

АДАПТИВНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В УПРАВЛЕНИИ ЧЕЛОВЕЧЕСКИМИ РЕСУРСАМИ

Хайдарова Малика Шокирджановна

Базовый докторант в ТГЭУ

Inson resurslarini boshqarishda moslashuvchan sun'iy intellekt

Haydarova Malika Shokirjon qizi

TDIU tayanch doktoranti

Adaptive Artificial Intelligence in Human Resource Management

Khaydarova M.Sh

Phd student in TSEU

Annotatsiya

Ko'pgina biznes jarayonlari va faoliyat sohalarida bo'lgani kabi, sun'iy intellekt (SI) inson resurslarini boshqarish (HRM) bilan tobora ko'proq integratsiyalashib bormoqda. Sun'iy intellekt tashkilotlarda inson resurslarini boshqarish faoliyatini yaxshilash uchun katta salohiyatga ega bo'lsa-da, odamlarni boshqarishni avtomatlashtirish xavfli va cheklovlarga to'la jarayonligicha qolmoqda. Ushbu risklarni aniqlash inson resurslarini boshqarishda SI dan mas'uliyatli foydalanishni rag'batlantirish uchun asosiy masala hisoblanadi. Shuning uchun biz inson resurslarini boshqarishda SI ning imkoniyatlari va mas'uliyati bo'yicha barcha fanlar bo'yicha empirik akademik adabiyotlarni ko'rib chiqdik. Shuningdek, inson resurslarini boshqarishda moslashuvchan sun'iy intellektning birinchi maqsadi ushbu hodisaning kuzatuvlari, o'lchovlari va sinovlarini o'z ichiga olgan tadqiqotlarga qaratilgan. Ko'p domenli va fanlararo yondashuv va empirik e'tibor inson resurslarini boshqarish sohasida SI ni ishlab chiqish, o'rganish va joriy etish haqiqatini yaxshiroq tushunish imkonini beradi va ular qanday mas'uliyat bilan amalga oshirilishiga oydinlik kiritadi. Xulosa qilib aytganda, biz eng kerakli va istiqbolli yo'nalishlar deb belgilaganimiz asosida tadqiqot o'tkazishga chaqiramiz.

Kalit so'zlar: Inson resurslarini boshqarish. Sun'iy intellekt. Moslashuvchan SI. Insonga moslashgan SI

Аннотация

Как и в случае со многими бизнес-процессами и дисциплинами деятельности, искусственный интеллект (ИИ) все чаще интегрируется в управление человеческими ресурсами (HRM). Хотя ИИ имеет большой потенциал для расширения деятельности по управлению человеческими ресурсами в организациях, автоматизация управления людьми не лишена рисков и ограничений. Выявление этих рисков имеет основополагающее значение для содействия ответственному использованию ИИ в управлении персоналом. Таким образом, мы провели обзор эмпирической академической литературы по всем дисциплинам, посвященной возможностям и принципам ответственности ИИ в управлении человеческими ресурсами. Это первый обзор ответственного ИИ в управлении человеческими ресурсами, в котором основное внимание уделяется исследованиям, содержащим наблюдения, измерения и тесты этого явления. Многодоменный и междисциплинарный подход, а также эмпирическая направленность обеспечивают лучшее понимание реальности разработки, изучения и внедрения ИИ в сфере управления персоналом и проливают свет на то, как они выполняются ответственно. В заключение мы призываем к проведению исследований, основанных на том, что мы определили как наиболее необходимые и многообещающие направления.

Ключевые слова: Управление людскими ресурсами. Искусственный интеллект. Ответственный ИИ. Человеко-ориентированный ИИ

Abstract

As is the case for many business processes and activities disciplines, artificial intelligence (AI) is increasingly integrated into human resources management (HRM). While AI has great potential to augment HRM activities in organizations, automating the management of humans is not without risks and limitations. The identification of these risks is fundamental to promoting the responsible use of AI in HRM. We thus conducted a review of the empirical academic literature across disciplines on the affordances and responsible principles of AI in HRM. This is the first review of responsible AI in HRM that focuses solely on studies containing observations, measurements, and tests about this phenomenon. The multi-domain and multidisciplinary approach and empirical focus provide a better understanding of the reality of the development, study, and deployment of AI in HRM and shed

light on how these are conducted responsibly. We conclude with a call for research based on what we identified as the most needed and promising avenues.

