

ЕВРОПЕЙСКА КОМИСИЯ ЗА ЕФЕКТИВНОСТ НА ПРАВОСЪДИЕТО (СЕРЕЈ)

Европейска етична харта за
използване на изкуствения интелект в
съдебните системи и тяхната среда



Приета на 31-вата пленарна сесия
на СЕРЕЈ (Страсбург, 3-4 декември 2018 г.)



ЕВРОПЕЙСКА КОМИСИЯ ЗА ЕФЕКТИВНОСТ НА ПРАВОСЪДИЕТО (СЕРЕЈ)

Европейска етична харта за
използване на изкуствения интелект в
съдебните системи и тяхната среда

Приета на 31-вата пленарна
сесия
на СЕРЕЈ (Страсбург, 3-4
декември 2018 г.)

Френско издание:
*Commission européenne pour
l'efficacité
de la justice (CEPEJ) – Charte
éthique
européenne d'utilisation de
l'intelligence
artificielle dans les systèmes
judiciaires et leur environnement*

*Становищата, изразени в
този труд, са отговорност на
авторите и не отразяват по
необходимост официалната
политика на Съвета на
Европа.*

Всички молби относно
възпроизвеждането или
превода на целия или част от
този документи следва да
бъдат отправяни към Дирекция
„Комуникация“ (F-67075
Strasbourg Cedex или
publishing@coe.int). Цялата
друга кореспонденция относно
този документ следва да бъде
отправяна към Европейската
комисия за ефективност на
правосъдието (CEPEJ)
cepej@coe.int

Корица и оформление:
Отдел „Изготвяне на документи
и публикации“ (SPDP),
Съвет на Европа
Снимка: Shutterstock

Тази публикация не е
преминала редакционен
преглед от Редакционното
звено на SPDP за коригиране
на технически и граматични
грешки.

© Съвет на Европа, февруари
2019 г.
Отпечатано в Съвета на
Европа

Съдържание

ВЪВЕДЕНИЕ	6
ПЕТТЕ ПРИНЦИПА НА ЕТИЧНАТА ХАРТА ЗА ИЗПОЛЗВАНЕ НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В СЪДЕБНИТЕ СИСТЕМИ И ТЯХНАТА СРЕДА	8
1. Принцип на спазване на основните права: да се гарантира, че проектирането и реализацията на използващи изкуствен интелект инструменти и услуги са съвместими с основните права.	9
2. Принцип на недискриминация: специално да се предотвратява развитието или засилването на каквато и да е дискриминация между физически лица или групи от физически лица.	10
3. Принцип на качество и сигурност: по отношение на обработването на съдебни решения и данни, да се използват удостоверени източници и невеществени данни с модели, изработени по мултидисциплинарен начин, в сигурна технологична среда.	11
4. Принцип на прозрачност, безпристрастност и справедливост: да се правят методите за обработване на данни достъпни и разбираеми, да се разрешават външни одити.	12
5. Принцип „под контрола на потребителя“: да се изключва подход на обвързване и да се гарантира, че потребителите са информирани участници, които имат контрол върху правените избори.	14
ПРИЛОЖЕНИЕ I – ЗАДЪЛБОЧЕНО ИЗСЛЕДВАНЕ НА ИЗПОЛЗВАНЕТО НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В СЪДЕБНИТЕ СИСТЕМИ, ПО-СПЕЦИАЛНО ПРИЛОЖЕНИЯ С ИЗКУСТВЕН ИНТЕЛЕКТ, ОБРАБОТВАЩИ СЪДЕБНИ РЕШЕНИЯ И ДАННИ	15
1. Състояние на използването на алгоритми с изкуствен интелект в съдебните системи на държавите членки на Съвета на Европа	18
2. Обзор на политиките за отворени данни във връзка със съдебните решения в съдебните системи на държавите членки на Съвета на Европа	20
3. Оперативни характеристики на изкуствения интелект (машинно самообучение), прилаган спрямо съдебни решения	32
4. Може ли изкуственият интелект предварително да моделира правно мотивиране?	39
5. Могат ли изкуствените интелекти да обяснят поведението на съдите в ретроспекция?	43
6. Как да се прилага изкуственият интелект в гражданското, търговското и административното правораздаване?	44
7. Въпроси, които са специфични за наказателното правораздаване: превенция на престъпленията, риск от рецидивизъм и оценка на нивото на опасност	52
8. Специфични въпроси във връзка със защитата на личните данни	61
9. Потенциалът и ограниченията на инструментите за прогностично правосъдие	62
10. Необходимостта от задълбочен обществен дебат относно тези	

инструменти преди реализацията на публични политики за разработването им. Спешната необходимост кибернетиката да предостави рамка за разработването на алгоритми за изкуствен интелект, като същевременно зачита основните права 64

ПРИЛОЖЕНИЕ II – КОИ НАЧИНИ НА ИЗПОЛЗВАНЕ НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ В ЕВРОПЕЙСКИТЕ СЪДЕБНИ СИСТЕМИ? 68

Насърчавани начини на използване 69

Възможни начини на използване, изискващи значителни методологични предпазни мерки 69

Начини на използване, които да бъдат обмислени след допълнителни научни проучвания 71

Начини на използване, които да бъдат обмислени с най-крайни резерви 72

ПРИЛОЖЕНИЕ III – РЕЧНИК 74

ПРИЛОЖЕНИЕ IV – КОНТРОЛЕН СПИСЪК ЗА ИНТЕГРИРАНЕ НА ПРИНЦИПИТЕ НА ХАРТАТА ВЪВ ВАШИЯ МЕТОД НА ОБРАБОТВАНЕ 81

КОНТРОЛЕН СПИСЪК ЗА ОЦЕНКА НА ВАШИТЕ МЕТОДИ НА ОБРАБОТВАНЕ 82

Въведение

Като признава нарастващата важност на изкуствения интелект¹ (AI) в нашите съвременни общества и очакваните ползи, които той ще бъде пълноценно използван в служба на ефективността и качеството на правосъдието, СЕРЕJ официално приема петте основни принципа, наименувани „Европейска етична харта за използване на изкуствения интелект в съдебните системи и тяхната среда”.

Хартата е предназначена за публични и частни заинтересовани страни, отговарящи за създаването и разгръщането на използващи изкуствен интелект инструменти и услуги, които включват обработването на съдебни решения и данни (машинно самообучение или всякакви други методи, произтичащи от науката за данните).

Освен това, тя касае вземащите решения в обществената сфера лица, отговарящи за законовата или подзаконовата нормативна уредба, за разработването, проверката или използването на тези инструменти и услуги.

Използването на тези инструменти и услуги в съдебните системи изглежда подобрява ефективността и качеството на правосъдието и следва да се насърчава. То обаче трябва да бъде осъществявано отговорно, като се вземат предвид основните права на лицата, предвидени в Европейската конвенция за правата на човека и Конвенцията за защита на личните данни, и в съответствие с други основни принципи, изложени по-долу, които следва да ръководят оформянето на публичните политики в сферата на правосъдието в тази област.

Обработването на съдебни решения от изкуствен интелект е вероятно, според неговите разработчици, по граждански, търговски и административни дела, да помогне за подобряване на предвидимостта на прилагането на закона и последователността на съдебните решения, при спазване на изложените по-долу принципи. По наказателни дела, използването му трябва да бъде обмислено с най-големи резерви, за да се предотврати дискриминация, основана на чувствителни данни, в съответствие с гаранциите за справедлив процес.

Независимо дали е създаден с цел подпомагане на предоставянето на правен съвет, оказване на помощ за изготвянето или в процеса на вземане на решение, или даване на съвет на потребителя, е от съществено значение обработването да се извършва с прозрачност, безпристрастност и равнопоставеност, удостоверени от външна и независима експертна оценка.

¹ За определението на изкуствен интелект, вж приложения Речник.

Приложение на Хартата

Принципите на Хартата следва да бъдат предмет на редовно приложение, наблюдение и оценка от публични и частни участници, с оглед на постоянно подобряване на практиките.

В това отношение е желателно тези участници да извършва редовен преглед на реализирането на принципите на Хартата, като обясняват, където е уместно, причините за нереализиране или частично реализиране, придружени с план за действие за въвеждане на необходимите мерки.

Независимите органи, посочени в Хартата, биха могли да са отговорни за периодична оценка на нивото на одобряване на принципите на Хартата от всички участници и за предлагане на подобрения за адаптирането ѝ към променящите се технологии и начини на използване на тези технологии.

Петте принципа на Етичната харта за използване на изкуствения интелект в съдебните системи и тяхната среда

- 1** **Принцип на спазване на основните права:** да се гарантира, че проектирането и реализацията на използващи изкуствен интелект инструменти и услуги са съвместими с основните права.
- 2** **Принцип на недискриминация:** специално да се предотвратява развитието или засилването на каквато и да е дискриминация между физически лица или групи от физически лица.
- 3** **Принцип на качество и сигурност:** по отношение на обработването на съдебни решения и данни, да се използват удостоверени източници и невеществени данни с модели, изработени по мултидисциплинарен начин, в сигурна технологична среда.
- 4** **Принцип на прозрачност, безпристрастност и справедливост:** да се правят методите за обработване на данни достъпни и разбираеми, да се разрешават външни одити.
- 5** **Принцип „под контрола на потребителя“:** да се изключва подход на обвързване и да се гарантира, че потребителите са информирани участници, които имат контрол върху правените избори.

1

Принцип на спазване на основните права: да се гарантира, че проектирането и реализацията на използващи изкуствен интелект инструменти и услуги са съвместими с основните права.

■ Обработването на съдебни решения и данни трябва да обслужва ясни цели, в пълно съответствие с основните права, гарантирани от Европейската конвенция за правата на човека (ЕКПЧ) и Конвенцията за защита на личните данни (Конвенцията за защита на лицата при автоматизираната обработка на лични данни, ETS № 108, изменена с протокол за изменение CETS № 223).

■ Когато използващи изкуствен интелект инструменти се използват за разрешаване на спор или като инструмент за оказване на съдействие при вземане на решения в областта на правосъдието или за даване на насоки на обществеността, е важно да се гарантира, че те не подкопават гаранциите за правото на достъп до съдията и правото на справедлив процес (равни процесуални възможности и зачитане на исковия процес).

■ Те следва да бъдат използвани и при надлежно зачитане на принципа на върховенство на закона и независимост на съдиите при вземане на тяхното решение.

■ Поради това следва да се предпочитат подходите *осигуряване на етичност още при проектирането (ethical-by-design)*² или *съобразяване на правата на човека още при проектирането (human-rights-by-design)*. Това означава, че още от фазите на проектиране и учене, изцяло се внедряват норми, забраняващи преки или косвени нарушения на основните ценности, защитавани от конвенциите.

² Етичният избор се прави нагоре по веригата от проектантите на програмата и поради това не е оставен на потребителя.

2

Принцип на недискриминация: специално да се предотвратява развитието или засилването на каквато и да е дискриминация между физически лица или групи от физически лица.

■ С оглед на способността на тези методи на обработване да разкрият съществуваща дискриминация, чрез групиране или класифициране на данни, свързани с физически лица или групи от физически лица, публичните и частните заинтересовани страни трябва да гарантират, че методите не възпроизвеждат или влошават тази дискриминация и че не водят до детерминистични анализи или начини на използване.

■ Особено внимание трябва да се обърне и във фазата на разработване и във фазата на разгръщане, особено когато обработването е пряко или косвено базирано на „чувствителни“ данни. Това би могло да включва твърдян расов или етнически произход, социално-икономическо положение, политически мнения, религиозни или философски вярвания, членство в профсъюз, генетични данни, биометрични данни, здравни данни или данни относно сексуалния живот или сексуалната ориентация. Когато е била установена такава дискриминация, трябва да се обмислят корективни мерки за ограничаване или, ако е възможно, неутрализиране на тези рискове и, също така, повишаване на информираността сред заинтересовани страни.

■ Следва обаче да се насърчи използването на машинно самообучение и мултидисциплинарни научни анализи за противодействие на такава дискриминация.

3

Принцип на качество и сигурност: по отношение на обработването на съдебни решения и данни, да се използват удостоверени източници и невеществени данни с модели, изработени по мултидисциплинарен начин, в сигурна технологична среда.

■ Проектантите на модели на машинно самообучение следва да са в състояние да се ползват в широки мащаби от опита на съответните специалисти от съдебната система (съдии, прокурори, адвокати и т.н.) и изследователи/лектори в областите на правото и обществените науки (например икономисти, социолози и философи).

■ Формирането на смесени проектни екипи в кратки проектни цикли за създаване на функционални модели е едни от организационните методи, които прави възможно да се извадят дивиденди от този мултидисциплинарен подход.

■ Съществуващите етични гаранции следва да бъдат постоянно споделяни от тези проектни екипи и подобрявани чрез обратна връзка.

■ Базираните на съдебни решени данни, които се въвеждат в софтуер, изпълняващ алгоритъм за машинно самообучение, следва да идват от удостоверени източници и не следва да се изменят докато не бъдат реално използвани от самообучителния механизъм. Поради това, целият процес трябва да бъде проследим, за да се гарантира, че не е настъпило изменение, което да промени съдържанието или значението на обработваното решение.

■ Създадените модели и алгоритми трябва също да са в състояние да бъдат съхранявани и изпълнявани в сигурни среди, така че да осигурят цялост и невещественост на системата.

4

Принцип на прозрачност, безпристрастност и справедливост: да се правят методите за обработване на данни достъпни и разбираеми, да се разрешават външни одити.

■ Трябва да бъде постигнат баланс³ между интелектуалната собственост на определени методи за обработване и необходимостта от прозрачност (достъп до процеса на проектиране), безпристрастност (липса на предубеждение)⁴, справедливост и интелектуална почтеност (приоритизиране на интересите на правосъдието), когато се използват инструменти, които може да имат правни последици или е възможно в значителна степен да засегнат човешки животи. Следва да бъде изяснено, че тези мерки се прилагат спрямо цялата верига на проектиране и експлоатация, тъй като процесът на подбор и качеството и организацията на данните оказват пряко влияние върху фазата на обучение.

■ Първата опция е пълната техническа прозрачност (например, отворен код и документация), което понякога се ограничава от защитата на търговските тайни. Системата би могла и да бъде обяснена на ясен и познат език (да се опише как се постигат резултати) чрез съобщаване, напр., на естеството на предлаганите услуги, инструментите, които са

³ От интерес в тази връзка е предложението, отправено на страница 38 на изследването на Комитета на експертите по интернет посредници (MSI-NET) на Съвета на Европа относно „Алгоритмите и правата на човека“ („Algorithms and Human Rights“): „Предоставянето на цели алгоритми или на основния софтуерен код на обществеността е малко вероятно решение в този контекст, тъй като частните дружества разглеждат алгоритма им като ключов собствен софтуер, който е защитен. Възможно е обаче да се изисква ключови подкласове информация относно алгоритмите да бъдат предоставяни на обществеността, например, кои променливи се използват, за какви цели се оптимизират алгоритмите, данните за обучението и средните стойности и стандартните отклонения от получените резултати, или количеството и типа данни, които се обработват от алгоритъма.” Или дори предложенията, появяващи се на страница 117 от доклада „Изкуствен интелект за човечеството“ („AI for humanity“), изготвен от г-н Седрик Вилани (Cedric Villani), член на френското Национално събрание, като част от мисия, възложена му от министър председателят на Френската република: „Проверяващите могат да бъдат удовлетворени като просто проверят справедливостта и равнопоставеността на една програма (правейки само онова, което се изискват от тях), подавайки различни невени изходни данни, например, или като създавайки голям брой профили на системи потребители съобразно точни указания.” В допълнение към това са и изявленията в доклада на Камарата на лордовете, „Изкуственият интелект в Обединеното кралство: готов, склонен и способен?“ („AI in the UK: ready, willing and able?“), параграфи 92, 96-99.

⁴ В тази връзка е интересно да се отбележат решенията, принципно разглеждани за осигуряване на неутралността на алгоритмите в горепосочения доклад на Камарата на лордовете (параграфи 114, 115, 116, 119, 120): по-разнообразни набори от данни, по-голямо разнообразие и повече мултидисциплинарни подходи, повече проверки на аспекти като обработване на данни и начина, по който е изградена машината.

разработени, работата и рисковете от грешка. Може да се възложи на независими органи или експерти да удостоверяват и проверяват методите на обработване или да дават предварително съвет. Публични органи биха могли да извършват сертифициране, което да е предмет на редовен преглед.

5

Принцип „под контрола на потребителя“: да се изключва подход на обвързване и да се гарантира, че потребителите са информирани участници, които имат контрол върху правените избори.

■ Трябва да бъде разширена автономията на потребителите и същата да не се ограничава чрез използването на използващи изкуствен интелект инструменти и услуги.

■ Професионалистите в системата на правосъдието следва, във всеки един момент, да бъдат в състояние да преглеждат съдебните решения и данните, използвани за постигане на даден резултат и да продължат да не са по необходимост обвързани от него в светлината на специфичните характеристиките на това конкретно дело.

■ Потребителят трябва да бъде информиран на ясен и разбираем език, независимо дали решенията, предлагани от използващите изкуствен интелект инструменти са обвързващи или не, за различните възможности, с които разполага и че той има правото да правен съвет и правото на достъп до съд. Той трябва да бъде внимателно информиран за каквото и да било предходно обработване на дело от изкуствен интелект преди или по време на съдебен процес и има правото да възрази, така че неговото дело може да бъде разгледано директно от съд по смисъла на член 6 от ЕКПЧ.

■ Като цяло, когато бъде реализирана каквато и да е основана на изкуствен интелект информационна система, следва да има програми за компютърна грамотност за потребителите и разисквания, включващи професионалисти от съдебната система.

Приложение I

Задълбочено изследване на използването на изкуствения интелект в съдебните системи, по-специално приложения с изкуствен интелект, обработващи съдебни решения и данни

Изготвено от г-н Ксавие Ронсин, Първи председател на Апелативен съд – Рен, научен експерт (Франция),

и

г-н Василеос Лампос, водещ изследовател в отдела за компютърни науки на Юнивърсити Колидж Лондон (UCL), научен експерт (Обединено кралство),

с участието на г-жа Агнес Метрпиер, съдия, член на Консултативния комитет на Конвенцията за защита на лицата при автоматизираната обработка на лични данни на Съвета на Европа (Франция)

Следните експерти също дадоха своя принос за фината настройка на Изследването:

Г-н Франческо Контини, Старши изследовател в Изследователския институт за съдебните системи – Национален изследователски съвет (IRSIG-CNR), Болоня (Италия)

Г-н Франческо Де Сантис, Професор по процедури във връзка с правата на човека, Университет на Неапол (Италия)

Г-н Жан Ласег, философ и епистемолог, изследовател в Centre National de Recherche Scientifique (CNRS) и асоцииран изследовател в Institut des Hautes Etudes sur la Justice (IHEJ) (Франция)

Г-жа Дори Райлинг, почетен старши съдия, Независим експерт по информационни технологии и съдебна реформа (Нидерландия)

Г-н Алеш Завршник, Главен изследовател в Института по криминология, доцент във Факултета по право на Университета на Любляна (Словения) и Изследовател на EURIAS 2017-18 в Collegium Helveticum в Цюрих (Швейцария)

ВЪВЕДЕНИЕ

1. Вълната на цифровата трансформация в нашите общества все още има неравномерно въздействие върху съдебните системи на държавите членки на Съвета на Европа. Много европейски държави изглежда вече са разработили изключително развит подход за използване на практически приложения (от гледна точка и на технологията и на правната поддръжка), докато за други, това все още е нов въпрос и фокусът е единствено върху ефективното ИТ управление.

2. Сред технологиите, работещи в тази голяма дигитална трансформация, изкуственият интелект (AI) изглежда като най-зрелищната и най-поразяващата. В Съединените американски щати, „роботите адвокати“ вече работят и изглежда разговарят на естествен език с хората. Правните технологични стартапи, които специализират в създаването на нови правни услуги, предлагат нови приложения на юридическите професии, основно на адвокатите, правните служби и застрахователите, позволяващи задълбочен достъп до правна информация и съдебна практика. Тези частни дружества дори се стремят да предвидят решенията на съдиите с инструментите на „прогностичното правосъдие“, въпреки че ще видим, че това може и да не е най-доброто описание за тях⁵.

3. Един първоначален преглед на този феномен обаче ни приканва да направим разграничение между това търговско разкриване и действителността на използването и разгръщането на тези технологии. Към настоящия момент, съдиите в държавите членки на Съвета на Европа не изглежда да използват на практика и ежедневно прогнозния софтуер. Проведени са изпитвания на място⁶ и академична работа⁷, за да се проучи потенциалът на тези приложения, но те все още не са приложени в широк мащаб. Инициативата за разработването на тези инструменти идва до голяма степен от частния сектор, чиято клиентела до този момент се състои основно от застрахователни дружества, адвокати и правни служби, които желаят да намалят правната несигурност и непредвидимостта на съдебните решения. Независимо от това, частният сектор все повече се обръща към вземащите решения в обществената сфера лица с желание да види тези инструменти – които са понякога „beta“ версии, т.е. ще се развиват във времето – интегрирани в публичните политики.

4. В линията на процеса на разсъждения, започнат в неговите

⁵ Вж по-специално въвеждащото каре към Глава 9.

⁶ Например, апелативните съдилища на Дуе и Рен във Франция проведоха тримесечно изпитване през 2017 г. със софтуерна програма, определена като „прогнозна“ чрез един съдебен състав.

