

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

УДК 004.383

Кудин
Никита Игоревич

**АВТОМАТИЗИРОВАННАЯ СИСТЕМА
КРАУДСОРСИНГА ДАННЫХ**

АВТОРЕФЕРАТ

диссертации на соискание степени
магистра информатики и вычислительной техники

по специальности 1-40 81 01 - Информатика и технологии
разработки программного обеспечения

Научный руководитель
Калугина М.А.
к.ф-м.н., доцент

Минск 2020

Работа выполнена на кафедре информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **КАЛУГИНА Марина Алексеевна**,
кандидат физико-математических наук, доцент,
доцент кафедры информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Рецензент: **ЧУБАРОВ Сергей Ильич**,
кандидат физико-математических наук, доцент,
автор более 150 научных работ, в том числе 3
учебных пособия, преподаватель кафедры информационных технологий учреждения образования «Белорусский государственный педагогический университет имени Максима Танка»

Защита диссертации состоится « » июня 2020 г. года в часов на заседании Государственной экзаменационной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, Минск, ул. Гикало, 9, копр. 4, ауд. 112, тел. 293-85-91, e-mail: inform@bsuir.by

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

ВВЕДЕНИЕ

В 2003 году Луис фон Ах (Luis von Ahn) вместе со своими коллегами впервые предложил понятие "человеческих вычислений" (human computation), которое оперирует возможностями человека для выполнения вычислительных задач, неподвластных компьютеру. Позже в 2006 году термин краудсорсинг (crowdsourcing) был сформулирован редактором журнала "Wired" Джеффом Хау (Jeff Howe) в статье под названием "Восход Краудсорсинга" (The Rise of Crowdsourcing). С тех пор было опубликовано множество работ, посвященных различным аспектам краудсорсинга, сферы его применимости и экономической эффективности.

Краудсорсинг — это мобилизация ресурсов людей посредством информационных технологий с целью решения задач, стоящих перед бизнесом, государством и обществом в целом.

Одним из первых примеров краудсорсинга было составление Оксфордского Английского Словаря (Oxford English Dictionary). Тогда издание обратилось к общественности с просьбой прислать варианты терминов с их потенциальными вариантами использования. В результате за период в 70 лет было получено более 6 млн. писем.

На сегодняшний день существует множество способов его применения, которые различаются как по тематике и виду получаемого результата, так и по категории привлекаемых людей. Так, краудсорсинг уже несколько лет используется различными западными компаниями для решения производственных задач, привлекая к этому широкий круг добровольцев. Примерами таких компаний являются British Gas, IAG, British Airways, Diageo, Argos, Virgin Trains и другие.

Краудсорсинг в настоящее время активно развивается в качестве модели для решения любого вида проблем и задач, стоящих как перед бизнесом, так и перед государством и обществом в целом. В рамках парадигмы краудсорсинга решение задачи передается распределенной и очень многочисленной группе людей, за счет чего стоимость и время достижения результата радикально снижаются.

Коммуникация и деятельность исполнителей осуществляется благодаря информационным технологиям, часто на безвозмездной основе. Для компаний-заказчиков это возможность привлечь человеческие ресурсы, ведь в выполнении задания участвует множество лиц. Добровольцы, выступая в качестве исполнителей, получают шанс реализовать свою идею, пополнить портфолио и приумножить опыт работы.

В итоге значение имеет лишь выбранный конечный продукт, а вот опыт работы, образование и квалификация исполнителей не так уж важны. Все права на полученный в результате продукт принадлежат компании-заказчику, которая имеет право по своему усмотрению продать этот товар по выгодной цене.

Несмотря на то, что краудсорсинг — это лишь разновидность аутсор-

синга, здесь есть свои характерные отличия: отсутствие юридических договоров между исполнителями и заказчиком, выполнение работы безвозмездно или за низкую цену, при наличии крупного проекта, дробление его на микрозадачи, которые затем формируются в отдельные проекты.

Конкурсные платформы — это один из видов площадок для краудсорсинга. Здесь заказчик создает конкурс, размещая определенное задание. Исполнители присылают свои варианты решения поставленной задачи, одновременно комментируя и оценивая работу друг друга. Таким образом, отбирается одна или несколько лучших работ. Работа-победитель приносит своему создателю денежное вознаграждение.

Краудсорсинг идей стал крайне популярен в период активного роста инновационных продуктов ввиду того, что с каждым годом на мировом рынке высокотехнологичных продуктов конкуренция становится все серьезнее.

Ввиду вышеуказанных фактов разработка краудсорсинговой платформы является не только экономически выгодным, но и крайне перспективной.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

В современное время существует множество проблем, связанных с поиском инновационных идей в крупных компаниях. Данные проблемы являются результатом серьезной конкуренции на рынке. С одной стороны, это прекрасная возможность для разработки новых продуктов для пользователей. С другой стороны, в крупных компаниях рождается серьезный кризис идей.

