Организация Объединенных Наций A/78/310

Генеральная Ассамблея

Distr.: General
30 August 2023
Russian
Original: Spanish

Семьдесят восьмая сессия
Пункт 73 b) предварительной повестки дня*
Поощрение и защита прав человека: вопросы прав человека, включая альтернативные подходы в деле содействия эффективному осуществлению прав человека и основных свобод

Право на неприкосновенность частной жизни

Записка Генерального секретаря

Генеральный секретарь имеет честь препроводить Генеральной Ассамблее доклад, подготовленный Специальным докладчиком по вопросу о праве на неприкосновенность частной жизни Аной Браян Нугререс и представленный в соответствии с резолюцией 28/16 Совета по правам человека.

* A/78/150.

23-15851 (R) 140923 210923

Доклад Специального докладчика по вопросу о праве на неприкосновенность частной жизни Аны Браян Нугререс

Принципы прозрачности и объяснимости при обработке персональных данных с помощью искусственного интеллекта

Резюме

В настоящем докладе Специальный докладчик по вопросу о праве на неприкосновенность частной жизни Ана Браян Нугререс подчеркивает важность принципов прозрачности и объяснимости при обработке персональных данных с помощью искусственного интеллекта. Повсеместное присутствие искусственного интеллекта во всех видах деятельности и принятие решений, затрагивающих людей, с помощью него требуют изучить вопрос об искусственном интеллекте и принять меры для обеспечения его этичного и ответственного использования, основанного на уважении прав человека.

Это важно, поскольку прозрачность и объяснимость способствуют не только повышению степени доверия к искусственному интеллекту и его надежности, но и защите прав человека. Эти принципы позволяют обеспечить, с одной стороны, чтобы люди своевременно получали полную, доступную для понимания и четкую информацию об основных аспектах, касающихся использования их персональных данных в процессах или проектах, связанных с применением искусственного интеллекта, и его последствий, и, с другой стороны, чтобы лица, подвергающиеся воздействию искусственного интеллекта, осознавали конкретные причины такого воздействия. Это предоставляет им возможность осуществлять свои права, такие как право на надлежащую правовую процедуру и право на защиту в случае принятия решений с использованием инструментов или технологий искусственного интеллекта.

I. Введение

1. Созданная Европейской комиссией Группа экспертов высокого уровня по искусственному интеллекту[1] отметила, что принципы прозрачности и объяснимости играют важную роль в распространении надежного искусственного интеллекта. В этой связи искусственный интеллект должен быть законным, этичным и надежным, «как с технической, так и с социальной точек зрения, поскольку даже при благих намерениях системы искусственного интеллекта могут причинить непреднамеренный вред[2]».

2. Аналогичным образом Организация Объединенных Наций по вопросам образования, науки и культуры (ЮНЕСКО) отметила, что «прозрачность и объяснимость тесно связаны с надлежащими мерами обеспечения ответственности и подотчетности, а также с надежностью систем [искусственного интеллекта][3]». Кроме того, она заявила, что «прозрачность и объяснимость систем искусственного интеллекта часто являются важнейшими условиями обеспечения уважения, защиты и поощрения прав человека, основных свобод и этических принципов[4]».

3. Искусственный интеллект занимает важное место в глобальной повестке дня. Так, ближе к концу декабря 2022 года Организация экономического сотрудничества и развития (ОЭСР) опубликовала декларацию о надежном, устойчивом и инклюзивном цифровом будущем[5], в котором обязалась работать, в частности, над содействием ориентированной на человека цифровой трансформации, предусматривающей поощрение прав человека как в реальной жизни, так и в Интернете, надежную защиту персональных данных, принятие законов и норм, соответствующих цифровой эпохе, а также надежное, безопасное, ответственное и устойчивое использование новых цифровых технологий и искусственного интеллекта.[6] Что касается искусственного интеллекта, то государства — члены ОЭСР призвали организацию способствовать разработке перспективной, последовательной и жизнеспособной нормативно-правовой основы для регулирования искусственного интеллекта и эффективного управления связанными с ним рисками, а также обеспечить наличие фактических данных, прогнозов и инструментов и отслеживание соответствующих инцидентов для эффективного планирования и реализации политики по внедрению надежного искусственного интеллекта[7].

4. 23 января 2023 года Европейский парламент, Совет Европы и Европейская комиссия приняли Европейскую декларацию о цифровых правах и принципах на цифровое десятилетие, в которой обязались:

---

[1] Группа независимых экспертов, сформированная Европейской комиссией в июне 2018 года.

[2] High-Level Expert Group on Artificial Intelligence, *Ethical guidelines for trustworthy artificial intelligence* (2019), p. 2. См. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[3] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2021, p. 22. См. https://unesdoc.unesco.org/ark:/48223/pf0000381137.

[4] Ibid.

[5] OCDE, *Declaration on a Trusted, Sustainable and Inclusive Digital Future* (2022). Декларация была принята по итогам совещания, проходившего на острове Гран-Канария, Испания, 14-15 декабря 2022 года. См. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0488.

[6] Ibid.

[7] Ibid.

a) способствовать созданию ориентированных на человека, надежных и этичных систем искусственного интеллекта на всех этапах их разработки, внедрения и использования в соответствии с ценностями [Европейского союза];

b) обеспечивать должный уровень прозрачности в отношении использования алгоритмов и искусственного интеллекта, а также предоставление людям возможности использовать их и их информирование при взаимодействии с ними;

c) обеспечивать, чтобы алгоритмические системы основывались на надлежащих массивах данных, во избежание дискриминации и в целях создания условий для человеческого контроля за всеми результатами, влияющими на безопасность и основные права людей;

d) гарантировать, чтобы такие технологии, как искусственный интеллект, не использовались для предопределения решений людей, например в сферах здравоохранения, образования, занятости и личной жизни;

e) предоставлять гарантии и принимать надлежащие меры, в том числе путем содействия установлению надежных норм, с тем чтобы искусственный интеллект и цифровые системы всегда были безопасными и использовались при всестороннем уважении основных прав человека;

f) принимать меры по обеспечению соблюдения в исследованиях, связанных с искусственным интеллектом, высочайших этических стандартов и соответствующего законодательства [Европейского союза][8].

5. В свете вышесказанного ниже приводятся некоторые соображения относительно искусственного интеллекта, в которых затрагиваются обозначенные ниже аспекты в целях разъяснения того, что понимается под принципами прозрачности и объяснимости в контексте обработки персональных данных в процессах или проектах, связанных с использованием искусственного интеллекта.

II. Искусственный интеллект и обработка персональных данных

6. Искусственный интеллект присутствует уже практически во всех сферах жизни общества — от мобильных устройств, которыми постоянно пользуются граждане, до сложнейших систем делового администрирования. Все большее распространение искусственного интеллекта открывает широкие возможности в различных сферах деятельности и секторах. Однако вместе с этими возможностями появляются проблемы и угрозы, к которым необходимо подходить ответственно, чтобы, помимо прочего, полностью реализовать потенциал искусственного интеллекта безопасно, этично и на основе уважения прав человека.

