

ЛИАНА ТУХВАТУЛИНА*

БЫСТРЫЕ ТЕХНОЛОГИИ И ИНЕРТНЫЕ ИНСТИТУТЫ**

О МЕСТЕ СОЦИАЛЬНО-ГУМАНИТАРНОГО ЗНАНИЯ
В ОСМЫСЛЕНИИ ПЕРСПЕКТИВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

DOI: 10.17323/2587-8719-2023-3-236-242.

Валентин Александрович Бажанов рассматривает в статье некоторые особенности развития искусственного интеллекта и больших данных, демонстрируя на множестве примеров неоднозначность этого процесса. При этом автор полагает, что преодоление возникающих противоречий требует «трансцендентализма алгоритмического типа», который связан с необходимостью

выхода за рамки жесткого *алгоритмического детерминизма* на своего рода *метауровень* для того, чтобы оценить и понять возможности и перспективы развития AI (Бажанов, 2023).

Такая позиция важна как для решения инженерных задач («очистки» базисных алгоритмов от «предрассудков и подсознательных установок программистов»), так и для философского осмысления природы и траектории развития новейших информационных технологий. Обоснование подобной метапозиции позволит, на мой взгляд, определить роль гуманитарной рефлексии в осмыслении научно-технического развития общества, а также форм проявления политической субъектности науки. Валентин Александрович не ставит вопроса о возможности

*Тухватулина Лиана Анваровна, к. филос. н., исследователь, Межрегиональная общественная организация «Русское общество истории и философии науки» (Москва), spero-meliora@bk.ru, ORCID: 0000-0002-9390-9701.

**© Тухватулина, Л. А. © Философия. Журнал Высшей школы экономики.

Благодарности: исследование выполнено при финансовой поддержке РФФ, проект № 21-18-00428 «Политическая субъектность современной науки: междисциплинарный анализ на перекрестье философии науки и философии политики» в Русском обществе истории и философии науки.

такой метапозиции в статье, поэтому в своей реплике я бы хотела остановиться на этой проблеме.

Отношение к связи между технологическими инновациями и обществом можно свести к трем точкам зрения (de Poel, 2020).

- (1) *Технологический детерминизм*, согласно которому технологии развиваются автономно, определяя при этом траекторию развития общества. Эта позиция объединяет и технопессимистов, опасющихся, например, того, что искусственный интеллект в конце концов поработит/уничтожит человечество, и технооптимистов, полагающих, что именно автономия технологий дает надежду на прогресс и благополучное будущее, поскольку главная угроза исходит именно от человека.
- (2) *Социальный детерминизм*, в рамках которого технологии рассматриваются как результат человеческой деятельности, воплощающий ценности и интересы их создателей. Согласно этой позиции, именно социальная укорененность технологий дает возможность контролировать инновации и проводить социотехнические преобразования, ориентируясь на образ желаемого будущего.
- (3) *Козволюция технологий и общества*, предполагающая, что в целом неподконтрольное развитие технологий ведет не только к непредсказуемым и непреднамеренным социальным последствиям, но и к изменению ценностей, интересов и ожиданий людей по мере технологических инноваций. В этом отношении «введение новых технологий оказывается также формой морального экспериментирования, в ходе которого мы узнаем, какие новые моральные проблемы приносят с собой новые технологии» (ibid.: 506).

Какая из названных позиций позволяет исследователю выйти на метауровень, чтобы «оценить и понять возможности и перспективы» той или иной технологии, в том числе AI?

На мой взгляд, технологический детерминизм явно не подходит для этой задачи, поскольку отводит субъекту пассивную роль наблюдателя саморазвертывающегося процесса развития технологий. Эта роль дает возможность получать знание лишь о предшествующих его стадиях, при этом существенно ограничивая возможности для прогнозирования будущего. А главное, такая позиция не предполагает влияния на характер этого развития. Оценка наблюдаемого процесса имеет здесь скорее эмоциональное значение (вызывает отчаяние или надежду), не создавая предпосылок для влияния на происходящее.

В то же время приверженность коэволюционной позиции предполагает, что мы не знаем, как именно будут меняться социальные институты вместе с технологическими инновациями, однако признаем неизбежность таких изменений. При этом тезис о коэволюции как бы изымает исследователя из ограничивающего контекста настоящего и предлагает взглянуть на процесс из точки «внеаходимости» (будущего). Однако такое перемещение все же мало дает для содержательного исследования перспектив развития технологий в силу отсутствия конкретного образа будущего.

