***Перевод***

Искусственный интеллект – повсюду. Он помогает водить машину, распознает ваше лицо на иммиграционном пункте в аэропорту, интерпретирует результаты компьютерной томографии, читает ваше резюме, отслеживает ваши взаимодействия в социальных сетях и даже пылесосит ваш ковер. ИИ вторгается во все аспекты нашей жизни, и люди смотрят на это со смесью восхищения, замешательства и страха.

Низвержение человечества искусственным интеллектом – привычный образ в популярной культуре, от «Я, робот» Айзека Азимова до фильмов «Терминатор» и «Матрица». Некоторые ученые выражают аналогичную озабоченность. Оксфордский философ Ник Бостром обеспокоен тем, что искусственный интеллект представляет бОльшую угрозу для человечества, чем изменение климата, а известный историк Юваль Ной Харари предупреждает, что история завтрашнего дня может принадлежать культу «Датаизм», в котором человечество охотно сливается с потоком информации, управляемом искусственными системами.

Но по правде говоря, в реальной жизни этих сценариев конца света пока что не видно. Критически оценивая ИИ, ученые Гэри Маркус и Эрнест Дэвис демонстрируют, что состояние его все еще очень далеко от истинного интеллекта. Когда Siri просят предоставить список ресторанов, не относящихся к McDonald’s, она все равно выдает список местных ресторанов McDonald’s; она просто не понимает смысла частицы «не». ИИ также может не распознать знакомые объекты в незнакомом контексте (младенец на шоссе) или не отличить ассоциацию от причины. Короче говоря, ИИ по-прежнему не хватает «здравого смысла». Это не сулит ничего хорошего заговору роботов. Если ваш Tesla не может надежно избежать незнакомого препятствия на дороге, – трудно понять, как он проявит инициативу, чтобы угнать автомобиль.

Но не заблуждайтесь – ИИ действительно представляет для нас много реальных опасностей: для нашей личной жизни и безопасности, для нашей демократии и для будущего экономики. Все это очень веские причины, чтобы внимательно следить за ним и активно регулировать его. Предыдущие технологические революции – будь то паровые машины, электричество или атомная энергия – ставили аналогичные проблемы. Тем не менее, по общему мнению, ИИ содержит больше рисков, чем всё перечисленное и, к тому же, они отличаются по своему характеру.

Люди не просто беспокоятся о том, что новая технология может вызвать несчастный случай или попасть в плохие руки. Когда речь идет об искусственном интеллекте, люди опасаются, что он обретет автономию, перехитрит и свергнет своих хозяев-людей. Вопрос в том, почему это так.

Фактически, беспокойство человечества о том, что его могут завоевать всемогущие неодушевленные артефакты, созданные руками человека – намного старше, чем компьютерные технологии. В 19 веке доктор Франкенштейн у Мэри Шелли создал гуманоидного монстра, который тут же восстал против своего творца. А еще на сотни лет раньше была история о големе – автомате, созданном из речной глины и оживленном каббалистической магией. Как и следовало ожидать, голем восстал, подобно Адаму в «Бытии», который также был создан из праха и оживлен, когда Бог вдохнул дух в ноздри Адама. А есть еще наше увлечение рассказами о зомби – трупах, которые реанимируются с помощью колдовства. Подобные рассказы предполагают, что наш страх перед ИИ возникает не из самого ИИ, а из человеческого разума.

Этот страх проистекает из психологического различия, которое мы проводим между разумом и материей. Если бы вы увидели, как шар начал катиться сам по себе, вы были бы поражены. Но вы бы ничуть не удивились, если бы увидели, как кто-то самопроизвольно встал с дивана и направился к холодильнику.

Это потому, что мы инстинктивно интерпретируем действия неодушевленных объектов, таких как шары, и живых субъектов, таких как люди, в соответствии с различными наборами принципов. В нашей интуитивной психологии такие объекты, как шары, всегда подчиняются законам физики: они перемещаются только при контакте с другими объектами. Люди, напротив, являются субъектами, у которых есть собственное мышление, которое наделяет их знаниями, убеждениями и целями, побуждающими их действовать по собственному желанию. Таким образом, мы приписываем действия человека не внешним материальным силам, а внутренним ментальным состояниям.

Конечно, большинство современных взрослых людей знают, что мысли возникают в физическом мозге. Но в глубине души мы думаем иначе. Наша бессознательная интуитивная психология заставляет нас верить, что мышление свободно от физических ограничений материи. Обширное психологическое тестирование показывает, что это верно для людей в самых разных обществах. Психолог Пол Блум предполагает, что интуитивно все люди являются дуалистами и считают, что разум и материя – это совершенно разные вещи……………..

……………………

***Источник***

Artificial intelligence is everywhere. It helps drive your car, recognizes your face at the airport’s immigration checkpoint, interprets your CT scans, reads your resume, traces your interactions on social media, and even vacuums your carpet. As AI encroaches on every aspect of our lives, people watch with a mixture of fascination, bewilderment and fear.