Keywords: Human resources management. Artificial intelligence. Responsible AI. Human-centered AI

Введение

Деятельность по управлению человеческими ресурсами (HRM) включает в себя несколько рутинных и трудоемких задач, пока они также зависят от человеческого восприятия, субъективности или предубеждений¹. По этим причинам управление человеческими ресурсами рассматривается как благодатная почва для использования искусственного интеллекта. Использование искусственного интеллекта (ИИ) в управлении человеческими ресурсами разрабатывается, тестируется, анализировались и исследовались эмпирически в различных областях исследований². Эмпирические исследования относятся к исследованиям, основанным на данных, связанных с явлением, которое наблюдали, измеряли и/или тестировали исследователи. Потому что нет единого мнения об определении ИИ во всех областях и внутри них из-за исторических дебатов о том, что именно такое «интеллект», а ИИ является обобщающим термином для различного подмножества технологий, имитирующих человеческий интеллект (т. е. компьютерное зрение, обработка естественного языка, машинное обучение, глубокое обучение), в этой статье будет использоваться относительно широкое, но четкое определение технологии, которая может применяться во всем мире. использование ИИ в HR. В частности, в этой статье ИИ определяется как «способность машины учиться на опыте, приспосабливаться к новым входным данным и выполнять человекоподобные задачи». Быстрый рост использования ИИ в управлении человеческими ресурсами отражается в публикации за последние несколько лет нескольких литературных и концептуальных обзоров по ИИ в управлении человеческими ресурсами.

Несмотря на важные достоинства этих обзоров, в нашем полном понимании возможностей и рисков интеллектуальных технологий в управлении

¹ 1. Абдурахманова Г.К. Управление человеческими ресурсами. Учебник. - Ташкент: Издательство «Фан» Академии наук Республики Узбекистан, 2023. - 396 с.

² Acharyya, S., Datta, A.K.: Matching formulation of the Staff Transfer Problem: meta-heuristic approaches. *Opsearch* 57(3), 629–668 (2020). <https://doi.org/10.1007/s12597-019-00432-w>

человеческими ресурсами остаются некоторые ограничения, что требует тщательного обзора литературы с другой точки зрения, чем предыдущие. В частности, уделяя больше внимания литературе соответствующих областей, таких как информатика или HRM, предыдущие обзоры не в полной мере учитывают многодоменную природу ИИ в HRM и сочетание как технических, так и социальных аспектов этого явления. Наше исследование преодолевает эти ограничения, рассматривая в разных областях как то, как ИИ используется в HRM (т. е. технический аспект), так и принципы адаптивного ИИ, применяемые в нашей выборке исследований (т. е. социальный аспект). Что касается технического аспекта, то здесь наблюдается определенная неточность в отношении изучаемой технологии (HRM с поддержкой искусственного интеллекта). В частности, поскольку в обзорах часто не удается четко указать и определить, что представляет собой рассматриваемая технология, в недавних обзорах описываются исследования различных и не обязательно связанных с ИИ технологий, используемых в управлении человеческими ресурсами (например, анализ больших данных, который представляет собой массивный анализ данных). количество детальных и исчерпывающих данных, но программное обеспечение ИИ не используется *ipso facto* для использования этих данных³. Наш текущий обзор преодолевает это ограничение, включив в него только исследования, которые подробно изучают использование ИИ в соответствии с вышеупомянутым определением и, таким образом, разъясняют технические аспекты использования ИИ в различных функциях управления персоналом.

Что касается социального аспекта, то в настоящее время нет обзора с сосредоточить внимание на принципах ответственного ИИ, применяемых в сфере управления персоналом. Именно текущие обзоры с учетом этого социального аспекта в основном обсуждают или предлагают концептуальные основы, предлагающие решения о том, как следует изучать и внедрять ИИ, или используются, но ни один из них эмпирически не наблюдает фактического применения таких рамок. Наше исследование способствует знания, взглянув изнутри на то, как принципы ответственности применяются при разработке, изучении и внедрении ИИ в HRM. Более того, существует необходимость рассмотреть возможность применения ответственных исследовательских практик, поскольку во многих исследованиях

³ Alola, U.V., Atsa'am, D.D.: Measuring employees' psychological capital using data mining approach. J. Public Affairs (2019). <https://doi.org/10.1002/pa.2050>

подчеркивается, что ответственность является ключевым элементом при изучении использования ИИ в управлении человеческими ресурсами. Насколько нам известно, это первый систематический обзор литературы, в котором конкретно рассматривается, какие или какие принципы составляют ответственный ИИ в управлении человеческими ресурсами и как они применяются в эмпирических исследованиях в различных областях. Однако, поскольку понятие ответственного использования технологий постоянно развивается в литературе, нет единого мнения относительно определения и применения ответственного ИИ в сфере управления персоналом. В этом исследовании мы адаптируем широкое определение ответственного ИИ, предложенное Барредо Арриетой и др. в котором говорится, что это «ряд принципов ИИ, которые обязательно должны соблюдаться при развертывании ИИ в реальных приложениях»⁴. Эта адаптация будет осуществлена путем включения ответственный способ изучения ИИ и определение ответственного ИИ как набора этических принципов, которым необходимо обязательно следовать при разработке, изучении и внедрении ИИ. Это определение будет служить основой для нашего обзора, а также даст исследователям, организациям и политикам необходимое общее понимание того, что подразумевается под ответственным ИИ.