Определенные перспективы в этой связи могут быть связаны с оптикой, которую предлагает социальный детерминизм. Как было сказано выше, технологии в его рамках рассматриваются как результат человеческой деятельности, который вбирает в себя черты его создателя. Отсюда негативные социальные эффекты и риски, сопряженные с внедрением инноваций, следует рассматривать в качестве новой проекции существовавших в обществе и ранее проблем. И действительно, многочисленные примеры, которые в статье Валентина Александровича иллюстрируют риски применения искусственного интеллекта, скорее подталкивают к выводу о том, что за новой цифровой оболочкой скрываются старые проблемы. Среди них влияние когнитивных искажений на формирование убеждений и межличностных отношений, неравенство в распределении ресурсов (экономических и информационных), а также угрозы превращения демократии в тиранию.

Как ни странно, то обстоятельство, что многие негативные эффекты цифровых технологий вовсе не представляют собой что-то принципиально новое, а лишь по-новому высвечивают глубинные социальные противоречия, несколько упрощает поиск концептуальных оснований для критики (анализа) и оценки перспектив развития технологий, что дает возможность контроля, пусть и ограниченную. Для решения этой задачи локус внимания, по-видимому, должен быть смещен с самих технологий на общество, в котором они развиваются. Критика технологий должна начинаться с социально-политической критики. Так, угрозы безопасности, связанные с утечкой персональных данных, зачастую оказываются следствием коррупции и/или всевластия силового аппарата в авторитарном государстве, которое использует технологические рычаги для проникновения в частную жизнь граждан. А использование цифровых технологий (например, в виде таргетированных бот-рассылок, о которых пишет Валентин Александрович) для агитации в рамках избирательной кампании едва ли само по себе представляет серьезную

угрозу, поскольку эта мера не сможет изменить взгляды избирателя в отсутствие каких бы то ни было реальных предпосылок для этого. Успешность такой агитации определяется все-таки не ее технологическим носителем, а социальным контекстом.

Обоснование возможности метапозиции сопряжено с разрешением дилеммы контроля над технологиями, обсуждавшейся в философии техники еще в конце XX века (Collingridge, 1980). Так, с одной стороны, контроль над той или иной технологией легко реализуем на этапе ее появления, однако в этот момент мы еще не обладаем достаточным знанием о ее внешних эффектах. Такое знание мы получим лишь после ее интеграции в общество — но на этом этапе мы окажемся неспособны ее контролировать. Отсутствие (или, по крайней мере, чрезвычайная ограниченность) контроля связано не со сложностью инженерной задачи, а с социальным контекстом, внутри которого существует та или иная технология. Организационно-административная структура, экономическая специфика, сети власти и институты (правила и способы их поддержания) во многом решают то, как проявят себя технологические инновации.

Представим ситуацию: в государстве *N* решили реформировать систему правосудия, отдав судебские функции искусственному интеллекту. Преимущество такого изменения реформаторы видят в том, что использование AI позволит нейтрализовать влияние случайных (человеческих) факторов на вынесение решений. Таким фактором может быть, например, оценка квалификации судьи в зависимости от того, какое количество его решений было впоследствии успешно опротестовано в вышестоящих инстанциях. При этом реформаторы предполагают, что решения по конкретным кейсам AI будет принимать, сопоставляя их с решениями по схожим случаям в предшествовавшей судебной практике. Однако в этой ситуации велик риск того, что искусственный интеллект будет решать так же предвзято, как и судьи. В то же время «освободить» AI от такого рода предвзятости может быть весьма непросто, поскольку процесс принятия судебных решений нельзя свести к автоматическому применению норм: он всегда требует интерпретации норм в контексте правовой культуры, укорененной в юридической практике. А значит, если обучение на предшествующем опыте неизбежно, внедрение AI может выявить своеобразный эффект колен, который воспроизводится при попытке сугубо технической модернизации права как социального института.

