

Аннотация.

Настоящая работа посвящена проектированию программного средства, позволяющего составлять на основании зафиксированных поисковых запросов пользователя обобщенную модель полученных компетенций и знаний. Представление о своих знаниях полезно при поиске работы и выстраивании образовательной траектории.

Введение.

Современные ИС, несмотря на большие возможности, часто оставляют вопросы эффективности использования и хранения информации конечному пользователю.

Люди ежедневно потребляют большие объемы информации и пытаются структурировать знания, не заботясь о том, чтобы сохранить информационные находки. Однако профессионалы и даже специфические комьюнити не решают проблему распространения информации.

Непрофессионалу или студенту требуется прилагать значительные усилия, чтобы наполнять базу знаний и управлять ею без автоматизированных программных средств поддержки. Также студенты полагаются на опыт и хранение (накопление) артефактов знания, и им нужно проверять сохранность и актуальность информации со временем. Проблема также опосредована появлением новых цифровых СМИ, усложнением цифровых экосистем, но минималистичностью стандартных сервисов. Другая сложность состоит в определении ценности, актуальности и достоверности полученной информации.

Крупные корпорации заинтересованы скорее в усложнении поисковых траекторий пользователей, чем в систематизации найденных ими полезных данных, так как это обусловлено логикой бизнес-модели, основанной на поиске. Поэтому они не поддерживают пользователей разработками для улучшения восприятия и появления структуры или надеются на самостоятельность пользователя. На рынке не обнаружено прямых аналогичных решений, а косвенные аналоги сильно оторваны от непосредственной работы с источниками, требуют дополнительных операций по обработке данных со стороны пользователя и (или) созданы с иными целями.

Исходя из этого, **целью работы** является проектирование инструмента для модификации и манипуляции дерева знаний на основе артефактов браузера, зафиксированных поисковых запросов.

Основная часть.

В работе описана архитектура и основные алгоритмы программного средства, которое анализирует на основе машинного поиска источников знаний, отбираемых и отмеченных небольшим количеством метаданных в процессе информационного поиска пользователем, и затем структурирует эти данные.

Инструмент представляет собой расширение для браузера Chrome и предполагает:

1. доступность и кроссплатформенность решения,
2. персонализацию и конфиденциальность данных,
3. встроенную рекомендательную систему,
4. оптимизированность хранимой структуры посредством автоматизации обработки данных машинным обучением.

Особенность решения состоит в использовании действительной классификации знаний в виде закладок браузера chrome, а также в графическом представлении классификации знаний.

Выбрана клиент-серверная архитектура, реализуемая посредством Chrome API, mongodb, python и js. Основной процесс декомпозирован в три процесса: сбор данных, структурирование, визуализация, соответствующие задачам разработки. Визуализировать дерево знаний предполагается в виде графа. Структурирование будет производиться классификацией лес решений, однако возможна и смена алгоритма в случае низкой точности. Для сбора данных настоящий момент предполагается использование одного алгоритма, который покажет наибольшую эффективность при прототипировании и эксперименте. Выбор подходящего алгоритма анализа текстовых данных производится среди алгоритмов: BERT, DeBERT, PLSA, LDA, ARTM. В результате эксперимента ожидается вывод о необходимости анализа содержимого URL.

Взаимодействие с расширением позволит пользователю самостоятельно выявить недостающие или лишние элементы структуры, а также позволит охватить взглядом смежные (терминологически связанные) непоследовательно следующие части структуры.

Выводы.

Работа представляет собой проектирование расширения браузера для анализа дерева знания, которое будет реализовано в ближайшем будущем. Изучена предметная область, на основе которой выявлены: миссия работы, поставлены цель, практические и исследовательские задачи. Также проведен обзор источников, анализ аналогичных решений, выделены и описаны косвенные источники, в том числе стандартные менеджеры закладок браузеров: Chrome, Mozilla, Vivaldi.

Никончук А.П. (автор)

Береснев А.Д. (научный руководитель)
