

ОБ ЭТИКЕ ПЕРСОНАЛИЗИРОВАННЫХ РЕКОМЕНДАЦИЙ В НАЦИОНАЛЬНОЙ БИБЛИОТЕКЕ ФРАНЦИИ

FOR AN ETHICS OF PERSONALIZED RECOMMENDATION AT THE FRENCH NATIONAL LIBRARY*

Селин Леклер (Céline Leclaire), *отдел научно-технических вопросов, Управление по услугам и сетям, Национальная библиотека Франции, Париж, Франция*

Люси Терминьон (Lucie Termignon), *отдел научно-технических вопросов, Управление по услугам и сетям, Национальная библиотека Франции, Париж, Франция*

Реферат. Проблема этического аспекта использования технологий искусственного интеллекта является предметом многочисленных обсуждений и работ, представленных исследователями, работниками соответствующих учреждений и операторами проектов. Авторы статьи вносят свой вклад в дискуссию, предлагая в качестве примера услугу, которую в настоящее время разрабатывает Национальная библиотека Франции. Какими вопросами и этическими принципами руководствуется библиотека, приступая к разработке системы персонализированных рекомендаций по контенту на основе искусственного интеллекта? Как этот подход отражает общую политику учреждения в отношении искусственного интеллекта?

Интеграция рекомендаций в платформу Gallica становится все более необходимым инструментом, который способствует повышению ценности ее богатых и разнообразных коллекций и удовлетворению потребностей пользователей, особенно исследователей. Помимо технических сложностей, при реализации данной опции необходимо учесть различные этические аспекты проблемы — доверие, прозрачность и избегание пузырей фильтров, что требует от библиотеки прояснить позицию, которую она стремится занять в документальном мире. На карту поставлены сама суть концепции рекомендаций, понимание профессии библиотекаря и ее роли.

Ключевые слова: искусственный интеллект, персонализированные рекомендации, электронная библиотека, этика.

Введение

В настоящее время, если вы выполните поиск на платформе Gallica¹ (цифровая библиотека Национальной библиотеки Франции (VnF) и ее партнеров), вы получите те же результаты, что

и любой другой интернет-пользователь, который выполнит тот же поиск, поскольку Gallica не располагает системой персонализированных рекомендаций. Это связано с этическим принципом, которого придерживается VnF: поисковая система Gallica не использует личные данные пользователей, особенно журналы трассировки, для выбора и ранжирования предоставляемых результатов. Платформа лишь предлагает библиографические данные, относящиеся к документам.

Однако обоснованность подобной позиции поставлена под вопрос несколькими факторами, которые могут способствовать стратегическому развитию VnF. Во-первых, многие технологии в настоящее время уже досконально разработаны, они все чаще позволяют системам обрабатывать большое количество данных в режиме реального времени. Во-вторых, требуется обеспечить новые точки доступа к коллекциям и новые способы представления контента в связи с эволюцией ожиданий пользователей и постоянно растущим объемом доступного культурного наследия, особенно если речь идет о франкоязычных областях, которые зачастую не включены или не полностью включены в алгоритмы обучающих корпусов.

С целью более эффективного учета этого «цифрового поворота» и разработки ответственной политики в области технологий искусственного интеллекта (ИИ), VnF создала глобальную структуру в рамках своей дорожной карты по развитию ИИ [1], опубликованной в 2022 году. В ней указано, каким образом сфера приложений ИИ охватывает все виды деятельности и услуг библиотеки. В ней также выделены ключевые проекты и приоритеты, а также предлагается твердая стратегическая основа и набор мер по управлению изменениями (обучение, оценка и т. д.).

Проект внедрения персонализированных рекомендаций для коллекций Gallica разработан как

* <https://repository.ifla.org/handle/123456789/2074>

часть данного перспективного и увлекательного процесса. Однако его сопровождают различные вопросы и проблемы: как деонтология государственных услуг может быть связана с возможностями, обусловленными технологическим развитием с точки зрения улучшения обслуживания?

Обсуждаемый проект является первым опытом VnF в области разработки проекта, который изначально создается с учетом этического аспекта будущей деятельности. Планируется постепенно сформулировать более точное определение того, в чем состоят персонализированные рекомендательные системы. На каждом этапе данное определение будет основано на обсуждении следующих этических вопросов: действительно ли библиотеке нужна такая услуга? Может ли она избежать систематического сбора персональных данных? Какие правила и гарантии учреждение способно установить?

Что такое персонализированная рекомендация на основе алгоритма и чем она полезна?

Начнем с того, что представим наши шаги с позиции дизайнера, т. е. с точки зрения библиотеки, действующей в соответствии с возложенными на нее обязанностями.

Принцип рекомендации

Любая работа с ИИ приводит к появлению нового подхода к повседневной деятельности, организации и уставу библиотеки. Например, при разработке чат-бота для пользователей мы задаемся вопросом, что собой представляет беседа между людьми, а когда мы пытаемся перевести на компьютерный язык правила, которые навязываем пользователям, мы часто начинаем осознавать, насколько они неуместны.