Я полагаю, что исток многих противоречий, связанных с расширением сферы применения AI (как, впрочем, и иных высоких технологий), — низкая скорость развития социальных институтов в сравнении со скоростью развития технологий. Эта особенность, по-видимому, характеризовала общественные отношения на всех этапах технологического овладения миром. Однако именно в связи с развитием искусственного интеллекта коллизия становится особенно драматичной. Вкладывая огромные усилия в совершенствование искусственного интеллекта, человек словно пытается обрести в нем лучшую версию себя. Искусственный интеллект должен обладать сопоставимыми с человеческими способностями (или даже превосходящими их) к анализу информации и формированию умозаключений, будучи при этом свободным от когнитивных искажений, предубеждений, эмоциональной предвзятости. Надежды, связанные с искусственным интеллектом, — это в конечном счете надежды на более справедливое, свободное и гуманное общество. Эти ожидания можно сравнить с ожиданиями родителя, который верит, что его ребенок будет успешнее/умнее/счастливее его, и игнорирует при этом колоссальное влияние его собственной личности на судьбу ребенка. Подобно тому как едва ли не единственный локус контроля родителя над будущим ребенка определяется силой личного примера, так и в случае отношений между человеком и его детищем (искусственным интеллектом) зона ответственности человека связана с самосовершенствованием — с модернизацией социальных институтов не путем их грубой технологизации, а через осмысление и преодоление фундаментальных противоречий, возникших задолго до появления высоких технологий.

В связи с этим критическая функция социогуманитарного знания может быть важна для преодоления основополагающего противоречия постнормальной науки,

когда решения не терпят отлагательства и могут иметь весьма серьезные последствия, но будущее крайне неопределенно и результаты предпринятых действий можно предсказать только в самом общем плане, если это вообще оказывается возможным (Бажанов, 2023).

Непредсказуемость долгосрочных последствий принимаемых сегодня решений определяет политическую субъектность современной науки, возросшую ответственность ученых за социальные эффекты от разработанных на базе научного знания технологий. Основной трудностью в связи с осмыслением политической субъектности науки становится именно фокусировка на будущем. Для инженера и ученого будущее

как понятие плохо концептуализируется, а будущее как феномен полно неопределенных и малопрогнозируемых рисков. Однако сила гумани-тария состоит в его способности предвидеть общие контуры будущего, которые в силу инертности социальных институтов в горизонте десяти-летий будут определяться происходящим здесь и сейчас. Это обстоятель-ство подталкивает к заключению, что неопределенность будущего как трудность на пути осмысления политической субъектности науки может быть компенсирована социально-критической функцией гуманитарного знания. На мой взгляд, ответственность за эту задачу ставит гума-нитариев в один ряд с инженерами, поскольку успешность внедрения инноваций будет определяться мерой учета инертности институтов.

ЛИТЕРАТУРА

- Бажанов В. А.* Искусственный интеллект, технологии Big Data (больших дан-ных) и особенности современного политического процесса // *Философия : Журнал Высшей школы экономики*. — 2023. — Т. 7, № 3.
- Collingridge D.* *The Social Control of Technology*. — London : Frances Pinter, 1980.
- de Poel I. van.* Three Philosophical Perspectives on the Relation between Tech- nology and Society, and How they Affect the Current Debate about Artificial Intelligence // *Human Affairs*. — 2020. — No. 30. — P. 499–511.

Tukhvatulina, L. A. 2023. “Bystryye tekhnologii i inertnyye instituty [Fast Technologies and Slow Institutions]: o meste sotsial’no-gumanitarnogo znaniya v osmyslenii perspektiv iskusstvennogo intellekta [About the Place of Social and Humanitarian Knowledge in Understanding of the Prospects of AI]” [in Russian]. *Filosofiya. Zhurnal Vysshey shkoly ekonomiki* [*Philosophy. Journal of the Higher School of Economics*] 7 (3), 236–242.

LIANA TUKHVATULINA

PHD IN PHILOSOPHY, RESEARCHER

RUSSIAN SOCIETY FOR THE HISTORY AND PHILOSOPHY OF SCIENCE (MOSCOW, RUSSIA);

ORCID: 0000-0002-9390-9701

FAST TECHNOLOGIES AND SLOW INSTITUTIONS

ABOUT THE PLACE OF SOCIAL AND HUMANITARIAN KNOWLEDGE IN UNDERSTANDING OF THE PROSPECTS OF AI

DOI: 10.17323/2587-8719-2023-3-236-242.

REFERENCES

- Bazhanov, V. A. 2023. “Iskusstvennyy intellekt, tekhnologii Big Data (bol’shikh dannykh) i osobennosti sovremennogo politicheskogo protsesssa [Artificial Intelligence, Big Data Technologies, and Features of Modern Politics]” [in Russian]. *Filosofiya* [*Philosophy*]: *Zhurnal Vysshey shkoly ekonomiki* [*Journal of the Higher School of Economics*] 7 (3).
- Collingridge, D. 1980. *The Social Control of Technology*. London: Frances Pinter.
- de Poel, I. van. 2020. “Three Philosophical Perspectives on the Relation between Technology and Society, and How they Affect the Current Debate about Artificial Intelligence.” *Human Affairs*, no. 30, 499–511.