⁷ Труд по извадка от 584 решения на Европейския съд по правата на човека: Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preotiuc-Pietro, Vasileios Lampos, „Predicting съдебни решения of the European Court of Human Rights: a Natural Language Processing perspective“, публикувано на 24 октомври 2016 г., [онлайн], <https://peerj.com/articles/cs-93/>

„Указания за предизвикване на промени към Киберправосъдие” („Guidelines on how to drive change towards Cyberjustice”),⁸ СЕРЕЈ предлага да бъдат предоставени ключове за по-добро разбиране на феномена „прогностично правосъдие” на вземащите решения в обществената сфера лица и на професионалистите в областта на правосъдието.

5. Първата задача е да се изяснят въпросите, свързани с присъщото естество на тези масови системи за обработване на данни от съдебната практика, заедно с техните технически и теоретични ограничения. Тези аспекти не са споменавани често в дебата по този въпрос в съдебната сфера, но те са много добре познати и обсъждани от специалистите в тези технологии (математици, статистици и компютърни специалисти) и заслужават определено внимание.

6. На второ място, този документ ще анализира ползите и рисковете от тези инструменти. Макар и поддръжниците им да обръщат специално внимание на тези активи от гледна точка на прозрачност, предвидимост и стандартизация на съдебната практика, техните критици отбелязват ограниченията и уклоните на мотивите на софтуера, който в момента е на пазара. Присъщите рискове на тези технологии могат дори да преминат границите на акта на осъждане и да засегнат съществени елементи на функционирането на правовата държава и съдебните системи, с които Съветът на Европа е особено обвързан.

7. Те включват принципи като върховенството на закона. Ефектът от тези инструменти може да не се изчерпва само до предоставяне на стимули, но и да бъдат по-императивни, създавайки нова форма на нормативност, която би могла да допълни закона, като уреди суверенното право на преценка на съдията и потенциално водейки, в дългосрочна перспектива, до стандартизацията на съдебните решения, основана вече не на размишляване по отделното дело от съдилищата, но на чисто статистическо изчисление, свързано със средното обезщетения, което преди това е било присъждано от други съдилища.

8. Необходимо е и да се обмисли дали тези решения са съвместими с отделните права, закрепени в Европейската конвенция за правата на човека (ЕКПЧ). Сред тях е правото на справедлив процес (по-специално правото естествен съдия, установено от закона, правото на независим и безпристрастен съд и равнопоставеност в съдебното производство) и, когато са положени недостатъчни грижи за защита на данните, съобщени като отворени данни, правото на зачитане на личния и семейния живот.

9. Макар да взема предвид тези въпроси, документът подчертава големият потенциал на изкуствения интелект да подпомогне юристите в работата им. Няма съмнение, че някои приложения с изкуствен интелект, които все още се разработват или тестват, като предназначението да

⁸ Вж по-специално параграф 51 от документа СЕРЕЈ(2016)13, Guidelines on how to drive change towards Cyberjustice.

подобряват правните проучвания, биха могли да бъдат много полезни при извършването на по-бързо и по-ефективно обработване на натовареността на съдиите. Документът подчертава тези позитивни примери и принципно застава зад използването на изкуствен интелект от юристите съобразно техните нужди, при условие че се обърне дължимото внимание на индивидуалните права, гарантирани от ЕКПЧ и стандартите на Съвета на Европа, особено по наказателно-правни въпроси. Съвсем без да е прост инструмент за подобряване на ефективността на съдебните системи, изкуственият интелект засилва гаранциите за върховенство на закона, заедно с качеството на публичното правораздаване.

10. На последно място, документът предлага средство за наблюдение на този феномен под формата на етична харта, като подчертава необходимостта от внимателен подход към интегрирането на тези инструменти в публични политики. От съществено значение е всеки публичен дебат да включват всички заинтересовани страни, независимо дали са юристи, правно-технически дружества или учени, за да им позволи да пренесат пълния обхват и възможно въздействие на въвеждането на приложения с изкуствен интелект в съдебните системи и да изработят етичната рамка, в която те трябва да работят. Впоследствие, този дебат би могъл да се пренесе отвъд чистата „бизнес“ рамка, включвайки самите граждани и така допринасяйки в някаква степен за общата компютърна грамотност, както е постигнато в Канада.⁹

1. Състояние на използването на алгоритми с изкуствен интелект в съдебните системи на държавите членки на Съвета на Европа

През 2018 г. използването на алгоритми за изкуствен интелект в европейските съдебни системи остава предимно търговска инициатива на частния сектор, насочена към застрахователни дружества, правни отдели, адвокати и физически лица.

11. Използването на изкуствен интелект в съдебната област изглежда доста популярно в Съединените американски щати, които са инвестирали в тези инструменти по един сравнително опростен начин – и по граждански и по наказателни дела.¹⁰

12. Определянето на случаи на инициативи за алгоритми за изкуствен интелект в съдебните системи на държавите членки на Съвета на Европа е по-трудна задача, тъй като повечето инициативи идват от частния сектор и не са често включени в публичните политики.

13. Въпросът за използването на изкуствен интелект в съдебните

⁹ Декларация от Монреал, налична на <https://www.declarationmontreal-iaresponsable.com/demarche>, прегледана на 16 юли 2018 г.

¹⁰ Вж алгоритми или инструменти на COMPAS като RAVEL LAW или ROSS чатбот

системи беше разгледан в едно специфично онлайн проучване, пуснато през април 2018 г. за представители на държавите на членки на СЕРЕJ и гражданското общество. Нивото на отговор беше сравнително ниско и не позволи установяването на ясни тенденции. Някои частни оператори не изглеждаха особено възприемчиви към това проучване и членовете на СЕРЕJ, които принадлежат основно към министерствата на правосъдието или висши съдебни съвети, можеха да посочат само инструментите, използвани от публичната сфера.

14. В резултат от това, списъкът по-долу е само частичен и базиран основно на проучване, извършено от експерти и секретариата, като се използва публично достъпна литература.¹¹

15. Може да бъде направена класификация съобразно предлаганата услуга. Включването на изкуствен интелект може да варира в голям диапазон съобразно приложенията. За илюстрация, основните категории са следните:

- ▶ Машините за търсене на съдебна практика с разширени възможности
- ▶ Онлайн разрешаване на спорове
- ▶ Съдействие при съставянето на актове
- ▶ Анализ (прогностичен, мащаби)
- ▶ Категоризация на договори съгласно различни критерии и установяване на отклоняващи се или несъвместими договорни клаузи
- ▶ „Чатботи”, които да информират страните в процеса или да ги подпомогнат в техните съдебни производства

16. **Латвия** посочи, че е проучвала възможностите за машинно самообучение за правораздаването. Основната цел би била да се обработва съдебна статистика, за да се съставят междинни прогнози за човешките и финансовите ресурси, които да бъдат разпределени.

17. Други дейности, осъществявани от правно-технически дружества, не са включени в тази класификация, тъй като те включват малко или никакво обработване с изкуствен интелект: някои сайтове предлагат достъп до правна информация, „облачни” решения, електронни подписи и т.н.

18. Неизчерпателен списък на правните служби, използващи изкуствен интелект в своите дейности, се съдържа по-долу:

¹¹ Вж обобщена библиография в Приложение IV – съществен принос от Беноа Шарпентие Venoit, както и Джузепе Контиса и Джовани Сартори (https://media.wix.com/ugd/c21db1_14b-04c49ba7f46bf9a5d88581cbda172.pdf) и Еманюел Барте (<http://www.precisement.org/blog/Intelligence-artificielle-en-droit-derriere-la-hype-la-realite.html#nb14>) (само на френски език)

Софтуер	Държава	Тип
Doctrine.fr	Франция	Машина за търсене
Prédicte	Франция	Анализ (с изключение на наказателни дела)
Case Law Analytics	Франция	Анализ (с изключение на наказателни дела)
JurisData Analytics (LexisNexis)	Франция	Машина за търсене, Анализ (с изключение на наказателни дела)
Luminance	Обединено кралство	Анализ
Watson/Ross (IBM)	САЩ	Анализ
HART	Обединено кралство	Анализ (наказателни, риск от повторно извършване на престъпление)
Lex Machina (LexisNexis)	САЩ	Анализ

2. Обзор на политиките за отворени данни във връзка със съдебните решения в съдебните системи на държавите членки на Съвета на Европа

Наличието на данни е съществено условие за разработването на изкуствени интелект, като му дава възможност да изпълнява определени задачи, които преди това са изпълнявани от хора по неавтоматизиран начин. Колкото повече данни са налични, толкова повече изкуственият интелект е в състояние да изчисти моделите, подобряващи тяхната способност да прогнозират. Поради това подходът с отворени данни към съдебните решения е предпоставка за работата на правно-техническите дружества, специализиращи в машини за търсене или анализ на тенденции („прогностично правосъдие“).

Обработването на тези данни поражда редица въпроси, като изменения във формирането на съдебната практика и защита на личните данни (включително имената на професионалистите).

19. Компютърно събраните данни се посочват като „нефта“ на 21-ви век, тъй като тяхното използване и взаимни препращане създават цяло ново богатство. Макар че някои заинтересовани страни и автори оспорват този аргумент, глобалните успехи на цифровата индустрия през последните десетилетия потвърдиха огромния потенциал за растеж на тази област на дейност.

20. Количественото определяне на човешките дейности, сега в глобален мащаб, не би могло да не засегне данните, създадени от публичния

сектор. Това наложи придвижването към отворени публични данни, на базата на много по-стари императиви, които са основополагащите принципи на нашите конституционни държави.

21. Главното предизвикателство през последните години е създадено от възникването на даващи възможност за сваляне публични данни (отворени данни), по-специално с контекста на „Партньорството за открито управление” („Partnership for Open Government”)(OGP). OGP е неправителствена организация, която свързва почти 70 държави членки (включително много от държавите членки на Съвета на Европа) с представители на гражданското общество и цифровите гиганти. Целта на тази откритост е да се повиши прозрачността на публичните дейности, да бъдат насърчени гражданите при разработването и оценката на публични политики и да се гарантира целостта на публичната услуга и тези, които я изпълняват, чрез обработване на значителни количества информация, организирана в бази данни (големи данни).

2.1. Определение за отворени данни по съдебни решения

22. Преди всичко, нека предефинираме понятието за отворени данни преди да обърнем внимание на въпроса за въздействието на позволяването на отворени данни върху съдебната дейност. На първо място, често се объркват достъп до информация и достъп до данни (точно, достъп до информация под формата на база данни)¹².

23. Определено количество публична информация, която изисква широка публичност, вече е разпространено чрез информация технология. Във Франция, правителственият сайт Legifrance.fr е основният онлайн източник на удостоверена публична информация, състоящ се не само от текстове на закони и подзакони нормативни актове, но и от съдебна практика и информация за назначения на публични позиции. Тази единна информация, макар и достъпна в интернет, се различава напълно от директния достъп до данни, организиран и включен като база данни, която може да бъде свалена и обработена от компютър.

24. Поради това, отворените данни включват само разпространението на „сурови” данни в структурирани компютърни бази данни. Тези данни, събрани като цяло или отчасти с други структурирани източници, представляват онова, което наричаме големи данни. Консултативният комитет по Конвенция 108 на Съвета на Европа определя големите данни

¹² Данните са безсмислени букви и цифри. Информацията е данни, включени в контекст. Контекстът е това, което дава смисъл на данните. Ние можем да познаем, че 2005 е година, но без контекст не знаем. Обаче в „през 2005 приключихме три 3 проекта” контекстът дава смисъл на цифрите. Поради това „отворените данни” не са данни в смисъла на определението, а информация. Също така, големи данни са също големи количества информация, а не данни.

като „нарастващата технологична възможност да се събират, обработват и извличат ново и прогностично знание с голям обем, скорост и разнообразие от данни. От гледна точка на защитата на данните, основните въпроси касаят не само обема, скоростта и разнообразието на обработваните данни, но и анализа на данни чрез използването на извличането на нови и прогностични знания за целите на вземането на решение по отношение на отделни лица или групи. Поради това, за целите на тези Указания, определението за Големи данни обхваща и Големи данни и Анализирани големи данни”.¹³

25. Както показва това определение, отворените данни следва да не бъдат бъркани с тяхното средство за обработване. Част от дискусиата по въпроса действително се отнася до обработване, извършвано с различни съвременни методи, които са принципно определяни като наука за данните. Прогностичното правосъдие, използващо изкуствен интелект, машините за търсене с разширени възможности, които прилагат изключително прецизни критерии и правните работи са все алгоритмични приложения, които се запазват с данни, но нямат нищо общо със самата политика за отворени данни.

26. Тази политика обаче трябва да бъде изследвана в светлината на възможностите, които тя предлага за допълнително обработване, независимо от нейния характер. Ако определени данни бъдат филтрирани нагоре по веригата, като се вземат предвид например, нуждата от поверителност и зачитане на неприкосновеността на личния живот, последващите рискове от злоупотреба изглеждат по-малки.

2.2. Състояние на развитие на отворените данни по съдебни решения в държавите членки на Съвета на Европа и последствия за развитието на съдебната практика

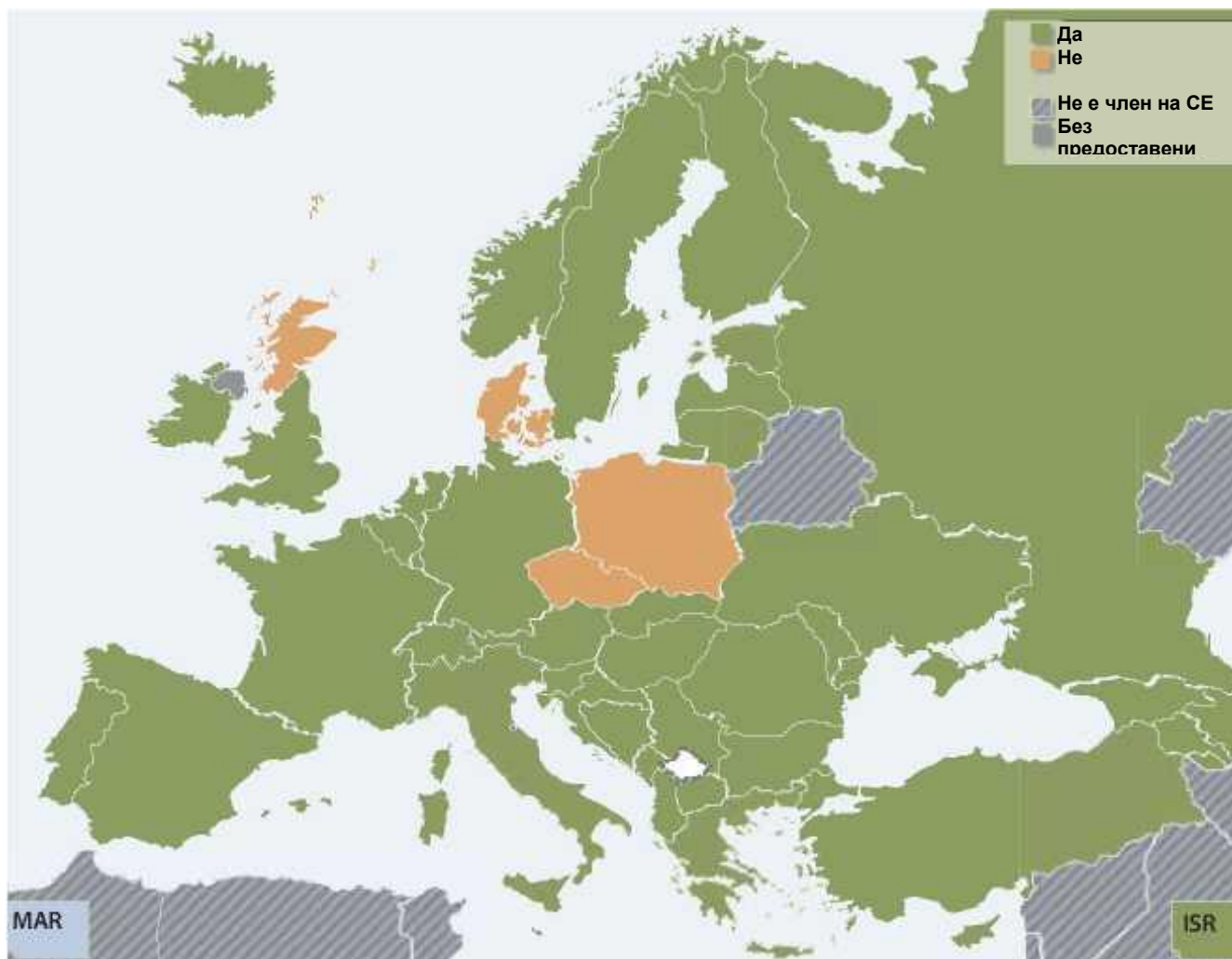
27. Каква е ситуацията в държавите членки на Съвета на Европа по отношение на отворени данни по съдебни решения? Цикълът на оценка на СЕРЕJ 2016-2018 се фокусира за пръв път върху въпроса за предоставянето на съдебните решения в отворени данни, за което се използва известно обработване с изкуствен интелект. Въпросът за анонимизирането или псевдонимизирането на данните в рамките на Европейската мрежа за защита на данните¹⁴, осигурена чрез Общия регламент за защита на данните (ОРЗД, Регламент на ЕС 2016/679) и

¹³ T-PD(2017)1, Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data (Указания за защита на лицата при обработка на лични данни в свят на големи данни).

¹⁴ Общи актове в областта защитата на данните (DGPS, Регламент на ЕС 2016/679, и Конвенцията № 108 на Съвета на Европа за защита на лицата при автоматизираната обработка на лични данни).

Конвенция № 108 на Съвета на Европа беше предмет на специфичен въпрос, предназначен да установи мерките, приложени от държавите членки и наблюдатели в тази особено чувствителна област.

Фигура 1. Таблица показваща, държавите членки на Съвета на Европа, в които има въведени отворени данни по съдебни решения в 2016 г. (Q62-4 от схемата за оценка)



28. От всички проучени държави и наблюдатели, само 5 декларираха, че не са реализирали политика на отворени данни за съдебни решения през 2016 г. Макар този процент отговори да не следва да бъде поставян в по-широк контекст, тъй като някои отговориха объркаха публичен достъп до решенията с отворени данни (**Армения, Белгия, Босна и Херцеговина, Люксембург, Руска федерация, Испания, Турция**), той разкрива, от една страна, желание за прозрачност от страна на европейските съдебни институции и, от друга страна, желание от страна на много държави да направят съдебните решения публични и така да направят възможна работата с тях по-късно чрез използване на инструменти с изкуствен интелект. Това изисква и усилия от съответните институции, тъй като трябва да бъдат въведени редица технически мерки в тази връзка. Във Франция по-специално, вече е налична известна административна съдебна практика за сваляне на данни на сайта *data.gouv.fr* (вж по-долу).

29. Що се отнася до защитата на личните данни, 23 държави са

декларирали, че псевдонимизират¹⁵ поне някои типове спорове (напр. личен статус, семейно положение), като заличават данни, позволяващи идентифицирането на страните или свидетелите (имена, адреси, телефонни номера, идентификационни номера, номера на банкови сметки, данъчни номера, здравен статус и т.н.). Тази работа изглежда отговорност на съдебни служители (напр. **Израел, Република Молдова**) или публични длъжностни лица (напр. **Босна и Херцеговина, Испания**). **Босна и Херцеговина и Унгария**, от друга страна, са посочили, че публикуват имената на професионалисти.

30. Съществува обаче реално затруднение при измерването на въздействието на отворените данни върху ефективността и качеството на правосъдието. Както е посочено по-горе, инициативата за повторно използване на тези данни е по същество частно, насочена към професионални клиенти (адвокати, правни отдели), и една изключително междуправителствена дейност вероятно не е най-добрият начин изцяло да се идентифицират тези положителни резултати.

31. Ситуацията във Франция е представителна за въпросите, повдигнати от този подход, и разкрива редица от разглежданите въпроси. На първо място, важно е да се подчертае, че Франция прие законодателство през 2016 г., налагащо задължителна рамка за разпространението на решения като отворени данни на своите съдилища.

32. Членове 20 и 21 от Закона за цифрова република¹⁶ скъса с предишната логика¹⁷ на подбор кои решения от съдебните и административните съдилища и трибунали да бъдат разпространявани, ако са „значителен интерес“. Съгласно новия френски закон обаче е заложен обратният принцип, че всичко може да бъде публикувано, освен в специфични случаи, определени от закона (за съдебни решения) и като се отчита неприкосновеността на личния живот на съответните лица. Въведена е обаче разпоредба за съдебни и административни решения, които се публикуват едва след като бъде направен анализ на риска от повторна идентификация на включените лица.

¹⁵ Както е определен от T-PD е неговите „Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data“ (Указания за защита на лицата при обработка на лични данни в свят на големи данни) T-PD(2017)1, псевдонимизирането се отнася до обработката на лични данни „по такъв начин ,че личните данни вече да не могат да бъде отнесени към конкретен субект на данни без използването на допълнителна информация, при условие че тази допълнителна информация се пази отделно и е предмет на технически и организационни мерки за гарантиране на това, че личните данни не са отнесени към идентифицирано или идентифицируемо физическо лице.“

¹⁶ Този закон е приет, за да приведе френското право в съответствие с Директива 2013/37/ЕС на Европейския парламент и на Европейския съюз от 26 юни 2013 г., която на свой ред измени Директивата на Съвета от 17 ноември 2003 г. относно повторната употреба на информацията в общественния сектор („Директивата за ИОС“).

¹⁷ Член R433-3 от Кодекса за организация на съдебната система.

33. Има множество предимства на широкото разпространение, промотирано по-специално от Касационния съд на две конференции през октомври 2016 г. и февруари 2018 г. Те включват по-голяма информираност за тенденциите в съдебната дейност и съдебната практика, повишеното качество на система на правораздаване, която знае, че е наблюдавана и създаването на напълно нова фактическа референтна база.

