На правах рукописи

**Вантяева Анастасия Сергеевна**

**СОЦИАЛЬНЫЕ РИСКИ ВНЕДРЕНИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

5.4.7 – Социология управления

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата социологических наук

Москва 2022

Диссертация «Социальные риски внедрения технологий искусственного интеллекта» выполнена на кафедре Социология знания факультета ВШССН МГУ имени М.В. Ломоносова.

**Научный руководитель** – Доктор экономических наук, Гребенюк Александр Александрович, заместитель директора по научной работе факультета Высшей школы современных социальных наук Московского Государственного университета имени М.В. Ломоносова.

**Официальные рецензенты:**

Соболева Юлия Юрьевна - Кандидат социологических наук, доцент кафедры социологии знания ВШССН МГУ имени М.В. Ломоносова.

Леденёва Виктория Юрьевна - Доктор социологических наук, профессор кафедры социологии управления РАНХиГС при Президенте РФ.

**ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

**Актуальность**. Развитие человеческой цивилизации претерпело ряд коренных изменений в ходе научно-технических революций начиная со второй половины XVIII века – с момента появления парового двигателя. Создание технического оборудования на основе использования воды и пара значительно увеличило производительность и сократило использование тяжелого физического труда. Появление электричества во второй половине XIX века дало возможность наладить поточное производство и создать систему разделения труда. Создание вычислительной техники в середине XX века помогло автоматизировать производство, а последующее развитие новых технологий повлекло еще более глубинные изменения.

Конец XX - начало XXI века характеризуется увеличением масштабов применения новых технологий во всех сферах жизнедеятельности человека. Происходят глубокие качественные изменения в сфере производства, науки, на рынке труда. В это время наступает следующая эпоха глобальных перемен, которая приходится на начало второго десятилетия XXI века, когда ученые начали говорить о четвертой научно-технической революции, для которой в 2011 году на «Ганноверской ярмарке» - крупнейшей промышленной выставке, был предложен термин «Индустрия 4.0». Новые технологии не стоят на месте и превращают наш мир в связанную смарт-среду. Происходит все более активное внедрение технологий искусственного интеллекта во все сферы жизнедеятельности современного общества. По данным консалтинговой компании Gartner, объем мирового рынка программного обеспечения, использующего алгоритмы ИИ, в 2021 составил порядка $51,5 млрд, увеличившись на 21,3% в сравнении с 2020 годом. Одним из наиболее развитых направлений и сегментов на рынке являются умные виртуальные помощники, расходы на них в 2021 году составили около $6,21 млрд, что на 12% больше, чем годом ранее.[[1]](#footnote-1) В 2022 году более высокие темпы роста ожидаются в развитии ПО для беспилотного транспорта и создания цифровых рабочих мест.

Новые технологии действительно открывают большое количество возможностей, но в то же время таят в себе потенциальные риски и опасности. Основными позитивными последствиями является облегчение труда благодаря процессу автоматизации, увеличение количества материальных благ на душу населения и т.п. Но пока что рано утверждать об альтернативах духовного развития общества, равенства, социальной справедливости, правильного распределения экономических благ и т.п. Во многом перспективы зависят от самих людей, и от той политической позиции и обстановки, в которой произойдёт следующий скачок в развитии технологий искусственного интеллекта.

Российское общество и сферы его жизнедеятельности также претерпели изменения в связи с появлением новых технологий, поэтому отечественные ученые попытались оценить социальные риски и последствия, к которым приводят инновации, а также проанализировать деятельность государства в вопросе поддержки развития технологий и минимизации рисков и опасностей, к которым они могут привести.

**Объект исследования –** социальные риски внедрения технологий искусственного интеллекта в жизнь современного общества.

**Предмет исследования –** государственная политика в сфере минимизации социальных рисков внедрения технологий искусственного интеллекта.

**Цель исследования –** выявление социальных рисков внедрения технологий искусственного интеллекта в различные сферы современного общества и способов их регулирования посредством государственного управления.