В целом, цель этой статьи — изучить объем существующей эмпирической литературы по ответственному ИИ в управлении человеческими ресурсами, одновременно пытаясь преодолеть ограничения предыдущей работы путем проведения систематического обзора литературы, включающей только эмпирические исследования, все типы журналов (не только в сфере управления человеческими ресурсами), и никакой априорной концептуальной основы. Целями данного обзора являются: выявить эмпирические исследования текущего использования ИИ в HRM, рассмотреть эмпирические знания о принципах ответственного ИИ в HRM и их применении, и оценить, в какой степени эти исследовательские практики способствуют сочетанию использования ИИ с этической, достойной и качественной работой.

Методология

⁴ Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F.: Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58, 82–115 (2020). <https://doi.org/10.1016/j.infus.2019.12.012>

В ходе нашего обзора мы руководствовались заявлением PRISMA 2020, которое позволяет прозрачно сообщать о нашей стратегии поиска и наших результатах. Чтобы быть включенными, статьи должны были: представлять собой эмпирическое исследование, пройти рецензирование, быть явно связаны с функцией управления человеческими ресурсами и явно включать технологию, основанную на искусственном интеллекте. на основе определения ИИ, представленного во введении.

Чтобы выявить исследования, мы провели поиск в следующих базах данных: Academic search Complete, Business Source Complete, PsycArticles, Web of Science и ABI/INFORM Collection. Широкий охват и разнообразие этих баз данных позволили нам оценить несколько областей исследований в нашем обзоре литературы.

Первым шагом было использование следующей структуры поиска: Термины, связанные с предметной областью «ИИ» ответственная практика, связанные термины. Этот поиск проводился в каждой базе данных. Было проверено пятьдесят случайных исследований в каждой базе данных (аннотация и название), чтобы выявить любые дополнительные условия поиска, которые могли быть пропущены. Второй шаг — использовать следующую структуру поиска: Термины, связанные с предметной областью «ИИ» термины, связанные с ИИ. Это снова было проверено в каждой базе данных, при этом было рассмотрено максимум 50 случайных исследований.

На данный момент количество возвращенных записей составило 2561.

Ссылки были организованы совместно с менеджером по библиографии Zotero (Корпорация цифровых стипендий), а управление данными осуществлялось с помощью Covidence (Covidence Inc., Австралия), онлайн-платформы для управления систематическими обзорами и множества электронных таблиц. Дубликаты были автоматически обнаружены Zotero и удалены вручную. Записи, не относящиеся к теме (например, в ветеринарной сфере), также удалялись вручную для всего удалено 1796 записей, осталось 765 записей. Затем мы использовали подход «снежного кома», чтобы добавить больше записей с соответствующими заголовками ($n = 259$). Метод «снежного кома» используется для обогащения систематических обзоров за счет использования ссылок на статьи в существующих выборках для выявления других статей, потенциально имеющих отношение к обзорам. Эта

техника была особенно важна для нашей статьи, поскольку литература по ИИ в управлении человеческими ресурсами быстро развивается, а недавно опубликованные работы или материалы конференций, возможно, медленно попадали в базы данных, которые мы искали. Затем 1024 идентифицированные записи были переданы в Covidence, которая автоматически удалила оставшиеся дубликаты, прошедшие в обход первого процесса (n=15).

Таким образом, на этапе проверки заголовка и реферата в Covidence было выявлено 1009 записей. Чтобы обеспечить согласованность перед тем, как приступить к проверке записей, показатель надежности между кодировщиками рассчитывался с использованием процентного согласия (мы заранее договорились, что если оно достигнет >75%, мы будем двигаться дальше). В частности, в ходе пилотного тестирования два исследователя независимо проанализировали название и аннотацию случайная выборка из 50 записей, основанная на четырех критериях отбора и указывающая, какие критерии не соблюдались, если исследование было исключено. Затем их работы сравнили. Потребовался только один раунд пилотного тестирования, при этом оба исследователя проверяли 42 из 50 записей совершенно одинаковым образом (т. е. получили оценку 84%).

Во-первых, после тщательного прочтения два участника исследования команда подробно перечислила каждый из текстов в сводной таблице, зафиксировав различные характеристики, связанные с рукописью, и сообщенные результаты. Эти синтезы были сравнены вручную и оказались очень похожими. Редкие различия, которые имели место, были устранены путем обсуждения внутри исследовательской группы. Мейеринк и Бондарук описывают возможности алгоритмов ИИ как планирование талантливой рабочей силы, а также вознаграждение и льготы⁵. Более того, чтобы внести большую ясность и детализацию в наш анализ технического аспекта ИИ в HRMR, мы дополнительно детализировали эту метакатегорию в зависимости от того, были ли связанные алгоритмы ИИ описательными, прогнозирующими и/или предписывающими согласно работе Лейхта-Деобальда и др.⁶. Таким образом, эти мета-категории и детализированные

⁵ Meijerink, J., Bondarouk, T.: The duality of algorithmic management: toward a research agenda on HRM algorithms, autonomy and value creation. *Hum. Resour. Manag. Rev.* 33(1), 1–14 (2023). <https://doi.org/10.1016/j.hrmmr.2021.100876>

⁶ Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., Kasper, G.: The challenges of algorithm-based HR decision-making for personal integrity. *J. Bus. Ethics* 160(2), 377–392 (2019).

