



Страсбург

CAHAI-COG(2021)01

Специальный комитет по искусственному интеллекту (CAHAI)

**Анализ проведенных консультаций с участием
заинтересованных сторон**

www.coe.int/cahai

Введение¹

В 2020 году Специальный комитет по Искусственному Интеллекту (САНАИ) провел детальное исследование, в котором подробно рассмотрены причины, почему сегодня необходима правовая база для защиты прав человека, демократии и верховенства права в контексте новых вызовов, связанных с интенсификацией использования в повседневной жизни общества системам искусственного интеллекта (ИИ). В 2021 году САНАИ начал рассматривать основные элементы такой правовой базы, которая будет основана на стандартах Совета Европы в области прав человека, демократии и верховенства права.

САНАИ принял решение о проведении консультаций с заинтересованными сторонами в 2021 году для сбора мнений институциональных субъектов (не отдельных лиц) о ключевых проблемах, связанных с созданием вышеуказанных элементов.

Консультации должны помочь Рабочей группе по созданию правовой базы (САНАИ-LFG), которая отвечает за подготовку основных элементов будущей правовой базы, в принятии решений, основанных на полученной в ходе консультаций обратной связи. По сути, ожидается, что элементы позволят найти решение для основных вопросов, такие, как ценности и принципы, на которых должны быть основаны проектирование, разработка и применение ИИ, сферы, где больше всего необходимы защитные и гарантийные механизмы, на выработку политики и различных решений, необходимых для соблюдения ценностей, установленных Советом Европы. В связи с этими и другими вопросами признано важным проведение обсуждения в наиболее широком формате, в ходе которого будут учтены точки зрения различных субъектов: представителей государственного аппарата и должностных лиц органов государственного управления, международных организаций, бизнеса, граждан, научного сообщества и технических специалистов.

Консультации проходили с 30 марта 2021 года по 9 мая 2021 года и основывались на анкете, утвержденной на 4-м пленарном заседании САНАИ. Анкета представлена в Приложении I к настоящему документу. Она включает, помимо вопросов предварительного отбора, 8 закрытых вопросов, 8 вопросов с несколькими вариантами ответов, 22 вопроса по шкале Лайкерта, 11 открытых вопросов и 8 дополнительных открытых вопросов, в

¹ Благодарим за организацию неофициального перевода с английского на русский язык К. Корнеева, Н. Чулкова

основном представляющих вариант «другое». За этот период САНАИ получил 260 ответов на консультации с участием многих заинтересованных сторон и 1 письменную публикацию.

САНАИ ранее решил, что прозрачность будет основным принципом консультаций, а что различные ответы, полученные в ходе консультации, будут включены в подборку ответов и послужат основой для составления этого отчета, который будет обсуждаться в первую очередь Рабочей группой по проведению консультаций САНАИ (САНАИ-COG), а затем самим САНАИ.

Оба документа должны быть опубликованы на сайте САНАИ. Настоящий аналитический отчет, рассмотренный САНАИ на его 6-м пленарном заседании 5-7 июля, будет затем предоставлен Рабочей группой созданию правовой базы САНАИ в качестве необязательного вспомогательного инструмента для использования в работе по подготовке элементов правовой базы.

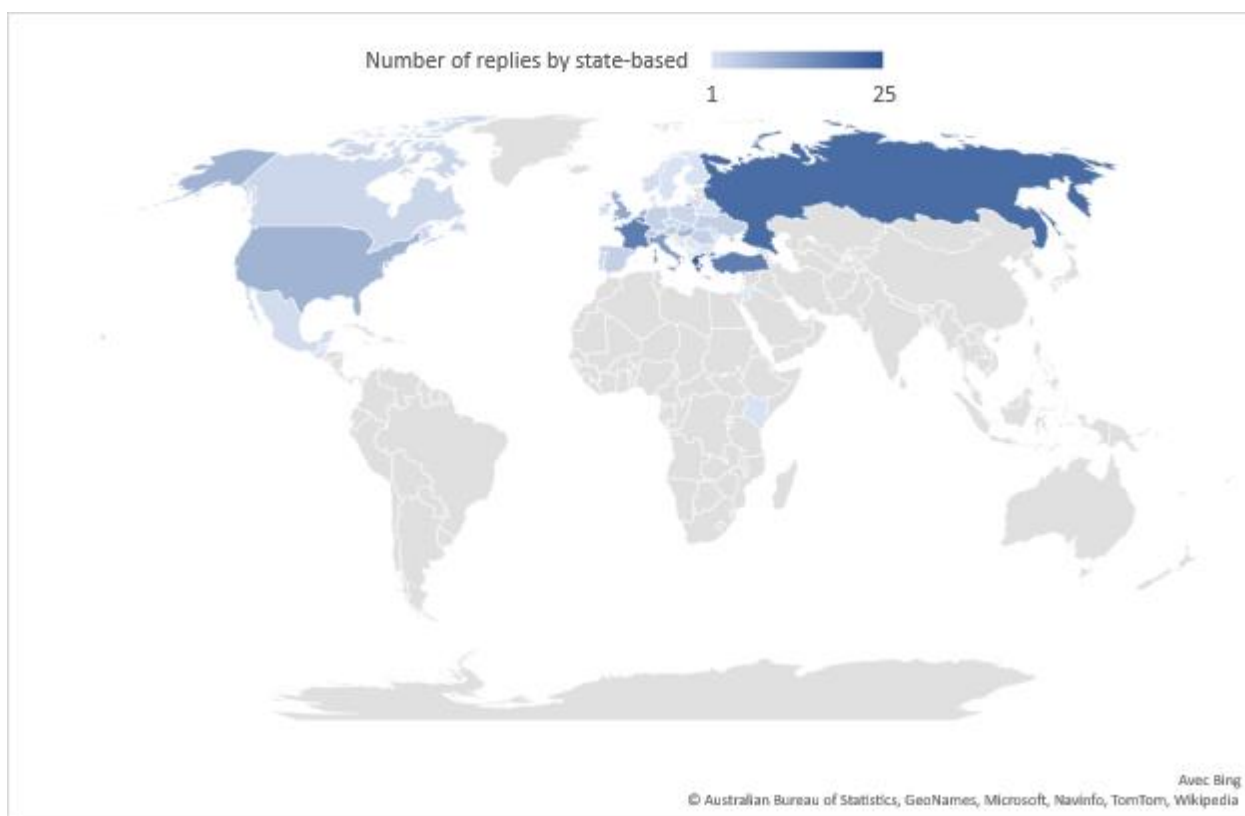
Анализ представляет в виде графиков и процентного распределения ответов различные позиции, выраженные участниками по каждому вопросу анкеты. Этот количественный анализ дополняется качественным анализом ответов на открытые вопросы. Кроме того, для этого типа анализа использовался подход, заключающийся в том, чтобы максимально нейтрально и всесторонне представить разнообразие и числовое соответствие различных выраженных позиций.

Анализ подготовлен секретариатом САНАИ при поддержке профессора Марка-Антуана Дильака, научного эксперта (Франция) и рассмотрен САНАИ-COG на его 3-м заседании 22-23 июня 2021 года.

Предварительный раздел: результаты предварительных вопросов

1. Предварительные вопросы

Географическая репрезентация (учреждения и организации государств)

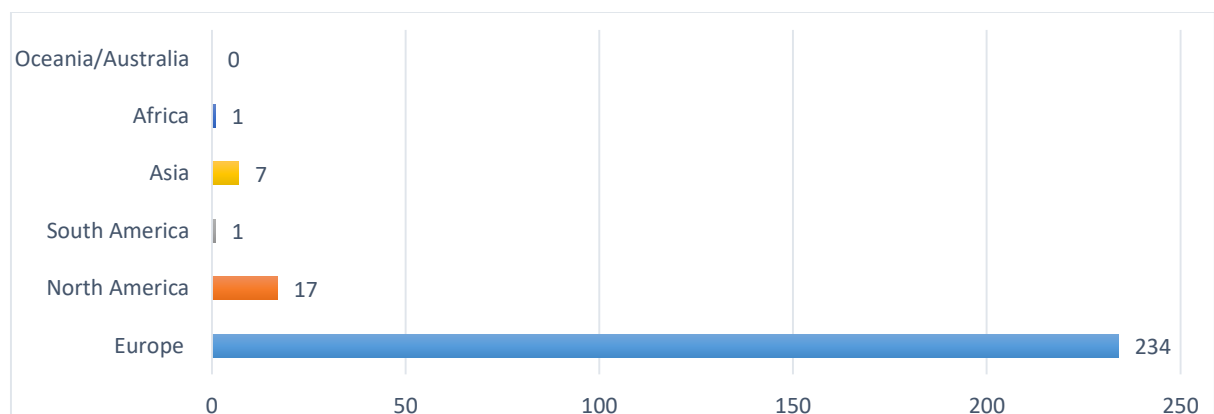


Как упоминалось ранее, общее количество ответов, полученных в рамках консультации, составляет 260. 49 респондентов из стран (Европы и других) приняли участие в консультации. Европейская комиссия направила в секретариат письмо с описанием своего предложения по регулированию ИИ.

234 респондента представляют учреждения или организации, базирующиеся в Европе, в частности из стран, перечисленных ниже.

▶ Andorra		▶ Lithuania	
▶ Austria		▶ Malta	
▶ Azerbaijan		▶ Netherlands	
▶ Belgium		▶ North Macedonia	
▶ Bosnia and Herzegovina		▶ Norway	
▶ Bulgaria		▶ Poland	
▶ Cyprus		▶ Portugal	
▶ Czech Republic		▶ Republic of Moldova	
▶ Denmark		▶ Romania	
▶ Estonia		▶ Russian Federation	
▶ Finland		▶ Serbia	
▶ France		▶ Slovak Republic	
▶ Georgia		▶ Slovenia	
▶ Germany		▶ Spain	
▶ Greece		▶ Sweden	
▶ Hungary		▶ Switzerland	
▶ Ireland		▶ Turkey	
▶ Italy		▶ Ukraine	
▶ Liechtenstein		▶ United Kingdom	

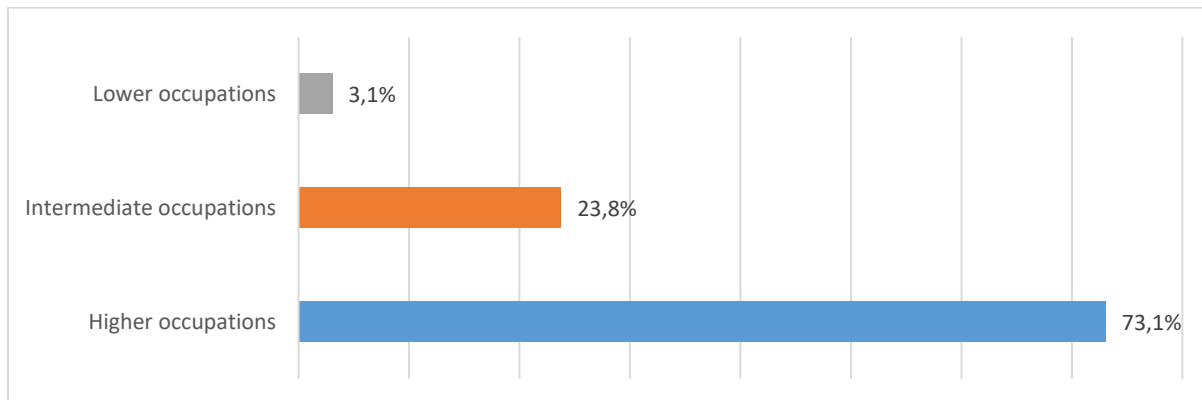
В рамках консультаций получена информация от не европейских стран, таких, как Кения, Мексика, Чили, Канада, США, Гватемала, Израиль, Иордания, Никарагуа и Гонконг.



Согласно данным графика: 0 организаций из Океании, 1 организация из Африки, 7 организаций из Азии, 1 организация из Южной Америки, 17 организаций из Северной Америки, 234 организации из Европы.

Необходимо отметить, что международные межгосударственные организации представляют большое количество государств за пределами Европы, которые не отражены в данном графике, так как учитывалось только принадлежность к государству.

Репрезентация социально-профессиональных категорий

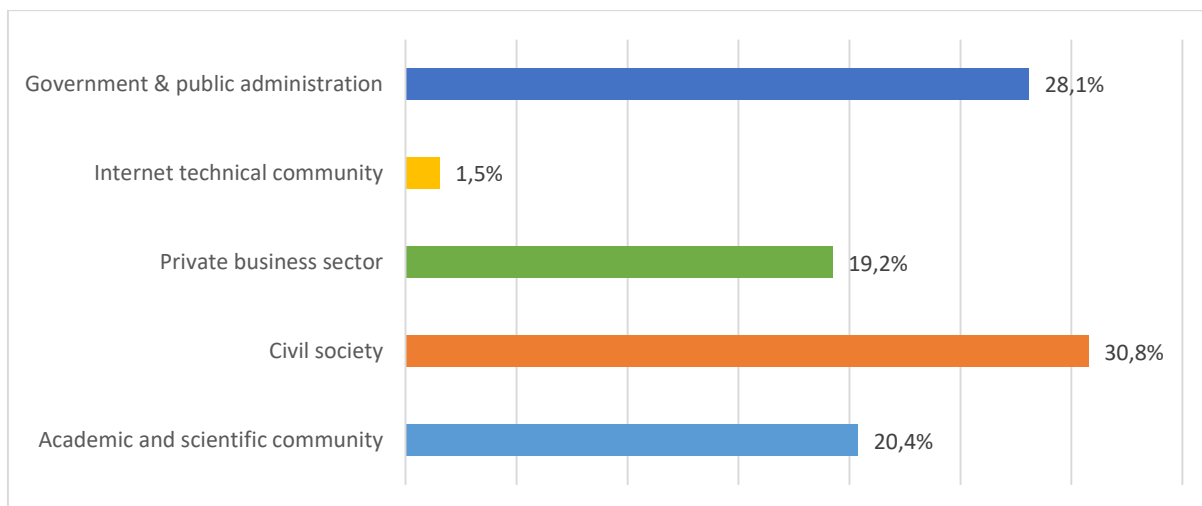


Пункты:

1. е Нижестоящие направления профессиональной деятельности – 3,1%.
2. Направления профессиональной деятельности среднего уровня – 23,8%.
3. Направления профессиональной деятельности высокого уровня – 73,1%.

Респонденты в основном работают по направлениям профессиональной деятельности высокого уровня (73%) и, в меньшей степени, по направлениям среднего уровня (24%). 3% респондентов работают по нижестоящим направлениям профессиональной деятельности.

Репрезентация групп заинтересованных сторон



Пункты:

1. Государственное управление – 28,1%.
2. Сообщество технических специалистов – 1,5%.

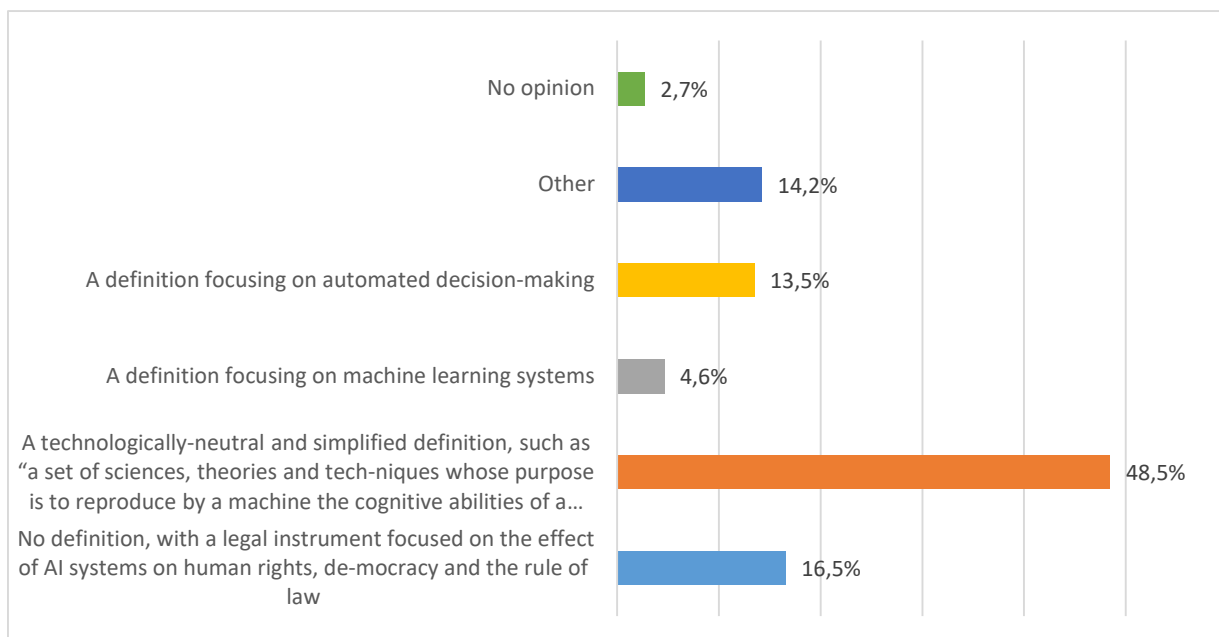
3. Частный бизнес – 19,2%.
4. Гражданское общество – 30,8%.
5. Научное сообщество – 20,4%.

Респонденты представлены государственным сектором (28%), частным бизнесом (19%), гражданским обществом (31%), научным сообществом (20%). Респонденты довольно равномерно представляют государственное управление (28%), частный бизнес (19%), гражданское общество (31%), академическое и научное сообщество (20%).

Следует уточнить, что представители сообщества технических специалистов представлены явно меньше (1,5%). Однако следует отметить, что число наиболее значимых игроков в этой группе весьма ограничено (6), и поэтому количество ответов, полученных в ходе этих консультаций, является достаточно репрезентативным для данной группы для целей настоящих консультаций.

Раздел 1: Определение систем ИИ

2. С учетом развития законодательной базы, регулирующей проектирование, разработку и применение ИИ, основывающейся на стандартах Совета Европы в области прав человека, демократии и верховенства права, какое определение искусственного интеллекта (ИИ) должно быть рассмотрено Специальным комитетом по искусственному интеллекту.



Пункты:

1. Мнение отсутствует – 2,7.
2. Другой вариант – 14,2%.
3. Определение, акцентирующее внимание на автоматическом принятии решений – 13,5%
4. Определение, акцентирующее внимание на системах машинного обучения – 4,6%.
5. Технологически нейтральное и упрощенное определение, например, «совокупность наук, теорий и методов, чье назначение заключается в воссоздании машиной когнитивных способностей человека» – 48,5%.
6. Отсутствие определения, правовой документ сфокусирован на воздействии систем ИИ на права человека, демократию и верховенство права – 16,5.

Данный вопрос подразумевал один вариант ответа из 5.

Вопрос был о видах определений ИИ, не о содержании таких определений. Были доступны три варианта:

- Технологически нейтральное и упрощенное определение, например, «совокупность наук, теорий и методов, чье назначение заключается в воссоздании машиной когнитивных способностей человека».
- Определение, акцентирующее внимание на системах машинного обучения.
- Определение, акцентирующее внимание на автоматическом принятии решений

Респонденты также могли выбрать вариант «Другое» или порекомендовать совсем не использовать какое-либо определение ИИ («Нет определения»).

В то время, как значительное количество респондентов предпочло отсутствие определения ИИ, подавляющее большинство выбрало технологически нейтральное и упрощённое определение ИИ. Многие респонденты дали пример нейтрального и упрощённого определения ИИ. Также было предложено много альтернативных определений. Несмотря на то, что это не было целью опроса, эти примеры укрепили позицию использования технологически нейтрального и упрощенного определения ИИ, даже при условии, что точное содержание определения продолжает оставаться темой для дискуссий и должно быть определено в дальнейшем. Респонденты на самом деле не предложили дополнительные виды определений для рассмотрения, но некоторые из респондентов, выбравших «Другое», предложили использование крайне широкого и всеохватывающего определения, распространяющегося на широкий спектр систем ИИ, например, «вычислительные методы».

3. Причины ваших предпочтений?

На фоне широкого спектра аргументов в пользу одного вида определения или отказа от какого-либо определения ИИ сложился консенсус по четырем нормативным критериям соответствующего определения ИИ для правовой базы:

1. Определение ИИ должно позволять применять правовую базу в реальных ситуациях;
2. Определение ИИ не должно чрезмерно сужать сферу применения правовой базы и должно гарантировать, что оно охватывает широкий спектр угроз и видов ущерба вызываемых компьютерными технологиями;
3. Определение ИИ не должно быть слишком ограничительным и должно допускать включение различных вычислительных технологий и систем;
4. Определение ИИ должно быть ориентировано на будущее и не должно приводить к потере актуальности правовой базы в случае, когда определение устареет;

Респонденты, утверждающие, что никакое определение ИИ не может соответствовать этим нормативным критериям, выбирали вариант «Нет определения». Их аргумент заключается в том, что ИИ - это быстро меняющийся набор технологий и что определение ИИ быстро сделает правовую базу устаревшей (критерий 4). Остальные респонденты считают, что определение ИИ может удовлетворять этим критериям и что определение ИИ,

которое идентифицирует некоторые вычислительные технологии среди других, является предварительным условием для правовой основы, применимой к использованию ИИ.

Эти респонденты разделились во мнениях относительно того, какой вид определения необходим. По мнению некоторых респондентов, слишком широкое определение ИИ (например, определение САНАИ) делает правовую базу расплывчатой и неприменимой (критерий 2). Поэтому они выбирают определение ИИ, ориентированное на автоматическое принятие решений. Хотя эти респонденты осознают ограничения этого определения, они утверждают, что оно позволяет регулировать системы, которые сегодня оказывают наиболее значительное влияние на права человека. Эти системы, даже если технологии ИИ будут развиваться, останутся. Однако против этой точки зрения были выдвинуты серьезные возражения: определение, касающееся автоматического принятия решений, вводит в заблуждение, поскольку система ИИ не принимает фактического решения.

В любом случае для большинства других респондентов слишком конкретное определение чрезмерно ограничивает сферу действия правовой базы и не охватывает угрозы, связанные с различными технологиями ИИ. Более широкое определение будет включать, среди прочего, автоматизированное принятие решений. Эти респонденты поддержали предложение технологически нейтрального и упрощенного определения. Важно отметить, что технологически нейтральное определение не означает, что технология нейтральна - такое определение не зависит от нейтральности технологии. Однако ссылка на нейтралитет сбивает с толку и может быть заменена менее спорным термином. Действительно, большинство респондентов остановились на простоте определения ИИ, которое следует принять САНАИ, а не на его нейтральности, а главный аргумент заключается в том, что определение ИИ должно быть доступно как можно большему количеству людей, как экспертов, так и непрофессионалов.

