Эксперты Инициатива открытого кода ищут единое определение на фоне бурного развития технологий.

В мире стремительно развивающихся технологий искусственный интеллект (ИИ) занял особое место. Генеративный ИИ, в частности, представляет собой нечто уникальное. Настолько уникальное, что даже определение термина «открытый код ИИ» вызывает затруднения у экспертов. Именно эту задачу взяла на себя Инициатива открытого кода (Open Source Initiative, OSI (Open Systems Interconnection) — концептуальная модель взаимодействия открытых систем, которая объединяет все коммуникационные функции вычислительных или телекоммуникационных систем. OSI демонстрирует, как компьютеры или другие типы систем коммуницируют друг с другом." data-html="true" data-original-title="OSI" >OSI), организовав международный тур для выработки определения, которое устроит большинство заинтересованных сторон.

Первым этапом этого путешествия стала конференция PyCon US, прошедшая в мае в Питтсбурге. Эта поездка OSI поддерживается Фондом Альфреда П. Слоуна, Amazon, Cisco и Google Open Source. В июне мероприятие пройдет в Париже (OW2) и Мадриде (OpenExpo Europe). Цель – завершить работу над определением к конференции All Things Open в Северной Каролине в конце октября.

В течение двух лет OSI разрабатывала проект определения, рассказал исполнительный директор организации Стефано Мафулли. Сейчас команда проходит этап «валидации», проверяя, что определение включает все аспекты, относящиеся к категории открытого кода. Параллельно разрабатывается FAQ для разъяснения ключевых моментов.

На семинаре PyCon около дюжины участников помогали команде формулировать вопросы и ответы для FAQ, чтобы охватить все важные аспекты.

Мафулли пояснил, что развитие открытого кода шло параллельно с развитием компьютерных наук, начиная с 1980-х годов. «Раньше все было просто: есть исходный код, есть двоичный код — два представления одного и того же артефакта. И долгие годы регуляторы не обращали на это внимания, пока экосистема открытого кода росла», — отметил он.

С появлением ИИ, особенно его нового поколения, появились новые артефакты – веса и параметры моделей. Они функциональны, влияют на работу систем, но не являются ни программным обеспечением, ни исходным кодом, ни данными. Более того, миллиарды людей уже их используют.

Все права защищены

Еще одной проблемой является то, что регуляторы во всем мире обеспокоены тем, что эти системы способны делать то, чего сами разработчики не могут объяснить или исправить.

Семинары PyCon для создания FAQ были оживленными, сообщила Мер Джойс, основательница консалтинговой компании Do Big Good, которая помогает OSI в этом проекте. По ее словам, наиболее острыми вопросами стали сертификация и определение данных в контексте ИИ.

Сертификация касается того, как анализировать и сертифицировать системы как открытый код ИИ. Другой важный вопрос — это определение данных в ИИ. В проекте определения OSI используется размытая формулировка, понятная юридическому сообществу: «Достаточно подробная информация о данных, использованных для обучения системы, чтобы квалифицированный специалист мог воспроизвести аналогичную систему, используя те же или аналогичные данные».

Мафулли пояснил, что необходимо четко объяснить термины «достаточно подробная информация», «квалифицированный специалист» и «аналогичная система». Примером данных могут быть синтетические данные, которые можно использовать вместо оригинальных данных, если последние защищены авторским правом или являются секретными.

Следующими этапами поездки OSI станут мероприятия в Париже и Мадриде. Европейский Союз опережает остальной мир в разработке политики управления ИИ, внедряя соответствующий закон, принятый в декабре.

Мафулли гордится двухлетней работой по сбору мнений от представителей различных стран, сфер и интересов. «Не было единого понимания этой среды, поэтому всем нам — академикам, индустриальным специалистам, исследователям, разработчикам, представителям гражданского общества, юристам — нужно было вести сложный и вдумчивый разговор, чтобы повысить общий уровень знаний», — сказал он. «Мы достигли этапа, когда я действительно впечатлен количеством людей, которых мы затронули этим процессом».

Работа OSI по созданию единого определения «открытого кода ИИ» продолжается, и ее результаты будут иметь значительное влияние на будущее этой технологии и всей экосистемы открытого кода.

Все права защищены