

ВЫСТУПЛЕНИЕ ЧЛЕНА-КОРРЕСПОНДЕНТА РАН Д.В. УШАКОВА

Материал поступил в редакцию 03.12.2018 г.

Принят к публикации 25.12.2018 г.

Ключевые слова: наука о человеке, искусственный интеллект, нейронные сети, глубинное научение, социальные технологии.**DOI:** <https://doi.org/10.31857/S0869-5873894376-378>

Взаимный интерес исследований в области искусственного интеллекта и социогуманитарных наук ярко проявился в рамках недавно завершившегося конкурса Российского фонда фундаментальных исследований "Конвергенция". В этом конкурсе участвовали проекты, выполняемые совместно естественно-научными и социогуманитарными коллективами. Конкурс открыл новую возможность для междисциплинарного сотрудничества, но он также аккумулировал сведения о том, какие естественно-научные и социогуманитарные дисциплины в наибольшей мере заинтересованы во взаимодействии друг с другом.

Абсолютным чемпионом по числу заявок оказалась информатика, которая более всего востребована для междисциплинарного сотрудничества с социогуманитарными науками. Это касается и искусственного интеллекта. Почему так происходит? Полагаю, что основные причины следующие: во-первых, взаимная полезность в сфере фундаментальных исследований; во-вторых, технологии искусственного интеллекта позволяют решать важные социогуманитарные задачи; в-третьих, сами технологии искусственного интеллекта нуждаются в социогуманитарном сопровождении.

Что касается фундаментальных исследований, в генезисе основных идей искусственного интеллекта можно обнаружить социогуманитарные корни. Например, большую отдачу в наши дни даёт такая технология искусственного интеллекта, как сети глубинного научения. Но вряд ли они сегодня существовали бы без интенсивной предварительной работы в рамках наук о человеке. Идея сети, перерабатывающей информацию, возникла из исследований человека. Философско-психологический ассоцианизм предложил принцип организации мышления человека на основе ассоциаций, которые вызывают в сознании некоторые идеи на базе связанных с ними других идей. Следующий важный шаг — позитивное и негативное подкрепление как фундамент для формирования ассоциации — был сделан в теории И.П. Павлова и в теориях оперантного научения. Затем было дано описание работы узлов сети в виде синапса

Д. Хебба и разработаны формальные нейронные сети Маккаллока—Питса. Наконец, первая нейронная сеть глубинного научения — перцептрон — была создана Ф. Розенблаттом, который, имея широкий круг интересов, в первую очередь всё же был психологом и нейрофизиологом, а затем уже специалистом в области информатики. Аналогичную "родословную" можно реконструировать и для многих других направлений и технологий искусственного интеллекта, например, эвристического программирования.

Безусловно, происходит движение и в обратную сторону. Наиболее яркий пример — когнитивная наука, которая обычно включает четыре дисциплины: психологию, лингвистику, нейрофизиологию и искусственный интеллект. Когнитивная наука, с одной стороны, опирается на компьютерную метафору, с другой — заимствует описание отдельных процессов из искусственных устройств (например, идея фреймов памяти, пришедшая в когнитивные исследования благодаря специалисту в области искусственного интеллекта Р. Шенку).

Второе направление сотрудничества — искусственный интеллект как средство создания технологий для социогуманитарной сферы. Эти технологии становятся всё более востребованными в фундаментальных исследованиях, но не менее важны и для практики.

Информатизация жизни общества приводит к формированию цифровых баз данных, характеризующих поведение людей и социальных групп. Классический пример — Интернет, однако им вопрос не исчерпывается. Так, всякого рода мобильные девайсы стали в последнее время потенциальным источником информации о поведении человека. Ожидается, что эта сфера будет только расширяться. Например, цифровизация образования со временем может привести к накоплению информации обо всех учебных действиях учеников, включая решение задач, написание сочинений и скетчей, рисование изображений и т.д., в масштабах отдельной школы, региона или страны в целом. Обработка этих данных методами искусственного интел-

лекта даёт колоссальные возможности оценки и предсказания поведения людей и различных социальных групп.

Уже сегодня системы искусственного интеллекта используются для выявления потенциально криминогенного поведения людей и установления рейтингов доверия, например, относительно банковских займов. Искусственный интеллект позволяет также автоматически оценивать переживания людей по их лицам, что может использоваться как в школах и офисах, так и при определении потенциальных источников опасного поведения в толпе. Цифровые следы личности, то есть зарегистрированные электронными устройствами действия людей, которые при помощи автоматического анализа могут стать основой выводов об индивидуально-личностных свойствах, уже в ближайшее время найдут применение, по меньшей мере, в сфере профотбора. Искусственный интеллект может быть полезен в образовании, например, для выработки индивидуальных образовательных траекторий учащихся. Вопросно-ответные системы, такие как SIRI, являются близкими родственниками уже появившихся на рынке электронных "ассистентов юриста" и "психотерапевтов". Роботы с искусственным интеллектом, способным распознавать эмоции по лицу и выдавать ответные квазиэмоциональные реакции, уже проявили себя в сфере работы с детьми-аутистами.