Когда дело доходит до предоставления персонализированной рекомендации, стоит отметить разнообразие и количество операций, которые данный процесс охватывает. Представим, что к библиотекарю, который находится в читальном зале или общается через сервис «Спросите библиотекаря», обращается пользователь и просит порекомендовать книгу. Библиотекарь попытается узнать больше о том, что именно хочет читатель, задаст уточняющие вопросы, спросит, что человек уже нашел или прочитал. Далее библиотекарь активизирует свои знания об интересующей пользователя области, тщательно проанализирует ее, сделает соответствующий запрос в каталоге, используя метаданные, и при необходимости задействует дополнительную помощь (базы данных, более опытные коллеги и т. д.). Наконец, библиотекарь адаптирует свой ответ под читателя, учиты-

вая время, которым тот располагает, и обеспечит ему доступ к самому контенту, поскольку библиотекарь понимает, что пользователь этого ожидает. На протяжении указанного процесса у библиотекаря есть возможность задаться рядом вопросов: не предвзят ли мой ответ в силу моих предубеждений о пользователе? Например, не недооценил ли я его способность разбираться в том или ином ресурсе? Где граница между рекомендацией и предписанием?

Рекомендация на основе алгоритма

Подобные сомнения не касаются алгоритмов, поскольку они содержат лишь вычисления и организованные приоритетные параметры. Цель рекомендательных систем состоит в том, чтобы предлагать элементы, наиболее соответствующие потребностям и предпочтениям пользователей, облегчая проведение исследований и доступ к информации.

Данные, используемые в работе, двойки: относящиеся к содержимому и относящиеся к пользователям. Первый тип касается описательных метаданных контента, связанных с потреблением контента, а также данных оценки контента, которые могут быть предоставлены пользователями. Пользовательские данные включают личные данные аутентификации и данные о поведении (данные просмотра и регистрации). Они могут быть собраны явным образом путем заполнения формы или анкеты или же с помощью косвенных методов отслеживания, таких как файлы cookie, веб-маяки, отпечаток браузера или когортный таргетинг (Федеративное обучение на основе когорт, FLoC; это инициатива компании Google, которая заключается в таргетинге группы пользователей со схожим поведением, а не одного пользователя, после определения общих центров интересов и журналов пользователей объединяют в анонимные группы, помеченные уникальным и постоянным идентификатором, который присваивается всем пользователям одной группы и управляется на уровне браузера или операционной системы).

Рекомендательные системы работают в три этапа: сбор пользовательских данных (1) позволяет создать профиль пользователя или модель предпочтений (2), которые затем эксплуатируются методами машинного обучения (т. е. семействами алгоритмов) (3) для определения рекомендаций, наиболее подходящих для конкретного пользователя. Алгоритмы, используемые в методах обучения, обычно подразделяются на два вида: совместная фильтрация и фильтрация на основе контента. Система рекомендаций на основе контента пытается угадать особенности или поведение пользователя с учетом характеристик предмета, на которые он положительно реагирует.

Ее работа не зависит от данных других пользователей в отличие от систем совместной фильтрации, которым не требуется указывать характеристики предмета, но которые учитывают реакцию других пользователей, давая рекомендации для конкретного пользователя.

Эти два подхода не являются взаимоисключающими и могут комбинироваться, чтобы избежать предвзятости и побочных эффектов. После выбора способа сбора и обработки данных и обучения модели участие оператора больше не требуется. Предложения относительно контента передаются пользователю через интерфейс без человеческого вмешательства. В связи с этим интересно отметить взаимодействие, которое устанавливается между пользователем и алгоритмом: чем больше отзывов алгоритм получает о том, что он предлагает, тем более эффективной становится его работа.

Привлекательность алгоритма в эпоху больших данных

Основная причина, по которой VnF стоит участвовать в проекте по разработке системы персонализированных рекомендаций (другими словами, потратить государственные средства на создание нового сервиса на основе алгоритмов), заключается в том, что библиотека стремится исполнить пожелания пользователей. По данным опроса, проведенного библиотекой в 2020 г. [2, р. 15–20], 90% пользователей VnF по всему миру удовлетворены работой платформы Gallica, однако 49% из них считают, что необходимо улучшить поисковую систему, лишь 11% ответили, что получают необходимую поддержку при поиске (41% аспирантов; почти треть аспирантов полагают, что результаты, предлагаемые поисковой машиной, не всегда релевантны). Приведем в качестве примера один из комментариев: «У меня мало времени, поэтому я не могу удовлетворить свое любопытство или по максимуму использовать возможности Gallica, чудесной библиотеки. Единственный минус: избыток такое, что мне иногда трудно, у меня недостаточно времени, чтобы точно определить сферы своих интересов» [2, р. 10–11]. Другой отзыв: «Поисковик часто выдает большое количество результатов, которые трудно отсортировать» [2, р. 18].

Потребность в простоте перекликается с этими ожиданиями: простой поиск — предпочитаемый способ получения доступа к контенту платформы Gallica (78% используют этот метод «каждый раз» или «часто»). Как выразился один из участников опроса, «просмотр цифрового контента VnF кажется запутанным» [2, р. 18]. Подобные ответы демонстрируют, насколько ошибочно положение о том, что пользователи имеют естественную способность применять цифровые инструменты (в частности,

работать с коллекциями Gallica). Как библиотекари, мы привыкли подвергать сомнению наши подходы к знанию, например к художественным произведениям и романам [3], поэтому подобные исследования наводят нас на размышления о том, является ли поисковая система с классическим дизайном лучшим инструментом, позволяющим направить пользователей к контенту, который они хотят (или жаждут) найти.

