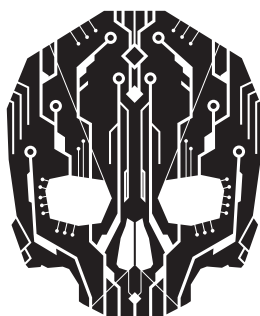


КЭТИ
О'НИЛ

КЭТИ
О'НИЛ



УБИЙСТВЕННЫЕ
БОЛЬШИЕ
ДАННЫЕ



Издательство АСТ
Москва

УДК 33:51
ББК 65+22.1
О-11

Серия «Цифровая экономика и цифровое будущее»

Cathy O'Neil

Weapons of Math Destruction:

How Big Data Increases Inequality and Threatens Democracy

Перевод с английского *Виктории Дегтяревой*

Печатается с разрешения литературных агентств
William Morris Endeavor Entertainment, LLC и Andrew Nurnberg.

О'Нил, Кэти.

О-11 Убийственные большие данные. Как математика превратилась в оружие массового поражения / Кэти О'Нил; [перевод с английского В. Дегтяревой]. — Москва: Издательство АСТ, 2018. — 320 с. — (Цифровая экономика и цифровое будущее).

ISBN 978-5-17-982583-8

Математические алгоритмы с каждым днем все сильнее подчиняют себе нашу жизнь. Более того: по мнению автора книги, профессора математики и финансового аналитика, эти алгоритмы уже превратились в опасное оружие в руках государства и корпораций — и это оружие нацелено в первую очередь на самые бедные и незащищенные слои населения. Новейшие математические приложения, с помощью которых банки и страховые компании отслеживают каждый наш шаг, претендуют на полную объективность, однако на самом деле в них заложены те же предрассудки и предубеждения, что свойственны их создателям — далеким от совершенства человеческим существам. При этом скрытые принципы работы математических моделей и их тайные критерии охраняются как величайшая коммерческая тайна, а их вердикты, подчас очевидно ошибочные и явно вредные, считаются окончательными и обжалованию не подлежат. Добро пожаловать в прекрасный новый мир — мир убийственных Больших данных!

УДК 33:51
ББК 65+22.1

© Cathy O'Neil, 2016, 2017
© Перевод. В. Дегтярева, 2017
© Издание на русском языке AST Publishers, 2018

Содержание

Благодарности	9
Вступление	11
Компоненты бомбы: что такое модель?	27
Посттравматический синдром: мое путешествие к утрате иллюзий	49
Гонка вооружений: поступление в колледж	72
Машина пропаганды: интернет-реклама.	96
Жертвы среди мирного населения: справедливость в эпоху Больших данных	116
К службе негоден: в поисках работы	143
Нервное расстройство: на работе	165
Небоевые потери: кредитная политика	188
В поисках укрытия: получение страховки	212
Гражданин под прицелом: общественная жизнь	236
Заключение	260
Примечания	286

Посвящается всем аутсайдерам

Благодарности

Спасибо моим мужу и детям за их невероятную поддержку. Также спасибо Джону Джонсону, Стиву Уолдману, Маки Инаде, Беки Джефф, Аарону Абрамсу, Джули Стил, Карен Бернс, Мэтту ЛаМаните, Марте Пун, Лизе Рэдклифф, Льюсу Дэниелу и Мелиссе Билски. И, наконец, спасибо людям, без которых эта книга не могла бы родиться: Лоре Стросфелд, Аманде Кук, Эмме Берри, Джордану Элленбергу, Стивену Бейкеру, Джею Мэнделу, Сэму Кансон-Бенанаву и Эрни Дэвису.

Вступление

В детстве я любила высматривать из окна машины номерные знаки других автомобилей. Я делила их на базовые элементы — числа, из которых они состояли. $45=3\times3\times5$. Этот процесс, который называется разложением на множители, был моим любимым познавательным времяпрепровождением. Математика уже тогда очень меня интересовала, и особенно я была заинтригована простыми числами.

Моя любовь к математике постепенно переросла в страсть. В четырнадцать лет я отправилась в математический лагерь и вернулась домой, прижимая к груди кубик Рубика. Математика дала мне аккуратное, уютное убежище, где можно было укрыться от хаотичности реального мира. Математика двигалась вперед, ее поле знаний постоянно расширялось, от доказательства к доказательству. Я чувствовала, что могу сделать и свой вклад. Я изучала математику в колледже, и степень доктора философии получила тоже в области математики. Темой моей диссертации была алгебраическая теория чисел — область, основанная на разложении на множители, которое так интересовало меня в детстве. В конце концов я стала профессором в Барнард-колледже, математический факультет которого объединен с Колумбийским университетом.

А затем я резко изменила свою жизнь: ушла из колледжа и устроилась на работу специалистом по количественному анализу в один из крупнейших хедж-фондов — фонд *David E. Shaw*. Уйдя из академического заведения в мир финансов, я перенесла математику из абстрактной теории в практику. Операции, которые мы производили с цифрами, воплощались в триллионы долларов, перетекающих с одного счета на другой. Сначала я была ошеломлена и взбудоражена работой в этой новой лаборатории глобальной экономики. Но осенью 2008 года, не успевшая проработать в этой области год с небольшим, экономика начала рушиться.

Это крушение продемонстрировало, что математика, когда-то бывшая моим убежищем, не только оказалась глубоко вовлечена в мировые проблемы, но и послужила причиной многих из них. Жилищный кризис, коллапс основных финансовых институтов, рост безработицы — все это произошло благодаря содействию и подстрекательству со стороны математиков, выдававших волшебные формулы. Более того — благодаря невероятной мощи, которую я так любила, математика смогла объединиться с технологиями, что привело к умножению хаоса и несчастий, а также к росту эффективности и масштабности систем, которые я теперь считаю ущербными.