**Задачи исследования:**

1. Изучить сущность понятия «риск», выделить его основные виды и факторы возникновения;
2. Проанализировать существующие подходы к изучению «риска» в социальном знании;
3. Описать основные научно – технические предпосылки появления технологий искусственного интеллекта;
4. Проанализировать особенности внедрения технологий искусственного интеллекта в жизнь современного российского общества;
5. Произвести анализ государственной политики РФ по развитию и внедрению технологий искусственного интеллекта;
6. Выделить основные риски внедрения технологий искусственного интеллекта и способы их регулирования на основе социологического исследования;
7. Сформировать рекомендации по оптимизации государственной политики по вопросу развития и внедрения технологий искусственного интеллекта.

**Методы исследования:** метод теоретического анализа, метод анализа вторичных данных, анализ документов, экспертный опрос, анкетный опрос, корреляционный анализ, сравнительный метод.

**Степень разработанности темы исследования**: Несмотря на то, что первые попытки изучения рисков были осуществлены вначале XX веке, более детальное изучение проблематике началось лишь с 70-х гг. Тема риска изучалась в основном в английской, французской и немецкой социологии. В этих странах риск анализировался на разных уровнях, в различных социальных группах и сферах деятельности.

Среди западных ученых, анализом рисков занимались – Э. Дюркгейм[[2]](#footnote-2), Э. Гидденс[[3]](#footnote-3), Н.Луман[[4]](#footnote-4), М. Дуглас[[5]](#footnote-5), М. Томпсон[[6]](#footnote-6), Г. Маркузе[[7]](#footnote-7), Дж. Фон Нейман[[8]](#footnote-8), Ф. Найт[[9]](#footnote-9), Дж. М Кейнс[[10]](#footnote-10), Ф. Буркардт[[11]](#footnote-11) и другие. Особый интерес вызывает концепция У. Бека «Общество риска»[[12]](#footnote-12), в рамках которой рассматривается состояние современного общества в условиях повышенной рискогенности в связи с наличием глобальных неопределенностей.

Со временем, отечественные ученые также начали изучать природу появления рисков, их влияние на различные сферы жизни. Среди наиболее известных авторов – Кравченко С.А.[[13]](#footnote-13), Смакотина Н.Л.[[14]](#footnote-14), Яницкий О.Н.[[15]](#footnote-15), Быков А.А[[16]](#footnote-16), Богоявленский С.Б[[17]](#footnote-17), Альгин А.П.[[18]](#footnote-18), Овсянников А.А.[[19]](#footnote-19), Мозговая А.В.[[20]](#footnote-20), Зубков В.[[21]](#footnote-21), Феофанов К.А.[[22]](#footnote-22), Богомолова С.Н.[[23]](#footnote-23), Матвеенко Ю.И.[[24]](#footnote-24) и другие.

Второй, анализируемой нами проблематикой, является появление и развитие технологий искусственного интеллекта и умных машин. Анализом данных процессов занимались С. Рассел.[[25]](#footnote-25), Д.А. Форсайт[[26]](#footnote-26), Дж. Н Нильсон[[27]](#footnote-27), Э. Хант[[28]](#footnote-28), Дж. Вейцбаум[[29]](#footnote-29), П.М. Черчленд[[30]](#footnote-30), А. Тьюринг[[31]](#footnote-31), А.М. Неймарк[[32]](#footnote-32).

Вопросами изменения состояния общества в связи с внедрением новых технологий занимались Белл Д.[[33]](#footnote-33), Кастельс М.[[34]](#footnote-34), Тоффлер Э.[[35]](#footnote-35), Фрике В[[36]](#footnote-36)., Хайдеггер М.[[37]](#footnote-37)

Среди отечественных ученых изучением новых технологий и процесса внедрения искусственного интеллекта в повседневную жизнь общества занимались Алексеев А.П.[[38]](#footnote-38), Макаров В.Л.[[39]](#footnote-39), Колесникова Г.И.[[40]](#footnote-40), Орлова И.Б.[[41]](#footnote-41), Осипов Г.С.[[42]](#footnote-42), Резаев А.В.[[43]](#footnote-43) и др.