подкатегории использовались для классификации типов алгоритмов HRM при извлечении данных. Что касается метакатегории «Адаптивный ИИ», мы сосредоточились на понятии ответственности, связанной с использованием системы, а не на понятии ответственности, связанной с целью системы. Например, независимо от того, можно ли считать качество системы ответственным, например, содействие благополучию сотрудников или устойчивому поведению, мы сосредоточили внимание на ответственном использовании систем. Это основано на аргументе, что система с благими намерениями все же может быть безответственной в его использование (например, система благосостояния может дискриминировать определенную группу населения). Что касается категоризации принципов ответственного (или этического) ИИ, интересно отметить, что на сегодняшний день в мире существует более 80 нормативных рамок по ответственному ИИ⁷. Эти концепции имеют несколько совпадений и общностей в принципах (например, прозрачность ИИ), но также и важные расхождения в терминологии (например, прозрачность; объяснимость; черный ящик; непрозрачность). Эта нечеткость побудила нас классифицировать ответственный ИИ, который вытекает из наших 107 избранных эмпирических исследований, в соответствии с наиболее общими принципами в литературе по ответственному ИИ (т. е. автономия и агентность, предвзятость и дискриминация, объяснимость и прозрачность, человеческая роль, воспринимаемая справедливость и доверие), конфиденциальность, подотчетность системы и условия труда) (в основном на основе⁸. Сначала мы проанализировали, включают ли исследования ответственные практики (Категория: Включение ответственных практик), а затем детализировали практику (Категория: Тип ответственных практик).

Во-вторых, после завершения этого этапа три члена исследовательской группы индивидуально проанализировали сводную таблицу с целью выявления точек соприкосновения внутри категорий. Отобранные эмпирические рецензируемые исследования существенно различались по используемой лексике, теоретическим подходам, целям, дисциплинам, углам анализа и методологии. Эта весьма разнообразная выборка исследований усложнила анализ результатов. Это побудило нас принять индуктивный

<https://doi.org/10.1007/s10551-019-04204-w>

⁷ Alethicist. (2022). <https://www.aiethicist.org/ai-principles>

⁸ Ashok, M., Madan, R., Joha, A., Sivarajah, U.: Ethical framework for artificial intelligence and digital technologies. *Int. J. Inf. Manag.* (2022). <https://doi.org/10.1016/j.ijinfomgt.2021.102433>

подход к анализу⁹. Этот подход направлен на генерирование знаний о концепциях в литературе, а не на подтверждение ранее существовавшей теории, а конечный результат достигается за счет обобщения всех наблюдений.

Руководствуясь целями исследования, мы уделили особое внимание возникающим темам о том, как ИИ в настоящее время используется в функциях управления персоналом (Метакатегория: Управление человеческими ресурсами) и как концепции ответственного ИИ применяются в этих эмпирических исследованиях (Метакатегория: : Адаптивный ИИ). Затем трое исследователей встретились, чтобы сравнить результаты своего анализа. Опять же, сходство было сильным, а некоторые различия обсуждались всей исследовательской группой и были согласованы.

Примечательно, что в отношении категории «Функция человеческих ресурсов» мы усовершенствовали концептуализацию Мейеринка и Бондарука, добавив функцию «Здоровье и благополучие», поскольку мы обнаружили несколько исследований, подпадающих под эту тему. Более того, что касается метакатегории «Адаптивный ИИ», в результате исследований выявилось только 5 принципов. Они будут подробно рассмотрены в следующем разделе.

Полученные результаты

Все наши 107 избранных эмпирических и рецензируемых исследований были опубликованы в период с 2004 по 2022 год. Средний год публикации — 2019. На рисунке 2 показано распределение наших 107 эмпирических исследований по годам их публикации. Важно иметь в виду, что 2022 год включает в себя только период с января по июнь, поскольку июнь 2022 года был месяцем извлечения данных.

Кроме того, наша выборка содержит 86 различных журналов или материалов конференций в различных областях (например, инженерия, этика, управление персоналом, информационные системы, менеджмент, математика и психология). Что касается дизайна исследования, выбранные эмпирические статьи включают 63 экспериментальных исследования, 15 полевых исследований, 24 исследования, сочетающих оба метода, три тематических

⁹ Aguinis, H., Ramani, R.S., Alabduljader, N.: Best-practice recommendations for producers, evaluators, and users of methodological literature reviews. *Organ. Res. Methods* 26(1), 46–76 (2020). <https://doi.org/10.1177/1094428120943281>

исследования и две этнографии. Более того, в 89 исследованиях использовались количественные данные, в 13 — качественные данные, а в пяти — и те, и другие. Кроме того, в 69 исследованиях изучалась разработка новой системы или модели искусственного интеллекта. В подавляющем большинстве этих исследований возможности и конструкция новых систем по сравнению со старыми были не обсуждаются. Скорее, они сосредоточились на том, как их новая система обеспечивает большую достоверность и производительность, чем предыдущие системы или люди-профессионалы. Кроме того, почти все эти исследования развития представляли собой лабораторные эксперименты и поэтому не были реализованы и применены на практике.