Один из важнейших факторов — это проблема оперативности. Данная проблема является одной из самых острых из-за того, что в условиях высокой конкуренции на рынке необходимо предоставлять новые продукты для пользователя оперативно и качественно, и, следовательно, на это необходимо выделять как можно больше ресурсов и времени.

Также немаловажным фактором решения проблемы является времязатратность. Крупные организации, которые имеют в своем штате несколько тысяч человек, могут затратить большое количество времени, чтобы узнать какие существуют идеи сотрудников. Опрос всех сотрудников может занять долгие месяцы или даже годы.

Для решения данного рода проблем созданы автоматизированные системы краудсорсинга, которые на основе настроек пользователя собирают и анализируют данные из открытых источников с помощью открытого API различных сервисов.

Цель и задачи исследования

Объект исследования – краудсорсинг идей как неотъемлемая составляющая процесса поиска инновационных идей.

Предмет исследования – краудсорсинг как механизм оптимизации рабочих процессов, его направления развития.

Цель работы: изучение теоретических основ и создание высоконагруженной системы в области краудсорсинга, а также исследование перспектив развития краудсорсинга.

Методы исследования: сравнительный анализ существующих аналогов, анализ перспективных ниш рынка.

Исследования и разработки: исследованы теоретические основы краудсорсинга, методы и приемы его анализа, рассмотрена практика его осуществления на примере существующих аналогов; разработана распределенная высоконагруженная система краудсорсинга данных с модулем автоответов и анализа данных в режиме реального времени.

Область возможного применения: деятельность ООО «Сайдвейс 6 Девелопмент», ООО «Польмед» по осуществлению синхронизации данных с Yammer и Office 365.

Автор работы подтверждает, что приведенный в данной работе расчетно-аналитический материал правильно и объективно отражает состояние исследуемого процесса, а все заимствованные из литературных и других источников теоретические, методические положения и концепции сопровождаются ссылками на их авторов.

Область исследования

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) специальности 1-40 81 01 «Информатика и технологии разработки программного обеспечения».

Теоретическая и методологическая основа исследования

В основе данной работы лежат теоретические и практические исследования экспертов в области разработки процесса краудсорсинга, а также статьи ведущих компаний.

Информационная база исследования сформирована на основе литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Основные положения, выносимые на защиту

1. Определение роли автоматизации краудсорсинга в контексте крупных организаций.
2. Использование открытых API для источников данных как средство для увеличения пользовательской активности и охвата аудитории разрабатываемой автоматизированной системы краудсорсинга.
3. Систематизация процесса проектирования и разработки высоконагруженного приложения.
4. Использование облачных платформ и сервисов как неотъемлемая часть современной разработки.

Апробация диссертации и информация об использовании ее результатов и публикации

Основные результаты диссертационного исследования были представлены на 56-ой «Научной Конференции Аспирантов, Магистрантов и Студентов БГУИР». Темой являлась «Автоматизированная система краудсорсинга данных». Данная тема раскрывала особенности использования данных систем в решении задач различных компаний и организаций. Выступление получило благодарность от комиссии конференции.

Внедрение результатов подтверждается актами о внедрении в организации ООО «Сайдвэйс 6 Девелопмент» и ООО «Польмед», которые применили систему краудсорсинга в своей деятельности, что в свою очередь принесло этим организациям как экономический, так и социальный эффект.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В общей характеристике работы описывается цель исследования и разработки, основные задачи, объект и предмет исследования, основные положения диссертации, выносимые на защиту, теоретическую и практическую значимость результатов исследования, а также личный вклад магистранта с апробацией результатов.

Во введении рассматривается процесс краудсорсинга с исторической точки зрения. Данный раздел диссертации содержат основные даты и фамилии, которые повлияли на формирование терминологии краудсорсинга. Также данный раздел раскрывает главные этапы формирования процесса автоматизации краудсорсинга.

В первой главе приведен пример исследовательской составляющей. Данная глава описывает различные обоснования актуальности темы диссертации, апробацию результатов, определение целей и задач диссертационной работы. Также глава включает анализ существующих аналогов на рынке и определяет плюсы и минусы каждого решения.

Также данная глава раскрывает особенности выбора инструментов разработки, хранения данных и хостинга автоматизированной системы. Данный раздел описывает основные плюсы выбранного языка программирования, а также его основных расширений, которое могут быть интегрированы в разработанное приложение.

Завершается глава краткими выводами о проделанной работе.

Во второй главе описывается моделирование предметной области с ее описанием. Ключевыми моментами данной главы являются проектирование модели приложения через различные диаграммы (IDEF0, вариантов использования). Кроме того, данная глава описывает правила декомпозиции диаграммы IDEF0. Данная часть представляет наиболее из важнейших шагов при построении системы, так как правильно созданные диаграммы IDEF0 верхнего уровня помогут корректно сделать декомпозицию системы на более низких уровнях.

