Опасность не в том, что ИИ принимает решения. А в том, что мы ему позволяем.

Отнимет ли ИИ у человека работу? Движемся ли мы к будущему, где ИИ принимает каждое решение, а человечество становится видом без самостоятельности, просто следуя нашему сверхразумному цифровому повелителю?

Это реальный страх, и он может быть оправдан: система с собственными целями заслуживает беспокойства. Но это не та опасность, которой мы сейчас подвергаемся больше всего. Задолго до того, как ИИ что-то захватит, мы отдадим это, потому что так проще.

Два вида конфликтов

Теория управления различает два типа конфликтов. Сходящийся конфликт разрешается через обсуждение, например, споры по терминологии или противоречивым фактам. Уточните термин или проведите небольшое исследование, и люди придут к общему пониманию.

Разходящийся конфликт нельзя разрешить таким образом: обе стороны правы и взаимоисключают. Например, стоит ли нам стремиться к увеличению доли рынка, что потребует увеличения предложения и снижения цен, или стремиться к увеличению прибыли, что потребует сокращения предложения и увеличения цен? И увеличение доли рынка, и увеличение прибыли — законные цели, но мы не можем добиваться обеих целей одновременно. Это структурный компромисс, а не разрыв в знаниях. Другой пример конфликта — конфликт интересов. Выигрыш для всех — это утопическое ожидание, и чаще всего, по крайней мере в краткосрочной перспективе, кто-то выигрывает, а кто-то проигрывает. Такова природа, и никакие разговоры этого не изменят.

Где ИИ помогает, а где нет

ИИ отлично справляется с конвергентными конфликтами: находит подходящее слово, выявляет недостающую информацию. Это проблема знаний.

Расходящиеся конфликты — нет. Никакое количество данных не решит противоречивые, взаимоисключающие цели или конфликтующие интересы. Задача — не оптимизировать одну сторону; это найти правильный баланс, и сегодняшний баланс не будет завтрашним. Это как вождение, постоянно корректируя руль. Эта коррекция — решение суждения, без правильного ответа,

только риск, и её нельзя передать тому, кто лучше всего умеет максимально использовать одну переменную.

Недавний пример

Недавно Фаучи снова был вызван в Конгресс по поводу решений во время пандемии. Стоит быть точным: как директор NIAID, он имел реальные полномочия над финансированием исследований и приоритетами институтов, но никогда не имел полномочий устанавливать национальные мандаты по общественному здравоохранению. В рабочей группе Белого дома по коронавирусу, которую возглавлял Трамп и HHS, он был ведущим научным советником, превращая эпидемиологию в политические варианты, а не определяя политику. Что касается важного решения, независимо от того, будет ли страна локдаун, он был советником, а не принимающим решения, именно такую роль ИИ будет играть для нас.

Как академик и государственный администратор, Фаучи смотрел на COVID с одной точки зрения: минимизировать риск того, что больницы будут переполнены. Это была законная цель, хотя её оптимизация означала не взвешивать социальные издержки локдаунов, экономический удар или цену личной свободы. Это не была ошибка Фаучи; взвешивать эти компромиссы никогда не было его обязанностью.

Это была власть Трампа. У Трампа была власть; Фаучи только имел влияние. Но Трамп вручил Фаучи трибуну и воспринял его рекомендации как устоявшийся факт, не потому что Фаучи переступил границу, а потому, что его пригласили, и мало кто отказывается от настоящего лидерства, когда его предлагают. Это лестно, подпитывает эго. Отречение даётся легко с обеих сторон: человек с властью избегает жёсткого решения, а тот, кому передали трибуну, редко возвращает её.

Более глубокая ошибка заключалась в том, что я не слушал Фаучи. Это было просто слушание Фаучи. Настоящий баланс достигается в том, чтобы объединить противоречивые точки зрения в одной комнате: эпидемиолог выступает за локдаун, бизнес-лидеры — о экономических издержках, психологи — о последствиях изоляции, защитники гражданских свобод — за личную свободу. Этот разговор неудобен, но именно так на самом деле находит баланс. Всё это не является критикой Фаучи, который выполнил свою работу. Провал был на Трампе, который никогда не делал более сложную работу — разобрался с конфликтом и не принял необходимое решение о правильном подходе. Вместо этого он уступил одному голосу и теперь обвиняет его, потому что ему не понравилось, как всё сложилось.

Настоящая опасность

Эта закономерность появилась ещё до появления ИИ. Именно поэтому компании платят McKinsey. Генеральный директор, сталкивающийся с политически опасным вызовом, например, увольнениями, на самом деле не покупает анализ; он покупает прикрытие. Пусть консультанты «рекомендуют» то, что он уже подозревал, и это становится «McKinsey велела нам это сделать», а не «я решил». Часто именно этот продукт продаётся.

ИИ вот-вот станет самой дешёвой и быстрой версией того же сервиса. Риск здесь не в том, что Fauci, McKinsey или ИИ захватят власть. Это человек с властью, который тихо отступает и позволяет советнику вести инициативу, потому что решение сложное, советник уверен, и если что-то пойдёт не так, виноват кто-то другой.

ИИ будет становиться всё лучше в конвергентных конфликтах, вызванных пропущенной информацией. Он никогда не будет правильным для конфликтов с расходами: это проблемы баланса, а баланс требует человека, готового сидеть с противоречивыми точками зрения и принимать решение без правильного ответа. Это не та работа, которую ИИ заберёт у нас. Это работа, которую мы передадьм, по одному удобному решению за раз.