 조치를 구현해야 합니다. 데이터 주체의 이익과 권리에 수반되는 잠재적 위험을 고려하고 특히 인종 또는 민족, 정치적 견해, 종교 또는 신념, 노동 조합 가입, 유전자에 근거한 자연인에 대한 차별적 영향을 방지 또는 건강

상태 또는 성적 취향, 또는 그러한 영향을 미치는 조치를 초래하는 결과

According to the Article 29 Working party, the input data must be shown to not be 'inaccurate or irrelevant, or taken out of context,' and to not violate 'the reasonable expectations of the data subjects', in relation to the purpose for which the data was collected.99

제 29조 작업반에 따르면, 입력 데이터는 '정확하지 않거나 관련이 없거나 문맥에서 벗어난'것이 아니며, 데이터 주체의 합리적인 기대치를 위반하지 않아야 한다. 데이터가 수집되었습니다.99

97 Rawls ([1971 1999, 63).
98 Article 29, WP251/2017 last revised 2018, 32
99 Article 29, WP251/2017 last revised 2018, 17

In approaches based on machine learning, this should apply not only to the data concerning the person involved in a particular decision, but also to the data in a training set, where the biases built into the training set may affect the learned algorithmic model,

and hence the accuracy the system's inferences.

머신러닝에 기반한 접근법에서, 이것은 특정 결정에 관련된 사람에 관한 데이터 뿐만 아니라 훈련 세트에 내장된 편향이 학습된 알고리즘 모델에 영향을 줄 수 있는 훈련 세트의 데이터에도 적용되어야 합니다. 따라서 시스템의 추론의 정확성.

Other measures pertain to the interaction with the data subjects, such the right to obtain human intervention and the right to challenge a decision. For instance, a link could be provided to 'an appeals process at the point the automated decision is delivered to the data subject, with agreed time scales for the review and a named contact point for any queries.'100 An appeals process is most significant with regard to AI applications, and especially when these applications are 'opaque', i.e., they are unable to provide human-understandable explanations and justifications.

다른 조치는 데이터 주체와의 상호 작용과 관련되며, 이는 사람의 개입을 얻을 권리 및 결정에 이의를 제기할 권리입니다. 예를 들어, '자동 결정이 데이터 주제에 전달되는 시점에 검토에 대한 합의된 시간 척도 및 모든 문의에 대해 지정된 연락처와 함께 이의 제기 프로세스에 대한 링크를 제공할 수 있습니다.'100 이의 제기 프로세스가 가장 중요합니다. AI 응용 프로그램과 관련하여, 특히

이러한 응용 프로그램이 '불투명'한 경우, 즉 사람이 이해할 수 있는 설명과 정당성을 제공할 수 없습니다.

3.6.4. Article 22(4) GDPR: Automated decision-making and sensitive data 자동화된 의사 결정 및 민감한 데이터

Article 22(4) introduces a prohibition, limited by an exception, to ground automated decisions on 'sensitive data', i.e., the special categories set out in Article 9(1):

제 22 (4) 조는 '민감한 데이터', 즉 제9 (1) 조에 규정된 특별 범주에 대한 자동화된 결정을 내리기 위해 예외로 제한되는 금지를 소개합니다.

> ***Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place***
>
> ***제 2 조에 언급된 결정은 제9 (2)조 (a) 또는 (g) 항이 적***

용되고 데이터 주체의 권리를 보호하기위한 적절한 조치가 아닌 한, 제9 (1) 조에 언급된 특정 범주의 개인 데이터에 근거하지 않아야 합니다. 자유와 정당한 이익이 있습니다

The exception concerns the cases in which the data subject has given explicit consent (Article 9(2)(a)) or processing is necessary for reason of public interest (Article 9(2)(g)). The role of the data subject's consent needs to be clarified since consent does not exclude that the method used for the decision is inacceptable (as when it is discriminatory).

데이터 주체가 명시적으로 동의한 경우 (제9 (2) (a) 항) 또는 공익상의 이유로 처리가 필요한 경우 (제9 (2) (g))는 예외입니다. 동의가 결정에 사용된 방법이 용납될 수 없다는 것을 배제하지 않기 때문에 데이터 주체의 동의의 역할을 분명히 해야 합니다 (차별적 일 때).

As noted above AI challenges the prohibition of processing sensitive data. First of all, sensitive data can be (probabilistically) inferred from non-sensitive data. For instance, sex orientation can be inferred from a data subject's Internet activity, likes or even facial features. In this case. the inference of sensitive data should count

as a processing of sensitive data, and therefore would have to be considered unlawful unless the conditions under Article 9 are met.

위에서 언급했듯이 AI는 민감한 데이터를 처리하는 것을 금지합니다. 우선, 중요하지 않은 데이터에서 민감한 데이터를 (비전문적으로) 추론할 수 있습니다. 예를 들어, 성 취향은 데이터 주체의 인터넷 활동, 좋아요 또는 얼굴 특징으로부터 추론될 수 있습니다. 이 경우 민감한 데이터의 추론은 민감한 데이터의 처리로 간주되므로 제9 조의 조건이 충족되지 않는 한 불법으로 간주되어야 합니다.

Secondly, non-sensitive data can work as proxies for sensitive data correlated to them, even though the latter are not inferred by the system. For instance, the place of residence can act as a proxy for ethnicity. In this case, an unlawful discrimination may take place.

둘째, 중요하지 않은 데이터는 시스템에 의해 추론되지 않더라도 민감한 데이터에 대한 프록시로 작동할 수 있습니다. 예를 들어, 거주지는 민족의 대리 역할을 할 수 있습니다. 이 경우 불법 차별이 발생할 수 있습니다.

3.6.5. A right to explanation?

To understand the GDPR ambiguous approach to the right to explanation we need to compare two provisions, Recital (71) and Article 22.

GDPR 설명에 대한 모호한 접근 방식을 이해하려면 Recital (71)과 22 조의 두 조항을 비교해야 합니다.

According to Recital (71), the safeguards to be provided to data subjects in case of automated decisions include all of the following:

Recital (71)에 따르면 자동화된 의사 결정의 경우 데이터 주체에게 제공되는 보호 조치에는 다음이 모두 포함됩니다.

- specific information
- the right to obtain human intervention,
- the right to express his or her point of view,
- the right to obtain an explanation of the decision reached after such assessment
- the right to challenge the decision.

- the right to obtain human intervention,
- the right to express his or her point of view,
- the right to challenge the decision.
- 특정 정보
- 인간의 개입을 얻을 권리,
- 자신의 관점을 표현할 권리
- 그러한 평가 후 도달한 결정에 대한 설명을 얻을 권리
- 결정에 이의를 제기할 권리.
- 인간의 개입을 얻을 권리,
- 자신의 관점을 표현할 권리
- 결정에 이의를 제기할 권리.

According to Article 22 the suitable safeguards to be provided include 'at least'

제 22 조에 따라 제공되는 적절한 보호 수단에는 '적어도'

Thus, two items are missing in article 22 relative to Recital (71): the provision of 'specific information' and the right to obtain an explanation of the decision reached after such assessment'. The first omission may not be very significative, since the obligation to provide information is already established by articles 13, 14 and 15 GDPR, as noted above, even though the requirement that the information be 'specific' is only spelled out in Recital (71). The second omission raises the issue of whether controllers are really required by law to provide an individualised explanation. Two interpretations are possible.

따라서 Recital (71)과 관련하여 제22 조에는 두 가지 항목, 즉 '특정 정보'의 제공과 그러한 평가 후 도달한 결정에 대한 설명을 얻을 수 있는 권리가 누락되어 있다. 정보를 제공할 의무는 이미 언급한 바와 같이 GDPR 13 조, 14조 및 15 조에 의해 이미 규정되어 있기 때문에 첫 번째 누락은 그다지 중요하지 않을 수 있습니다.). 두 번째 생략은 법률에 따라 컨트롤러가 개별화된 설명을 제공해야 하는지의 문제를 제기합니다. 두 가지 해석이 가능합니다.

According to the first one, the European legislator, by only including the request for specific explanation in the recitals and omitting it from the articles of the GDPR, intended to convey a double message: to exclude an enforceable legal obligation to provide individual explanations, while recommending that data controllers provide such explanations when convenient, according to their discretionary determinations. Following this interpretation, providing individualised explanation would only be a good practice, and not a legally enforceable requirement.

첫 번째에 따르면 유럽 의회 의원은 리사이틀에 구체적인 설명 요청을 포함시키고 GDPR 조항에서 생략함으로써 이중 메시지를 전달하기 위해 다음과 같은 두 가지 메시지를 전달하고자 한다. 데이터 컨트롤러는 재량에 따라 편리할 때 이러한 설명을 제공할 것을 권장합니다. 이러한 해석에 이어 개별화된 설명을 제공하는 것은 좋은 관행 일 뿐이며 법적으로 시행 가능한 요건은 아닙니다.

According to the second interpretation, the European legislator intended on the contrary to establish an enforceable legal obligation to provide individual explanation, though without

unduly burdening controllers. This interpretation is hinted at by the qualifier 'at least', which precedes the reference made to a 'right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.' The qualifier seems to suggest that some providers are legally required to adopt further safeguards, possibly including individualised explanations, as indicated in Recital 71. On this second approach, an explanation would be legally needed, whenever it is practically possible, i.e., whenever it is compatible with technologies, costs, and business practices.

두 번째 해석에 따르면, 유럽 의회 의원은 과도하게 부담을 가하는 통제자없이 개인의 설명을 제공할 수 있는 법적 의무를 수립하려는 의도를 가지고 있지 않았다. 이 해석은 '적어도'한정자에 의해 암시되는데, 이는 '제어기 측에서 사람의 개입을 얻을 권리, 자신의 견해를 표현하고 결정에 이의를 제기할 권리'에 대한 언급보다 우선합니다. 한정자는 Recital 71에 표시된 대로 일부 제공업체는 개별화된 설명을 포함하여 추가 보호 조치를 법적으로 채택해야 한다고 제안하는 것 같습니다. 이 두 번째 방법에서는 실제로 가능할 때마다, 즉 가능할 때마다 설명이 필요합니다. 기술, 비용 및 비즈니스 관행과 호환됩니다.

Both readings of these provisions – the combination of Article 13,

14, 15 and 22 – seems possible. The reason for this ambiguous language is likely to be that the legislator was unsure as to whether individualised explanations should be made into a legal requirement. As noted by some commentators, the view that data subjects have a right to individualised explanations under the GDPR may in the future be endorsed by data protection authorities and courts, perhaps viewing individualised explanation as a precondition for the data subjects' ability to effectively contest automated decisions.

이 조항들에 대한 제13 조, 제14 조, 제15조 및 제22 조의 조합을 읽는 것이 가능해 보입니다. 이 모호한 언어의 이유는 입법자가 개별 설명을 법적 요구 사항으로 만들어야 하는지 확실하지 않기 때문일 수 있습니다. 일부 의견가들에 의해 언급된 바와 같이, 데이터 주체가 GDPR에 따라 개별화된 설명에 대한 권리를 가지고 있다는 견해는 미래에 데이터 보호 당국과 법원에 의해 승인될 수 있으며, 아마도 개별화된 설명을 데이터 주체가 효과적으로 자동으로 경쟁하는 능력의 전제 조건으로 간주할 수 있습니다 결정.

A broad reading of Article 22(3), according to which an explanation is required to contest a decision, would strengthen the right to contest. In this case, the

argument for a right to explanation of specific decisions could be further buttressed by drawing on the rights to fair trial and effective remedy enshrined in Articles 6 and 13 of the European Convention on Human Rights.101

결정에 이의를 제기하기 위해 설명이 필요한 22조 3 항을 광범위하게 읽으면 이의를 제기할 권리가 강화될 것입니다. 이 경우, 유럽 인권 협약 제6 조와 13 조에 규정된 공정한 재판 및 효과적인 구제권에 대한 권리를 바탕으로 특정 결정에 대한 설명 권에 대한 주장은 더욱 강화될 수 있다.101

However, we should be cautioned against overemphasising a right to individualised explanations as a general remedy to the biases, malfunctions, and inappropriate applications of AI and big data technologies.102 A parallel may be drawn between consent and individualised explanation, as both rely on the data subject's informed initiative. It has often been observed that consent provides no effective protection, given the disparity in knowledge and power between controllers and data subjects, and also the limited time and energy available to the latter, and their inability to pool their interests and resources and coordinate their activities.

그러나 AI와 빅 데이터 기술의 편견, 오작동 및 부적절한적용에 대한 일반적인 구제책으로서 개별화된 설명에 대한 권리를 지나치게 강조하지 않도록 주의해야 한다.102 동의와 개별화된 설명 사이에 유사점이 있을 수 있다. 데이터 주체의 정보에 입각한 주도권. 컨트롤러와 데이터 주체 사이의 지식과 능력의 불균형, 후자에게 이용 가능한 시간과 에너지의 제한, 그리고 관심과 자원을 모으고 활동을 조정할 수 없다는 점에서 동의가 효과적인 보호를 제공하지 않는 경우가 종종 있었습니다..

101 Wachter et al (2016).
102 Edwards and Veal (2019).

The same may also apply to the right to an explanation, which is likely to remain underused by the data subjects, given that they may lack a sufficient understanding of technologies and applicable normative standards. Moreover, even when an explanation elicits potential defects, the data subjects may be unable to obtain a new, more satisfactory decision.

기술에 대한 이해와 적용 가능한 규범적 표준에 대한 충분한 이해가 부족한 점을 감안할 때, 설명에 대한 권리에도 동일하게 적용될 수 있다. 더욱이, 설명이 잠재적 결함을 유발하더라도 데이터 주체는 새롭고 더 만족스러운 결정을 얻지 못할 수 있습니다.

3.6.6. What rights to information and explanation?. 정보와 설명에 대한 권리는 무엇입니까?

Our analysis of the right to information and explanation to data subject end up with puzzling results.

데이터 주제에 대한 정보 및 설명에 대한 우리의 분석은 수수께끼의 결과로 끝납니다.

Let us summarise the main references in the GDPR:

GDPR의 주요 참고 자료를 요약 해 보겠습니다.

- According to Article 13 and 14 (on the right to information and Article 15 (on the right to access), the controller should provide information on 'the existence of automated decision-making, including profiling, referred to in Article 22(1)' and 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'.

  제13조 및 제14조 (정보 권 및 제15조 (액세스권))에 따

라, 컨트롤러는 '제 22조 제1 항에 언급된 프로파일링을 포함한 자동화된 의사 결정의 존재에 관한 정보를 제공해야 한다. '및'데이터 주제에 대한 이러한 처리의 의미 및 예상 결과 뿐만 아니라 관련된 논리에 대한 의미있는 정보 '.

- According to Article 22, the data subject has at least the right to obtain human intervention, the right to express his or her point of view, and the right to challenge the decision.

  제22조에 따르면, 데이터 주체는 최소한 사람의 개입을 받을 권리, 자신의 견해를 표현할 권리 및 결정에 이의를 제기할 권리가 있습니다.

- According to Recital (71), the data subject should also have the right to obtain an explanation of the decision reached after the assessment of his or her circumstances.

  Recital (71)에 따르면 데이터 주체는 자신의 상황을 평가한 후 도달한 결정에 대한 설명을 얻을 권리가 있어야 합니다.

We have also observed that according to the European Data Protection Board, controllers should provide data subject, in simple

ways, with the 'rationale behind or the criteria relied on in reaching the decision.' This information should be so comprehensive as to 'enable data subjects to understand the reasons for the decision.'103

또한 유럽 데이터 보호위원회 (European Data Protection Board)에 따르면, 컨트롤러는 간단한 방법으로 '이론의 근거 또는 결정에 도달하는 기준에 의존하는'간단한 방법으로 데이터 주제를 제공해야 한다는 것을 관찰했습니다. 이 정보는 '데이터 주체가 결정 이유를 이해할 수 있도록'포괄적이어야 합니다.103

Finally, Article 7(4)(a) of the Directive on Consumer Rights104 addresses information to be provided to consumers with regard to online offers, which often are based on profiling. It establishes that the supplier should indicate 'the main parameters determining ranking [...] of offers presented to the consumer' as well as 'the relative importance of those parameters as opposed to other parameters".