34. Това теоретично желание за „регулиране“ на съдебната практика чрез цифрово усилване повдига обаче редица общи въпроси, които се отнасят до всички държави, обмислящи подобен подход. На първо място, то трябва да бъде поставено в контекста на някои от принципите, изложени от Европейския съд по правата на човека в случай на различия в националната съдебна практика. Съдът ясно подчертава необходимостта да се постигне баланс между правната сигурност, която право решенията по-предвидими, и жизнеността в съдебното тълкуване.¹⁸

35. Това желание се охлажда и няколко аспекта от чисто техническо естество, на първо място:

а. Събирането на всички съдебни решения, отговарящи на изискванията за публикуване не е по необходимост добре координиране между всички съдебни инстанции: някои от търговските приложения в европейските съдилища не са предназначени за това, особено що се отнася до решенията на първоинстанционните съдилища, а някои държави ще трябва да създадат нови процеси за събиране и подреждане на решения, ако искат събирането да е изчерпателно;

б. Въпреки някои обещаващи събития, все още не е създаден напълно ефективен автоматизиран механизъм след идентификация, който да може да предотврати какъвто и да е риск от идентификация или повторна идентификация;

И също толкова съществени аспекти:

в. Следва да се обмисли и трансформирането на самата логика на създаването на съдебна практика. Каква е стойността на „стандарта“, произтичащ от броя решения, произнесени по конкретен въпрос? Допълва ли този „стандарт“ закона? Ако да, това нов източник на правото ли е?

г. Следва ли от съдиите да се изисква да предоставят допълнително разяснения за решения да се отклонят от тази норма?¹⁹ Този въпрос не е без значение и не води до еднакъв отговор във всички съдебни системи. Във френското гражданско

¹⁸ *Greek Catholic parish Lupeni and Others v. Romania* [GC]. No. 76943/11, 29/11/2016, § 116.

¹⁹ Eloi Buat-Menard and Paolo Giambiasi, «La mémoire numérique des décisions judiciaires», Dalloz Reports, 2017, p. 1483. (само на френски език)

право,²⁰ „съдиите уреждат спорове в съответствие с правилата, приложими спрямо тях“. Във френския контекст, ако те бъдат приканени да обосноват решенията си, като предоставят всички причини, поради които са се отклонили от предложената тенденция на преобладаваща съдебна практика за начина на решаване на спора (като спазват съответните правни норми), това не би ли представлявало отстраняването им от длъжност? Съдиите не биха решавали спорове само в съответствие с правилата на закона, но и в съответствие с тенденциите на съдебната практика, изведени от статистика, съставена с цифров инструмент (който също би могъл да е тенденциозен (пристрастен) или разработен без външен контрол от частен оператор, вж раздели 6.1 и 9 по-долу). Въпросът, който възниква, отива много отвъд специфична национална характеристика, свързана по-широко с мястото на използващите изкуствен интелект инструменти в процесите на вземане на решения от съдиите. Отбележете, че този анализ е ограничен до обработване на съдебна практика чрез алгоритми и по никакъв начин не се отнася до ролята на съдебната практика в източниците на правото или авторитета на прецедентите в съдебната практика, които – освен това – са добре установени принципи на европейско равнище.²¹

- д. Освен това, ако бъдат установени норми съобразно преобладаващата тенденция, съдебните решения не биха ли станали еднообразни и вече не биха били подредени съобразно йерархията на съдилищата, от които произлизат, като по този начин се пренебрегне значимостта на решенията на върховните съдилища, които са гарант за еднообразното тълкуване на правото в много европейски държави? Какво би било отношението между нормите и гражданското право? Тъй като те произтичат от болшинството, биха ли станали те критерий за тези съдилища, когато се определя тяхната собствена съдебна практика, която те би следвало на свой ред да обосноват когато се отклоняват от виждането на мнозинството?**
- е. Накрая, не съществува ли в действителност риск съдебните решения да се пишат според една репродуктивна логика? Макар да е вероятно съдебните решения да се развиват успоредно с развитието на нормативната рамка (национална, европейска или международна), съдебната практика, която е на разположение за справка (например, от върховни съдилища и европейски съдилища)**

²⁰ Член 12 от френския Граждански процесуален кодекс

²¹ Съдът на ЕС е посочил, по отношение на условията, уреждащи отговорността на една държава членка за съдържанието на дадено решение на върховен национален съд, че „нарушение на правото на Общността ще бъде достатъчно сериозно“ и трябва да породи обезщетение за уреждането, „когато съответното решение е било постановено в явно нарушение на съдебната практика на Съда по въпроса“ (Съд на ЕС, Дело C-224/01, Koebler, §56).

или социално-икономическия контекст, нормата, произтичаща от мнозинството, не би ли станала стандарт, към който съдиите биха били насърчавани да се обръщат безкритично, с индуциран ефект от прекомерното стандартизиране на съдебните решения?

36. **Може да възникнат и съмнения относно последствията за потребители, които се предполага, че ще се ползват директно от прозрачността на дейностите:** ще се възползват ли те реално от публикуването на всички съдебни решения в интернет или е по-вероятно да се окажат затрупани от масата решения, без задължително да се възползват от качествени ползи от тях, поради липсата на правни знания и критичния апарат за анализа им?

37. Колкото и да е парадоксално, някой може да се чуди дали отворените данни са в състояние да предоставят смислена информация, тъй като със сигурност ще бъде създадена възможност за сваляне на значителен набор от сурови данни с едно кликуване на даден линк, но смисълът ще остане напълно неясен за голямата част от гражданите.

38. Първото изискване, за да бъдат отворените данни прозрачни и информативни, е третите лица да могат да ги използват за анализ и разработване. Поради това, един икономически модел, в който публични данни от съдебната практика, които предметът на работа на съдебните органи, за да бъдат направени технически „разчетими“ за изкуствения интелект и съответстващи на юридическите изисквания за защита на личните данни, биха били обработвани безплатно от частния сектор и след това продавани на съдилищата, професионалистите и гражданите, заслужава да бъде поставен под въпрос и дори критикуван.

2.3. Защита на личните данни в политиките за отворени данни за съдебни решения

2.3.1. Имената на страните и свидетелите

39. За да се постигне справедлив баланс в цифровата ера между необходимостта съдебните решения да бъдат направени публични и зачитането на основните права на страните или свидетелите, техните имена и адреси трябва да не се появяват в публикуваните решения, по-специално с оглед на риска от присвояване и повторно използване на тази лична информация и особената чувствителност на данните, които е вероятно да се съдържат в решенията.²² За систематично прикриване на

²² ЕКПЧ, *Z. v. Finland* No. 22009/93, §§95 и сл. и Зелената книга на Европейската комисия за информацията в обществеността в информационното общество (COM(1998)585) („Ако не бъдат предприети специални предпазни мерки, базите данни със съдебна практика, които са инструменти за правна документация, могат да станат информационни файлове за отделни лица, ако в тези бази данни се

тази информация могат да бъдат използвани автоматизирани процеси.

40. Може да бъде прикритие и друга идентификационна информация (например телефонни номера, адреси на електронна поща, дати на раждане, лични имена на деца, редки лични имена, прякори и имена на места). От гледна точка на принципите на защита на личните данни, това прикриване представлява просто псевдонимизиране на данните, а не пълно анонимизиране. Обемът и разнообразието от информация, съдържаща се в съдебни решения, съчетана с все по-лесното взаимно препращане с други бази данни, прави практически невъзможно да се гарантира, че съответното лице не може да бъде идентифицирано отново. При отсъствието на такава гаранция, тези данни не могат да бъдат квалифицирани като анонимни и поради това попадат в приложното поле на правилата за защита на личните данни.

41. Някои особено чувствителни елементи от лични данни гарантират специално внимание, както е предвидено в Член 6 от Конвенция 108. Това се прилага спрямо данни, разкриващи етнически или расов произход, политически становища, членство в профсъюз, религиозни или други вярвания, физическо или психическо здраве или сексуален живот, които се счита за интимни подробности.

42. Съдебните решения могат да съдържат други, много разнообразни, типове лични данни, които попадат в тази категория чувствителни данни. Особено вероятно е съдилищата, работещи с наказателни дела, да обработват чувствителни данни, като тези по наказателни производства и присъди. Поради това, всички тези чувствителни данни заслужават специална бдителност. Тяхното масово разпространение бе създадо сериозни рискове от дискриминация, профилиране²³ и нарушаване на човешкото достойнство.

2.3.2. Имената на професионалистите, включително съдиите

43. Очевидно, да се знае как ще се стигне до едно решение е елемент от съществено значение за адвокатите при прогнозирането на изхода от дадено дело и те смятат, че познаването на съдията по делото е понякога също толкова важно колкото е познаването на закона. Те дълго време правят опити да правят сравнения между съдебни състави, повече или по-малко емпирично, така че да дадат по-добър съвет на клиенти, работещи с конкретен съдия или съдебен състав.

правят справки за получаване на списък от съдебни решения за конкретно лице, а не за намиране на информация за съдебната практика”).

²³ Профилирането е определено в член 4 от ОРЗД. То е обработване на лични данни на физическо лице с цел анализиране и прогнозиране на поведението му или състоянието му, като определяне на изпълнението на служебните му задължения, финансовото му състояние, здраве, предпочитания, навици в живота и т.н.

44. Този метод беше достатъчен, когато един адвокат говори само пред ограничен брой съдилища, но постепенното намаляване на ограниченията на място спрямо адвокатските колегии в много държави и свободата на придвижване и работа в рамките на Европейския съюз дава основания на всеки национален или дори европейски адвокат да иска да познава съдебната практика на всяка национална или европейска юрисдикция, в която е вероятно да пледира, в пълни подробности.

45. Поради това не можем да изключим възможността, в бъдеще, едни много полезни и – поради това – много скъпи приложения за машинно самообучение да бъдат много по-ефективни от опита и „здравия разум“ на съдебните адвокати, работещи по дела по традиционния начин. Използването на такива приложения би могло допълнително да подчертае изкривяването на конкуренцията и неравнопоставеност между адвокатските кантори, които са или не са използвали такъв софтуер за „прогностичен“ анализ на съдебната практика.

46. Съществува реален риск, заради едно такова конкурентно предимство, принципът на справедлив процес, установен от закона, да бъде подкопан. Възможността от профилиране на съдия чрез взаимни препращания на публични и частни данни, би могло да позволи на частните дружества и техните адвокати да се ангажират дори в повече практики на пренасочване на дела към удобен съд. Тази тактика вече е наблюдавана в продължение на дълъг период от време в Съединените американски щати и във Франция за свързани с пресата престъпления и нарушения на правото на личен живот в пресата, когато за ищците вече се знае, че избират съда, който изглежда ще присъди обезщетения за вреди и лихва в най-висок размер.

47. В допълнение към това, голяма част от дискусията по този предмет бърка отворените данни с необходимостта да се публикува определено количество публична информация. В резултат от това, понякога се привеждат аргументи, че имената на професионалистите следва да фигурират в отворени данни с цел публичност и прозрачност.

48. Въпреки това, предоставянето на компютризиранни данни от съдебна практика данни е напълно отделен въпрос от принципите на публикуване на оригинални или сертифицирани копия на решения. Целта на отворените данни е да позволи автоматизирано обработване на съдебна практика съгласно евтин оперативен лиценз. Както беше посочено по-рано, това предоставяне се извършва под формата на обхватна компютърна база данни, която е непрозрачна и не е директно разбираема за гражданите.

49. Това предоставяне очевидно не отговаря на нуждата да се публикуват имената на професионалистите, които са допринесли за конкретно решение. Следва да бъде отбелязано, че, по закон, посочването на имената на съдиите в съдебните решения е общо задължение за държавите членки, във връзка с принципа за публичен процес, предвиден

в Член 6, параграф 1 от Европейската конвенция за правата на човека, за да се гарантира обективната безпристрастност на съдиите (които трябва да бъдат идентифицируеми и законно назначени и натоварени със задълженията, които изпълняват) и спазването на процесуалните правила (напр. публичност и колегиалност).²⁴

50. Поради това, отговорът на въпроса за легитимността или нелегитимността на публикуването на имената на професионалистите²⁵ в отворени данни няма никаква връзка със задължението за публикуване на имената на професионалистите в решенията. По-скоро изглежда, че предизвикателството се състои в съчетаването на често пъти противоречащи си изисквания: публични дейности да бъдат направени прозрачни, като се даде възможност на гражданите да познават и да оценяват своите съдии, от една страна, докато се защитава неприкосновеността на личния живот на професионалистите (чиито функции следва да не ограничават фундаменталните им гаранции в тази област), от друга страна. Налице са сериозни предизвикателства, когато става дума за гарантиране на безпристрастността на съдиите и дори на съдебните институции, като цяло, на които политиките за отворени данни реално са създадени да отговорят²⁶. Какви практически мерки могат да бъдат предприети за да се защитят те от потенциални опити за дестабилизиране с лични данни на съдии от препращания, включени в бази данни с други източници (социални мрежи, търговски обекти) с цел да се установят хипотетични политически, религиозни и други пристрастия?

51. Тези въпроси не възникват в същата форма навсякъде в Европа и зависят от конкретните характеристики на съответната съдебна система (и от естеството на кадровия орган на съдебната система), колегиалния характер или липсата на такъв на решението и нивото на съответния съд. В Швейцария, например, където съдиите се избират, публикуването е гаранция за прозрачност и социална отговорност на съдиите спрямо гражданите и политическите групи. Тази информация е вече налична в онлайн машините за търсене (които не отворени данни в строгия смисъл на думите).²⁷

²⁴ ЕКПЧ, *Vernes v. Франция*, № 30183/06 от отношение на идентифицирането на съдиите, които са постановили решението, *Pretto v. Италия*, № 7984/77 за публикуването на решението, *Kontalexis v. Гърция*, No. 59000/08, § 38, *DMD GROUP, a.s, v. Словакия*, № 19334/03, § 66, *Miracle Europe KFT v. Унгария*, № 57774/13, § 58 за правото на съд, установено по право във връзка с правото на безпристрастен съдия.

²⁵ Тези въпроси могат да се прилагат им спрямо *Rechtspfleger*, който постановява съдебни решения и спрямо чиновниците, асистентите на съдията, посочени във формирания състав на съда (макар и в по-малка степен).

²⁶ Вж ЕКПЧ *Previti v. Италия*, № 45291/06, §§ 249 и сл., който припомня принципите на обективна безпристрастност на съдията.

²⁷ Вж пример от Швейцарския федерален съд, чиято съдебна практика може да бъде свалена: <https://www.bger.ch/fr/index/jurisdiction/jurisdiction-inherit-template/jurisdiction-recht.htm>; или, за кантоните: <http://ge.ch/justice/dans-la>

52. Тези въпроси не възникват в същата форма и в зависимост от нивото на юрисдикцията. Ценността на характеризирането на съдебната практика на съдилищата от по-ниско ниво, може да не е същата като за съдиите от върховните съдилища или международните съдилища. Например, Европейският съд по правата на човека разрешава търсения за на решения по имената на съдиите членове на решаващия състав, но не позволява изчисляването на статистика във връзка с конкретен съдия.²⁸ От друга страна, в държавите, в които съдебните органи са незапознати с практиката на особени мнения (съществуващи в рамките на този международен съд), може да изглежда несправедливо да се предвиди за даден съдия лична отговорност за решение, с което не е бил съгласен по време на обсъждането в един колегиален съд.

53. Тези дебати са добре определени от проучвателна мисия, проведена във Франция от професор Лои Кадие. Заключениета на мисията остават ограничени, тъй като те не препоръчват забрана на публикуването, но го запазват за определени типове съдебни процеси и го изключват за други (например, за специализирани наказателни дела). Беше предложена възможността за публикуване само на имената на върховните съдии, макар че беше прието, че това може да доведе до „еднопосочно пътуване“.

54. Към настоящия момент, може да бъде приложен един прост предохранителен принцип, за да се оцени интересът, по тип съдебен процес и степен на юрисдикцията, при публикуването на имената на професионалистите в подлежаща на сваляне база данни. Също така, ние не можем да изключим възможността самите съдебни институции или упълномощени трети лица, използващи тази информация извън контекста на отворените данни, да узнаят за съдебна практика.

55. Публикация от този вид във всеки случай все пак би трябвало да бъде проучена в светлината на приложимите международни европейски правила за защита на личните данни, като тези на Конвенция № 108 и ОРЗД и Директива 680/2016 на Европейския съюз.

3. Оперативни характеристики на изкуствения интелект (машинно самообучение), прилаган спрямо съдебни решения

Обработването на естествения език и машинното самообучение са двете техники в сърцевината на обработването на съдебни решения с използване на изкуствен интелект.

В повечето случаи, целта на тези системи не е да се възпроизвеждат правни мотиви,

jurisprudence (напр. Кантон Женева).

²⁸ Решенията на Европейския съд по правата на човека са колегиални. Публикациите включват всякакви особени мнения.

а да се установят взаимовръзките между различните параметри на едно решение (например при иск за развод продължителността на брака, доходът на съпрузите, съществуването на изневяра, размерът на постановената издръжка и т.н.) и, чрез използването на машинно самообучение, да се достигне до един или повече модели. Тези модели след това биха били използвани, за да се „прогнозира“ или „предвиди“ бъдещо съдебно решение.

Целта на тази глава е да се разясни свойствената природа на софтуера, описан като „прогностичен“ – понякога „beta“ версии, т.е. във фаза на разработване – и като потенциал и като ограничения на същия. Тя се стреми да предостави просто обяснение на алгоритмите на машинно самообучение, които са в сърцевината на автоматизирания анализ на съдебна практика.

3.1. Теоретичните функционалности на софтуера за „прогностично правосъдие“

56. Като представяне, следва да направим кратък обзор на характеристиките, обещавани от „прогностичния“ софтуер. Те предлагат установяването на вероятностите за успех (или успех) на дадено дело пред даден съд. Тези вероятности се установяват чрез статистическо моделиране на предишни решения чрез използване на методи от две обширни области на компютърната наука: обработване на естествен език и машинно самообучение. Тези подходи на моделиране често са наричани изкуствен интелект; в действителност те са „слаб“ изкуствен интелект (вж речника, стр.

57. Следва да бъде директно изяснено, че терминът изкуствен интелект е подложен на дебати от експертите, тъй като той води до много двусмислия. Терминът изкуствен интелект сега е навлязъл в нашия ежедневен език, за да опише разнообразен обхват от науки и технологии, позволяващи на компютрите да победят най-добрите шампиони в играта Го,²⁹ да управляват кола, да разговарят с хора и т.н. Изследователите предпочитат да идентифицират различните приложения чрез точните технологии, стоящи в основата им, включително машинно самообучение, и понякога наричат всички тези високоспециализирани ресурси на изкуствен интелект „слаби“ (или „умерени“) изкуствени интелекти. Те се разграничават от една крайна – все още напълно теоретична – цел да се създаде „силен“ изкуствен интелект, т.е. самообучаваща се машина, която е в състояние автоматично да разбира света като цяло, в цялата му сложност.

58. Във връзка специално с правосъдието, системите за прогностично правосъдие са създадени за използване от правни отдели, застрахователи (както за техните нужди, така и за притежателите на техни

²⁹ <https://www.nature.com/articles/nature16961>

полиции), както и адвокати, за да очакват резултата от съденото дело. Теоретично, те биха могли да помагат и на съдии в процеса и на взимане на решение.

59. Те предоставят графично представяне на вероятността за успех за всеки резултат от спор на база на критерии, въведени от потребителя (специфичен за всеки тип спор). Тези системи претендират, че са в състояние да изчислят вероятния размер на обезщетението, разпределяно от съдилищата.

3.2. Практическо функциониране на изкуствения интелект: статистически машини, конструиращи модели, базирани на миналото

60. Изначално трябва да се прави разлика между това, което е „предсказание“ и това, което е „прогноза“. Предсказание е актът на обявяване на това какво ще се случи (*prae*, преди – *dictare*, казвам) преди бъдещите събития (чрез свръхестествено вдъхновение, чрез ясновидство или предчувствие). Прогнозирането, от друга страна, е резултат от наблюдение (*насочване*, виждане) на набори от данни, за да се предвиди бъдеща ситуация. Тази злоупотреба с езика и разпространението ѝ изглежда се обяснява с прехвърляне на термина от природните науки, където той се отнася до редица техники от науката за данни, изведени от математиката, статистиката и теорията на игрите, които анализират настоящи и минали факти, за да създадат хипотези за съдържанието на бъдещи събития.

61. Също така, следва да бъде отбелязано, че логиката на работата на софтуера за прогностично правосъдие софтуер е основана по същество на генеративни (общо наричани Байесовски) или дискриминативни методи, които в крайна сметка се стремят да преценят настоящия или бъдещ диапазон от стойности на една променлива (напр. резултатът от съдебен процес) от анализа на примери от миналото.

62. Като цяло, важно е и да се има предвид антропоморфичната концепция, че изчислителните машини са умни и че техните създатели са успели да вкарат съзнание в механизмите им.³⁰ За съжаление, тази идея все още прониква в много анализи на прогностичното правосъдие, които дават на въпросните устройства непосредствени или бъдещи способности за близкото копиране на човешкия интелект. Поради това, този контекст, подхранван ежедневно от допълнителна поредица от революционни постижения в областта на изкуствения интелект,³¹ води всички ни към това

³⁰ Dominique Cardon, *A quoi rêvent les algorithmes, nos vies à l'heure des big data*, La République des idées, Editions du Seuil, 2015, стр.58.

³¹ "AlphaZero : l'IA de DeepMind apprend et maîtrise le jeu d'échecs en moins de 4 heures", *Génération NT*, статия, публикувана на 8 декември 2017 г., [онлайн], <https://www.generation-nt.com/alphazero-ia-deepmind-maitrise-echecs-4-heures->

да подхождаме към тези прогностични инструменти с определена доза – съзнателно или не – мистицизъм, заявявайки понякога, че това, което не е напълно възможно днес, неизбежно ще бъде възможно утре.

