

Algo. Rules

Правила разработки
алгоритмических систем

Преамбула

Алгоритмические системы внедряются во всё большее число областей, от которых зависят решения, оказывающие на нашу жизнь значимое воздействие. Они предлагают возможности, но сопряжены и с рисками. Только мы можем позаботиться о том, чтобы алгоритмические системы разрабатывались бы на благо общества. Индивидуальные и коллективные свободы и права, установленные в концепции «прав человека», с помощью алгоритмических систем должны быть укреплены, а не ограничены. Правила, разработанные для защиты этих норм, должны стать частью реальности. Для достижения этой цели вместе с различными экспертами и заинтересованной общественностью мы разработали следующие правила Algo.Rules.

Что такое Algo.Rules?

Правила Algo.Rules представляют собой каталог формальных критериев для разработки и надзора за алгоритмическими системами в интересах общества. Они определяют этические основы, а также правовые рамки для реализации систем. Эти критерии должны быть изначально интегрированы при разработке любой системы и, следовательно, внедрены в нее по умолчанию. Отдельные Algo.Rules взаимосвязаны, их следует рассматривать как единый комплекс. Мы приглашаем заинтересованные стороны и экспертов для дальнейшего развития Algo.Rules, их принятия, адаптации, расширения и, прежде всего, для изучения возможностей применения на практике. Динамические по своему дизайну, Algo.Rules должны быть отлажены, особенно с точки зрения практической реализации.

На кого нацелены Algo.Rules?

Правила Algo.Rules адресованы всем лицам, которые оказывают существенное влияние на создание, разработку, программирование, использование или воздействие алгоритмических систем, а также всем, кто заказал их разработку или интеграцию. Это подразумевает, в частности (но не исключительно):

- Людей, которые исследуют разработку и использование алгоритмических систем;
- Людей, которые собирают, сопоставляют, классифицируют, объединяют и обрабатывают данные;
- Людей, которые определяют использование и цели алгоритмической системы;
- Людей на руководящих должностях в институтах или предприятиях, которые используют или разрабатывают алгоритмические системы;

- Людей, которые программируют алгоритмические системы;
- Людей, которые тестируют, развивают и адаптируют алгоритмические системы;
- Людей, которые разрабатывают дизайн того, в каком виде выводятся результаты алгоритмического процесса;
- Людей, которые принимают решения с помощью алгоритмических систем или определяют их применение в своей повседневной работе.

Что такое “алгоритмическая система”?

«Алгоритм» – это однозначная инструкция к действию для решения исходно очерченной проблемы. Алгоритмическая система – это система, состоящая из одного или нескольких алгоритмов, используемых в программном обеспечении для сбора и анализа данных, а также для выводов, необходимых для решения заранее заданной проблемы. Система при этом может быть самообучаемой или следовать изначально запрограммированным правилам для принятия решений. Использование Algo.Rules для оценки алгоритмической системы включает в себя учёт и более широкого социально-технического контекста, в который встроено программное обеспечение, например, понимания и интерпретации пользователями результатов и решений, которые они на ее основе принимают. Правила Algo.Rules применяются ко всему процессу разработки алгоритмических систем, а также к их внедрению в социальный контекст.

Правила Algo.Rules ориентированы на алгоритмические системы, которые оказывают прямо или косвенно значительное влияние на жизнь общества или отдельных людей. Поэтому речь идёт о наборе правил не для всех алгоритмических систем, а, скорее, для имеющих социальную значимость. Чтобы определить, отвечает ли алгоритмическая система этим требованиям, необходимо до её разработки оценить ее воздействие. Чем сильнее потенциальное влияние алгоритмической системы на общество или жизнь людей, тем более тщательно она должна соответствовать Algo.Rules.

1 Усилить компетентности

Функционирование и потенциальное воздействие алгоритмической системы должны быть осознаны

Те, кто разрабатывают, используют алгоритмические системы и / или принимают решения по их применению, должны обладать необходимой профессиональной компетенцией и соответствующим нюансированным пониманием того, как работает технология и какие возможные последствия она за собой влечет. Обмен индивидуальными и институциональными знаниями, а также междисциплинарный обмен между разными областями задач так же важны, как и меры по развитию соответствующих навыков. Последние должны быть внедрены в обучение, а также в тренинг новых сотрудников. Кроме того, междисциплинарный обмен должен быть поставлен на постоянную основу и оставаться открытым для всех заинтересованных и причастных лиц.

2 Определить ответственность

Физическое или юридическое лицо всегда должно нести ответственность за последствия, связанные с использованием алгоритмической системы

Ответственное лицо должно быть однозначно определено. Ему должны быть ясны сопряженные с этим задачи. Это также относится к обязанностям, которые разделены между несколькими людьми или организациями. Распределение ответственности должно быть полностью задокументировано и прозрачно как изнутри, так и извне. Ответственность не может быть передана самой алгоритмической системе, пользователям или тем людям, на которых направлена работа алгоритмической системы.