Что касается контекста всех 107 исследований, они включали сбор данных из 23 различных стран: Австралии, Бангладеш, Бельгии, Канады, Китая, Колумбии, Франции, Германии, Индии, Индонезии, Ирана, Иордании, Кореи, Новой Зеландии, Нигерии, Норвегия, Палестина, Португалия, Россия, Швейцария, Тайвань, Турция и США. При этом в 29 исследованиях нашей выборки не указана страна сбора данных. Страной, которая встречается чаще всего, являются США¹⁰, и ни одно исследование не включает межстрановой анализ. Что касается сектора деятельности, то в 49 исследованиях исследуемый сектор не был указан или он был неприменим. Наиболее изученным сектором были государственные или общественные услуги (например, учителя) с 13 исследованиями, за которым следовал сектор информационных технологий (ИТ) с 10 исследованиями. Другими секторами в нашей выборке являются услуги (7 исследований), производство (4), научные круги (4), электроснабжение (3), военное дело (2), телекоммуникации (2), строительство (1), продажи (1), розничная торговля (1) и некоммерческая организация (1). Кроме того, девять исследований не были специфичны для изучаемого сектора и не сообщали о популяции работников различных профессий, и поэтому их нельзя было классифицировать конкретно по их сектору обучения.

Более того, многие исследуемые организации были крупными или многонациональными¹¹. Это согласуется с размерами выборки 69

¹⁰ Anoaica, A., Ben Hassine, A., Deleris, L.A.: Equal pay for equal competences: a statistical approach to address equal pay gap. *ECAI 2020*, 2949–2955 (2020). <https://doi.org/10.3233/FAIA200468>

¹¹ Altemeyer, B.: Making the business case for AI in HR: two case studies. *Strateg. HR Rev.* 18(2), 66–70 (2019). <https://doi.org/10.1108/shr-12-2018-0101>

исследований, в которых была разработана новая система или модель искусственного интеллекта, которым часто приходилось использовать огромные наборы данных. Например, Авраами и др.¹² использовали набор продольных архивных данных, включающий более 700 000 сотрудников крупной общественной организации, чтобы разработать инструмент, который прогнозирует уровень текучести кадров.

Цель этой статьи — представить обзор основных направлений использования ИИ в сфере управления человеческими ресурсами на основе эмпирических исследований, включенных в наш обзор. Афордансы относятся к использованию или назначению, которое может иметь вещь, которую люди замечают как часть того, как они ее видят или ощущают. Наши 107 избранных эмпирических и рецензируемых исследований включают 79 исследований, которые описывают, как ИИ используется в конкретных функциях управления персоналом и типы задействованных алгоритмов ИИ (30 описательных алгоритмов, 31 прогнозирующий алгоритм и 27 предписывающих алгоритмов), а также 28 исследований, которые не были конкретными. нам достаточно для классификации, поэтому они не подробно описаны в этом разделе (например, исследования общего восприятия ИИ или общего использования ИИ в управлении человеческими ресурсами)¹³. Примечательно, что некоторые включенные исследования содержат более одного типа алгоритмов ИИ и/или более одной функции HRM.

Из нашей выборки из 107 эмпирических исследований в 63 явно не применялся принцип ответственного ИИ. В 27 из этих 63 исследований предполагалось, что система ИИ уменьшит предвзятость и дискриминацию, поскольку она уменьшит или устранил человеческую субъективность. Хотя это предположение согласуется с некоторыми концептуальными разработками, оно не было эмпирически проверено в 27 выявленных исследованиях.

Концепция справедливости ИИ в управлении человеческими ресурсами, похоже, не имеет общепринятого определения в эмпирической литературе. Вместо этого мы обнаружили, что это скорее общий термин, который

¹² Avrahami, D., Pessach, D., Singer, G., Chalutz Ben-Gal, H.: A human resources analytics and machine-learning examination of turnover: implications for theory and practice. *Int. J. Manpow.* 43(6), 1405–1424 (2022). <https://doi.org/10.1108/ijm-12-2020-0548>

¹³ Bigman, Y.E., Wilson, D., Arnestad, M.N., Waytz, A., Gray, K.: Algorithmic discrimination causes less moral outrage than human discrimination. *J. Exp. Psychol. Gen.* (2022). <https://doi.org/10.1037/xge0001250>