7. Единого мнения по поводу определения искусственного интеллекта нет, но существует несколько подходов к тому, что под ними понимается. В одном из справочных текстов по данной тематике была предложена следующая таксономия[9]:

[8] European Parliament, Council of Europe and European Commission, “European Declaration on Digital Rights and Principles for the Digital Decade”, *Official Journal of the European Union*, 2023/C 23/01, 23 January 2023. См. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOC_2023_023_R_0001.

[9] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Essex, England, Pearson, 2009).

- системы, мыслящие подобно человеку (например, когнитивные архитектуры и нейронные сети);
- системы, действующие подобно человеку (например, автоматизированное рассуждение и обучение).
- системы, мыслящие рационально (например, умозаключения);
- системы, действующие рационально (например, интеллектуальные программные агенты и встроенные роботы, которые достигают целей посредством восприятия, планирования, рассуждения, обучения, коммуникации, принятия решений и действия).

8. Все эти системы для получения результатов обрабатывают информацию, которая содержит, помимо прочего, персональные данные. В этой связи Европейская комиссия уточнила:

> «Для целей настоящей “белой книги”, а также для любых возможных будущих обсуждений по политическим инициативам представляется важным пояснить, какие основные элементы образуют искусственный интеллект — “данные” и “алгоритмы”. Искусственный интеллект может быть интегрирован в аппаратное обеспечение. Что касается методов машинного обучения, представляющих собой подсистему искусственного интеллекта, то алгоритмы обучаются выявлять определенные закономерности на основе набора данных, чтобы устанавливать действия, необходимые для достижения поставленной цели[10]».

9. Иными словами, для разработки искусственного интеллекта собираются, хранятся, анализируются и обрабатываются внушительные объемы информации, используемой для получения различных результатов, действий или моделей поведения машин или пользователей таких машин. Однако, как отмечается в упомянутой выше рекомендации ЮНЕСКО, «неприкосновенность частной жизни, являющаяся правом, необходимым для защиты достоинства, автономии и субъектности человека, должна уважаться, защищаться и поощряться на протяжении всего жизненного цикла систем [искусственного интеллекта][11]».

10. Надлежащая обработка персональных данных имеет принципиальное значение для обеспечения того, чтобы — сообразно обстоятельствам — с развитием искусственного интеллекта правам человека не причинялся вред и они не ставились под угрозу. Существует несколько инициатив и организаций, деятельность в рамках которых направлена на обеспечение разработки искусственного интеллекта на основе уважения прав человека, и некоторые примеры таких инициатив и организаций приводятся ниже.

11. Во-первых, в октябре 2020 года Глобальная ассамблея по вопросам неприкосновенности частной жизни приняла резолюцию о подотчетности при разработке и использовании искусственного интеллекта[12]. В этой резолюции, помимо прочего, организациям, разрабатывающим или использующим системы искусственного интеллекта, настоятельно рекомендуется рассмотреть возможность принятия следующих мер:

---

[10] European Commission, *White Paper on Artificial Intelligence - a European approach to excellence and trust*, COM (2020) 65 final. См. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1603192201335&uri=CELEX%3A52020DC0065.

[11] См. https://unesdoc.unesco.org/ark:/48223/pf0000381137 с. 21.

[12] См. https://globalprivacyassembly.org/wp-content/uploads/2020/11/GPA-Resolution-on-Accountability-in-the-Development-and-Use-of-AI-EN.pdf, с. 3.

- оценивать потенциальное воздействие на права человека (включая право на защиту данных и право на неприкосновенность частной жизни) до разработки и/или использования искусственного интеллекта;
- проверять устойчивость, надежность, точность и защищенность данных искусственного интеллекта до его использования, включая выявление и устранение субъективности в системах и используемых ими данных, которая может привести к получению несправедливых результатов;
- принимать меры по обеспечению подотчетности, соответствующие рискам нарушения прав человека.

12. В этой связи ЮНЕСКО в своей рекомендации отметила:

«Алгоритмические системы требуют надлежащей оценки воздействия на неприкосновенность частной жизни, включающей также социальные и этические аспекты их использования и инновационное применение концепции проектируемой конфиденциальности. Участники [деятельности, связанной с искусственным интеллектом] должны взять на себя ответственность за проектирование и применение систем [искусственного интеллекта] таким образом, чтобы гарантировать защиту персональных данных на протяжении всего жизненного цикла системы [искусственного интеллекта][13]».

13. До этого, в июне 2019 года, Иберо-американская сеть защиты данных опубликовала документ под названием “Recomendaciones generales para el tratamiento de datos personales en la inteligencia artificial” («Общие рекомендации по обработке персональных данных с помощью искусственного интеллекта»)[14]. В этом документе выдвигается ряд предложений в качестве руководства для разработчиков продуктов искусственного интеллекта, с тем чтобы уже на этапе проектирования этих продуктов учитывались нормативные положения по обработке персональных данных. Были вынесены следующие рекомендации:

- соблюдать местные нормативные положения по обработке персональных данных;
- проводить оценки воздействия на неприкосновенность частной жизни;
- внедрить принципы конфиденциальности, этики и защищенности по замыслу и по умолчанию;
- установить принцип подотчетности;
- подготовить соответствующие схемы управления обработкой персональных данных в организациях, разрабатывающих продукты искусственного интеллекта;
- принять меры по обеспечению соблюдения принципов обработки персональных данных в проектах, связанных с использованием искусственного интеллекта;
- уважать права владельцев данных и внедрять эффективные механизмы осуществления этих прав;

[13] См. https://unesdoc.unesco.org/ark:/48223/pf0000381137, сс. 21–22.

[14] Red Iberoamericana de Protección de Datos, “Recomendaciones generales para el tratamiento de datos personales en la inteligencia artificial” (2019). Текст утвержден структурами, являющимися членами Сети, на заседании, состоявшемся 21 июня 2019 года в Наукалпан-де-Хуарес, Мексика. См. https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf.

- обеспечивать качество персональных данных;
- использовать инструменты анонимизации;
- повышать степень доверия и прозрачности в отношениях с владельцами персональных данных.

14. Для предоставления более подробной информации о выполнении некоторых из этих рекомендаций Иберо-американская сеть защиты данных подготовила дополнительные и более детальные указания, содержащиеся в документе под названием “Orientaciones específicas para el cumplimiento de los principios y derechos que rigen la protección de los datos personales en los proyectos de inteligencia artificial” («Конкретные указания относительно соблюдения принципов и прав, регулирующих защиту персональных данных в проектах, связанных с использованием искусственного интеллекта»)[15]. Особенно подробно в настоящем докладе рассматривается принцип прозрачности, о котором будет говориться позднее.

III. Риски, присущие искусственному интеллекту

15. Общество и его цифровая трансформация подвергаются влиянию искусственного интеллекта, который присутствует в различных сферах повседневной жизни, в экономике, науке, образовании, здравоохранении и многих других отраслях и сферах деятельности.