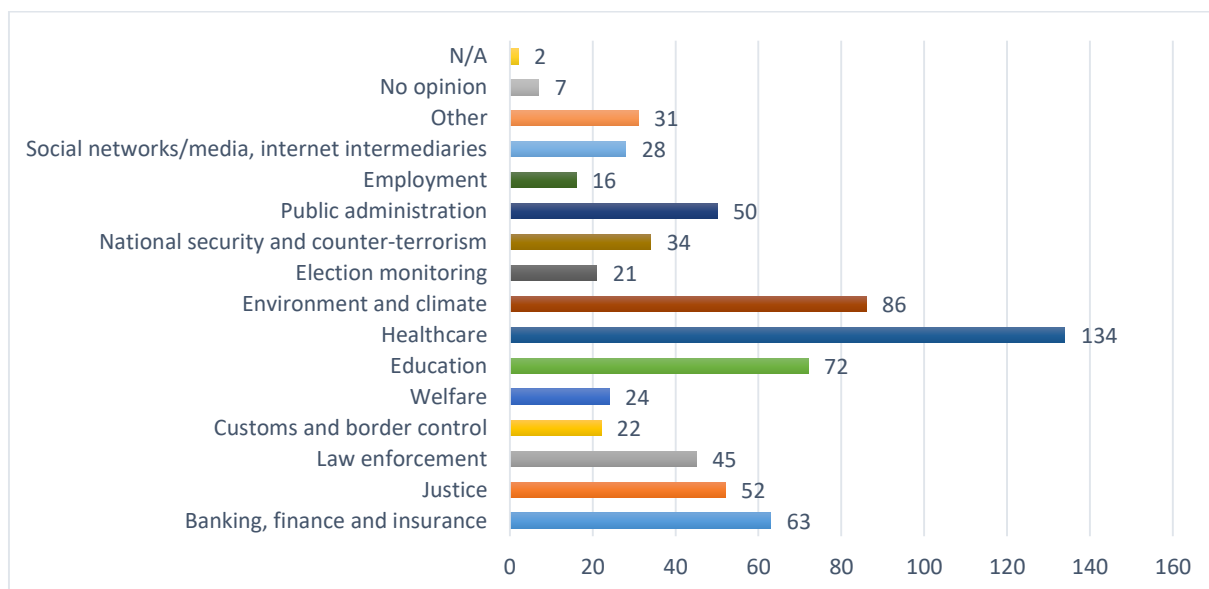
Наконец, несколько респондентов призывают САНАИ избегать антропоморфизма в описании технологий искусственного интеллекта. В частности, утверждалось, что ИИ не воспроизводит когнитивные способности человека, в лучшем случае может их имитировать.

Респонденты ссылаются на три других определения ИИ, которые могут заменить или улучшить определение САНАИ: определение, предложенное Группой экспертов высокого уровня Европейской комиссии по ИИ (Руководящие принципы этики для надежного ИИ, 2019), определение, разработанное некоммерческой организацией AlgorithmWatch и другое, разработанное Всемирной комиссией ЮНЕСКО по этике научных знаний и технологий (COMEST).

Раздел 2: Возможности и риски, создаваемые системами ИИ

Возможности, возникающие благодаря ИИ

4. Пожалуйста, выберите сферы, в которых системы ИИ наиболее перспективны в защите прав человека, демократии и верховенства права



Пункты:

5. Нет ответа – 2.
6. Мнение отсутствует – 7.
7. Другое – 31.
8. Социальные сети/медиа, интернет посредники – 28.
9. Трудоустройство – 16.
10. Государственное управление – 50.
11. Национальная безопасность и контртерроризм – 34.
12. Мониторинг выборов – 21.
13. Окружающая среда и климат – 86.
14. Здравоохранение – 134.
15. Образование – 72.
16. Социальное обеспечение – 24.
17. Таможенный и пограничный контроль – 22.
18. Правовое принуждение – 45.
19. Правосудие – 52.
20. Банковская, финансовая и страховая деятельность – 63.

Каждое опрашиваемое лицо могло выбрать максимум 3 варианта из 15.

3 самых выбираемых сферы, в которых системы ИИ являются наиболее перспективными для защиты прав человека, демократии и верховенства права: Здравоохранение (134), Окружающая среда и климат (86), Образование (72).

За ними следуют Банковская, финансовая и страховая деятельность (64), Правосудие (52), Государственное управление (50).

Трудоустройство – наименее выбираемая опрашиваемыми сфера (16), в которой системы ИИ являются наиболее перспективными для защиты прав человека, демократии и верховенства права.

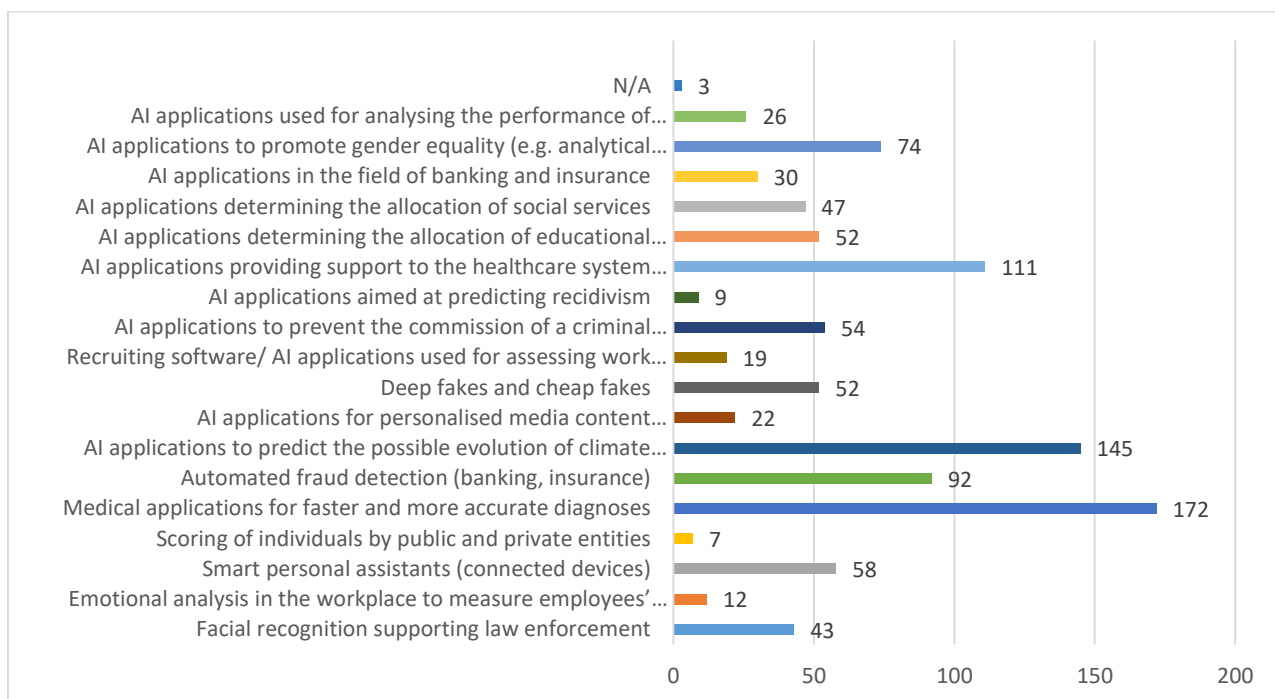
7 респондентов выбрали вариант «Мнение отсутствует», 31 выбрали «Другое», при этом предложив свой вариант с пояснением.

Многие респонденты отметили сложность ответа на данный вопрос, так как одна и та же система ИИ может оказывать как позитивное, так и негативное воздействие вне зависимости от сферы применения; также во всех сферах имеется обширное разнообразие систем ИИ, часть из которых принесет полезные результаты, в то время как другие могут оказывать негативное влияние на права человека. Например, в правовом принуждении, системы ИИ могут помочь в предотвращении преступлений с помощью прогнозирования, когда и где существует наибольшая вероятность их совершения, но также такие системы могут усиливать дискриминацию в отношении людей, проживающих в районах с наибольшей вероятностью совершения преступлений. Все зависит от преследуемых целей, условий и контекста применения, а также возможности правильного использования таких систем.

С учетом этой пояснения, некоторые области кажутся более перспективными, чем другие, и независимо от сектора, сложился широкий консенсус среди респондентов, что развитие ИИ предполагает больше возможностей или меньше рисков для защиты прав человека, демократии и верховенства закона в области здравоохранения, окружающей среды и образования. Однако представители государственных структур склонны выбирать «государственное управление» или «финансы» (банковское дело, страхование), а не «окружающую среду». Для такого выбора есть причины, которые будут понятны в ответе на следующий вопрос.

Некоторые респонденты предложили добавить в список научные открытия и журналистику (или информацию).

5. Укажите какие из предложенных вариантов приложений систем ИИ, по вашему мнению, обладают наибольшим потенциалом для развития и защиты прав человека, демократии и верховенства права.?



Пункты:

1. Нет ответа – 3.
2. Приложения ИИ, используемые для анализа успеваемости учеников/студентов в образовательных учреждениях, таких, как школы и университеты – 26.
3. Приложения ИИ, направленные на развитие гендерного равенства (аналитические инструменты) – 74.
4. Приложения ИИ в банковской и страховой сфере – 30.
5. Приложения ИИ, определяющие распределение социальных услуг/обеспечения – 47.
6. Приложения ИИ, определяющие распределение образовательных услуг – 52.
7. Приложения ИИ, помогающие в сфере здравоохранения (медицинская сортировка, оказание медицинских услуг) - 111.
8. Приложения ИИ, предназначенные для прогнозирования рецидива – 9.
9. Приложения ИИ, предназначенные для предотвращения совершения преступлений (в том числе приложения ИИ, предназначенные для борьбы с отмыванием денег) – 54.
10. Программное обеспечение для найма на работу / приложения ИИ, предназначенные для оценки показателей работы – 19.
11. «Дипфейки» и «чипфейки» - 52.
12. Приложения ИИ, персонализирующие контент (системы рекомендации) – 22.
13. Приложения ИИ, предназначенные для прогнозирования потенциального развития изменения климата и/или природных катаклизмов – 145.

14. Автоматизированное обнаружение мошенничества (банковская и страховая сфера) – 92.
15. Медицинские приложения для быстрого и точного постановления диагноза – 172.
16. Скоринг физических лиц публичными и частными учреждениями – 7.
17. Умные(Smart) персональные ассистенты (подключенные устройства) – 58.
18. Эмоциональный анализ на рабочем месте для оценки вовлеченности в работу – 12.
19. Распознавание лиц для правового принуждения – 43.

Было разрешено выбирать максимум 5 вариантов из предложенных 18.

5 наиболее популярных вариантов приложений систем ИИ, обладающих наибольшим потенциалом для развития и защиты прав человека, демократии и верховенства права:

1. Медицинские приложения для быстрой и точной постановки диагноза – 172.
2. Приложения ИИ, предназначенные для прогнозирования потенциального развития изменения климата и/или природных катаклизмов – 145.
3. Приложения ИИ, помогающие в сфере здравоохранения (медицинская сортировка , оказание медицинских услуг) - 111.
4. Автоматизированное обнаружение мошенничества (банковская и страховая сфера) – 92.
5. Приложения ИИ, направленные на развитие гендерного равенства (аналитические инструменты) – 74.

5 наименее популярных вариантов приложений систем ИИ, обладающих наибольшим потенциалом для развития и защиты прав человека, демократии и верховенства права:

1. Скоринг физических лиц публичными – частными учреждениями – 7.
2. Приложения ИИ, предназначенные для прогнозирования рецидива – 9.
3. Эмоциональный анализ на рабочем месте для оценки вовлеченности в работу – 12.
4. Программное обеспечения для найма на работу / приложения ИИ, предназначенные для оценки показателей работы – 19.
5. Приложения ИИ, персонализирующие контент (рекомендательные системы) – 22.

Также, один респондент предоставил другой ответ, отсутствующий в списке вариантов ответа.

В соответствии с их ответом на предыдущий вопрос респонденты отдают предпочтение приложениям в области здравоохранения, окружающей среды и финансов (банковское дело и страхование). Образование не оценивается высоко, вероятно, в связи с тем, что единственная система ИИ, указанная в образовании, была скорее инструментом государственного управления, чем образовательным инструментом.

6. Пожалуйста кратко объясните как такие приложения будут полезны правам человека, демократии и верховенству права.

Следует отметить, что, хотя многие респонденты назвали ИИ в здравоохранении, окружающей среде или финансах полезным, лишь немногие объяснили, как эти приложения ИИ конкретно способствуют продвижению прав человека, демократии и верховенства права. Этот, казалось бы, анекдотический вывод может выявить у респондентов сложность в том, чтобы различать ИИ во благо (или ИИ для человечества, продвигаемый во многих этических регламентах, которые до недавнего времени лежали в основе дебатов по управлению ИИ) и ИИ для прав человека и демократии. и верховенства права. Например, приложение, которое может предсказывать неминуемое стихийное бедствие и эвакуировать население, находящееся под угрозой, способствует благу, но связь с правами человека, демократией и верховенством права сама по себе не является ни прямой, ни очевидной.

Из большого количества выдвинутых аргументов следует, что приложения ИИ могут способствовать защите прав человека, демократии и верховенства права двумя способами:

1. Путем борьбы с практикой, которая подрывает осуществление процессуальных и гражданских прав, таких как право на справедливое судебное разбирательство и недискриминационное обращение или право на безопасность и свободу слова;
2. Путем улучшения доступа к государственным услугам, которые обуславливают осуществление основными и социальными правами, такими как право на образование.

Приложения ИИ в здравоохранении и образовании чаще попадают во вторую категорию приложений, улучшающих доступ к основным государственным услугам.

Медицинские приложения, которые позволяют быстрее и точнее ставить диагноз, были названы респондентами одними из самых полезных для прав человека, демократии и верховенства права. Действительно, помимо именно медицинских возможностей ИИ, многие респонденты подчеркнули, что использование ИИ облегчит доступ к здравоохранению для наибольшего числа людей за счет снижения затрат на диагностику, терапевтическое лечение и медицинское отслеживание, а также за счет высвобождения человеческих ресурсов, которое приведет к улучшению ухода за пациентами и достойному к ним отношению. Это лучше защитит фундаментальное право на здоровье и право на здоровую жизнь, право, признанное различными хартиями и декларациями, такими как Всеобщая декларация прав человека, Конституция Всемирной организации здравоохранения или Международный пакт об экономических, социальных и культурных правах и Европейская социальная хартия.

Приложения искусственного интеллекта также могут способствовать продвижению права на образование за счет индивидуализации обучения и, таким образом, борьбы с явлениями прекращения обучения и отсева, а также путем обеспечения доступности образовательных услуг для уязвимых, маргинализированных или изолированных групп населения (например, сельских общин). Это право на образование признано Европейской конвенцией о правах человека.

Респонденты также утверждали, что системы искусственного интеллекта, в частности персональные помощники, могут улучшить доступ к государственным услугам, делая информацию более доступной, облегчая и ускоряя административные процедуры.

Эти приложения искусственного интеллекта также помогают бороться с дискриминацией, поскольку они улучшают доступность государственных услуг и социальных благ для всех. Однако, как отметили многие участники, это может быть только в том случае, если эти алгоритмические системы находятся под надзором и регулируются, и если используемые ими данные непредвзяты. Исходя из предположения, что ИИ будет или мог бы быть менее предвзятым, чем люди, большинство приложений ИИ в списке поддерживают права человека, демократию и верховенство права. Многие респонденты считают приложения по предупреждению преступности полезными при условии, что они не являются предвзятыми и используются правильно. Однако ответы на вопрос 6 противоречат ответам на предыдущие вопросы, поскольку эти приложения относятся к областям применения (правовое принуждение и правосудие), которые считались наименее перспективными для прав человека, демократии и верховенства закона - ответы на последующие вопросы покажут, что они также считаются наиболее опасными.

Обобщая ответы на этот вопрос, можно сказать, что полезное использование ИИ для защиты прав человека, демократии и верховенства права зависит от двух условий:

- Вычислительные методы для уменьшения предвзятости систем ИИ или обеспечения того, чтобы они не воспроизводили человеческую предвзятость;
- Правовая база, направленная на недопущение неправильного использования ИИ.

7. Какие еще приложения могут внести значительный вклад в укрепление прав человека, демократии и верховенства закона?

Многие респонденты комментировали свои предыдущие аргументы о преимуществах ИИ, и многие подтвердили свою заинтересованность в аналитических возможностях ИИ,

включая отслеживание мошенничества, кибератак, нарушений авторских прав, пропаганды ненависти, фальшивых новостей и эмоциональных манипуляций.

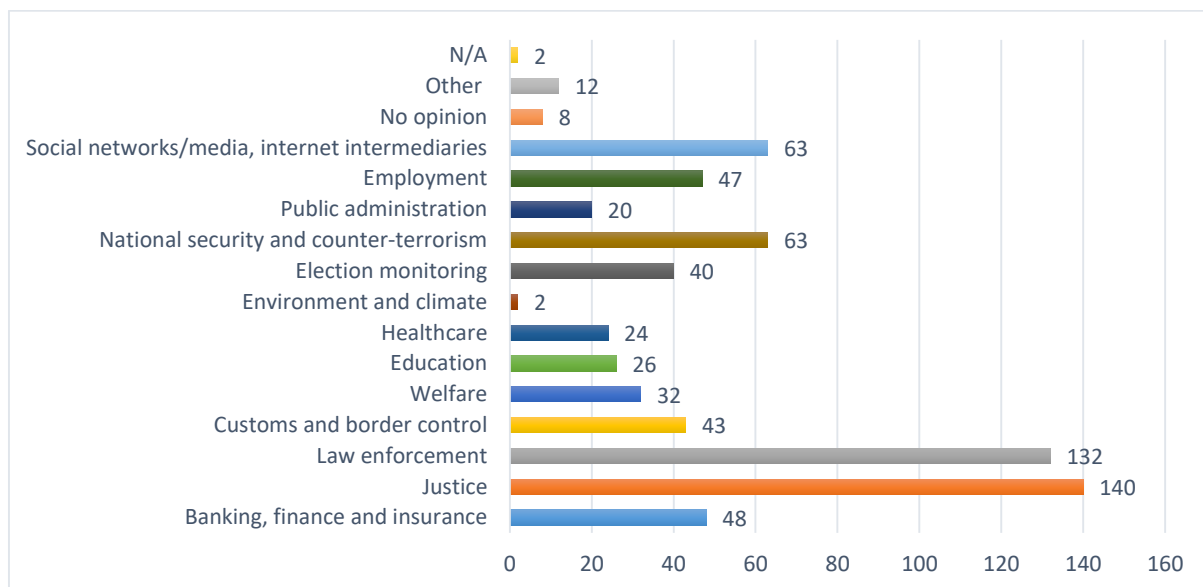
Кроме того, большинство предложений дублируют приложения, перечисленные в предыдущем вопросе, или не демонстрируют преимуществ для прав человека, демократии и верховенства закона. Среди актуальных предложений можно выделить следующие:

- Приложения, способствующие получению самостоятельности лицами с инвалидностью;
- Продвинутая аналитика для исследовательской журналистики;
- Автоматизированный перевод, способствующий активному участию в дебатах, принятии решения и обсуждениях лиц, разговаривающих на языках этнического меньшинства;
- Продвинутая аналитика для исторических и культурных исследований. Такие приложения ИИ способствуют защите прав человека двумя способами: они развивают взаимопонимание и продвигают мирные отношения между культурными сообществами; они помогают людям воссоединиться с своим обществом или лучше свое место в нем, что является первоначальным условием для развития чувства собственного достоинства и социальной автономии.

Последние два предложения указывают на игнорируемую сферу, которая должны быть затронута правовой базой: культурная сфера.

Влияние на права человека, демократию и верховенство права

8. Выберите сферы, в которых использование систем ИИ представляет наибольшую опасность нарушения прав человека, демократии и верховенства права



Пункты:

1. Нет ответа – 2.
2. Другое – 12.
3. Мнение отсутствует – 8.
4. Социальные сети/медиа, интернет посредники – 63.
5. Трудоустройство – 47.
6. Государственное управление – 20.
7. Национальная безопасность и контртерроризм – 63.
8. Мониторинг выборов – 40.
9. Окружающая среда и климат – 2.
10. Здравоохранение – 24.
11. Образование – 26.
12. Социальная поддержка – 32.
13. Таможенные и пограничный контроль – 43.
14. Правовое принуждение – 132.
15. Правосудие – 140.
16. Банковская, финансовая и страховая деятельность – 48.

Можно было выбрать максимум 3 варианта из 15.

3 наиболее выбираемых сферы, в которых использование ИИ представляет наибольшую угрозу нарушения прав человека, демократии, верховенства права: Правосудие (140), Правоохранительная деятельность (132), Национальная безопасность и контртерроризм (63) и Социальные сети/медиа и интернет посредники (63).

За ними идут Банковская, финансовая и страховая деятельность (48) и Трудоустройство (47).

Вариант Окружающая среда и климат является наименее выбираемой респондентами (2) в качестве сферы, в которой использование ИИ может представлять наибольшие риски нарушения прав человека, демократии и верховенства права.

8 опрошенных выбрали вариант «мнение отсутствует» и 12 респондентов выбрали вариант «Другое», предложив свои варианты.

В то время как «здравоохранение», «окружающая среда» и «образование» были наиболее перспективными областями для защиты прав человека, демократии и верховенства закона, неудивительно, что сейчас они относятся к областям, вызывающим наименьшее беспокойство. Напротив, «правосудие» (20%), «правовое принуждение» (19%) и «национальная безопасность и контртерроризм» вместе с «социальными сетями» (9%) считаются очень опасными для прав человека, демократии и верховенства закона.

9. Пожалуйста кратко объясните, как такие приложения могут нарушать права человека, демократию и верховенство права.

Неудивительно, что этот вопрос получил наибольшее количество комментариев, поскольку он очень актуален для разработки будущей правовой базы для ИИ, основанной на стандартах в области прав человека, демократии и верховенства закона. Хотя некоторые респонденты повторяют, что ИИ может быть, как вредным, так и полезным, в зависимости от его использования во всех областях списка, подавляющему большинству респондентов очевидно, что некоторые сферы представляют большую угрозу для защиты прав человека, демократии и верховенства права, чем другие: правосудие, полиция, национальная безопасность. Выявление этих областей основывается на ясных и здравых рассуждениях:

1. Первые три области являются исключительной компетенцией государства; это области, в которых государство применяет принуждение для защиты прав, сохранения условий демократии и обеспечения верховенства закона для всех. Как отмечают некоторые респонденты, это также области, где дисбаланс сил таков, что граждане наиболее уязвимы и подвержены влиянию третьей стороны; те, кто применяет приложения ИИ в этих областях, обладают большей властью над теми, на кого распространяется действие этих приложений.
2. Более того, несмотря на то, что приложения ИИ могут быть использованы не по назначению и причинить вред во всех рассматриваемых областях, угрозы и

характер вреда не одинаковы: при управлении людьми и обществом риск совершения ошибки и возникновения предвзятости считаются намного выше, чем в случае прогнозирования климатической катастрофы или развития опухоли. Более того, вред, причиняемый приложениями искусственного интеллекта, неодинаков по своей природе, хотя их влияние может быть значительнее в здравоохранении или окружающей среде. В случае ошибочного диагноза нельзя утверждать, что право на здоровье было напрямую нарушено. С другой стороны, в случае приложения, которое может привести к тюремному заключению лица, и которое не соответствует требованиям прозрачности или не дает объяснений, право лица на справедливое судебное разбирательство прямо нарушается.