마지막으로, 소비자 권리에 관한 지침 104의 7 (4) (a) 조는 종종 프로파일링을 기반으로 하는 온라인 오퍼와 관련하여 소비자에게 제공되는 정보를 다룹니다. 공급 업체는 '소비자에게 제공되는 제안의 순위 [...]를 결정하는 주요 매개 변수'와 '다른 매개 변수와

대조적으로 해당 매개 변수의 상대적 중요성'을 표시해야 합니다.

Based on this set of norms, the obligation to provide information to the profiled data subject can take very different content:

이 규범에 따라 프로파일링된 데이터 주제에 정보를 제공할 의무는 매우 다른 내용을 취할 수 있습니다.

1. information on the existence of profiling, i.e., on the fact that the data subject will be profiled or is already being profiled;

   프로파일링의 존재에 관한 정보, 즉 데이터 주체가 프로파일링 되거나 이미 프로파일링 되고 있다는 사실에 관한 정보;

2. general information on the purposes of the profiling and decision-making;

   프로파일링 및 의사 결정의 목적에 대한 일반적인 정보;

3. general information on the kind of approach and technology that is adopted;

채택된 접근 및 기술의 종류에 대한 일반 정보

4. general information on what inputs factors (predictors) and outcomes (targets/predictions), of what categories are being considered;

   어떤 카테고리가 고려되고 있는지에 대한 입력 요소 (예측 자) 및 결과 (목표/예측)에 대한 일반적인 정보;

5. general information on the relative importance of such input factors in determining the outcomes;

   결과를 결정할 때 그러한 입력 요소의 상대적 중요성에 대한 일반적인 정보;

6. specific information on what data have been collected about the data subject and used for profiling him or her;

   데이터 주제에 대해 수집된 데이터에 대한 구체적인 정보.

7. specific information on what values for the features of the data subject determined the outcome concerning him or her;

   데이터 주체의 특징에 대한 어떤 가치가 그에 관한 결과를 결정했는지에 대한 특정 정보;

8. specific information on what data have been inferred about

the data subject;

데이터 주체에 대해 어떤 데이터가 추론되었는지에 대한 특정 정보;

9. specific information on the inference process through which certain values for the features of the data subject have determined a certain outcome concerning him or her.

   데이터 주체의 특징에 대한 특정 값이 그에 관한 특정 결과를 결정한 추론 프로세스에 대한 특정 정보.

103 Guidelines of the European Data Protection Board of 3 October 2017 on Automated individual decision-making and Profling, p. 25.
104 Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, as amended by Directive 2019/2161/EU of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules

In this list, items from (1) to (5) concern information ex ante, to be provided before the data are collected or anyway processed, while

items from (5) to (9) concern information to be provided ex post.

With regard to the ex-ante information, it is sure that the controller is required to provide the information under (1) and (2). Information under (3) may also be required, when the adopted technology makes a relevant difference (e.g., it may be inappropriate or lead to errors and biases). Information under (4) should also be provided, as a minimal account of the 'logic' of the processing, at least relative to the categories into which the input factors can be classified. This idea is explicitly adopted in the California Consumer Privacy Act, which at Section 1798.100 (b) requires controllers to 'inform consumers as to the categories of personal information to be collected.' We may wonder whether also some information under (5) should be provided, as an aspect of the information about the 'logic' of the processing, though it may not easy to determine in the abstract (without reference to a specific case) the importance of a certain input factor.

이 목록에서 (1)에서 (5)까지의 항목은 데이터가 수집되거나 어쨌든 처리되기 전에 제공되어야 하며, (5)에서 (9)의 항목은 사후에 제공되어야 합니다.

사전 정보와 관련하여 컨트롤러는 (1) 및 (2)에 정보를 제공해야 합니다. 채택된 기술이적절한 차이를 만들 때 (3)의 정보가 필요할 수도 있습니다 (예 : 부적절하거나 오류 및 편견으로 이어질 수 있음). (4)의 정보는 최소한 입력 요소가 분류될 수 있는 범주와 관련하여 처리의 '논리'에 대한 최소한의 설명으로 제공되어야 한다. 이 아이디어는 캘리포니아 소비자 프라이버시 법 (California Consumer Privacy Act)에 명시적으로 채택되었으며, 섹션 1798.100 (b)에서 컨트롤러는 '수집할 개인 정보의 범주에 대해 소비자에게 정보를 제공해야 합니다. 처리의 '논리'에 대한 정보의 한 측면으로서 (5)의 일부 정보를 제공해야 하는지 궁금할 수도 있지만, 초록 (특정 사례를 참조하지 않고)에서 중요성을 결정하는 것은 쉽지 않습니다. 특정 입력 계수.

With regard to the ex-post information, all data under (6) should be provided, as they are the object of the right to access. Information about (7) should also be provided, if we assume that there is right to individualised explanation. An individualised explanation may also require information about (8), when the intermediate conclusions by the system play a decisive role. Finally, information about (9) might also be provided, though information on (7) and (8) should generally be sufficient to provide adequate individualised explanations.

사후 정보와 관련하여 (6)의 모든 데이터는 액세스 권한의 대상이므로 제공해야 합니다. 개별 설명에 대한 권리가 있다고 가정하면 (7)에 대한 정보도 제공해야 합니다. 개별 설명은 시스템에 의한 중간 결론이 결정적인 역할을 할 때 (8)에 대한 정보를 요구할 수도 있습니다. 마지막으로 (9)에 대한 정보도 제공할 수 있지만 (7)과 (8)에 대한 정보는 일반적으로 적절한 개별 설명을 제공하기에 충분해야 합니다.

The information above needs to be complemented with further information in the case of decisions by public authorities, in which case also a reference to the norms being applied and the powers being exercised is needed, based on principles concerning the required justification for administrative acts.

위의 정보는 공공 당국의 결정에 따라 추가 정보로 보완되어야 하며, 이 경우 행정 행위에 대한 정당화에 대한 원칙에 근거하여 적용되는 규범 및 행사되는 권력에 대한 참조가 필요합니다.

Given the variety of ways in which automated decision-making can take place, it is hard to specify in precise and general terms what information should be provided. What information the controller may be reasonably required to deliver will indeed depend on the

importance of the decision, on the space of discretion that is being used, and on technological feasibility. However, it seems that data subjects who did not obtain the decision they hoped for should be provided with the specific information that most matters to them, namely, with the information on what values for their features determined in their case an unfavourable outcome. The relevant causal factors could possibly be identified by looking at the non - normal values that may explain the outcome. Consider for instance the case of person having an average income, and an ongoing mortgage to repay, whose application for an additional mortgage is rejected. Assume both of the following hypotheticals: (a) if the person had had a much higher income her application would have been accepted, regardless of her ongoing mortgage, and (b) if she had had no ongoing mortgage, her application would have been accepted, given her average income. Under such circumstances, we would say that it was the previous mortgage, rather than the average income, the key reason or cause explaining why the mortgage application was rejected, since it is what explains the departure from the standard outcome for such a case. 105

자동화된 의사 결정이 수행될 수 있는 다양한 방법을 고려할 때 어떤 정보를 제공해야 하는지 정확하고 일반적인 용어로 지정하기는 어렵습니다. 컨트롤러가 합리적으로 제공하기 위해 필요한

정보는 실제로 결정의 중요성, 사용되는 재량 공간 및 기술적 타당성에 달려 있습니다. 그러나, 그들이 원하는 결정을 얻지 못한 데이터 주체는 그들에게 가장 중요한 특정 정보, 즉 그들의 특징에 대한 가치가 그들의 경우에 불리한 결과로 결정된 정보와 함께 제공되어야 하는 것으로 보인다. 결과를 설명할 수 있는 비정규 값을 보면 관련 원인 요인을 식별할 수 있습니다. 예를 들어 평균 수입이 있는 사람과 상환할 모기지 상환의 경우 추가 모기지 신청이 거부되는 경우를 고려하십시오. 다음 가설을 모두 가정해 봅시다 : (a) 그 사람이 진행중인 모기지와 상관없이 신청이 훨씬 더 많은 수입을 받았다면, (b) 진행중인 모기지를 가지고 있지 않은 경우, 그녀의 신청이 수락되었을 것입니다 그녀의 평균 소득을 감안할 때. 그러한 상황에서 우리는 그것이 평균 소득이 아닌 이전의 모기지라고 말하고, 모기지 신청이 거부된 이유를 설명하는 주요 이유 또는 원인은 그러한 경우에 대한 표준 결과에서 벗어난 것을 설명하기 때문입니다. 105

3.7. AI and privacy by design 설계에 따른 AI 및 개인 정보

Two different legal perspective, complementary rather than incompatible, may inspire data protection law, a right based and a risk-based approach. Though the focus of the GDPR is on the right-based approach, there are abundant references to the risk prevention in the GDPR that can be used to address AI-related risks.106

양립할 수 없는 것이 아니라 보완적인 두 가지 법적 관점이 올바른 보호 및 위험 기반 접근 방식 인 데이터 보호법에 영향을 줄 수 있습니다. GDPR의 초점은 올바른 접근 방식에 있지만 AI 관련 위험을 처리하는 데 사용할 수 있는 GDPR의 위험 예방에 대한 참조가 풍부합니다.106

3.7.1. Right-based and risk-based approaches to data protection

데이터 보호에 대한 올바른 기반 및 위험 기반 접근 방식

The right-based approach to data protection, which underlies in

particular European law, views data protection as a matter of individual rights. These rights are organised in two layers. The top layer includes the fundamental rights to privacy and data protection, which are synergetic to other fundamental rights and principles: dignity, freedom of thought, conscience and religion, freedom of assembly and association, freedom to choose an occupation and right to engage in work, non-discrimination, etc. The lower tier is constituted by the data protection rights granted to individuals by the GDPR, such as the power to consent and withdraw consent (to processing not having other legal bases), the right to information, access, erasure, and the right to object. The focus is on the harm to individuals and on legal measure empowering their initiatives.

특히 유럽 법의 기초가되는 데이터 보호에 대한 올바른 접근 방식은 데이터 보호를 개인의 권리 문제로 간주합니다. 이러한 권리는 두 가지 계층으로 구성됩니다. 최상위 계층에는 개인 정보 보호 및 데이터 보호에 대한 기본 권리가 포함되며, 이는 개인의 존엄성, 사고의 자유, 양심과 종교, 집회 및 결사의 자유, 직업 선택의 자유 및 업무에 대한 권리와 같은 다른 기본 권리 및 원칙과 시너지 하위 계층은 GDPR이 개인에게 부여한 데이터 보호 권한, 예를 들어 동의 및 철회 권한 (다른 법적 근거가 없는 처리에 대한 권한), 정보, 액세스 권한, 소거 및 반대의 권리. 개인에 대한

피해와 그들의 주도권을 강화하는 법적 조치에 중점을 둡니다.

The risk-based approach, rather than granting individual entitlements, focuses on creating a sustainable ecology of information, where harm is prevented by appropriate organisational and technological measures. Data protection, when seen from the latter perspective appears to be as a risk-regulation discipline, similar to environmental protection, food safety, or even the regulation of medical devices or financial markets. In these domains the emphasis in on preventive measures, certification, private and public expertise, and on the way in which not only individuals by also society and groups are affected.

개인 자격을 부여하는 대신 위험 기반 접근 방식은 적절한 조직 및 기술 조치로 피해를 방지할 수 있는 지속 가능한 정보 생태를 창출하는 데 중점을 둡니다. 후자의 관점에서 볼 때 데이터 보호는 환경 보호, 식품 안전 또는 의료 기기 또는 금융 시장의 규제와 유사한 위험 규제 원칙으로 보입니다. 이러한 영역에서 예방 조치, 인증, 개인 및 공공 전문 지식, 그리고 사회와 그룹의 개인만이 영향을 받는 방식에 중점을 둡니다.

3.7.2. A risk-based approach to AI AI에 대한 위험 기반 접근법

With regard to AI, both the right -based and the risk-based approaches are meaningful, but the second is particularly significant. It has been noted that in the US a risk-based approach to data protection has emerged in the public sector. A 'Big Data due progress',107 has been argued for, which requires agencies to educate officers on biases and fallacies of automation, to appoint hearing officers tasked with reviewing automated decisions, to test regularly computer systems, to ensure that audit trails are kept, etc.108 For instance, it has been argued that the US Federal Trade Commission should play a key role in ensuring fairness and accuracy of credit scoring systems, given the huge impact that a bad credit score may have on people's life. Other suggested remedies include auditing, noticing consumer, and enabling consumers not only to access their data, but also to test the system by submitting hypotheticals.109

AI와 관련하여 올바른 접근 방식과 위험 기반 접근 방식은 모두 의미가 있지만 두 번째 방법은 특히 중요합니다. 미국에서는 공공 부문에서 데이터 보호에 대한 위험 기반 접근 방식이 등장했다. 기관들이 자동화의 편견과 오류에 대해 임원들을 교육하고, 자동화된 의사 결정을 검토하는 청각 담당자를 임명하고, 정기적으로 컴퓨터 시스템을 테스트하고, 감사 추적이 유지되도록 하는 '빅

데이터 마감 진행', 107이 주장되었다. 예를 들어, 미국 연방 무역 위원회는 신용 점수 불량이 사람들의 삶에 미칠 수 있는 큰 영향을 고려할 때 신용 점수 시스템의 공정성과 정확성을 보장하는 데 핵심적인 역할을 수행해야 한다고 주장 해 왔습니다. 제안된 다른 해결책으로는 감사, 소비자 인식 및 소비자가 자신의 데이터에 액세스할 수 있을 뿐만 아니라 가설을 제출하여 시스템을 테스트할 수 있습니다.109

105 On the connection between causal explanations and (ab)normality, see Halpern and Hitchcock (2013)
106 Edwards and Veal (2019).
107 Edwards and Veal (2019).
108 Citron (2008).
109 Citron and Pasquale (2014).

The GDPR also contains a number of provisions that contribute to prevent the misuse of AI, in particular, in connection with the idea of 'privacy by design and by default', namely, with preventive technological and organisational measures.110

GDPR에는 AI의 오용을 방지하는 데 도움이 되는 많은 조항들, 특히 '설계에 의한 개인 정보 보호 및 기본적으로'라는 개념, 즉 예

방적 기술적 및 조직적 조치와 관련하여 제공됩니다.110

A serious issue pertaining to risk-prevention and mitigation measures concerns whether the same measures should be required by all controllers engaging in similar processings or whether a differentiated approach is needed, that takes into account the size of controllers and their financial and technical capacity of adopting the most effective precautions. More precisely, should the same standards be applied both to the Internet giants, which have huge assets and powerful technologies and profit of monopolistic rents, and to small start-ups, which are trying to develop innovative solutions with scanty resources. Possibly a solution to this issue can be found by considering that risk prevention and mitigation measures are the object of best effort obligations, having a stringency that is scalable, depending not only on the seriousness of the risk, but also the capacity of the address of the obligation. Thus, more stringent risk preventions measures may be required to the extent that the controller both causes a more serious social risk, by processing a larger quantity of personal data on larger set of individuals and has superior ability to respond to risk in effective and financially sustainable ways.

위험 방지 및 완화 조치와 관련된 심각한 문제는 유사한 처리에

관여하는 모든 컨트롤러가 동일한 조치를 취해야 하는지 또는 컨트롤러의 크기 및 채택의 재무 및 기술 능력을 고려한 차별화된 접근 방식이 필요한지 여부와 관련이 있습니다. 가장 효과적인 예방책. 보다 정확하게는 거대한 자산과 강력한 기술과 독점 임대료의 이익을 가진 인터넷 거인과 소규모 자원을 가진 혁신적인 솔루션을 개발하려는 소규모 신생 기업 모두에 동일한 표준을 적용해야 합니다. 이 문제에 대한 해결책은 위험 예방 및 완화 조치가 최선의 노력 의무의 대상이며 위험의 심각성 뿐만 아니라 주소의 용량에 따라 확장 성이 엄격함을 고려하여 찾을 수 있습니다. 의무. 따라서, 더 큰 규모의 개인에 대해 많은 양의 개인 데이터를 처리함으로써 컨트롤러가 더 심각한 사회적 위험을 야기하는 정도까지 더욱 엄격한 위험 예방 조치가 필요할 수 있으며 효과적이고 재정적으로 지속 가능한 위험에 대한 대응 능력이 우수합니다. 방법.