16. Хотя возможности и преимущества искусственного интеллекта в обществе в целом неоспоримы, не стоит забывать о том, что он также может скрывать в себе проблемы, риски и угрозы. Они могут включать в себя, в частности, неэтичную разработку или использование искусственного интеллекта, а также принятие субъективных, непрозрачных или неверных решений в отношении людей.

17. Степень риска зависит от каждой конкретной ситуации.

> По мнению Европейской комиссии, как правило, должно считаться, что то или иное применение [искусственного интеллекта] сопряжено с повышенным риском, в зависимости от того, что поставлено на карту, и с учетом того, что как сектор, так и предполагаемое использование связаны со значительными рисками, в частности с точки зрения защиты безопасности, прав потребителей и основных прав. Если говорить более конкретно, то должно считаться, что применение [искусственного интеллекта] сопряжено с повышенным риском, если оно в совокупности отвечает следующим двум критериям:
>
> a) во-первых, применение [искусственного интеллекта] происходит в секторе, в котором ввиду его специфики или характера обычно осуществляемой в нем деятельности можно ожидать возникновения значительных рисков. [...]. Например, в данном случае имеются в виду здравоохранение, транспорт, энергетика и некоторые элементы государственного сектора [...];
>
> b) во-вторых, [искусственный интеллект] в рассматриваемом секторе также применяется таким образом, что существует вероятность

[15] Red Iberoamericana de Protección de Datos, “Orientaciones específicas para el cumplimiento de los principios y derechos que rigen la protección de los datos personales en los proyectos de inteligencia artificial” (2019). См. https://www.redipd.org/sites/default/files/2020-02/guia-orientaciones-espec%C3%ADficas-proteccion-datos-ia.pdf.

возникновения значительных рисков. [...]. Оценка степени риска того или иного вида использования может быть основана на воздействии на затрагиваемые стороны. Например, имеется в виду применение [искусственного интеллекта], приводящее к юридическим или столь же значимым последствиям для прав физического или юридического лица, представляющее опасность травмирования, смерти или нанесения значительного материального или нематериального ущерба, а также влекущее за собой последствия, которых физические или юридические лица объективно не могут избежать[16].

18. Искусственный интеллект сопряжен с различными рисками. К непредвиденным обстоятельствам, которые необходимо учитывать, относятся риски, присущие работе с алгоритмами (человеческая субъективность, технические недочеты, факторы уязвимости в плане безопасности и сбои в функционировании), а также риски, связанные с их разработкой. В этой связи были определены аспекты, оказывающие влияние на управление связанными с алгоритмами рисками; эти аспекты отражены на приведенном ниже рисунке[17].

19. В этой связи дается следующее пояснение:

«На исходные данные влияют в основном две переменные: субъективность (включение неполных, недостаточных, устаревших или подвергшихся манипулированию данных) и уместность (релевантность, противоречивость или полнота данных); при этом на разработку алгоритмов могут влиять шаблоны (субъективность логики программирования, включение непредвиденных функций и присущие функциям, используемым для их кодирования, сбои) и ошибки (условия эксплуатации, отражающие порядок работы, отличающийся от запланированного и идущий вразрез с тем, что было предусмотрено при проектировании). Наконец, риски при принятии

[16] См. https://eur-lex.europa.eu/legal-content/ES/TXT/?qid=1603192201335&uri=CELEX%3A52020DC0065.

[17] См. https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf , с. 18.

итоговых решений связаны с уместностью и точностью исполнения алгоритма как реакции на анализ исходных данных[18]».

IV. Принцип прозрачности при обработке персональных данных

20. Понятие прозрачности используется в различных дисциплинах, включая информатику, доступ к информации, право и обработку персональных данных. По мнению ЮНЕСКО, «прозрачность направлена на предоставление надлежащей информации соответствующим адресатам для обеспечения их понимания и укрепления доверия[19]».

21. Не существует единого мнения о том, что понимается под прозрачностью в каждом конкретном случае, и в зависимости от ситуации этот термин может приобретать разные оттенки смысла. Например, принцип прозрачности означает одно, когда используется в отношении обработки персональных данных в целом, и другое, когда применяется в контексте искусственного интеллекта. В настоящем докладе говорится о прозрачности при обработке персональных данных в целом и при их обработке с помощью искусственного интеллекта в частности.

22. Принцип прозрачности рассматривается во многих документах организаций из различных уголков мира[20]. Как было ранее отмечено Специальным докладчиком в этой связи, в соответствии с принципом прозрачности все операторы данных должны информировать владельцев данных об условиях обработки их личной информации на протяжении всего процесса ее обработки, то есть с момента ее сбора, с тем чтобы владельцы данных располагали необходимыми сведениями для осуществления надлежащего контроля[21].

---

[18] Alejandro Useche and Jeimy Cano, *Robo-Advisors: Asesoría automatizada en el mercado de valores*, Universidad del Rosario y Autorregulador del Mercado de Valores de Colombia (2019), págs. 9 y 10. См. https://www.researchgate.net/publication/331358231_Robo-Advisors_Asesoria_automatizada_en_el_mercado_de_valores.

[19] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2021, p. 22.

[20] Organisation for Economic Cooperation and Development (OECD), *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (принято 23 сентября 1980 года и пересмотрено в июле 2013 года); Совет Европы, Конвенция № 108 о защите физических лиц при автоматизированной обработке персональных данных, 28 января 1981 года; Организация Объединенных Наций, Руководящие принципы регламентации компьютеризированных картотек, содержащих данные личного характера, 14 декабря 1990 года; Совет Европы, Дополнительный протокол к Конвенции о защите физических лиц при автоматизированной обработке персональных данных, касающийся надзорных органов и трансграничных потоков данных, 8 ноября 2001 года; Asia-Pacific Economic Cooperation Forum, *Asia-Pacific Economic Cooperation Forum Privacy Framework*, 2004; Agencia Española de Protección de Datos, *Propuesta Conjunta para la Redacción de Estándares Internacionales para la Protección de la Privacidad en relación con el Tratamiento de Datos de Carácter Personal*, Madrid, 5 de noviembre de 2009; Regulation (European Union) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 27 April 2016; Red Iberoamericana de Protección de Datos, Directrices para la Armonización de la Protección de Datos en la Comunidad Iberoamericana, de 2017; Совет Европы, Протокол о внесении изменений в Конвенцию о защите физических лиц при автоматизированной обработке персональных данных, октябрь 2018 года; и Organización de los Estados Americanos, Comité Jurídico Interamericano, *Principios Actualizados sobre la Privacidad y la Protección de Datos Personales*, de 2021.

[21] A/77/196, п. 45.

23. В вышеупомянутом докладе проводится анализ принципа прозрачности на основе следующих международных документов, касающихся неприкосновенности частной жизни и обработки персональных данных: a) Общего регламента Европейского союза по защите данных; b) Конвенции о защите физических лиц при автоматизированной обработке персональных данных; c) Стандартов защиты персональных данных для иберо-американских государств, принятых Иберо-американской сетью защиты данных; d) Рекомендаций Совета, касающихся Руководства по защите неприкосновенности частной жизни и трансграничной передаче персональных данных Организации экономического сотрудничества и развития; e) Рамочной основы неприкосновенности частной жизни форума «Азиатско-Тихоокеанское экономическое сотрудничество»; и f) Обновленных принципов Организации американских государств по неприкосновенности частной жизни и защите персональных данных, с аннотациями.