63. „Силните” изкуствени интелекти от научно-фантастичната литература не съществуват. Този тип изкуствен интелект, който би бил снабден не само с интелект, но и със съзнание, остава чисто въображаем. Системите за машинно самообучение, които се разработват понастоящем, са описвани като „слаби” изкуствени интелекти и **са в състояние да извеждат сложни модели и да се учат от големи обеми данни ефективно и често с високи нива на прогностична точност.**

64. За да се преодолеят всякакви инстинктивни или въображаеми съображения, тези технологии за обработване и анализ на информация трябва да бъдат разбирани на базата на следните три концепции.

- **изкуственият интелект не е един, хомогенен обект:** той в действителност е сбор от науки и техники (математика, статистика и компютърни науки), които могат да обработват данни за разработване на много сложни задачи по компютърна обработка.
- **машините за изкуствен интелект не създават интелект сами по себе си, но използват индуктивен подход:** идеята е да се свърже по почти автоматизиран начин набор от наблюдения (изходни данни) с набор от възможни резултати (продукти), използващи различни предварително конфигурирани свойства. Специално за прогностичното правосъдие, машината изгражда връзки между различните лексикални групи, като съставя съдебни решения. Тези групи са взаимосвързани между установените на етапа на въвеждане (факти и мотиви) и тези на етапа на създаване на резултата (диспозитива на решението), които са класифицирани след това.
- Надеждността на изградения модел (или функцията) силно зависи от качеството на използваните данни и избора на техника за машинно самообучение.

65. Струва си да се върнем за кратко към предшествениците на тези системи – експертни системи – които, от своя страна, са разчитали на правилата за обработване, написани от компютърни учени. Експертните системи (ЕСм) се развиват в края на 80-те години и 90-те години на 20 век, специално в медицината и финансите.³² Тези системи са в състояние да отговорят на специални въпроси и да разсъждават, като използват известни факти, изпълняват предварително определени правила за

actualite-1948919.html (страницата е прегледана на 14 декември 2017 г.).

³² Пример за това е високочестотната търговия (HFT), която е вид финансова транзакция, която се извършва на висока скорост от софтуер, базиран на алгоритми. През 2013 г. и в Европа, 35 % от транзакциите вече се извършваха с HFT. В САЩ, HFT представлява 70 % от обемите търговия на капиталовия пазар. С оглед на автоматизацията на транзакциите, пропадането е дори по-бързо в случай на сринове, което се случи по време на финансовата криза през 2007 г.

кодирание в машина. Въпреки успеха на експертните системи, като Deep Blue срещу Гари Каспаров в поредица от шахматни игри през 1997 г., тези системи се провалят, по-специално тъй като не са в състояние да тълкуват „безкрайното разнообразие от ситуации и контексти“³³ и стават неефективни между 200 до 300 правила за кодирание, от гледна точка и на изпълнение и на поддръжка (мотивите, следвани от системата, станаха почти невъзможни за разбиране от нейните създатели).

Фиг. 2: Старите експертни системи бяха програмирани с правила, възпроизвеждащи логиката на правните мотиви



66. Днес, идеята е вече да не се пишат правила за мотивиране, които отразяват човешкото разсъждаване, както при по-старите експертни системи,³⁴ а да се оставят системите за машинно самообучение сами да установят съществуващи статистически модели в данните и да ги съчетаят с конкретни резултати.

Фиг. 3: Машинното самообучение само създава модели, като автоматично търси резултати от взаимосвързки



67. Правейки това, тези системи не възпроизвеждат или моделират разсъждаване (като правно аргументиране). Например, днешните онлайн преводачи не извършват абстрактно мислене. Те подсказват вероятна преценка на най-доброто попадение между други от лексикални структури и преводи, които са вече направени. Действията на потребителите

³³ Hubert Dreyfus, *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992 цитиран от Dominique Cardon в *A quoi rêvent les algorithmes, nos vies à l'heure des Big Data*, стр.59.

³⁴ Писмените правила за обработване биха могли да са базирани на специфичен код на различни езици за програмиране, като LISP или редактори, за да моделират правила. Вж например Radboud Winkels, „CLIME : Un projet de développement de serveurs juridiques intelligents”, в Danièle Bourcier, Patricia Hasset and Christophe Roquilly (eds.), *Droit et intelligence artificielle*, 2000, Romillat, стр.59.

очевидно допринасят за подобряването на търсенето на попадение, но алгоритмът за машинно самообучение реално не извършва превод, като разбира значението на обработваните изречения.³⁵

68. Този пример показва как подходът на адвокатите и изследователите може да бъде различен: един адвокат ще се стреми да разбере ефективността на заварените правила, които изследователят може да предостави само чрез езика на своята наука, и това понякога се ограничава до тълкуване на хиляди параметри от една система за машинно самообучение.

69. Машинното самообучение е област от компютърната наука, в която компютърни програми се учат от опит. Чрез алгоритми машината осъществява процес на обучение, като дете, което се учи в своята среда. В обобщение, тези техники на учение може да са или да не са надзиравани от човек. Най-популярната категория машинно самообучение е понастоящем тази на обучението с утвърждение: машината сама жъне виртуални „награди“, ако действието доведе до очаквания резултат в дадена среда. Методите на машинно самообучение включват невронни мрежи (или тяхната по-сложна версия, известна като дълбоко обучение), които са получили доста широка публичност с оглед на тяхната автономност и техните доста шокиращи приложения, като това, което успява да постигне високи резултати на стари видео игри Atari 2600³⁶ (на база само на позиционирането на пикселите и резултатите). Други практически приложения за тези технологии вече засягат ежедневието ни и започват да се появяват в професионалния свят на правораздаването.³⁷

70. Тези алгоритми обаче остават силно специализирани в една конкретна задача и поставят проблеми с разбирането, когато са изправени пред хаотични ситуации или пред недостатъчни данни, за да може да се направи прогноза (като действителното разбиране на естествения език). В обществените науки, към които принадлежат правото и правораздаването, неуспехът дори би изглеждал неизбежен в отсъствието на убедителен модел на познавателната способност. За Пиер Леви, изкуственият интелект всъщност е съдържание, което предоставят разнороден инструментариум (логически правила, формални синтаксиси, статистически методи, невронни или социално-биологични симулации и т.н.), който не предлага общо решение на проблема с математическото

³⁵ Li Gong, "La traduction automatique statistique, comment ça marche ?" Interstices.info, публикувано на 29 октомври 2013 г., [онлайн], https://interstices.info/jcms/nn_72253/la-traduction-automatique-statistique-comment-ca-marche (страницата е прегледана на 14 декември 2017 г.).

³⁶ <https://www.nature.com/articles/nature14236>

³⁷ Изкуствен интелект by IBM Watson providing a service to the medical field and providing a search tool called "Ross", presented as a virtual lawyer - Roos, "Do more than humanly possible" [Online], <http://rossintelligence.com> (страницата е прегледана на 14 декември 2017).

моделиране на човешката познавателна способност.³⁸ Така, действителното прогностично обучение следва всъщност да бъде основани на добро системно представяне на света, което изследователят на изкуствения интелект Ян ЛеКун смята за фундаментален научен и математически въпрос, а не въпрос на технология.³⁹

71. Освен това, уникалността на настоящите системи за обработване на големи данни е, че те не се опитват да възпроизведат нашият модел на познание, но създават контекстуална статистика на база на безпрецедентна обем от данни, без никаква реална гаранция за изключване на грешните взаимовръзки.⁴⁰

72. Въпреки че е общоприето, че статистическите и вероятностните модели са подобрени чрез увеличаване на данните, с които същите се хранят, някои математици предупреждават за риска от увеличаване на грешните взаимовръзки (т.е. връзки между фактори, които нямат абсолютни никаква причинно-следствена връзка) в големите данни. Математиците Кристиян Сорин Калуд и Джузепе Лонго посочват риска от потоп от грешни взаимовръзки в големите данни: колкото по-голяма е базата данни, използвана за взаимовръзките, толкова по-големи са шансовете да се намерят повтарящи се модели и по-голяма е възможността да се допуснат грешки.⁴¹ Онова, което може да изглежда като закономерност за един изкуствен интелект (повтарящи се връзки между различни данни, концепции, контексти или лексикални групи) може всъщност да е произволно. Дори аргументът на двамата математици да не следва да се генерализира твърде прибързано, те отбелязват, че в определени огромни набори от числа, точки или обекти, се появяват редовни произволни модели и изглежда невъзможно те да бъдат алгоритмично разграничени от модели, разкриващи причинно-следствени връзки.⁴²

73. Накрая, алгоритмите, независимо от настоящото им равнище на сложност, все още автоматично се свеждат до взаимодействия, установени във фазата на обучение и – следователно – до тяхното

³⁸ Pierre Lévy, "Intelligence artificielle et sciences humaines", *Pierre Levy's blog*. 8 октомври 2014 г. [Онлайн], <https://pierrelevyblog.com/2014/10/08/intelligence-artificielle-et-scienc-es-humaines/> (страницата е прегледана на 30 декември 2017 г.).

³⁹ Yann LeCun, "Qu'est-ce que l'intelligence artificielle", изследване, публикувано на уебсайта на Collège de France, [онлайн], https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_Artificielle_Y._LeCun.pdf (страницата е прегледана на 14 декември 2017).

⁴⁰ Dominique Cardon, *цитирано съчинение*, стр.60.

⁴¹ Cristian Sorin Calude, Giuseppe Longo, "Le déluge des corrélations fallacieuses dans le big data", в Bernard Stiegler (dir.) (2017), *La toile que nous voulons*, FYP editions, стр.156.

⁴² Теория на „Памзи“, Ronald L. Graham, Joel H. Spencer -"Ramsey Theory", *Scientific American*, vol.263, No.1, July 1990, стр.112-117, цитирано от by Cristian Sorin Calude, Giuseppe Longo, "Le déluge des corrélations fallacieuses dans le big data " *пак там*.

минало. Поради това, съдържанието и качеството на потоците от данни, използвани в съставянето на изчисленията, е от основополагащо значение за разбирането на получените резултати и при установяването на възможни аналитични пристрастия. Тук отново съществуват много предизвикателства, тъй като, при един изчерпателен подход, анализът на най-голямото възможно количество данни, свързани с дейност, ще доведе до резултати, чието значение трябва да бъде разяснено във връзка с всички фактори, които са имали значение. При един по-ограничителен подход, в който се взема мостра от входящите данни, ще възникнат рискове и от уклона към преценка на плюсовете и минусите, които се изискват, за да се подбират едни данни спрямо други.

4. Може ли изкуственият интелект предварително да моделира правно мотивиране?

Изкуственият интелект заобикаля трудностите, на които се натъкват по-старите експертни системи: те не се опитват ръчно да отразят правното мотивиране, чието възпроизвеждане не само по себе си цел за тях. Машинното самообучение води до категоризиране между различните параметри, установени от създателите, или тези, които са открити от машината.

74. Континенталните правни системи са изместени далеч от идеала на рационализма, който олицетворява, напр. Гражданския кодекс на Франция от 1804 г. Има многобройни източници, които не се съчетават перфектно и които се отнасят до набор от правила, чието значение остава неопределено, което правният теоретик Хърбърт Л. А. Харт е нарекъл „открита структура на правото” („open texture of law”).⁴³ Системите на общото право, макар и да се разглеждат като икономически по-ефективни, тъй като се адаптират динамично към новите потребности,⁴⁴ също се развиват и предлагат не по-голяма правна сигурност. Поради това, правните норми не се развиват линейно, което ги разграничава от емпиричните закони (тези на „точните науки”), където всяко ново правило принципно допълва предишните и не обезсилва пълен набор от аргументи.

75. Следва да бъде подчертано, че, изправени пред тази обща сложност на правните системи, старите ИТ експертни системи бързо са достигнали своите лимити след полагането на 200 до 300 логически правила. Разделянето на правото на правила за създаване не беше достатъчно, за да осигури валиден образ на съвкупността от знания и методи, от които се ръководи един адвокат.

⁴³ Herbert L. A. Hart (1976), *Le concept de droit*, Saint-Louis university departments, Brussels.

⁴⁴ <https://www.contrepoints.org/2014/08/15/177160-common-law-contre-droit-civil-lexperience-francaise-de-lancien-regime>

76. Следователно, се е оказало невъзможно да се моделира мотивирането на съдиите на компютър, тъй като то трябва да моделира една позитивна правна система. Както подчертава теорията на правото, съдебното мотивиране е преди всичко въпрос на оценка и тълкуване на доказаните и относимите факти по едно дело, на приложимите норми на закона (писани или прецеденти) – значението на които остава, както беше казано, неопределено⁴⁵ – и на субективно тълкуване от съдиите на концепцията за справедливост, която следва да претърпи нови изменения в Европа с изискването за преглед на съразмерността, насърчен от Европейския съд по правата на човека.

77. Ксавие Линан дьо Белефонс подчертава, че сложността на правото е заложена в неговия телеологичен и спорен характер: два последователни аргумента могат да доведат до различни решения според два различни приоритета.⁴⁶

78. Това е така, понеже известният правен силогизъм е по-скоро начин за представяне на правното мотивиране отколкото негов формален правен. Той не отразява целият начин на мотивиране на съдията, който всъщност се състои от много на брой фактори за вземане на решения, не може да бъде формализиран *a priori*, и понякога е основан на неговата преценка: какви са относимите факти? Доказани ли са тези факти? Коя норма се прилага спрямо тях? Какво е значението на тази норма по отношение на делото, което ще бъде решавано? Кой източник следва да преобладава измежду множество противоречащи си източници? Общата последователност на съдебните решения никога не се постига и в по-голяма степен е въпрос на последващ отчет, който съдиите използват в своите мотиви, чиято функция е по-скоро да ги убеди във валидността на конкретно решение, отколкото да описва стриктно и обективно всички етапи, довели до постановеното решение.

79. Тази работа по тълкуване обаче е именно това, което техниките за машинно самообучение не правят – и не се опитват да правят – днес, тъй като, както виждаме, те извършват автоматизирано обработване, при което взаимовръзката на големи обеми информация се предполага да замести разбирането на реалните причинно-следствени връзки при едно решение. Те не се опитват да формализират правното мотивиране, но се надяват уловените от тях модели да могат да предвидят вероятните решения на даден съдия при подобни ситуации.

80. Резултатите, постигнати от изкуствените интелекти, са в действителност несвързани с въпроса за съответствието със закона на определено решение и не могат да разграничат законосъобразните от незаконосъобразните аргументи.

⁴⁵ По тези въпроси, Michel Troper (2001), *La théorie du droit, le droit, l'Etat*, PUF, Paris, спец. стр. 69-84.

⁴⁶ Xavier Linant de Bellefonds (1994), “*L'utilisation des systèmes experts en droit comparé*”, *Revue internationale de droit comparé*, Vol. 46, No. 2, стр. 703-718.

81. Преглед на труда на University College of London (UCL) по съдебната практика на Европейският съд по правата на човека потвърждава тази диагноза. Изследването на UCL е допуснало, че един прост автоматичен модел на обучени би могъл да предскаже изхода от едно дело със 79 % точност за този конкретен съд. Моделът на машинно самообучение се доказал като по-точен по отношение на частта, съдържаща описание на фактическата обстановка, от изследваните решения отколкото по отношение на мотивите, свързани с прилагането на Конвенцията по въпросния случай.⁴⁷

82. Разглеждането на условията, свързани с намирането на нарушение (с положително претегляне), и тези, свързани с ненарушаване на Европейската конвенция за правата на човека (с отрицателно претегляне), недвусмислено осветлява използваният механизъм, който по никакъв начин не може да бъде сравнен с правно мотивиране.

Фиг. 4: Илюстрация на труда на UCL – Теоретично тегло, което се дава на думите или термините съобразно тяхната връзка с констатациите на нарушение или липса на нарушение на Европейската конвенция за правата на човека

⁴⁷ Труд, базиран на 584 решения на Европейския съд по правата на човека: Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preoțiu-Pietro, Vasileios Lampos, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", публикуван на 24 октомври 2016 г., [онлайн], <https://peerj.com/articles/cs-93/> (страницата е прегледана на 14 декември 2017).



83. Изборът на съответните факти и тяхното тълкуване представлява един от елементите на решението на съдията. С други думи, изследването на UCL реално е могло да създаде вероятност само с лексикалния материал, който е до голяма степен изведен от обосновката и мотивите на съдията, а не със събрания от жалбоподателя на база само на честотите. По този начин, техният изкуствен интелект е създавал висока вероятност от съвпадане между групите от думи и решение, което вече е било формализирано, и би могло да породи само ограничен брой евентуални изходи. Той по никакъв начин не може самостоятелно да възпроизведе мотивите на европейските съдии, нито, преди всичко, да предскаже резултат на базата, например, на груб отчет за бъдещ жалбоподател пред съда на Страсбург, чиято жалба ще подлежи на много строг преглед за допустимост (почти 70 356 жалби са обявени за недопустими или са заличени от списъка през 2017 г.)⁴⁸ до голяма степен на база на прилагането на стандарти за оценка (важност и сериозност на оплакването и т.н.), оставящи значителна свобода на преценка при вземането на решение.

⁴⁸ Източник: Статистически анализ по ЕКПЧ за 2017 г., януари 2018 г.

5. Могат ли изкуствените интелекти да обяснят поведението на съдите в ретроспекция?

Едно последващо обяснение на поведението на съдия, по-специално разкриването на пристрастие, би изисквало всички потенциално причинни фактори да бъдат установени чрез тълкувателна рамка и контекстуализиран анализ. Фактът, че – статистически – грижата за децата по-често се възлага на майките отколкото на бащите, не показва пристрастие от страна на съдиите, а разкрива необходимостта да се мобилизират различни дисциплини от обществените науки, за да внесат яснота по този феномен.

84. От научна гледна точка, разясняването на един феномен или, доколкото става дума за нас, елемент от човешко поведение, представлява определяне на причинно-следствени механизми, довели до това поведение, което се използва определено количество контекстуални данни.

85. Това изисква по един много схематичен начин, предварителното създаване на тълкувателна рамка, която сама е изведена от повторното наблюдение на този тип събитие или поведение в присъствието на определени фактори или елементи. Тълкувателната рамка се състои от хипотезите или възгледите, възприети от различните дисциплини на обществените науки. Това е допълнителна аналитична стъпка, която може да бъде подадена в алгоритми, но която те не могат да изпълнят сами.

86. Някои правно-технически дружества отидоха още по-далеч и счетоха, че могат да установят евентуални лични пристрастия на съдиите и да подадат подозрения за пристрастие. Отворените данни на имената на определени председатели на административни съдилища и административни апелативни съдилища във Франция направиха възможно разработването на показател за процента отхвърлени жалби срещу задължения за напускане на френската територия, наложени от административните органи. Някои коментатори ревностно спорят, че по този начин безпристрастността на съдиите е поставена под съмнение от изкуствения интелект.

87. Могат ли обаче такива тълкувания действително да бъдат постигнати на базата на алгоритмично обработване на съдебни решения? За да има лично пристрастие в процесите по вземане на решение на съдиите (което се различава от техните лични и публични изявления в конкретния случай), тяхното поведение, или в този случай – тяхното решение – трябва да бъде определено от техния характер, мнения или религия. Както обаче беше казано, едно такова причинно-следствено обяснение не може просто да бъде изведено от вероятностния резултат, предоставен от алгоритмите. Обратното, то изисква допълнителна аналитична работа, за да се изолират, наред с многобройните взаимосвързани фактори (включително самоличността на членовете на съдебния състав), тези,

които са действително причинно-следствени. Например, фактът, че семеен съд статистически решава по-често децата да живеят с майка си, не отразява задължително пристрастието на съдията в полза на жените, а по-скоро съществуването на философски, икономически и дори културни фактори, които са специфични за юрисдикцията, като работно време на всеки от родителите, техния доход, наличието на колективна грижа за деца на местно ниво, независимо дали детето е в училище или не, независимо дали единият родител е или не е в нова връзка или дори просто липсата на интерес от който и да е родител да се грижи за малко дете.

88. Също така, решения за експулсиране от държава, постановено от административен съд, разположен близо до голям център за лишаване от свобода, не могат да бъдат сравнявани справедливо с тези на съд, който работи с такива спорове само от време на време.

89. В допълнение към това, независимо от местоположението на съда, особено интересен е въпросът за съдебната практика на еднолично заседаващ съдия, който само инцидентно работи с определен тип спорове, но който използва (или пренебрегва) съдебната практика на своите колеги, и той може да породи въпроса за равноправието на гражданите в съдебното производство. Фокусът обаче трябва да остане върху пре моделирането или запазването на колегиалния характер на съдебната система, а не класифицирането или дамгосването чрез инструменти за машинно самообучение.

90. Какво може да бъде извлечено от личността на председателя на съдебен състав в колегиален състав, когато името му е единствената номинална информация, видима в публично достъпни административни съдебни решения?

91. Освен това, как можем да отчетем два философски и културни подхода към съдебните решения, където в някои европейски държави, включително Франция, има култура на прецедента и подробни знания по съдии за фактическите бази данни на всички решения на първа и втора инстанция (база данни Ariane) в областта на административното правораздаване, докато други държави или системи предпочитат интелектуалната независимост на всеки съд, заедно с желанието към всяка ситуация да се подхожда индивидуално?

92. Поради това, предоставянето на точно обяснение на едно решение изисква много по-подробен анализ на условните данни във всеки случай и приложимите норми на закона, а не поддържане на напразна надежда, че множеството връзки ще се окаже смислено.

6. Как да се прилага изкуственият интелект в гражданското, търговското и административното правораздаване?