3.7.3. Article 24 GDPR: Responsibility of the controller

컨트롤러의 책임

Article 24, on 'Responsibility of the controller', requires the controller to

제 24조 '컨트롤러의 책임'에 대해서는 컨트롤러가

> ***implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation'.***
>
> ***본 규정에 따라 처리가 수행됨을 보장하고 입증할 수 있는 적절한 기술적 및 조직적 조치를 시행해야 합니다.***

Such measures are to be 'reviewed and updated where necessary.' With regard to AI applications, the measures include controls over the adequacy and completeness of training sets, over reasonableness of the inferences, over the existence of causes of bias and unfairness.

이러한 조치는 '필요한 경우 검토 및 업데이트'되어야 합니다. AI 적용과 관련하여 측정에는 훈련 세트의 적절성 및 완전성, 추론의 합리성, 편견 및 불공정의 원인에 대한 통제가 포함됩니다.

3.7.4. Article 25 GDPR: Data protection by design and by default

설계 및 기본적으로 데이터 보호

Article 25 (1) on 'Data protection by design and by default', specifies that both 'at the time of the determination of the means for processing and at the time of the processing' the controller should

'설계 및 기본적으로 데이터 보호'에 관한 제25조 (1)은 '처리 수단 결정 시점 및 처리 시점'에 컨트롤러가

> ***implement appropriate technical and organisational measures which are designed to implement data-protection principles [...] in an effective manner and to integrate the necessary safeguards into the processing.***
>
> ***효과적인 방법으로 데이터 보호 원칙 [...]을 구현하고 필요한 보호 수단을 처리에 통합하도록 설계된적절한 기술적 및 조직적 조치를 구현합니다.***

Article 25(2) addresses data minimisation. It is relevant to AI and big data applications as it requires the implementation of

제 25 (2) 조는 데이터 최소화를 다루고있다. AI 및 빅 데이터 응용 프로그램과 관련이 있습니다.

appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed.

기본적으로 처리의 각 특정 목적에 필요한 개인 데이터 만 처리되도록 하는 적절한 기술적 및 조직적 조치.

Such measures should address 'the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility'. Article 25(2) questions the possibility to retain the data in consideration of future still undetermined purposes, unless the scope the future uses is defined (e.g. scientific or market research).

이러한 조치는 '수집된 개인 데이터의 양, 처리 범위, 저장 기간 및 접근성'을 다루어야 합니다. 제25 (2) 조는 미래의 사용 범위가 정의되지 않는 한 (예 : 과학 또는 시장 조사) 미래의 미결정 목적을 고려하여 데이터를 유지할 가능성에 의문을 제기합니다.

110 Edwards and Veal 2019.

3.7.5. Article 35 and 36 GDPR: Data protection impact assessment

데이터 보호 영향 평가

Article 35 requires that a data protection impact assessment is preventively carried out relatively to processing that is likely to result in a high risk to the rights and freedoms of natural persons. The assessment is required in particular when the processing involves

제35 조는 데이터 보호 영향 평가가 자연인의 권리와 자유에 높은 위험을 초래할 수 있는 처리에 비해 상대적으로 예방적으로 수행되도록 요구하고 있다. 특히 처리 과정에서 평가가 필요합니다.

> ***a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person.***

프로파일링을 포함한 자동화된 처리를 기반으로 하고 자연인에 관한 법적 영향을 초래하거나 자연인에게도 비슷한 영향을 미치는 결정에 근거한 자연인과 관련된 개인적 측면에 대한 체계적이고 광범위한 평가.

Thus, an impact assessment is usually required when AI-based profiling contributes to automated decision-making affecting individuals, since such profiling is likely to be 'systematic and extensive.'

따라서 AI 기반 프로파일링 이 개인에게 영향을 미치는 자동화된 의사 결정에 기여할 때 영향 평가가 필요합니다. 이러한 프로파일링은 '체계적이고 광범위합니다'.

When the assessment determines that a processing involves 'high risk', according to Article 36 (1) the controller should preventively ask the supervisory authority (the national data protection authority) for advice.

평가 결과 처리에 '고 위험성'이 포함된 것으로 판단되면, 제36조 (1)에 따라, 감독 당국은 감독 당국 (국가 데이터 보호 당국)에 조언을 구해야 한다.

The controller shall consult the supervisory authority prior to processing where a data protection impact assessment under Article 35 indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk.

관리자는 제35 조에 따른 데이터 보호 영향 평가에서 위험을 완화하기 위해 조치를 취하지 않은 조치가 없을 경우 처리가 높은 위험을 초래할 것임을 나타내는 처리 전에 감독 기관에 문의해야 합니다.

The impact assessment must be shared with the supervisory authority. The authority must provide written advice to the controller where

영향 평가는 감독 기관과 공유해야 합니다. 권한은 컨트롤러에게 서면 조언을 제공해야 합니다.

the supervisory authority is of the opinion that the intended processing referred to in paragraph 1 would infringe this Regulation, in particular where the controller has insufficiently identified or mitigated the

risk.

감독 당국은 제1 항에 언급된 의도된 처리가 특히 본 규제 기관이 위험을 불충분하게 식별하거나 완화한 경우이 규정을 위반할 것이라고 생각한다.

The authority may also use its investigative and corrective powers. In particular it may (article 50(2)(d)):

당국은 또한 조사 및 수정 권한을 사용할 수 있습니다. 특히 (제 50조 (2) (d))는 다음과 같습니다.

order the controller or processor to bring processing operations into compliance with the provisions of this Regulation

컨트롤러 또는 프로세서가 이 규정의 조항을 준수하도록 처리 작업을 수행하도록 명령

The authority may even temporarily or permanently ban the use of the system (article 50(2)(f)).

당국은 심지어 시스템의 사용을 일시적 또는 영구적으로 금지할

수 있다 (제50조 제2 항 (f)).

Articles 35 and 36 are particularly important to the development of data-protection compliant AI application, since may enable cooperation and mutual learning between data protection authorities and controllers.

제35 조와 제36 조는 데이터 보호 준수 AI 응용 프로그램 개발에 특히 중요합니다. 데이터 보호 기관과 컨트롤러 간의 협력 및 상호 학습이 가능하기 때문입니다.

3.7.6. Article 37 GDPR: Data protection officers

데이터 보호 책임자

Article 37 requires controllers to designate a data protection officer when they engage in

제37 조는 컨트롤러가 데이터 보호 책임자를 지정할 때

processing operations which, by virtue of their nature,

their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale, or when they process on a large-scale sensitive data or data concerning criminal convictions.

본질 상 그 범위 및/또는 목적으로 인해 대규모로 데이터 주제를 정기적으로 그리고 체계적으로 모니터링해야 하거나 대규모의 민감한 데이터 또는 범죄 유죄 관련 데이터를 처리할 때 처리 작업.

This provision is relevant to AI, since various AI-based applications are based on data sets collected by the monitoring the behaviour of data subject (e.g., their online behaviour, or their driving behaviour, etc.). A specialised and impartial internal review would arguably be useful in such cases.

이 조항은 다양한 AI 기반 애플리케이션이 데이터 주제의 동작(예 : 온라인 동작 또는 운전 동작 등)을 모니터링하여 수집한 데이터 세트를 기반으로 하기 때문에 AI와 관련이 있습니다. 이러한 경우에는 전문적이고 공정한 내부 검토가 유용할 것입니다.

3.7.7. Articles 40-43 GDPR: Codes of conduct and certification

행동 강령 및 인증

Articles 40-43 address codes of conduct and certification. While these provisions do not make explicit reference to AI, codes and conduct and certification procedure may be highly relevant to AI, given the risks involved in AI application, and the limited guidance provided by legal provisions.

40-43 조는 행동 강령 및 인증을 다루고 있습니다. 이러한 조항들이 AI를 명시적으로 언급하지는 않지만, AI적용과 관련된 위험 및 법적 조항에 의해 제공된 제한된 지침을 고려할 때, 코드 및 행동 및 인증 절차는 AI와 매우 관련이 있을 수 있습니다.

Adherence to codes of conduct and certification mechanisms, according to Articles 24 and 25 may contribute to demonstrate compliance with the obligations of the controller and with the requirements of privacy by design. The idea of a certification for AI applications has been endorsed by the European Economic and Social Committee (EESC) which 'calls for the development of a robust certification system based on test procedures that enable

companies to state that their AI systems are reliable and safe.' Thus, it suggests developing a 'European trusted-AI Business Certificate based partly on the assessment list put forward by the High-Level Experts' group on AI.' On the other hand, some perplexities on a general framework for certification have also been raised, based on the complexity of AI technologies, their diversity, and their rapid evolution.111

제 24조 및 제25 조에 따라 행동 강령 및 인증 메커니즘을 준수하는 것은 컨트롤러의 의무와 설계에 의한 프라이버시 요구 사항의 준수를 입증하는 데 기여할 수 있습니다. AI 응용 프로그램에 대한 인증 아이디어는 유럽 경제 사회위원회 (EESC)에 의해 승인되었으며, 이는 기업이 AI 시스템이 신뢰할 수 있고 안전하다고 진술할 수 있는 테스트 절차를 기반으로 강력한 인증 시스템의 개발을 요구합니다. ' 따라서 AI에 대한 고급 전문가 그룹이 제시한 평가 목록에 부분적으로 기초하여 유럽의 신뢰할 수 있는 AI 비즈니스 인증서를 개발할 것을 제안합니다. 다른 한편으로, AI 기술의 복잡성, 다양성 및 빠른 진화에 기초하여 인증을 위한 일반적인 프레임 워크에 대한 일부 난관도 제기되었습니다.111

Certification and code of conducts could address both algorithms as such (in particular with regard to their technical quality and accuracy) as well as the context of their application (training sets,

input data, intended outcomes and their uses). They could enable sectorial approaches and the rapid adaptation to technological and social changes.

인증 및 행동 강령은 (특히 기술 품질 및 정확성과 관련하여) 알고리즘뿐만 아니라 적용 상황 (훈련 세트, 입력 데이터, 의도된 결과 및 용도)을 모두 다룰 수 있습니다. 그들은 부문 별 접근과 기술 및 사회적 변화에 대한 빠른 적응을 가능하게 할 수 있다.

On the other hand, it has been observed that 'voluntary self-or co-regulation by privacy seal has had a bad track record in privacy, with recurring issues around regulatory and stakeholder capture.'112 Certification and codes of conduct – in combination with the requirement to demonstrate compliance, according to accountability – may lead to formalistic practices, rather than to the real protection of the interests of data subject.113 Much will depend on the extent to which data protection authorities will supervise the adequacy of these soft law instruments, and the effectiveness of their application.

반면, '개인 정보 보호에 의한 자발적 자체 또는 공동 규제는 규제 및 이해 관계자 확보와 관련하여 되풀이되는 문제로 인해 개인 정보 보호에 나쁜 기록을 가지고 있습니다.'112 인증 및 행동 강

령 – 책임에 따라 준수를 입증해야 하는 요구 사항 – 데이터 주체의 이익을 실제로 보호하기보다는 공식적인 관행으로 이어질 수 있다.113 데이터 보호 당국이 이러한 법률의 적절성을 감독하는 정도에 따라 크게 달라질 것이다. 그리고 그들의 적용 효과.

3.7.8. The role of data protection authorities

데이터 보호 기관의 역할

As shown in the previous sections, there are various references in the GDPR that support a proactive risk-based approach towards AI and big data. It will be up to the creativity of technological and legal experts, in particular those having the role of data protection officers, to provide adequate solutions. An important role can also be played by data protection authorities, in enforcing data protection law, but also in proposing and promoting appropriate standards. The GDPR makes explicit reference both to National data protection authorities and to the European Data Protection Board, to which is confers an important role.

이전 섹션에서 볼 수 있듯이 GDPR에는 AI 및 빅 데이터에 대한 사전 위험 기반 접근 방식을 지원하는 다양한 참조가 있습니다.

적절한 솔루션을 제공하는 것은 기술 및 법률 전문가, 특히 데이터 보호 책임자의 역할을 담당하는 전문가의 창의성에 달려 있습니다. 데이터 보호 당국은 데이터 보호법을 시행 할 뿐만 아니라 적절한 표준을 제안하고 홍보하는 데 중요한 역할을 수행할 수도 있습니다. GDPR은 국가 데이터 보호 당국과 유럽 데이터 보호위원회를 모두 명시적으로 언급하며 중요한 역할을 합니다.

The European Data Protection Board is the continuation or the Article 29 Working Party, established by the 1995 Data Protection Directive. It includes representatives of the Member States' data protection authorities and of the European data protection supervisors is meant to ensure the consistent application of the Regulation. According to Recital (77) the Board is supposed to provide guidance on the implementation of the GDPR through guidelines:

유럽 데이터 보호위원회는 1995년 데이터 보호 지침에 의해 제정된 지속 또는 제29조 작업반입니다. 여기에는 회원국의 데이터 보호 당국과 유럽 데이터 보호 감독관의 대표가 포함되어 있어 규정의 일관성 있는 적용을 보장합니다. Recital (77)에 따르면 이 사회는 지침을 통해 GDPR 이행에 대한 지침을 제공해야 합니다.

Guidance on the implementation of appropriate measures and on the demonstration of compliance by the controller or the processor, especially as regards the identification of the risk related to the processing, their assessment in terms of origin, nature, likelihood and severity, and the identification of best practices to mitigate the risk, could be provided in particular by means of approved codes of conduct, approved certifications, guidelines provided by the Board or indications provided by a data protection officer.

특히 처리와 관련된 위험의 식별, 원산지, 자연, 가능성 및 심각성 측면에서의 평가, 그리고 식별과 관련하여 적절한 조치의 구현 및 컨트롤러 또는 프로세서에 의한 준수 입증에 대한 지침 위험을 완화하기위한 모범 사례는 특히 승인된 행동 강령, 승인된 인증, 이사회가 제공한 지침 또는 데이터 보호 담당자가 제공한 표시를 통해 제공될 수 있습니다.

111 AI Now (2018) report

112 Edwards and Veal (2019, 80).

113 Edwards and Veal (2019, 80).

The Board is entrusted with the task of determining whether certain processing operations do not involve high risks, and of indicating what measures may be appropriate in such cases:

IASB는 특정 처리 작업에 높은 위험이 수반되지 않는지 여부를 결정하고 이러한 경우 어떤 조치가 적절한 지 표시하는 임무를 맡고 있습니다.

> ***The Board may also issue guidelines on processing operations that are considered to be unlikely to result in a high risk to the rights and freedoms of natural persons and indicate what measures may be sufficient in such cases to address such risk.***
>
> ***IASB는 또한 자연인의 권리와 자유에 대한 높은 위험을 초래할 가능성이 없는 것으로 간주되는 처리 작업에 대한 지침을 발행할 수 있으며 그러한 경우 그러한 위험을 해결하기 위해 어떤 조치가 충분한 지 표시할 수 있습니다.***

An explicit reference to automated decision-making is contained in Article 70 (1)(f) GDPR, which lists the tasks of Board. With regard to automated decision-making the Board should

자동화된 의사 결정에 대한 명시적 언급은 이사회의 업무를 열거한 GDPR 70 (1) (f) GDPR에 포함되어 있습니다. 자동화된 의사 결정과 관련하여 이사회는

> ***on its own initiative or, where relevant, at the request of the Commission, issue guidelines, recommendations and best practices [...] for further specifying the criteria and conditions for decisions based on profiling pursuant to Article 22(2)***
>
> ***위원회 자체 요청에 따라 위원회의 요청에 따라 가이드 라인, 권고 사항 및 모범 사례 [...]를 추가하여 제22 (2) 조에 따른 프로파일링에 근거한 결정의 기준과 조건을 구체적으로 명시한다.***

3.8. AI, statistical processing and scientific research

AI, 통계 처리 및 과학 연구

AI and big data provide not only risks but also great opportunities. In particular, they offer new avenues to gain knowledge about nature and society that can be used for beneficial purposes.