24. По итогам проведенного анализа был сделан вывод о том, что, как правило, необходимо предоставлять следующую информацию:

- личность и адрес оператора или его представителя, а также цели или задачи обработки — данные, составляющие основу прозрачности;
- права владельца данных и способы их осуществления, а также получатели данных или категория получателей;
- правовое основание для обработки, а также факт осуществления обработки и/или ее основные характеристики;
- категория обрабатываемых данных и их происхождение, если они не были получены непосредственно от владельца данных.

25. Следует подчеркнуть, что для соблюдения принципа прозрачности информация, предоставляемая владельцу данных, должна быть изложена ясным, простым, доходчивым, доступным и понятным языком. Это должно обеспечиваться и в случаях, связанных с детьми и подростками, с внесением необходимых корректировок.

26. Не все упомянутые выше нормативные документы требуют предоставления одной и той же информации, поскольку некоторые из них содержат более обширные перечни данных, подлежащих раскрытию. В частности, в соответствии с Общим регламентом Европейского союза по защите данных должна быть предоставлена, помимо прочего, следующая информация[22]: контактная информация лица, ответственного за защиту данных; срок хранения персональных данных или критерии его определения; намерение оператора сообщать или передавать данные, а также нормативный акт, которым это может быть предусмотрено; право на подачу жалобы в надзорный орган; является ли сообщение данных юридическим или договорным требованием или необходимо для заключения договора, а также обязан ли субъект данных предоставить свои персональные данные и каковы последствия отказа от их предоставления; наличие автоматизированных процессов принятия решений, включая профилирование, когда должна предоставляться содержательная информация о применяемой логике, значении и предполагаемых последствиях такой обработки данных, и информация о назначении, если планируется дальнейшая обработка с целью, отличной от той, для которой были собраны данные.

[22] См. https://eur-lex.europa.eu/legal-content/ES/TXT/?qid=1532348683434&uri=CELEX%3A02016R0679-20160504.

V. Принцип прозрачности при обработке персональных данных с помощью искусственного интеллекта

27. Крайне важно обеспечить прозрачность при использовании искусственного интеллекта, так как неосведомленность о его применении или опущение этой информации могут привести к негативным последствиям. В этой связи Европейская комиссия отметила:

> «Непрозрачность (неявность [искусственного интеллекта]) затрудняет выявление и доказывание возможных нарушений законодательства, в том числе правовых норм, обеспечивающих защиту основных прав, установление ответственности и возможность требования компенсации[23]».

28. Степень потенциальной непрозрачности искусственного интеллекта может быть снижена, если будет предусмотрена необходимость соблюдения минимальных стандартов прозрачности. В этой связи были установлены следующие требования:

> «Обеспечивать предоставление четкой информации о возможностях и ограничениях систем [искусственного интеллекта], в частности о предназначении этих систем, условиях, при которых предполагается их ожидаемое функционирование, и ожидаемом уровне точности достижения указанной цели [...]. Отдельно граждане должны четко информироваться о том, что они имеют дело с системой [искусственного интеллекта], а не с человеком [...]. Кроме того, важно, чтобы предоставляемая информация была объективной, краткой и понятной[24]».

29. В рекомендации ЮНЕСКО говорится:

> «В случае систем [искусственного интеллекта] обеспечение прозрачности может позволить людям понять, что происходит на каждом этапе работы системы [искусственного интеллекта], сообразно обстоятельствам и в зависимости от восприимчивости соответствующей системы. Также может быть предоставлена информация о факторах, влияющих на конкретный прогноз или решение, а также о наличии или отсутствии соответствующих гарантий (например, мер обеспечения безопасности или справедливости)[25]».

30. Группа экспертов высокого уровня по искусственному интеллекту отметила, что для достижения надежности искусственного интеллекта необходимо соблюдение определенных требований, в том числе требования об обеспечении прозрачности, в связи с чем следует:

> «четко и в упреждающем порядке доводить до сведения заинтересованных сторон информацию о возможностях и ограничениях систем [искусственного интеллекта], позволяющую формировать реалистичные ожидания, а также о порядке выполнения требований, и прозрачно обозначать, что они имеют дело с системой [искусственного интеллекта][26]».

В указанной рекомендации ЮНЕСКО также говорится о том, что «участники [деятельности, связанной с искусственным интеллектом] должны надлежащим и своевременным образом информировать пользователей, когда продукт или

---

[23] European Commission, *White Paper on Artificial Intelligence – a European approach to excellence and trust*, 2020, p. 14.

[24] Ibid., pp. 23–24.

[25] См. https://unesdoc.unesco.org/ark:/48223/pf0000381137, с. 22.

[26] См. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, с. 2 и 3.

услуга предоставляются напрямую или с помощью систем [искусственного интеллекта][27]».

31. Кроме того, в рекомендации ЮНЕСКО отмечается:

> «С прозрачностью тесно связана объяснимость, поскольку результаты и подпроцессы, приводящие к результатам, должны — сообразно обстоятельствам — быть понятными и прослеживаемыми. Участники [деятельности, связанной с искусственным интеллектом] должны взять на себя обязательства по обеспечению объяснимости разрабатываемых алгоритмов. В случае применения [искусственного интеллекта], влияние которого на конечного пользователя не носит временного характера, не является легко обратимым и сопряжено со значительными рисками, следует обеспечить предоставление обстоятельного разъяснения любого решения, которое привело к предпринятому действию, чтобы результат считался прозрачным[28]».

32. Группа экспертов высокого уровня по искусственному интеллекту пояснила, что прозрачность «тесно связана с принципом объяснимости и подразумевает прозрачность элементов, имеющих отношение к системе [искусственного интеллекта]: данных, системы и принципов работы[29]». Она также подчеркнула важность прослеживаемости, объяснимости и информирования следующим образом:

- «прослеживаемость: наборы данных и процессы, на основе которых принимается решение в системе [искусственного интеллекта], включая сбор и маркировку данных, а также используемые алгоритмы, должны быть задокументированы в соответствии с как можно более строгими стандартами, чтобы обеспечить прослеживаемость и повышение прозрачности. Это относится и к решениям, принимаемым системой [искусственного интеллекта]. Это позволяет выявить причины, по которым принятое с помощью системы решение оказалось ошибочным, что, в свою очередь, может помочь предотвратить ошибки в будущем. Следовательно, прослеживаемость способствует как контролируемости, так и объяснимости;

- объяснимость: под объяснимостью понимается способность объяснить как технические процессы системы [искусственного интеллекта], так и соответствующие решения человека (например, области применения системы [искусственного интеллекта]). Техническая объяснимость предполагает обеспечение того, чтобы решения, принимаемые системой [искусственного интеллекта], были понятными для человека и могли прослеживаться им. Кроме того, может быть необходимо искать компромисс между повышением объяснимости системы (что может снизить ее точность) и повышением точности (в ущерб объяснимости). Когда система [искусственного интеллекта] оказывает существенное влияние на жизнь людей, должна быть возможность потребовать соответствующего разъяснения процесса принятия решений системой [искусственного интеллекта]. Такое разъяснение должно быть своевременным и отражать степень осведомленности заинтересованных сторон (к которым могут относиться люди, не являющиеся специалистами в этой области, осуществляющие надзор или занимающиеся исследовательской работой). Кроме того, должна быть обеспечена возможность получения разъяснений относительно степени воздействия и влияния системы [искусственного интеллекта] на процесс принятия

[27] См. https://unesdoc.unesco.org/ark:/48223/pf0000381137, с. 22.
[28] Ibid., с. 22.
[29] См. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, с. 18.