Наряду с изложенными заключениями, стоит отметить, что подход к управлению контентом, ориентированный на пользователя, постоянно развивается, привлекая все больше партнеров и самих пользователей. В течение последних нескольких лет партнеры VnF имели возможность вносить свой вклад в блог Gallica. Например, Национальный музей естественной истории выложил в блоге серию публикаций о растениях под названием «Herbier de Gallica»², а Университет г. Лиля опубликовал серию «Tissus et Textiles», посвященную текстилю и тканям³. Благодаря сотрудничеству с партнерами в 2020 г. VnF и Международный центр комиксов и изображений выпустили на платформе Gallica тематическое руководство по комиксам⁴. Пользователям были предоставлены новые тематические или географические подборки и идеи. VnF также взаимодействует в социальных сетях с пользователями платформы, так называемыми галликанавами (Gallicanauts), и привлекает пользователей YouTube к адаптации контента, связанного с культурным наследием, для различных типов аудитории. В течение долгого времени искусство библиотекаря основывалось на высоком качестве собранных библиографических метаданных (и подобные метаданные все еще служат фундаментом библиотечной деятельности). VnF является ключевым участником процесса распространения библиографических метаданных, особенно идентификаторов (ISNI и пр.), в 2011 г. она запустила портал⁵, который функционирует как справочный веб-сайт [4]. В настоящее время библиотека разрабатывает новый инструмент каталогизации, который призван еще больше расширить потенциал этих данных, и налаживает новые механизмы взаимодействия с культурными и креативными индустриями, чему способствует переход на цифровой обязательный экземпляр.

Предполагалось, что пользователь должен уметь выполнять поиск в каталоге или электронной библиотеке, при этом библиотекари могли лишь догадываться о тех навыках и знаниях, которыми он действительно располагал. Подобного пользователя все чаще рассматривают как «эксперта», поскольку он просматривает другие веб-сайты, особенно коммерческие (к этому положению мы еще вернемся).

Также отметим, что на фоне увеличения числа документов и расширения их разнообразия стано-

вится недостаточно тех руководств по проведению исследований и советов по поиску материалов, которые в настоящее время предлагает Gallica (подборки, всплывающие блоки, разработанные специалистами по модели инструмента Knowledge Graph компании Google, отчеты о результатах поиска, поиск на основе географического положения или по смежным запросам и пр.). Большинство пользователей даже не подозревают, что в библиотеке Gallica можно найти не только книги или газеты, но и фотографии, звукозаписи, видеоматериалы и 3D-объекты.

Таким образом, перед VnF стоит двойная задача, а именно: помочь пользователям ориентироваться в чрезвычайно больших коллекциях и повысить видимость часто пропускаемых материалов — при этом весь процесс должен работать в режиме реального времени, в любое время суток и с учетом разнообразных и меняющихся ожиданий пользователей.

Что такое этическая рекомендация?

Использование алгоритма рекомендаций оправдано необходимостью повысить уровень удовлетворенности пользователей, в то же время существует риск вызвать их недовольство, если предлагаемые рекомендательной системой результаты покажутся пользователям нерелевантными и малопонятными, а также если методы работы системы будут слишком навязчивыми. Доверие пользователей к алгоритму основано на общих этических принципах, которые часто упоминаются в многочисленных кодексах этики и характеризуются рядом аспектов: объяснимость, надежность, справедливость, ответственность, прозрачность, человеческий контроль (управляемость), безопасность, сохранность (конфиденциальность и приватность). Эти принципы соответствуют ключевым характеристикам доверия, перечисленным Французским агентством по стандартизации (AFNOR) при формировании французской стратегии в области стандартизации ИИ⁶. Другие характеристики, связанные с доверием, например возможность отслеживания или подталкивание к действию, также регулярно упоминаются в деонтологических хартиях национальных и международных организаций: CNIL⁷, ARCOM⁸, Европейская комиссия⁹, ЮНЕСКО¹⁰. Заявление о соблюдении этических принципов при разработке рекомендаций является ответом на проблемы, выявленные с тех пор, как использование рекомендательных систем получило широкое распространение во многих сферах (социальные сети, торговые платформы, платформы доставки потокового контента и др.), а именно: неправомерное использование персональных данных, манипулирование поведением, предубеждения, дискриминация и т. д.

Осведомленность об указанных проблемах объясняет, почему примеры использования алгоритма персонализированных рекомендаций для некоммерческого веб-сайта, посвященного цифровому культурному контенту (например, в библиотеках), встречаются нечасто, а если и существуют, то редко выходят за рамки стадии индустриализации (среди учреждений, хранящих коллекции, упомянем рекомендательную систему, основанную на библиографических записях MARC21, разработанную Национальной библиотекой и архивами Квебека (VAnQ) в 2017 г., онлайн-каталог Совета Национальной библиотеки Сингапура, сервис Maken, разработанный Национальной библиотекой Норвегии, или рекомендательный инструмент Bookarang, приобретенный в 2020 г. службой NDB Biblion).