**Практическая значимость**: основные результаты работы могут быть использованы при разработке федеральных и региональных программ по развитию и внедрению технологий искусственного интеллекта в жизнь современного российского общества.

**Научная новизна**: уточнен перечень социальных рисков внедрения технологий искусственного интеллекта в различные сферы общественной жизни, предложены механизмы минимизации данных рисков и сформированы рекомендации по оптимизации государственной политики в сфере развития технологий искусственного интеллекта.

**Положения, выносимые на защиту:**

1. По результатам количественного социологического исследования был сформирован перечень общественных сфер, где технологии искусственного интеллекта применяются наиболее активно и успешно: информационные технологии, маркетинг и реклама, медицина и здравоохранение, финансы и банковское дело, сельское хозяйство.
2. По результатам качественного и количественного социологических исследований определен перечень наиболее рискогенных сфер общественной жизни, в которых внедряются технологии искусственного интеллекта. Перечень включает в себя такие сферы как медицина, судебное дело, государственное управление, транспорт и военная сфера.
3. На основе результатов количественного социологического исследования были определены основные меры по минимизации рисков внедрения технологий искусственного интеллекта, среди которых: совершенствование и всестороннее тестирование систем перед их внедрением, контроль процесса внедрения технологий, получение обратной связи от пользователей, просветительская деятельность и обучение граждан, а также разработка актуальной нормативно-правовой базы для регулирования технологий.
4. По результатам экспертного опроса, в котором приняло участие 10 специалистов из ведущих отечественных и зарубежных компаний, занимающихся развитием и внедрением технологий искусственного интеллекта, было выявлено, что финансирование, оказываемое государством для поддержки развития и внедрения технологий искусственного интеллекта, является недостаточным.
5. По результатам количественного и качественного социологических исследований был определен перечень позитивных и негативных эффектов внедрения технологий искусственного интеллекта. К позитивным относятся: автоматизация и оптимизация производственных процессов, увеличение производительности труда, минимизация ошибок, вызванных человеческим фактором, улучшение качества выпускаемой продукции, ускорение развития социальных процессов. К негативным относятся: излишний формализм и усложнение процедур взаимодействия между субъектами управления; ошибки систем искусственного интеллекта, влекущие финансовые убытки; риски для жизни людей, вызванные ошибками искусственного интеллекта, сокращение рабочих мест, злоупотребление и манипуляция данными; избыточный контроль; снижение уровня грамотности населения.
6. На основе результатов, полученных в ходе исследования, были сформированы рекомендации по оптимизации государственной политики в сфере развития искусственного интеллекта, включающие: публичное освещение позитивных эффектов применения искусственного интеллекта; создание общеобразовательных курсов по искусственному интеллекту и курсов переподготовки для специалистов, которые в этом нуждаются; совершенствование образовательных программ для бакалавров и магистров по искусственному интеллекту; развитие инфраструктуры для тестирования технологий искусственного интеллекта; привлечение на работу специалистов ведущих зарубежных компаний для обмена опытом; увеличение объема финансирования сферы разработки технологий искусственного интеллекта; развитие венчурного рынка для ускорения развития технологий искусственного интеллекта; привлечение высококвалифицированных специалистов сферы искусственного интеллекта к решению управленческих задач государственного уровня.

**ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

Во введении обосновывается актуальность выбранной темы диссертационного исследования, определяются степень разработанности темы, объект, предмет, цель, задачи и методы исследования, научная новизна исследования, практическая значимость, основные положения, выносимые на защиту, описывается структура работы.

В **первой главе** автором описаны теоретико-методологические аспекты изучения социальных рисков. Проанализирована сущность понятия, основные виды рисков, причины и факторы их появления. Кроме того, автором описаны различные подходы к изучению рисков в социальном знании, среди них – социологический, экономический, психологический, политический и управленческий подходы.