3 Документировать цели и ожидаемое воздействие

Цели и ожидаемое воздействие от внедрения алгоритмической системы должны быть задокументированы и оценены до ее запуска

Цели алгоритмической системы должны быть четко определены, а сведения о ее запуске задокументированы. Это включает в себя как лежащие в ее основе данные, так и модели расчета. Перед запуском алгоритмической системы должна быть проведена задокументированная оценка ее последствий. Особенно в случае самообучающихся систем и в

динамических областях применения, которые подвержены частым изменениям, оценку воздействия следует повторять через регулярные промежутки времени. При этом следует обращать особое внимание на риски дискриминации и другие последствия, имеющие отношение к индивидууму и общему благу. Ценностные параметры при целеполагании и запуске алгоритмических систем должны быть закреплены.

4 Гарантия безопасности.

Безопасность алгоритмической системы должна быть протестирована до и во время ее внедрения

Надежность и прочность алгоритмической системы, а также лежащих в ее основе данных против атак, доступа и манипуляций должны быть гарантированы. Безопасность должна быть изначально встроена в структуру алгоритмической системы (безопасность по умолчанию). Система должна быть проверена перед ее внедрением в защищенной среде. Все меры предосторожности должны быть задокументированы.

5 Провести маркировку

Использование алгоритмической системы должно быть маркировано

Благодаря соответствующей маркировке, люди, взаимодействующие с алгоритмическими системами, должны иметь возможность определить, что решение или прогноз основаны на работе алгоритмической системы. Это особенно важно в тех случаях, когда система имитирует человека и его способы взаимодействия (язык, внешность и т.д.).

6 Обеспечить понятность

Процессы принятия решений в алгоритмической системе всегда должны быть понятными

Алгоритмическая система в ее прямых и косвенных воздействиях или в принципах ее работы должна быть легко понятна для человека, он должен иметь возможность подвергать ее сомнению и перепроверять ее решения. Для этого информация о данных и моделях, на которых основана система, ее архитектура и ее потенциальные последствия должны быть объяснены легким и доступным языком. Кроме того, необходимо постоянно проверять, не может ли быть достигнута та же цель без существенной

потери качества за счет использования менее сложной алгоритмической системы, которая предполагает более понятный режим работы.

7 Гарантия управляемости

Алгоритмическая система должна быть управляемой в течение всего срока ее использования.

Чтобы алгоритмическая система оставалась управляемой, все, кто участвует в ее разработке и реализации, должны обладать постоянным общим контролем над системой. При этом необходим гарантированный доступ к постоянному обзору всей системы в целом, даже когда задачи распределены между отдельными лицами и участками работ. Функционирование системы никогда не должно достигать такого уровня сложности или непрозрачности, чтобы человек терял над ней контроль или не мог бы вносить в нее изменения. Это относится в особенности к самообучающимся системам. Если эта управляемость не может быть гарантирована, алгоритмическую систему не следует вводить.

8 Мониторинг воздействия

Последствия работы алгоритмической системы должны регулярно анализироваться.

Алгоритмическая система должна подвергаться активному мониторингу для выяснения того, действительно ли достигнуты поставленные ей цели, и не нарушает ли система действующее законодательство. Внешние контролирующие инстанции, принимая соответствующие технологические меры предосторожности, чтобы не ставить под угрозу законную конфиденциальность бизнеса, должны иметь возможность проводить независимый, всеобъемлющий и эффективный аудит алгоритмической системы. В случае выявления негативного воздействия системы, необходимо определить его причину и соответствующим образом ее адаптировать.

9 Установка механизмов подачи жалоб

Если алгоритмическая система приводит к сомнительному решению или решению, которое затрагивает права человека, оно должно быть объяснено или обжаловано

Организация, запускающая алгоритмическую систему, должна обеспечить легко доступные способы обратной связи. Лица, подверженные воздействию системы, должны, во-первых, иметь возможность запрашивать соответствующую подробную информацию о конкретном решении и соображениях, на которых оно основано. Это должно быть возможным и для организаций, действующих от их лица, и для ситуаций, в которых организация исходит из собственных законных интересов. Во-вторых, должен быть внедрен легкодоступный и эффективный способ подачи жалобы. Жалобы и предпринятые меры также должны быть задокументированы.

Лицензия

Текст данной публикации защищен авторским правом и лицензируется по лицензии «Attribution-ShareAlike» («Атрибуция-СохранениеУсловий») 4.0 Всемирная (CC BY-SA 4.0). Вы можете свободно делиться (обмениваться) и адаптировать (создавать производные материалы) материал, но вы должны обеспечить соответствующее указание авторства, предоставить ссылку на лицензию, и обозначить изменения, если таковые были сделаны. Если вы перерабатываете, преобразовываете материал или берёте его за основу для производного произведения, вы должны распространять переделанные вами части материала на условиях той же лицензии, в соответствии с которой распространяется оригинал. Полный текст лицензии можно найти по адресу: <https://creativecommons.org/licenses/by-sa/4.0/legalcode.ru>