решений в организации, проектирования системы и обоснования ее внедрения (чем обеспечивается прозрачность принципов работы);

- информирование: системы [искусственного интеллекта] не должны казаться пользователям людьми; люди имеют право знать, что они имеют дело с системой [искусственного интеллекта]. Это означает, что системы [искусственного интеллекта] должны опознаваться как таковые. Кроме того, при необходимости для обеспечения соблюдения основных прав должна быть предусмотрена возможность пользователей отказаться от взаимодействия с системой [искусственного интеллекта] в пользу взаимодействия с человеком. Помимо этого, возможности и ограничения системы [искусственного интеллекта] должны быть доведены до сведения соответствующих специалистов или конечных пользователей в форме, соответствующей конкретным условиям использования, с обязательным включением информации о степени точности системы [искусственного интеллекта], а также об ее ограничениях[30]».

33. Европейский совет по защите данных и Европейский инспектор по защите данных опубликовали совместное заключение по этому вопросу, в котором заявили:

> «Субъекты данных всегда должны уведомляться, когда их данные используются для обучения или прогнозирования [в связи с искусственным интеллектом], о правовом основании для такого использования их данных, об общем объяснении логики (процедуры) и об области применения системы [искусственного интеллекта]. В связи с этим в таких случаях всегда должно быть гарантировано право физических лиц на ограничение обработки данных (статья 18 [Общего регламента Европейского союза № 2016/679 по защите данных] и статья 20 Регламента, касающегося защиты физических лиц при обработке персональных данных, осуществляемой учреждениями, органами, службами и агентствами Союза, и о свободном обращении таких данных), а также на уничтожение или удаление данных (статья 16 [Общего регламента Европейского союза № 2016/679 по защите данных] и статья 19 Регламента, касающегося защиты физических лиц при обработке персональных данных, осуществляемой учреждениями, органами, службами и агентствами Союза, и о свободном обращении таких данных). Кроме того, оператор данных должен иметь явное обязательство информировать субъекта данных об установленных сроках для возражения, ограничения, удаления данных и т. д. Система [искусственного интеллекта] должна быть способна выполнять все требования по защите данных с помощью надлежащих технических и организационных мер. Право на объяснение должно обеспечивать дополнительную прозрачность[31]».

34. В вышеупомянутом докладе[32] подчеркивается, что в случаях, когда субъекты данных подвергаются автоматизированному принятию решений или профилированию, они должны понимать, каким образом будет обрабатываться касающаяся их информация (в частности, будет ли использован искусственный

[30] Ibid.

[31] European Data Protection Board and the European Data Protection Supervisor, Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, p. 17. См. https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf .

[32] A/77/196, п. 55.

интеллект), и получать содержательную информацию о применяемой логике и значении и предполагаемых последствиях такой обработки.

35. В этой связи Испанское агентство по защите данных отметило, что «под “содержательной информацией” [...] следует понимать информацию, которая, будучи предоставленной субъекту данных, позволяет ему понять, какой обработке подвергаются его данные, и обеспечивает определенность и доверие в отношении соответствующих результатов[33]».

36. Агентство также подчеркнуло:

> «Выполнение этого обязательства путем технического указания на исполнение алгоритма может порождать неясность, вводить в заблуждение или приводить к информационному перенасыщению. Необходимо предоставлять информацию, позволяющую понять процесс обработки. Хотя это будет зависеть от типа используемого компонента [системы искусственного интеллекта], под информацией, которая может иметь отношение к субъекту данных, подразумевается следующее:
>
> • подробные сведения об используемых для принятия решений данных, не ограничивающиеся указанием категории, особенно информация о продолжительности использования данных (давность данных);
>
> • относительная значимость, придаваемая всем данным при принятии решения;
>
> • качество обучающих данных и тип используемых шаблонов;
>
> • проведенные мероприятия по профилированию и их последствия;
>
> • значения точности или ошибки согласно соответствующим количественным показателям, используемым для измерения достоверности сделанного вывода;
>
> • наличие или отсутствие квалифицированного человеческого контроля;
>
> • указания на проверки, особенно на проверки возможных отклонений результатов вывода, а также на сертификацию системы [искусственного интеллекта]. В случае адаптивных или эволюционных систем — последняя проведенная проверка;
>
> • если система [искусственного интеллекта] содержит информацию, относящуюся к идентифицируемым третьим лицам, — запрет на несанкционированную обработку такой информации и ее последствия[34]».

37. Европейский инспектор по защите данных опубликовал заключение, согласно которому, если Комиссия предложит новую нормативную базу непосредственно по вопросу об искусственном интеллекте, то ко всем приложениям, работающим на базе искусственного интеллекта, независимо от степени риска, должен применяться определенный набор обоснованных гарантий, таких как принятие технических и организационных мер (включая документацию); полная прозрачность в отношении целей, использования и проектирования внедряемых алгоритмических систем; и обеспечение надежности системы искусственного интеллекта, внедрение имеющихся механизмов обеспечения

[33] Spanish Data Protection Agency, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, February 2020. pág. 24. См. https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf.

[34] Ibid.

подотчетности, возмещения ущерба и независимого надзора и обеспечение прозрачности в их отношении[35].

38. Европейский совет по защите данных и Европейский инспектор по защите данных также особо отметили необходимость распространять следующее:

> «новые, в большей степени упредительные и актуальные методы информирования пользователей систем [искусственного интеллекта] об этапе процесса (принятия решений), на котором находится система в тот или иной момент времени, — методы, обеспечивающие раннее предупреждение о потенциально опасных результатах, с тем чтобы лица, чьи права и свободы могут быть ущемлены автономными решениями машин, могли отреагировать или исправить принятое решение[36]».