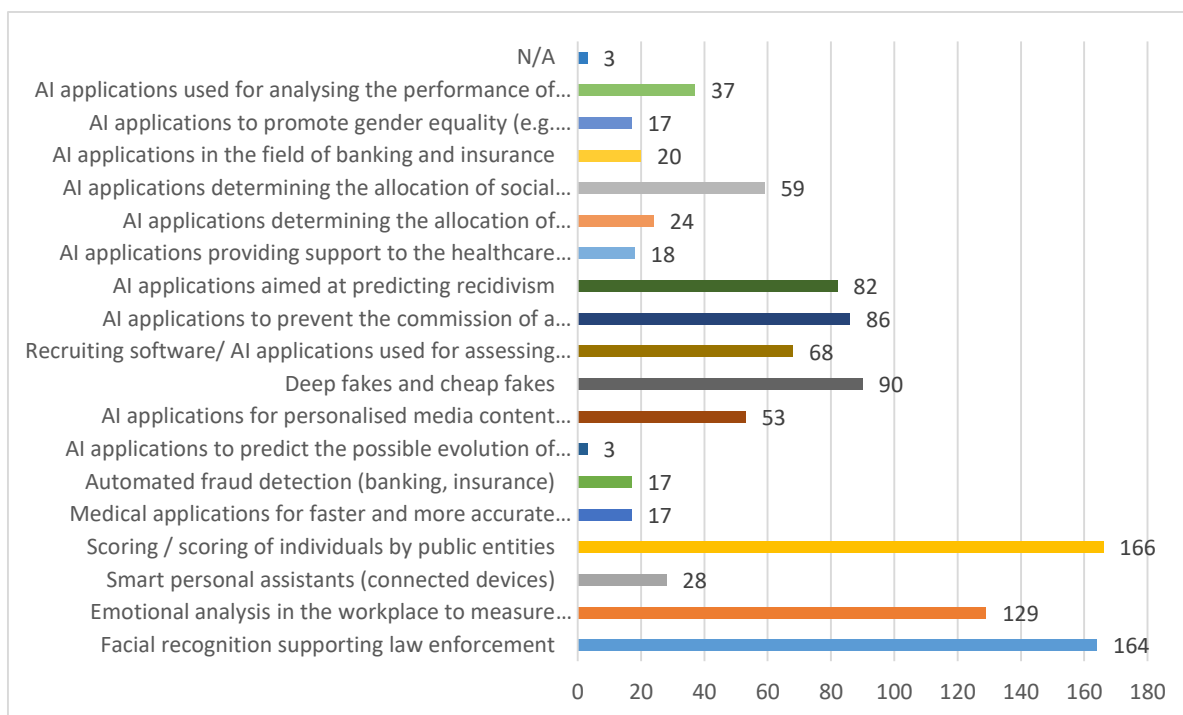
3. Наконец, использование систем искусственного интеллекта в сферах правосудия, правового принуждения (включая пограничный контроль) и национальной безопасности - это не просто инструменты, а способы отправления правосудия и управления.

После указания причин, по которым эти области являются наиболее опасными, респонденты определяют вред, который могут понести люди, и причины этого вреда. Причины - в основном ошибки в алгоритмической обработке, ее неточности и предвзятости, а также низкое качество данных. Основные недостатки следующие:

- Неопределенность судебного процесса для граждан, их невозможность защитить свои права, отсутствие справедливого судебного разбирательства с соблюдением процедуры.
- Неравное обращение и дискриминация (по расовому, религиозному, социально-экономическому, половому и гендерному признакам), ведущие к исключению из системы социального обеспечения.
- Нарушение психологического состояния и благополучия, в том числе путем слежки, манипулирования мнениями и эмоциями.

Большинство респондентов ожидают следующих двух вопросов о типах приложений, наиболее опасных для защиты прав человека, демократии и верховенства права. Поэтому мы полагаемся на анализ вопроса 11, чтобы изучить их аргументы о типах приложений ИИ.

10. Пожалуйста определите виды систем ИИ, которые представляют наибольшую угрозу правам человека, демократии и верховенству права



Пункты:

1. Нет ответа – 3.
2. Приложения ИИ, используемые для анализа успеваемости учеников/студентов в образовательных учреждениях, таких, как школы и университеты – 37.
3. Приложения ИИ, направленные на развитие гендерного равенства (аналитические инструменты) – 17.
4. Приложения ИИ в банковской и страховой сфере – 20.
5. Приложения ИИ, определяющие распределение социальных услуг/обеспечения – 59.
6. Приложения ИИ, определяющие распределение образовательных услуг – 24.
7. Приложения ИИ, помогающие в сфере здравоохранения (медицинская сортировка, оказание медицинских услуг) - 18.
8. Приложения ИИ, предназначенные для прогнозирования рецидива – 82.
9. Приложения ИИ, предназначенные для предотвращения совершения преступлений (в том числе приложения ИИ, предназначенные для борьбы с отмыванием денег) – 86.
10. Программное обеспечения для найма на работу / приложения ИИ, предназначенные для оценки показателей работы – 68.
11. «Дипфейки» и «чипфейки» - 90.
12. Приложения ИИ, персонализирующие контент (рекомендательные системы) – 53.
13. Приложения ИИ, предназначенные для прогнозирования потенциального развития изменения климата и/или природных катаклизмов – 3.

14. Автоматизированное обнаружение мошенничества (банковская и страховая сфера) – 17.
15. Медицинские приложения для быстрого и точного постановления диагноза – 17.
16. Скоринг физических лиц публичными и частными учреждениями – 166.
17. Умные персональные ассистенты (подключенные устройства) – 28.
18. Эмоциональный анализ на рабочем месте для оценки вовлеченности в работу – 129.
19. Распознавание лиц для правового принуждения – 164.

Можно было выбрать максимум пять вариантов из 18 доступных.

5 наиболее выбираемых видов систем ИИ, представляющих наибольшую угрозу правам человека, демократии, верховенству права:

1. Скоринг физических лиц публичными и частными учреждениями – 166.
2. Распознавание лиц для правового принуждения – 164.
3. Эмоциональный анализ на рабочем месте для оценки вовлеченности в работу – 129.
4. «Дипфейки» и «чипфейки» - 90.
5. Приложения ИИ, предназначенные для предотвращения совершения преступлений (в том числе приложения ИИ, предназначенные для борьбы с отмыванием денег) – 86

4 наименее выбираемых видов систем ИИ, представляющих наибольшую угрозу правам человека, демократии, верховенству права:

1. Приложения ИИ, предназначенные для прогнозирования потенциального развития изменения климата и/или природных катаклизмов – 3.
2. Равнозначно, Медицинские приложения для быстрого и точного постановления диагноза (17), Приложения ИИ, направленные на развитие гендерного равенства (аналитические инструменты) (17), Автоматизированное обнаружение мошенничества (банковская и страховая сфера) (17).
3. Приложения ИИ, помогающие в сфере здравоохранения (медицинская сортировка, оказание медицинских услуг) - 18.
4. Приложения ИИ в банковской и страховой сфере – 20.

11. Пожалуйста кратко объясните, как такие приложения могут нарушать права человека, демократию и верховенство права.

Выбор респондентов в целом соответствует их предыдущим ответам. Системы ИИ в сферах правосудия и правового принуждения выходят на первое место, вместе с приложениями в СМИ, но также и в сфере социального обеспечения, области, которая не была отмечена как рискованная. В своих комментариях респонденты сосредотачиваются на трех типах проблемных приложений:

1. Распознавание лиц;
2. Прогнозирующие приложения в правосудии;
3. Приложения социального скоринга.

Во всех случаях респонденты обращают внимание на сквозной вопрос о предвзятости и дискриминации: для большинства респондентов распознавание лиц по своей сути предвзято и гораздо менее точно определяет лица темнокожих людей, чем, например, белых людей. Как прогнозирующие приложения в сфере правосудия, так и приложения для социального скоринга склонны воспроизводить и усиливать дискриминацию, с которой уже сталкиваются обездоленные, стигматизированные и маргинализированные группы населения, поскольку эти приложения основаны на исторических данных, отражающих существующие предубеждения в отношении этих групп населения. Наконец, приложения для эмоционального анализа не могут правильно интерпретировать выражения лиц людей из культур, менее представленных в данных; их результаты, как правило, дискриминируют этих людей, так же, как и людей, которые, как правило, ведут себя не в соответствии с преобладающими социальными нормами. Помимо этих четырех типов приложений, проблема дискриминации возникает и во всех остальных приложениях.

Технологии распознавания лиц получают больше всего комментариев, что очевидно указывает на то, что они являются тестовым примером для разработки правовой базы для использования ИИ. Хотя угроза ошибок, ведущих к судебному преследованию и необоснованному превентивному лишению свободы, четко определены, именно массовое наблюдение вызывает основную озабоченность респондентов. Поскольку проблема слежки не является такой всеобъемлющей, как дискриминация, количество случаев употребления термина «слежка» впечатляет. Но следует отметить, что существует значительный дисбаланс между заинтересованными сторонами в различных секторах: наблюдение - это, в основном, забота организаций гражданского общества с почти двумя сотнями случаев использования этого термина, в сравнении с едва ли десятком в частном секторе и чуть менее двадцати в государственном секторе. Эти данные подчеркивают большой разрыв между ожиданиями гражданского общества и частного и государственного секторов в отношении правовой базы.

По словам респондентов, системы массового наблюдения с искусственным интеллектом, использующие распознавание лиц, нарушают фундаментальное право на неприкосновенность частной жизни и подвергают всех, особенно людей с различным цветом кожи, ошибочно идентифицированных системами распознавания лиц, риску произвольного задержания. Но, как отметили несколько респондентов из разных секторов, даже когда наблюдение не приводит к репрессивному вмешательству, оно наносит психологический вред и снижает активность тех, кто чувствует угрозу, то есть способность действовать свободно, пользуясь своими гарантированными правами. Массовое наблюдение оказывает «сдерживающий эффект на гражданское общество и активизм» и подрывает осуществление основных свобод, таких как свобода выражения мнений, ассоциаций или протестов.

Респонденты также обеспокоены использованием ИИ в судебной системе с приложениями, направленными на прогнозирование рецидивизма и помощь судьям в принятии решений. Некоторые респонденты считают, что, хотя ИИ может сократить время обработки штрафов ПДД и незначительных споров, недостаточная точность и надежность этих систем порождают ошибки (ложные положительные и отрицательные результаты), с которыми гражданам трудно бороться. Отсутствие объяснимости и прозрачности, делающее эти системы объектами с неизвестными свойствами как для граждан, так и для судей, увеличивает ненадежность судебной системы. Это серьезно подрывает право на защиту, принцип состязательности и, следовательно, основное право на справедливое судебное разбирательство.

Приложения социального скоринга обращают внимание на эти же проблемы:

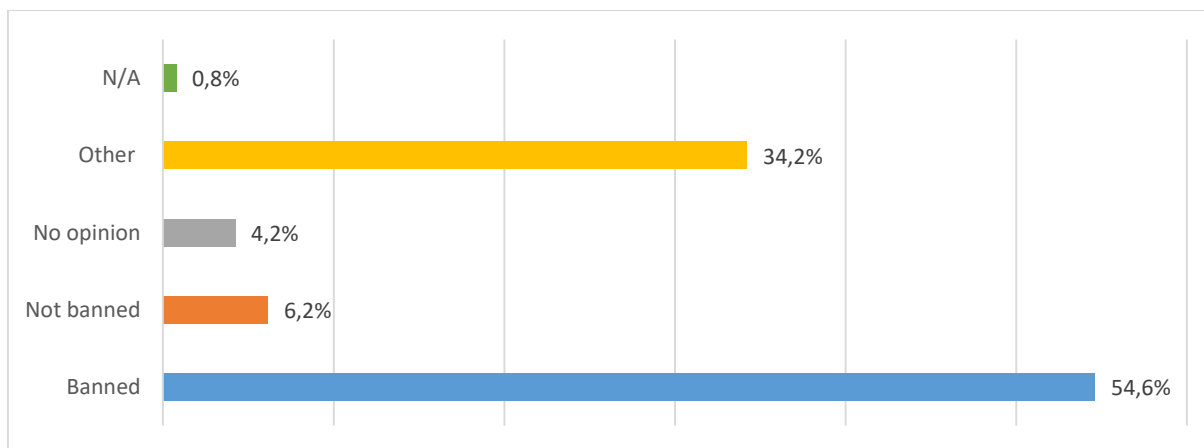
- Они не избавлены от ошибок и предвзятости.
- Они усугубляют положение тех, у кого низкие баллы, потенциально лишая их основных услуг; в зависимости от представленного анализа идентификация наиболее уязвимых слоев населения различается: молодежь, бедные, расовые меньшинства и т. д. Межсекторальный анализ показал бы, что дискриминация носит многофакторный характер и что наиболее обездоленными являются те, кто сталкивается с несколькими видами дискриминации в результате принадлежности в несколько групп.
- Они поддерживают устоявшееся неравенство.
- И, наконец, не давая обоснованных объяснений, они лишают людей возможности обратиться за защитой прав в случае плохой оценки и низких скоринговых баллов.

Для полноты картины, в силу того, что в этом обзоре рассматриваются только те приложения, которые получили наибольшее количество комментариев или неодобрения, мы должны добавить несколько слов о приложениях для эмоционального анализа. Эти приложения занимают третье место в списке самых вредоносных приложений, но, как ни странно, их мало комментируют. К этим приложениям применимо большинство приведенных выше аргументов: предвзятость, ошибки и дискриминация; наблюдение и нарушение конфиденциальности, которую необходимо защищать даже на рабочем месте; и, в особенности, они могут нанести вред продвижению по службе сотрудников, подпадающих под действие этих методов управления.

12. Какие другие приложения могут представлять существенную угрозу правам человека, демократии и верховенства права?

Некоторые респонденты предлагают приложения для микротаргетинга, предназначенные для влияния на политические взгляды и ход избирательных кампаний; другие упоминают приложения для кибератак, но эти приложения де-факто являются незаконными. Большинство респондентов предлагают добавить автономные системы летального оружия, применяемые военными и потенциально используемое для правового принуждения. Однако следует помнить, что использование в военной сфере не входит в компетенцию Совета Европы.

13. По вашему мнению, разработка, развертывание и использование систем ИИ, которые, как было доказано, нарушают права человека или подрывают демократию и верховенство права должны быть:



Пункты:

- 21. Нет ответа – 0,8%.
- 22. Другое – 34,2%.
- 23. Мнение отсутствует – 4,2%.
- 24. Не запрещены – 6,2%.
- 25. Запрещены – 54,6%.

Большинство респондентов считают, что следует запретить разработку, применение и использование систем ИИ, которые, как было доказано, нарушают права человека, подрывают демократию или верховенство закона (55%).

Только 6% респондентов считают, что их не надо запрещать, а 4% не имеют мнения по этому вопросу.

34% опрошенных выбрали вариант «Другое», предложив другие варианты.

Хотя абсолютное большинство ответов поддержали запрет на системы ИИ, которые, как было доказано, нарушают права человека, подрывают демократию или верховенство закона, значительная часть опрошенных выбрала вариант «другое». К сожалению, слишком мало респондентов объяснили этот выбор или предложили альтернативу. Тем не менее, предлагается надежная альтернатива простому запрету, основанная на поэтапном подходе:

1. Системы искусственного интеллекта, разработанные таким образом, что они нарушают права человека, должны быть запрещены. Их назначение обесценивает их.
2. Системы искусственного интеллекта, которые непреднамеренно или косвенно нарушают права человека, должны быть исправлены. Следует принять меры предосторожности и создать механизмы возмещения ущерба, чтобы остановить нарушение прав человека до внесения поправок в приложение.
3. Если разработчики не могут исправить приложение (независимо от того, намереваются ли они это сделать), система должна быть убрана.
4. Должны быть определены исключения из общего запрета для определенного использования и при наличии веских причин.

Более того, любой общий запрет должен быть обоснован недопустимым и постоянным нарушением прав человека, демократии и верховенства права. Поэтому крайне важно иметь четко определенный список запрещенных приложений ИИ, который не подлежит интерпретации.

Вопросы 14, 15, 16

Следующие три вопроса образуют единый блок. Они нацелены на оценку того, насколько респонденты поддерживают или отвергают меры регулирования, в зависимости от уровня риска, создаваемого системами искусственного интеллекта. Предлагается рассмотреть три уровня риска:

- Высокий риск нарушения с высокой вероятностью возникновения
- Низкий риск нарушения с высокой вероятностью возникновения

- Высокий риск нарушения в низкой вероятностью возникновения

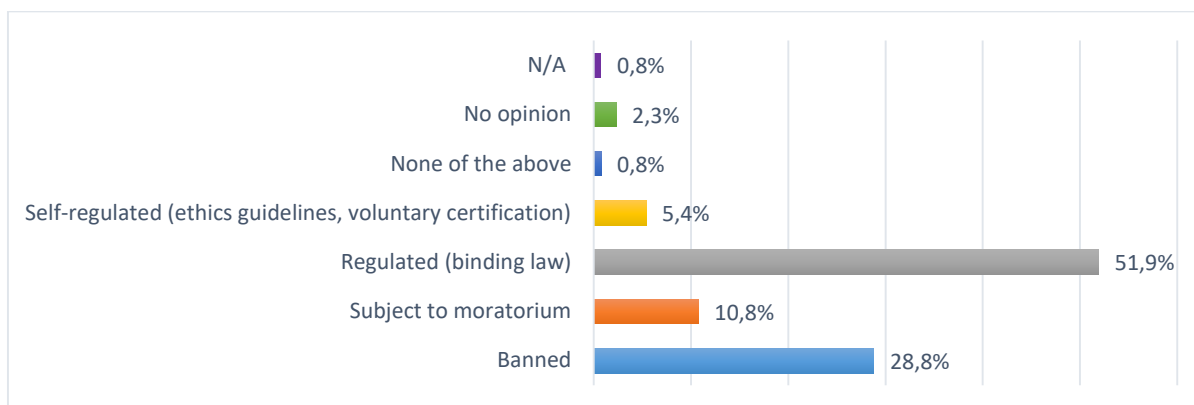
Что касается приложений, представляющих высокий риск нарушения прав человека, демократии и верховенства закона, с высокой вероятностью возникновения (Q14), подавляющее большинство респондентов (52%) поддерживают регулирование с помощью обязательных законов, а значительная часть (29%) даже поддерживают их запрет или введение моратория (11%). Саморегулирование в данном случае не представляется возможным (5%).

Саморегулирование является актуальным вариантом для 28% респондентов в случае с приложениями с низким уровнем риска и высокой вероятностью возникновения (Q15). Мы можем предположить, что этот процент увеличился бы, если бы вопрос касался приложений с низким уровнем риска с низкой вероятностью. Поскольку этот случай не является проблемным, вопрос не задавался. Однако поддержка императивного регулирования остается преобладающей с 58% ответов.

Наконец, респонденты снова массово поддерживают императивное регулирование для приложений с высокой степенью риска с низкой вероятностью (Q16). В этом случае выбор саморегулирования снижается до 14%, а выбор моратория возрастает до 14%.

Ответы очень последовательны и дают четкие указания для создания правовой основы. Главный вывод, который следует сделать, заключается в том, что респонденты поддерживают введение императивного регулирования для приложений искусственного интеллекта, которые представляют риски нарушения прав человека, демократии и верховенства права, независимо от уровня риска. Саморегулирование не представляется надежным подходом к устранению этих рисков.

14. По вашему мнению, разработка, развертывание и использование систем ИИ, которые с высокой вероятностью² являются серьезной угрозой³ для прав человека, демократии и верховенства права, должны быть:



Пункты:

1. Нет ответа – 0,8%.
2. Мнение отсутствует – 2,3%.
3. Ничего из перечисленного – 0,8%.
4. Саморегулируемы (этические руководства, добровольная сертификация) – 5,4%.
5. Урегулированы (императивная норма) – 51,9%.
6. Подвергнуты наложению моратория – 10,8%.
7. Запрещены – 28,8%.

Большинство респондентов считают, что разработка, развертывание и использование систем ИИ, которые с высокой вероятностью являются серьезной угрозой для прав человека, демократии и верховенства закона, должны регулироваться обязательным к соблюдению законом (52%).

29% считает, что такие системы ИИ должны быть запрещены, и 11% считает, что на них должен быть наложен мораторий.

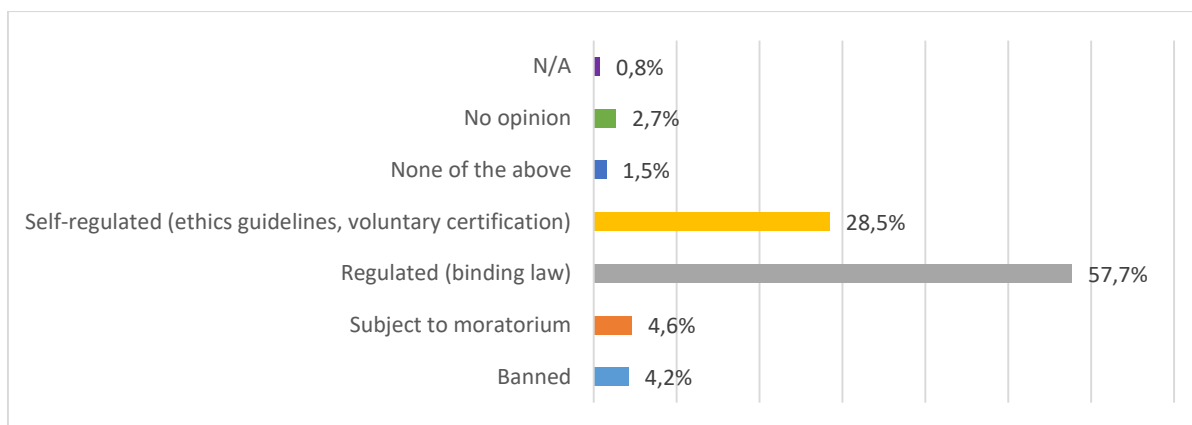
Только 5% опрошенных считают, что они должны быть саморегулируемыми с помощью «мягкого» права, то есть этических руководств и добровольной сертификации.

Менее 1% опрошенных выбрали вариант «ничего из перечисленного», и 2% респондентов выбрали «мнение отсутствует» в данном опросе.

² Высокая вероятность возникновения таких угроз

³ Негативное воздействие на права человека, демократию и верховенство права

15. По вашему мнению, разработка, развертывание и использование систем ИИ, с высокой вероятностью представляющих слабую угрозу правам человека, демократии и верховенству права должны быть:



Пункты:

1. Нет ответа – 0,8%.
2. Мнение отсутствует – 2,7%.
3. Ничего из перечисленного – 1,5%.
4. Саморегулируемы (этические руководства, добровольная сертификация) – 28,5%.
5. Урегулированы (императивная норма) – 27,7%
6. Подвергнуты наложению моратория – 4,6%.
7. Запрещены – 4,2%.