Эта тенденция отражает специфику задач, возникающих при разработке подобного алгоритма. Это касается сингулярности целевой аудитории (широкая общественность и экспертная общественность), данных, которые подпитывают алгоритм, и вопросов, связанных с обслуживанием населения, характерных для данного типа учреждений. Все это побуждает нас задаться вопросом об этических проблемах, сопутствующих использованию новых технологий в области рекомендаций: какие базовые технические требования направлены на противодействие подобным опасностям, какие показатели и средства контроля позволяют оценивать рекомендации на основе алгоритмов?

Этические проблемы, связанные с использованием системы персонализированных рекомендаций в библиотеке

Сбор, хранение и использование пользовательских данных вызывают множество этических вопросов, главный из которых заключается в том, чтобы убедиться, что пользователь доверяет рекомендательной системе. Эти вопросы также отражают двойную проблему необходимости обеспечения документальной релевантности результатов, предложенных алгоритмом, и соблюдения конфиденциальности пользователей. С одной стороны, речь идет о том, чтобы избежать создания для пользователя закрытых пузырей фильтров (filter bubble, «фильтр интересов»), а с другой — о гарантии соблюдения конфиденциальности с целью сохранения информации и сведений, полученных в процессе индивидуального поиска.

Сбор, хранение и использование данных пользователей: проблема безопасности и прозрачности

Алгоритмы фильтрации, развернутые в рамках рекомендательных систем, основаны на построении «профилей пользователей». Эти

профили в той или иной степени персонализированы в зависимости от данных, на которых они основаны (персональные данные, внесенные пользователем в форму, телеметрия, журналы трассировки или даже относящиеся к пользователю данные из внешних источников). Чтобы составить список предложений в соответствии с запросом пользователя, на основе его данных (собранных явным или косвенным образом) строится матрица для сравнения их с данными, связанными с использованием контента других пользователей. С помощью алгоритмических решений – криптографии, обезличивания данных (VnF уже сталкивалась с проблемами и техническими решениями в области обезличивания данных в рамках исследовательского проекта, проведенного совместно с компанией Télécom ParisTech, который включал эксперименты с отслеживанием использования платформы Gallica [5]), протоколов или стандартов (Регламент Европейского парламента и Совета Европы от 27 апреля 2016 г. о защите физических лиц в отношении обработки персональных данных и о свободном перемещении таких данных, а также об отмене Директивы 95/46/ЕС¹¹) – были разработаны различные методы для предотвращения злонамеренного использования персональных данных на разных уровнях рекомендательной системы. Однако существующие подходы зачастую либо медлительны, либо не масштабируемы, поэтому с целью решения этой проблемы необходимо и далее поощрять проведение экспериментов и изучать полученный опыт в области совместной фильтрации и выявления возможностей защитить право на частную жизнь.

Помимо проблемы неприкосновенности частной жизни, использование поведенческих данных вызывает озабоченность с точки зрения конфиденциальности. Уникальность пользователей VnF заключается в разнообразии вариантов обращения с материалами, начиная от чрезвычайно целенаправленного поиска известного читателю документа, который является частью доступного на платформе Gallica контента, до свободного изучения коллекций, руководствуясь исключительно любопытством. В то время как профили ученых и специалистов будут привязаны к конфиденциальности их исследований, чтобы гарантировать объективность их работы, начинающие пользователи руководствуются различными существующими формами рекомендаций и действительно в них нуждаются. Интеграция системы алгоритмических рекомендаций в поисковую систему такого веб-сайта, как Gallica, должна отвечать требованиям этих двух противоположных типов аудитории (литература по методам совместной фильтрации и публика-

ции данных с сохранением неприкосновенности частной жизни требует тщательного изучения [6; 7]).

Кроме того, алгоритмы создают эффект «черного ящика», т. е. их работа не всегда поддается объяснению, что вызывает у пользователей недоверие и страх по отношению к системе, которая может манипулировать ими, используя личные сведения без их ведома. Предоставление пользователям возможности контролировать процесс профилирования вручную и выбирать критерии, на которых основаны рекомендации алгоритма, – один из способов повысить не только прозрачность системы рекомендаций, но и удовлетворенность пользователей от ее работы, а также заручиться их доверием [7].

Пузыри фильтров, релевантность и разнообразие

В зависимости от того, каким образом разработаны алгоритмы, системы рекомендаций могут привести к сокрытию от пользователей части результатов и созданию ограничительных рамок, обусловленных информацией об интересах, полученной на основе анализа истории просмотров. Замыкание пользователей в пузырях фильтров влечет за собой риск стандартизации и обеднения сферы исследований. Так, если алгоритм не принимает во внимание диверсификацию, исследователи с похожими областями поиска получают одинаковые результаты, что вряд ли приведет их к другим направлениям мысли, отличным от тех, в которых они привыкли развиваться [8, p. 194]. Наблюдения (женщина, которая указала на негативный аспект своего опыта поиска информации на платформе Gallica, сразу же смягчила свои слова, добавив, что это дает ей возможность обнаружить новые или неизвестные ей области) и несколько исследований показали, что диверсификация оказывает положительное влияние на уровень удовлетворенности пользователей. Однако это может привести к недоверию и непониманию, если связь с их обозначенными предпочтениями будет слишком отдаленной [9].