39. По мнению Иберо-американской сети защиты данных, для соблюдения принципа прозрачности необходимо следующее[37]:

- «доводить до сведения владельцев данных основные характеристики обработки, которой будут подвергнуты их персональные данные;
- явно информировать владельцев данных о том, что при обработке их персональных данных будут использоваться процессы автоматизации;
- включать в метод, выбранный операторами для соблюдения принципа прозрачности, все цели обработки данных;
- обозначать происхождение персональных данных, когда данные были получены путем передачи, а в случаях, когда предполагается использование искусственного интеллекта, — подтверждать, что об этой цели было сообщено первым оператором, который получил данные для их использования с этой целью;
- разрабатывать инновационные способы информирования владельцев данных об основных характеристиках обработки и степени риска с точки зрения повышения или снижения ожиданий в отношении конфиденциальности;
- охранять право на информационное самоопределение, обеспечивая, чтобы владельцы данных всегда должным образом и своевременно информировались, что они будут непосредственно взаимодействовать с системой искусственного интеллекта или когда их информация будет обрабатываться с помощью искусственного интеллекта;
- предоставлять содержательную информацию о цели и последствиях использования систем искусственного интеллекта для проверки постоянного соответствия ожиданиям владельцев данных в отношении конфиденциальности, обеспечивая им возможность в любое время осуществлять контроль за обработкой своих персональных данных;

---

[35] European Data Protection Supervisor, *Opinion 4/2020, European Data Protection Supervisor Opinion on the European Commission's White Paper on Artificial Intelligence – a European approach to excellence and trust*, 29 June 2020, p. 14. См. https://edps.europa.eu/sites/edp/files/publication/20-06-19_opinion_ai_white_paper_en.pdf.

[36] European Data Protection Board and the European Data Protection Supervisor, "Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)", 18 June 2021, p. 22.

[37] См. https://www.redipd.org/sites/default/files/2020–02/guia-orientaciones-espec%C3%ADficas-proteccion-datos-ia.pdf, сс. 17–19.

- выявить часто используемые термины, дать им определения и создать базу данных, чтобы эти термины можно было повторно использовать в различных контекстах, со стандартными пиктограммами для доведения информации до сведения владельцев данных;
- постоянно информировать владельцев данных, чтобы они знали, как автоматизированное принятие решений может повлиять на них и как при необходимости запросить вмешательство человека, для принятия осознанного решения о том, соглашаться или нет на обработку их данных».

40. Иберо-американская сеть защиты данных уточнила:

> «Информация, предоставляемая в отношении логики модели [искусственного интеллекта], должна по меньшей мере включать сведения об основных аспектах ее работы, а также весе и корреляции данных, изложенные ясным, простым и понятным языком, и при этом нет необходимости приводить полное разъяснение используемых алгоритмов или даже включать их[38]».

41. Иберо-американская сеть защиты данных призвала лиц, ответственных за обработку данных с помощью искусственного интеллекта, к новаторству для передачи информации в простой и лаконичной форме. Она отметила, что «существует несколько инновационных подходов к предоставлению уведомлений относительно конфиденциальности, включая использование видеороликов, мультфильмов и стандартных пиктограмм. Чтобы облегчить понимание сложной информации об [искусственном интеллекте] для владельцев персональных данных, можно одновременно использовать несколько подходов[39]».

42. В следующих пунктах приводится примерный и неисчерпывающий перечень стран, которые явно или неявно отразили в своем законодательстве принцип прозрачности при обработке персональных данных с помощью искусственного интеллекта.

43. В Эквадоре в пунктах 14 и 17 статьи 12 Органического закона о защите данных, принятого в 2021 году, закреплено право на получение информации о существовании права не подпадать под действие решения, основанного исключительно на автоматизированных оценках, о способах реализации этого права и о существовании автоматизированных оценок и решений, включая профилирование.

44. Данный закон также предусматривает, что в случаях получения данных непосредственно от их владельцев информация должна быть предоставлена заранее — в момент сбора персональных данных. Кроме того, статья 12 этого закона гласит:

> «Если персональные данные не поступили непосредственно от владельцев данных или были получены из общедоступных источников, владельцы данных должны быть проинформированы в течение тридцати (30) дней или в первом полученном ими сообщении, в зависимости от того, что произойдет раньше. Владельцам данных должна предоставляться ясная, недвусмысленная, прозрачная, понятная, краткая и точная информация без технических препятствий».

[38] См. https://www.redipd.org/es/documentos/guia, сс. 17–19.
[39] Ibid.

45. В Перу право на объективную обработку персональных данных рассматривается в статье 72 Положения о применении Закона № 29733 о защите персональных данных, которая гласит:

> «Для обеспечения реализации права на объективную обработку в соответствии со статьей 23 Закона[40], когда персональные данные обрабатываются в рамках процесса принятия решений, в котором не участвует владелец персональных данных, лицо, ответственное за ведение базы персональных данных, или лицо, ответственное за их обработку, должно незамедлительно сообщить об этом субъекту данных, если положениями об осуществлении других прав, закрепленных в Законе и Положении о его применении, не предусмотрено иное».

46. В Сан-Томе и Принсипи Закон № 3/2016 от 2 мая 2016 года о защите персональных данных физических лиц уникален тем, что в его статье 21 предусмотрено, что операторы или их представители должны письменно и не более чем за восемь дней до начала обработки данных уведомить Национальное агентство по защите персональных данных о начале полностью или частично автоматизированной обработки или пакетной обработки для достижения одной или нескольких взаимосвязанных целей, за некоторыми исключениями. Кроме того, статья 11 Закона предусматривает, что субъекты данных при осуществлении своего права на доступ имеют право на получение от операторов информации об основании для автоматизированной обработки касающихся их данных.

47. В Уругвае статья 13 Закона № 18831 от 11 августа 2008 года о защите персональных данных гласит, что владельцы данных имеют право на получение — до сбора данных — информации в явной, четкой и недвусмысленной форме о критериях оценки, применяемых процессах и используемых технологиях или программах в случаях автоматизированной обработки данных для анализа определенных аспектов их личности, таких как трудовая деятельность, кредитоспособность, надежность и поведение, в целях принятия решений, имеющих юридические последствия, которые могут существенно повлиять на владельцев данных. Закон гласит также, что «когда персональные данные не собираются непосредственно у владельцев данных, информация [...] должна быть предоставлена им в течение пяти рабочих дней с даты получения запроса операторами».

VI. Принцип объяснимости при обработке персональных данных в проектах, связанных с использованием искусственного интеллекта

48. Все более распространенным становится создание «виртуальных профилей» людей на основе имеющейся информации. Кроме того, зачастую затрагивающие людей решения принимаются по итогам автоматизированной обработки их данных с использованием различных технологических инструментов.

49. На человека могут положительно или отрицательно влиять решения, принимаемые в отношении него на основе использования и обработки данных в

[40] «Статья 23. Право на объективную обработку данных. Владельцы персональных данных имеют право не подпадать под действие решения, которое имеет для них юридические последствия или оказывает на них существенное влияние и которое основывается исключительно на обработке персональных данных, предназначенных для оценки определенных аспектов их личности или поведения, если только это не происходит в рамках переговоров, заключения или исполнения договора или оценки для целей занятия должности в государственном учреждении, в соответствии с законом, без ущерба для возможности отстаивания своей точки зрения для защиты своих законных интересов».