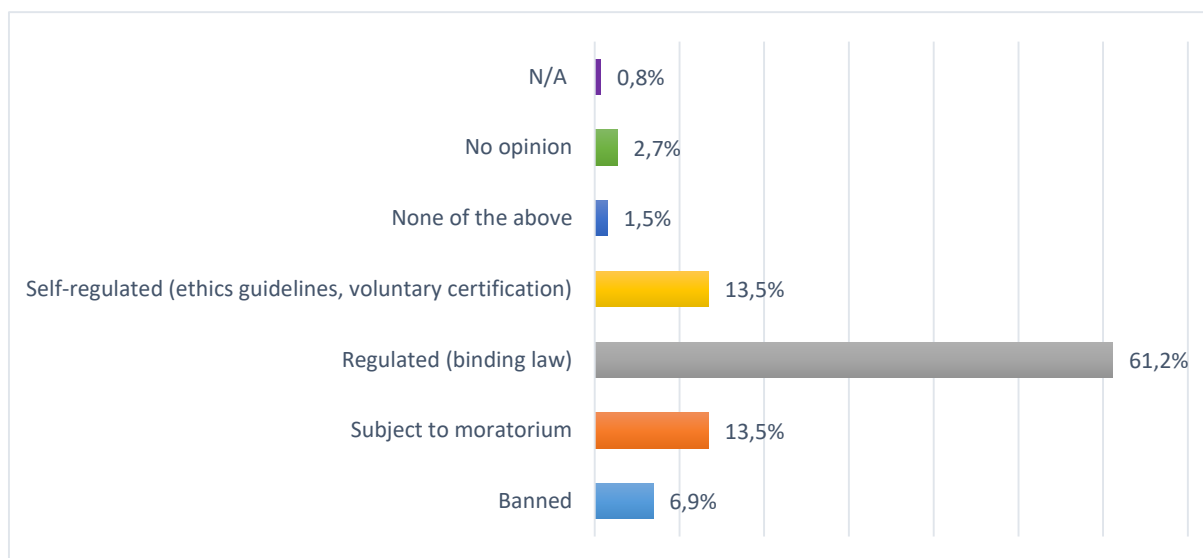
Большинство опрошенных считают, что разработка и использование систем ИИ, которые с высокой вероятностью представляют слабую угрозу правам человека, демократии и верховенству права должны быть урегулированы обязательным к соблюдению законом (58%).

28% считает, что такие системы должны быть саморегулируемы с помощью «мягкого» права, то есть этических руководств и добровольной сертификации.

Только 5% респондентов считает, что такие системы ИИ должны быть подвергнуты наложению моратория, и 4% опрошенных считает, что такие системы ИИ должны быть запрещены.

2% респондентов выбрали вариант «ничего из перечисленного», и 3% опрошенных не имеют мнения по этому вопросу.

16. По вашему мнению, разработка, развертывание и использование систем искусственного интеллекта, которые представляют высокие риски с низкой вероятностью для прав человека, демократии и верховенства закона, должны быть:



Нет данных – 0,8%

Нет мнения – 2,7%

Ничего из вышеперечисленного – 1,5%

Саморегулируемы (этические руководства и добровольная сертификация) – 13,5%

Урегулированы (императивная норма) – 61,2%

Подлежит мораторию – 13,5%

Запрет – 6,9%

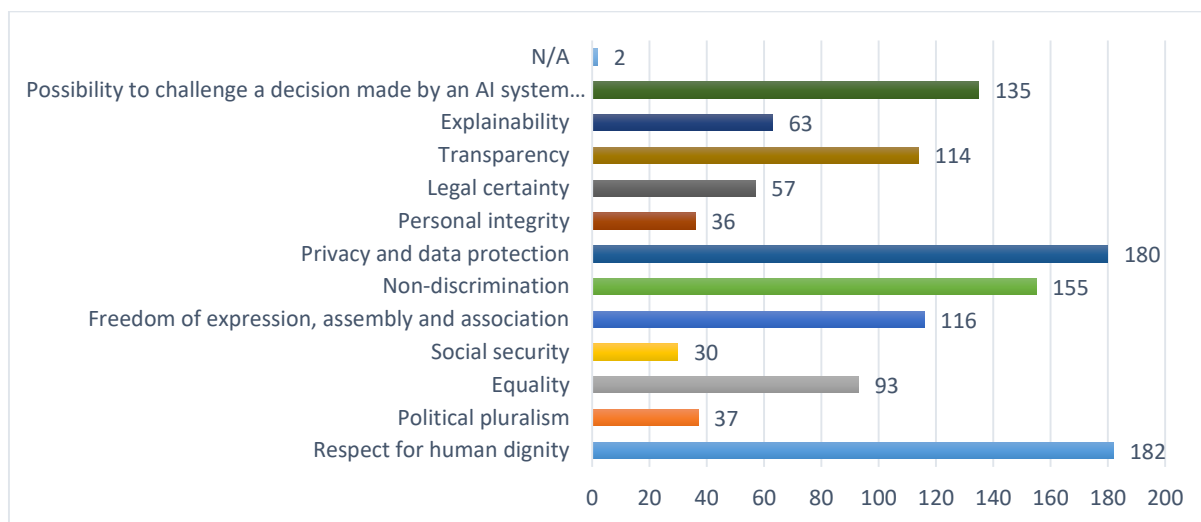
Большинство респондентов считают, что разработка и использование систем искусственного интеллекта, которые представляют высокие риски с низкой вероятностью для прав человека, демократии и верховенства закона, должны регулироваться императивными нормами (61,2%).

13,5% респондентов считают, что необходимо саморегулирование с помощью инструментов мягкого права, таких как этические руководства или добровольная сертификация, и столько же респондентов - что необходим мораторий.

Только 6,9% респондентов считают, что необходим запрет.

1,5% респондентов выбрали вариант «ничего из вышеперечисленного» и 2,7% не имеют мнения по этому вопросу.

17. Каковы наиболее важные юридические принципы, права и интересы, ые необходимо рассмотреть и, следовательно, обосновать регулирование разработки, внедрения и использования систем искусственного интеллекта?



На этот вопрос респонденту было позволено выбрать не более 5 из 12.

Пять наиболее выбираемых юридических принципов, прав и интересов, которые необходимо учитывать и, следовательно, оправдать регулирование разработки, развертывания и использования систем ИИ:

1. Уважение человеческого достоинства (182)
2. Конфиденциальность и защита данных (180)
3. Недискриминация (155)
4. Возможность оспорить решение, принятое системой ИИ, и доступ к эффективному средству правовой защиты (135)
5. Прозрачность (114)

Тремя менее избранными правовыми принципами, правами и интересами, которые необходимо рассмотреть и, следовательно, оправдать регулирование разработки, развертывания и использования систем ИИ, являются:

1. Социальное обеспечение (30)
2. Неприкосновенность личности (36)
3. Политический плюрализм (37)

Среди предложенных принципов некоторые из них являются общими этическими, политическими и правовыми принципами, другие более конкретно применяются к управлению данными и ИИ. Принцип уважения человеческого достоинства является

самым популярным, его выбрали 182 респондента. Этот всеобъемлющий принцип международного и европейского права закреплён в преамбуле и статье 1 Всеобщей декларации прав человека 1948 года и в преамбуле Европейской конвенции о правах человека. Несмотря на различия в интерпретации, наибольшее количество ответов предполагает, что этот принцип должен занять центральное место в новой правовой базе ИИ, основанной на правах человека, верховенстве закона и демократии.

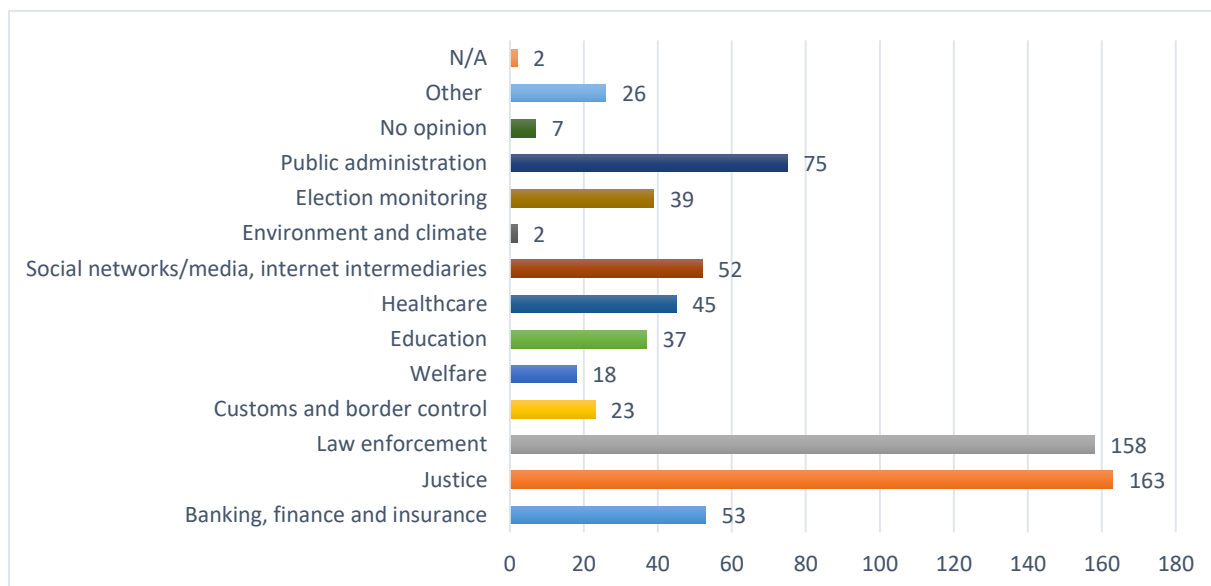
Более того, признание равного человеческого достоинства является основанием для принципа равного обращения или недискриминации, который выбрали 155 респондентов. Этот результат, вероятно, был бы более важным, если бы принцип недискриминации сочетался с принципом равенства.

Признание равного достоинства также подразумевает уважение к частной жизни, второй по популярности принцип: понятие, появившееся в наше время для защиты людей от вмешательства государства, затем в девятнадцатом веке от вторжения прессы, а теперь в наши дни - от социальных сетей. Конфиденциальность - это условие физического и психологического благополучия и неприкосновенности.

Четвертый общий принцип признает право людей оспаривать решение, которое их касается. Этот процедурный принцип лежит в основе верховенства закона, а значит, и демократии. Признание этого принципа в качестве основы для нормативной базы ИИ имеет очень важные последствия: право на оспаривание обусловлено правом на доступ к соответствующей информации, получение значимого объяснения, возможность быть услышанным судьями-людьми, требовать проверки со стороны человека, чтобы получить компенсацию или возмещение, если это применимо. В остальной части анкеты респонденты предлагают ответы, соответствующие этому выбору.

Пятый выбранный вариант - прозрачность. Прозрачность не всегда легко интерпретировать, и респонденты, часто ссылающиеся на нее, редко ее объясняют. Прозрачность — это демократический принцип, который требует от публичных лиц (1) обнародовать соответствующую информацию об исполнении своих служебных обязанностей, (2) объяснять свой выбор или решения, поскольку по определению они влияют на общественность. Применительно к ИИ этот принцип означает, что государственные учреждения и частные субъекты предоставляют информацию об использовании алгоритмов и объясняют решения, которые они принимают, на основе алгоритмических рекомендаций.

18. По вашему мнению, в каких секторах / областях необходим обязательный правовой инструмент для защиты прав человека, демократии и верховенства закона?



На этот вопрос респонденту было разрешено не более 3 выбранных вариантов из 13 вариантов.

Три наиболее выбираемых сектора / области, в которых требуется обязательный правовой инструмент для защиты прав человека, демократии и верховенства закона, — это правосудие (163), правовое принуждение (158) и государственное управление (75).

За ним следуют банковское дело, финансы и страхование (53), социальные сети / СМИ, интернет-посредники (52) и здравоохранение (45).

Вариант «Окружающая среда и климат» - наименее выбранный респондентами (2) сектор/область, в которой необходим обязательный правовой инструмент для защиты прав человека, демократии и верховенства закона.

Наконец, 7 респондентов выбрали вариант «нет мнения», а 26 респондентов выбрали вариант «другое» и представили другое предложение.

Результаты повторяют предыдущие ответы, но некоторые респонденты утверждают, что все области должны подлежать императивному регулированию, поскольку на карту поставлены основные права во всех из них, хотя в некоторых областях регулирование должно быть более строгим из-за более высокого риска нарушений прав человека.

Раздел 3: Возможные пробелы в существующих обязательных юридических документах, применимых к ИИ.

В следующем разделе респондентов попросили указать, в какой степени вы согласны или не согласны со следующими утверждениями, или если у вас нет мнения по данному вопросу.

Ключевые результаты

Следующий раздел направлен на оценку потребности в новом правовом инструменте для развития ИИ. В частности, вопрос заключается в том, является ли обязательное регулирование ИИ наиболее эффективным способом предотвращения и снижения риска нарушений прав человека, демократии и верховенства закона, или же другие подходы к регулированию могут быть реализованы с более конкретными результатами. Например, подход, основанный на саморегулировании компаний, может быть как эффективным, так и достаточным. Компании имеют возможность применять различные инструменты саморегулирования, такие как добровольная сертификация или этические принципы, последние отдают предпочтение 31,5% респондентов, как показано в результатах Q21. Респонденты упоминают несколько этических рамок: наиболее цитируемыми являются Рекомендации по этике надежного ИИ (2019 г.) Группы экспертов высокого уровня по ИИ (Европейская комиссия), а также проект Рекомендации ЮНЕСКО по этике ИИ (2020 г.) и Европейская этическая хартия по использованию искусственного интеллекта в судебных системах, разработанная Европейской комиссией по эффективности правосудия (CEPEJ), за которой следует «Распаковка искусственного интеллекта: 10 шагов по защите прав человека» (Совет Европы, 2019 г.) и Монреальская декларация Ответственное развитие ИИ (2018). Два других документа упоминаются один раз: Rome Call For AI Ethics: A Human-Centric Artificial Intelligence (2020) и Towards Trustworthy AI: Malta Ethical AI Framework for Public Consulting (2019).

Затем предлагаются другие инструменты саморегулирования, такие как кодекс поведения (например, Кодекс этики данных, Россия), отраслевые стандарты (например, IEEE P7000), оценочные списки (например, оценочный лист для надежного искусственного интеллекта группой экспертов высокого уровня по ИИ, Европейская комиссия) и принципы надлежащего управления как в государственном, так и в частном секторе (например, Рекомендация ОЭСР Совета по искусственному интеллекту, 2019 г.). Но большинство заинтересованных сторон в различных секторах, 75,5% и 85% соответственно, считают,

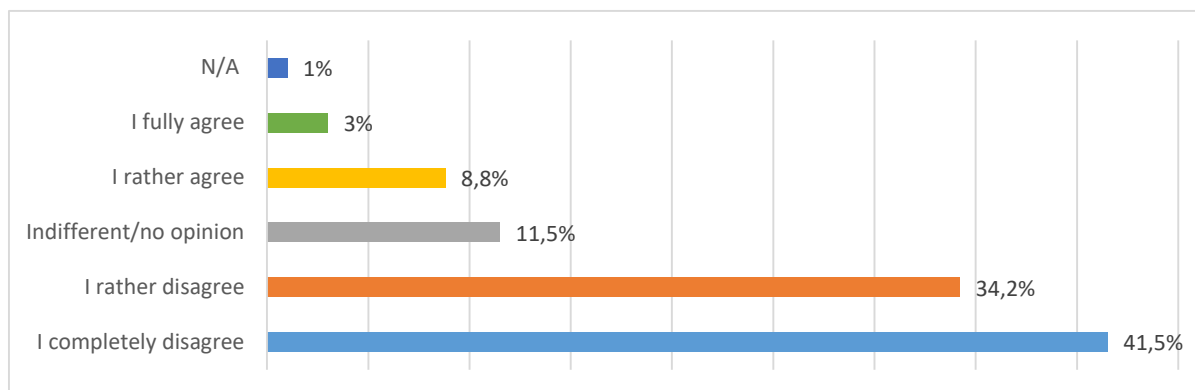
что саморегулирование не более эффективно, чем государственное регулирование (Q19), и не является достаточным (Q20) для предотвращения и снижения риска нарушений прав человека. права, демократия и верховенство закона. И некоторые респонденты прямо сожалеют о том, что Европейская этическая хартия по использованию искусственного интеллекта в судебных системах, Рекомендация Совета по искусственному интеллекту ОЭСР или Монреальская декларация по ответственному развитию ИИ не имеют обязательной силы.

На самом деле, как показывают результаты Q22, даже существующие юридические инструменты, имеющие обязательную силу, недостаточны для регулирования систем искусственного интеллекта для 74% респондентов. Хотя многие респонденты с разных континентов и правовых культур хвалят Европейский общий регламент по защите данных (EU GDPR) как наиболее эффективный юридический инструмент, имеющий обязательную силу, и предлагают его в качестве модели регулирования ИИ, они не думают, что существующие правовые инструменты также применяются конкретно. в AI или эффективно предотвращать нарушения прав человека, демократии и верховенства закона (Q24). Примеры существующих инструментов регулирования, представленных в ответ на Q23, подпадают под ту или иную категорию.

Например, Всеобщая декларация прав человека, Хартия основных прав Европейского Союза или Европейская конвенция о правах человека являются юридически обязательными документами, но не применяются конкретно к AI, если вообще применяются. Напротив, Белая книга по искусственному интеллекту (Европейская комиссия, 2020) или Модель управления ИИ (Сингапур, 2019) конкретно рассматривают управление ИИ, но не являются обязательными правовыми инструментами.

При этом GDPR ЕС и Конвенция 108 (Конвенция о защите частных лиц в отношении автоматической обработки персональных данных, 1981 г.) предлагаются в качестве наиболее близких правовых инструментов к обязательной нормативной базе для ИИ. Конвенция 108+ (модернизированная Конвенция 108 с Протоколом о внесении поправок) еще ближе, поскольку включает в себя соображения, касающиеся алгоритмических процессов принятия решений. Респонденты также обращают внимание на существующие правовые инструменты за пределами Европы, такие как структура FDA США применительно к алгоритмам (нормативно-правовая база для модификации программного обеспечения на основе искусственного интеллекта / машинного обучения (AI / ML) в качестве медицинского устройства), Справедливая кредитная отчетность США. Закон, который применяется к кредитным алгоритмическим системам, и Директива об автоматизированном принятии решений правительства Канады (2020).

19. Саморегулирование компаний более эффективно, чем государственное регулирование, для предотвращения и снижения риска нарушений прав человека, демократии и верховенства закона.

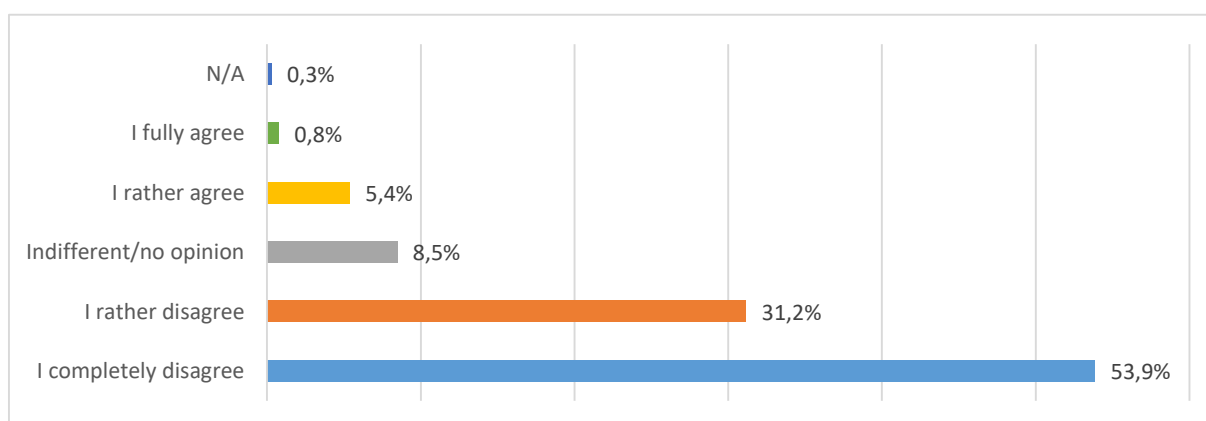


Подавляющее большинство респондентов (75,5%) не согласны с утверждением, что саморегулирование компаний более эффективно, чем государственное регулирование, для предотвращения и снижения риска нарушений прав человека, демократии и верховенства закона (108). Скорее всего, с этим не согласны 89 респондентов.

8 респондентов полностью согласны и 23 скорее согласны.

30 респондентов не имеют мнения или равнодушны к этому утверждению.

20. Саморегулирования компаний достаточно для предотвращения и снижения риска нарушений прав человека, демократии и верховенства закона.



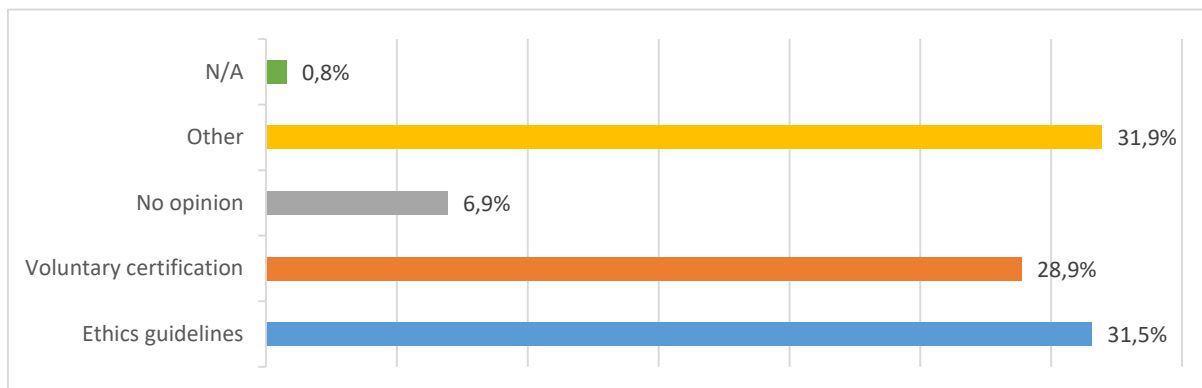
Подавляющее большинство респондентов (85%) не согласны с утверждением, что саморегулирования компаний достаточно для предотвращения и снижения риска нарушений прав человека, демократии и верховенства закона:

- 140 респондентов полностью не согласны с утверждением.
- 81 респондент скорее не согласен с утверждением.

Только 2 респондента полностью согласны и 14 скорее согласны.

22 респондента не имеют мнения или нейтральны к этому утверждению.

21. Какие из следующих инструментов саморегулирования вы считаете наиболее эффективными?



Инструмент саморегулирования, который незначительное большинство респондентов (не выбравших вариант «другое») считает наиболее эффективным - это этические рекомендации (31,5%).

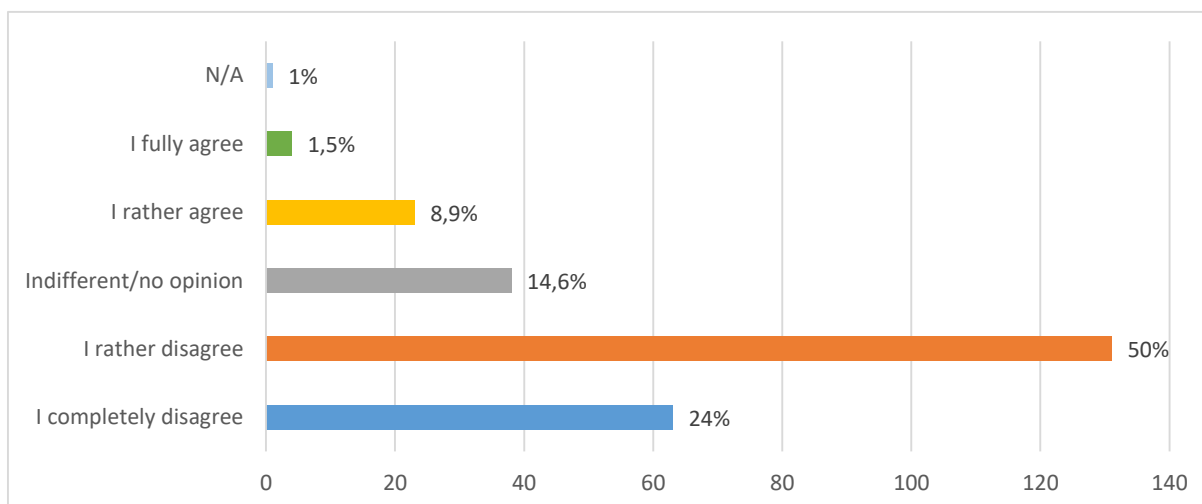
75 респондентов считают добровольную сертификацию наиболее эффективным инструментом саморегулирования.