AI와 빅 데이터는 위험 뿐만 아니라 큰 기회도 제공합니다. 특히, 그들은 유익한 목적으로 사용될 수 있는 자연과 사회에 대한 지식을 얻을 수 있는 새로운 길을 제공합니다.

Consider for instance the huge importance of applying AI to medical data, to improve the accuracy of medical tests, to assess connection between symptoms and pathologies, to analyse the effectiveness of therapies. Similar considerations also concern the AI and big data applications to social and economic data, to better plan and optimise private and public activities. As note in Section 2.3.2, big data analytics can lead to unexpected discoveries, which may result from combining data collected for different purposes. Thus, the traditional principles of data protection, such as data minimisation and purpose limitation are challenged, since they may preclude some useful applications and technological development. The problem is aggravated by the fact that many non-European countries seem to offer normative environments that are more facilitative to the full development and deployment of AI systems.

예를 들어 AI를 의료 데이터에 적용하고, 의료 테스트의 정확성을 개선하고, 증상과 병리 사이의 연결을 평가하고, 치료의 효과를 분석하는 것이 매우 중요하다는 것을 고려하십시오. 개인 및 공공 활동을 더 잘 계획하고 최적화하기 위해 사회 및 경제 데이터에

대한 AI 및 빅 데이터 응용 프로그램과 비슷한 고려 사항도 있습니다. 섹션 2.3.2에서 언급한 바와 같이, 빅 데이터 분석은 예기치 않은 발견으로 이어질 수 있으며, 이는 다른 목적으로 수집된 데이터를 결합하여 발생할 수 있습니다. 따라서, 데이터 최소화 및 목적 제한과 같은 전통적인 데이터 보호 원칙은 일부 유용한 응용 프로그램 및 기술 개발을 배제할 수 있기 때문에 어려운 과제입니다. 유럽 이외의 많은 국가들이 AI 시스템의 완전한 개발 및 배포에 보다 용이한 규범적인 환경을 제공하는 것처럼 보임에 따라 문제가 더욱 악화됩니다.

3.8.1. The concept of statistical processing

통계 처리의 개념

It has been argued that the way forward, to enable the use of big data analytics also in Europe is to refer to the discipline for scientific and statistical purposes.114 In particular, Recital (162) GDPR refers to further EU or National law for the regulation of processing for statistical purposes:

유럽에서도 빅 데이터 분석을 사용할 수 있게 하려면 과학과 통계적 목적을 위해 규율을 참조해야 합니다.114 특히, Recital (162) GDPR은 EU 또는 국가 법률을 통계 목적의 처리 규정 :

Union or Member State law should, within the limits of this Regulation, determine statistical content, control of access, specifications for the processing of personal data for statistical purposes and appropriate measures to safeguard the rights and freedoms of the data subject and for ensuring statistical confidentiality.

노조 또는 회원 국법은이 규정의 범위 내에서 통계적 내용, 접근 통제, 통계 목적의 개인 데이터 처리 사양 및 데이터 주체의 권리와 자유를 보호하고 통계적 기밀성을 보장하기 위한 적절한 조치를 결정해야합니다..

In the same Recital, processing for statistical purposes is positively characterised by the objective of producing statistical surveys and results and negatively by the fact that their outcomes are not used for measures or decisions concerning particular individuals:

동일한 리사이틀에서 통계 목적의 처리는 통계 조사 및 결과를 생성한다는 목표와 그 결과가 특정 개인에 관한 측정 또는 결정에 사용되지 않는다는 사실에 의해 긍정적으로 특징 지워집니다.

Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. Those statistical results may further be used for different purposes, including a scientific research purpose. The statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person.

통계 목적은 통계 조사 또는 통계 결과 생성에 필요한 개인 데이터 수집 및 처리 작업을 의미합니다. 이러한 통계 결과는 과학적 연구 목적을 포함하여 다른 목적으로 추가로 사용될 수 있습니다. 통계적 목적은 통계적 목적에 대한 처리 결과가 개인 데이터가 아니라 집계 데이터이며 이 결과 또는 개인 데이터가 특정 자연인에 관한 조치 또는 결정을 지원하는 데 사용되지 않음을 의미합니다.

As it emerges from this characterisation, the meaning of statistical purpose in the GDPR is not narrowly defined and may be constructed as including not only uses for the public interest, but also by private companies for commercial goals.115

이 특징에서 나온 바와 같이, GDPR의 통계적 목적의 의미는 좁게 정의되지 않으며 공익 목적 뿐만 아니라 민간 기업의 상업적 목표를 포함하여 구성될 수 있다.115

114 Mayer-Schonberger and Padova 2016.
115 Mayer-Schoeberger and Padova 2016, 326-7

3.8.2. Article 5(1)(b) GDPR: Repurposing for research and statistical processing 연구 및 통계 처리를 위한 용도 변경

According to Article 5(1)(b) repurposing data for statistical purposes is in principle admissible, as it will 'not be considered to be incompatible with the initial purposes.' Similarly, at 5(1)(e) data retention limits are relaxed with regard to processing for research and statistical purposes. However, processing for research and statistical purposes requires appropriate safeguards, including in particular pseudonymisation. On the other hand, EU or National law may provide for derogation from the data subjects' rights, when needed to achieve scientific or statistical purposes.

제 5 (1) (b) 조에 따르면 통계 목적을 위한 용도 변경 데이터는 원칙적으로 '초기 목적과 호환되지 않는 것으로 간주되지 않기 때

문에 허용된다'. 유사하게, 5 (1) (e)에서 연구 및 통계 목적의 처리와 관련하여 데이터 보유 한계가 완화된다. 그러나 연구 및 통계 목적의 처리에는 특히 가명 화를 포함한적절한 보호 조치가 필요합니다. 반면에, EU 또는 국가 법은 과학적 또는 통계적 목적을 달성하기 위해 필요할 때 데이터 주체의 권리를 박탈할 수 있습니다.

3.8.3. Article 89(1,2) GDPR: Safeguards for research of statistical processing 통계 처리 연구를 위한 보호

Statistical processing is addressed in Article 89(1), requiring that appropriate safeguards are adopted for processing for archiving, research or statistical purposes and that in particular that the data be pseudonymised or anonymised when these purposes can be achieved in this manner.

통계 처리는 제89 (1) 조에서 다루어지며, 보관, 연구 또는 통계 목적으로 처리하기 위해적절한 보호 수단을 채택해야 하며 특히 이러한 방식으로 이러한 목적을 달성할 수 있을 때 데이터를 가명 화하거나 익명화해야 합니다.

Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject.

공공 이익, 과학적 또는 역사적 연구 목적 또는 통계적 목적으로 아카이빙 목적으로 처리하는 것은 이 규정에 따라 데이터 주체의 권리와 자유에 대해적절한 보호 조치를 받아야합니다.

The safeguards are linked to data minimisation, though a reference is made not only to anonymisation but also to pseudonymisation (which does not involve a reduction in the amount of personal data).

보호는 데이터 최소화와 관련이 있지만 익명화 뿐만 아니라 가명화 (개인 데이터 량의 감소를 포함하지 않음)도 참조합니다.

Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data

minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.

이러한 보호 조치는 특히 데이터 최소화 원칙에 대한 존중을 보장하기 위해 기술적 및 조직적 조치가 마련되어 있는지 확인해야 합니다. 그러한 조치에는 가명 화가 포함될 수 있으며, 그러한 목적으로 그러한 목적을 달성할 수 있는 경우. 데이터 주체의 식별을 허용하지 않거나 더 이상 허용하지 않는 추가 처리를 통해 이러한 목적을 달성할 수 있는 경우, 그러한 목적을 달성해야 합니다.

Finally, Article 89 (2) allows for derogations from certain data subjects' rights – to access (Article 15 GDPR), to rectification (16), to restriction of processing (18), to object (21)– in the case of processing for research or statistical purposes.

마지막으로, 제89조 제2 항은 처리의 경우 특정 데이터 주체의 권리-접근 (제15조 GDPR), 정류 (16), 처리 제한 (18), 이의 제기 (21)에 대한 철회를 허용합니다. 연구 또는 통계 목적.

Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

개인 정보가 과학적 또는 역사적 연구 목적 또는 통계적 목적으로 처리되는 경우, 연합 또는 회원국 법률은 제15 조, 제16 조, 제18조 및 제21 조에 언급된 권리에서 제1 항에 언급된 조건 및 보호 조치에 따라 철회할 수 있습니다. 이 조는 그러한 권리가 특정 목적의 달성을 불가능하게 하거나 심각하게 손상시킬 가능성이 있는 한, 그러한 목적을 달성하기 위해서는 그러한 규범이 필요하다.

It has been argued that the EU and member States have a strong interest in enabling statistical processing, to support economic and

technological development. Thus, they may use the provisions above to enable this processing on a large scale, while establishing the required safeguards and derogations. This would provide the opportunity for an EU approach to data analytics, which is compatible with effective data protection:

EU와 회원국은 경제 및 기술 개발을 지원하기 위해 통계 처리를 가능하게 하는 데 큰 관심을 가지고 있다고 주장했다. 따라서 그들은 위의 규정을 사용하여 이 처리를 대규모로 수행하면서 필요한 보호 조치와 폐기를 설정할 수 있습니다. 이는 효과적인 데이터 보호와 호환되는 EU의 데이터 분석 접근 방식을 제공합니다.

GDPR is making small, but noteworthy steps towards enabling Big Data in Europe. It is a peculiar kind of Big Data, though, that European policymakers are facilitating: one that emphasizes reuse and permits some retention of personal data, but that at the same time remains very cautious when collecting data.116

GDPR은 유럽에서 빅 데이터를 활성화하기위한 작지만 주목할만한 조치를 취하고 있습니다. 그러나 유럽의 정책 입안자들은 재사용을 강조하고 개인 데이터의 일부 보존을 허용하지만 동시에 데이터를 수집할 때 매우 신중한 태도

를 취하는 것은 일종의 빅 데이터입니다.116

The facilitations for scientific and statistical processing, however, may extend beyond reuse and retention: these kinds of processing may also be justified by legitimate interests according to 6(1)(f), as long as the processing is done in such a way as to duly fulfil the that data subjects' data protection interests, including their interests in not being subject to risks because of unauthorised uses of their data.

그러나 과학적 및 통계적 처리에 대한 촉진은 재사용 및 보유를 넘어 확장될 수 있다. 이러한 종류의 처리는 처리가 다음과 같은 방식으로 수행되는 한 6 (1) (f)에 따라 정당한 이익에 의해 정당화 될 수 있다. 데이터의 무단 사용으로 인해 위험에 노출되지 않는 것에 대한 관심사를 포함하여 해당 데이터 주체의 데이터 보호 이익을 정식으로 이행합니다.

A difficult issue concerns whether access to the data sets of personal information supporting statistical inferences (e.g., to predict consumer preferences, or market trends) should be limited to the companies or public bodies who have collected the data. On the one hand, allowing, or even requiring, that the original

controllers do not make the data accessible to third parties, may affect competition and prevent beneficial uses of the data. On the other hand, requiring the original controllers to make their data sets available to third parties would cause additional data protection risks.

어려운 문제는 통계적 추론을 지원하는 개인 정보의 데이터 세트(예 : 소비자 선호도 예측 또는 시장 추세)에 대한 액세스가 데이터를 수집한 회사 나 공공 기관으로 제한되어야 하는지 여부와 관련이 있습니다. 한편으로, 원래의 컨트롤러가 제3자가 데이터에 액세스할 수 없도록 함으로써 경쟁에 영향을 미치고 데이터의 유익한 사용을 막을 수 있습니다. 반면, 원래 컨트롤러가 데이터 세트를 타사에 제공하도록 요구하면 추가 데이터 보호 위험이 발생할 수 있습니다.

116 Mayer-Schoenberger and Padova 2016, 331

4. Policy options: How to reconcile AI-based innovation with individual rights & social values, and ensure the adoption of data protection rules and principles

정책 옵션 : 개인의 권리와 사회적 가치로 AI 기반 혁신을 조정하고 데이터 보호 규칙 및 원칙의 채택을 보장하는 방법

In this section, the main results of the report will be summarised, pointing out the main conclusions reached and proposing some policy indications.

이 섹션에서는 보고서의 주요 결과를 요약하여 도달한 주요 결론을 지적하고 일부 정책 표시를 제안합니다.

4.1. AI and personal data

In Section 2 the social and legal issues pertaining to the application of AI to personal data have been discussed. First opportunities and risks have been illustrated, and then the key ethical and legal issues

have been considered.

섹션 2에서는 개인 데이터에 AI를적용하는 것과 관련된 사회적 및 법적 문제가 논의되었습니다. 첫 번째 기회와 위험이 설명된 다음 주요 윤리 및 법적 문제가 고려되었습니다.

4.1.1. Opportunities and risks

First, the concept of AI has been introduced and the development of AI research and applications have been presented, focusing particularly on the recent successes of machine learning based models for narrow AI.

먼저 AI 개념을 소개하고 AI 연구 및 응용 프로그램 개발을 발표했으며 특히 좁은 AI를위한 머신러닝 기반 모델의 성공에 초점을 맞추고 있습니다.

Then, the ways in which AI-based systems may use personal data have been described and the resulting opportunities and risks have been illustrated. It has been observed that personal data can be used to predict human behaviour, to learn the propensities and

attitudes of individuals, to exercise influence over behaviour. The feedback relations between AI and big (personal) data have also been considered: the possibility of using AI stimulates the collection of vast sets of personal data, and the availability of big data sets, in its turn, stimulates novel applications of AI.

그런 다음 AI 기반 시스템이 개인 데이터를 사용하는 방법을 설명하고 그 결과로 얻은 기회와 위험을 설명했습니다. 개인 데이터는 인간 행동을 예측하고, 개인의 성향과 태도를 배우고, 행동에 대한 영향력을 행사하기 위해 사용될 수 있음이 관찰되었습니다. AI와 빅 (개인) 데이터 간의 피드백 관계도 고려되었습니다. AI를 사용할 가능성은 방대한 개인 데이터의 수집을 자극하고 빅 데이터 세트의 가용성은 AI의 새로운 응용을 자극합니다.

Benefits and risks concerning the deployment of AI have been examined. The combination of AI and big data offers great opportunities for scientific research, welfare, governance and administration, but it also engenders serious risks for individuals and society: intensified surveillance, control, manipulation, unfairness and discrimination. Even when the processing of data is non-discriminatory and based on reliable technologies, it may lead to unacceptable levels of surveillance, control and nudging, which affect individual autonomy, cause psychological harm, and impair

genuine social interactions and the formation of public opinion.

AI 배포와 관련된 이점과 위험이 조사되었습니다. AI와 빅 데이터의 결합은 과학 연구, 복지, 거버넌스 및 관리에 큰 기회를 제공하지만 강화된 감시, 통제, 조작, 불공정 및 차별과 같은 개인과 사회에 심각한 위험을 초래합니다. 데이터 처리가 비 차별적이며 신뢰할 수 있는 기술을 기반으로 하더라도 수용할 수 없는 수준의 감시, 제어 및 노출이 발생할 수 있으며 이는 개별 자율성에 영향을 미치고 심리적 해를 끼치며 진정한 사회적 상호 작용 및 여론 형성에 영향을 미칩니다.

4.1.2. Normative foundations. 규범적 기초.

The normative foundations of a human-centred regulation of AI have been considered. It has been observed that a framework is emerging, in which traditional ethical ideas, such as respect for human autonomy, prevention of harm, and fairness are combined with specific and somehow technical requirements concerning transparency, explicability, robustness and safety.

