Рациональные и убедительные доводы Ника Бострома об опасности создания искусственного интеллекта заставят задуматься даже заядлых скептиков.

Евгений Касперский,
генеральный директор «Лаборатории Касперского»

Ник Бостром

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

ЭТАПЫ. УГРОЗЫ. СТРАТЕГИИ

Оглавление

Предисловие партнера

> …У меня есть один знакомый, — сказал Эдик. — Он утверждает, будто человек — промежуточное звено, необходимое природе для создания венца творения: рюмки коньяка с ломтиком лимона.
>
> *Аркадий и Борис Стругацкие. Понедельник начинается в субботу*

Компьютеры, а точнее алгоритмы, опирающиеся на непрерывно растущие вычислительные мощности, лучше людей играют в шахматы, шашки и нарды. Они очень неплохо водят самолеты. Они смогли пройти тест Тьюринга, убедив судей в своей «человечности». Однажды таксист в Дублине — городе, где расположены европейские штаб-квартиры многих глобальных IT-компаний, — сказал мне, что приветствует бурное развитие технологического сектора своей страны, но потом с сожалением добавил: «Одна беда — из-за этих умных ребят довольно скоро таксисты будут не нужны». Автомобили без водителей, управляемые компьютерами, уже проходят испытания на обычных дорогах в нескольких странах. По мнению философа Ника Бострома, чью книгу вы держите в руках, — все это звенья одной цепи и довольно скоро из-за развития компьютерных технологий нам всем, человеческому роду, может прийти конец.

Автор считает, что смертельная угроза связана с возможностью создания искусственного интеллекта, превосходящего человеческий разум. Катастрофа может разразиться как в конце XXI века, так и в ближайшие десятилетия. Вся история человечества показывает: когда происходит столкновение представителя нашего вида, человека разумного, и любого другого, населяющего нашу планету, побеждает тот, кто умнее. До сих пор умнейшими были мы, но у нас нет гарантий, что так будет длиться вечно.

Ник Бостром пишет, что если умные компьютерные алгоритмы научатся самостоятельно делать еще более умные алгоритмы, а те, в свою очередь, еще более умные, случится взрывной рост искусственного интеллекта, по сравнению с которым люди будут выглядеть приблизительно как сейчас муравьи рядом с людьми, в интеллектуальном смысле, конечно. В мире появится новый, хотя и искусственный, но сверхразумный вид. Неважно, что ему «придет в голову», попытка сделать всех людей счастливыми или решение остановить антропогенное загрязнение мирового океана наиболее эффективным путем, то есть уничтожив человечество, — все равно сопротивляться этому у людей возможности не будет. Никаких шансов на противостояние в духе кинофильма про Терминатора, никаких перестрелок с железными киборгами. Нас ждет шах и мат — как в поединке шахматного компьютера «Дип Блю» с первоклассником.

За последнюю сотню-другую лет достижения науки у одних пробуждали надежду на решение всех проблем человечества, у других вызывали и вызывают безудержный страх. При этом, надо сказать, обе точки зрения выглядят вполне оправданными. Благодаря науке побеждены страшные болезни, человечество способно сегодня прокормить невиданное прежде количество людей, а из одной точки земного шара можно попасть в противоположную меньше чем за сутки. Однако по милости той же науки люди, используя новейшие военные технологии, уничтожают друг друга с чудовищной скоростью и эффективностью.

Подобную тенденцию — когда быстрое развитие технологий не только приводит к образованию новых возможностей, но и формирует небывалые угрозы, — мы наблюдаем и в области информационной безопасности. Вся наша отрасль возникла и существует исключительно потому, что создание и массовое распространение таких замечательных вещей, как компьютеры и интернет, породило проблемы, которые было бы невозможно вообразить в докомпьютерную эру. В результате появления информационных технологий произошла революция в человеческих коммуникациях. В том числе ею воспользовались разного рода киберпреступники. И только сейчас человечество начинает постепенно осознавать новые риски: все больше объектов физического мира управляются с помощью компьютеров и программного обеспечения, часто несовершенного, дырявого и уязвимого; все большее число таких объектов имеют связь с интернетом, и угрозы

кибермира быстро становятся проблемами физической безопасности, а потенциально — жизни и смерти.

Именно поэтому книга Ника Бострома кажется такой интересной. Первый шаг для предотвращения кошмарных сценариев (для отдельной компьютерной сети или всего человечества) — понять, в чем они могут состоять. Бостром делает очень много оговорок, что создание искусственного интеллекта, сравнимого с человеческим разумом или превосходящего его, — искусственного интеллекта, способного уничтожить человечество, — это лишь вероятный сценарий, который может и не реализоваться. Конечно, вариантов много, и развитие компьютерных технологий, возможно, не уничтожит человечество, а даст нам ответ на «главный вопрос жизни, Вселенной и всего такого» (возможно, это и впрямь окажется число 42, как в романе «Автостопом по Галактике»). Надежда есть, но опасность очень серьезная — предупреждает нас Бостром. На мой взгляд, если вероятность такой экзистенциальной угрозы человечеству существует, то отнестись к ней надо соответственно и, чтобы предотвратить ее и защититься от нее, следует предпринять совместные усилия в общемировом масштабе.

Завершить свое вступление хочется цитатой из книги Михаила Веллера «Человек в системе»:

> Когда фантастика, то бишь оформленная в образы и сюжеты мысль человеческая, долго и детально что-то повторяет — ну так дыма без огня не бывает. Банальные голливудские боевики о войнах людей с цивилизацией роботов несут в себе под шелухой коммерческого смотрива горькое зернышко истины.
>
> Когда в роботы будет встроена передаваемая программа инстинктов, и удовлетворение этих инстинктов будет встроено как безусловная и базовая потребность, и это пойдет на уровень самовоспроизводства — вот тогда, ребята, кончай бороться с курением и алкоголем, потому что будет самое время выпить и закурить перед ханой всем нам.

Евгений Касперский,
генеральный директор «Лаборатории Касперского»

Обратите внимание!

[Перегляд позитивного мислення]

[Психология. Люди, концепции, эксперименты]

[Парадокс страсти. Она его любит, а он ее нет]

[Санта действительно существует? Философское расследование]

[Голая статистика. Самая интересная книга о самой скучной науке]

[Чарунки долі]

[В активном поиске]

[Женщина. Руководство для мужчин]

[Книга радости. Как быть счастливым в меняющемся мире]