18 респондентов не придерживаются мнения по данному вопросу.

Однако большинство респондентов (32%) выбрали вариант «другое».

Другие инструменты включают стандарты, оценочные листы, кодексы этики и кодексы поведения.

22. Существующие международные, региональные и/или национальные юридические инструменты, имеющие обязательную и/или необязательную силу, достаточны для регулирования систем ИИ в целях обеспечения защиты прав человека, демократии и верховенства закона.



Подавляющее большинство респондентов (74%) не согласны с утверждением о том, что существующие международные, региональные и / или национальные юридические инструменты, имеющие обязательную и / или необязательную силу, достаточны для регулирования систем искусственного интеллекта с целью обеспечения защиты прав человека, демократии и верховенство закона:

- С утверждением скорее не согласны 50% респондентов.
- 24% респондентов полностью не согласны с утверждением.

Полностью согласны только 1,5% респондентов и скорее согласны 9%.

15% респондентов не имеют мнения или нейтральны к этому утверждению.

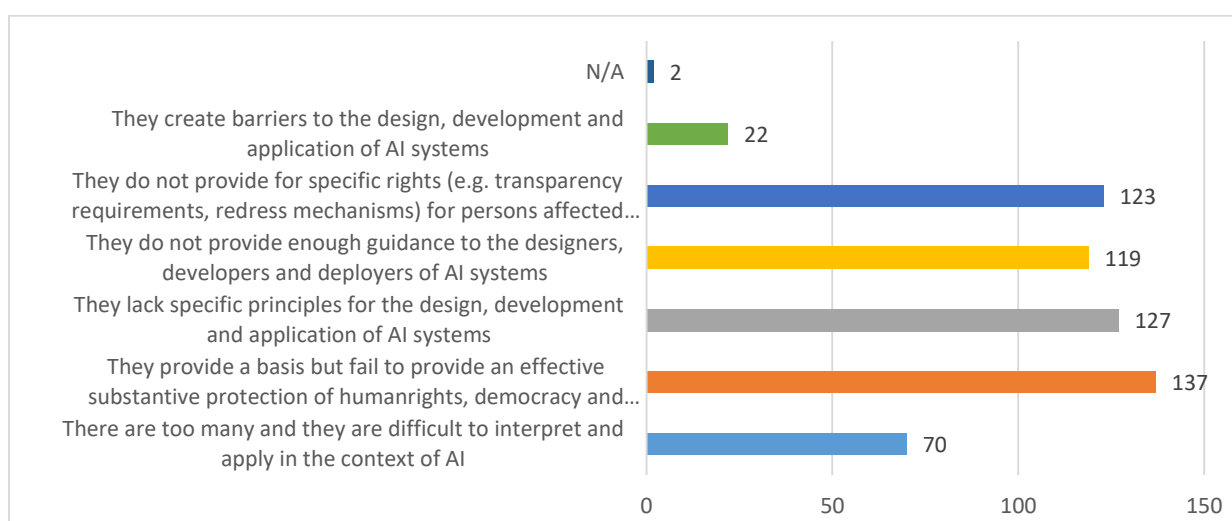
23. Приведите примеры существующих международных, региональных и / или национальных (обязательных и / или необязательных) инструментов, которые, по вашему мнению, эффективны в руководстве и регулировании проектирования, разработки и использования систем искусственного интеллекта для обеспечения совместимости со стандартами прав человека, демократии и верховенства закона.

Не имеющие обязательной силы инструменты	Инструменты, имеющие обязательную силу
Руководящие принципы этики для надежного ИИ (2019 г.) - Группа экспертов высокого уровня по ИИ (Европейская комиссия)	Европейский общий регламент по защите данных (2018 г.) - ЕС
Специальная группа экспертов (Ad Hoc Expert Group - ANEG) по подготовке проекта текста рекомендации по этике искусственного интеллекта (2020 г.) И ОБНОВЛЕНИЕ: Заключительный отчет по проекту текста Рекомендации по этике искусственного интеллекта (2021 г.) - ЮНЕСКО	Всеобщая декларация прав человека (1948) - ООН
Европейская этическая хартия использования искусственного интеллекта в судебных системах (2019 г.) - Европейская комиссия по эффективности правосудия	Хартия основных прав Европейского Союза (2000 г.) - ЕС
Распаковка искусственного интеллекта: 10 шагов по защите прав человека (2019 г.) - Совет Европы	Конвенция о защите прав человека и основных свобод (1950 г., в редакции 2010 г.) - UE
Монреальская декларация об ответственном развитии искусственного интеллекта (2018 г.) - Université de Montréal	«Конвенция о защите частных лиц в отношении автоматической обработки персональных данных» (Конвенция 108-1981) И «Конвенция 108+. Конвенция о защите частных лиц при автоматической обработке персональных данных »(2018 г.) - COE

Стратегия и видение искусственного интеллекта в Мальте до 2030 года (2019 г.) - Мальта	SB-1121 Закон штата Калифорния о защите прав потребителей от 2018 г. (2018 г.) - California Gov.
На пути к надежному ИИ: Мальтийская этическая структура ИИ для общественных консультаций (2019)	Закон о справедливой кредитной отчетности (2018 г.) - FTC US Gov.
Этически согласованная разработка (2019) - IEEE	Директива об автоматизированном принятии решений (2020 г.) - Правительство Канады
Оценочный лист надежного искусственного интеллекта, подготовленный группой экспертов высокого уровня по ИИ (2020 г.) - Европейская комиссия	Директива (ЕС) 2016/680 Европейского парламента и Совета от 27 апреля 2016 г. о защите физических лиц в отношении обработки персональных данных компетентными органами в целях предотвращения, расследования, обнаружения или преследования преступников, правонарушений или исполнения уголовных наказаний, а также о свободном перемещении таких данных, а также об отмене Рамочного решения Совета 2008/977 / JHA (2016) - UE
Рекомендация Совета по искусственному интеллекту (2019 г.) - OECD	Директива Совета 2000/43/ЕС от 29 июня 2000 г., реализующая принцип равного обращения с людьми независимо от расового или этнического происхождения (2000 г.) - UE
Белая книга по искусственному интеллекту. Европейский подход к совершенству и доверию (2020 г.) - Европейская комиссия	Директива Совета 2000/78 / ЕС от 27 ноября 2000 г., устанавливающая общие рамки для равного обращения в сфере труда и занятий (2000 г.) - UE
Модельная структура управления искусственным интеллектом (2019 г.) - Сингапур	Регламент (ЕС) 2017/745 Европейского парламента и Совета от 5 апреля 2017 г. о медицинских изделиях (2017 г.) - UE
Торонтская декларация - Защита права на равенство и недискриминацию в системах машинного обучения (2018 г.) - Human Rights Watch и другие группы	Декларация об основополагающих принципах и правах в сфере труда и последующих мерах (1998 г., пересмотрена в 2010 г.) - МОТ
Нормативно-правовая база для модификаций программного обеспечения на основе искусственного интеллекта / машинного обучения (AI / ML) в качестве медицинского устройства (2020 г.) - FDA US Gov.	Трехсторонняя декларация принципов, касающихся многонациональных корпораций и социальной политики (2017 г.) - МОТ
ЛИДЕРСТВО США В ОБЛАСТИ ИИ: план федерального участия в разработке технических стандартов и связанных с ними инструментов (2019) - NIST	Регламент (ЕС) 2019/881 Европейского парламента и Совета от 17 апреля 2019 года о ENISA (Агентстве Европейского союза по кибербезопасности) и о сертификации кибербезопасности информационных и коммуникационных технологий и отмене Регламента (ЕС) № 526/2013 (Закон о кибербезопасности)) (Текст, имеющий отношение к ЕЭЗ) (2019) - UE
Руководящие принципы предпринимательской деятельности в аспекте прав человека (2011 г.) - ООН	Директива об автоматическом принятии решений (2020 г.) - Канада

Руководство по комплексной проверке ответственного делового поведения (2018 г.) - OECD	Кодекс этики данных (2019) - Council for Data Ethics Russia
Руководящие принципы по многонациональным предприятиям (2011 г.) - ОЭСР	Закон № 78-17 от 6 января 1978 г. об обработке данных, файлах и свободах (1978 г.) - Франция
Рамочное соглашение о цифровизации европейских социальных партнеров 2020 г. (2020 г.) - Конфедерация европейского бизнеса или Business Europe	ЗАКОН № 2016-1321 от 7 октября 2016 г. о цифровой республике (2016 г.) - Франция
Руководство по кибербезопасности (2018) - IBA	Закон о защите данных 2018 (2018) - Великобритания
Этика Римского CallForAI; Искусственный интеллект, ориентированный на человека (2020) - Rome Call	
Стратегия искусственного интеллекта в Австрии до 2030 года (2018 г.) - Австрия	
Ответственное использование искусственного интеллекта (2016-2021) - Канада	
Инструмент оценки алгоритмического воздействия (2021 г.) - Канада	
ExpiAI (2019) - ICO и Институт Алана Тьюринга	
Модель управления ИИ (2019 г., обновлено в 2020 г.) - Комиссия по защите персональных данных	

24. Если вы ответили "не согласен/полностью не согласен" на вопрос 22, укажите, почему существующих международных, региональных и/или национальных (обязательных и / или необязательных) правовых инструментов недостаточно для регулирования систем ИИ (выберите все, с чем вы согласны):



Ограничений по выбору количества ответов на этот вопрос не было.

Четыре наиболее популярных варианта, часто вместе:

- Они служат основой, но не обеспечивают действенной реальной защиты прав человека, демократии и верховенства закона от рисков, создаваемых системами искусственного интеллекта (137).
- В них отсутствуют конкретные принципы проектирования, разработки и применения систем искусственного интеллекта (127).
- Они не предоставляют достаточных руководств для проектировщиков, разработчиков и разработчиков систем искусственного интеллекта (119).
- Они не предусматривают конкретных прав (например, требований прозрачности, механизмов возмещения ущерба) для лиц, затронутых ИИ (123).

Менее выбранный вариант - «Они создают препятствия для проектирования, разработки и применения систем искусственного интеллекта» (22).

Один респондент дал другой ответ, которого нет в списке вариантов.

25. Укажите другие конкретные правовые пробелы, которые, по вашему мнению, необходимо устранить на уровне Совета Европы.

Вклад респондентов в толкование термина «правовые пробелы» сильно различается. Их можно классифицировать следующим образом:

а. Нормативный контекст:

Некоторые сосредотачиваются на нормативном контексте и подчеркивают тот факт, что нормативные инструменты слишком многочисленны, разнородны и чаще всего не имеют обязательной силы. Это оставляет слишком много места для интерпретации государствами-членами, создавая риск фрагментации (европейского) внутреннего рынка. В этом контексте регуляторной конкуренции и разрозненной разработки законодательства приоритетом является создание общей правовой базы, которая объединяет различные инициативы на европейском и международном уровнях. Это также требует от Совета Европы сотрудничества с другими межправительственными организациями, такими как ОЭСР.

б. Юридические методы:

В нескольких статьях приводятся рекомендации по методам права, таким как присвоение бремени доказывания. Необходимо переложить бремя доказывания, особенно в случаях дискриминации, чтобы организация, разрабатывающая или развертывающая систему искусственного интеллекта, продемонстрировала, что она не оказывает негативного воздействия (дискриминации, несправедливости) на людей. Авторы этого запроса также хотят, чтобы ответственность лежала на разработчиках и разработчиках, а не на жертвах,

которым придется защищаться, или на системах ИИ, которые не могут нести ответственность.

Предлагаются еще две методологические рекомендации: во-первых, правовая база должна предусматривать метод оценки справедливости, основанный на (постфактум) сравнении результатов систем ИИ по группам, на которые влияют эти результаты (прогнозы, решения); во-вторых, Совет Европы должен принять подход, основанный на оценке риска.

с. Права:

Респонденты также предлагают добавить права, которые на данный момент недостаточно охвачены. Это будет включать право на психическую неприкосновенность и безопасность. Это особенно важно для молодых людей, и некоторые респонденты призывают Совет Европы признать особые права детей в рамках своего законодательства и признать большую уязвимость детей. Наконец, при определенных обстоятельствах должно быть гарантировано право отказаться от использования системы искусственного интеллекта (право на отказ), и должны быть предложены альтернативные варианты без ущерба для тех, кто отказывается.

d. Механизмы

В большинстве ответов «правовые пробелы» называются отсутствующими механизмами для обеспечения эффективного регулирования ИИ и поощрения прав человека, демократии и верховенства закона. Первый тип механизма — это участие граждан в управлении ИИ и создание платформы для взаимодействия властей и компаний с организациями гражданского общества и маргинализированными группами. Еще одним важным механизмом обеспечения демократической прозрачности является создание публичных реестров систем искусственного интеллекта, используемых государственными учреждениями. Можно создать обсерваторию ИИ для мониторинга законодательства и государственной политики в области ИИ и обмена передовым опытом.

Некоторые жалуются, что отсутствие регламентации процессов сертификации приводит к гонке на худой конец в предоставлении аудиторских услуг, если не к конфликту интересов. Они призывают Совет Европы предоставить юридические требования для сертификации систем искусственного интеллекта. Наконец, правовая база должна включать положения о механизмах возмещения и компенсации в случае причинения вреда.

Раздел 4: Элементы правовой базы для систем искусственного интеллекта

В отношении некоторых систем искусственного интеллекта мы можем разумно предвидеть значительный риск для прав человека, демократии и верховенства закона. Учитывая это, в следующем разделе респондентам было предложено указать, в какой степени респонденты согласны или не согласны со следующими утверждениями или нет ли у них мнения по данному вопросу.

Ключевые результаты

В этом разделе вопросы направлены на выявление консенсуса по элементам, которые должна включать правовая база по системам искусственного интеллекта. Для каждого заявления предлагается пять вариантов, чтобы оценить поддержку или отклонение представленных предложений. В целом респонденты достигли высокого уровня консенсуса, при этом большинство предложений получили поддержку более 80% (я склонен согласиться и полностью согласен). Два предложения получают уровень поддержки 67% (Q36) и 73% (Q37), что делает их относительно согласованными предложениями, но которые следует обсудить дополнительно. Наконец, одно предложение (Q40) разделяет участников с уровнем поддержки 53% и уровнем неодобрения 25%.

Респонденты в целом согласны с тем, что люди всегда должны быть проинформированы, когда они взаимодействуют с системой ИИ при любых обстоятельствах (Q26), когда решение, которое затрагивает их лично, принимается системой ИИ (Q27) и когда система ИИ используется для принятия решения. -процесс изготовления, который затрагивает их лично (Q28). Более того, они поддерживают право на осмысленное объяснение решений, основанных на алгоритмах (Q29), и право на то, чтобы алгоритмические решения или решения, основанные на алгоритмических результатах, проверялись и анализировались людьми. (30, 31 и 32 вопросы). Точнее, они согласны с тем, что люди всегда должны иметь право на то, чтобы любое решение, принятое системой ИИ в рамках судебного разбирательства, рассматривалось «человеческим» судьей (93% поддержки), и право требовать пересмотра принятого алгоритмом решения человеком (85% поддержки). С этой целью всегда должно быть лицо, ответственное за анализ решений, основанных на алгоритмах, в государственном секторе и частных компаниях.

Государственные учреждения обязаны не использовать системы искусственного интеллекта для продвижения или дискредитации определенного образа жизни или мнения (например, «социальный скоринг»), а также разрабатывать и применять устойчивые

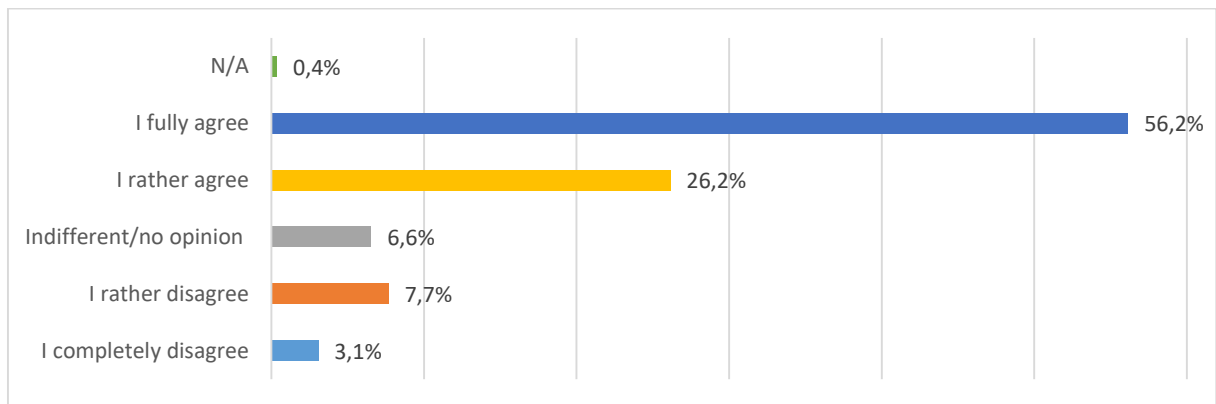
системы искусственного интеллекта, которые соблюдают применимые стандарты защиты окружающей среды, согласно подавляющему большинству, респонденты в вопросах 33 и 34.

Респонденты постоянно сходятся во мнении о механизмах надзора, которые необходимо создать. Для 85% респондентов государства-члены должны создать механизмы общественного надзора за системами ИИ, которые могут нарушать юридически обязательные нормы в сфере прав человека, демократии и верховенства закона (Q38), а 94% респондентов поддерживают механизм отчетности: ошибки и об обнаруженных в системах ИИ проблемах, которые привели или могут привести к нарушению прав человека, демократии и верховенства закона, необходимо сообщать компетентным органам (Q39). Однако, хотя 82% респондентов согласны с тем, что код, лежащий в основе систем искусственного интеллекта, всегда должен быть доступен компетентным государственным органам для целей внешнего аудита (Q35), они в меньшей степени (67%) поддерживают различия в стандартах прозрачности, применимых к общественности, юридическим и частным лицам (Q36). Различия в юридических требованиях между секторами, как правило, менее одобряются, что подтверждается результатами вопроса Q37, где 73% респондентов считают, что должны быть более высокие стандарты доступа к эффективным средствам правовой защиты для отдельных лиц в отношении решений, проинформированных и принятых ИИ. система в области правосудия, чем в области защиты потребителей.

Из предыдущих ответов ясно, что респондентов очень беспокоят системы искусственного интеллекта для распознавания лиц, и многие из них призывают к запрету. Но предложение запретить использование распознавания лиц в общественных местах (Q40) является более спорным и поляризующим: в то время как 53% респондентов поддерживают эту точку зрения, 25% возражают против нее, а 22% не имеют определенного мнения по этому вопросу. Однако на вопрос о том, всегда ли информация, полученная с помощью систем распознавания лиц, должна просматриваться человеком перед использованием в целях, влияющих на личную свободу (Q41), процент неодобрения снижается до 9%.

Результаты по этим двум вопросам соответствуют результатам Q13 и Q14. Большинство респондентов считают, что использование распознавания лиц в общественных местах или для целей наблюдения нарушает права человека и должно быть запрещено. Большая часть респондентов считает, что это технология с высоким уровнем риска, и она поддерживает регулирование, обеспечивающее надлежащее использование ИИ. В том же духе респонденты согласны с тем, что использование систем ИИ в демократических процессах (например, выборах) должно строго регулироваться.

26. Люди всегда должны быть проинформированы, когда они взаимодействуют с системой ИИ при любых обстоятельствах.



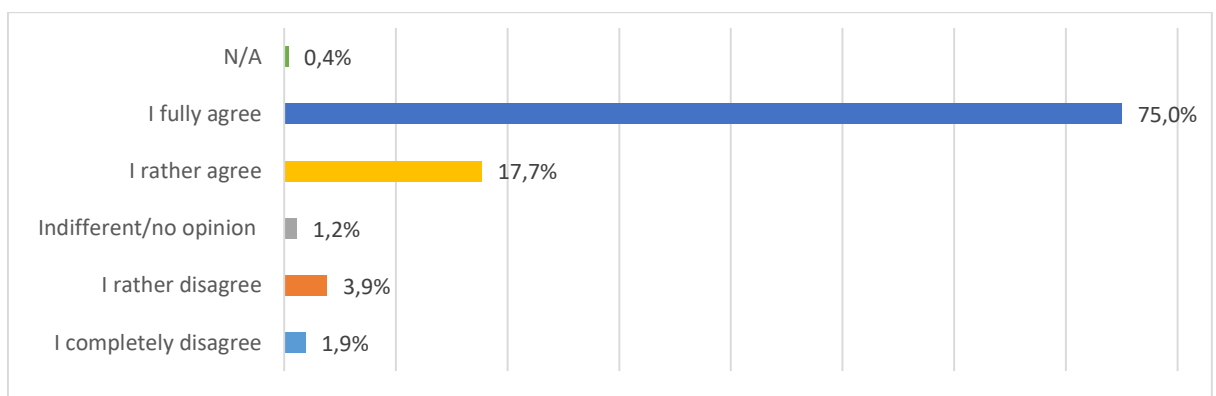
Подавляющее большинство респондентов согласны с утверждением, что люди всегда должны быть проинформированы, когда они взаимодействуют с системой ИИ при любых обстоятельствах:

- 146 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 68 респондентов.

Только 8 респондентов полностью не согласны и 20 скорее не согласны.

17 респондентов не имеют мнения или нейтральны к этому утверждению.

27. Люди всегда должны быть проинформированы, когда решение, которое затрагивает их лично, принимается системой ИИ.



Подавляющее большинство респондентов согласны с утверждением, что люди всегда должны быть проинформированы, когда решение, которое затрагивает их лично, принимается системой ИИ:

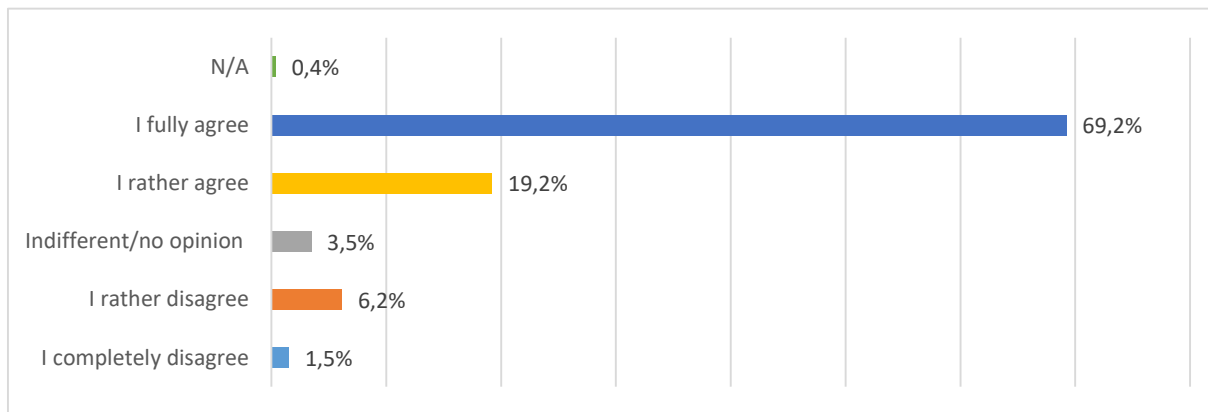
- 195 респондентов полностью согласны с утверждением.