인간 중심의 인공 지능 규제의 규범적인 기초가 고려되었습니다. 인간의 자율성 존중, 피해 방지 및 공정성과 같은 전통적인 윤리

적 아이디어가 투명성, 설명 성, 견고성 및 안전에 관한 특정적이고 기술적인 요구 사항과 결합된 프레임 워크가 등장하고 있음이 관찰되었습니다.

Turning from ethics to law, it has been claimed that AI relates to the law at different levels. As a pervasive and multifaceted technology, AI may either enhance or impair the exercise of multiple fundamental rights: privacy and data protection, civil freedoms and social rights. It can also contribute to, or detract from, the realisation of different social values, such as democracy, welfare, or solidarity. Correspondingly, promoting the opportunities of AI and countering its risks falls within the purview of multiple areas of the law, from data protection, to consumer protection, competition law, labour law, constitutional and administrative law. Different interests are at stake: the interests in data protection, in a fair algorithmic treatment, in transparency and accountability, in not being misled or manipulated, in the trustworthiness of AI systems, in algorithmic competition, among others.

윤리에서 법으로 전환하면서 인공 지능은 다른 수준의 법과 관련이 있다고 주장되었습니다. 광범위하고 다면적인 기술인 AI는 개인 정보 및 데이터 보호, 시민의 자유 및 사회적 권리와 같은 여러 가지 기본 권리의 행사를 강화하거나 손상시킬 수 있습니다.

또한 민주주의, 복지 또는 연대와 같은 다른 사회적 가치의 실현에 기여하거나 방해할 수 있습니다. 이에 따라 AI의 기회를 홍보하고 그 위험에 대응하는 것은 데이터 보호에서 소비자 보호, 경쟁법, 노동법, 헌법 및 행정법에 이르기까지 법률의 여러 영역에 속합니다. 데이터 보호, 공정한 알고리즘 처리, 투명성 및 책임 성, AI 시스템의 신뢰성, 알고리즘 경쟁 등의 오해 또는 조작에 대한 관심은 서로 다른 관심을 끌고 있습니다.

4.2. AI in the GDPR

Based on this analysis, the provisions in the GDPR have been analysed to determine to what extent they adequately address AI applications. Does the GDPR contribute make it possible to enjoy the opportunities enabled by AI while preventing the attendant risks, or does it rather fail in this mission, either by establishing barriers to the beneficial deployment of AI, or conversely failing to prevent avoidable risks?

이 분석에 기초하여, GDPR의 조항들은 AI적용을 어느정도 적절하게 다루는 지 결정하기 위해 분석되었습니다. GDPR이 AI에 의한 기회를 즐기면서 수반되는 위험을 방지하거나 AI의 유익한 배치에 대한 장벽을 설정하거나 반대로 피할 수 있는 위험을 방지하

지 못함으로써 이 임무에서 실패할 수 있도록 기여합니까?

4.2.1. Personal data in re-identification and inferences

재식별 및 추론의 개인 데이터

First of all, AI raises issues pertaining to the very nature of personal data, concerning in particular the possibility of reconnecting the data subjects with their de-identified data, and the possibility of inferring new personal data from existing data. In this regard the notion of personal data in the GDPR does not provide clear answers. It would be advisable to clarify, possibly in a soft-law instrument, such as an opinion of the Article 29 Working Party, that re-identification consists of a processing of personal data, and indeed can be assimilated to collection of new personal data. Therefore, re-identification is fully subject to all GDPR requirements (including obligations to inform the data subject and the need for a legal basis).

우선, AI는 특히 개인 정보를 비 식별 처리된 데이터와 재 연결할 가능성 및 기존 데이터로부터 새로운 개인 정보를 유추할 수 있는 가능성과 관련하여 개인 정보의 본질과 관련된 문제를 제기합

니다. 이와 관련하여 GDPR의 개인 데이터 개념은 명확한 답변을 제공하지 않습니다. 제29조 작업반의 의견과 같은 연약한 법률 문서에서 재식별은 개인 데이터의 처리로 구성되며 실제로 새로운 개인 데이터의 수집에 동화될 수 있음을 명확히 하는 것이 좋습니다. 따라서 재식별은 모든 GDPR 요건 (데이터 주체에게 정보를 제공할 의무 및 법적 근거의 필요성 포함)을 전적으로 준수합니다.

Special considerations apply to the inference of personal data. A possible approach could consist in distinguishing the cases in which an inference of personal data is accomplished without engaging in consequential activities, i.e., the inferred personal data are merely the output of a computation which does not trigger consequential actions, and the cases in which the inferred data are also used as input for making assessment and decisions. In the latter case, the data should definitely count as newly collected personal data.

개인 정보 유추에는 특별한 고려 사항이 적용됩니다. 가능한 접근법은 결과적 활동에 관여하지 않고 개인 데이터의 추론이 이루어지는 경우를 구별하는 것으로 구성될 수 있다. 즉, 추론된 개인 데이터는 단지 결과적인 행동을 유발하지 않는 계산의 결과 일 뿐이다. 유추된 데이터는 평가 및 결정을 위한 입력으로도 사용됩

니다. 후자의 경우 데이터는 새로 수집된 개인 데이터로 간주됩니다.

4.2.2. Profiling

Profiling is at the core of the application of AI to personal data: it consists in inferring new personal data (expanding a person's profile) on the basis of the available personal data. Profiling provides the necessary precondition for automated decision-making, as specifically regulated in the GDPR. A key issue is the extent to which the law may govern and constrain such inferences, and the extent of the data subject's rights in relation to them. This aspect is also not clearly worked out in the GDPR.

프로파일링은 AI를 개인 데이터에 적용하는 핵심입니다. 사용 가능한 개인 데이터를 기반으로 새로운 개인 데이터 (사람의 프로필 확장)를 유추하는 것으로 구성됩니다. 프로파일링은 GDPR에서 구체적으로 규제되는 자동화된 의사 결정에 필요한 전제 조건을 제공합니다. 주요 이슈는 법률이 그러한 추론을 지배하고 제한할 수 있는 정도와 관련하여 데이터 주체의 권리의 범위입니다. 이 측면은 GDPR에서도 명확하게 해결되지 않습니다.

Neither is the extent to which the data subject may have a right to reasonable automated inferences clear, even when these inferences provide a basis for making assessments or decisions.

이러한 추론이 평가 나 결정을 위한 근거를 제공하더라도 데이터 주체가 합리적인 자동 추론에 대해 명확한 권리를 가질 수 있는 정도도 아닙니다.

4.2.3. Consent

The requirement of specificity, granularity and freedom of consent are difficult to realise in connection with AI applications. Thus, in general, consent will be insufficient to support an AI application, unless it appears that the application pursues a legitimate interest and does not unduly sacrifice the data subject's rights and interests under Article 6 (1)(f). There are, however, cases in which consent by the data subject would be the decisive criterion by which to determine whether his or her interests have been sufficiently taken into consideration by the controller (e.g., consent to profiling in the interest of the data subject).

AI 응용 프로그램과 관련하여 구체적, 세분성 및 동의의 자유 요구 사항을 실현하기는 어렵습니다. 따라서, 애플리케이션이 합법적인 이익을 추구하고 제6조 (1) (f)에 따라 데이터 주체의 권리와 이익을 과도하게 희생하지 않는 한, AI 애플리케이션을 지원하기에는 동의가 불충분합니다. 그러나 데이터 주체의 동의가 컨트롤러가 자신의 관심사를 충분히 고려했는지 여부를 결정하는 결정적인 기준이 되는 경우가 있습니다 (예 : 데이터 주체의 이해에 대한 프로파일링에 동의)..

4.2.4. AI and transparency

The report distinguishes between information to be provided before the data subject's data are processed for the purpose of profiling and automated decision-making (ex-ante information), and the information to be provided after the data have been processed (ex-post information).

이 보고서는 프로파일링 및 자동화된 의사 결정을 위해 데이터 주체의 데이터가 처리되기 전에 제공될 정보 (예외 정보)와 데이터가 처리된 후 제공될 정보 (사후 정보)를 구분합니다..

Ex-ante information is addressed by the right to information established by Articles 13(2)(f) and 14(2)(g) requiring two kinds of information to be provided: information on the existence of automated decision-making and meaningful information on its logic and envisaged consequences.

사전 정보는 두 가지 종류의 정보 제공을 요구하는 제13 (2) (f) 및 14 (2) (g) 조에 의해 확립된 정보에 대한 권리에 의해 다루어진다 : 자동 의사 결정의 존재에 관한 정보 및 의미있는 정보논리와 예상 결과에.

There is an uncertainty as to what is meant by the logic and consequences of an automated decision. With regard to complex AI processing, there is a conflict between the need for the information to be concise and understandable on the one hand, and the need for it to be precise and in-depth on the other.

자동화된 의사 결정의 논리와 결과가 무엇을 의미하는지에 대해서는 불확실성이 있습니다. 복잡한 AI 처리와 관련하여 정보가 간결하고 이해하기 쉬울 필요와 정보가 정확하고 깊이 있어야하는 것은 갈등이 있습니다.

Ex-post information is addressed by Article 15(1), which reiterates the same information requirements in Articles 13 and 14. It remains to be determined whether the controller is required to provide the data subject with only general information or also with an individualised explanation.

사후 정보는 제15조 1 항에 의해 다루어지며, 제15조 제1 항은 제13 조와 제14 조에서 동일한 정보 요구 사항을 반복한다. 컨트롤러가 데이터 주체에게 일반적인 정보 만 제공해야 하는지 또는 개별화된 설명을 제공해야 하는지 결정해야 한다..

4.2.5. The rights to erasure and portability

The GDPR provisions on the rights to erasure and portability do not specifically address AI-based processing. However, some important issues emerge concerning the scope of such rights. With regard to the right to erasure, we may ask whether it may also cover inferred information and with regard to the right to portability, whether it also includes information collected by tracking the individuals concerned. The scope of the right to erasure, as distinguished from the right to object, depends on the extent to which the processing is unlawful. Thus, uncertainties

about the unlawfulness of the processing will likely also affect the right to erasure.

소거 및 이동권에 관한 GDPR 조항은 AI 기반 처리를 구체적으로 다루지 않습니다. 그러나 그러한 권리의 범위와 관련하여 몇 가지 중요한 문제가 발생합니다. 삭제 권한과 관련하여, 우리는 추론된 정보를 포함할 수 있는지 여부와 이식성에 대한 권리와 관련있는 개인을 추적하여 수집한 정보를 포함하는지 여부를 요청할 수 있습니다. 소거 권과 구별되는 소거 권의 범위는 처리가 불법 인 정도에 달려 있습니다. 따라서, 처리의 불법성에 대한 불확실성 또한 소거 권에 영향을 줄 수 있습니다.

4.2.6. The right to object

Article 21 specifically addresses the ability to object to profiling, on personal grounds, when the processing is based on public interests (Article 6 (1)(e)), or on legitimate private interests (Article 6 (1)(f)). Data subjects have an unconditioned right to object to profiling for purposes of direct marketing. Data subjects can also object to profiling for statistical purposes. The right to object should have a vast scope with regard to AI-based processing. The key issue would be to make it easier to exercise this right.

제 21 조는 처리가 공공의 이익 (제6조 (1) (e))에 근거하거나 합법적인 사적 이익에 근거할 때 (개인적 근거), 프로파일링에 반대하는 능력을 구체적으로 다루고있다 (제6조 (1) (f)). 데이터 주체는 직접 마케팅을 목적으로 프로파일링에 대해 이의를 제기할 수 있는 조건이 없습니다. 데이터 주제는 통계 목적으로 프로파일링에 반대할 수도 있습니다. 이의 제기 권리는 AI 기반 처리와 관련하여 광범위한 범위를 가져야합니다. 중요한 문제는 이 권리를 보다 쉽게 행사할 수 있도록 하는 것입니다.

4.2.7. Automated decision-making

Article 22 on automated decision-making is highly relevant to AI, since automated decisions today are indeed taken through AI-based systems. According to the interpretation suggested above, Article 22(1) prohibits any completely automated decisions based on profiling and having legal or significant effects on the data subject. Article 22(2) introduces broad exceptions to the prohibition, allowing for automated decisions to be introduced by contract, law or consent.

자동화된 의사 결정에 관한 제22 조는 오늘날 AI 기반 시스템을 통해 자동화된 의사 결정을 하기 때문에 AI와 매우 관련이 있습

니다. 위에서 제안한 해석에 따르면, 제22 (1) 조는 데이터 주체에 대한 프로파일링 및 법적 또는 중대한 영향을 미치는 완전히 자동화된 결정을 금지합니다. 제22 (2) 조는 금지에 대한 광범위한 예외를 도입하여 계약, 법률 또는 동의에 의해 자동 결정을 도입할 수 있습니다.

This provision raises a number of issues, from determining when a decision is 'based solely on automated processing' to establishing whether its effects 'significantly' affect the data subject, to establishing when exceptions apply. Article 22(3) requires suitable safeguard measures to be adopted, 'at least' concerning the data subject's right to obtain human intervention, to express his or her point of view and to contest the decision. This list omits the safeguard consisting of the right to obtain an individualised explanation, which specifies the reasons why an unfavourable decision has been adopted. It also leaves out the requirement that the decision be 'reasonable,' meaning that its input factors and aims are acceptable and its method reliable (see Section 3.1.2 above). Reasonableness also requires that the extent to which certain input factors influence the decision should be proportionate to the causal or at least predictive importance of such factors relative to the legitimate goals being pursued.

이 규정은 의사 결정이 '자동 처리에만 기초한'시기 결정에서 그 영향이 데이터 주체에 '중요하게'영향을 미치는지 여부를 설정하는 것, 예외가 적용되는 시기를 설정하는 것에 이르기까지 많은 문제를 제기합니다. 제22 (3) 조는 인간의 개입을 얻을 수 있는 데이터 주체의 권리에 관한 '적어도'적절한 관점의 보호 조치가 채택되어야 하며, 자신의 관점을 표현하고 결정에 이의를 제기할 것을 요구하고 있다. 이 목록에는 개별 설명을 얻을 수 있는 권리로 구성되는 안전 조치가 생략되어 있으며, 이는 바람직하지 않은 결정이 채택된 이유를 지정합니다. 또한 결정이 '합리적'이어야 한다는 요구 사항을 제외하고, 입력 요소와 목표가 수용 가능하고 그 방법이 신뢰할 수 있음을 의미합니다 (위의 3.1.2 절 참조). 또한 합리성을 위해서는 특정 입력 요소가 결정에 영향을 미치는 정도가 추구하는 합법적인 목표와 관련하여 그러한 요소의 인과적 또는 적어도 예측 중요성에 비례해야 합니다.

4.2.8. AI and privacy by design

A risk-based approach to data-protection focuses on preventing harm, rather than on providing individual data subjects with legal powers over the processing of their data. A key role in this regard is played by Article 25, which, under the heading 'Data protection

by design and by default', requires that technical and organisational measures be adopted to implement data protection principles and integrate safeguards in the processing. With regard to AI, these measures should include controls over the representativeness of training sets, over the reasonableness of the inferences (including the logical and statistical methods adopted) and over the absence of unfairness and discrimination.

데이터 보호에 대한 위험 기반 접근 방식은 개인 데이터 대상에게 데이터 처리에 대한 법적 권한을 제공하는 것이 아니라 피해 예방에 중점을 둡니다. 이와 관련하여 핵심적인 역할은 제25 조에 의해 수행되며, '설계 및 기본 데이터 보호'라는 제목하에 기술 및 조직적 조치를 채택하여 데이터 보호 원칙을 구현하고 처리 과정에서 보호 기능을 통합해야 합니다. AI와 관련하여 이러한 조치에는 훈련 세트의 대표성, 추론의 합리성 (적용된 논리적 및 통계적 방법 포함) 및 불공정 및 차별의 부재에 대한 통제가 포함되어야 합니다.