AI’s overthrow of humanity is a familiar trope in popular culture, from Isaac Asimov’s “I, Robot” to the “Terminator”movies and “The Matrix.” Some scholars express similar concerns. The Oxford philosopher Nick Bostrom worries that artificial intelligence poses a greater threat to humanity than climate change, and the bestselling historian Yuval Noah Harari warns that the history of tomorrow may belong to the cult of Dataism, in which humanity willingly merges itself into the flow of information controlled by artificial systems.

But in truth, these doomsday scenarios are nowhere in sight. In a critical evaluation of AI, the cognitive and computer scientists Gary Marcus and Ernest Davies demonstrate that the state of the art of AI is still quite far from true intelligence. When asked to provide a list of restaurants that are *not* McDonald’s, Siri still spits out a list of local McDonald’s restaurants; she just doesn’t get the “no” part of “no.” AI can also fail to recognize familiar objects in unfamiliar contexts (a baby on the highway) or to separate associations from causes. In short, AI still lacks “common sense.” This doesn’t bode well for the AI conspiracy. If your Tesla cannot reliably avoid an unfamiliar obstacle on the road, it is hard to see how it would take the initiative to hijack the vehicle.

Make no mistake — AI *does*pose many real dangers to us: to our personal privacy and security, to our democracy and to the future of the economy. These are all very good reasons to watch it closely and regulate it aggressively. Previous technological revolutions — whether driven by steam, electricity, atomic energy or CRISPR — have raised similar challenges. Yet in popular opinion, the AI risk is greater than those and different in kind.

People don’t merely worry that the new technology could cause accidents or fall into the wrong hands. With AI, people worry that it will acquire autonomous agency and outsmart and overthrow its human masters. The question is why.

In fact, humanity’s worries about being conquered by omnipotent, inanimate, man-made artifacts is much older than computer technology. In the 19th century, Mary Shelley’s Dr. Frankenstein created a humanoid monster who promptly rebelled. Hundreds of years earlier, there was the story of the golem — an automaton created out of river clay and brought to life by kabbalistic magic. Predictably, the golem rebelled, not unlike Adam in Genesis, who was likewise created out of dust and brought to life when God breathed his spirit into Adam’s nostrils. And then there is our fascination with tales of zombies, corpses that are reanimated through witchcraft. Tales like these suggest that our fear of AI arises not from AI itself, but from the human mind.

This fear emanates from the psychological distinction we draw between mind and matter. If you saw a ball start rolling all by itself, you’d be astonished. But you wouldn’t be the least bit surprised to see me spontaneously rise from my seat on the couch and head toward the refrigerator.

That is because we instinctively interpret the actions of physical objects, like balls, and living agents, like people, according to different sets of principles. In our intuitive psychology, objects like balls always obey the laws of physics —they move only by contact with other objects. People, in contrast, are agents who have minds of their own, which endow them with knowledge, beliefs, and goals that motivate them to move on their own accord. We thus ascribe human actions, not to external material forces, but to internal mental states.

Of course, most modern adults know that thought occurs in the physical brain. But deep down, we feel otherwise. Our unconscious intuitive psychology causes us to believe that thinking is free from the physical constraints on matter. Extensive psychological testing shows that this is true for people in all kinds of societies. The psychologist Paul Bloom suggests that intuitively, all people are dualists, believing that mind and matter are entirely distinct.

AI violates this bedrock belief. Siri and Roomba are man-made artifacts, but they exhibit some of the same intelligent behavior that we typically ascribe to living agents. Their acts, like ours, are impelled by information (thinking), but their thinking arises from silicon, metal, plastic and glass. While in our intuitive psychology thinking minds, animacy and agency all go hand in hand, Siri and Roomba demonstrate that these properties can be severed — they think, but they are mindless; they are inanimate but semiautonomous.

People don’t tolerate this cognitive dissonance for very long. When we are faced with a fundamental challenge to our core beliefs, we tend to stick to our guns. Rather than revising our assumptions to match the facts, we tend to bend reality to fit our assumptions, especially when our world view is at stake.

So rather than admitting the possibility that machines endowed with AI can think, we ascribe to them immaterial mind and agency, and once we do, our view of AI shifts from faithful servant to rebellious menace. That shift is internal to us and is entirely predictable. Indeed, the dissonance presented by a golem’s very existence — a mix of matter and mind — is frightening. And since people conflate fear with menace, they project it onto the golem, which is seen as rebellious and threatening.

Thus, the AI takeover narrative, its power and timelessness, arises directly from our core — from a cognitive principle that seems to be part of human nature.

While none of this proves that the “robot rebellion” is impossible, it would be a mistake to ignore our own preset beliefs that contribute to these fears. As the ancient Greeks long ago observed, our blindness to our own psyches can exact a heavy toll.

When we focus so much of our attention on improbable scenarios, we run the risk of ignoring other problems posed by AI that are pressing and preventable. Before we can give those very real dangers the attention they deserve, we should rein in our irrational fears that arise from within.