Секция «40.5 Этика»

Теория контекстуальной целостности Х. Ниссенбаум и приватность в условиях развития генеративного искусственного интеллекта

Научный руководитель – Скворцов Алексей Алексеевич

Косырева Анастасия Сергеевна

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Философский факультет, Кафедра этики, Москва, Россия

E-mail: astiko21@yandex.ru

Быстрое внедрение генеративного искусственного интеллекта (ИИ), в том числе больших языковых моделей, обостряет этические проблемы приватности. Обострение происходит не столько из-за нарастающего объема данных, сколько из-за неизбежного размывания границ между разными средами обращения информации (образование, работа, здоровье, семейная жизнь и др.). Важно и то, что в цифровой среде приватность всё меньше определяется тем, что именно человек сообщил или скрыл, и всё больше — тем, что о нём можно правдоподобно вывести, самим «потоком» информации. Генеративный ИИ усиливает этот сдвиг: системы способны не только объединять сведения, но и производить правдоподобные интерпретации, выводы, связывая разрозненные цифровые следы в единый нарратив, соединяя множество разнородных контекстов в единую информационную структуру, профиль, размывая те самые границы. Но, по мнению Х. Ниссенбаум, было бы ошибкой впадать в «нигилизм приватности» — отказ от каких-либо норм на том основании, что ИИ может вывести «всё из всего» [2].

Теория контекстуальной целостности Х. Ниссенбаум (contextual integrity, далее — CI) предлагает для описания этой проблемы и формулирования нормативной рамки достаточно прикладной язык: под приватностью понимается соответствие потоков информации контекстным нормам, а нарушение — как разлад «надлежащей передачи» (appropriate flow). Контекст в данном подходе подразумевает не просто тематическую сферу, а вполне конкретные социальные практики: контекст задается социальными ролями, целями, ценностями, а нормативная структура потока описывается конкретными параметрами: акторы, атрибуты и принципы передачи [5]. Применение этой схемы к использованию генеративного ИИ позволяет сместить дискуссию от абстрактных постулатов о необходимости защищать персональные данные к проверяемым вопросам: какие именно контекстные нормы нарушаются, кому это дает власть и какие организационные решения нужны для восстановления целостности.

Подход Х. Ниссенбаум в связи с этим больше отвечает современным реалиям в решении проблемы приватности, чем классическое контролирующее право, которое чаще всего представляет данные как отдельные элементы, которые можно удалить или скрыть, и уместно дополняет правовые режимы, которые пытаются эту проблему решить (например, GDPR). К примеру, по теории CI, публичность в одном социальном смысле не равна разрешению на машинную переработку и повторное предъявление в другом [4]. Здесь уместна и критика Ниссенбаум модели уведомления и информационного согласия пользователя: при масштабном вторичном использовании данных согласие либо фиктивно (условия слишком сложны для понимания), либо не описывает реальную структуру будущих потоков, потому что технически это описать действительно сложно [1]. Практика информационного согласия также демонстрирует свою недостаточность в условиях нарушения групповой приватности, когда субъект может стать объектом для выводов только на основе данных о его социальной группе [3]. Генеративный ИИ усиливает этот разрыв,

потому что будущие получатели и режимы использования информации непредсказуемы зачастую даже для поставщика сервиса.

Теория CI позволяет провести первичную гуманитарную экспертизу на уровне потоков, ролей и обязанностей. Это и сильная, и слабая сторона теории CI, т.к. требует достройки для работы с первичными техническими данными (логи, сигналы и др.), которые сложно перевести в семантические значения. Но минимальная процедура гуманитарной экспертизы может быть сформулирована как требование к организациям, внедряющим ИИ-ассистентов: (1) явным образом определить контексты использования (например, внутреннее консультирование студентов и внешняя коммуникация с абитуриентами — разные контексты); (2) перечислить акторов, включая часто невидимых (провайдер, асессоры, сторонние получатели, а также третьи лица, чьи данные пользователь вводит в запрос); (3) задать атрибуты информации, с которыми допускается работа (чувствительные сведения, авторские тексты, данные о здоровье, внутренние документы); (4) согласовать принципы передачи: например, запрет на хранение запросов, ограничение повторного использования результатов ответа модели, правила цитирования, сроки удаления и т.п. Эта рамка удобна тем, что переводит этику в язык институциональных обязанностей и ролевых ожиданий, а не призывает к ответственности одного только пользователя.

Важно, что генеративный ИИ меняет саму моральную сторону приватности: в повседневной коммуникации он выступает как квази-субъект, но фактически является каналом к целой инфраструктуре с множеством актором (платформе, экосистеме и пр.). Теория CI позволяет артикулировать это как проблему подмены ролей: пользователь ведет себя так, будто общается с собеседником или помощником, тогда как реально вступает в отношение с инфраструктурой. Отсюда прикладные рекомендации: интерфейс и регламенты должны делать роль провайдера видимой, а не маскировать ее антропоморфным дизайном; для чувствительных контекстов допустимость ИИ должна оцениваться не только по точности, но и по совместимости с принципами передачи и дальнейшего использования.

Таким образом, теория Х. Ниссенбаум показывает, что этика приватности в условиях развития генеративного ИИ требует не универсальных запретов и попытки контроля информации, а проектирования и поддержания социотехнических границ — организационных, интерфейсных и культурных — которые помогают сохранять целостность социальных практик.

Источники и литература

1) Barocas[U+202F]S., Nissenbaum[U+202F]H. Computing ethics: Big Data's end run around procedural privacy protections // Communications of the ACM. – 2014. – Vol.[U+202F]57, No[U+202F]11. – P.[U+202F]31-33.

2) Engelmann[U+202F]S., Nissenbaum[U+202F]H. Countering privacy nihilism // Conceptions of data protection and privacy. Legal and philosophical perspective. – London: Hart Publishing, 2025. – DOI:[U+202F]10.48550/arXiv.2507.18253.

3) Floridi[U+202F]L. The ethics of information. – 1st[U+202F]ed. – Oxford: Oxford University Press, 2013. – 376[U+202F]p.

4) Nissenbaum[U+202F]H. Contextual integrity up and down the data food chain // Theoretical Inquiries in Law. – 2019. – Vol.[U+202F]20, No[U+202F]1. – P.[U+202F]221-256.

5) Nissenbaum[U+202F]H. Privacy in context: Technology, policy, and the integrity of social life. – Stanford, CA: Stanford University Press, 2010. – 384[U+202F]p.

6) Selbst [U+202F] A. [U+202F] D. Fairness and abstraction in sociotechnical systems // Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT '19). – New [U+202F] York: ACM, 2019. – P. [U+202F] 59-68.