Appropriate security measures, such as encryption or pseudonymisation, should also prevent unauthorised uses of the data (Article 32 (1)). High risk processing operations are subject to mandatory data protection assessment (Article 35 (1)), a requirement that applies in particular to the 'systematic and

extensive evaluation of personal aspects' for the purpose of automated decision-making including profiling (Article 35 (3)(a)). Article 37 requires that a data protection officer be designated when a 'regular and systematic monitoring of data subjects on a large scale' is envisaged. Articles 40-43, on codes of conduct and certification, although not specifically addressing AI, identify procedures for anticipating and countering risks, and incentivise the adoption of preventive measures that are highly significant to AI.

암호화 또는 가명 화와 같은 적절한 보안 조치는 데이터의 무단 사용을 방지해야 합니다 (제32조 제1 항). 위험도가 높은 처리 작업은 필수 데이터 보호 평가 (제35조 제1 항)의 적용을 받으며, 특히 프로파일링을 포함한 자동화된 의사 결정을 위해 '개인적 측면의 체계적이고 광범위한 평가'에 적용됩니다 (제35조 (3) (a)). 제 37 조는 '대규모의 데이터 주체에 대한 규칙적이고 체계적인 모니터링'이 구상될 때 데이터 보호 책임자를 지명하도록 요구하고 있다. 행동 강령 및 인증에 관한 조항 40-43은 AI를 구체적으로 다루지는 않지만 위험을 예측하고 대응하는 절차를 식별하며 AI 에 매우 중요한 예방 조치의 채택을 장려합니다.

4.2.9. AI, statistical processing and scientific research

In combination with big data, AI can provide useful results for science and statistical purposes (e.g. in medicine for diagnosis or prognosis, in the social sciences for understanding economic or political behaviour, in business for detecting consumer tastes and trends). These results have a general nature (they are not attached to particular individuals); therefore, they do not count as personal data. However, statistical and scientific processing also affects individuals, by exposing their data to security risks and abuse. Moreover, statistical results may indirectly affect individuals, since they provide information – possibly inaccurate or misleading – concerning the groups to which an individual belongs. The GDPR allows repurposing for scientific and statistical processing (under appropriate safeguards). The permission to engage in scientific and in particular statistical processing may enable beneficial uses of AI and big data in Europe, even though we need to take the implications for data subjects' rights and for competition into account.

AI는 빅 데이터와 결합하여 과학 및 통계 목적에 유용한 결과를 제공할 수 있습니다 (예 : 진단 또는 예후 의학, 사회 과학, 경제 또는 정치 행동 이해, 소비자의 취향 및 동향 감지를 위한 비즈니스). 이러한 결과는 일반적인 성격을 지닙니다 (특정 개인에게는

첨부되지 않음). 따라서 개인 데이터로 계산되지 않습니다. 그러나 통계 및 과학적 처리는 데이터를 보안 위험 및 남용에 노출시킴으로써 개인에게 영향을 미칩니다. 또한 통계 결과는 개인이 속한 그룹에 관한 정보 (정확하지 않거나 오해의 소지가 있음)를 제공하므로 개인에 간접적으로 영향을 줄 수 있습니다. GDPR은 과학적 및 통계적 처리를 위한 용도 변경을 허용합니다 (적절한 보호 조치하에). 과학 및 특히 통계 처리에 참여할 수 있는 권한은 비록 데이터 주체의 권리와 경쟁에 대한 영향을 고려해야 하더라도 유럽에서 AI와 빅 데이터의 유익한 사용을 가능하게 할 수 있습니다.

4.3. AI and GDPR compatibility

In this section, the main results of the foregoing review will be summarised. It will be argued that policy options exist for ensuring that innovation in the field of AI is not stifled and remains responsible. Guidelines for controllers are needed, though there is no urgent need to make broad changes to the GDPR

이 섹션에서는 전술한 검토의 주요 결과가 요약됩니다. AI 분야의 혁신이 방해받지 않고 책임을 유지하기 위해 정책 옵션이 존재한다고 주장할 것이다. GDPR을 광범위하게 변경할 필요는 없지만,

컨트롤러에 대한 지침이 필요합니다.

4.3.1. No incompatibility between the GDPR and AI and big data

GDPR과 AI 및 빅 데이터 간에 비 호환성 없음

It has been argued that the GDPR would be incompatible with AI and big data, given that the GDPR is based on principles – purpose limitation, data minimisation, the special treatment of 'sensitive data', the limitation on automated decisions – that are incompatible with the extensive use of AI, as applied to big data. As a consequence, the EU would be forced to either renounce application of the GDPR or lose the race against those information-based economies – such as the USA and China – that are able make full use of AI and big data.117

GDPR이 목적 제한, 데이터 최소화, '민감한 데이터'의 특수한 처리, 자동 결정에 대한 제한 등의 원칙에 기반한다는 점을 감안할 때 GDPR은 AI 및 빅 데이터와 호환되지 않을 것이라고 주장했습니다. 빅 데이터에 적용되는 AI의 광범위한 사용. 결과적으로 EU는 GDPR적용을 포기하거나 AI 및 빅 데이터를 최대한 활용할 수 있는 미국 및 중국과 같은 정보 기반 경제와의 경쟁에서 벗어날 수

밖에 없었습니다.117

Contrary to this opinion, this report shows that it is possible – and indeed likely – that the GDPR will be interpreted in such a way as to reconcile both desiderata: protecting data subjects and enabling useful applications of AI. It is true that the full deployment of the power of AI and big data requires collecting vast quantities of data concerning individuals and their social relations, and that it also requires processing of such data for purposes that were not fully determined at the time the data were collected. However, there are ways to understand and apply the data protection principles that are consistent with the beneficial uses of AI and big data.

이 견해와는 반대로, 이 보고서는 GDPR이 데이터 주제를 보호하고 AI의 유용한 응용을 가능하게 하는 데 필요한 두 가지를 조정하는 방식으로 해석될 수 있음을 보여줍니다. 인공 지능과 빅 데이터의 모든 기능을 완전히 활용하려면 개인 및 사회적 관계에 관한 방대한 양의 데이터를 수집해야 하며 데이터를 수집할 당시 완전히 결정되지 않은 목적으로 이러한 데이터를 처리해야 합니다. 모은. 그러나 AI 및 빅 데이터의 유익한 사용과 일치하는 데이터 보호 원칙을 이해하고 적용하는 방법이 있습니다.

117 Zarsky (2017), Hildebrandt (2015)

The requirement that consent be specific and purpose limitation be respected should be linked to a flexible application of the idea of compatibility, that allows for the reuse of personal data when this is not incompatible with the purpose for which the data were collected. As noted above, the legal basis laid down in Article (6)(1)(f), namely, that the processing should serve a legitimate interest that is not outweighed by the interests of the data subjects, in combination with a compatibility assessment of the new uses, may provide sufficient grounds on which to make reuse permissible. Moreover, as noted above, reuse for statistical purposes is assumed to be compatible, and thus would in general be admissible (unless it involves unacceptable risks for the data subject).

동의가 구체적이고 목적 제한이 존중되어야 한다는 요구 사항은 개인 정보가 데이터 수집 목적과 호환되지 않을 때 개인 데이터를 재사용할 수 있는 호환성 아이디어의 유연한적용과 연계되어야 합니다. 위에서 언급한 바와 같이, 제6조 (1) (f) 조에 규정된 법적 근거, 즉, 처리는 데이터 주체의 이익보다 중요하지 않은 합법적인 이익을 제공해야 한다. 새로운 용도는 재사용이 가능하도록 충분한 근거를 제공할 수 있습니다. 더욱이, 위에서 언급한 바

와 같이, 통계 목적의 재사용은 호환되는 것으로 가정되며, 따라서 (데이터 주체에 대해 수용할 수 없는 위험이 없는 한) 일반적으로 허용될 수 있다.

Even the principle of data-minimisation can be understood in such a way as to enable a beneficial application of AI. This may involve in some context reducing the 'personality' of the data, namely the ease with which they can be connected to the individuals concerned, with measures such as pseudonymisation, rather than focusing on the amount of personal data to be preserved. This also applies to re-identification, the possibility of which should not exclude the processing of data which can be re-identified, but rather requires viewing re-identification as the creation of new personal data, which should be subject to all applicable rules, and strictly prohibited unless all conditions for the lawful collection of personal data are met, and should also be subject to the compatibility test.

데이터 최소화의 원칙조차도 AI의 유익한적용을 가능하게 하는 방식으로 이해될 수 있습니다. 이것은 어떤 맥락에서 데이터의 '인격', 즉 보존될 개인 데이터의 양에 초점을 맞추기보다는 가명화와 같은 조치를 통해 관련 개인과 쉽게 연결할 수 있는 것을 감소시키는 것을 포함할 수 있다. 이는 재식별에도 적용되며, 재

식별될 수 있는 데이터의 처리를 배제해서는 안되며, 새로운 개인 데이터의 생성으로 재식별을 검토해야 하며, 적용가능한 모든 규칙에 따라야합니다. 합법적인 개인 데이터 수집 조건을 모두 충족하지 않으면 엄격히 금지되며 호환성 테스트를 받아야합니다.

The information requirements established by the GDPR can also be met with regard to AI-based processing, even though the complexity of AI systems represents a difficult challenge. The information concerning AI-based applications should enable the data subjects to understand the purpose of the processing and its limits, without going into technical details.

AI 시스템의 복잡성이 어려운 과제 일지라도 GDPR에 의해 확립된 정보 요구 사항은 AI 기반 처리와 관련하여 충족될 수 있습니다. AI 기반 응용 프로그램에 관한 정보는 데이터 주체가 기술적인 세부 사항으로 가지 않고 처리의 목적과 한계를 이해할 수 있도록 해야 합니다.

The GDPR allows for inferences based on personal data, including profiling, but only under certain conditions and so long as the appropriate safeguards are adopted.

GDPR은 프로파일링을 포함하여 개인 데이터를 기반으로 추론을 허용하지만 특정 조건 하에서 적절한 보호 조치가 채택되는 한에만 가능합니다.

The GDPR does not exclude automated decision-making, as it provides for ample exceptions – contract, law or consent – to the general prohibition set forth in Article 22(1). Uncertainties exist concerning the extent to which an individual explanation should be provided to the data subject. Uncertainties also exist about the extent to which reasonableness criteria may apply to automated decisions.

GDPR은 자동화된 의사 결정을 배제하지 않습니다. 계약, 법률 또는 동의 등 제22조 (1)에 명시된 일반 금지에 대한 예외는 충분합니다. 데이터 주제에 개별 설명을 제공해야하는 정도와 관련하여 불확실성이 존재합니다. 자동화된 결정에 합리성 기준이 적용될 수 있는 정도에 대한 불확실성이 존재합니다.

The GDPR provisions on preventive measures, and in particular those concerning privacy by design and by default should also not hinder the development of AI applications, if correctly designed and implemented, although they may entail some additional costs.

Finally, the possibility of using the data for statistical purposes – with appropriate security measures, proportionate to the risks, which should include at least pseudonymisation – opens wide spaces for the processing of personal data in ways that do not involve the inference of personal data.

예방 조치에 대한 GDPR 조항, 특히 설계 및 기본적으로 개인 정보 보호와 관련된 조항은 AI 응용 프로그램의 개발을 방해하지 않아야 합니다.

마지막으로, 가명 화를 포함해야 하는 위험에 비례한적절한 보안 조치와 함께 통계 목적으로 데이터를 사용할 가능성은 개인 데이터의 추론을 포함하지 않는 방식으로 개인 데이터를 처리할 수 있는 넓은 공간을 열어줍니다.

4.3.2. GDPR prescriptions are often vague and open-ended

GDPR 처방전은 종종 모호하고 개방적입니다.

In the previous sections it has been argued that the GDPR allows for the development of AI and big data applications that successfully balance data protection and other social and economic interests.

이전 섹션에서는 GDPR을 통해 데이터 보호와 기타 사회적, 경제적 이익의 균형을 맞추는 AI 및 빅 데이터 응용 프로그램을 개발할 수 있다고 주장했습니다.

However, this does not mean that such a balance can be found by referring to the GDPR alone. The GDPR rules need to be interpreted and consistently implemented, and appropriate guidance needs to be provided on concrete implication of the GDPR for particular processing activities.

그러나 이것이 GDPR만을 언급함으로써 그러한 균형을 찾을 수 있다는 것을 의미하지는 않습니다. GDPR 규칙은 해석되고 일관되게 구현되어야 하며 특정 처리 활동에 대한 GDPR의 구체적인 의미에 대한적절한 지침이 제공되어야 합니다.

The GDPR indeed abounds in vague clauses and open standards. Among those pertaining to the issues here addressed, the

following can be mentioned: the identifiability of the data subject (Article 4 (1)), the freeness of consent (Article (4)(11), the compatibility of further processing with the original (Article 5(1)(c)), the necessity of the data relative to their purpose (Article 5 (1)(c)), the legitimacy of the controller's interests and their non-overridden importance (Article 6(1)(f)), the meaningfulness of the information about the logic involved in automated decision-making (Articles 13(2)(f) and 14 (2)(g)), the suitableness of the safeguard measures to be adopted for automated decision-making (Article 22 (2)), and the appropriateness of the technical and organisational measures for data protection by design and by default (Article 25).

GDPR은 실제로 모호한 조항과 공개 표준이 풍부합니다. 여기에 언급된 문제들과 관련된 것들 중에서, 데이터 주체의 식별 가능성 (제4조 (1)), 동의의 자유 (제4조 (11), 원본과 추가 처리의 호환성) (제5조 (1) (c)), 목적에 관한 자료의 필요성 (제5조 (1) (c)), 지배인의 이익의 정당성 및 재정의되지 않은 중요성 (제6조 (1) (f)), 자동화된 의사 결정과 관련된 논리에 관한 정보의 의미 (제13조 (2) (f) 및 14 (2) (g)), 자동화된 의사 결정을 위해 채택된 보호조치의 적합성 설계 및 기본적으로 데이터 보호를 위한 기술 및 조직적 조치의 적절성 (제22조 제2 항).

In various cases, the interpretation of undefined GDPR standards requires balancing competing interests: it requires determination of whether a certain processing activity, and the measures adopted are justified on balance, i.e., whether the controller's interests in processing the data and in (not) adopting certain measures are outweighed by the data subjects' interests in not being subject to the processing or in being protected by additional or stricter measures. These assessments depend on both (a) uncertain normative judgements on the comparative importance of the impacts on the interests at stake and (b) uncertain forecasts concerning potential future risks. In the case of AI and big data applications the uncertainties involved in applying indeterminate concepts and balancing competing interests are aggravated by the novelty of the technologies, their complexities, the broad scope of their individual and social effects.

다양한 경우에, 정의되지 않은 GDPR 표준의 해석은 경쟁적 이해관계의 균형을 필요로 한다 : 특정 처리 활동과 채택된 측정치가 균형에 따라 정당화되는지, 즉 데이터 처리에 대한 제어기의 이해와 채택 (비) 특정 조치는 처리 대상이 아니거나 추가적이거나 엄격한 조치로 보호되는 데이터 주체의 이익보다 중요합니다. 이러한 평가는 (a) 이해 관계자 이익에 미치는 영향의 비교 중요성에 대한 불확실한 규범적 판단과 (b) 잠재적 미래 위험에 관한 불확

실한 예측 모두에 달려 있습니다. AI 및 빅 데이터 애플리케이션의 경우, 기술의 복잡성, 복잡성, 개인 및 사회적 영향의 넓은 범위로 인해 불확실한 개념을 적용하고 경쟁 관계의 균형을 맞추는 데 따른 불확실성이 더욱 심화됩니다.

It is true that the principles of risk-prevention and accountability potentially direct the processing of personal data toward being a 'positive sum' game (where the advantages of the processing, when constrained by appropriate risk-mitigation measures, outweigh its possible disadvantages), and enable experimentation and learning, avoiding the over- and under-inclusiveness issues involved in the applications of strict rules. On the other hand, by requiring controllers to apply these principles, the GDPR offloads the task of establishing how to manage risk and find optimal solutions onto controllers, a task which may be both challenging and costly. The stiff penalties for non-compliance, when combined with the uncertainty as to what is required for compliance, may constitute a novel risk, which, rather than incentivising the adoption of adequate compliance measure, may prevent small companies from engaging in new ventures.