Более того, поведение пользователей постоянно развивается и склонно меняться от одного сеанса к другому: в зависимости от стадии исследовательской деятельности ученый может, например, испытывать потребность скорее в разнообразии, нежели в точности полученных результатов. Замыкание пользователей в рамках стандартных фиксированных профилей рискует привести к ситуации, когда ожидания пользователей не будут удовлетворены. Алгоритмы, используемые в рекомендательных системах, должны учитывать изменяющийся характер подходов, поведения и ожиданий пользователей.

Этические принципы разработки рекомендательной системы

Проектируемая этичность

Моральные дилеммы (конфиденциальность пользователей или персонализация, обезличивание данных или полезность данных и т. д.) возникают на каждом этапе разработки рекомендательной системы (профилирование пользователей, разработка алгоритма, создание пользовательского интерфейса, экспериментирование и тестирование). Решение этических вопросов является сложной задачей и требует глобального видения проекта. Работа по обезличиванию данных, параметризация самого алгоритма — внесение элемента случайности, отклонение от запроса в предложенных результатах или даже результаты, противоположные предполагаемым вкусам пользователя; использование данных о просмотрах на уровне агрегации (пример применения правила: если только один пользователь делает определенный выбор в списке результатов, это будет считаться особым подходом, но если 20 человек ведут себя одинаковым образом, это будет считаться поведением, которое, вероятно, окажется общим) — все это служит мерами защиты от этических рисков, возникающих на различных

этапах разработки рекомендательной системы. Интеграция этических показателей в индикаторы оценки рекомендательных систем также является частью «проектируемого» этического подхода. Например, параметры диверсификации были объединены с параметрами релевантности для оценки алгоритма Culture Pass, разработанного Министерством культуры Франции. Четыре индикатора этой метрики позволяют количественно определить обнаруживаемость и оценить алгоритм с точки зрения диверсификации предлагаемого контента.

Пользователи и библиотекари как активные участники работы рекомендательной системы

Рекомендательные системы, которые внедряют на некоммерческом веб-сайте, посвященном культурному контенту, не предназначены заменить консультационную роль библиотекаря. Параметры близости, используемые алгоритмом для составления рекомендаций, могут быть представлены на двухосевой диаграмме, где они расположены в зависимости от степени автономии, оставленной алгоритму, и персонализации рекомендации, предлагаемой пользователю (рис.).



Рис. Несколько возможных рычагов настройки рекомендательного алгоритма и их основные ограничения с этической точки зрения

Рисунок иллюстрирует тот факт, что риск быть запертым в пузырях фильтров выше, когда алгоритм максимально использует личные и поведенческие данные пользователя, а консультативная роль библиотекарей отходит на второй план. Это также говорит о том, что риск в меньшей степени осознается пользователями, когда их данные собирают без запроса на указание каких-либо предпочтений.

Помимо дихотомии между конфиденциальностью пользователей (выбор в пользу «ручных» рекомендаций библиотекаря по тематике) и персонализацией, основанной на массовом учете пользовательских данных, стоит рассмотреть гибридные промежуточные сценарии, когда пользователи могут применять настраиваемые параметры в механизме рекомендаций [7], а библиотекари задействуют алгоритмические рекомендации, полученные на основе агрегированных данных, в качестве вспомогательного средства для управления контентом на платформе без прямого взаимодействия между пользователями и алгоритмами рекомендаций. Национальный аудиовизуальный институт (INA) движется к этому решению, мобилизуя технологии искусственного интеллекта для укрепления своей миссии по продвижению и сохранению аудиовизуального наследия (см.: «Как искусственный интеллект дает INA возможность усилить свою миссию по продвижению и сохранению аудиовизуального наследия»¹²).

Повышение осведомленности через дизайн интерфейса

Способ повышения доверия пользователей к системе рекомендаций и решения проблем конфиденциальности и персонализации заключается в дизайне интерфейса и его эргономике. Перспективными представляются следующие пути развития: явное и четкое указание критериев рекомендации («люди, заинтересованные в этом документе, также смотрели тот...», «неожиданные предложения по результатам», «результаты, противоположные вашим поискам», «то, что VnF рекомендует всем»), которые позволяют пользователям настраивать свои предпочтения (не только за счет их объявления через заполнение формы, но и возможность балансировки веса личных предпочтений в списке предлагаемых пользователю результатов), обновление их и определение степени желаемой релевантности (в отличие от диверсификации) или даже предложение двух режимов просмотра — с рекомендацией или без нее. Так, в центре внимания находятся эргономика интерфейса, такие направления, как UX-дизайн и взаимодействие человека с компьютером (Human-Computer Interaction), а также социологический подход.

Несмотря на то что при разработке рекомендательной системы необходимо пойти на различные

компромиссные уступки, исключение этических соображений приведет лишь к страху и недоверию со стороны пользователей. Таким образом, этические вопросы необходимо решать на каждом этапе разработки системы рекомендаций и с разных точек зрения, начиная от инженерии данных и наук о данных и заканчивая дизайном пользовательского интерфейса, взаимодействием человека с компьютером и социологией.