Състоянието развитие на техните на машинно самообучение не позволява днес да се постигнат надеждни резултати по отношение на „предсказването“ на съдебни решения. От друга страна, тяхното прилагане в областта на гражданското, търговското и административното правораздаване следва да се разглежда за създаването на мащаби или онлайн разрешаването при процеса на спора, когато остане възможна по-късна жалба пред съдията.

93. Както можем да видим, първият въпрос, повдигнат от това използване на изкуствен интелект, не е толкова дали е полезен или вреден, желателен или друг, но дали предложените алгоритми могат да постигнат търсения тип резултат. Заключениета от експериментите, проведени в апелативните съдилища в Дие и Рен във Франция, ясно показват, че в присъствието на провокиращо размисъл разискване, популяризирането на продукт на изкуствения интелект може да прикрие неприемливи недостатъци в проектирането и напълни грешни резултати от анализа.

94. Независимо от качеството на тествания софтуер, предугаждането на решенията на съдиите по граждански, търговски и административни дела би изглежда потенциално желана облага, макар и понякога поради много различни причини - и за отговорните за съдебната публична политика и за частните юристи.

95. Независимо каква е правната традиция на държавата, правната несигурност, т.е. рискът нечия правна претенция да бъде потвърдена или отхвърлена, извиква желанието да могат да се оценят количествено тези фактори, като се използват тези нови технологични приложения.

96. Адвокатите (или правния отдел на дружество) виждат възможността да използват тази технология, за да осигурят на своите клиенти по-добре информиран съвет, като емпирично и систематично оценяват шансовете за успех в процедурата, както и насърчават сключването на сделки, които, ако е необходимо, правят възможно избягването на дълъг и скъп процес. Някои застрахователи вече предлагат използването на прогностични системи на своите клиенти, за да оценят същината на тяхната дейност.⁴⁹

97. Същевременно, вземащите решения в обществената сфера лица виждат в това възможност по-добре да регулират потока на нови производства през съдилищата и да им предоставят лост за намаляване на съдебните оперативни разходи. Това се смята, че насърчава страните по съдебни спорове да използват алтернативни методи за разрешаване на спорове (помирение, медиация или арбитраж).⁵⁰

⁴⁹ Прогностичният софтуер се предлага на притежателите на полици на Allianz, например.

⁵⁰ В тази връзка, вж информационен доклад № 495 (2016-2017), изготвен от името на Правната комисия на Сената и внесен на 4 април 2017 г. от сенатор Филип Бас.

98. Подходът, който вече съществува в много съдебни системи за хармонизиране на решения по много дела чрез използване на скали (развод, уволнение, обезщетение за телесно увреждане), би могъл да бъде съживен чрез вероятностен или актюерски подход.⁵¹ Услугите за алтернативно онлайн разрешаване на спорове дори са създадени, за да помогнат да се оцени размер на обезщетението за малки спорове, наред с другото. Тези интересни подходи обаче не са безпристрастни и не трябва да лишават гражданите от достъп до съдия или да поставят под въпрос принципа на състезателност.

Експерименти, проведени във Франция

По инициатива на министерството на правосъдието, двете апелативни съдилища в Рен и Дуе се съгласиха да направят тест на софтуер за прогностично правосъдие по няколко съдебни жалби през пролетта на 2017 г., което всъщност беше анализ на граждански, социални и търговски решения на всички френски апелативни съдилища.

Макар тези вътрешни и изчерпателни данни за съдебната практика вече да бяха достъпни за тях безплатно от много години (базата данни JURICA), министерството ги предостави специално на разположение на публикуващото дружество, когато то предложи да оцени стойността на един количествено определен (иновативен) анализ на сумите, определяни от двете съдилища, в допълнение към географска класификация на несъответствията, отбелязани за подобни заявления и процеси.

Поради това, посочената цел на софтуера беше да създаде инструмент за вземане на решения, за да намали, ако е необходимо, прекомерната променливост в съдебните решения, в името на принципа на равенство на гражданите пред закона. Резултатът от експеримента, противоречиво дебатиран между двете апелативни съдилища, министерството на правосъдието и правно-техническото дружество, което е създадо продукта, за съжаление заяви липса на добавена стойност на тестваната версия на софтуера за работата на магистратите по отразяване и вземане на решения.

Което е по-съществено, бяха разкрити пристрастия в мотивите на софтуера, които доведоха да отклоняващи се или неподходящи резултати поради объркване между прости лексикални оформлени на съдебните мотиви и причинно-следствените връзки, които са били решаващи в мотивите на съдиите.

6.1. Нов изчислителен инструмент за изчисляване на скали

99. Процедурите за изчисляване на скали по различни граждански дела (например, обезщетение за телесно увреждане, добавка за компенсирани загуби и обезщетение при уволнение) изглеждат значително подобрени, когато са съчетани с други техники за обработване, при условията на многобройни мерки за проектиране и начини на използване

⁵¹ По отношение на актюерския подход, предложението на Case Law Analytics обявява повече оценка на риск отколкото прогноза за решение на спора.

(ефект на показност).⁵²

100. Важно е да се подчертае онова, което Жан-Пол Жан, председател на работната група за оценка на СЕРЕЈ, описва като качествено предизвикателство по време на конференция за отворените данни, проведена през 2016 г. във Франция: процедурата, извършена чрез машинно самообучение или какъвто и да е друг метод на обработване, следва да използва заверени оригинали, чиято непокътнатост е била проверена и които са били обогатени, за да се разграничат важните от маловажните.⁵³

101. Другият риск, на който вече са се натъкнали създателите на тези инструменти, е рискът от „data-snooping“ („лов на данни“), а именно избиране на данни само нагоре по веригата, които са от съществено значение за предварително определени мрежи за анализ, например чрез изключване от извадки на решения, които не служат добре като взаимовръзки на лингвистични последователности чрез машинно самообучение или какъвто и да е друг метод (например решения без представяне на процеса или малко мотиви).

102. Ако обаче изчисляваме скала, вече не правим ли малко предсказание? Граничната линия може да изглежда относително бегла, освен ако направим ясно разграничение между целта на процеса: целта тук не е да се предостави директивна информация, а да се предостави информация за фактически обстоятелства.

103. При условията на тези методологични и оперативни предпазни мерки, скалите са на разположение за разглеждане на определени спорове и са мощен инструмент за хармонизиране на съдебната практика. Преди скалите се изчисляваха на база на повече или по-малко ограничени извадки от решения, но инструментите с изкуствен интелект дават възможност да се проучват повече решения и могат да представят по-точни скали, като прилагат определен стандарт и оттам дават по-голяма тежест на резултатите.

6.2. Онлайн разрешаване на спорове

104. Всички европейски съдилища се изправят в различна степен пред граждански спорове на ниска стойност. Идеята за улесняване на

⁵² Ефектът на показност или само-осъществяване е рискът една система да създава един и същи резултат прогресивно, като влияе на онези, които подават изходна информация; този ефект често се споменава по отношение на съдебните скали, които, когато са информирани от решения, базирани на тези скали, са склонни да са представителни само за тях.

⁵³ J-P. Jean, “Penser les finalités de la nécessaire ouverture des bases de données de jurisprudence”, Конференция от 14 октомври 2016 г., проведена в Касационния съд, <https://www.courdecassation.fr/IMG///Open%20data,%20par%20Jean-Paul%20Jean.pdf>, страницата е прегледана на 17 март 2018 г.

процедурата чрез информационни технологии и/или възлагането им от съдилищата намира широка подкрепа. Великобритания, Нидерландия и Латвия са примери за държави, които вече са приложили или ще приложат тези типове повече или по-малко автоматизирани решения⁵⁴. За презгранични претенции, Европейският съюз е създал, чрез Регламент № 524/2013, обща рамка, която е на разположение в интернет (Европейски искиове с малък материален интерес).

105. Обхватът обаче на тези услуги за онлайн разрешаване на спорове (ОРС) изглежда, че постепенно се е разширил. Те са преминали от ограничени онлайн услуги към мерки за алтернативно разрешаване на спорове преди оплакването да е подадено пред съд и сега се въвеждат все повече в самия съдебен процес до момента на предлагане на електронни съдебни услуги.⁵⁵ Те не просто засягат спорове с нисък материален интерес, но и данъчни спорове или спорове, свързани с обществено-осигурителни услуги или производства по развод.

106. За онези, които се застъпват за тези решения, които са от интерес за редица юридически професии и частния сектор, достъпът до правосъдие би могъл да бъде значително подобрен с едно широко решение, съчетаващо ОРС и изкуствен интелект (или поне експертни системи, вж раздел 3 по-горе за разграничение). Идеята е да бъдат преведени оплакващите се лица през автоматично диагностициране на спора чрез поставяне на редица въпроси, които след това се обработват от машина, което води до предложението за решение. Работата на лабораторията Cyberjustice de Montréal, която обединява различните досъдебни и съдебни фази в компютърно базиран процес за спорове с ниска интензивност (например съдилища по дела с малък материален интерес в Квебек), е добра илюстрация за хибридизация.⁵⁶ Според създателите, има ясни ползи от гледна точка на ефективност и качество.

107. На каква база обаче би било изчислено каквото и да е обезщетение, предложено от такава система? Какъв метод? Алгоритъмът обработва ли

⁵⁴ Вж в този смисъл онлайн разрешаването на спорове, достъпно в Обединеното кралство - <https://www.judiciary.gov.uk/wp-content/uploads/2015/02/Online-Dispute-Resolution-Final-Web-Version1.pdf>. Виж също системата PAs в Нидерландия, която издава автоматични решения на база на дадени по-рано разрешения и която е породила съдебен спор на национално ниво и пред Съда на ЕС: има две дела (с-293/17 и с-294/17), заведени пред Държавния съвет (Council of State) на Нидерландия (фермери/ природозащитници срещу Нидерландия) за това да се определи дали е позволено на една система (Програма за регулиране на азота) да реши дали фермерите, наред с другото, нарушават директивата за хабитатите или не я нарушават. Наскоро Съдът на Европейския съд в Люксембург отговори на преюдициалните запитвания относно тези обединени дела (ECLI:EU:C:2018:882)

⁵⁵ Darin Thompson, "Creating new pathways to justice using simple artificial intelligence and online dispute resolution", Osgoode Hall Law School of York University.

⁵⁶ <http://www.cyberjustice.ca/projets/odr-plateforme-daide-au-reglement-en-ligne-de-litiges/>

информацията справедливо? Предвижда ли се предложението да бъде обсъдено на състезателна база с помощта на обучено и сертифицирано трето лице? Възможен ли е винаги достъп до съдия? Някои автори дори виждат широко разпространеното използване на тези методи за разрешаване на спорове като ново проявление на идеята, че за всеки проблем съществува цифрово решение (digital „solutionism“), т.е. систематичното използването на технологиите в опит да се разрешават проблеми, които не попадат задължително в рамките на техния обхват.⁵⁷ Следва да бъде отбелязано, също така, че в Европа наскоро беше въведено по-протективна регулаторна рамка, която е обвързваща за държавите членки: Член 22 от Плана за наблюдение на безопасността на данните изрично предвижда лицата да могат да откажат да бъдат субект на решение, което е основано изключително на автоматизирано обработване, с определени изключения.⁵⁸

⁵⁷ Evgeny Morozov, „Pour tout résoudre cliquez ici“ FYP editions, цитиран от David Larrousserie, „Contre le ‘solutionnisme’ numérique“, Le Monde, 6 октомври 2014 г., https://www.lemonde.fr/sciences/article/2014/10/06/contre-le-solutionnisme-numerique_4501225_1650684.html

⁵⁸ Член 22(1) от Регламент на ЕС 2016/679: „Субектът на данните има право да не бъде обект на решение, основаващо се единствено на автоматизирано обработване; предвидени са изключения (като съгласието на субекта на данни), но администраторът трябва да приложи „подходящи мерки за защита на правата и свободите, и легитимните интереси на субекта на данните“, включително „правото на субекта на данни на човешка намеса от страна на администратора, правото да изрази гледната си точка и да оспори решението“.

Вж, в същия смисъл, Конвенцията на Съвета на Европа за защита на лицата при автоматизираната обработка на лични данни, изменена с Протокола, приет през май 2018 г., когато последният влиза в сила. Член 9, параграф (1), буква а) предвижда принципа, че „Всеки има право да не бъде обект на решение, засягащо го в значителна степен, което се взема се единствено на основата на автоматизирано обработване данни, без да бъде отчетена неговата гледна точка. Независимо от този принцип на забрана, Член 9, параграф (2) посочва, че „параграф 1, буква а) не се прилага, ако решението е разрешено от закон, на който е подчинен администраторът и който предвижда и подходящи мерки за предпазване на правата, свободите и законните интереси на субекта на данни“.

Обяснителният доклад гласи следното (§75): „От съществено значение е всяко лице, което е вероятно да бъде обект на чисто автоматизирано решено да има правото да оспори това решение, като ефективно представи своята гледна точка и своите аргументи. По-специално, субектът на данни трябва да има възможността да докаже потенциалната неточност на личните данни преди тяхното използване, неадекватността на профила, който ще бъде приложен към тази конкретна ситуация или други фактори, които ще имат въздействие върху изхода от автоматизираното решение. Това е особено валидно, когато приложението на алгоритмично мотивиране, водейки до ограничаването на правото, отказа на социална помощ или оценката на капацитета изцяло на база на софтуера, води до дамгосване на лицата. Субектът на данни обаче не упражнява това права, ако автоматизираното решение е предвидено от закона, на който се подчинява администраторът, който предвижда подходящи мерки за защита на правата и свободите и законните интереси на субекта на данни“.

Вж също Член 9, параграф 1, буква в) от осъвременената Конвенция 108, който

108. Поради тази причина, потенциалните ползи от една система за ОРС, нейната степен на интеграция в пълен съдебен процес (от досъдебната фаза до същинската съдебна фаза) и почти решаващата роля на изкуствения интелект при изпълнението на процеса, трябва да бъдат надлежно оценени във всеки отделен случай.

109. ОРС вече предлага познание нагоре по веригата за съдебните процеси. Неговата роля явно е да допринесе за реализирането на услуги по помирение, медиация и арбитраж извън съдебната зала. Тези услуги могат да бъдат използвани и по време на състезателно производство под надзора на съдии преди те да вземат решение за изхода от споровете по същество (за някои спорове, тази фаза се счита за задължителна).

110. От друга страна, действителният принос на изкуствения интелект следва да бъде оценен. Дали е само въпрос за използване на машинно самообучение да се установят ориентировъчни скали или да се предприеме решение? Дали става дума наистина за изкуствения интелект, който се използва или експертна система или просто веригата на логическото правило? Във всеки случай, следва да е възможно да се съчетават тези системи с изискванията за прозрачност, неутралност и лоялност.⁵⁹

111. Накрая, важно е да се изследва и начинът, по който жалбоподателите се насърчават да използват системата: има ли потенциално объркване в самото име на това, което се предлага? Ако става дума за съд, това трябва да е формата на организация, определена от Европейската конвенция за правата на човека, а не просто частна правораздавателна институция, която само изглежда като държавно правораздаване⁶⁰. Ясно ли е осигурена възможност за обръщане към съдия? В Нидерландия, частните здравноосигурителни договори изглежда

предвижда правото на субекта на данни да „получи, по негово искане, информация за мотивите за обработването на данни, когато резултатите от това обработване се прилагат спрямо него”. Обяснителният доклад на осъвременената Конвенция (§77) гласи следното: „Субектите на данни имат правото да получат информация за мотивите, стоящи в основата на обработването на данни, включително последиците от тези мотиви и заключенията, които може да са били направени от тях, по-специално, когато се използват алгоритми за автоматизирано вземане на решение, по-специално в контекста на профилиране. Например, в случай на система за кредитен рейтинг, кредитополучателите имат правото да знаят каква е логиката зад обработването на техните данни, която води до решението да се отпусне или да се откаже кредит, а не просто да бъдат информирани за самото решение. Разбирането на тези елементи допринася за ефективното упражняване на други съществени гаранции, като правото на възражение и правото на жалба пред компетентния орган”. Това „трябва да получи полезна информация относно стоящата в основата логика” се намира и в ОРЗД (Член 13, параграф 1, буква е); Член 14, параграф 2, буква ж); Член 15, параграф 1, буква з)).

⁵⁹ Charlotte Pavillon, “Concerns over a digital judge”, nrc.nl, <https://www.nrc.nl/nieuws/2018/01/19/zorgen-om-populaire-digitale-rechter-a1588963>

⁶⁰ Scarlett-May Ferrié, Тествани алгоритми спрямо справедлив процес, документ, прегледан на <http://lexis360.fr>, свален на 09/07/2018 г., § 27-38

предвиждат автоматично прибегване към ОРС преди завеждането на какъвто и да е съдебен иск.

112. Работната група по медиация на СЕРЕЈ (СЕРЕЈ-GT-MED), стартирана през 2018 г., предложи първите си размишления по приноса на информационните технологии за методите за алтернативно разрешаване на спорове. Европейският комитет за правно сътрудничество (CDCJ) понастоящем извършва задълбочена работа по методите за ОРС, за да установи потенциала на тези инструменти, но също така всеки проблем сочи възможни нарушения на членове 6, 8 и 13 от Европейската конвенция за правата на човека.

6.3. Основните гаранции, които трябва да бъдат препотвърдени в гражданското, търговското и административното производство

Право на достъп до съд

113. Предоставянето на инструменти за онлайн разрешаване на спорове следва да не засяга правото на достъп по смисъла на член 6, дори макар това право да не е абсолютно и да се подчинява на имплицитни ограничения.⁶¹ По граждански дела, например, всяка страна по спор има правото да подава пред съда всеки спор, отнасящ се до нейните „граждански права и задължения“, разглеждани от съд.⁶² През 2015 г., Парламентарната асамблея на Съвета на Европа прие резолюция за „Достъп до правосъдие и интернет: потенциал и предизвикателства“, в която призовава да се гарантира, че „страните, ангажирани в процедури по ОРС запазват правото на достъп до процедура по съдебно обжалване, отговаряща на изискванията на справедлив процес съгласно член 6 от Конвенцията”.⁶³

Принцип на състезателност

114. Изглежда задължително определено количество количествена информация (например брой обработени решения, за да се получи скалата) и качествена информация (произход на решения, представителност на подбраните извадки, разпределение на решенията

⁶¹ Член 6, параграф 1 „1. Всяко лице, при решаването на... основателността на каквото и да е наказателно обвинение срещу него, има право на справедливо и публично гледане... от независим и безпристрастен съд; за ограниченията вж Deweer c. Belgique, § 49 ; Kart c. Turquie [GC], § 67.

⁶² CEDH, Golder c. Royaume-Uni, §§ 28-36

⁶³ Резолюция 2054 (2015) на Парламентарната асамблея на Съвета на Европа (ПАСЕ), 10 ноември 2015 г., <http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-EN.asp?fileid=22245&lang=en>

между различни критерии, като икономически и социален контекст) да бъде предоставена на разположение на гражданите и, преди всичко, на страните по процес, за да се разбере как са създадени скалите, за да се измерят възможните им граници и а да може да се обсъдят същите пред съдия.

Равенство

115. Използването на технологични средства не следва да причинява дисбаланси между страните, тъй като използването на цифрови средства би могло наистина да улесни производството за определени оператори (институции, дружества със средства, компютърно грамотни лица) и – обратното – да създаде трудности и за определени групи от населението, които са по-несигурни или по-слабо запознати с компютрите. Важно е някои лица да не бъдат оставени сами пред техните екрани и да бъдат информирани, че могат да потърсят правен съвет и че получават съдействие, когато такова им е необходимо.

Безпристрастност и независимост на съдиите

116. Възприето е като постулат, че нормата, произтичаща от преобладаващата тенденция, посочена по-горе в раздел 2.2., може да има непреки ефекти върху независимостта и безпристрастността на съдебната власт, по-специално в системи, в които независимостта на съдебната власт не е напълно постигната. В тези системи, ние не можем да изключим риска тези норми да окажат индиректно натиск върху съдиите, когато се вземат решения и искат тяхното одобрение, или изпълнителната власт да наблюдава онези, които се отклоняват от нормата.

Право на адвокат

В началото на тази глава споменахме предимствата, извлечени от прилагането на инструментите на прогностичното правосъдие за адвокати и, по-специално, възможността да се предоставя на техните клиенти по-добре информиран съвет чрез емпирично или систематично извършване на оценка на шансовете за успех на процедурата. Нека обаче си представим случай, в който шансовете за успех за страната по съдебния спор са крайно малки: може ли това да засегне решението на адвоката да помогне на своя клиент? Професионалната практика следва да се стреми да сведе до минимум риска лицата, които се нуждаят от правен съвет накрая да бъдат лишени от него.

7. Въпроси, които са специфични за наказателното правораздаване: превенция на престъпленията, риск от рецидивизъм и оценка на нивото на опасност

Макар да не са специално предвидени да бъдат дискриминативни, използването на статистика и изкуствен интелект в наказателното производство е показало риск от това да предизвика възраждане на детерминистични доктрини в ущърб на доктрините за индивидуализация на санкциите, които са широко възприети след 1945 г. в повечето европейски съдебни системи.