위험 예방 및 책임의 원칙은 잠재적으로 개인 데이터의 처리를 '양의 합'게임으로 향하게 할 가능성이 있습니다 (처리의 장점은

적절한 위험 완화 조치로 제한될 때 가능한 단점보다 중요 함). 엄격한 규칙의 적용과 관련된 과도 및 불완전 성 문제를 피하면서 실험과 학습을 가능하게 합니다. 반면에 GDPR은 컨트롤러가 이러한 원칙을 적용하도록 요구함으로써 위험을 관리하고 컨트롤러에 대한 최적의 솔루션을 찾는 방법을 수립하는 작업을 오프로드합니다. 이 작업은 까다 롭고 비용이 많이 드는 작업입니다. 비준수에 대한 엄격한 처벌은 규정 준수에 필요한 것에 대한 불확실성과 결합될 때, 새로운 규정 준수 조치의 채택을 장려하기보다는 소기업이 새로운 벤처에 참여하는 것을 방해할 수 있는 새로운 위험을 구성할 수 있습니다.

No easy solution is available in the hyper-complex and rapidly evolving domain of AI technologies: rules may fail to enable opportunities and counter risks, but the private implementation of open standard, in the absence of adequate legal guidance, may also be unsatisfactory:

복잡하고 빠르게 진화하는 AI 기술 영역에서는 쉬운 솔루션을 사용할 수 없습니다. 규칙은 기회를 제공하지 못하고 위험을 방지할 수 있지만 적절한 법적 지침이 없으면 공개 표준의 개인 구현도 만족스럽지 않을 수 있습니다.

[Giving] appropriate content to the law often requires effort, whether in analysing a problem, resolving value conflicts, or acquiring empirical knowledge. […] [I]ndividuals contemplating behavior that may be subject to the law will find it more costly to comply with standards, because it generally is more difficult to predict the outcome of a future inquiry (by the adjudicator, into the law's content) than to examine the result of a past inquiry. They must either spend more to be guided properly or act without as much guidance as under rules.118

법에 적합한 내용을 제공하는 것은 종종 문제 분석, 가치 갈등 해결 또는 경험적 지식 습득에 노력이 필요합니다. […] [I] 법의적용을 받는 행동을 고려하는 개인들은 표준에 따르는 것이 비용이 많이 든다는 것을 알게 될 것입니다. 왜냐하면 일반적으로 미래의 질의 결과를 판단하기가 더 어렵 기 때문입니다 (심사관이 법의 내용에 대해) 과거의 문의 결과를 검토하는 것보다 그들은 제대로 인도하기 위해 더 많은 비용을 지출하거나 규칙에 따라 많은 지도없이 행동해야 합니다.118

Thus, the way in which the GDPR will affect successful applications

of AI and big data in Europe will also depend on what guidance data protection bodies – and more generally the legal system – will be able to provide to controllers and data subjects. This would diminish the cost of legal uncertainty and would direct companies – in particular small ones that mostly need advice – to efficient and data protection-compliant solutions. Appropriate mechanisms may need to be devised, such as an obligation to notify data protection authorities when new applications based on profiling are introduced, but also the possibility to ask for preventive, non-binding, indications on whether and how such applications should be developed, and with what safeguards.

따라서 GDPR이 AI의 성공적인 적용 및 유럽의 빅 데이터에 영향을 미치는 방식은 또한 데이터 보호 기관 (일반적으로 법률 시스템)이 컨트롤러 및 데이터 주제에 제공할 수 있는 지침에 따라 달라집니다. 이는 법적 불확실성 비용을 감소시키고 회사, 특히 조언이 필요한 소규모 기업을 효율적이고 데이터 보호 준수 솔루션으로 안내합니다. 프로파일링을 기반으로 하는 새로운 응용 프로그램이 도입될 때 데이터 보호 기관에 알리는 의무와 같은 적절한 메커니즘을 고안해야 할 뿐만 아니라 그러한 응용 프로그램의 개발 여부와 방법에 대한 예방적, 구속력이 없는 표시를 요구할 수 있는 가능성, 그리고 무엇을 보호하는가.

4.3.3. Providing for oversight and enforcement

감독 및 집행 제공

As noted above, AI applications may affect not only the concerned individuals but also society at large. Even applications based on correct statistical principles, which do not target protected categories, and which adopt appropriate security measures may still impose undue burden on certain categories of citizens, or anyway have negative social impacts. Oversight by competent authorities needs to be complemented by the support of civil society.

위에서 언급했듯이 AI 응용 프로그램은 관련 개인 뿐만 아니라 사회에도 영향을 줄 수 있습니다. 올바른 통계 원칙에 근거한 보호 범주를 대상으로 하지 않고 적절한 보안 조치를 채택하는 응용 프로그램조차도 특정 범주의 시민에게 과도한 부담을 주거나 어쨌든 사회적 영향을 미칠 수 있습니다. 권한있는 당국의 감독은 시민 사회의 지원으로 보완되어야 합니다.

118 Kaplow (1992, 621).

As collective interests, power relations, and societal arrangements are at stake, a broad public debate and the involvement of representative institutions is also needed.

집단의 이익, 권력 관계 및 사회적 약정이 위태로워짐에 따라 광범위한 대중 토론과 대표 기관의 참여도 필요하다.

Collective enforcement is also a key issue that is not answered by the GDPR, which still relies on individual action by the concerned data subjects. An important improvement toward an effective protection could consist in enabling collective actions for injunctions and compensation. It has indeed been observed that US courts have been unable so far to deal satisfactorily with privacy harms, since on the one hand they rely on old-fashioned theories requiring compensable harms to be concrete, actual and directly caused by the defendant, and on the other hand they are unable to address a very high numbers of similar claims, each having small monetary value.119 In Europe, data protection authorities can provide an alternative and easier avenue to enforcement, but nevertheless, the damaged parties have to rely on the judiciary to obtain compensation from privacy harms, which also includes non-material harm (Article 82). Thus, effective protection is dependent on the data subject's ability to engage in lawsuits. The possibility

for multiple data subjects to merge similar claims to share cost and engage more effectively with the law is necessary to make legal remedies available to data subjects.

집단 집행은 GDPR에 의해 답변되지 않은 주요 문제이며, 여전히 관련 데이터 주체의 개별 행동에 의존합니다. 효과적인 보호를 향한 중요한 개선은 금지 명령과 보상에 대한 집단 행동을 가능하게 하는 것으로 구성될 수 있습니다. 실제로 미국 법원은 지금까지 프라이버시 피해를 만족스럽게 처리할 수 없었으며, 한편으로는 피고에 의해 구체적이고 실제적이며 직접적으로 야기되는 보상 가능한 피해를 요구하는 구식 이론에 의존하기 때문이다. 반면에, 그들은 각각 작은 금전적 가치를 갖는 매우 많은 수의 유사한 주장을 다룰 수 없다.119 유럽에서, 데이터 보호 당국은 집행에 대안적이고 더 쉬운 길을 제공할 수 있지만 그럼에도 불구하고 피해 당사자는 사법부에 의존해야 한다 비 물질적 피해를 포함하여 프라이버시 피해로부터 보상을 얻는 것 (제82조). 따라서 효과적인 보호는 데이터 주체가 소송에 참여할 수 있는 능력에 달려 있습니다. 데이터 주체가 법적 구제책을 이용할 수 있도록 하려면 여러 데이터 주체가 유사한 클레임을 병합하여 비용을 공유하고 법률에 보다 효과적으로 참여할 수 있어야 합니다.

The Court of Justice has recently denied that a consumer can combine his or her individual data protection claim with claims

concerning other consumers involved in similar cases.120 In particular, it has affirmed that Max Schrems could exercise, in the courts of his domicile, only his individual claim against Facebook for data protection violations. He could not bring, before the same court, claims for similar violations that had been assigned to him by other data subjects. Perhaps the proposed directive on collective redress for consumers,121 currently under interinstitutional negotiation122, could present an opportunity to enable collective actions in the context of data protection.

법원은 최근 소비자가 자신의 개인 정보 보호 주장을 유사한 경우에 관련된 다른 소비자에 대한 청구와 결합할 수 있음을 거부했습니다. 데이터 보호 위반에 대한 Facebook의 개인 주장. 같은 법원에 다른 데이터 주체에 의해 부여된 유사한 위반에 대한 주장을 제기할 수 없었습니다. 아마도 현재 기관 간 협상 (122) 하에서 소비자에 대한 집단적 보상에 대한 제안된 지침은 데이터 보호의 맥락에서 집단적 행동을 가능하게 하는 기회를 제시할 수 있다.

4.4. Final considerations: some policy proposals on AI and the GDPR

최종 고려 사항 : AI와 GDPR에 대한 일부 정책 제안

In the following, the main conclusions of this report on the relations between AI and the processing of personal data are summarised.

다음에는 AI와 개인 데이터 처리 간의 관계에 대한이 보고서의 주요 결론이 요약되어 있습니다.

- The GDPR generally provides meaningful indications for data protection relative to AI applications.

  GDPR은 일반적으로 AI 응용 프로그램과 관련하여 데이터 보호에 대한 의미있는 표시를 제공합니다.

- The GDPR can be interpreted and applied in such a way that it does not hinder beneficial application of AI to personal data, and that it does not place EU companies at a disadvantage in comparison with non-European competitors.

  GDPR은 AI가 개인 데이터에 유리하게 적용되는 것을 방해하지 않으며 EU 기업이 비 유럽 경쟁 업체와 비교할 때 불리한 점이 되지 않도록 해석하고 적용할 수 있습니다.

- Thus, GDPR does not seem to require any major change in order to address AI.

  따라서 GDPR은 AI를 다루기 위해 큰 변화가 필요하지 않은 것으로 보인다.

- That said, a number of AI-related data protections issues are not explicitly answered in the GDPR, which may lead to uncertainties and costs, and may needlessly hamper the development of AI applications.

  GDPR에서 다수의 AI 관련 데이터 보호 문제가 명시적으로 답변되지 않아 불확실성과 비용이 발생할 수 있으며 AI 응용 프로그램의 개발을 불필요하게 방해할 수 있습니다.

- Controllers and data subjects should be provided with guidance on how AI can be applied to personal data consistently with the GDPR, and on the available technologies for doing so. This can prevent costs linked to legal uncertainty, while enhancing compliance.

  컨트롤러와 데이터 주체는 AI가 GDPR과 일관되게 개인 데이터에 어떻게 적용될 수 있는지, 그리고 사용 가능한 기술에 대한 지침을 제공해야 합니다. 이를 통해 법적 불확실성과 관련된 비용을 방지하면서 규정 준수를 강화할

수 있습니다.

- Providing adequate guidance requires a multilevel approach, which involves civil society, representative bodies, specialised agencies, and all stakeholders.

  적절한 지침을 제공하기 위해서는 시민 사회, 대표 단체, 전문 기관 및 모든 이해 관계자가 참여하는 다단계 접근이 필요합니다.

- A broad debate is needed, involving not only political and administrative authorities, but also civil society and academia. This debate needs to address the issues of determining what standards should apply to AI processing of personal data, particularly to ensure the acceptability, fairness and reasonability of decisions on individuals

  정치 및 행정 당국 뿐만 아니라 시민 사회 및 학계와 관련된 광범위한 토론이 필요합니다. 이 토론은 개인 데이터의 AI 처리에 적용할 표준을 결정하는 문제, 특히 개인에 대한 의사 결정의 수용 가능성, 공정성 및 합리성을 보장하는 문제를 해결해야 합니다..

- The political debate should also address what applications are to be barred unconditionally, and which may instead be admitted only under specific circumstances. Legally

binding rules are needed to this effect, since the GDPR is focused on individual entitlements and does not take the broader social impacts of mass processing into account.

정치적 논쟁은 무조건적으로 어떤 신청이 금지되어야 하며, 특정 상황에서만 허용될 수 있는 문제를 다루어야 합니다. GDPR은 개인 자격에 초점을 맞추고 대량 처리의 사회적 영향이 더 크지 않기 때문에 법적 구속력이 있는 규칙이 이 효과에 필요합니다.

- Discussion of a large set of realistic examples is needed to clarify which AI applications are on balance socially acceptable, under what circumstances and with what constraints. The debate on AI can also provide an opportunity to reconsider in depth, more precisely and concretely, some basic ideas of European law and ethics, such as acceptable and practicable ideas of fairness and non-discrimination.

  어떤 상황에서 어떤 제약 조건으로 사회적으로 수용 가능한 AI 응용 프로그램인지를 명확하게 설명하기 위해 현실적인 일련의 사례에 대한 논의가 필요합니다. AI에 대한 토론은 또한 공정하고 차별이 없는 수용 가능하고 실용 가능한 아이디어와 같은 유럽 법과 윤리에 관한 몇 가지 기본 아이디어를 보다 정밀하고 구체적으로 재고할 수 있

는 기회를 제공할 수 있습니다.

- Political authorities, such as the European Parliament, the European Commission and the Council could provide general open-ended soft law indications about the values at stake and ways to achieve them.

  유럽 의회, 유럽위원회, 협의회와 같은 정치 당국은 위기에 처한 가치와 이를 달성하는 방법에 대한 일반 개방형 소프트 법률 표시를 제공할 수 있습니다.

- Data protection authorities, and in particular the Data Protection Board, should provide controllers with guidance on the many issues for which no precise answer can be found in the GDPR, which could also take the form of soft law instruments designed with a dual legal and technical competence.

  데이터 보호 당국, 특히 데이터 보호위원회는 컨트롤러에 GDPR에서 정확한 답변을 찾을 수 없는 많은 문제에 대한 지침을 제공해야 하며, 이중 법률과 기술적 능력.

- National Data Protection Authorities should also provide guidance, in particular when contacted for advice by controllers, or in response to data subjects' queries.

National Data Protection Authorities는 특히 컨트롤러의 조언을 구하거나 데이터 주체의 질문에 대한 답변을 받을 때 지침을 제공해야 합니다.

- The fundamental data protection principles – especially purpose limitation and minimisation – should be interpreted in such a way that they do not exclude the use of personal data for machine learning purposes. They should not preclude forming training sets and building algorithmic models, whenever the resulting AI systems are socially beneficial, and compliant with data protection rights.

  기본 데이터 보호 원칙, 특히 목적 제한 및 최소화는 머신러닝 목적으로 개인 데이터의 사용을 배제하지 않는 방식으로 해석되어야 합니다. 결과 AI 시스템이 사회적으로 유리하고 데이터 보호 권한을 준수할 때마다 훈련 세트 형성 및 알고리즘 모델 구축을 배제해서는 안됩니다.

- The use of personal data in a training set, for the purpose of learning general correlations and connection, should be distinguished from their use for individual profiling, which is about making assessments of individuals.

  일반적인 상관 관계 및 연결을 학습하기 위해 훈련 세트

에서 개인 데이터를 사용하는 것은 개인을 평가하는 데 사용되는 개인 프로파일링 과는 구별되어야 합니다.

- The inference of new personal data, as is done in profiling, should be considered as creation of new personal data, when providing an input for making assessments and decisions. The same should apply to the re-identification of anonymous or pseudonymous data.

  프로파일링에서와 같이 새로운 개인 데이터의 추론은 평가 및 결정을 위한 입력을 제공할 때 새로운 개인 데이터의 생성으로 간주되어야 합니다. 익명 또는 유사 데이터의 재식별에도 동일하게 적용됩니다.

- Both should be subject to the GDPR constraints on the collection of new data.

  둘 다 새로운 데이터 수집에 대한 GDPR 제약 조건을 따라야합니다.

- Guidance is needed on profiling and automated decision-making. It seems that an obligation of reasonableness – including normative and reliability aspects – should be imposed on controllers engaging in profiling, mostly, but not only when profiling is aimed at automated decision-making. Controllers should also be under an obligation to

provide individual explanations, to the extent that this is possible according to the adopted AI technology and reasonable according to costs and benefits. The explanations may be high-level, but they should still enable users to contest detrimental outcomes.