В рамкам экономического подхода автором были проанализированы теории А. Смита, Ф. Найта, Дж. М. Кейнса, М. Алле. В рамках психологического - концепции С.Н. Богомоловой, М. Цукермана, В. А. Петровского, Т.В. Корниловой, Дж. Стойнера, Ю. Козельского, Дж. Уальда, Ф. Буркардта. В рамках политического подхода описаны идеи О’Лири и В. Коплина, В.В. Вялых и О. Моторина, Ю.И. Матвеенко и Г. Кнудсена. В рамках управленческого подхода были рассмотрены теории Г.Ю. Кулика, И.А. Бронникова, Н.В. Блинова. Большая часть внимания была уделена социологическому подходу, а именно теориям М. Вебера, А. Шюца, С.А. Кравченко, Э. Гидденса, Н. Лумана, У. Бека, М. Дуглас, О.Н. Яницкого, В.И. Игнатьева и Е.Е. Хабировой.

Во **второй главе** диссертационной работы автором описываются социальные риски внедрения технологий искусственного интеллекта в различные сферы общества. Первым делом, описываются научно – технические предпосылки появления технологий искусственного интеллекта, а также даются определения данному понятию, среди которых «направление научных исследований, целью которых является «интеллектуализация» вычислительных машин, моделирование процессов познания и мышления и т.п.», «различные устройства, системы, программы, названные по тем или иным критериям, «интеллектуальными»» и др. Однако, после проведенного автором анализа, было выявлено, что на сегодняшний день под искусственным интеллектом правильнее понимать именно машинное обучение. Кроме того, в рамках второй главы были рассмотрены различные практики внедрения технологий искусственного интеллекта в различные общественные сферы: образование, промышленность, транспортная сфера, здравоохранение, информационная сфера, а также государственный сектор. Помимо этого, в рамках второй главы описываются риски внедрения технологий ИИ на основе различных зарубежных и отечественных исследований.

В **третьей главе** работы автором описана государственная политика РФ по развитию и внедрению технологий ИИ на основе существующих стратегий, концепций, стандартов и законодательных актов. Раскрыты результаты проведенных качественного и количественного исследований, которые показывают отношение граждан к технологиям, наиболее рискогенные сферы общественной жизни с точки зрения внедрения технологий ИИ, позитивные и негативные эффекты внедрения технологий ИИ, возможные меры минимизации рисков, степень участия государства в развитии сферы. Кроме того, на основе полученных в ходе исследования результатов, автором были сформированы рекомендации по оптимизации государственной политики в области развития ИИ.

В Заключении диссертации подведены итоги исследования, сформулированы основные выводы.

**По теме диссертации автором опубликованы следующие работы:**

*Статьи в изданиях, рекомендованных ВАК Минобрнауки России:*

- Публикация статьи «Технологии искусственного интеллекта в сфере медицины и отечественного здравоохранения: социологический аспект» в журнале Общество: социология, психология, педагогика, издательство ООО Издательский дом ХОРС (Краснодар). — 2022. № 6, с. 70-75.

- Публикация статьи «Антропологические аспекты и социальные последствия синтеза человека и новых технологий» в журнале Alma mater. Вестник высшей школы, издательство РУДН (М.). — 2022. № 5, с. 22-25.

- Публикация статьи «Социальные риски внедрения технологий искусственного интеллекта» в журнале Теория и практика общественного развития. — 2022. № 7. С. 67–71.

- Публикация статьи «Обзор отечественных концепций трансформации общества в условиях внедрения интеллектуальных робототехнических систем» в журнале Социально-гуманитарные знания. — 2022. — № 4. — С. 165–176.