117. Използването на науката и технологията за изкуствен интелект по наказателни дела създава специфични предизвикателства, тъй като прилагането му може да отрази някои текущи публични дебати относно твърдяната предвидимост на престъпното поведение. Този дебат обаче изглежда да е напълно утихнал за около тридесет години в редица европейски държави. Например в Италия, член 220, алинея 2 от Наказателно-процесуалния кодекс изрично изключва използването на експертно заключение, за да се установят обичайните или професионални престъпни характеристики, тенденцията към извършване на престъпление, характерът и личността на обвиняемия и, принципно, психологическите качества на обвиняемия, независимо от патологичните причини. Във Франция, например, доктрината за „нова социална защита“, разработена от Марк Ансел, е базата на наказателното право: вместо един чисто наказателен и детерминистичен подход, е въведен система на обществена реабилитация, за да се предотврати извършването на престъпление чрез избягване на условията за престъпление. Този подход се споделя от редица европейски инструменти на наказателната политика, които се фокусират върху целта да се превъзпитат и реинтегрират извършителите на престъпления.⁶⁴

118. Поради това са създадени инструменти на наказателното правораздаване в съответствие с тези основополагащи принципи на реабилитация,⁶⁵ включително ролята на съдията в индивидуализацията на присъдата, на базата на обективни елементи на личностите (образование, заетост, редовни медицински и социални грижи) без никаква друга форма на анализ освен извършвания от специално обучени професионалисти като пробационни служители. Техники за анализ на големи данни биха могли да бъдат използвани от тези професионалисти, за да централизират и съберат информация за обвинения в леко наказуемо престъпление, която след това може да бъде съхранена от различни институции и агенции и след това би трябвало да бъде разгледана от съдия, понякога в рамките на много кратък период от време (например в контекста на съкратени съдебни процедури).

7.1. Инструменти, използвани от следствените органи

⁶⁴ Вж Европейски съд по правата на човека, Голям състав, *Vinter and Others vs. Обединено кралство*, параграфи 114-118

⁶⁵ От друга страна, следва да се насърчи използването на изкуствен интелект за лечение и реабилитация (напр. за събиране на данни за предписано лечение или методи на реинтеграция в затвора).

преди наказателното дело

119. Инструментите, описвани като „предотвратяващ контрол” (преди съдебния процес или преди препращане към съд), вече нарастват бързо и започват да стават познати от широката общественост (например, да помислим за списъка лица, които не се допускат по полет, който е всъщност приложени на анализ на големи данни, при което се събират и анализират данни за потенциални терористи, за да се предотврати извършването на действия, или алгоритми, използвани, за да се установи измама или изпиране на пари).

120. Като цяло, голям брой компютърни инструменти са широко използвани за предотвратяване на извършването на престъпни деяния (чрез установяване на възможните места, където това би могло да се случи или техните автори) или за по-ефективното им преследване.⁶⁶ Първата категория включва инструменти за „предотвратяващ контрол”, които се използват за превенция на определени типове престъпления с елементи на редовност в тяхното възникване, като кражба с взлом, насилие на улицата, кражба от/на автомобили. Наименованието на тези инструменти идва от способността им да определят точно къде и кога биха могли да бъдат извършени тези престъпления и да възпроизведат тази информация на географска карта под формата на горещи точки, които се наблюдават в реално време от полицейски патрули. Този процес се нарича създаване на прогнозна карта на престъпленията. По-голямата част от софтуера, използван в тази област, се базира на исторически доказателства за местоположенията на престъпленията, като полицейски доклади, но се тестват дори по-мощни и нови технологии, съчетаващи различни данни и от различни източници.⁶⁷ Тези инструменти, които имат

⁶⁶ Вж Ales Završnik, *Big Data, crime and social control*, стр.194 и сл., който изброява подробно редица инструменти, използвани от полицейските служби в Европа и Съединените американски щати.

⁶⁷ Например, като част от проекта „Електронна сигурност – ИКТ за основана на знанието и прогностична градска сигурност” (<http://www.esecurity.trento.it/>), който се провежда в италианския град Тренто в периода между ноември 2012 г. и май 2015 г., база данни, събираща информация за престъпления, сигнализирани на полицията, резултатите от проучвания, проведени от градския съвет за виктимизацията и реалната и възприеманата сигурност от гражданите, информация за физически и социални граждански размирици от полицията, както и други променливи във връзка със „SmartCity” („Интелигентен град”) (напр. информация във връзка със социално-демографския контекст, градската среда, нощното осветление, наличието на камери за наблюдение и обществен транспорт). Тя е създадена за подпомагане на работата по превенцията на престъпността и по подобряване на градската сигурност. Ръководителите на проекта свидетелстват за надеждността на използваните техники, които се казва, че са направили възможно да се предвидят престъпни деяния с успеваемост от приблизително 60-65 % и които се казва, че са помогнали за по-добра борба с престъпността когато има на разположение ограничени ресурси. В допълнение към това, тестове, проведени в Обединеното кралство като част от пилотен проект за предвиждане на евентуални локации на взлом, кражба и нападение,

много убедителни проценти на ефективност, също се претендира, че имат възпиращо въздействие спрямо извършването на престъпления в области около горещи точки, което води до положително мнение за публичните политики.⁶⁸

121. Въпреки това, трябва да бъдат реално оценени прогностичните възможности на тези инструменти, които показват техните ограничения по отношение на престъпления от по-нередовен характер или насочен към различни местоположения, като тероризъм. В допълнение към това, една от техните слабости е ефектът на „омагьосаните кръгове” и „самоизпълняващите се пророчества”: кварталите, които се считат за рискови, привличат по-голямо внимание от страна на полицията и полицията открива повече престъпления, което води до прекомерно полицейско наблюдение на общностите, живеещи в тях.⁶⁹ Накрая, въпросите относно евентуална „тирания на алгоритъма”, който би могъл да сведе до минимум или дори прогресивно да замени човешката преценка, не са напълно отсъстващи в рамките на самите полицейски служби, дори ако понастоящем технологията се представя като оставаща в служба на хората с цел да им осигури по-добра база за вземане на решение.⁷⁰

122. В допълнение към това, анализът на големи данни се прилага все повече в наказателното преследване на престъпленията. Инструменти като Connect, който се използва от полицията на Обединеното кралство, за да анализира милиарди данни, генерирани при финансови сделки, за да се намери взаимовръзки или модели на операции, или базата данни за Международна сексуална експлоатация на деца (ICSE DB), управлявана от Интерпол, която помага да се установят жертвите и/или извършителите чрез анализа, например на мебели и други предмети в изображения, показващи неправомерни действия, или анализа на фоновия шум във видеоклипове, са се оказали особено ефективни в борбата с престъпността. С Connect например, издирвания, които преди са изисквали месеци разследване, сега могат да бъдат изпълнени за минути, с много високо ниво на сложност и много голям обем от данни.

123. Все пак доктрината поставя под въпрос управленската логика на реакцията на престъпността от страна на тези прогностични инструменти,

показват, че използваните софтуерни предвиждания, наречени PREDPOL, са били точни в 78 % от случаите, в сравнение с 51 % чрез използването на традиционни техники.

⁶⁸ Един показател за географска концентрация на престъпността би помогнал на полицейските сили по-добре да отчетат факторите на околната среда, които повишават вероятността за престъплението в установената област (осветление, наличие на магазини и т.н.) и да планират адекватни реакции в консултация с други партньори.

⁶⁹ „Predicting crime, LAPD style”, The Guardian, 25 юни 2014 г.

⁷⁰ „How technology is allowing police to predict where and when crime will happen”, The Independent, 7 октомври 2017 г.

в която цялостният анализ на причините за престъплението става по-маловажен от това да се направи нещо тук и сега. Това се случва в момент, когато разполагаемите бюджети се свиват и полицията трябва да осигури същото ниво на защита на обществото, но с ограничен персонал, оборудване и ресурси.⁷¹

7.2. Инструменти по време на наказателното дело

124. Използването на прогностични инструменти⁷² от съдиите в наказателни дела е много рядко срещано в Европа.

125. HART (Harm Assessment Risk Tool – Рисков инструмент за оценка на вредите) е разработен в партньорство с Университета Кеймбридж и сега се тества в Обединеното кралство. Тази технология, базирана на машинно самообучение, е обучена като са използвани архивите на полицията в Дърам от 2008 г. до 2012 г. Учейки се от решенията, взети от полицейските служители през този период и независимо дали определени заподозрени са извършили повторно престъпление или не, машината се очаква да може да оцени риска – нисък, среден или висок – заподозрените да извършат повторно престъпление, на база на около тридесет фактора, някои от които не са свързани с извършеното престъпление (напр. пощенски код и пол).

126. При тестове, първоначално проведени през 2013 г., по време на които е наблюдавано поведението на заподозрените за двугодишен период след извършването на престъплението, се установи, че предсказанията на HART са с 98 % ефективност при предсказване на нисък риск и 88 % ефективност при висок риск от рецидивизъм. В тази експериментална фаза, HART ще има чисто консултантска стойност за съдията. В допълнение, полицията редовно ще извършва проверки на функционирането на HART и надеждността на неговите заключения.

127. Макар да е единственият прогностичен инструмент, установен в Европа към днешна дата, той предоставя възможността да се вземат предвид предизвикателствата, на които вземащите решения в обществената сфера лица биха могли да се натъкнат в близко бъдеще, ако този тип приложение бъде тестван в по-голям мащаб, особено в светлината на констатациите в Америка.

128. В Съединените американски щати,⁷³ НПО ProPublica е разкрила дискриминативните ефекти на алгоритъма, използван в софтуера COMPAS (Correctional Offender Management Profiling for Alternative

⁷¹ Ales Završnik, *Big data, crime and social control*, страница 196.

⁷² В литературата, тези инструменти често са наричани „алгоритмично правораздаване” или „автоматизирано правораздаване” или „симулирано правораздаване”.

⁷³ Изследване от 2015 г. е установило около шестдесет прогностични инструмента в Съединените американски щати.

Sanctions –Профилиране за корективно управление на извършители на престъпления за алтернативни санкции), чиято цел е да оцени риска от рецидивизъм, когато съдията трябва да определи присъдата за дадено лице.

129. Този алгоритъм, който е разработен от частно дружество и който трябва да бъде използван от съдиите в определени американски федерални щати, включва 137 въпроса, на които отговаря ответника, или информация, извлечена от криминални архиви. Въпросите са доста разнообразни и включват наличието на телефон в у дома, затруднения при плащането на сметки, семейна история, съдебно минало на обвиняемия и т.н.⁷⁴ Алгоритмът дава рейтинг на лицето по скала от 1 (нисък риск) до 10 (висок риск). Това е от помощ за вземането на решение от съда, като заключенията му са само една от променливите, които се вземат предвид от съдията при вземане на решение за присъдата.

130. На афро-американските популации е приписан висок риск от рецидивизъм – два пъти по-висок от този на другите популации, в рамките на две години от осъждането – без този ефект да е бил естествено търсен от създателите⁷⁵. За разлика от тях, алгоритмът е счел, че при другите популации изглежда много по-малко вероятно да извършат повторно едно престъпление⁷⁶. Трябва, разбира се, да се отбележи, че този тип

⁷⁴ Съществуват други алгоритми, които са разработени като се използват критични наблюдения, изразени от доктрината (вж следващата глава), които са основани на по-малки променливи, по-пряко свързани с извършеното престъпление и по-малко свързани с расата, пола или социално-икономическия статус. Пример за това е Инструмента за оценка на публичната безопасност (Public Safety Assessment Tool), използван в 30 американски юрисдикции.

⁷⁵ Този чисти дискриминативен ефект може всъщност да бъде обяснен с относително избираемата „калибровка“ на модела на алгоритма, който създава много „фалшиви позитивни лица“.

⁷⁶ Чернокожи популации по-често са класифицирани като високо-рискови, въпреки че те не са извършили повторно престъпление в рамките на две години от осъждането; бели популации, от друга страна, които по-често са класифицирани като ниско-рискови, са извършвали престъпления в рамките на следващите две години. Накратко, алгоритмът е надценил риска от рецидив за чернокожите и го е подценил за белите („фалшивите позитивни“ са били основно чернокожи, докато „фалшивите отрицателни“ са били основно бели). В отговорна твърденията на ProPublica, NorthPointe (сега Equivant след този спор) отговаря, че белите и чернокожите популации са еднакво представени при разглеждането на „действителните позитивни“, т.е. тези които реално са извършили повторно престъпление. **Въпросът как да се съчетаят точността на алгоритма при откриването на рецидив и необходимостта да се избегнат дискриминативни ефекти към чернокожото население е източник на интензивни дебати в литературата;** вж по-специално Chouldechova A (2016), „A fair prediction with a disparate impact: a study on bias in recidivism prediction instruments“, достъпна на <http://arxiv.org/abs/1610.07524>; както и „Bias in criminal risks scores is mathematically inevitable, Researchers say“, достъпна на <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>. **Този дебат също отразява въпроса доколкото е легитимно едно частно дружество, без**

подвеждащо тълкуване всъщност разкрива само обществената и икономическа крехкост на определени групи от населението, които очевидно не са криминогенни по своя характер. Изследователите в Dartmouth College се показали, също така, че този тип алгоритъм не създава добавена стойност, тъй като хора без никакво престъпно минало могат да възпроизведат точно същата оценка само като отговорят на въпросника.

131. В допълнение към това, липсата на прозрачност в оперативния процес на алгоритма, създаден от частни дружества (които претендират интелектуална собственост) беше още една причина за притеснения. Ако вземем предвид факта, че те вземат изходните си данни от самите държавни органи, липсата на отчетност от тяхна страна към гражданите създава голям демократичен проблем. Отчетите показват, че обществеността се информира за операциите с големи данни случайно, спорадично и когато има големи изтичания или грешки: пример за това е ситуацията, в която ProPublica разкри грешките в алгоритма на COMPAS след отказа на дружеството собственик да го сподели. Неправителствената организация трябваше да подаде жалба пред публичните органи за достъп до данните и да наеме свои собствен учен, който да прегледа алгоритма.

7.3. Предизвикателствата на „предсказването“ по наказателни дела

132. Ние видяхме в предишните раздели, че степента на развитие на прогностичните инструменти в Европа в областта на наказателното право е много разнообразна. Макар че инструментите, описвани като „предотвратяващ контрол“ бързо нарастват и дори започват да стават известни на широката общественост, не така стоят нещата при прилагането на този тип инструмент от съдиите по наказателни дела. Що се отнася до инструментите, с които разполагат прокуратурите, вече са изразени размишления за техните предимства и недостатъци. Нека сега да проучим инструментите, специфични за наказателните дела.

133. На първо място, важно е да се изключат аргументи, базирани изключително на ефективността или неефективността на тези инструменти. Дадените по-горе примери показват, че може да съществуват огромни възможности, но има и реални рискове при прилагането на нови технологии, които се използват без необходимите предпазни мерки. Вземащите решения в обществената сфера лица и съдебните заинтересовани страни трябва да бъдат особено бдителни и да играят активна роля в разработването на тези технологии; необходимо е

никакъв институционален контрол, да арбитрира между две противостоящи изисквания: това за защита на обществото, от една страна, и това за спазване на правата на лице, от друга страна.

постоянно наблюдение, за да се определи тяхната реална ефективност и да се избегнат непредвиден последствия. Това е дори по-важно в наказателното производство заради неговото пряко въздействие върху личните свободи на лицата.⁷⁷

134. Това означава, че и ползите и недостатъците на прилагането на тези инструменти в съдебната област следва да бъдат внимателно измерени.

135. Поддръжниците често твърдят, че те са неутрални и разчитат на фактически и обективни методи, които могат да помогнат правосъдието да стане по-точно и прозрачно. Твърди се, че друг голям техен актив е тяхната ефективност, която понякога надхвърля човешките възможности и може да бъде изключително ценна само в един общ контекст на намалени публични средства или дори недостиг на ресурси.

136. Включването на алгоритмични променливи, като история на съдимостта и информация за семейството, означава, че поведението в миналото на определена група може да определи съдбата на едно лице, което, разбира се, е уникално човешко същество със специфичен социален произход, образование, умения, степен на вина и отличителни мотивации за извършването на престъпление⁷⁸. Освен това, те спорят, че решенията на човека могат да бъдат основани на ценности и съображения (напр. обществени), които биха били запомнени от машината. Например, един съдия може да реши да наложи освобождаване под гаранция на извършител-жена, при която съществува риск от рецидив, на база на йерархия на ценностите, например като отдаде по-голямо значение на ролята ѝ на майка и защитник на децата ѝ, докато алгоритмът би бил в състояние да определи риска от повторно извършване на престъпление по-точно, но не би могъл да работи с тази йерархия на приоритети.

137. По наказателни дела, има потенциални рискове от дискриминация и когато някой счита, че тези инструменти, които се тълкуват и интерпретират от хора, могат да възпроизведат необосновани и вече съществуващи неравенства в съответната система на наказателно правораздаване; вместо да коригира определени проблематични политики, технологията може накрая да ги узакони. Както вече беше

⁷⁷ Извадка от решение на Върховния съд на Уисконсин в *Wisconsin v. Loomis* може също да даде вдъхновение на европейско ниво: „Важно е да се отчете, че инструменти като COMPAS продължават да се променят и развиват. Притесненията, за които говорим днес, могат да бъдат значително облекчени в бъдеще. Системата на наказателното правораздаване трябва да признае, че в идните месеци и години, ще станат налични допълнителни данни от изследвания. Могат да бъдат разработени различни и по-добри инструменти. Докато данните се променят, използването от наша страна на базирани на доказателства инструменти ще трябва също да се променя. Системата на правораздаване трябва да не изоставя от изследванията и да оценява постоянно използване на тези инструменти.”

⁷⁸ Aleš Završnik, “Big data, crime and social control”, страница 196.

посочено, НПО ProPublica⁷⁹ ясно разкри дискриминативните ефекти на използвания в COMPAS алгоритъм,⁸⁰ който предвиждаше, че при популациите на чернокожи съществува два пъти по-голям риск да извършат повторно престъпление в сравнение с популациите на белите, в срок от две години от осъждането, докато в същото време считаше, че при популациите от бели е много по-малко вероятно да бъде повторено престъплението. Те биха могли обаче да помогнат за разкриването на грешки при вземането на решения, така че същите да могат да бъдат коригирани.⁸¹ Освен това, липсата на прозрачност в процеса на изграждане на алгоритма от частни дружества и тяхната отговорност към обществото е основание за притеснение, още повече ако те са част от стъпки, предприемани от държавни органи, за да бъдат направени данни достъпни за обществото.

138. В светлината на гореизложеното, когато се използват алгоритми в контекста на наказателен процес, изглежда, че е от съществено значение напълно да се гарантира зачитане на принципа на равенство и презумпцията за невинност, установени от Член 6 от ЕКПЧ. Съответната страна следва да има достъп до и да може да оспори научната валидност на даден алгоритъм, теглото дадено на неговите различни елементи и всяко грешно заключение, до което достига когато даден съдия предположи, че може да го използва преди вземе своето решение. Освен това, въпросното право на достъп е обхванато и от основния принцип на защита на личните данни. Всички хора имат правото да не бъдат обекти на решения, засягащи ги в значителна степен, които са взети единствено на основата на автоматизирана обработка на данни, без тяхната гледна точка да е била взета предвид преди това.

139. В това отношение, съществува разлика между Европа и Съединените американски щати по отношение на достъпа до алгоритми: докато в Съединените американски щати съдебните органи все още не са склонни да признаят това право изцяло и претеглят частните интереси (особено защитата на интелектуалната собственост) спрямо правата на защита, в Европа рамката е по-протективна заради ОРЗД, който установява право на информация за логиката на решенията, взети посредством използване на алгоритми.⁸²

⁷⁹ www.propublica.org/article/technical-response-to-northpointe.

⁸⁰ Други алгоритми се фокусират върху други елементи, които са по-пряко свързани с извършеното престъпление.

⁸¹ Mojca M. Plesnicar and Katja Sugman Stubbs, „Subjectivity, algorithms and the courtroom”.

⁸² Член 15, параграф 1, буква з) от Регламент на ЕС 2016/679: „Субектът на данните има право да получи от администратора”... „следната информация:”... „съществуването на автоматизирано вземане на решения, включително профилирането, посочено в член 22, параграфи 1 и 4, и поне в тези случаи съществена информация относно използваната логика, както и значението и предвидените последствия от това обработване за субекта на данните”.

140. Изказаните по-рано съображения относно потенциално отрицателните ефекти на тези инструменти върху безпристрастността на съдията са валидни и по наказателни дела: съдия, който отсъди нещо различно от предсказаното от алгоритъма, е в вероятно да поема рискове, тъй като той поема по-голяма отговорност. Не изглежда нереалистично да си представим, че съдиите не биха желали да поемат това допълнително бреме, особено в системи, където техните мандати не са постоянни, а подлежат на пряко избиране,⁸³ или в които е вероятно да бъде ангажирана тяхната лична отговорност (дисциплинарна, гражданска или дори наказателна), особено ако техните законоустановени гаранции по дисциплинарни дела са недостатъчни.

8. Специфични въпроси във връзка със защитата на личните данни

Използването на алгоритми повдига въпроса за защитата на личните данни, когато се обработват. Принципът на вземане на предпазни мерки следва да бъде прилаган спрямо политиките за оценка на риска.

141. За да бъде използван пълният потенциал на алгоритмите при същевременно придържане към принципите на защита на данните, принципът на вземане на предпазни мерки следва да бъде прилаган и следва да бъдат въведени превантивни политики за борба с потенциалните рискове, свързани с използването на данните, обработени от тези алгоритми, и въздействието на използването им върху отделните лица и обществото като цяло.

142. Принципът на законност при обработването на лични данни и задължението за предотвратяване или свеждане до минимум на въздействието на обработването на данни върху правата и основните свободи на субектите на данни следва да предизвикат предварителна оценка на риска. Това следва да направи възможно внедряването на подходящите мерки, по-специално по време на етапа на проектиране (и оттук по проект) и по подразбиране, за да се намалят установените рискове.

143. Тъй като личните данни трябва да бъдат обработвани за посочени и законни цели, те не трябва да бъдат използвани по начин, който е несъвместим с тези цели, и не трябва да бъдат последващо обработвани по начин, който субектът на данни може да сметне за неочакван, неподходящ или спорен (принцип на лоялността). Поради това, въпросът за повторното използване на лични данни, което ги прави широко достъпни, трябва да бъде разгледан с най-голямо внимание.