- С утверждением скорее согласны 46 респондентов.

Только 5 респондентов полностью не согласны и 10 скорее не согласны.

3 респондента не имеют мнения или нейтральны к этому утверждению.

28. Люди всегда должны быть проинформированы, когда система ИИ используется в процессе принятия решений, которые затрагивают их лично.



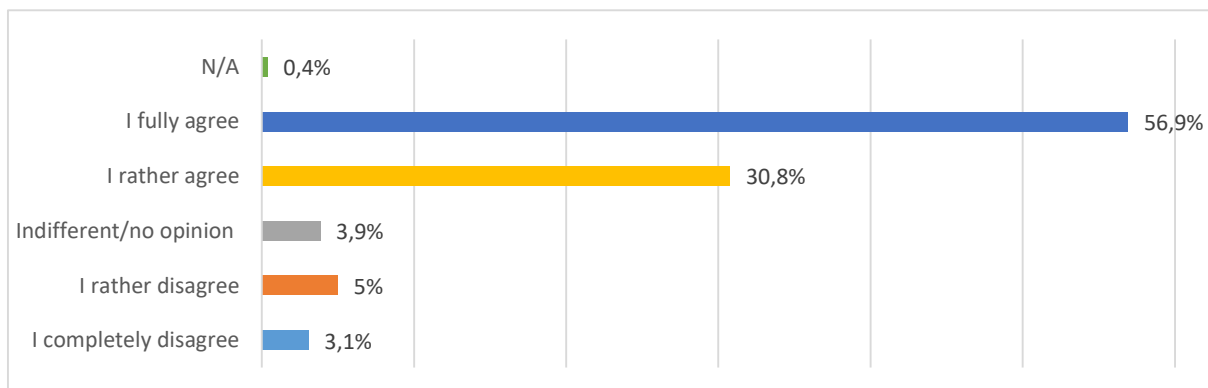
Подавляющее большинство респондентов согласны с утверждением, что люди всегда должны быть проинформированы, когда решение, которое затрагивает их лично, принимается системой ИИ:

- 180 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 50 респондентов.

Только 4 респондента полностью не согласны и 16 скорее не согласны.

9 респондентов не имеют мнения или нейтральны к этому утверждению.

29. Люди должны иметь право на содержательное объяснение решений, основанных на алгоритмах, в частности, как алгоритм достиг своих результатов.



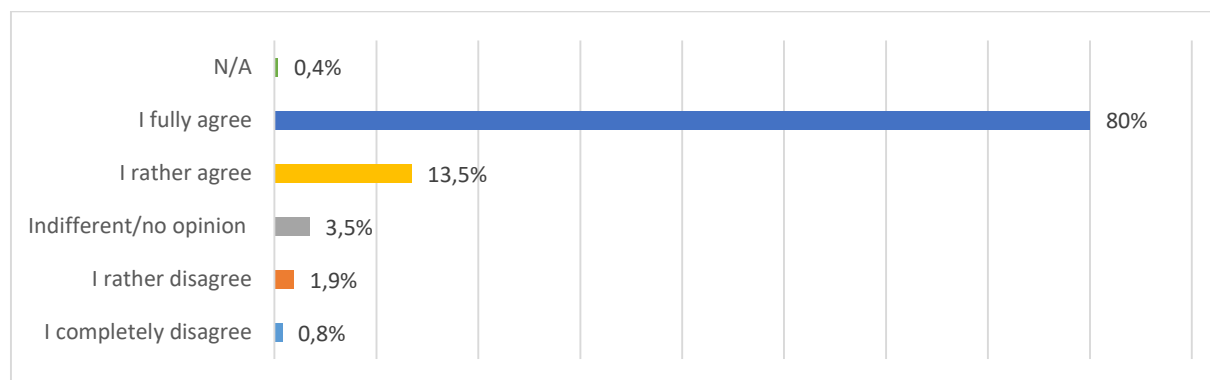
Подавляющее большинство респондентов согласны с утверждением, что отдельные лица должны иметь право на содержательное объяснение решений, основанных на алгоритмах, в частности, как алгоритм достиг своего результата:

- 148 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 80 респондентов.

Только 8 респондентов полностью не согласны и 13 скорее не согласны.

10 респондентов не имеют мнения или нейтральны к этому утверждению.

30. Физические лица всегда должны иметь право на то, чтобы любое решение, принятое системой искусственного интеллекта в рамках судебного разбирательства, было пересмотрено судьей-человеком.



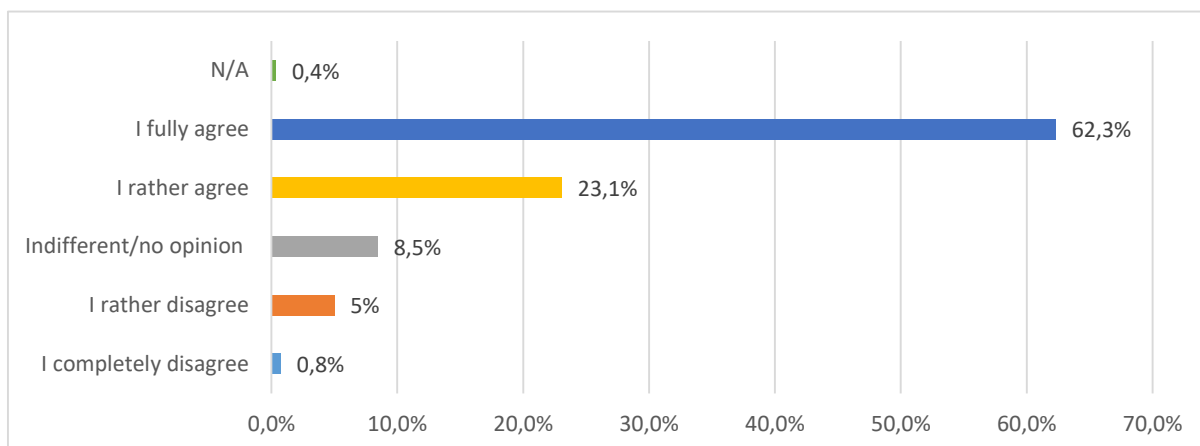
Подавляющее большинство респондентов согласны с утверждением, что физические лица всегда должны иметь право на то, чтобы любое решение, принятое системой ИИ в рамках судебного разбирательства, было пересмотрено судьей-человеком:

- 208 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 35 респондентов.

Только 2 респондента полностью не согласны и 5 скорее не согласны.

9 респондентов не имеют мнения или нейтральны к этому утверждению.

31. Люди должны иметь право требовать пересмотра человеком решения, принятого на основе алгоритмов.



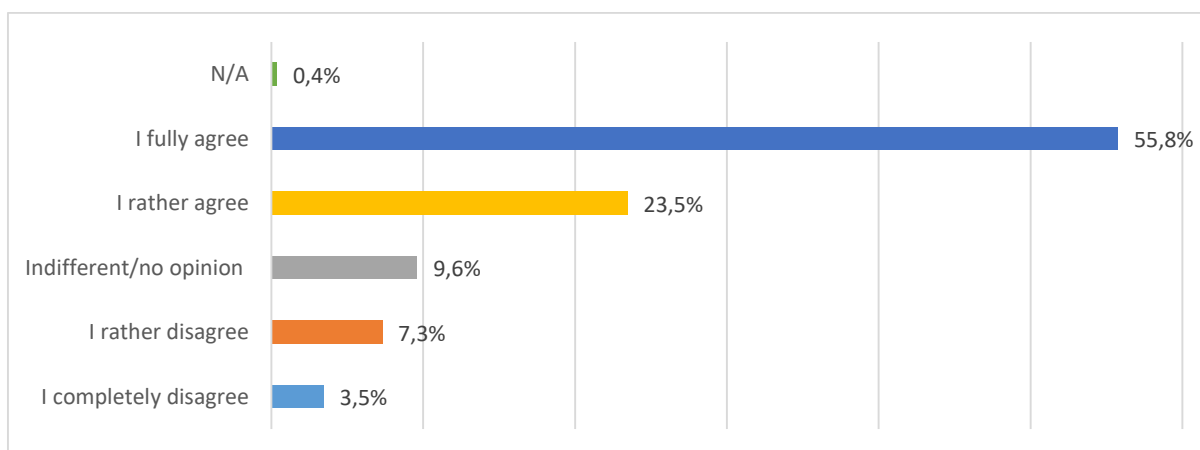
Подавляющее большинство респондентов согласны с утверждением, что люди должны иметь право требовать пересмотра человеком решения, принятого на основе алгоритмов:

- 162 респондента полностью согласны с утверждением;
- С утверждением скорее согласны 60 респондентов.

Только 2 респондента полностью не согласны и 13 скорее не согласны.

22 респондента не имеют мнения или нейтральны к этому утверждению.

32. Всегда должно быть лицо, ответственное за анализ решений, принятых на основе алгоритмах, в государственном секторе и частных компаниях.



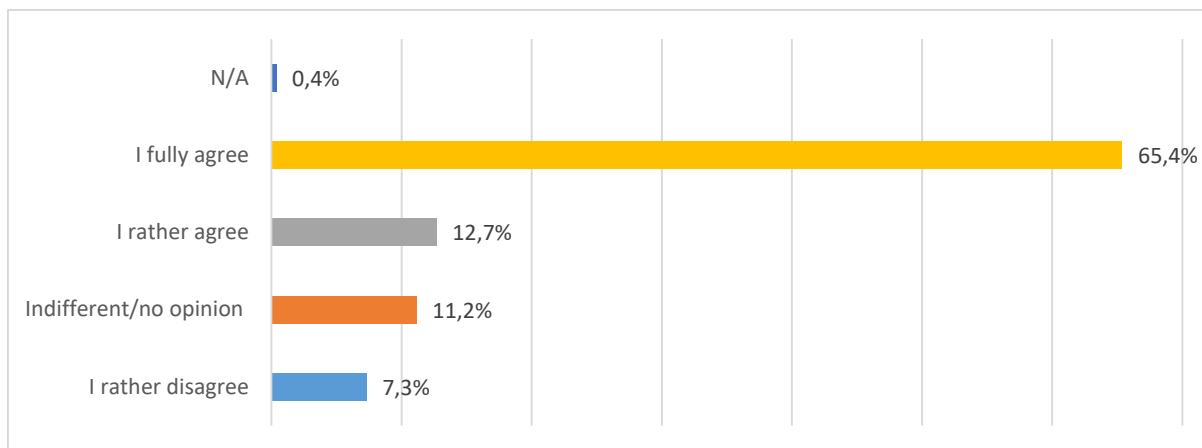
Подавляющее большинство респондентов согласны с утверждением, что всегда должен быть человек, ответственный за анализ алгоритмических решений в государственном секторе и частных компаниях:

- 145 респондентов полностью согласны с утверждением;
- 61 респондент скорее согласен с утверждением.

Только 9 респондентов полностью не согласны и 19 скорее не согласны.

25 респондентов не имеют мнения или нейтральны к этому утверждению.

33. Государственные учреждения не должны использовать системы искусственного интеллекта для продвижения или дискредитации определенного образа жизни или мнения (например, «социальный скоринг »).



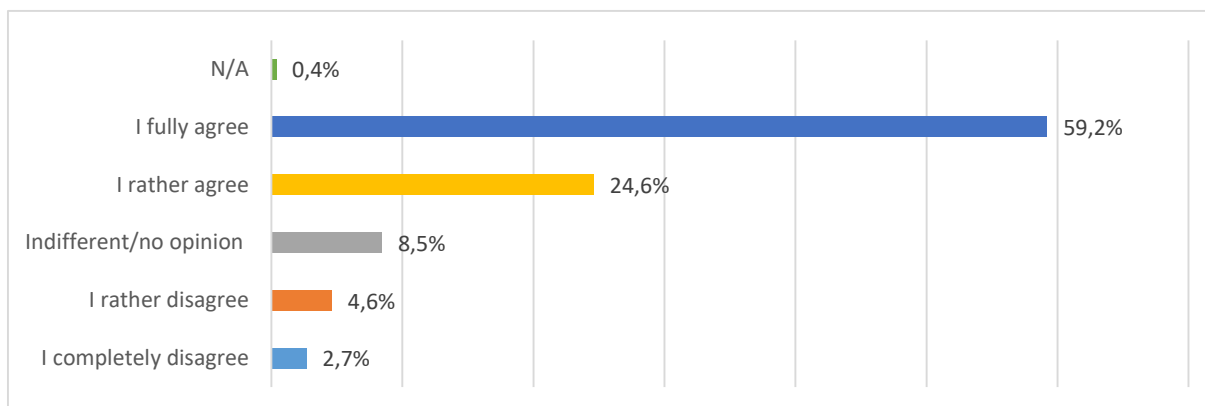
Подавляющее большинство респондентов согласны с утверждением, что государственные учреждения не должны использовать системы искусственного интеллекта для продвижения или дискредитации определенного образа жизни или мнения (например, «социальный скоринг»):

- 170 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 33 респондента.

Только 8 респондентов полностью не согласны и 19 скорее не согласны.

29 респондентов не имеют мнения или нейтральны к этому утверждению.

34. Государства должны быть обязаны разрабатывать и применять устойчивые системы искусственного интеллекта, соблюдающие применимые стандарты защиты окружающей среды.



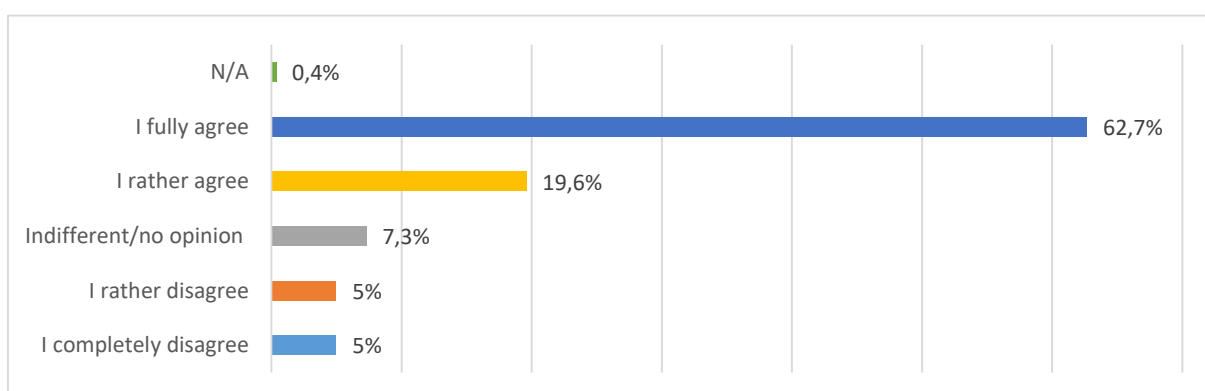
Подавляющее большинство респондентов согласны с утверждением, что государства должны быть обязаны разрабатывать и применять устойчивые системы искусственного интеллекта, соблюдающие применимые стандарты защиты окружающей среды:

- 154 респондента полностью согласны с утверждением;
- С утверждением скорее согласны 64 респондента.

Только 7 респондентов полностью не согласны и 12 скорее не согласны.

22 респондента не имеют мнения или нейтральны к этому утверждению.

35. Код, лежащий в основе систем искусственного интеллекта, используемых в государственном и частном секторах, всегда должен быть доступен компетентным государственным органам для целей внешнего аудита.



Подавляющее большинство респондентов согласны с утверждением, что код систем искусственного интеллекта, используемых в государственном и частном секторах, всегда должен быть доступен компетентным государственным органам для целей внешнего аудита:

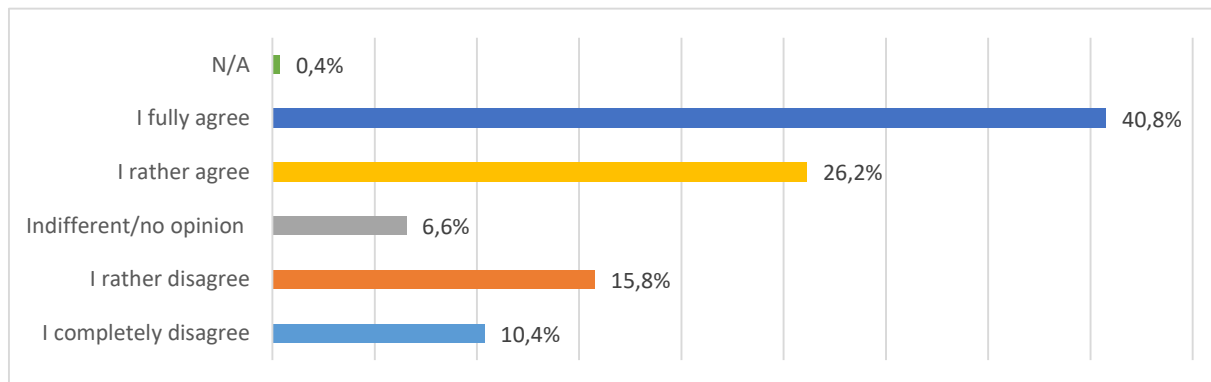
- 163 респондента полностью согласны с утверждением.

- С утверждением скорее согласен 51 респондент.

Полностью не согласны только 13 респондентов и еще 13 - скорее, не согласны.

19 респондентов не имеют мнения или нейтральны к этому утверждению.

36. Должны быть более высокие стандарты прозрачности для государственных организаций, использующих ИИ, чем для частных.



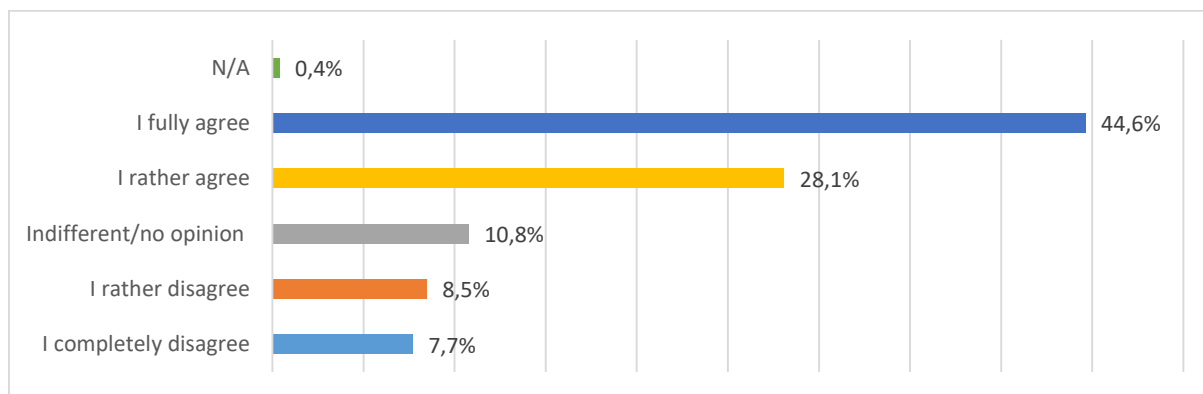
Большинство респондентов согласны с утверждением о том, что для государственных организаций, использующих ИИ, должны быть более высокие стандарты прозрачности, чем для частных:

- 106 респондентов полностью согласны с утверждением;
- С утверждением скорее согласны 68 респондентов.

27 респондентов полностью не согласны и 41 скорее не согласны.

17 респондентов не имеют мнения или нейтральны к этому утверждению.

37. Должны быть более высокие стандарты доступа к эффективным средствам правовой защиты для лиц в отношении решений, принятых системой искусственного интеллекта в области правосудия, чем в области защиты потребителей.



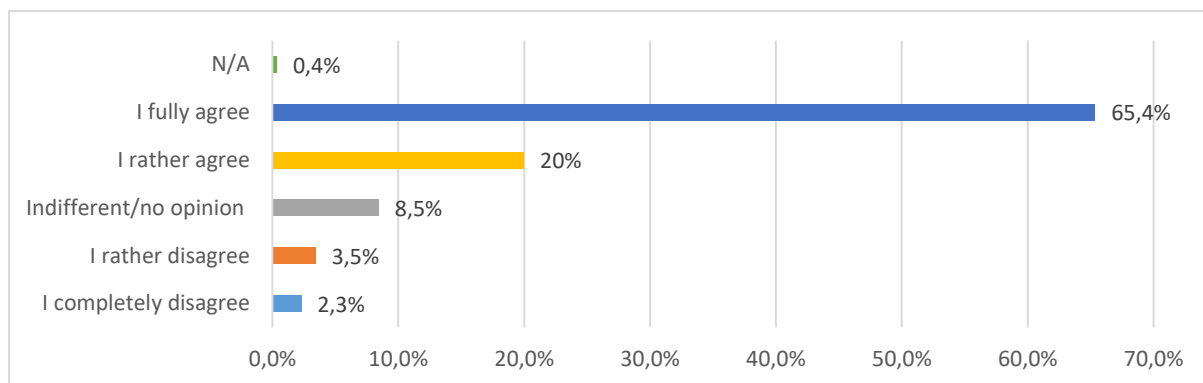
Подавляющее большинство респондентов согласны с утверждением о том, что должны быть более высокие стандарты доступа к эффективным средствам правовой защиты для отдельных лиц в отношении решений, основанных на информации и принимаемых системой искусственного интеллекта в области правосудия, чем в области защиты потребителей:

- 116 респондентов полностью согласны с утверждением.
- С утверждением скорее согласны 73 респондента.

20 респондентов полностью не согласны и 22 скорее не согласны.

28 респондентов не имеют мнения или нейтральны к этому утверждению.

38. Государства-члены должны создать механизмы государственного надзора за системами ИИ, которые могут нарушать императивные нормы в сфере прав человека, демократии и верховенства закона.



Нет данных – 0,4%

Я полностью согласен – 65,4%

Я скорее согласен – 20%

Нейтрально/нет мнения – 8,5%

Я скорее не согласен – 3,5%

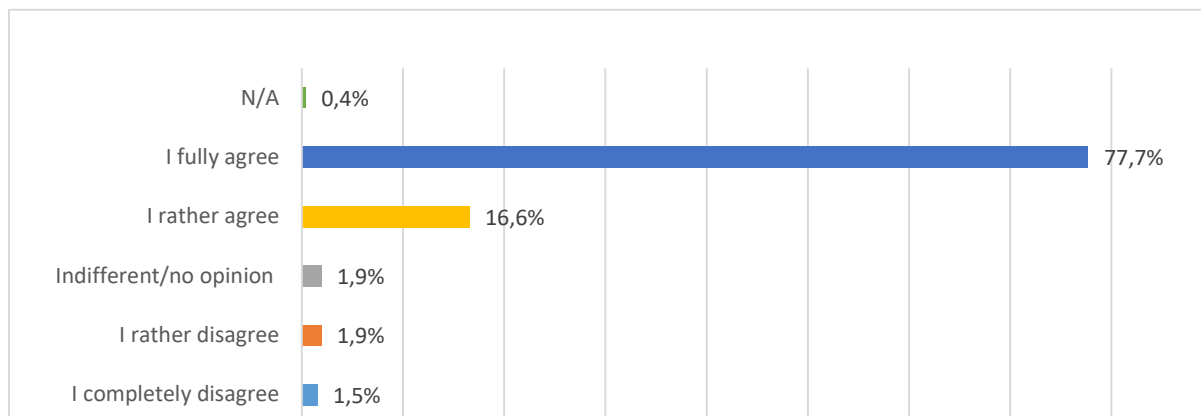
Я совершенно не согласен – 2,3%

Подавляющее большинство респондентов согласны с утверждением, что государства-члены должны создать механизмы общественного надзора за системами ИИ, которые могут нарушать императивные нормы в сфере прав человека, демократии и верховенства закона:

- 170 респондентов полностью согласны с утверждением;
- 52 респондента скорее согласны с утверждением;

- Только 6 респондентов совершенно не согласны и 9 скорее не согласны;
- 22 респондента нейтральны или у них нет мнения на это утверждение.