охватывает три выявленных нами принципа, а именно предвзятость и дискриминация в ИИ, ориентированном на управление персоналом, воспринимаемую справедливость и доверие к решениям и результатам, а также проблемы конфиденциальности (или навязчивость), связанную с использованием ИИ. Двадцать исследований были сосредоточены на выявлении или смягчении предвзятости и дискриминации в системе искусственного интеллекта для управления человеческими ресурсами. Действительно, решения, принимаемые с помощью ИИ, на самом деле могут быть предвзятыми и дискриминационными, поскольку они отражают данные, на которых они основаны¹⁴. В некоторых исследованиях изучалось, как можно проверять инструменты HRM AI и как эта возможность аудита может способствовать выявлению и смягчению предвзятости¹⁵. Например, что касается искусственного интеллекта в привлечении талантов, Кёхлинг и др. [76] показывают, что ИИ воспроизводит (и может даже amplify) существующее неравенство в наборе данных и недостаточная представленность определенных групп приводит к непредсказуемой вероятности приглашения кандидатов из этих групп на собеседования. Другие включают принцип предвзятости, добавляя этап проверки или тест при разработке своей системы ИИ, чтобы продемонстрировать, что система, разработанная в исследовании, не дискриминирует¹⁶. Наконец, принцип предвзятости и дискриминации применяется в эмпирических исследованиях, в которых были разработаны системы искусственного интеллекта, единственной целью которых является обнаружение и смягчение предвзятости и дискриминации¹⁷. Например, Хангартнер и др. разработали инструмент на базе искусственного интеллекта для постоянного мониторинга дискриминации при приеме на работу на онлайн-платформах по подбору персонала.

¹⁴ Buolamwini, J., Gebru, T.: Gender shades: intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 2018 ACM Conference on Fairness, Accountability and Transparency, (2018)

¹⁵ Booth, B.M., Hickman, L., Subburaj, S.K., Tay, L., Woo, S.E., D’Mello, S.K.: Integrating psychometrics and computing perspectives on bias and fairness in affective computing: a case study of automated video interviews. IEEE Signal Process. Mag. 38(6), 84–95 (2021). <https://doi.org/10.1109/msp.2021.3106615>

¹⁶ Campion, M.C., Campion, M.A., Campion, E.D., Reider, M.H.: Initial investigation into computer scoring of candidate essays for personnel selection. J. Appl. Psychol. 101(7), 958–975 (2016). <https://doi.org/10.1037/apl0000108>

¹⁷ Anoaica, A., Ben Hassine, A., Deleris, L.A.: Equal pay for equal competences: a statistical approach to address equal pay gap. ECAI 2020, 2949–2955 (2020). <https://doi.org/10.3233/FAIA200468>

Наконец, 17 исследований либо применяли, либо изучали важность участия людей (например, разработчиков, менеджеров, HR-практиков или сотрудников) в разработке, внедрении и использовании системы искусственного интеллекта в HRM. Характер исследуемой роли человека в первую очередь включал уровень контроля заинтересованных сторон над системой (например, изменять или принимать окончательное решение, задавать вопросы, апеллировать или вносить данные в алгоритм). Уровень контроля или участия пользователей в системах ИИ кажется важным для содействия ответственному использованию и даже принятию¹⁸, поскольку «люди в конечном итоге должны сохранить роль лиц, принимающих решения» [10, с. 66]. Например, Аноаиса и др.¹⁹ внедрили механизмы (в основном с точки зрения объяснимости), чтобы предоставить отделу кадров свободу выносить собственные суждения, а Фалиагга и др. предостерегают от слепого доверия к автоматизированной системе. С точки зрения взаимодействия человека и компьютера, как и в других областях, обеспечение некоторой степени контроля кажется полезным. Эти результаты отражают принцип подотчетности, согласно которому люди должны оставаться ответственными и подотчетными за свои решения, даже если они поддерживаются системами искусственного интеллекта.

Заключение

В данной статье представлен обзор литературы по эмпирическим и рецензируемое исследование по ответственному ИИ в сфере управления персоналом доменах, принимая во внимание сложность этого явления, рассматривая как технический аспект (т. е. то, как ИИ используется в управлении персоналом), так и социальный аспект (т. е. принципы ответственного ИИ). Мы вносим свой вклад в литературу, показывая, как ИИ используется в HRM, изучаем, как принципы ответственности применяются в эмпирических исследованиях ИИ в HRM, и оцениваем степень, в которой эти исследовательские практики способствуют развитию ответственного ИИ.