Практическая реализация концепции в VnF: к новой схеме, требующей образцового подхода

В 2023 г. VnF начала разработку прототипа для экспериментов с рекомендациями на основе алгоритмов с учетом этических соображений. Дизайн основывается на трех компонентах: инженерии данных и обучение алгоритма, разработка прототипа в соответствии с принципами UX-дизайна и изучение использования, основанное на социологической науке.

Как конкретно в библиотеке учитываются поднятые вопросы? Какой должна быть новая услуга, чтобы она отвечала интересам и ценностям не только граждан, которым она может быть полезна, но также сотрудников и партнеров? Подобно тому как деятельность исследователей претерпела значительные изменения в силу увеличения объема контента, доступного в текстовом формате, точно так же внедрение основанных на алгоритмах рекомендаций приведет к настоящим новаторским изменениям в возможностях, предлагаемых библиотекой ученым, к глубокому перевороту в стратегиях проведения исследований. Так, будут объединены несколько средств: институциональные, организационные, технические и юридические.

Место в глобальной стратегии

Как изложено выше, персонализированные рекомендательные системы не должны заменять другие формы адаптации контента, ориентированные на пользователя, или же сбивать библиотеку с ее пути разработки новых методов интеллектуального анализа и нового механизма визуального поиска на платформе Gallica.

Кроме того, этот проект упоминается в нескольких дорожных картах, уточняющих стратегию библиотеки, что гарантирует разработку последовательной политики в отношении основных задач учреждения: программа MISAOA (взаимная поддержка и инновации для обеспечения сохранности и доступа к французским аудиовизуальным произведениям), работа в области цифрового обязательного экземпляра, планы по улучшению цифровой пользовательской среды и их взаимодействия с «интеллектуальными» рекоменда-

ями в отношении аудиовизуальных документов. Отметим, что заявление о целях и результатах работы библиотеки на 2022—2026 гг. перекликается с основными целями ее дорожной карты в отношении искусственного интеллекта [1], где персонализированные рекомендации считаются ключевым проектом. Эта дорожная карта уже привела к проведению ряда внутренних семинаров и совещаний по этике, которые были направлены на то, чтобы оглянуться назад, точнее, выделить общие критические моменты, возникающие при разработке проектов, и изучать передовой опыт. Такой подход является частью глобальной политики подотчетности в отношении цифровых технологий (устойчивость, оценка цифрового воздействия, прозрачное и уважительное использование cookie-файлов, предоставление большинства услуг и контента без необходимости входа в систему и т. д.).

Помимо собственных интересов в области управления или организации библиотечной деятельности, ВпФ также выполняет свою миссию в качестве государственного учреждения, участвуя в обучении населения принципам критического мышления и развития навыков медиаграмотности: открытые общественные обсуждения по проблеме искусственного интеллекта (например, дискуссия на французском языке «Существует ли искусственный интеллект?»¹³), партнерские отношения с заинтересованными сторонами в сфере образования и пр. Недостаточно разработать этические цифровые услуги, поскольку пользователи будут больше доверять учреждению, учитывая набор всех его различных обязательств, когда, внедряя систему персонализированных рекомендаций, оно также отображает эффект и видимость работы алгоритмов, чтобы повысить осведомленность граждан о различных способах, посредством которых они могут участвовать в поиске ресурсов и информации. Подобная осведомленность может быть обусловлена как эргономикой платформ, предназначенных для оказания государственных услуг, так и настройками алгоритма: таким образом, цели рекомендательной системы становятся явными, объяснимыми и управляемыми, какими бы многочисленными и развивающимися они ни были (например, экономия времени за счет возможности быстро найти наиболее релевантные результаты или, напротив, устроить «прогулку» по коллекциям).

Наконец, ВпФ руководствуется государственной политикой, особенно политикой Министерства культуры Франции, от которого библиотека напрямую зависит. В документе, где изложена цифровая стратегия [10], Министерство культуры Франции подчеркивает необходимость постоянного отслеживания эволюции в области использования информации, чтобы максимально увеличить возможность обнаружения культурного контента

в Интернете, и указывает на то, как определенные алгоритмы угрожают культурному разнообразию, питающему художников и творцов. На международном уровне, если привести только один пример, Франция и Квебек установили несколько приоритетов в рамках общего заявления о расширении видимости франкоязычного культурного контента. С одной стороны, речь идет об усилении мер, направленных на развитие цифровых навыков и распространение самой идеи о повышении видимости контента, а с другой — о просветительской работе с населением, чтобы граждане осознали роль крупных платформ и, в частности, рекомендательных механизмов в доступе к культурному контенту¹⁴. Что касается этических вопросов и ответственности в цифровой среде, ВпФ вдохновляется опытом других международных организаций: ЮНЕСКО¹⁵, Европейский союз¹⁶ и др.