Всё большее внедрение в жизнь человека и общества технологий искусственного интеллекта придаёт огромную важность третьему из указанных направлений взаимодействия искусственного интеллекта и социогуманитарных наук — социогуманитарному сопровождению этих технологий.

Возьмем, например, такую сферу применения искусственного интеллекта, как автономный автомобиль, где возникает целая серия социогуманитарных проблем. К области юриспруденции относится вопрос об ответственности в случае приводящих к ущербу аварий автоматического транспорта. Далее возникают непростые дилеммы такого рода: имеет ли право автономный автомобиль задавить взрослого человека, чтобы не нанести ущерба ребёнку; можно ли свернуть на тротуар и сбить там пешехода, чтобы предотвратить неизбежный наезд на находящихся на проезжей части трёх человек и т.д.? Кто и на каком основании должен включать в программы управления автомобилями соответствующие решения?

Контроль за тем, какие действия искусственный интеллект способен совершить в будущем, может оказаться обратно пропорциональным его эффективности. Уже сегодня программы игры

в шахматы заведомо способны делать непредвиденные и более сильные ходы, чем могли бы ожидать люди. Эти программы, однако, не имеют доступа к управлению реальными механическими системами. Автомобили же, с одной стороны, позволяют удовлетворять человеческую потребность в перемещении, а с другой — представляют потенциальную угрозу жизни и здоровью. При соединении этих двух начал возникает реальная опасность неподконтрольности. Вместе с тем попытка создателей предвидеть возможные действия искусственного интеллекта ограничивает его потолок когнитивными способностями конструкторов.

Автономный автомобиль — лишь один из примеров внедрения искусственного интеллекта в нашу жизнь и далеко не самый драматичный. Ещё большей ответственности требуют технологии автоматического пилотирования самолёта, учитывая возможные последствия авиационных аварий. Об этих проблемах на сессии говорил академик С.Ю. Желтов.

Передача машинам контроля над сферами жизни, которые всегда зависели только от людей, приводит к непростым последствиям для общества. Так, обсуждается проблема уничтожения рабочих мест и возможной безработицы, вызванной автоматизацией. В ближайшие десятилетия работу могут потерять профессиональные водители такси, общественного и грузового транспорта. А если искусственный интеллект окажется способен работать в области бухгалтерского учёта и логистики, торговли и жилищного проектирования, перевода и юридического консультирования? Всё это с неизбежностью будет означать перераспределение сфер занятости, причём не вполне понятно, где экспансия искусственного интеллекта остановится. Подобный сценарий развития информационных технологий фактически грозит повторением на новом витке тех же последствий, к каким в начале XIX в. привело широкое внедрение машин, а именно протестов теряющих трудовые места работников. Два столетия назад это вылилось в формирование движения луддитов, которое на своей родине в Англии составило для правительства проблему, соизмеримую по масштабам с проходившими примерно в то же время наполеоновскими войнами. По сравнению с событиями 200-летней давности научная рефлексия над жизнью общества сегодня развита намного больше, поэтому есть возможность предвидеть подобные перспективы. Задача социогуманитарных наук в этом контексте заключается в разработке таких моделей, которые минимизируют отрицательные последствия грядущих изменений.

Технологии искусственного интеллекта могут применяться и с "тёмными" целями — для веде-

ния разрушительной войны или установления господства одних групп лиц над другими. Искусственный интеллект вызывает опасения в связи с возможностью его использования "Большим братом" для выявления неблагонадёжных и чрезмерно свободомыслящих граждан. Впрочем, пока эти технологии, скорее, используются коммерческими предприятиями в целях определения потенциального интереса людей к совершению тех или иных покупок и последующего таргетирования рекламы.

Средства влияния на людей посредством технологий искусственного интеллекта также раз-

виваются. Например, сформировалось понятие "глубинная фальсификация" (deep fakes), которое обозначает технологию, использующую глубинное научение (deep learning) в целях создания достаточно правдоподобной ложной информации (fakes). Глубинная фальсификация основывается на возможности искусственного интеллекта генерировать изображения и ролики, практически не отличимые от настоящих.

В целом искусственный интеллект вызывает к жизни силы, для обуздания которых и направления на благо человечества необходима мощная социогуманитарная проработка.

SPEECH OF CORRESPONDING MEMBER OF RAS D.V. USHAKOV

Received: 03.12.2018

Accepted: 25.12.2018

Keywords: socio-humanistic sciences, artificial intelligence, cognitive science, socio-humanistic support of artificial intelligence technologies.