Во-первых, наши результаты показывают, что ИИ в HRM представляет собой многодоменную систему. Тема исследования изучается во всем мире и в различных секторах, поскольку наша выборка из 107 эмпирических и

¹⁸ Altemeyer, B.: Making the business case for AI in HR: two case studies. *Strateg. HR Rev.* 18(2), 66–70 (2019). <https://doi.org/10.1108/shr-12-2018-0101>

¹⁹ Anoaica, A., Ben Hassine, A., Deleris, L.A.: Equal pay for equal competences: a statistical approach to address equal pay gap. *ECAI 2020*, 2949–2955 (2020). <https://doi.org/10.3233/FAIA200468>

рецензируемых исследований содержит 86 различных журналов или материалов конференций в различных областях, 23 разных странах и 12 различных секторах. Более того, наши описательные результаты показывают, что популярность этой темы исследования значительно возросла за последнее десятилетие. Наши результаты также показывают, что об использовании трех типов алгоритмов ИИ (т. е. описательных, прогнозирующих и предписывающих) сообщается в соответствии с шестью различными функциями HRM (т. е. 1 — привлечение талантов, 2 — оценка эффективности, 3 — управление талантами, 4 — планирование рабочей силы, 5 — здоровье и благополучие и 6 — оплата труда), при этом системы ИИ по привлечению талантов являются наиболее эмпирически изученной функцией управления человеческими ресурсами и, по-видимому, являются наиболее хорошо реализовано. Можно сосуществовать несколько объяснений, объясняющих значительный дисбаланс с точки зрения интереса к использованию ИИ в различных функциях управления персоналом. Например, системы искусственного интеллекта могут более склонны к привлечению талантов, поскольку эта задача, как известно, требует много времени, является избыточной и подвержена человеческой предвзятости, а данные, доступные для обучения систем, включают как реальных, так и потенциальных кандидатов, поэтому количество данные обычно намного больше.

Учитывая ранее подчеркнутые элементы, тем более тревожно, что некоторые исследования представили системы искусственного интеллекта в управлении человеческими ресурсами как более этические практики, чем традиционное управление человеческими ресурсами, основываясь на теоретическом аргументе о том, что системы искусственного интеллекта смягчают человеческую субъективность в практике и, следовательно, уменьшают предвзятость. Этот аргумент часто основывается на представлении о том, что системы могут достичь «справедливости за счет неосведомленности», что означает способность систем не использовать явно защищенные атрибуты или исключать чувствительные функции в процессе прогнозирования²⁰. Однако мы не нашли эмпирических исследований, проверяющих, действительно ли системы ИИ менее предвзяты, чем традиционные практики

²⁰ Chen, J., Kallus, N., Mao, X., Svacha, G., Udell, M.: Fairness under unawareness. In: Proceedings of the AMC Conference on Fairness, Accountability, and Transparency, pp. 339–348. (2019). <https://doi.org/10.1145/3287560.3287594>

HRM. Таким образом, мы бы не одобряли любые заявления о том, что системы ИИ менее предвзяты, чем практики, если только дальнейшие исследования эмпирически исследуют и демонстрируют обоснованность этого утверждения. Действительно, мы считаем, что поддержка таких предположений без научных испытаний была бы безответственной, поскольку они могли бы ошибочно побудить практиков применять технологии искусственного интеллекта для уменьшения предвзятости и дискриминации.

Наш обзор ясно показывает, что как использование ИИ в управлении человеческими ресурсами, так и применение принципов ответственности требуют дальнейшего изучения. Нашим первым и главным стимулом для будущих исследований будет разработка более разнообразных протоколов исследований, основанных на обширных полевых исследованиях и реальных условиях. Действительно, поскольку большинство исследований в нашей выборке основаны на экспериментальных планах, кажется трудным обобщить их вклад в реальность организационного контекста. Поэтому их вклад в

практики остаются несколько ограниченными. Более того, поскольку наши результаты показывают серьезный разрыв между концептуальными и эмпирическими исследованиями ответственного ИИ в управлении человеческими ресурсами, мы настоятельно призываем будущие исследования либо применять принципы ответственности в качестве концептуальной основы при проведении эмпирической работы, либо исследовать влияние ответственных принципов ИИ в управлении человеческими ресурсами. Мы обнаружили мало эмпирических исследований по таким темам, как объяснимость и прозрачность (на самом деле в нашей выборке нет исследований по субъективной прозрачности) или конфиденциальность. Даже наиболее эмпирически изученный принцип в нашей выборке (т.е. предвзятость и дискриминация) получил мало эмпирического изучения по сравнению с общественной и академической дискуссией вокруг него. Более того, воспринимаемая справедливость и/или доверие в основном изучались в экспериментах и гипотетических сценариях, и мы призываем к методологической диверсификации, например, к большему количеству исследований в контексте реальной жизни. Что касается роли людей в ответственном ИИ в управлении человеческими ресурсами, мы

считаем, что этот принцип может быть одним из самых сложных для исследования, поскольку степень и характер человеческой роли могут сильно различаться от одной ситуации к другой, и, следовательно, требуют дополнительных исследований этого принципа. Говоря более конкретно, хотя роль практиков управления человеческими ресурсами была документально подтверждена, количество исследований было небольшим, а знания о результатах высокой или низкой роли управления человеческими ресурсами были незначительными, практиков по адаптивному ИИ по-прежнему мало. В том же духе, хотя это в основном обсуждается в теоретических или концептуальных статьях, мы очень мало знаем о навыках, которые следует развивать среди специалистов по управлению человеческими ресурсами, чтобы в полной мере позволить им играть эту роль. Мы также обнаружили, что некоторые ответственные принципы, представленные в литературе, отсутствовали в нашей эмпирической и рецензируемой выборке. В частности, наши исследования не включали эмпирические исследования влияния ИИ на автономию и агентность заинтересованных сторон или подотчетность системы. Таким образом, мы призываем будущие исследования ИИ в управлении человеческими ресурсами разнообразить подход, используемый для дальнейшего изучения этих принципов ответственности.