1. Искусственный интеллект (мировой рынок). Tadviser. Государство, бизнес, технологии. URL: [https://www.tadviser.ru/index.php/Статья:Искусственный\_интеллект\_(мировой\_рынок)](https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F%3A%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D0%BB%D0%BB%D0%B5%D0%BA%D1%82_%28%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D0%B9_%D1%80%D1%8B%D0%BD%D0%BE%D0%BA%29) (Дата обращения: 10.03.2020) [↑](#footnote-ref-1)
2. Дюргкейм Э. О разделении общественного труда. – М.Канон, 1996. [↑](#footnote-ref-2)
3. Giddens A. The Consequences of Modernity. Stanford, CA: Stanford University Press, 1990. [↑](#footnote-ref-3)
4. Луман Н. Социальные системы. Очерк общей теории : пер. с англ. / Н. Луман. – СПб.: Наука, 2007. [↑](#footnote-ref-4)
5. Douglas M., Wildavsky A. Risk and Culture. Berkeley, CA: University of California Press, 1982. [↑](#footnote-ref-5)
6. Thompson V., Ellis W., Wildavsky A. Cultural Theory. Boulder, CO: West-view, 1990. [↑](#footnote-ref-6)
7. Маркузе Г. Одномерный человек. М.: REFL-book, 1994. [↑](#footnote-ref-7)
8. Нейман Дж. Фон, Моргенштерн О. Теория игр и экономическое поведение / пер. с англ. Под ред. и с дор. Н.Н. Воробьева. М.: Наука, 1970. [↑](#footnote-ref-8)
9. Найт Ф. Понятие риска и неопределенности // Альманах THESIS. 1994. [↑](#footnote-ref-9)
10. Кейнс Дж.М. Общая теория занятости, про- цента и денег. Антология экономической классики. – М.: ЭКОНОВ, Ключ, 1993. [↑](#footnote-ref-10)
11. Burkardt F. Arbeitssicherheit // Mayer A., Herwig B. (Hg.). Betriebspsychologie. Handbuch der Psychologie, Bd. 9. Gottingen, 1996. [↑](#footnote-ref-11)
12. Бек У. Общество риска. На пути к другому модерну. М.: Прогресс- Традиция, 2000. [↑](#footnote-ref-12)
13. Кравченко С.А. Риски грядущей модернизации страны: нужны социологи-рискологи Вестник МГИМО-Университета. 2009;(6(9)):15-26. https://doi.org/10.24833/2071-8160-2009-6-9-15-26 [↑](#footnote-ref-13)
14. Смакотина Н.Л. Основы социологии нестабильности и риска. М., 1999. [↑](#footnote-ref-14)
15. Яницкий О.Н. Социология риска: ключевые идеи.- Мир России. 2003. № 1 [↑](#footnote-ref-15)
16. Быков А.А. Эффективное управление рисками предполагает мониторинг системных свойств мира. Проблемы анализа риска. 2016;13(2):4-5. https://doi.org/10.32686/1812-5220-2016-13-2-4-5 [↑](#footnote-ref-16)
17. Богоявленский, С.Б. "Управление риском в социально-экономических системах": Учебное пособие / С.Б. Богоявленский - СПб: Изд-во СПБГУЭФ, 2010. [↑](#footnote-ref-17)
18. Альгин А.П. Риск и его роль в общественной жизни. М.: Мысль, 1989 [↑](#footnote-ref-18)
19. Овсянников Анатолий Александрович Социология риска и безопасности. Учебник и практикум. М. , 2015// СНиСП. 2016. [↑](#footnote-ref-19)
20. Мозговая А.В. Допустимость ущерба как одно из социологических измерений отношения к риску // Социологические координаты риска. М.: Изд- во Ин-та социологии РАН, 2008. [↑](#footnote-ref-20)
21. Зубков В. Социологическая теория риска. М.: Издательство Руди, 2003. [↑](#footnote-ref-21)
22. Феофанов К.А. Социальные риски в современной социологии. М.: Луч, 2001. [↑](#footnote-ref-22)