⁸³ Mojca M.Plesnicar and Katja Sugman Stubbs, „Subjectivity, algorithms and the courtroom”.

144. Проектът на методите за обработване на данни, използвани от алгоритми, следва да сведе до минимум наличието на излишни или маргинални данни и да избегне всяка потенциална скрита грешка и всеки риск от дискриминация или отрицателно въздействие върху основните права и свободи на субектите на данни.

145. Когато се използва изкуствен интелект, правата на субектите на данни са от особена важност и контролът, който всеки от нас трябва да има върху нашата лична информация, подсказва, че трябва да е възможно да се упражнят следните права: правото на субектите на данни да не бъдат обект на автоматизирани решения, засягащи ги в значителна степен, без тяхната гледна точка да е била взета предвид, правото да получат информация относно мотивите за извършеното чрез алгоритми обработване на данни, правото да се противопоставят на това обработване и правото на правно средство за защита.

9. Потенциалът и ограниченията на инструментите за прогностично правосъдие

Терминът прогностично правосъдие следва да бъде изоставен, тъй като той е двусмислен и подвеждащ. Тези инструменти са основани на методи на анализ на съдебна практика, използващи статистически методи, които по никакъв начин не възпроизвеждат правни мотиви, но може да се опитват да ги описват. Аналитичните грешки, ако не могат да бъдат напълно отстранени, трябва да бъдат установени. Процесът на проектиране (създаване) и използването на инструмента, трябва да бъдат поставени в ясна етична рамка.

146. В раздел 3, вече подчертахме двусмислието и заблудата на концепцията за прогностично правосъдие и как тя извършва бавно изместване в колективното съзнание, като ни води да вярваме, че машините, лишени от всякаква емоция, един ден ще бъдат по-способни да направят акта на отсъждане по-надежден. Повече от всякога, неговите обещания трябва да бъдат разгледани по един обективен и научен начин, стъпил на здравата основа на фундаментално проучване, за да се установят възможните ограничения. В тази връзка, следва да бъде отбелязано, че рисковете от изопачени тълкувания на значението на съдебните решения са крайно високи, когато са базирани само на статистическо моделиране. Това наблюдение е допълнително подкрепено от липсата на прецизно разбиране на връзките между данните и очевидното присъствие на грешни взаимовръзки, които не могат да бъдат разграничени в големи масиви от данни.

147. Освен това, неутралността на алгоритмите е мит, тъй като техните създатели съзнателно или непреднамерено прехвърлят собствените си ценностни системи в тях. Философът, Ерик Садин, отбелязва, че зад ефективната им и неперсонифицирана фасада, алгоритмичните системи недоловимо отразяват намеренията на своите създатели или спонсори,

като предизвикват функциониране и асиметрична власт върху животите на други хора. Също така, изследователят Аурелиен Гросдидие счита, че един алгоритъм, сам по себе си, не е в състояние да направи нищо повече от това не позволи – в най-добрия случай – да уловим част от намерението на проектанта и поставя под съмнение цялата верига на обработване на информацията (намерение на проектанта, създаване на компютърен код, изпълнение на компютърен код и контекст на изпълнение и после поддръжка). Това наблюдение се споделя и от криминолога Алеш Завршник, който подчертава как етапите на изграждане и тълкуване на алгоритми се извършват от човека, за човека, и не могат да избегнат грешки, предразсъдъци, ценности, човешки интереси и човешко представяне на света, независимо как са създадени.

148. Въпреки тези значителни ограничения, следва ли да подминем приноса на технология с несравнима мощ? Самите математици К. С. Калуд и Д. Лонго подчертават в своето изследване за големи данни, че ограничителният или отрицателен обхват на техните резултати, както често се случва, не разрушава науката за данни, но павира пътя за повече размисъл, включително предизвикателството на един нов, по-разширен научен метод, който е в състояние да включи нови алгоритмични инструменти и класически инструменти, като придружи обработването със строга оценка на доказателствата. Както е подчертано във въведението, използването на изкуствен интелект е вероятно да предложи особено съществена подкрепа за професионалистите, включително съдии и адвокати, но и за широката общественост, особено ако един ден те направят възможно създаването на несравними инструменти за изследване и анализ на документи по законодателни, регулаторни, юриспруденциални и доктринални въпроси и създадат динамични връзки между всички тези източници. Но този тип приложение отива отвъд обхвата на този член, тъй като той не е създаден да предскаже изхода от спор, а да анализира съдебна практика в определен момент и пространство.

149. Както е обсъдено в раздел 6, при условията на представителност на избраните и обработени извадки, изкуственият интелект е помогнал за изработването на много по-прецизни скали на средните парични суми, разпределяни, съответно, в различни области (финансова издръжка, компенсаторни плащания, обезщетение за телесно увреждане, плащане при уволнение и т.н.). Тези скали, които са основани в по-голяма степен на консенсус, отколкото на осреднен анализ на това, което вече съществува, вече предоставят значителна подкрепа за вземането на решения и насока, без да могат да заменят самото право. Както беше посочено по-рано, рискът се състои в това, в отсъствието на статистическо представяне на действителността или на възможност да се предскаже каквото и да е, резултатите от софтуера за прогностично правосъдие да бъдат заложени като стандарти без никакво потвърждаване от правната система и в противоречие с нея.

150. Накрая, нека обмислим идеята да можем по свое желание да отстъпим от прогностичните системи. Вместо да заключваме потребителите в една вероятност (или набор от вероятности), идеята би била да им позволим да навигират през взаимовръзките, които са довели системата до предложението на нейната оценка и да могат да се дистанцират, като изберат други по-относими концепции или други от думи, или да изключат неверни взаимовръзки. Ако използваме примера на UCL, това би се състояло от предлагане на графично представяне на различните термини, запомнени от системата (с техните съответни тегла), за да се установи, че е имало нарушение (или липса на нарушение) и за да се разреши да бъдат поети други пътеки чрез предлагането на избор на други термини или лексикални групи.

151. Независимо колко смело и изкусително може да е това предложение, то предполага, че самите професионалисти (съдии, адвокати, университети) го поемат колективно, за да тестват възможността за прилагането му и че те не позволяват само на частните оператори, с изключение на няколко непроверени учени, да проектират софтуер и неясни или заключени режими на мотивиране или изчисляване.

152. Амбициозните (и неизпълнени обещания на някои правно-технически дружества не трябва да скриват необятния потенциал на технологиите и необходимостта от приложения, адаптирани и внедрени директно в научните и академичните научни среди, както и при всички юристи, като магистрати, деловодители, адвокати, нотариуси, съдебни изпълнители и експерти в определени области. Ще се окаже, че редица мерки могат да се възползват изцяло от тези нови инструменти чрез приложения, адаптирани и изработени в пряка връзка с изследователи и всички юристи, включително съдии, прокурори, деловодители, адвокати, нотариуси, съдебни изпълнители и експерти в областта.

153. В този динамичен контекст, изглежда от съществено значение, на първо място, да не се вземат прибързани решения и да се отдели време предварително да се обсъдят рисковете и практическите приложения на тези инструменти за съдебните системи, и същите да се тестват на първия етап. Съдебна система, която е в крак със своето време, би била такава, която може да установи, администрира и гарантира истинска кибернетика и за публичния и за частния сектор, и настоява за пълна прозрачност и справедливост при функционирането на алгоритми, които могат един ден да помагат в процеса на приемане на съдебни решения.

10. Необходимостта от задълбочен обществен дебат относно тези инструменти преди реализацията на публични политики за разработването им. Спешната необходимост кибернетиката да предостави рамка за разработването на алгоритми за изкуствен интелект, като

същевременно защита основните права

Предизвикателството да се интегрират тези инструменти в процес на приемане на съдебни решения обосновава опростяването на концепциите за съответната целева група. Трябва да бъде създадена етична рамка, за да насърчава бързото развитие на форма на изкуствен интелект, която включва механизми, предотвратяващи пристрастие и дискриминация в самия процес на проектирането му.

10.1. Важността на обсъждането, тестването и постоянното преглеждане на тези инструменти преди реализирането на публични политики

154. Въпросите, свързани с осъществяването на инструменти на прогностично правосъдие, са толкова многобройни и многоаспектни, че те изискват балансиран подход от вземащите решения в обществената сфера лица.

155. На първо място, е от съществено значение да се проведе публичен дебат по тези въпроси, който да събере на едно място създателите на инструментите и юристите. Съдебните съвети, професионалните асоциации на съдиите и адвокатските колегии могат несъмнено да допринесат за това и да помогнат да се намерят възможности и повече противоречиви аспекти. В допълнение към това, съдебното обучение и юридическите факултети могат да играят ключова роля в повишаването на информираността сред юристите по тези въпроси, така че да могат по-добре да разберат и да дадат своя практически принос за текущите развития.

156. Освен това, е от съществено значение да се извърши проучване на предложените приложения и те да бъдат тествани, за да се разберат техният потенциал и слабостите им, както и за може да бъдат доразработвани и адаптирани към нашите нужди. Правото на преглед на компонентите и характеристиките на инструментите, предлагани от частния сектор (или разработените от независими и специализирани публични институти, решение, което следва да бъде насърчено), изглежда също толкова важно, така че услугата по правораздаване да може ефективно да изпълни своята мисия. Следва да бъде проведена строга оценка на резултатите от тестовете преди по-широко разгръщане и интегриране в публична политика. Освен това изглежда силно препоръчително редовно да се оценява въздействието на тези инструменти върху работата на професионалистите в областта на правораздаването.

10.2. Създаването на етична рамка

157. Преди всичко, простото приемане на законодателна или регулаторна

рамка за изкуствения интелект изглежда напразно в един цифров контекст, който е присъщо транснационален по своя обхват. От друга страна, отделянето на сериозно внимание на естеството и качеството на отворените данни е вероятно да сведе до минимум риска от неадекватни препратки и да укрепят относимостта на резултатите от автоматизираното обработване. Що се отнася до имената на професионалистите, една проста предпазна мярка би била да се забрани публичното им разпространяване в структурирани сурови бази данни с оглед на рисковете от злоупотреба. Не би било въпрос за ограничаване на достъпа до информация, която е вече обработена (например, състава на съдебен състав), а за филтриране на суровите данни, които са безплатно предоставени. Накратко, трябва да се направи разграничение между достъп до информация и достъп до бази данни, които могат да бъдат произволно манипулирани.

158. Изследователите Бутарели и Мар подчертават как големите данни трябва да бъдат стриктно контролирани и защитавани. Други изследователи (Паскуале и Морозов) подчертават необходимостта да се установят прозрачни процедури за разгръщането на големи данни и, по-общо, изкуствен интелект в съдебната сфера, тъй като предлаганите решения никога не могат да представят живота в сложността му.

159. Развитието на правилата на кибернетиката за насочване на дейността на заинтересованите страни в сектора и за насърчване на горепосочените принципи на прозрачност, справедливост и неутралност на инструмента е от съществено значение. Редовното наблюдение от независими експерти следва да гарантира, че стимулите за изкуствения интелект, използвани за подпомагане на съдиите в техните решения, не са пристрастни. Не е неподходящо да се очаква прилагането, директно или не, на платени справочни системи (основани на модела за рекламиране на машини за търсене на Google), които позволяват на определени оператори да дадат по-малко тежест на решения, които са неблагоприятни за тях. Тези правила ще играят ключова роля при повишаването на доверието на гражданите в техните съдебни системи.

160. В тази връзка, качеството на най-добрите системи би могло да бъде признато чрез предоставянето на класификация или сертифициране. По-специално, те трябва да гарантират общата прозрачност и перфектната справедливост по начина, по който се обработва информация, както за професионалисти, така и за граждани, за да се предотврати повтарянето на грешки, като горепосочения алгоритъм на COMPAS. Юристите трябва да бъдат силно ангажирани, за да са в състояние надлежно да оценят рисковете и въздействието на тези приложения върху съдебните системи.

161. В днешно време, всички експерти, участващи в разработването на изкуствен интелект, включително изследователи, инженери и компютърни разработчици, имат едни изключителни и безпрецедентни отговорности. Тяхната работа би могла да бъде придружена от дори по-голямо

засилване на човешките науки. Примерът с някои иновативни училища на компютърни разработчици показва, че зад волята „да се хакне системата“ се крие, според някой наблюдател, всъщност прагматизъм, лишен от всякаква контекстуализация на отговорността, която сега тежи върху техниците с квази-демиургични сили. Хипократовата клетва със сигурност има своите граници в медицината, но ритуализира отговорността и осигурява етична рамка.

162. Накрая, кибернетиката трябва да бъде придружена от мащабно обучение на заинтересованите страни, от създателите на алгоритми и правно-технически дружества, на техните ползватели. Следва на всички да бъдат предоставени нови трансдисциплинарни човешки науки, така че изкуственият интелект да стане вектор на положително развитие за човечеството.

Приложение II

Кои са начините на използване на изкуствения интелект в европейските съдебни системи?

Това приложение към Хартата разглежда различните начини на използване на изкуствения интелект в европейските системи и насърчава в различна степен прилагането им в светлината на принципите на и ценностите, изложени в Етичната харта.

Използването на машинно самообучение за създаване на машини за търсене за подобряване на съдебната практика е възможност, която ще бъде предприета за всички юристи. Следва да бъдат обмислени допълнителни приложения (изготвяне на скали, поддръжка за алтернативни мерки за разрешаване на спорове и т.н.), но трябва да се вземат съответните грижи (по-специално, качеството на източника на данни, а не масово обработване на целия въпросен спор). Други приложения („прогностично правосъдие“) следва да бъдат прехвърлени в областта на изследването и допълнителното разработване (в консултация с юристи, за да се гарантира, че те изцяло се свързват с действителните потребности) преди да планират използване със значителен размах в публичната сфера.

По наказателни дела, това е много чувствителен въпрос, но той не следва да се пренебрегва. В светлината на много съществуващи въпроси относно тяхната съвместимост с определен брой основни права, използването на алгоритми за изчисляване на потенциалните рискове от рецидив на отделно лице, което е изправено пред съд, следва да бъде отчетен с възможно най-големи резерви. От друга страна, обработването на глобални количествени данни за превенция на престъпността е път, който трябва да бъде допълнително изследван с тези нови техники, като се отчитат известните грешки (ефекти на показност, качество на данните и т.н.). Също така, използването на алгоритми за формиране на по-добра връзка между типа налична обществена услуга и личността на лицето може да бъдат фактор за ефективността на една мярка от този вид.

Насърчавани начини на използване

- ▶ **Подобряване на съдебната практика:** техниките на машинно самообучение са все по-използвани в областта на обработването на естествения език през последните години (това включва първоначални усилия в разбирането на естествения език) и значителен актив за намиране на опции за търсене за допълване на настоящото пълнотекстово търсене от клавиатура. Тези инструменти биха могли да свържат различни източници (напр. конституции и конвенции, закони, съдебна практика и правна теория). Техните за визуализация на данни биха могли да илюстрират резултатите от търсения.
- ▶ **Достъп до право:** без да се замества човешката намеса, могат да бъдат създадени чатботи, които да улеснят достъпа до различните съществуващи източници на информация, като използват естествен език. Шаблони на документи (съдебни заявления, договори за наем и т.н.) също могат да се генерират онлайн.
- ▶ **Създаване на нови стратегически инструменти:** използването на науката за данни и техниките за изкуствен интелект спрямо данни за съдебна дейност може да помогне да се подобри ефективността на правосъдието, като даде възможност, например, да се извършват количествени и качествени оценки и да се правят прогнози (напр. бъдещи човешки и бюджетни ресурси). На тази база могат да бъдат изготвени ключови показатели за изпълнение. Препоръчва се юристите, и по-специално съдиите, да бъдат ангажирани във внедряване на тези инструменти, от гледна точка на поемане на отговорността за тези инструменти и за анализа на резултатите във връзка с фактори, свързани със специфичните характеристики на въпросния съд или качеството на правораздаването (например, необходимостта да се запази достъпът до правосъдие).

Възможни начини на използване, изискващи значителни методологични предпазни мерки

- ▶ **Помощ при изготвянето на скали в определени граждански спорове:** анализът на всички съдебни решения не е статистически от значение, ако не бъдат установени всички причинно-следствени фактори (изрични и подразбиращи се в решенията). Знанието, че средното обезщетение, присъждано в определена географска област, е по-високо отколкото в друга, може да бъде обяснено не с оглед поведението на съдиите, но в светлината на характеристиките на въпросната област. Поради

това машинното самообучение може да бъде полезно за установяването на решения (вж *Подобряване на съдебната практика* по-горе), но само автоматизираното обработване на данни не може да предостави смислена информация. Една съществена предпоставка е събирането на относима извадка от решения, които да бъдат обработени (например, чрез проучвания).

- ▶ **Подкрепа за мерки за алтернативно разрешаване на спорове по граждански дела:** в някои европейски държави инструментите на „прогностичното правосъдие“ се използват от застрахователни дружества, за да преценят шансовете за успех на даден спор и за да насочат страните по спора към друг метод на разрешаване на спора, когато е налице усещането, че има малки шансове за успех. Освен това, някои системи в чужбина предлагат суми за обезщетение без никаква реална прозрачност във връзка с правилата за изчисляване. Тези системи обаче не могат да бъдат разглеждани като безпристрастни и надеждни (вж раздела за техники за машинно самообучение). Вземат се решения за даден гражданин, като се използват окастрени бази. В други случаи, страна по спор може да бъде посъветвана, чрез виртуален агент (чатбот), да избере определена мярка за алтернативно разрешаване на спорове след предварителен преглед на критериите, въведени от самата страна по спора, при посещение на уебсайт на съд или търсене на правна информация онлайн. Когато е подходящо, виртуалният агент може да препоръча и страната да потърси консултация със служба за медиация или адвокат. Във всички тези случаи, присъствието на обучено трето лице (медиатор, използващ не само техники, но може би скали, изчислени по посочения по-горе начин, или адвокат) би изглеждало най-подходящото решение на този етап.
- ▶ **Онлайн разрешаване на спорове:** когато страните отидат на онлайн платформа за разрешаване на спорове, те следва да бъдат информирани по ясен и разбираем начин дали обработването на техния спор се извършва по изцяло автоматизиран начин или с участието на медиатор или арбитър. В допълнение към това, информацията, дадена на страните, трябва да бъде честна и трябва да се избягва да се създава у тях впечатлението, че участва съд (в тази връзка, терминът „онлайн съд“ се използва често за този тип платформи, докато технически тяхната цел е да предоставят услуги по алтернативно разрешаване на спорове). Това са два много съществени фактори, когато се дава възможност на страните по спора да направят информиран избор, като евентуално не се съгласят със съвета и решат да отидат на истински съд по смисъла на член 6

от ЕКПЧ. Освен това, с оглед на изискванията на членове 6 и 13 от ЕКПЧ, следва винаги да се обмисли въвеждането на форми на контрол върху процедурата за онлайн разрешаване на спорове и резултата от нея от държавните съдилища, особено когато страната се е съгласила с напълно автоматизирано онлайн разрешаване на спорове.

- ▶ **Използването на алгоритми в криминални разследвания, за да се установи къде се извършват престъпления:** този вид приложение би могъл да касае не само полицията, но и прокурорите в органите за превенция на престъпността, от които те са част. В Съединените американски щати са използвани системи, за да насочват полицейски патрули в реално време към евентуални места, където се извършват престъпления. Този тип количествен подход обаче може да генерира много силен „ефект на показност“ (на дадено място, има по-голям шанс от откриване на престъпление и това следователно подсилва системата). Наказателно-правен анализ чрез подходи, съчетаващи географски информационни системи (ГИС) и големи обеми от данни за процедури, би могъл по-скоро да бъде споделен с прокурорите и би могъл със сигурност да се ползва от значителен принос от машинно самообучение. Звената за борба с изпирането на пари вече използват „прогностични“ системи, за да установят подозрителни финансови потоци, но в случая на количествена (финансова) информация, машините демонстрират по-голяма способност да генерират надеждни резултати. Изследователите следва също така да имат по-добър достъп до тези данни, за да представят съответни изследвания за създаващите политики лица.

Начини на използване, които да бъдат обмислени след допълнителни научни проучвания

- ▶ **Профилиране на съдии:** количественото определяне на дейността на съдия ще разкрие по-малко за каквито и да е потенциални пристрастия отколкото за каквито и да е външни фактори, влияещи на решенията му. Самият съдия не е причината, поради която съдебната дейност в една бедна област не дава същите резултати както на друга територия, независимо каква е личността му. Когато решението е постановено колегиално и без възможност съдия да изрази разминаващо се мнение, е безсмислено да се профилира всеки от съдиите от колегията. От друга страна, предлагането на съдиите на по-подробна количествена и качествена оценка на техните дейности, благодарение на нови инструменти, но с чисто информативна цел за подпомагане на вземането на решение и

за тяхно изключително ползване, следва да се насърчи.

- ▶ **Предугаждане на съдебните решения:** самостоятелното статистическо обработване на лексикални групи разкрива честотата на използване на определени групи от думи, но не установява реалните причини за едно решение и не извършва правен анализ (виж проучването, извършено по отношение на решенията на ЕКПЧ от University College of London, което представя по-добри резултати по фактите, отколкото по анализа на закона). Хибридни системи, основани на тълкуването на математически модели, които се предполага че представляват разнообразния диапазон от мотиви на съдиите, не са много по-ефективни, тъй като все пак са ограничени от пристрастието в извадката от данни, които те са обработили, и трябва да започнат отново от първото квадратче, ако се измени закон или ако има отмяна в съдебна практика.