39. Об ошибках и недостатках, обнаруженных в системах искусственного интеллекта, которые привели или могут привести к нарушению прав человека, демократии и верховенства закона, необходимо сообщать компетентным органам.



Нет данных – 0,4%

Я полностью согласен – 77,7%

Я скорее согласен – 16,6%

Нейтрально/нет мнения – 1,9%

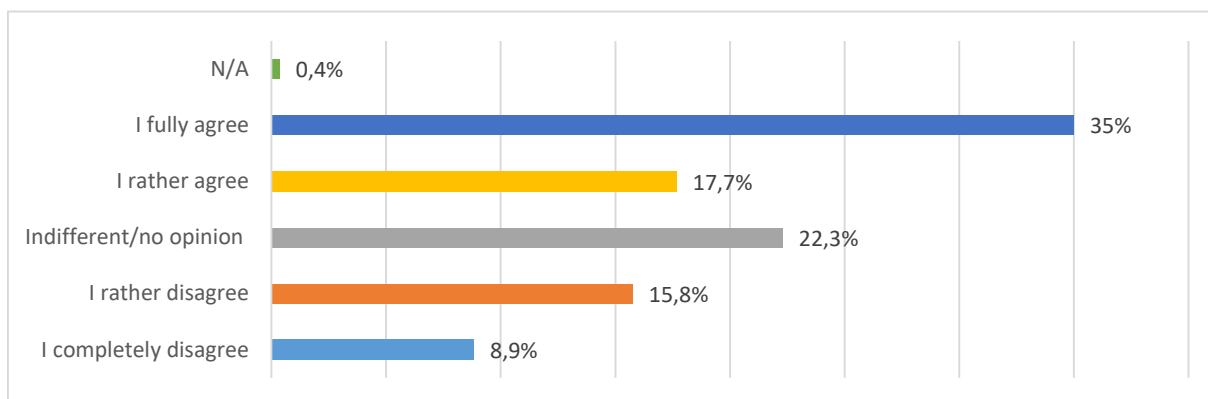
Я скорее не согласен – 1,9%

Я совершенно не согласен – 1,5%

Подавляющее большинство респондентов согласны с утверждением, что об ошибках и недостатках, которые привели или могут привести к нарушению прав человека, демократии и верховенства закона, необходимо сообщать компетентным органам.

- 202 респондента полностью согласны с утверждением;
- 43 респондента скорее согласны с утверждением;
- Только 4 респондента совершенно не согласны и 5 скорее не согласны;
- 5 респондентов нейтральны или у них нет мнения на это утверждение.

40 Использование технологии распознавания лиц следует запретить в общественных местах.



Нет данных – 0,4%

Я полностью согласен – 35%

Я скорее согласен – 17,7%

Нейтрально/нет мнения – 22,3%

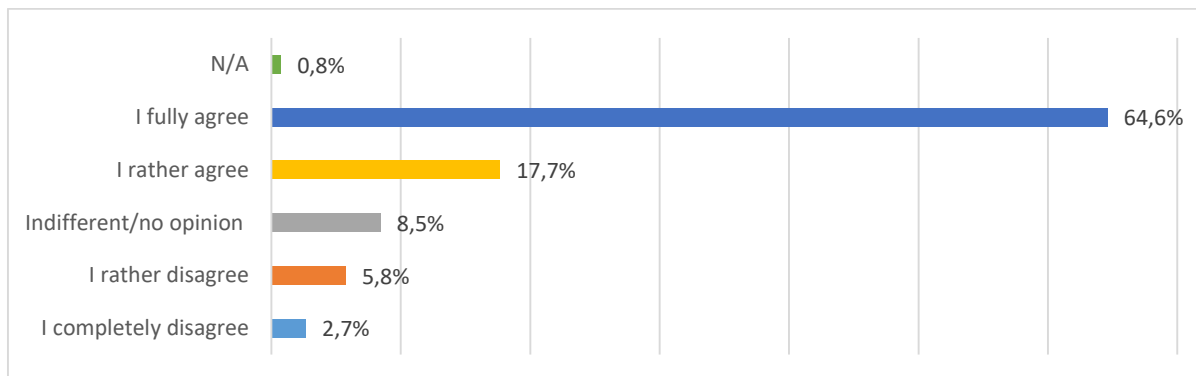
Я скорее не согласен – 15,8%

Я совершенно не согласен – 8,9%

Большинство респондентов полностью согласно с утверждением, что использование технологии распознавания лиц в общественных местах должно быть запрещено.

- 91 респондент полностью согласны с утверждением;
- 46 респондентов скорее согласны с утверждением;
- 23 респондента совершенно не согласны и 41 скорее не согласны;
- 58 респондентов нейтральны или у них нет мнения на это утверждение.

41. Информация, полученная с помощью систем распознавания лиц, всегда должна просматриваться человеком перед использованием в целях, влияющих на индивидуальную свободу, например, в отношении человека, сажающегося в самолет, после ареста полицией или в полицейском участке в рамках судебного разбирательства.



Нет данных – 0,8%

Я полностью согласен – 64,6%

Я скорее согласен – 17,7%

Нейтрально/нет мнения – 8,5%

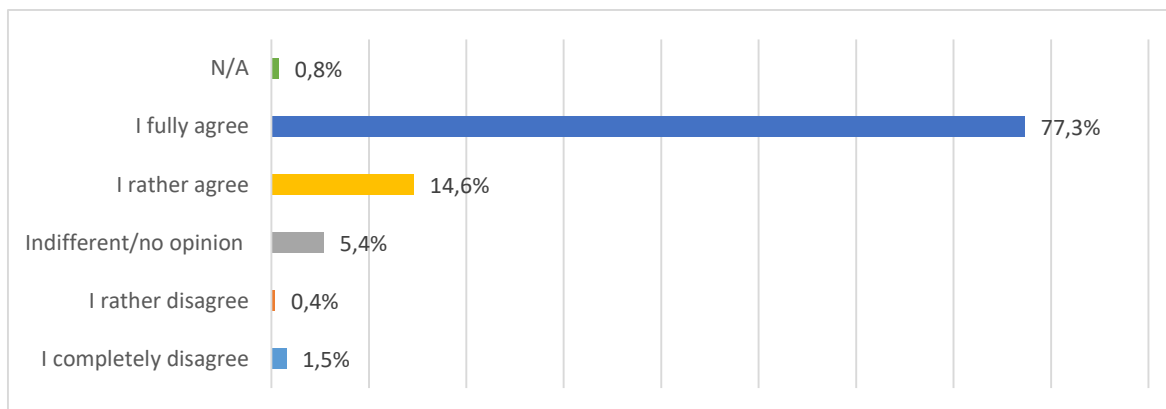
Я скорее не согласен – 5,8%

Я совершенно не согласен – 2,7%

Подавляющее большинство респондентов согласно с утверждением, что информация, полученная с помощью систем распознавания лиц, всегда должна просматриваться человеком перед использованием в целях, влияющих на индивидуальную свободу, например, в отношении человека, сажающегося в самолет, после ареста полицией или в полицейском участке в рамках судебного разбирательства.

- 168 респондентов полностью согласны с утверждением;
- 46 респондентов скорее согласны с утверждением;
- Только 7 респондентов совершенно не согласны и 15 скорее не согласны;
- 22 респондента нейтральны или у них нет мнения на это утверждение.

42. Использование систем ИИ в демократических процессах (например, на выборах) должно строго регулироваться.



Нет данных – 0,8%

Я полностью согласен – 77,3%

Я скорее согласен – 14,6%

Нейтрально/нет мнения – 5,4%

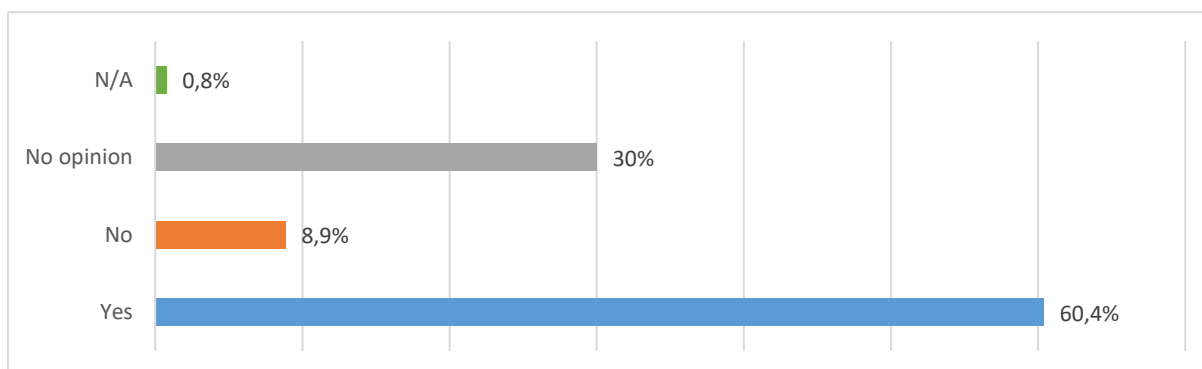
Я скорее не согласен – 0,4%

Я совершенно не согласен – 1,5%

Подавляющее большинство респондентов согласны с утверждением, что использование систем искусственного интеллекта в демократических процессах (например, на выборах) должно строго регулироваться:

- 201 респондент полностью согласны с утверждением;
- 38 респондентов скорее согласны с утверждением;
- Только 4 респондента совершенно не согласны и 1 скорее не согласен;
- 14 респондентов нейтральны или у них нет мнения на это утверждение.

43. Должна ли будущая правовая база на уровне Совета Европы включать особый режим ответственности в отношении приложений ИИ?



Большинство респондентов (60%) согласны с тем, что будущая правовая база на уровне Совета Европы должна включать особый режим ответственности в отношении приложений ИИ.

44. Если да, то какие аспекты следует охватить?

Этот вопрос получил мало комментариев по сравнению с вопросами в первом и втором разделах. Несмотря на то, что данный вопрос был адресован тем, кто ответил «да» на предыдущий вопрос, несколько респондентов выразили сомнения в необходимости создания особого режима ответственности, поскольку использование ИИ носит сквозной характер, а обязанности не являются или не должны быть одинаковыми для разных областей использования. Для некоторых более целесообразно включить ответственность за проблемные виды использования в существующие режимы ответственности.

Для тех, кто считает, что необходим новый режим ответственности, первая задача состоит в том, чтобы определить тех, кто контролирует системы ИИ и кто может быть привлечен к ответственности в случае причинения вреда пострадавшему. Одна из специфических проблем с ответственностью в ИИ, особенно с глубоким обучением, заключается в том, что ущерб может быть вызван алгоритмическим решением без каких-либо дефектов в конструкции алгоритма, в его работе и в используемых базах данных.

Некоторые предлагают объединить безвиновную ответственность (с пересмотренными мерами защиты и законодательными исключениями) и виновную ответственность. Это поможет избежать «пробелов в ответственности». Другие указывают, что для систем, которые поддерживают человеческие решения, именно люди, принимающие решения, должны нести ответственность; а для систем, которые принимают свои собственные решения или действуют на основе собственного анализа и прогнозов, ответственность должны нести люди, которые выбрали и использовали систему. И, наконец, он должен включать организации, выдающие сертификаты соответствия, а также конечных пользователей, которые могут неправильно использовать системы ИИ.

Здесь также возникает вопрос о бремени доказывания. Не только люди должны нести ответственность, но и бремя доказательства должно лежать на разработчиках систем искусственного интеллекта, чтобы продемонстрировать, что их системы не причиняют вреда.

Наконец, по мнению многих респондентов из разных секторов, адекватный режим ответственности должен включать механизмы оспаривания, восстановления в правах, компенсации и даже восстановления. Один респондент из частного сектора предложил создать европейский фонд для быстрой компенсации жертвам.

Раздел 5: Политика и меры в области развития.

В этом разделе вопросы 45-49 посвящены полезности для респондентов механизмов соблюдения (5 вариантов) и последующих действий (4 варианта). Последний вопрос (Q 50) - открытый вопрос, предназначенный для респондентов, желающих привлечь внимание САНАИ к другим вопросам.

Большинство респондентов считает очень полезными следующие механизмы соблюдения: оценка воздействия на права человека, демократию и верховенство закона, аудит и межсекторальный аудит, непрерывный автоматический мониторинг (Q45).

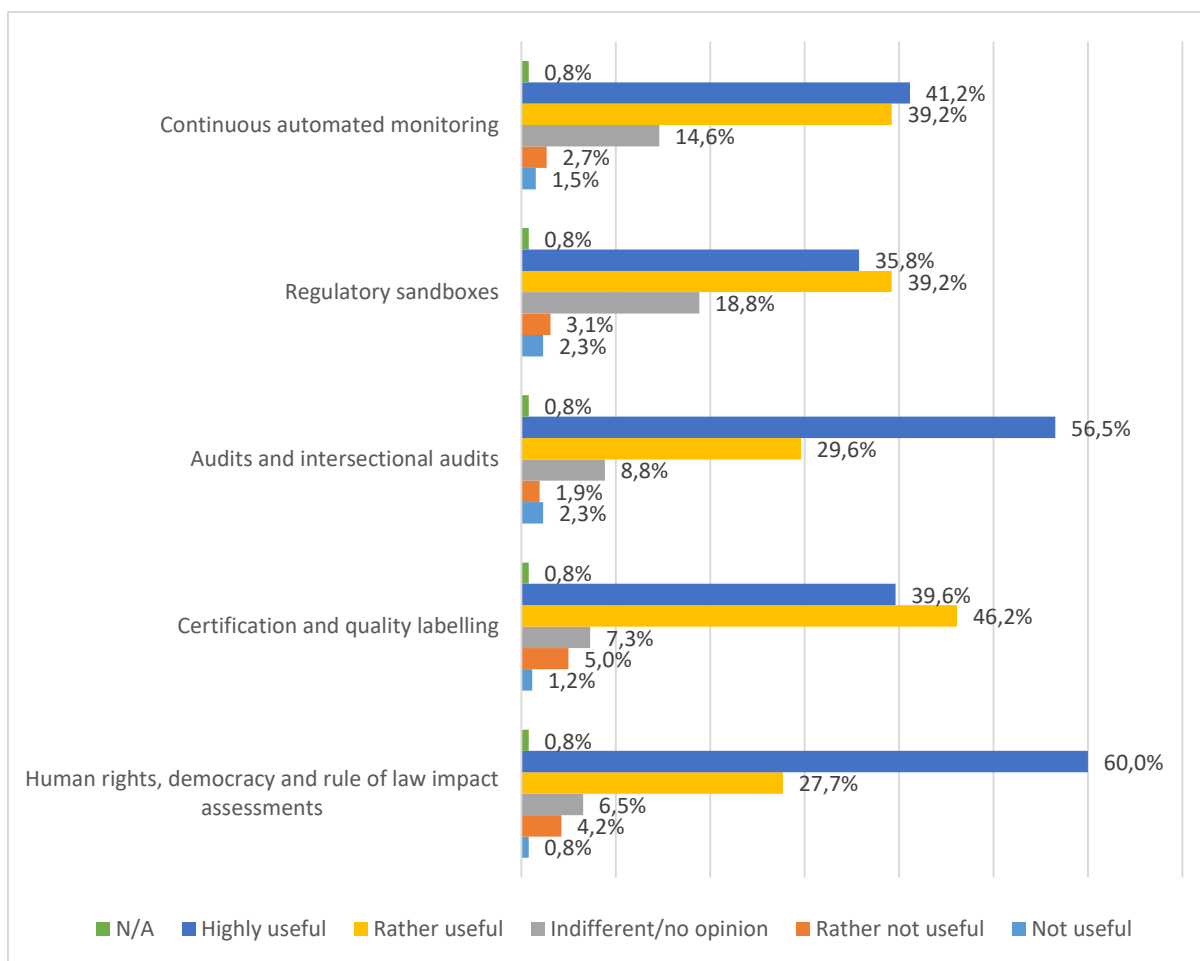
Для большинства респондентов механизмы комплаенс должны быть частью обязательного документа (Q 47). Однако это предпочтение менее очевидно для регуляторных «песочниц», где 40% респондентов считают, что регуляторные «песочницы» должны быть частью обязательного документа, по сравнению с 38% в пользу необязательного документа. Кроме того, респонденты, выбравшие вариант «другое» на этот вопрос, предложили несколько механизмов, которые могли бы дополнить предлагаемый список.

Четыре механизма последующих действий, предложенные в вопросе 48: Мониторинг законодательства и политики в области ИИ в государствах-членах, Наращивание потенциала в отношении инструментов Совета Европы, включая помощь в содействии ратификации и реализации соответствующих инструментов Совета Европы, Обсерватория ИИ для обмена передовым опытом и обмен информацией о юридических, политических и технологических разработках, связанных с системами ИИ, и Создание центра экспертизы по ИИ и правам человека, по мнению большинства респондентов, считаются полезными мероприятиями, которые должны быть реализованы Советом Европы.

На вопрос «Какие еще механизмы, если таковые имеются, следует рассмотреть?» (Q 49), многие респонденты вспомнили предложения, сделанные в предыдущих вопросах. Однако выделяются три предложения: механизм разрешения споров, механизм сотрудничества и участие общественности.

Наконец, ответы на вопрос 50 были очень содержательными и информативными. Респонденты призывают Совет Европы быть ясным, последовательным, педагогическим и решительным в принятии обязательных мер, даже если они могут или должны сосуществовать с мягкими и необязательными правилами.

45. По вашему мнению, насколько полезными будут следующие механизмы соблюдения для предотвращения и снижения рисков для прав человека, демократии и верховенства закона, возникающих в результате проектирования, разработки и применения ИИ?



Непрерывный автоматический мониторинг:

Нет данных – 0,8%

Очень полезны – 41,2%

Довольно полезны – 39,2%

Нейтрально/нет мнения – 14,6%

Скорее не полезны – 2,7%

Не полезны – 1,5%

Регуляторные песочницы:

Нет данных – 0,8%

Очень полезны – 35,8%

Довольно полезны – 39,2%

Нейтрально/нет мнения – 18,8%

Скорее не полезны – 3,1%

Не полезны – 2,3%

Аудиты и межотраслевые аудиты:

Нет данных – 0,8%

Очень полезны – 56,5%

Довольно полезны – 29,6%

Нейтрально/нет мнения – 8,8%

Скорее не полезны – 1,9%

Не полезны – 2,3 %

Сертификация и маркировка качества:

Нет данных – 0,8%

Очень полезны – 39,6%

Довольно полезны – 46,2%

Нейтрально/нет мнения – 7,3%

Скорее не полезны – 5,0%

Не полезны – 1,2%

Оценка влияния на права человека, демократии и верховенства закона:

Нет данных – 0,8%

Очень полезны – 60,0%

Довольно полезны – 27,7%

Нейтрально/нет мнения – 6,5%

Скорее не полезны – 4,2%

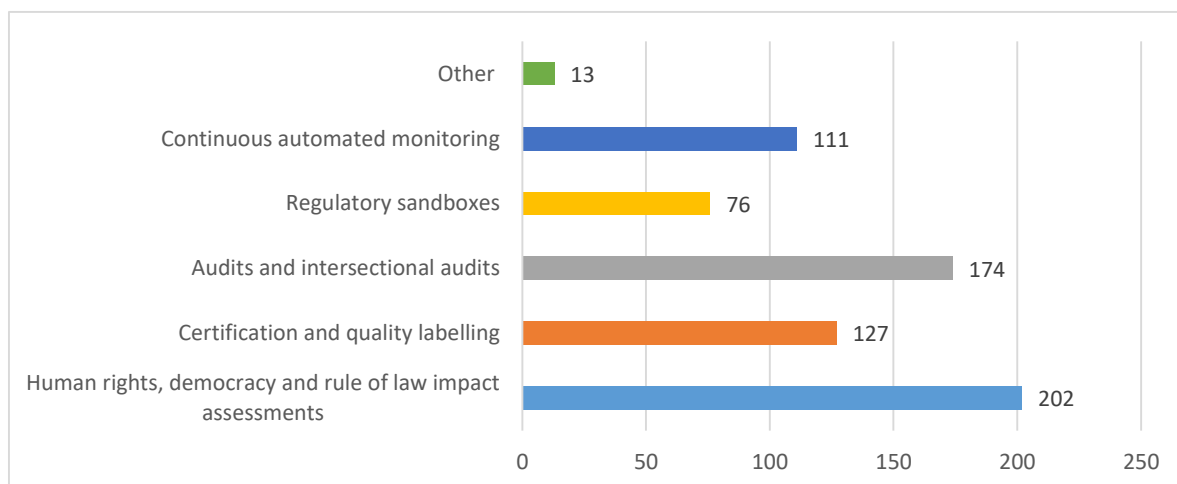
Не полезны – 0,8%

Все следующие предложенные механизмы комплаенс были сочтены большинством респондентов либо очень полезными, либо довольно полезными.

Три механизма комплаенс, которые большинство респондентов считают очень полезными, - это оценка воздействия на права человека, демократию и верховенство закона (60%), аудиты и межсекторальные аудиты (56,5%) и непрерывный автоматический мониторинг (41%).

Сертификация и маркировка качества (46%) и регуляторные «песочницы» (39%) были признаны большинством респондентов весьма полезными.

46. Пожалуйста укажите, какую комбинацию механизмов следует предпочесть для эффективной защиты прав человека, демократии и верховенства закона.