Кроме того, мы призываем будущих исследователей быть откровенными и предоставлять как можно больше подробностей об изучаемых алгоритмах ИИ, их характеристиках и возможностях, поскольку это позволит лучше понять, как различные типы, функции и возможности ИИ обязанности принципы влияют на разные результаты. Этому можно способствовать с помощью многопрофильных исследовательских групп. В связи с этим мы также призываем будущих исследователей учитывать многодоменную природу Адаптивного ИИ в управлении человеческими ресурсами, создавая междисциплинарные исследовательские группы и устраняя разрозненность между областями исследований. То есть с исследователями с передовыми техническими знаниями в области искусственного интеллекта, а также с исследователями с глубокими знаниями в области управления человеческими ресурсами. Эти комбинации позволят лучше понять сложные явления ответственного ИИ в управлении человеческими ресурсами. Поскольку наши результаты показывают концептуальную путаницу в отношении принципов ответственности, мы также призываем к будущим исследованиям с

использованием знания из концептуальной литературы и подробно детализируйте изучаемый принцип ответственности. Мы обнаружили, что в некоторых эмпирических исследованиях используются термины из области ответственности ИИ, такие как прозрачность или дискриминация, без определения этого термина или использования его в соответствии с литературой. Например, в некоторых исследованиях прозрачности ИИ это привело к концептуальной путанице, поскольку исследователи фактически изучая концепцию объяснимости. Кроме того, мы обнаружили, что ответственный ИИ в сфере управления персоналом изучается во многих странах, но не нашли межстранового анализа. Мы призываем к будущим исследованиям для проведения такого анализа, чтобы углубить наше понимание Адаптивного Искусственного Интеллекта в управлении человеческими ресурсами и его различия в разных странах. Наконец, наши результаты показывают, что область применения ИИ в управлении человеческими ресурсами быстро развивается, при этом количество исследований значительно увеличилось за последнее десятилетие. С тех пор уже были опубликованы дополнительные эмпирические работы по адаптивному ИИ в управлении человеческими ресурсами.

Использованная литература

1. Абдурахманова Г.К. Управление человеческими ресурсами. Учебник. - Ташкент: Издательство «Фан» Академии наук Республики Узбекистан, 2023. - 396 с.
2. Acharyya, S., Datta, A.K.: Matching formulation of the Staff Transfer Problem: meta-heuristic approaches. *Opsearch* **57**(3), 629–668 (2020). <https://doi.org/10.1007/s12597-019-00432-w>
3. Acikgoz, Y., Davison, K.H., Compagnone, M., Laske, M.: Justice perceptions of artificial intelligence in selection. *Int. J. Sel. Assess.* **28**(4), 399–416 (2020). <https://doi.org/10.1111/ijsa.12306>
4. Aguinis, H., Ramani, R.S., Alabduljader, N.: Best-practice recommendations for producers, evaluators, and users of methodological literature reviews. *Organ. Res. Methods* **26**(1), 46–76 (2020). <https://doi.org/10.1177/1094428120943281>
5. *AIethicist*. (2022). <https://www.aiethicist.org/ai-principles>.

6. Aizenberg, E., van den Hoven, J.: Designing for human rights in AI. *Big Data Soc* (2020). <https://doi.org/10.1177/2053951720949566>
7. Albert, E.T.: AI in talent acquisition: a review of AI-applications used in recruitment and selection. *Strateg. HR Rev.* **18**(5), 215–221 (2019). <https://doi.org/10.1108/shr-04-2019-0024>
8. Allal-Chérif, O., YelaAránega, A., Castaño Sánchez, R.: Intelligent recruitment: how to identify, select, and retain talents from around the world using artificial intelligence. *Technol. Forecast Soc. Change* (2021). <https://doi.org/10.1016/j.techfore.2021.120822>
9. Alola, U.V., Atsa'am, D.D.: Measuring employees' psychological capital using data mining approach. *J. Public Affairs* (2019). <https://doi.org/10.1002/pa.2050>
10. Altemeyer, B.: Making the business case for AI in HR: two case studies. *Strateg. HR Rev.* **18**(2), 66–70 (2019). <https://doi.org/10.1108/shr-12-2018-0101>