проектах, связанных с применением искусственного интеллекта. Вызывает обеспокоенность вопрос о том, как защитить права людей, затрагиваемых решениями, которые принимаются в отношении них с помощью инструментов или технологий искусственного интеллекта. В частности, в «белой книге» по искусственному интеллекту отмечается: «Использование [искусственного интеллекта], как и любой другой новой технологии, несет в себе как возможности, так и риски. Граждане боятся оказаться бессильными в защите своих прав и обеспечении своей безопасности перед лицом информационной асимметрии алгоритмического принятия решений[41]».

50. В свете вышесказанного следует отметить, что люди должны знать, исходя из каких данных было принято то или иное затрагивающее их решение, а также какая для этого применялась логика. Наличие доступа к этой информации, в частности, позволит соответствующему лицу узнать, правильно ли решение, принятое в отношении него, и, если нет, защитить себя. Иными словами, такая информация необходима для обеспечения надлежащей правовой процедуры, поскольку она будет служить доказательством возможных неточностей или проявлений несправедливости в отношении того или иного лица при обработке его персональных данных с помощью искусственного интеллекта. В этой связи упомянутая выше Группа экспертов высокого уровня по искусственному интеллекту подчеркнула:

> «Принцип объяснимости имеет решающее значение для формирования и поддержания доверия пользователей к системам [искусственного интеллекта]. Это означает, что процессы должны быть прозрачными, что необходимо открыто сообщать о возможностях и назначении систем [искусственного интеллекта] и что решения, насколько это возможно, должны быть объяснимыми для тех, кого они прямо или косвенно касаются. Без такой информации решение не может быть должным образом оспорено [...]. Степень необходимости объяснения сильно зависит от конкретной ситуации и серьезности последствий в случае ошибочного или ненадлежащего результата[42]».

51. Все это объясняет, почему при использовании искусственного интеллекта важна прозрачность, — в связи с ним не должно быть случаев неясности, секретности или введения в заблуждение. Именно поэтому в упомянутой выше Европейской декларации было отмечено:

> «Каждый человек должен иметь возможность пользоваться преимуществами алгоритмических систем и систем искусственного интеллекта, прежде всего чтобы принимать собственные, осознанные решения в цифровой среде, будучи при этом защищенным от рисков и вреда для своего здоровья, безопасности и основных прав[43]».

52. В соответствии с вышеизложенным в 2019 году Иберо-американская сеть защиты данных рекомендовала повысить прозрачность при работе с владельцами персональных данных[44].

---

[41] См. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1603192201335&uri=CELEX%3A52020DC0065, с. 9.

[42] См. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, с. 13.

[43] См. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOC_2023_023_R_0001.

[44] См. https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf, сс. 23 и 24.

53. Также в соответствии с вышеизложенным впоследствии в своей упомянутой выше резолюции 2020 года Глобальная ассамблея по вопросам неприкосновенности частной жизни подчеркнула, что организации, разрабатывающие или использующие системы искусственного интеллекта, должны принимать во внимание следующие меры: a) обеспечение прозрачности и открытости путем раскрытия информации о применении искусственного интеллекта, используемых данных и логики, лежащей в основе работы искусственного интеллекта; b) обеспечение того, чтобы было обозначено ответственное лицо, которому можно было бы выразить обеспокоенность в связи с автоматизированными решениями и с помощью которого можно было реализовать соответствующие права, а также который мог бы инициировать оценку процесса принятия решений и вмешательство человека; c) предоставление по запросу разъяснений на ясном и понятном людям языке в отношении автоматизированных решений, принимаемых с помощью искусственного интеллекта; и d) обеспечение по запросу вмешательства человека в автоматизированный процесс принятия решений с помощью искусственного интеллекта[45].

54. Все вышесказанное частично согласуется с положениями Общего регламента по защите данных, в котором, в частности, говорится:

> «Если персональные данные не были получены от субъекта данных, оператор должен предоставить субъекту данных следующую информацию: [...] 2. […] g) информацию об осуществлении автоматизированного процесса принятия решений, включая профилирование, о котором говорится в пунктах 1 и 4 статьи 22, и — по крайней мере в этих случаях — содержательную информацию о применяемой логике, а также о значении и предполагаемых последствиях такой обработки для субъекта данных[46]».

Кроме того, субъект данных или владелец данных имеет право:

> «получить от оператора подтверждение того, обрабатываются ли касающиеся его персональные данные, и, если это так, доступ к персональным данным и следующей информации: [...] h) информации об осуществлении автоматизированного процесса принятия решений, включая профилирование, о котором говорится в пунктах 1 и 4 статьи 22, и — по крайней мере в этих случаях — содержательную информацию о применяемой логике, а также о значении и предполагаемых последствиях такой обработки для субъекта данных[47]».

55. Национальный институт стандартов и технологий резюмирует описание сферы действия этого принципа на приведенном ниже рисунке[48].

[45] См. https://globalprivacyassembly.org/document-archive/adopted-resolutions/, с. 3.

[46] См. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679, art. 14, para. 2 (g).

[47] Ibid., art. 15 (1).

[48] National Institute of Standards and Technology, *Four Principles of Explainable Artificial Intelligence*, - NISTIR 8312 (2021), p. 3. См. https://doi.org/10.1017/bhj.2023.10.

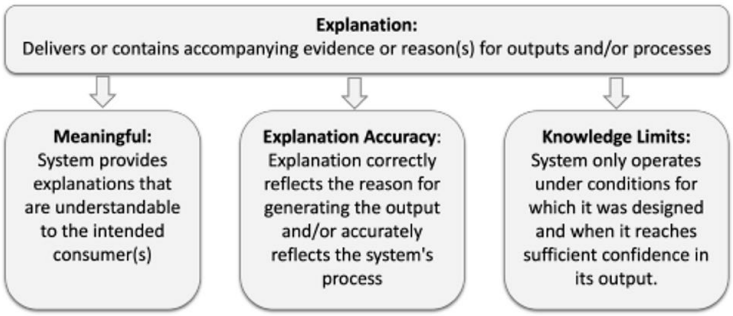

Fig. 1. Illustration of the four principles of explainable artificial intelligence. Arrows indicate that for a system to be explainable, it must provide an explanation. The remaining three principles are the fundamental properties of those explanations.

56. В приведенной ниже таблице разъясняются наиболее значимые аспекты каждого принципа в соответствии с документом Национального института стандартов и технологий[49].

Принцип	*Значение или сфера действия*
Объяснение	Фактические данные, подтверждение или обоснование в отношении результатов или процессов системы [искусственного интеллекта].
Содержательность	Объяснение с использованием понятных предполагаемому потребителю терминов. Иными словами, объяснение должно быть доступным для восприятия целевой аудитории. Качество объяснения зависит от множества факторов, в связи с чем необходимо учитывать целевую аудиторию или аудиторию, которой адресовано объяснение.
Точность объяснения	Технические пояснения должны быть четкими, точными и исчерпывающими.
Пределы знаний	Определение и обозначение пределов знаний подразумевает демонстрацию того, что система не является ни совершенной, ни непогрешимой, поскольку [искусственный интеллект] действует в определенных рамках и пределах, в которых он был запрограммирован. Они также зависят от качества и количества обрабатываемой информации, а также от других факторов.