Другие – 13

Непрерывный автоматический мониторинг – 111

Регуляторные песочницы – 76

Аудиты и межотраслевые аудиты – 174

Сертификация и маркировка качества – 127

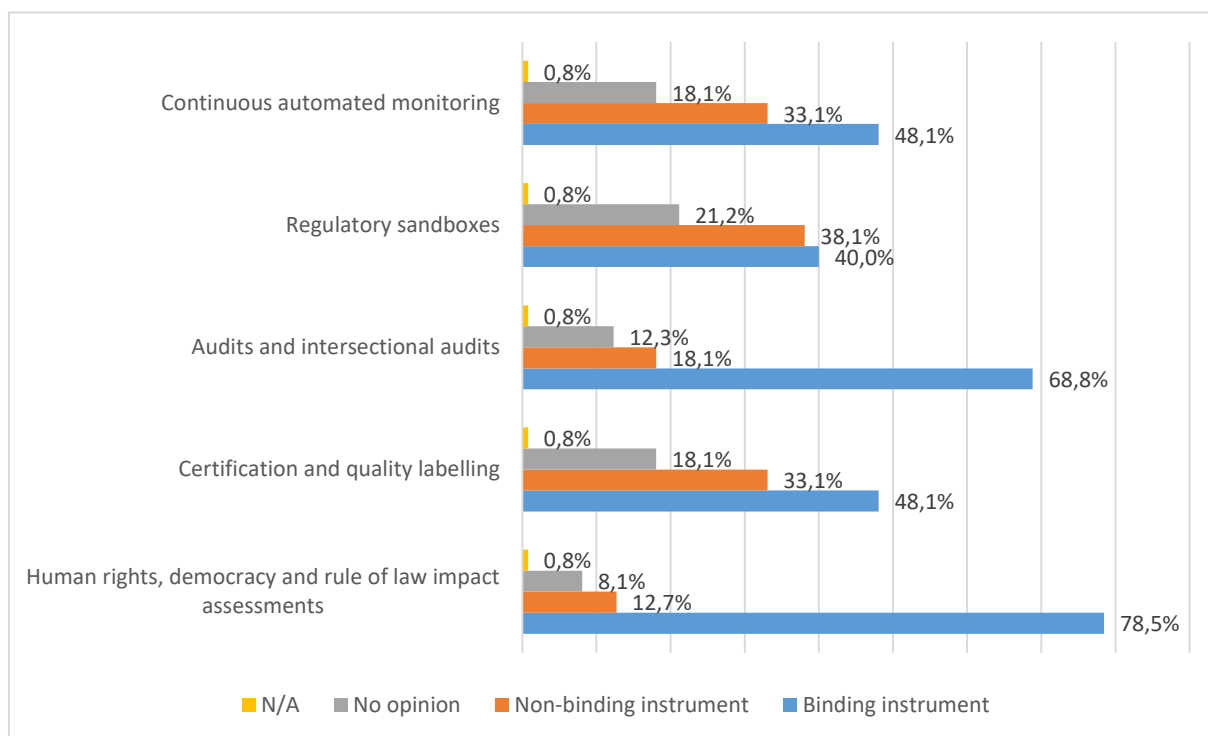
Оценка влияния на права человека, демократии и верховенства закона – 202

Максимум три выбранных ответа на этот вопрос было позволено на одного респондента.

Наиболее избранная комбинация объединяет механизмы оценки влияния на права человека, демократию и верховенство закона (202), аудиты и межотраслевые аудиты (174) и механизмы сертификации и маркировки качества (127).

47. Пожалуйста, выберите, какие механизмы должны быть частью обязательного или необязательного документа для наилучшей защиты прав

человека, демократии и верховенства закона.



Непрерывный автоматический мониторинг:

Нет данных – 0,8%

Нет мнения - 18,1%

Необязательный документ – 33,1%

Обязательный документ – 48,1%

Регуляторные песочницы:

Нет данных – 0,8%

Нет мнения – 21,2%

Необязательный документ – 38,1%

Обязательный документ – 40%

Аудиты и межотраслевые аудиты:

Нет данных – 0,8%

Нет мнения – 12,3%

Необязательный документ – 18,1%

Обязательный документ – 68,8%

Сертификация и маркировка качества:

Нет данных – 0,8%

Нет мнения – 18,1%

Необязательный документ – 33,1%

Обязательный документ – 48,1%

Оценка влияния на права человека, демократии и верховенства закона:

Нет данных – 0,8%

Нет мнения – 8,1%

Необязательный документ – 12,7%

Обязательный документ – 78,5%

Большинство респондентов считают, что оценка воздействия на права человека, демократию и верховенство закона должна быть частью обязательного документа (79%). Это также относится к аудитам и межсекторальным аудитам (69%).

48% респондентов выступают за обязательный документ для сертификации и маркировки качества, а также за непрерывный автоматический мониторинг.

40% респондентов считают, что «регуляторные песочницы» должны быть частью обязательного документа, против 38% в пользу необязательного документа.

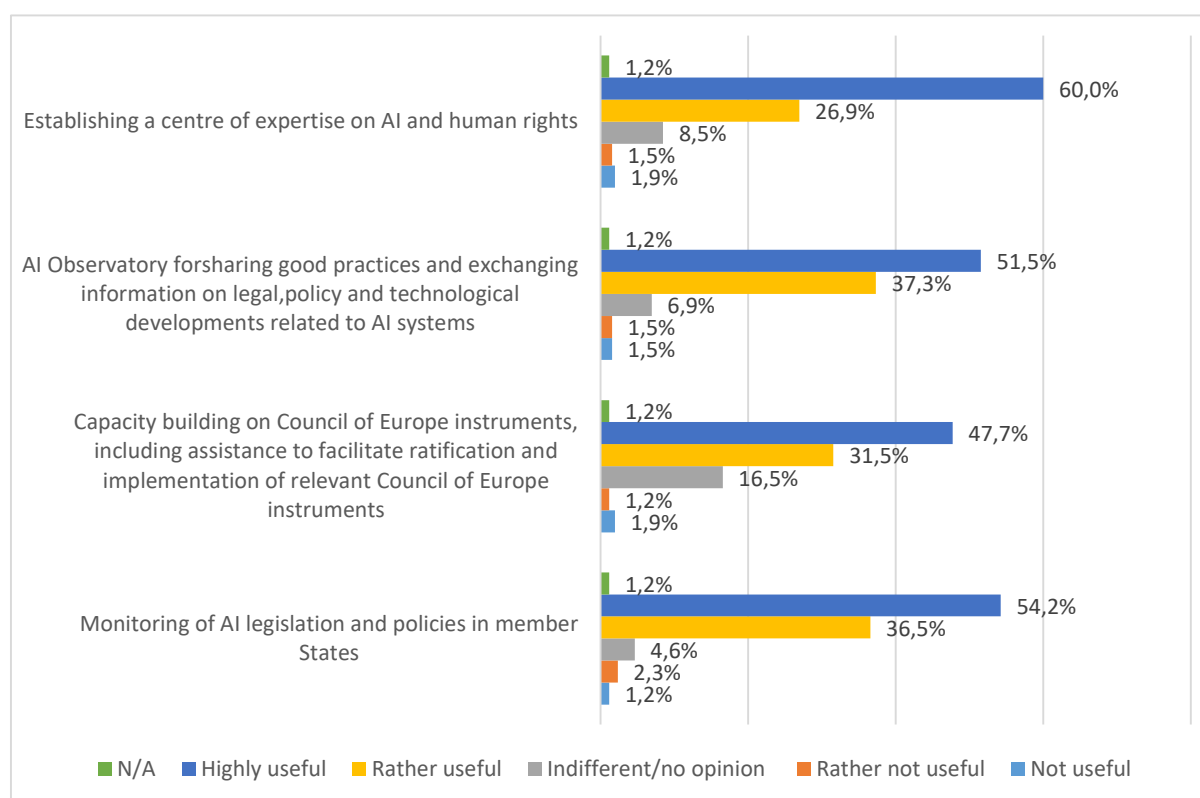
47bis. Другие

По мнению респондентов, предложенный список можно было бы дополнить несколькими механизмами, такими как кодексы этики и кодексы добросовестного поведения, упомянутые в вопросе 21, или механизм участия граждан в разработке систем искусственного интеллекта на этапе проектирования; Сторонники последнего предложения подчеркивают важность включения наиболее уязвимых слоев населения. Это позволит разработчикам оценить социальную значимость системы искусственного интеллекта, которую они проектируют. Таким образом, некоторые респонденты предлагают механизм мониторинга полезности ИИ до принятия решения о замене текущих процедур процессами на основе ИИ.

Другие предлагают технические протоколы, обеспечивающие доступ к методам сбора и обработки данных в стандартизированной форме (машиночитаемые механизмы раскрытия информации, такие как Open Ethics Transparency Protocol).

Но предложения, которые возникают чаще всего, касаются мониторинга алгоритмов. Некоторые отмечают, что сертификация имеет смысл только в том случае, если сертифицированный продукт не меняется, что трудно обеспечить для автономных систем AI. Вместо того, чтобы устанавливать сертификаты, которые не являются устойчивыми, более полезно создать механизмы отчетности с регулярными отчетами об использовании и влиянии автоматизированных систем на права человека. Мониторинг систем ИИ также подразумевает создание (и инвестирование) независимую экосистему аудита ИИ.

48. По вашему мнению, насколько полезными были бы следующие мероприятия во исполнение принятых решений, если бы они были реализованы Советом Европы?



Установление центра экспертиз по ИИ и права человека:

Нет данных – 1,2%

Очень полезны – 60,0%

Довольно полезны – 26,9%

Нейтрально/нет мнения – 8,5%

Скорее не полезны – 1,5%

Не полезны – 1,9%

Обсерватория искусственного интеллекта для обмена передовым опытом и информацией о юридических, политических и технологических разработках, связанных с системами искусственного интеллекта:

Нет данных – 1,2%

Очень полезны – 51,5%

Довольно полезны – 37,3%

Нейтрально/нет мнения – 6,9%

Скорее не полезны – 1,5%

Не полезны – 1,5%

Наращивание потенциала по инструментам Совета Европы, включая содействие в ратификации и применении соответствующих инструментов Совета Европы:

Нет данных – 1,2%

Очень полезны – 47,7%

Довольно полезны – 31,5%

Нейтрально/нет мнения – 16,5%

Скорее не полезны – 1,2%

Не полезны – 1,9%

Мониторинг законодательства и политики в области ИИ в государствах-членах:

Нет данных – 1,2%

Очень полезны – 54,2%

Довольно полезны – 36,5%

Нейтрально/нет мнения – 4,6%

Скорее не полезны – 2,3%

Не полезны – 1,2%

90% респондентов считают полезным мониторинг законодательства и политики в области ИИ в государствах-членах. Только 3 респондента считают его бесполезным, а 12 не имеют мнения по данному вопросу.

80% респондентов считают создание потенциала по инструментам Совета Европы, включая содействие в ратификации и реализации соответствующих инструментов Совета

Европы, полезным. Только 5 респондентов считают его бесполезным, а 43 не имеют мнения по данному вопросу.

89% респондентов считают, что Обсерватория ИИ для обмена передовым опытом и обмена информацией о юридических, политических и технологических разработках, связанных с системами ИИ, является полезной. Только 4 респондента считают его бесполезным, а 18 не имеют мнения по данному вопросу.

87% респондентов считают создание экспертного центра по ИИ и правам человека полезным. Только 5 респондентов считают его бесполезным, а 22 не имеют мнения по данному вопросу.

Вкратце, результат демонстрирует, что четыре предложенных мероприятия во исполнение принятых решений рассматриваются как полезные мероприятия, которые должны быть реализованы Советом Европы по мнению большинства респондентов.

49. Какие еще механизмы, если таковые имеются, следует рассмотреть?

Многие ответы повторяют предложения, сделанные в предыдущих. Но выделяются три предложения:

1. Механизм разрешения споров

Правовой инструмент для значимого регулирования ИИ не может быть полным, если он содержит только принципы и правила и не содержит механизмов обеспечения прав в случае возникновения споров при использовании ИИ. Поэтому Совет Европы должен разработать механизм разрешения споров на европейском и международном уровне. Управлять им будут сотрудники, обладающие опытом в области искусственного интеллекта и права. Поэтому Совет Европы должен способствовать созданию специальных судов или центров посредничества и арбитража, которые существуют в других областях.

2. Механизм сотрудничества

Несколько респондентов из всех секторов призывают к сотрудничеству между европейскими и международными организациями, чтобы избежать фрагментации нормативных требований и улучшить правовой надзор и мониторинг систем искусственного интеллекта. Некоторые хотят, чтобы Совет Европы с ОЭСР и другими организациями работал вместе над созданием глобальной политики в области ИИ, задействовал общие ресурсы и объединял различный опыт, накопленный этими организациями. Другие заинтересованные стороны выступают за формат с участием многих заинтересованных сторон, который включает компании, чтобы иметь возможность

адаптировать нормативно-правовые инструменты к развивающейся реальности технологий искусственного интеллекта.

Хотя следующее предложение было сделано в ответ на следующий вопрос, оно хорошо дополняет эту рекомендацию, и поэтому мы позволяем себе упомянуть его здесь. Один респондент приводит конкретный пример сотрудничества, который может вдохновить САНАИ на работу. В нем упоминается проект партнерства третьего поколения (3GPP) как механизм управления с участием многих заинтересованных сторон в отрасли информационных и коммуникационных технологий (ИКТ).

3. Участие общественности

Многие респонденты рекомендуют не только сотрудничество между международными организациями, но и совместное участие всех заинтересованных сторон, включая общественность, граждан и особенно группы, подверженные негативному воздействию ИИ. Цель состоит в том, чтобы привлечь и расширить возможности гражданского общества в контексте массового использования систем искусственного интеллекта. Поэтому они предлагают создать платформу, которая будет способствовать обмену передовым опытом, выявлению тенденций в развитии ИИ и предупреждению этических и юридических проблем. Эта платформа также обеспечит участие общественности, особенно групп, недопредставленных в государственных учреждениях и в разработке политики в области ИИ. Это позволит получать отзывы об использовании ИИ и сообщать о серьезных проблемах. Это предложение близко к предложению обсерватории ИИ, но отличается от него.

50. Есть ли какие-либо другие вопросы в отношении проектирования, разработки и применения систем искусственного интеллекта в контексте прав человека, демократии и верховенства закона, на которые вы хотите обратить внимание САНАИ?

Ответы на этот вопрос были очень содержательными. Мы не можем уделить им достаточно внимания в этом резюме. По сути, респонденты призывают Совет Европы быть прозрачным, последовательным, педагогическим и решительным в принятии обязательных мер, даже если они могут или должны сосуществовать с мягкими и необязательными правилами.

- Прозрачный :

Участники призывают к однозначной и хорошо объясненной формулировке, в частности, чтобы непрофессионалы, организации гражданского общества, которые не являются передовыми заинтересованными сторонами в области ИИ, и даже профессионалы,

использующие ИИ, понимали правовую базу и могли применять ее и ссылаться на нее для жалоб.

- **Последовательный:**

Совету Европы также необходимо создать свою правовую базу согласованным образом, то есть принять во внимание другие существующие правовые и политические рамки. Таким образом, любая потенциальная конвенция по ИИ должна учитывать взаимодействие со структурами защиты данных, разработанными в рамках Совета Европы, а также с Законом ЕС о цифровых услугах, GDPR, Европейской структурой Положения о медицинских устройствах и другими постановлениями Европейской комиссии по ИИ. Без целостного подхода к регулированию ИИ существует серьезный риск регуляторной фрагментации, о чем часто говорилось в анкете, но также и риск отсутствия доверия и привлечения клиентов к продуктам и услугам ИИ, разработанным европейской индустрией ИИ.

- **Педагогический:**

Многие комментарии касались необходимости повышения осведомленности об ИИ среди непрофессионалов и государственных служащих, а также о повышении этической осведомленности среди разработчиков, разработчиков ИИ. Для некоторых респондентов этическое понимание должно быть обязательной частью обучения профессионалов в области искусственного интеллекта, что характерно для регулируемых профессий, связанных со значительной степенью риска для общества.

- **Напористый:**

Хотя этические дискурсы полезны и необходимы для руководства, многие респонденты призывают Совет Европы защищать права человека и демократию посредством обязательных правовых положений. Это очень ясно из общих результатов анкеты. Сфера правосудия стала наиболее рискованной областью для защиты прав людей, и необходимость регулирования ИИ в этой области кажется более насущной. Некоторые респонденты, представляющие юридические профессии, представили конкретные протоколы, регулирующие использование ИИ в судебной сфере. Например, чтобы снизить потенциальные риски и влияние инструментов искусственного интеллекта в судебных системах, необходимо реализовать следующие принципы:

- **Возможность идентифицировать использование ИИ (принцип идентификации):** все стороны, участвующие в судебном процессе, всегда должны иметь возможность идентифицировать, до и в рамках судебного решения, элементы, возникающие в результате внедрения инструмента ИИ.
- **Неделегирование полномочий судьи по принятию решений (принцип недопустимости делегирования):** ни при каких обстоятельствах судья не должен

делегировать все или часть своих полномочий по принятию решений инструменту искусственного интеллекта. В любом случае право на человека-судью должно быть гарантировано на любой стадии судебного разбирательства.

- Возможность для сторон проверить ввод данных и аргументацию инструмента AI (Принцип прозрачности).
- Возможность для сторон обсуждать и оспаривать результаты ИИ (принцип обсуждения) в состязательной манере вне фазы обсуждения и в разумные сроки.
- Нейтральность и объективность инструментов искусственного интеллекта (принцип нейтральности), используемых судебной системой, должны быть гарантированы и проверены.

Давайте подведем итог по парадоксу, который проявляется в нескольких статьях и остался незамеченным респондентами: ИИ следует использовать для разработки систем ИИ, направленных на защиту прав человека.

Это указывает на важность для Совета Европы разработки ответных политических мер и любых других подходящих мер, которые могут поддержать разработку государственными и частными организациями систем искусственного интеллекта в соответствии с требованиями прав человека, верховенства закона и демократии.

ЗАКЛЮЧЕНИЕ

Целью консультаций являлось получение обратной связи от широкого круга заинтересованных сторон - представителей различных секторов (правительство и государственное управление, гражданское общество, частный сектор и научные круги) по основным элементам новой правовой базы для проектирования, разработки и применения ИИ, основанной на стандартах Совета Европы в области прав человека, демократии и верховенства закона. Эта обратная связь, в свою очередь, может помочь измерить уровень поддержки различных вариантов регулирования, рассматриваемых Рабочей группой по созданию правовой базы (CAHAI-LFG), и, следовательно, помочь этой Рабочей группе сформировать свои нормативные предложения.

Многие респонденты, внося важные и актуальные идеи в вопросник, жаловались на жесткость формата консультаций либо потому, что диапазон вариантов был слишком ограничен, либо потому, что открытые вопросы позволяли давать только короткие ответы.

Однако следует иметь в виду, что ограничение количества вариантов требует тщательной расстановки приоритетов и дает CAHAI-LFG наиболее полезную информацию для выявления тенденций и точного определения областей где возможен консенсус, вопросов для обсуждения и неразрешимых вопросов. В этом отношении, благодаря заинтересованности респондентов, консультации прошли успешно.

Большинство сконцентрировались на нескольких пунктах, начиная с типа определения ИИ, которое следует использовать в правовой базе. Для 48% участников технологически нейтральное и упрощенное определение ИИ предпочтительнее узких определений или отсутствия определения вообще. Определение, данное в технико-экономическом обосновании, кажется подходящим определением для создания правовой базы. Кроме того, сложился консенсус в отношении областей, в которых ИИ представляет наибольший риск нарушения прав человека, демократии и верховенства закона. Четко выделяются две области: правосудие и правовое принуждение, за которыми следует национальная безопасность. Остальные области менее четко обозначены. Очевидно, что, хотя возможная будущая правовая база, вероятно, будет носить сквозной характер и, следовательно, будет применяться к использованию ИИ в различных областях и секторах, полученные отзывы указывают на то, что возможная будущая правовая база должна решать как общие, так и конкретные вопросы, связанные с широким использованием систем ИИ в этих областях.

В наличии доказательств нарушения прав человека, демократии и верховенства закона системой ИИ большинство (55%) поддерживает ее запрет. Однако ответ государственного сектора предусматривает другой вариант: вместо немедленного запрета систем ИИ следует принять поэтапный подход, увеличивая давление на разработчиков с целью

исправления дефектных систем ИИ (код, модель или база данных) и увеличивая штрафы. Этот поэтапный подход не применяется к системам ИИ, цель которых по замыслу или в ходе широкого использования противоречит правам человека, например, социальный скоринг, который должен быть запрещен. Следует отметить, что вариант «моратория» не был предусмотрен, и, основываясь на результатах следующего вопроса о системах ИИ с высоким риском нарушения прав человека с высокой вероятностью, мы можем обоснованно сделать вывод, что вариант строгого регулирования, возможно, включая этот вариант, мог быть поддержан большинством выбравших ответ «другое».

Конечно, остается вопрос, какие критерии следует использовать для определения того, заложено ли изначально в системе ИИ нарушение прав человека. Это вызовет серьезные дискуссии, о чем свидетельствуют комментарии о регулировании распознавания лиц в общественных местах. По этому поводу нет единого мнения. Распознавание лиц - это проблема, которая глубоко разделяет заинтересованные стороны.

Позиции иногда очень четкие, без возможности сближения, что свидетельствует о важности поиска благодатной почвы для обсуждения, чтобы разработать предложение, которое без учета самых отдаленных позиций будет поддержано твердым большинством. Действительно, ответы на вопросы выявили очень широкий консенсус в отношении необходимости обязательного регулирования технологий, которые представляют высокий риск нарушений прав человека, независимо от уровня вероятности.

Из ответов на вопросы следует, что ожидается, что государственные органы и государства будут осуществлять обязательное регулирование и что саморегулирование или мягкое регулирование с помощью этических руководств является недостаточным, когда на карту поставлены права человека, демократия и верховенство закона.

Результаты консультаций с участием многих заинтересованных сторон ясно указывают на то, что юридически обязательный инструмент, регулирующий использование ИИ, считается необходимым для обеспечения эффективной защиты прав человека, верховенства закона и демократии в свете конкретных вопросов, вызванных ИИ и которые нельзя решить с помощью существующих обязательных норм. Этот вариант, как указано в технико-экономическом обосновании, должен получить поддержку Комитета министров. Основываясь на ответах на вопросы, можно констатировать, что заинтересованные стороны из всех секторов, включая государственные организации, в подавляющем большинстве поддерживают этот вариант (74%).

Возможная новая правовая база для проектирования, разработки и применения ИИ, основанная на принципах Совета Европы, должна, согласно многосторонним консультациям, уделять особое внимание следующим принципам и правам: уважение человеческого достоинства, неприкосновенность частной жизни и защита данных, а также

недискриминация. Возможность оспорить решение, принятое системой ИИ, и доступ к эффективным средствам правовой защиты, а также обеспечение прозрачности приложений ИИ также рассматриваются большинством респондентов как важные аспекты, требующие решения.

Консультации также позволили зафиксировать высокий уровень консенсуса в отношении конкретных регуляторных выборов, которые необходимо сделать, таких как, например, обязанность всегда информировать людей, когда они взаимодействуют с системой ИИ при любых обстоятельствах, когда принимается решение, которое затрагивает их лично. системой ИИ, и когда система ИИ используется в процессе принятия решений, которые затрагивают их лично. Сильная поддержка также зафиксирована в отношении права на человеческий анализ алгоритмических решений, особенно в деликатных контекстах, таких как правосудие.

Респонденты постоянно сходятся во мнении о механизмах надзора, которые необходимо создать. По мнению 85% респондентов, государства должны создать механизмы общественного надзора за системами ИИ, которые могут нарушать юридически обязательные нормы в сфере прав человека, демократии и верховенства закона, а 94% респондентов поддерживают механизм отчетности: об ошибках и недостатках, обнаруженных в системах ИИ, которые привели или могут привести к нарушению прав человека, демократии и верховенства закона, необходимо сообщать компетентным органам, должны быть созданы надзорные механизмы. Кроме того, консультации подтвердили полезность механизмов комплаенс для предотвращения и снижения рисков для прав человека, верховенства закона и демократии. По мнению подавляющего большинства респондентов, некоторые из этих механизмов, такие как оценка воздействия на права человека, демократию и верховенство закона, а также аудиты и межсекторальные аудиты, должны стать частью будущего юридически обязательного инструмента, в то время как другие могут быть лучше развиты с помощью необязательного регулирования.