Эксперименты, оценка

Эксперименты, проводимые в условиях полной прозрачности с помощью специального интерфейса, с предварительным информированием пользователей и без ущерба для них (подробнее об этическом подходе к экспериментам см. [7, р. 5—6]) позволят оценить релевантность результатов, предложенных рекомендательной системой, доступность и пригодность услуги для пользователей, а также спрогнозировать ее долгосрочное воздействие (например, существует ли вероятность, что этот процесс повлияет на политику оцифровки как таковую?). Эксперименты не будут пытаться охватить все вопросы сразу, а сосредоточатся на некоторых аспектах, чтобы учитывать те или иные возникающие проблемы: алгоритмы, интерфейс и т. д. Путем определения соответствующих индикаторов оценки для моделей и в целом введения практики поэтапной проверки на протяжении всего процесса разработки библиотека сможет обеспечить эффективную реализацию проекта и обезопасить пользователей.

Обмен идеями и мнениями, привлечение партнеров

В ожидании времени, когда функционирующие и надежные алгоритмы рекомендаций станут общедоступными в соответствии с принципами открытости и прозрачности, а также для экономии инвестиционных затрат обсуждаемый проект уже готовят к работе в рамках объединенных и сетевых подходов. Старшие и младшие научные сотрудники и специалисты, ученые, библиотекари с техническими навыками и без них, менеджеры цифровых библиотек, партнеры Gallica, юристы, специалисты по кибербезопасности — все эти люди, к счастью, в настоящее время собираются

для дискуссий в BnF, однако библиотека также стремится узнать сторонние мнения, чтобы лучше изучить проблему доступности, особенно неанглоязычного контента. Международные проекты, в частности европейские, предлагают подходящую основу для разработки системы рекомендаций с проектируемой этичностью.

Заключение

Изложенные соображения выявляют классическую дихотомию спроса и предложения, которая определяет развитие библиотечных услуг. Рассмотрение концепции персонализации, основанной на искусственном интеллекте, позволяет нам пролить свет на то, как организация себя позиционирует. Так, если миссия библиотеки, в частности BnF, заключается в расширении и развитии разнообразия и открытости, она не может удовлетвориться подходом, основанным лишь на спросе, и должна также развивать предложение. Невыполнение этого требования подвергает библиотеку риску лишиться того, что представляет суть ее фондов, какими бы разнообразными и энциклопедическими они ни были, и своей документальной (или цифровой) политики, т. е. основы своего существования. Невыполнение этого требования ставит под угрозу стратегию в области культурного наследия, которая направлена на открытие неизведанных территорий: вслед за эфемерными или забытыми публикациями прошлых веков, терпеливо собираемыми библиотекой, новые объекты, которые теперь входят в сферу ее интересов (посты в социальных сетях, блоги, интернет-архивы в целом, NFT и т. д.), вероятно, однажды присоединятся к существующим фондам цифровой библиотеки без какого-либо явного или общего спроса на начальных этапах. В узкой области, ориентированной на удовлетворение спроса, библиотекарь стал бы простым статистиком, анализирующим особенности использования и поведения.

В этом сочетании спроса и предложения роль пользователя приобретает новое измерение. Интеграция системы рекомендаций в поисковую систему позволяет адаптировать предложение в режиме реального времени под запрос каждого конкретного пользователя. Задача состоит в том, чтобы сообщить пользователю о его новом статусе. Цель — сделать его активным и незаменимым партнером в процессе взаимодействия.

Таким образом, библиотека должна проявлять уважение к пользователю, учитывать его особенности и не забывать о своих обязательствах перед ним. Начинается новый этап в истории, который обещает быть интересным и напряженным: во имя соблюдения того же разнообразия, что и в области фондов, мы не должны усугублять цифровой

разрыв между опытными пользователями, которые много читали и изучали, и начинающими, которым также необходимо предоставить доступ к большому объему захватывающих материалов.

Примечания

- ¹ <https://gallica.bnf.fr/>
- ² <https://c.bnf.fr/MrN>
- ³ <https://c.bnf.fr/Ogx>
- ⁴ <https://gallica.bnf.fr/html/und/litteratures/bandedessinee>
- ⁵ <https://data.bnf.fr/>
- ⁶ <https://normalisation.afnor.org/thematiques/intelligence-artificielle>
- ⁷ <https://www.cnil.fr/fr/ethique-et-intelligence-artificielle>
- ⁸ <https://www.arcom.fr/>
- ⁹ https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en
- ¹⁰ <https://en.unesco.org/artificial-intelligence/ethics>
- ¹¹ [https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679R\(02\)&qid=1528814703534&from=en%22](https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679R(02)&qid=1528814703534&from=en%22)
- ¹² <https://www.bnf.fr/fr/mediatheque/en-quoi-lintelligence-artificielle-constitue-pour-lina-une-opportunit-e-pour-renforcer>
- ¹³ <https://www.bnf.fr/fr/agenda/y-t-il-une-intelligence-artificielle>
- ¹⁴ <https://www.culture.gouv.fr/en/Presse/Communique-de-presse/Mission-franco-quebecoise-sur-la-decouvrabilite-lancement-d-un-appel-a-projet-sur-la-decouvrabilite-numerique-des-contenus-culturels-francophones>
- ¹⁵ <https://unesdoc.unesco.org/ark:/48223/pf0000368028>
- ¹⁶ <https://artificialintelligenceact.eu>