프로파일링 및 자동 의사 결정에 지침이 필요합니다. 규범 및 신뢰성 측면을 포함한 합리성의 의무는 프로파일링이 자동화된 의사 결정을 목표로 할 때 뿐만 아니라 프로파일링에 관여하는 컨트롤러에 적용되어야 하는 것으로 보입니다. 또한 컨트롤러는 채택된 AI 기술에 따라 가능하고 비용과 혜택에 따라 합리적으로 개인 설명을 제공할 의무가 있어야 합니다. 설명은 수준이 높을 수 있지만 사용자는 여전히 해로운 결과에 이의를 제기할 수 있어야 합니다.

- It may be useful to establish obligations to notify data protection authorities of applications involving individualised profiling and decision-making, possibly accompanied with the possibility of requesting indications on data-protection compliance.

  개인화된 프로파일링 및 의사 결정과 관련된 응용 프로그램에 대해 데이터 보호 기관에 통지해야 할 의무를 설정하는 것이 유용할 수 있으며, 데이터 보호 준수에 대한

표시를 요청할 가능성이 있습니다.

- The content of the controllers' obligation to provide information (and the corresponding rights of data subjects) about the 'logic' of an AI system need to be specified, with appropriate examples, with regard to different technologies.

  AI 시스템의 '논리'에 관한 정보 (및 데이터 주체의 해당 권리)를 제공해야하는 컨트롤러의 의무 내용은 적절한 기술과 함께 다른 기술과 관련하여 명시되어야 한다.

- It needs to be ensured that the right to opt out of profiling and data transfers can easily be exercised through appropriate user interfaces, possibly in standardised formats.

  표준화된 형식으로 적절한 사용자 인터페이스를 통해 프로파일링 및 데이터 전송을 거부할 수 있는 권한을 쉽게 행사할 수 있어야 합니다.

- Normative and technological requirement concerning AI by design and by defaults need to be specified.

  설계 및 기본적으로 AI에 관한 규범적 및 기술적 요구 사항을 지정해야 합니다.

- The possibility of repurposing data for AI applications that

do not involve profiling – scientific and statistical ones – may be broad, as long as appropriate precautions are in place preventing abusive uses of personal data.

과학적 및 통계적 프로파일링을 포함하지 않는 AI 응용 프로그램의 데이터 용도 변경 가능성은 개인 정보의 남용을 방지하는 적절한 예방 조치가 마련되어 있는 한 광범위할 수 있습니다.

- Strong measures need to be adopted against companies and public authorities that intentionally abuse the trust of data subjects by misusing their personal data, to engage in applications that manipulate data subjects against their interests.

  개인 정보를 오용하여 의도적으로 데이터 주체의 신뢰를 남용하는 회사 및 공공 기관에 대해 강력한 조치를 취하여 관심사에 대해 데이터 주체를 조작하는 응용 프로그램에 참여해야 합니다.

- Collective enforcement in the data protection domain should be enabled and facilitated.

  데이터 보호 영역에서의 집단 시행이 활성화되고 촉진되어야 합니다.

In conclusion, controllers engaging in AI-based processing should endorse the values of the GDPR and adopt a responsible and risk-oriented approach, and they should be able to do so in a way that is compatible with the available technologies and with economic profitability (or the sustainable achievement of public interests). However, given the complexity of the matter and the gaps, vagueness and ambiguities present in the GDPR, controllers should not be left alone in this exercise.

결론적으로 AI 기반 처리에 관여하는 컨트롤러는 GDPR의 가치를 인정하고 책임감 있고 위험 지향적인 접근 방식을 채택해야 하며, 가용 기술과 경제적 수익성과 호환되는 방식으로 그렇게 할 수 있어야 합니다 (또는 공익의 지속 가능한 달성). 그러나 GDPR에 존재하는 사안의 복잡성과 격차, 모호함 및 모호함을 감안할 때, 컨트롤러는 이 연습에서 혼자서는 안됩니다.

Institutions need to promote a broad social debate on AI applications, and should provide high level indications. Data protection authorities need to actively engage a dialogue with all stakeholders, including controllers, processors, and civil society, to develop appropriate responses, based on shared values and

effective technologies. Consistent application of data protection principles, when combined with the ability to use AI technology efficiently, can contribute to the success of AI applications, by generating trust and preventing risks.

기관은 AI 애플리케이션에 대한 광범위한 사회적 토론을 장려해야 하며 높은 수준의 적응증을 제공해야 합니다. 데이터 보호 당국은 공유 가치와 효과적인 기술을 기반으로 적절한 대응을 개발하기 위해 컨트롤러, 프로세서 및 시민 사회를 포함한 모든 이해관계자와 적극적으로 대화해야 합니다. AI 기술을 효율적으로 사용하는 능력과 결합된 데이터 보호 원칙의 일관된 적용은 신뢰를 생성하고 위험을 예방함으로써 AI 응용 프로그램의 성공에 기여할 수 있습니다.

119 Cohen (2019, Ch. 5).

120 Judgment in Case C-498/16 Maximilian Schrems v Facebook Ireland Limited, of 25 January 2018.

121 Proposal for a directive of the European Parliament and of the Council on representative actions for the protection of the collective interests of consumers, COM(2018) 184 final.

122 See European Parliament Legislative train schedule, Area of Justice and Fundamental Rights, Representative actions for the protection of the collective interests of consumers - a New deal for consumers at https://www.europarl.europa.eu/legislative-train/theme-area-of-justice-and-fundamental -rights/file-representative-actions-for-consumers

5. References

AI-HLEG, High-Level Expert Group on Artificial Intelligence (2019). A definition of AI: Main capabilities and scientific disciplines.

AI-HLEG, High-Level Expert Group on Artificial Intelligence (2019). Ethics guidelines for trustworthy AI.

Ashley, K. D. (2017). Artificial Intelligence and Legal Analytics. Cambridge University Press.

Balkin, J. M. (2008). The constitution in the national surveillance state. Minnesota Law Review 93, 1– 25.

Balkin, J. M. (2017). The three laws of robotics in the age of big data. Ohio State Journal Law Journal 78, 1217–241.

Barocas, S. and A. D. Selbst (2016). Big data's disparate impact. California Law Re view 104, 671–732.

Bayer, J., Bitiukova, N., Bard, P., Szakacs, J., Alemanno, A., and Uszkiewicz, E. (2019). Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its member states. Study, Policy Department for Citizens' Rights and Constitutional Affairs, European Parliament.

Bhuta, N., S. Beck, R. Geiss, C. Kress, and H. Y. Liu (2015). Autonomous Weapons Systems: Law, Ethics, Policy. Cambridge University Press. Bostrom, N. (2014). Superintelligence. Oxford University Press.

Bosco, F., Creemers, N., Ferraris, V., Guagnin, D., & Koops, B. J. (2015). Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. In Reforming European data protection law (pp. 3-33). Springer, Dordrecht.

Brynjolfsson, E. and A. McAfee (2011). Race Against the Machine. Digital Frontier Press.

Burr, C. and Cristianini, N. (2019). Can machines read our minds? Minds and Machines 29:461–494.

Calo, M. R. (2012). Against notice skepticism in privacy (and elsewhere). Notre Dame Law Review, 87:1027– 72.

Cate, F. H., P. Cullen, and V. Mayer-Schönberger (2014). Data Protection Principles for the 21st Century:

Revising the 1980 OECD Guidelines. Oxford Internet Institute.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., and Floridi, L. (2018). Artificial intelligence and the 'good society': the US, EU, and UK approach. Science and Engineering Ethics 24:505–528.

Cohen, J. D. (2019). Between Truth and Power. The Legal Constructions of Informational Capitalism. Oxford University Press.

Cristianini, N. (2016a, 23 November). Intelligence rethought: AIs know us, but don't think like us. New Scientist.

Cristianini, N. (2016b, 26 October). The road to artificial intelligence: A case of data over theory. New Scientist.

Cristianini, N. and T. Scantamburlo (2019). On social machines for algorithmic regulation. AI and Society.

De Hert, P. and Gutwirth, S. (2009). Data protection in the case law of Strasbourg and Luxemburg: Constitutionalisation in action. In Gutwirth, S., Poullet, Y., De Hert, P., de Terwangne, C., and Nouwt, S., editors, Reinventing Data Protection? 3–44. Springer.

Edwards, L. and Veale, M. (2019). Slave to the algorithm? Why a

'right to an explanation' is probably not the remedy you are looking for. Duke Law and Technology Review, 16-84.

Floridi, L., J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena (2018). Ai4people– an ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. Minds and Machines 28, 689–707.

Galbraith, J. K. ([1952]1956). American Capitalism: The Concept of Countervailing Power. Houghton Mifflin.

Guidotti, R., A. Monreale, F. Turini, D. Pedreschi, and F. Giannotti (2018). A survey of methods for explaining black box models. ACM Computer Surveys 51 (5) Article 93, 1–4.

Halpern, J. Y. and Hitchcock, C. (2013). Graded causation and defaults. The British Journal for the Philosophy of Science, 1–45.

Harel, D. and Y. Feldman (2004). Algorithmics: The Spirit of Computing. Addison-Wesley.

Hildebrandt, M. (2009). Profiling and AML. In Rannenberg, K., Royer, D., and Deuker, A., editors, The Future of Identity in the Information Society. Challenges and Opportunities. Springer.

Hildebrandt, M. (2014). Location data, purpose binding and contextual integrity: What's the message? In Floridi, L., editor, The protection of information and the right to privacy, 31–62. Springer.

Hildebrandt, M. (2015). Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology. Edgar.

Jobin, A., Ienca, M., and Vayena, E. (2019). Artificial intelligence: the global landscape of ethics guidelines.

Nature Machine Intelligence, 1: 389–399.

Kahneman, D. (2011). Thinking: fast and slow. Allen Lane.

Kamara, I. and De Hert, P. (2019). Understanding the balancing act behind the legitimate interest of the controller ground: A pragmatic approach. In Seligner, E., Polonetsky, J., and Tene, O., editors, The Cambridge Handbook of Consumer Privacy. Cambridge University Press.

Kaplow, L. (1992). Rule vs standards: An economical analysis. Duke Law Journal, 42: 557–629.

Kleinberg, J., J. Ludwig, S. Mullainathan, and C. R. Sunstein (2018). Discrimination in the age of algorithm. Journal of Legal Analysis 10, 113–174.

Kurzweil, R. (1990). The Age of Intelligent Machines. MIT. Kurzweil, R. (2012). How to Create a Mind. Viking.

Licklider, J. C. R. (1960). Man-computer symbiosis. IRE Transactions on Human Factors in Electronics HFE-1 (March), 4–11.

Lippi, M., P. Palka, G. Contissa, F. Lagioia, H.-W. Micklitz, Y. Panagis, G. Sartor, and P. Torroni (2019). Claudette: an automated detector of potentially unfair clauses in online terms of service. Artificial Intelligence and Law.

Lippi, M., Contissa, G., Jablonowska, A., Lagioia, F., Micklitz, H.-W., Palka, P., Sartor, G., and Torroni, P. (2020). The force awakens: Artificial intelligence for consumer law. The journal of Artificial Intelligence Research 67:169 – 190.

Mantelero, A. (2017). Regulating Big Data. The guidelines of the Council of Europe in the context of the European data protection framework. Computer Law and Security Review 33, 584–602.

Mayer-Schönberger, V. and K. Cukier (2013). Big Data. Harcourt.

Mayer-Schönberger, V. and Y. Padova (2016). Regime change? enabling Big Data through Europe's new data protection regulation. Columbia Science and Technology Law Review 17, 315–35.

McAfee, A. and E. Brynjolfsson (2019). Machine, Platform, Crowd. Norton.

Marcus, G. and Davis, E. (2019). Rebooting AI: building artificial intelligence we can trust. Pantheon Books.

Mindell, D. A. (2015). Our Robots, Ourselves: Robotics and the Myths of Autonomy. Penguin.

Nilsson, N. (2010). The Quest for Artificial Intelligence. Cambridge University Press. O'Neil, C. (2016).

Weapons of math destruction: how Big Data increases inequality

and threatens democracy. Crown Business. Pariser, E. (2011). The Filter Bubble. Penguin.

O'Neil, C. (2016). Weapons of math destruction: how big data increases inequality and threatens democracy. Crown Business.

Pariser, E. (2011). The Filter Bubble. Penguin.

Parkin, S. (14 June 2015). Science fiction no more? channel 4's humans and our rogue ai obsessions.

The Guardian.

Pasquale, F. (2019). The second wave of algorithmic accountability. Law and Political Economy.

Pentland, A. (2015). Social Physics: How Social Networks Can

Make Us Smarter. Penguin.

Polanyi, K. ([1944] 2001). The Great Transformation. Beacon Press.

Powles, J. and Nissenbaum, H. (2018). The seductive diversion of 'solving' bias in artificial intelligence.Medium.

Prakken, H. and G. Sartor (2015). Law and logic: A review from an argumentation perspective. Artificial Intelligence 227, 214–45.

Rawls, J. ([1971] 1999). A Theory of Justice. Oxford University Press.

Ruggeri, S., D. Pedreschi, and F. Turini (2010). Integrating induction and deduction for finding evidence of discrimination. Artificial Intelligence and Law 18, 1–43.

Russell, S. J. and P. Norvig (2016). Artificial Intelligence. A Modern

Approach (3 ed.). Prentice Hall.

Sartor, G. (2017). Human rights and information technologies. In R. Brownsword, E. Scotford, and K. Yeung (Eds.), The Oxford Handbook on the Law and Regulation of Technology, pp. 424–450. Oxford University Press.

Stiglitz, J. (2019). People, Power, and Profits. Progressive Capitalism for an Age of Discontent. Norton.

Sunstein, C. R. (2007). Republic.com 2.0. Princeton University Press.

Turing, A. M. ([1951] 1996). Intelligent machinery, a heretical theory. Philosophia Mathematica 4, 256– 60.

van Harmelen, F., V. Lifschitz, and B. Porter (2008). Handbook of Knowledge Representation. Elsevier.

Varian, H. R. (2010). Computer mediated transactions. American Economic Review (2): 100, 1–10.

Varian, H. R. (2014). Beyond Big Data. Business Economics (49), 27–31.

Wachter, S. and B. Mittelstadt (2017). A right to reasonable inferences: Re-thinking data protection law in the age of Big Data and AI. Columbia Business Law Review, 1–130.

Wachter, S., B. Mittelstadt, and L. Floridi (2016). Why a right to explanation of automated decision-

making does not exist in the General Data Protection Regulation. International Data Privacy Law 7, 76–99.

Yeung, K. (2018). 'Hypernudge': Big data as a mode of regulation by design. Communication and Society 20, 118–36.

Zarsky, T. Z. (2017). Incompatible: The GDPR in the age of Big Data. Seton Hall Law Review, 47:995–1020.

Zuboff, S. (2019). The Age of Surveillance Capitalism. Hachette.

This study addresses the relation between the EU General Data Protection Regulation (GDPR) and artificial intelligence (AI). It considers challenges and opportunities for individuals and society, and the ways in which risks can be countered and opportunities enabled through law and technology.

The study discusses the tensions and proximities between AI and data protection principles, such as in particular purpose limitation and data minimisation. It makes a thorough analysis of automated decision-making, considering the extent to which it is admissible, the safeguard measures to be adopted, and whether data subjects have a right to individual explanations. The study then considers the extent to which the GDPR provides for a preventive risk-based approach, focused on data protection by design and by default.

QA-02-20-399-EN-N

This is a publication of the Scientific Foresight Unit (STOA)
EPRS | European Parliamentary Research Service

ISBN: 978-92-846-6771-0
doi:10.2861/293
QA-QA-02-20-399-EN-N

구글번역 편집 : jason Min V1 (20200704)

mikado22001@yahoo.co.kr http://www.kaail.org/ http://aitimes.org/