57. Было указано, что объяснение должно: a) «быть понятным и убедительным для пользователя; b) точно отражать рассуждения системы; c) быть исчерпывающим; и d) быть конкретным в том смысле, что разные пользователи с разными обстоятельствами или разными результатами должны получать разные пояснения[50]». Кроме того, было отмечено:

[49] Приведенные в таблице пояснения представляют собой адаптированный текст и краткое изложение содержания оригинального текста на английском языке, с которым можно ознакомиться по адресу: https://doi.org/10.6028/NIST.IR.8312.

[50] Gavilán, Ignacio, “Cuatro principios para una buena explicabilidad de los algoritmos” (2022). См. https://ignaciogavilan.com/cuatro-principios-para-una-buena-explicabilidad-de-los-algoritmos/.

> «Стремление к обеспечению объяснимости искусственного интеллекта понятно с этической и даже юридической точки зрения, но оно сопряжено со значительными техническими трудностями, которые стоит знать, и, вероятно, существенная часть решения также будет носить технический характер в той мере, в какой возможно перепроектирование существующих алгоритмов или выявление новых алгоритмов, соответствующих этическим и нормативным задачам[51]».

ЮНЕСКО, со своей стороны, указала:

> «Под объяснимостью понимается придание понятного характера результатам систем [искусственного интеллекта] и предоставление возможности получить о них представление. Объяснимость систем [искусственного интеллекта] также связана с понятностью ввода, вывода и функционирования каждого алгоритмического составляющего и его вклада в результаты систем[52]».

58. Чтобы определить сферу действия принципа объяснимости, следует помнить о его цели и, отталкиваясь от нее, установить, что необходимо для ее достижения. В свете вышесказанного было отмечено:

> «Если принцип объяснимости предназначен для того, чтобы любой человек мог знать, почему то или иное решение принимается на основе обработки его данных с помощью инструментов [искусственного интеллекта], то объяснение должно быть по меньшей мере ясным, простым, полным, достоверным и понятным для запрашивающего его лица. Недостаточно сообщить о данных, использованных в качестве исходных для принятия решения, — необходимо представить логику или методологию, использованную для принятия решения. Эта задача сложна, но выполнима, если есть желание доходчиво объяснить людям, почему то или иное решение было принято на основе обработки их персональных данных[53]».

59. Ниже приведены примеры местных законов в странах, которые явно или неявно включили принцип объяснимости в свою правовую базу.

60. В Колумбии законодательство запрещает обработку данных, которая «вводит в заблуждение[54]», и, если говорить непосредственно о решениях, принимаемых по заявлениям на выдачу кредита, требует, чтобы лица, отклоняющие такие заявления, в случае необходимости письменно информировали соответствующее лицо об «объективных причинах отказа[55]».

61. В Эквадоре в статье 20 Органического закона о защите данных предусматривается, что владельцы данных, столкнувшиеся с решением, основанным исключительно или частично на оценках, которые были получены по итогам автоматизированных процессов, включая профилирование, влекут за собой юридические последствия для них или нарушают их основные права и свободы, могут потребовать аргументированного объяснения решения, получить критерии

---

[51] Ibid.

[52] См. https://unesdoc.unesco.org/ark:/48223/pf0000381137, с. 23.

[53] Nelson Remolina Angarita, “Del principio de explicabilidad en la inteligencia artificial (notas preliminares)”, in *Protección de datos personales: doctrina y jurisprudencia*, Pablo Palazzi, ed., vol. III (Centre for Technology and Society, University of San Andrés, Buenos Aires, 2023).

[54] Законодательный акт № 1581 от 2012 года, содержащий общие положения о защите персональных данных, ст. 4 d).

[55] Закон № 2157 от 2021 года, которым был изменен и дополнен Закон № 1266 от 2008 года и в котором содержатся общие положения о принципе «хабеас дата» в отношении финансовой, кредитной, коммерческой, служебной информации и информации о третьих странах и другие положения, ст. 5, п. 1.

оценки автоматизированной программы, представить замечания, запросить информацию о типах использованных данных и их источнике, а также оспорить решение перед ответственным или уполномоченным лицом (за некоторыми исключениями).

62. В Уругвае статья 16 Закона № 18331 гласит:

> «Физические лица имеют право не подпадать под действие решения, влекущего за собой юридические последствия для них и существенно влияющего на них, — решения, которое основано на автоматизированной обработке данных, предназначенной для оценки определенных аспектов их личности, например их трудовой деятельности, кредитоспособности, надежности и поведения. Тот, кого это касается, имеет право получить от лица, ответственного за ведение базы данных, информацию как о критериях оценки, так и о программе, использовавшейся при обработке, на основе которой было принято решение, упомянутое в законе».

VII. Выводы

63. **Из вышесказанного можно сделать следующие выводы:**

a) **прозрачность и объяснимость способствуют укреплению доверия к искусственному интеллекту и соблюдению прав человека;**

b) **разработчики систем искусственного интеллекта должны обеспечивать прозрачность в отношении того, как обрабатываются данные (как они собираются, хранятся и используются), а также в отношении того, как с помощью искусственного интеллекта принимаются решения, надежности таких решений и защищенности информации;**

c) **людям, которых затрагивают решения, принимаемые с помощью искусственного интеллекта, должно предоставляться ясное, простое, полное, достоверное и понятное разъяснение причин принятия такого решения. В этой связи принцип объяснимости имеет решающее значение не только потому, что он согласуется с принципом прозрачности, но и потому, что он позволяет обеспечить осуществление права таких людей на защиту и надлежащую правовую процедуру;**

d) **объяснимость и прозрачность требуют ясности, полноты, достоверности, беспристрастности и открытости решений, принимаемых с помощью искусственного интеллекта, а также логики, метода или аргументации принятия решений в отношении человека на основе информации, в частности персональных данных. Разумеется, объяснимость и прозрачность противоположны неявности, неясности, обману, лжи и злоупотреблению вычислительными возможностями, которые относятся к числу признаков незаконной и неэтичной обработки данных, отражающей отсутствие уважения к человеку и его достоинству.**

VIII. Рекомендации

64. **В свете вышесказанного Специальный докладчик настоятельно призывает государства:**

a) **способствовать повышению прозрачности при использовании искусственного интеллекта в целях снижения рисков, которые может породить ее недостаточность в обществе, особенно в связи с защитой прав человека;**

b) **включить в законодательство принцип объяснимости, чтобы люди могли не только понимать, как принимаются затрагивающие их решения, но и получать доступ к инструментам для защиты своих прав человека в случае использования искусственного интеллекта;**

c) **способствовать распространению этичных методов работы, обеспечивающих прозрачность и объяснимость при обработке персональных данных в проектах или процессах, связанных с использованием искусственного интеллекта;**

d) **поощрять, поддерживать и облегчать просвещение и повышение цифровой грамотности, с тем чтобы граждане лучше знали понятия, связанные с искусственным интеллектом, прозрачностью и объяснимостью, и имели возможность требовать уважения своих прав.**