Список источников

1. BnF and Artificial Intelligence // Bibliothèque nationale de France. URL: <https://www.bnf.fr/en/feuille-de-route-ia> (дата обращения: 21.03.2024).
2. *Bastard I.* Les publics de la BnF — Synthèse de l'Observatoire 2020 — pour diffusion BnF / Bibliothèque nationale de France. URL: https://www.bnf.fr/sites/default/files/2022-01/Les_publics_de_la_BnF_Synthese_de_l_observatoire_2020.pdf (дата обращения: 21.03.2024).
3. *Pernoo M.* Quelles classifications et quels classements pour les oeuvres de fiction dans les bibliothèques? La question des frontières // Bulletin des bibliothèques de France. 2001. № 1. P. 47–53. URL: <https://bbf.enssib.fr/consulter/bbf-2001-01-0047-003> (дата обращения: 21.03.2024).
4. *Bermès E., Boulet V., Leclaire C.* Making the Most of Library Data on the Web: the data.bnf.fr Example // IFLA World Library and Information Congress, 82nd IFLA General Conference and Assembly. Columbus, 2016. 8 p. URL: <http://library.ifla.org/id/eprint/1447/7/081-bermes-en.pdf> (дата обращения: 21.03.2024).
5. *Alché-Buc, de F., Beaudouin V., Bermès E., Chevallier P., Le Moulllec-Rieux A., Nouvellet A., Prieur C., Roueff F.* Analysis of Gallica and Data BnF logs and Modelling

- of Behaviour Patterns: Presentation of the Main Results. [Research Report] Bibliothèque nationale de France (Paris) ; Télécom ParisTech. 2017. URL: <https://hal.archives-ouvertes.fr/hal-01968742v2> (дата обращения: 21.03.2024).
6. *Majeed A., Lee S.* Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey // IEEE Access. 2021. Vol. 9. P. 8512–8545. DOI: 10.1109/ACCESS.2020.3045700.
 7. *Paraschakis D.* Towards an Ethical Recommendation Framework // 11th IEEE International Conference on Research Challenges in Information Science (10–12 May 2017). DOI: 10.1109/RCIS.2017.7956539.
 8. *Farchy J., Méadel C., Anciaux A.* Une question de comportement. Recommandation des contenus audiovisuels et transformations numériques // tic&société. 2016–2017. Vol. 10, № 2–3. P. 168–198. DOI: 10.4000/ticetsociete.2136.
 9. *Castagnos S., Brun A., Boyer A.* La diversité : entre besoin et méfiance dans les systèmes de recommandation // Revue I3 – Information Interaction Intelligence. 2014. URL: <https://hal.inria.fr/hal-01108998/> (дата обращения: 21.03.2024).
 10. *Stratégie numérique culturelle* // Ministère de la Culture. URL: <https://www.culture.gouv.fr/Thematiques/Innovation-numerique/Strategie-numerique-culturelle> (дата обращения: 21.03.2024).

Перевод **Марии Федотовой**,
Российская государственная библиотека

АНОНС

Конференция BOBCATSSS – 2025: «Искусственный интеллект в библиотечно-информационной науке: исследование взаимосвязи»

Дата проведения: 21–23 января 2025 года
Место проведения: Стамбул, Турция
Веб-сайт: <https://bobcatsss2025.org.tr>

Конференция BOBCATSSS – это ежегодное международное научное мероприятие, посвященное библиотечно-информационной науке и организованное студентами из различных университетов Европы.

Стремительный прогресс в области искусственного интеллекта (ИИ) меняет ландшафт библиотечного дела, открывая беспрецедентные возможности и ставя новые задачи. Этот информационный форум направлен на изучение многогранных взаимоотношений между ИИ и библиотечной отраслью, а также на сбор широкого спектра научных материалов, практических идей и перспективных взглядов на пересечение этих двух динамично развивающихся областей.

Цель конференции – предоставить студентам, исследователям и профессионалам в области библиотечного дела платформу для изложения результатов своих исследований, обмена идеями и обсуждения новых тенденций и проблем отрасли.

Заявки принимаются по указанным ниже направлениям, но не ограничиваются ими:

- применение ИИ в библиотеках и информационных центрах;
- этические соображения и последствия для библиотечной деятельности;
- новые технологии и будущие тенденции;
- ИИ и информационная грамотность;
- ИИ и управление информацией;
- преобразование библиотечных услуг и операций с помощью ИИ;
- ИИ и управление знаниями;

открытая категория: приветствуется подача материалов, которые, возможно, не совсем вписываются в вышеуказанные направления, но вносят существенный вклад в изучение библиотечного дела, в том числе сравнительные исследования, историческая справка об ИИ и библиотеках, междисциплинарные исследования или проекты дизайна будущих библиотек.

Заявки на участие с докладом принимаются до 30 октября 2024 г. по адресу: <https://bobcatsss2025.sciencesconf.org>.

Организаторы:

- Стамбульский университет, факультет информации и управления документами;
- Университет г. Анкары, факультет информации и управления документами;
- Университет им. Гюстава Эйфеля, факультет библиотечно-информационной науки.

BOBCATSSS – Budapest, Oslo, Barcelona, Copenhagen, Amsterdam, Tampere, Stuttgart, Szombathely, Sheffield – Будапешт, Осло, Барселона, Копенгаген, Амстердам, Тампере, Штутгарт, Сомбатхей, Шеффилд.