23. Богомолова С. Н. О принятии решения в ситуации неопределенности и риска // Личность преступника как объект психологического исследования: сб. научн. тр. / под ред. А. Р. Ратинова. М.: Изд- во Всесоюзного ин-та по изучению причин и разработке мер предупреждения преступности, 1979. [↑](#footnote-ref-23)
24. Матвеенко Ю.И. Политический риск: о некоторых методах и прикладных моделях его анализа и изучения // Социология власти. 2012. [↑](#footnote-ref-24)
25. Рассел С., Норвиг П. "Искусственный интеллект. Современный подход" (2-е издание). 2006. [↑](#footnote-ref-25)
26. Форсайт Д. А., Понс Ж. Компьютерное зрение. Современный подход = Computer Vision. A Modern Approach / Под ред. А. В. Назаренко. — М.: Вильямс, 2004. [↑](#footnote-ref-26)
27. Нильсон Н. Дж. Искусственный интеллект: методы поиска решений = Problem-Solving Methods in Artificial Intelligence / Пер с англ. В. Л. Стефанюка, под ред. С. В. Фомина. — М.: Мир, 1973. [↑](#footnote-ref-27)
28. Хант Э. Искусственный интеллект = Artificial intelligence / Под ред. В. Л. Стефанюка. — М.: Мир, 1978. [↑](#footnote-ref-28)
29. Вейценбаум Дж. Возможности вычислительных машин и человеческий разум. От суждений к вычислениям, 1976. [↑](#footnote-ref-29)
30. Черчленд П.М., Черчленд П.С. Искусственный интеллект: Может ли машина мыслить, 1990. [↑](#footnote-ref-30)
31. Тьюринг А. Могут ли машины мыслить? 1950. [↑](#footnote-ref-31)
32. Неймарк А.М. Роботы на службе человека. - Ленинград: Наука, 1982 [↑](#footnote-ref-32)
33. Bell D. The coming of post-industrial society: A venture of social forecasting. — N.Y.: Basic Books, 1973. [↑](#footnote-ref-33)
34. Castells, M. The Rise of the Network Society. Information Age, vol. 1; 2nd Edition with a New Preface edition. Wiley-Blackwell, 2009. [↑](#footnote-ref-34)
35. Тоффлер Э. Шок будущего: Пер. с англ. / Э. Тоффлер. — М.: ООО «Издательство ACT», 2002. [↑](#footnote-ref-35)
36. Фрике В.В. Социология техники: Становление гуманистической парадигмы. — Перевод с немецкого: Е. З. Мирская. — Журнал «Социологические исследования», № 6–7, 1993. [↑](#footnote-ref-36)
37. Хайдеггер М. Вопрос о технике // Хайдеггер М. Время и бытие: статьи и выступления / пер. с нем. В. В. Бибихина. М.: Республика, 1993. [↑](#footnote-ref-37)
38. Алексеев А. П., Алексеева И. Ю. Судьба интеллекта и миссия разума: философия перед вызовами эпохи цифровизации: Монография. М.: Проспект, 2021. [↑](#footnote-ref-38)
39. Макаров В.Л. Становится ли человеческое общество стабильнее?// Естественный и искусственный интеллект: методологические и социальные проблемы/ Под ред. Д.И. Дубровского и В.А. Лекторского. - М., 2011. [↑](#footnote-ref-39)
40. Колесникова Г.И. Искусственный интеллект: проблемы и перспективы. М.:Видеонаука. 2018. [↑](#footnote-ref-40)
41. Орлова И.Б. Влияние искусственного интеллекта на социальную жизнь человека. Ж. «Alma Mater» №5, 2019. [↑](#footnote-ref-41)
42. Осипов Г.С. Методы искусственного интеллекта [Текст] / Г. С. Осипов. - Москва : Физматлит, 2011. [↑](#footnote-ref-42)
43. Резаев А. В., Трегубова Н. Д. Искусственный интеллект и искусственная социальность: новые явления и проблемы для развития медицинских наук // Эпистемология и философия науки. 2020. [↑](#footnote-ref-43)