Начини на използване, които да бъдат обмислени с най-крайни резерви

- ▶ **Използване на алгоритми по наказателни дела, за да се профилират лица:** експериментите в други държави (COMPAS в Съединените американски щати и HART в Обединено кралство) са критикувани от неправителствените организации (вж труд на ProPublica в Съединените американски щати и Big Brother Watch в Обединеното кралство). С оглед на ограниченията на използваната методология, този чисто статистически подход е довел до погрешни резултати: констатацията, че някои лица от афро-американски произход са по-често замесени в престъпни деяния е довел до по-висок рисков фактор за цялото афро-американско население. По този начин, дори за дребни престъпления, тези системи са направили отрицателно претегляне на ответниците от афро-американски произход, резултатът от което е несправедливото увеличаване на размера на техните присъди. Този подход, който има дискриминативни и детерминистични ефекти, трябва да бъде заменен с подход, който проявява повече уважение към европейските стандарти за наказателно-правните санкции и който трябва да предлага на лицето възможността за рехабилитация и реинтеграция. Ако алгоритмичните системи помагат за подобряване на събирането на информация за пробационните служби, например, и правят възможно събирането на съответната информация по по-бърз начин за последващо обработване от човек, тогава определено би бил постигнат напредък (по-специално в съкратени производства). Всяко друго използване е предразположено към пристрастия, което ще влезе в конфликт с определени

национални и наднационални основни принципи.

- ▶ **Основана на количеството норма:** тук не става дума само за създаване на скали, които биха могли да бъдат законни, но за предоставяне на всеки съдия на съдържанието на решенията, постановени от всички други съдии и претенция да се заключи бъдещият му избор в обхвата на тези „прецеденти“. Този подход следва да бъде отхвърлен, тъй като този голям брой не може да допълни или да замести закона. Поради представените по-горе причини (*Помощ при изготвянето на скали*), основаният на количество подход не е начина, по който да се действа. Изследването на СЕРЕЈ подчертава и опасностите от кристализирането на съдебна практика и потенциално отрицателните ефекти върху безпристрастността и независимостта на съдиите.

Приложение III

Речник

Този речник предоставя определение на термините, използвани от Етичната харта и документа по изследването. Предпочетено е да се даде тясно определение за всички използвани термини. Всички документи трябва да бъдат четени и разбирани в светлината на тези определения.

АЛГОРИТЪМ Крайна последователност от формални правила (логически операции и инструкции), които правят възможно да се получи резултат от първоначално въвеждане на информация. Тази последователност може да е част от автоматизиран процес на изпълнение и стъпва на модели, създадени чрез машинно самообучение.

АНОНИМИЗИРАНЕ Метод за обработване на лични данни, за да се предотврати напълно и необратимо идентификацията на физическо или юридическо лице. Поради това анонимизирането означава, че вече няма никаква възможна връзка между съответната информация и лицето, за което се отнася тя. След това идентификацията става напълно невъзможна.⁸⁴ Тъй като принципите, свързани със защита на данните, се прилагат спрямо цялата информация във връзка с идентифицирано или идентифицируемо физическо лице, те не се прилагат спрямо анонимизирани данни.

ИЗКУСТВЕН ИНТЕЛЕКТ (ИИ) Набор от научни методи, теории и техники, чиято цел е да възпроизведат чрез машина, познавателните способности на хората. Настоящите разработки имат за цел машините да извършват сложни задачи, които преди това са изпълнявани от хора.

Терминът изкуствен интелект обаче е критикуван от експертите, които правят разграничение между „силни“ изкуствени интелекти (вече способни да контекстуализират специализирани и разнообразни проблеми по напълно автономен начин) и „слаби“ или „умерени“ изкуствени интелекти (високи резултати в областта, в която са обучени). Според някои експерти „силните“ изкуствени интелекти биха изисквали значителен напредък в базовото изследване, а не само прости подобрения в работата на съществуващите системи, за да могат да моделират света като цяло.

⁸⁴ Работна група по Член 29 от Становище 05/2014 относно Техническите способности за анонимизиране. Вж също № 26 от Регламент (ЕС) 2016/679 на Европейския парламент и на Съвета от 27 април 2016 г.

Инструментите, идентифицирани в този документ, са разработени посредством методите на машинно самообучение, т.е. „слаби“ изкуствени интелекти.

ГОЛЕМИ ДАННИ (метаданни, комплекти от големи данни) Терминът големи данни се отнася до големи набори от данни от смесени източници (напр. отворени данни, служебни данни и търговски закупени данни). За данни, изведени от съдебна дейност, големите данни могат да бъдат съчетание на статистически данни, записи за свързвания на бизнес софтуер (дневници за приложения), бази данни със съдебни решения и т.н.

ЧАТБОТ (агент за разговори) Агентът за разговори, който разговаря със своя ползвател (например, емпатични роботи, използвани за подпомагане на хора със заболявания, или автоматични услуги за разговори в обслужването на клиенти).⁸⁵

ДАННИ Представяне на информация за автоматично обработване. Когато е казано, че алгоритми могат да бъдат „прилагани“ спрямо най-разнообразните реалности в правния свят или другаде, се предполага „възможност за цифровизация“ на всяка среда под формата на „данни“. Но от физиката е ясно, че нищо не ни казва, че физичните процеси могат да бъдат адекватно преведени в „данни“ (и интегрирани във входния/изходния цикъл на алгоритми). Ако във физиката вече е така, няма причина това да не е приложимо и в социалните отношения. Поради това ние трябва да бъдем внимателни с идеята за „данни“, която винаги допуска, че реалността, която се опитваме да опишем има такъв формат, който по естествен начин може да бъде обработен алгоритмично.

БАЗА ДАННИ Базата данни е „контейнер“, който съхранява данни, като цифри, дати или думи, които могат да бъдат повторно обработени чрез използване на компютър за представяне на информация, например, събиране и сортиране на цифри и имена за формиране на директория.

ИЗВЛИЧАНЕ НА ДАННИ Извличането на данни дава възможност да се анализира голям обем данни и да се подчертаят модели, взаимовръзки и тенденции.

НАУКА ЗА ДАННИТЕ Голяма област, която събира заедно математици, статистици, вероятности, обработване на данни и визуализиране на данни, за да се получи разбиране от смесен набор от данни (образи, звук, текст, данни за генома, връзки между социални мрежи, физически измервания и т.н.).

Методите и инструментите, разработени от изкуствен интелект, попадат в тази категория.

⁸⁵ Доклад на CNIL за декември 2017 г.: Как могат хората да запазят своето преимущество? Етичните въпроси, повдигнати от алгоритмите и изкуствения интелект.

ДЪЛБОКО ОБУЧЕНИЕ Вж *Машинно самообучение* и *Неврони*

ЕКСПЕРТНА СИСТЕМА Това е един от начините да се постигне изкуствен интелект. Експертната система е инструмент, който е в състояние да възпроизведе познавателните механизми на даден експерт в конкретна област. По-точно, това е софтуер, който може да отговори на въпроси, с аргументация, основана на известни факти и правила. Той се състои от 3 части:

- ▶ фактическа база;
- ▶ нормативна база;
- ▶ машина за заключения.

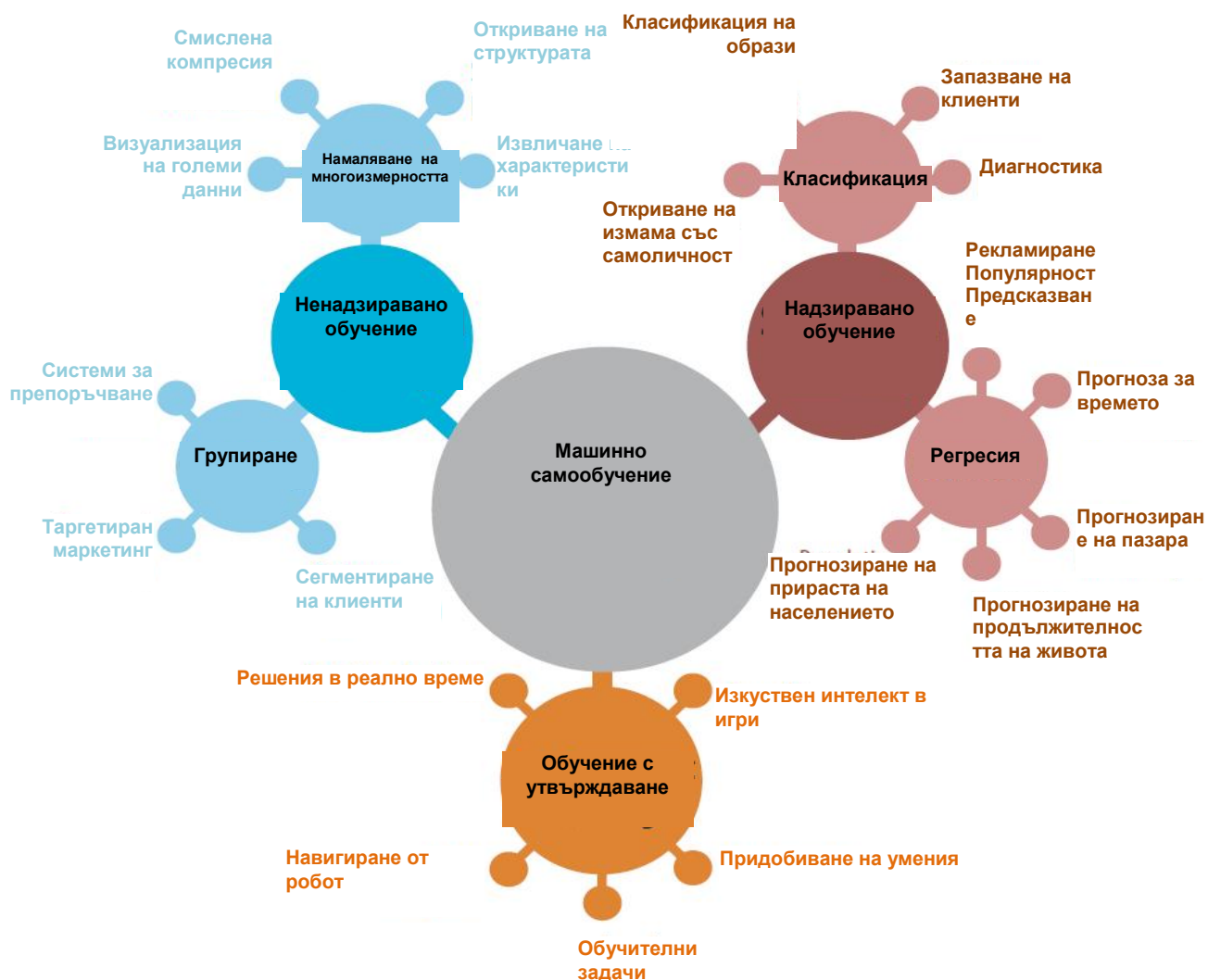
Машината за заключения е в състояние да използва факти и правила, за да създава нови факти, докато стигне до отговора на зададения експертен въпрос.

Повечето съществуващи експертни системи са основани на механизмите на формалната логика (Аристотелова логика) и използват дедуктивно мотивиране.

ПРАВНО-ТЕХНИЧЕСКИТЕ дружества, използващи информационни технологии в областта на правото, за да предлагат иновативни правни услуги. Тези дружества са стартъпи, специализирани в правото. Появяват се и други термини, изведени от бизнес сектори, като *Fintechs* (*Финансово-технически*) за стартъпи, предлагащи финансови услуги и *Medtechs* (*Медицинско-технически*) в медицинската област.

МАШИННО САМООБУЧЕНИЕ Машинното самообучение прави възможно създаването на математически модел от данни, включващи голям брой променливи, които не са предварително известни. Параметрите са конфигурирани постепенно по време на фазата на обучение, която използва набори от обучителни данни, за да намери и класифицира връзки. Различните методи на машинно самообучение са избрани от създателите в зависимост от характера на задачите, които трябва да бъдат изпълнени (групиране). Тези методи са обичайно класифицирани в три категории: надзиравано (от човек) обучение, ненадзиравано обучение и обучение с утвърждение. Тези три категории се събират различни методи, включително невронни мрежи, дълбоко обучение и т.н.

Графиката по-долу илюстрира различните категории машинно самообучение:



МЕТАДАНИ Данни, които дават възможност да се определят, контекстуализират или опишат други данни. В повечето от техните компютърни ползватели, мета префиксът означава „референтно определение или описание”.

Метаданните синтезират основна информация за данни, те улесняват търсенето и манипулирането на конкретни прояви на данни. Авторът, датата на създаване, дата на изменение и размерът на файла са примери. Метаданните и техният резултат, филтриране на данни, помощ за намиране на специфичен документ.

НЕВРОНИ/НЕВРОННА МРЕЖА Невронните мрежи са компютризирани системи, които са смътно вдъхновени от биологичните невронни мрежи, съставляващи животинските мозъци.[1] Тези системи „се обучават” да изпълняват задачи, като вземат предвид примери, принципно без да бъдат програмирани с каквито и да е специфични за задачата правила. Например, при разпознаването на образ, те могат да се обучават да идентифицират образи, които съдържат котки, като анализират примерни образи, които са били ръчно етикетирани като „котка” или „различно от

котка” и използват резултатите, за да идентифицират котки в други образи. Те правят това без никакво предходно познание относно котки, напр., че те имат козина, опашки, мустаци и подобни на котки лица. Вместо това, те автоматично генерират идентифициращи характеристики от обучителния материал, който те обработват.

ANN (ИНМ) е основана на събирането на свързани звена или възли, наричащи се изкуствени неврони, които са схематичен модел на невроните в един биологичен мозък. Всяко свързване, като синапсите в един биологичен мозък, могат да предават сигнал от един изкуствен неврон към друг. Един изкуствен неврон, който получава сигнал, може да го обработва и след това да подаде сигнал на допълнителни изкуствени неврони, които са свързани с него.

Първоначалната цел на подходът с ИНМ е бил да се разрешават проблеми по същия начин, по който би го разрешил човешкият мозък. Във времето обаче вниманието се е изместило към изпълнението на специфични задачи, което е довело до отклонения от биологията. Изкуствените невронни мрежи са използвани по разнообразни задачи, включително компютърно виждане, разпознаване на глас, машинен превод, филтриране в социална мрежа, настолни и видео игри и медицинска диагноза.

ОТВОРЕНИ ДАННИ Терминът се отнася до предоставянето на структурирани бази данни за публично сваляне. Тези данни могат да бъдат използвани повторно на относително ниска цена, при условията на конкретен лиценз, който може – по-специално, да предвижда или да забранява определени цели на повторно използване.

Отворените данни следва да не се бъркат с единна публична информация, която може да бъде намерена на уебсайтове, когато цялата база данни не може да бъде свалена (например, база данни със съдебни решения). Отворените данни не заменят задължителното публикуване на специфични административни или съдебни решения или мерки, които вече са заложили от определени закони или подзаконови нормативни актове.

Накрая, понякога има объркване между данни (в строг смисъл отворени данни) и методите за тяхното обработване (машинно самообучение, наука за данните) за различни цели (машини за търсене, съдействие при изготвянето на документи, анализ на тенденции на решения, предсказване на съдебни решения и т.н.).

СОФТУЕР С ОТВОРЕН КОД Софтуер, програмният код за който е достъпен за всеки. Поради това, софтуерът може да бъде свободно използван, изменян и редистрибутиран.

ЛИЧНИ ДАННИ Всяка информация за идентифицирано или идентифицируемо физическо лице („съответното лице”), пряко или косвено.

Тук се включват чувствителни данни, свързани с генетични данни, биометрични данни, които по уникален начин идентифицират едно физическо лице, данни във връзка с престъпления, наказателно производство и присъди и свързани мерки за сигурност, и всякакви данни за информация, която те разкриват относно расов или етнически произход, политически мнения, членство в профсъюз, религиозни или други вярвания, здраве или сексуален живот.

ПРОГНОСТИЧНО ПРАВОСЪДИЕ Прогностичното правосъдие е анализът на големи количества съдебни решения чрез технологии за изкуствен интелект, за да се направят предвиждания за резултата от определени видове специализирани спорове (например, плащания при уволнение или пенсия за издръжка).

Терминът „прогностично“, използван от правно-технически дружества, идва от клоновете на науката (основно статистика), което дава възможност да се предскажат бъдещи резултати чрез индуктивен анализ. Съдебните решения се обработват с оглед на откриването на взаимовръзки между изходните данни (критерии, изложени в законодателството, фактите по делото и мотивите) и резултатните данни (формално решение като размера на обезщетението).

Взаимовръзките, които се считат за относими, дават възможност да се създадат модели, които – когато се използват с нови изходни данни (нови факти или точности, описани като параметър, напр. продължителност на договорните взаимоотношения) – създават според разработчиците им прогноза за решенето (например, диапазона на обезщетението).

Някои автори критикуват и формата и същината на този подход. Те твърдят, че, по принцип, математическото моделиране на определени социални феномени не е задача, съпоставима с други по-лесно определими количествено дейности (изолирането на действително причинно-следствените фактори на едно съдебно решение е безкрайно по-сложно от играенето на играта Го или разпознаването например на изображение):

тук, съществува много по-висок риск от фалшиви взаимовръзки. В допълнение към това, в правната теория, две противоречащи си решения могат да се окажат валидни, ако правните мотиви са достатъчно логични. Следователно, даването на прогнози би било едно чисто информативно упражнение без никаква претенция за обвързващ характер.

ПРОФИЛИРАНЕ Техника на автоматизирано обработване на данни, която се състои от прилагане на „профил“ спрямо физическо лице, по-специално, за да се приеме решение за него или за да се анализират или предскажат лични предпочитания, поведението и нагласи.

ОБРАБОТВАНЕ НА ЛИЧНИ ДАННИ Съгласно член 2 от ревизираната Конвенция 108, „обработване на данни“ означава всяка операция или набор от операции, извършено спрямо лични данни, като събирането,

съхранението, запазването, изменението, извличането, разкриването, предоставянето, изтриването или унищожаването на такива данни, или извършването с тях на логични и/или аритметични операции.

ПСЕВДОНИМИЗИРАНЕ Съгласно член 4 от ОРЗД, това е обработването на лични данни по такъв начин, че те повече да не могат да бъдат свързани с конкретен субект на данни без да се използва допълнителна информация, при условие че тази допълнителна информация се държи отделно и е предмет на технически и организационни мерки, за да се гарантира, че личните данни не са свързани с идентифицирано или идентифицируемо физическо лице.⁸⁶

⁸⁶ Член 4 от Регламент (ЕС) 2016/679 на Европейския парламент и на Съвета от 27 април 2016 г.

Приложение IV

Контролен списък за интегриране на принципите на Хартата във вашия метод на обработване

За да се оцени съвместимостта на Вашия метод за обработване с Хартата, има скала за самооценка за всеки от изброените принципи.

За всеки принцип, поставете отметка в кутията, която съответства на Вашите методи на обработване.

Кутията, разположена най-отляво, указва пълно интегриране, кутията най-вдясно указва липса на интегриране.

В дъното на листа с оценка (в полето, обозначено с „Общо“), въведете сумата на отметнатите кутии. Колоната с най-висок резултат показва нивото на съвместимост на Вашия метод на обработване с Хартата.

Тази оценка, разбира се, е изцяло информативна и по никакъв начин не се равнява на никакво сертифициране.



Контролен списък за оценка на вашите методи на обработване

Принцип на основните права:

Да се гарантира, че проектирането и реализацията на използващи изкуствен интелект инструменти и услуги са съвместими с основните права, включително правото на защита на личните данни

Контролен списък за оценка на Вашите методи за обработване

Принцип на недискриминация:

Специално да се предотвратява развитието или засилването на каквато и да е дискриминация между физически лица или групи от физически лица

Контролен списък за оценка на Вашите методи за обработване

Принцип на качество и сигурност:

По отношение на обработването на съдебни решения и данни, да се използват удостоверени източници и невеществени данни с модели, изработени в сигурна технологична среда

Контролен списък за оценка на Вашите методи за обработване

Принцип на прозрачност, безпристрастност и справедливост:

Да се правят методите за обработване на данни достъпни и разбираеми, да се разрешават външни одити

Контролен списък за оценка на Вашите методи за обработване

Принцип “под контрола на потребителя”: да се изключва подход на обвързване и да се гарантира, че потребителите са информирани участници, които имат контрол върху своите избори.



Хартата предоставя рамка на принципите, които могат да ръководят създаващите политика лица, законодателите и професионалистите в областта на правосъдието, когато те се борят да не изостанат от бързото развитие на Изкуствения интелект в националните правосъдни процеси.

Гледната точка на СЕРЕЈ е, че приложението на Изкуствения интелект в областта на правосъдието може да допринесе за подобряването на ефективността и качеството. Той трябва да се прилага по един отговорен начин, който е в съответствие с основните права, гарантирани по-специално в Европейската конвенция за правата на човека (ЕКПЧ) и Конвенцията на Съвета на Европа за защита на личните данни. От съществено значение е да се гарантира, че Изкуственият интелект остава инструмент в служба на общия интерес и използването му да зачита правата на лицата.

Хартата определя пет основни принципа, които трябва да бъдат зачитани в областта на Изкуствения интелект и правораздаването: зачитане на основните права; недискриминация; качество и сигурност; прозрачност, непристрастност и справедливост; „под контрола на потребителя“. Хартата е придружена от задълбочено изследване на използването на Изкуствен интелект в съдебните системи.



www.coe.int/cepej

PREMS 005419

ENG

www.coe.int

Съветът на Европа е водещата организация в Европа в областта на човешките права. Тя включва 47 държави членки, включително всички държави членки на Европейския съюз. Всички държави членки на Съвета на Европа са подписали Европейската конвенция за правата на човека, споразумение, предназначено да защитава правата на човека, демокрацията и върховенството на закона. Европейският съд по правата на човека осъществява надзор върху прилагането на Конвенцията в държавите членки.