Если бы мы проявили больше здравомыслия, мы отступили бы в тот момент на шаг назад, выяснили, как именно математика была использована во вред и что нужно сделать, чтобы предотвратить подобную катастрофу в будущем. Но вместо этого сразу после кризиса появились еще более крутые математические технологии, распространившиеся на новые области. С помощью этих технологий мы круглосуточно перерабатывали петабайты информации, большая часть которой была взята из социальных сетей или с сайтов интернет-магазинов. И анализ этих данных все больше сосредоточивался не

на движениях глобальных финансовых рынков, а на людях — то есть на нас с вами. Математики и статистики изучали наши желания, перемещения и финансовые возможности. Они прогнозировали нашу кредитоспособность и оценивали наш потенциал в качестве студентов, работников, любовников, преступников-рецидивистов.

Это была экономика Больших данных, и она сулила потрясающие возможности. Компьютерная программа могла просмотреть тысячи резюме или кредитных заявок и за одну-две секунды разложить их по аккуратным спискам, поместив наверх самых многообещающих кандидатов. Это не только сэкономило время, но и считалось справедливым и объективным. Ведь в основе этой системы — не полные предубеждений человеческие существа, копающиеся в бумажных кипах, а машины, обрабатывающие бесстрастные цифры. Примерно к 2010 году математика утвердила себя в человеческих делах с небывалым ранее запасом прочности — и люди в большинстве своем это приветствовали.

Однако я увидела здесь проблему. Усиленные математикой приложения, на которых работала экономика данных, были основаны на выборе, сделанном несовершенными человеческими существами. Иногда выбор, без сомнения, делался из лучших побуждений. Но многие из этих моделей вставляли предубеждения, непонимание и предвзятость, свойственные каждому живому человеку, в программное обеспечение, которое во все большей степени управляло нашими жизнями. Подобно богам, эти математические модели были загадочными, а принцип их работы — непонятным ни для кого, кроме высшей касты — математиков и специалистов в области компьютерных наук. Их вердикты, даже неправильные и наносящие вред, оказывались вне рамок обсуждения и за пределами обжалования. И в их работе проявлялась тенденция к наказанию бедных и угнетенных слоев населения

в нашем обществе, в то время как богатые становились еще богаче.

Я подобрала название для таких вредоносных моделей: оружие математического поражения, или сокращенно ОМП. Сейчас мы с вами разберем один пример, и я продемонстрирую разрушительные возможности этого оружия.

Как часто бывает, все началось с благой цели. В 2007 году Адриан Фенти, новый мэр Вашингтона, решительно взялся за улучшение недостаточно успешных школ. Перед ним стояла непростая задача: в то время половина старшеклассников едва дотягивала до конца девятого класса и лишь 8 % восьмиклассников осваивали соответствующую их возрасту программу по математике. Фенти нанял реформатора в области образования Мишель Ри и назначил ее на высокую должность председателя городского школьного совета Вашингтона.

Основной причиной, по которой ученики не получали знаний в достаточном количестве, считалось низкое качество преподавания. Поэтому в 2009 году Ри привела в действие план по «выкорчевыванию» плохих учителей. Этот процесс вскоре распространился на всю страну, затронув районы с наиболее проблемным образованием. И с точки зрения системного проектирования план был идеальным: давайте оценим учителей, избавимся от самых плохих и поместим самых лучших туда, где они смогут принести больше всего пользы. Говоря языком специалистов по анализу данных, это «оптимизирует» школьную систему, предположительно обеспечив лучшие результаты для детей.

Кто стал бы с этим спорить, кроме самих «плохих» учителей? Ри разработала инструмент оценки учителей под названием IMPACT, и к концу 2009–2010 учебного года федеральный округ уволил всех преподавателей, чьи результаты оказались в нижних 2 %. К концу следующего

учебного года еще 5 %, то есть 205 учителей, были уволены из школ.

У Сары Высоцки, учительницы пятого класса, казалось бы, не было повода для беспокойства. Она преподавала в средней школе Макфарланда всего два года, но уже удостоилась высоких оценок своей работы и от директора школы, и от родителей своих учеников. В одном отзыве подчеркивалось ее внимание к детям, в другом было отмечено, что «она — одна из лучших учительниц, с которыми мне когда-либо доводилось сталкиваться».

Однако в конце 2010–2011 учебного года Высоцки получила крайне низкий результат в тесте IMPACT. Проблема заключалась в новой системе подсчета, известной как моделирование коэффициента улучшения знаний учеников, которая оценивала эффективность преподавания математики и языковых навыков. Этот результат, выданный алгоритмом, составлял половину общей оценки Сары Высоцки — и он перевесил положительные оценки со стороны школьной администрации и местного сообщества. У округа не осталось выбора, кроме как уволить ее вместе с другими 205 учителями, результат теста у которых оказался ниже минимального порога.

Это не выглядело ни охотой на ведьм, ни сведением счетов. В таком подходе городского школьного совета есть логика. В конце концов, администраторы могут просто дружить с очень плохими учителями. Они могут восхищаться их стилем или лояльностью. Плохие учителя вполне могут производить впечатление хороших. И в Вашингтоне, как и во многих других школьных округах, минимизировали фактор человеческой предвзятости и сосредоточились на результатах, основанных на объективных данных: достижения детей в математике и чтении. Официальные лица пообещали, что цифры будут говорить сами за себя, что цифры более справедливы.