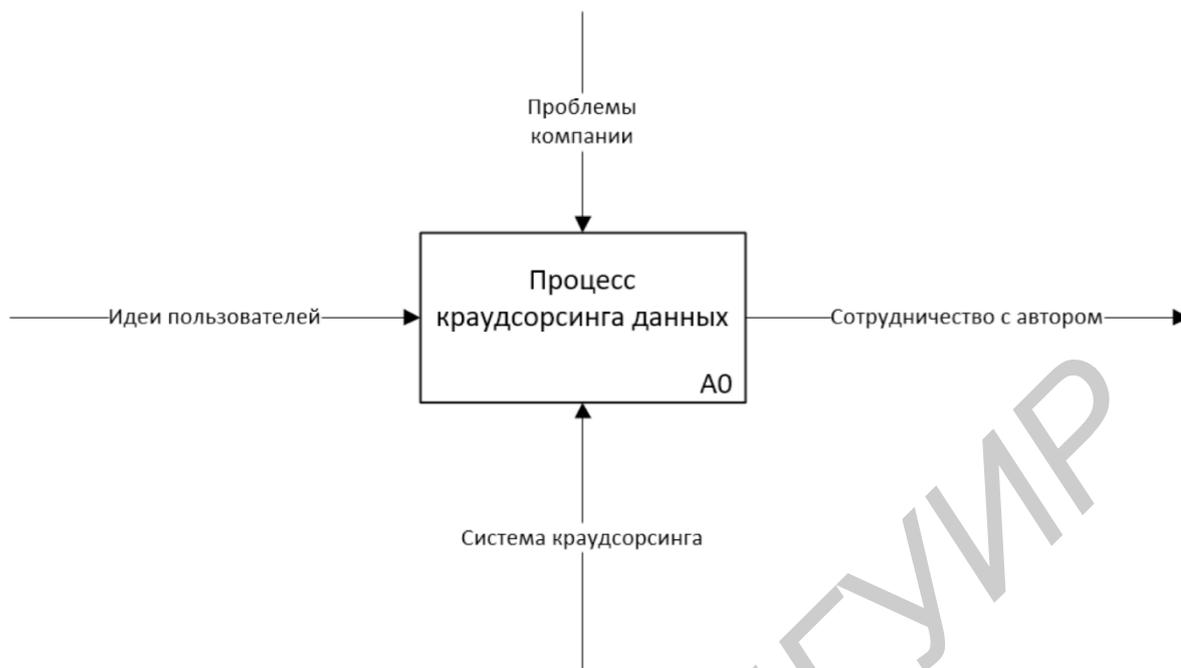


Рисунок 1 – Разработанная диаграмма IDEF0 для автоматизированной системы краудсорсинга

Кроме того, данный раздел содержит в себе основную терминологию информационной модели предметной области, с помощью которой производится построение реляционной базы данных на основе связей между сущностями.

Диаграмма вариантов использования, построение которой описывается в данной главе, помогает графически изобразить логические действия различных пользователей системы. На данном этапе проясняется структура приложения и его логика. В результате были получены следующие сущности системы:

1. Пользователь системы краудсорсинга.
2. Модуль задач.
3. Автор идеи.

Диаграмма вариантов использования содержит конечное множество вариантов использования, которые в целом должны определять все возможные стороны ожидаемого поведения системы. Для удобства множество вариантов использования может рассматриваться как отдельный пакет. Применение вариантов использования на всех этапах работы над проектом позволяет не только достичь требуемого уровня унификации обозначений для представления функциональности подсистем и системы в целом, но и является мощным средством последовательного уточнения требований к проектируемой системе на основе их итеративного обсуждения со всеми заинтересованными специалистами.

ЗАКЛЮЧЕНИЕ

Задачи и цели, которые были достигнуты в период проектирования и реализации автоматизированной системы краудсорсинга:

1. Произведен анализ влияния автоматизированных систем краудсорсинга в крупных организациях. Выявлено, что внедрение подобных систем способно благоприятно повлиять на бизнес-климат компании, а также улучшить социальную составляющую.

2. Выполнено внедрение внешних программных интерфейсов приложений (API) в систему краудсорсинга. Это помогло многократно увеличить количество активных пользователей, которые отправляют идеи в случае грамотного информирования пользователей о предстоящем процессе краудсорсинга.

3. Были систематизированы знания процесса разработки, которые включает в себя различные этапы проектирования, написания выполняющегося кода и тестирования.

4. Были использованы облачные платформы, такие как Microsoft Azure. Это является логическим шагом при построении различных высоконагруженных систем, когда необходимо получить высокие показатели доступности сервисов и оптимизировать расходы разработки.

Были изучены материалы о краудсорсинге данных, проанализированы существующие аналоги. Таким образом, была разработана система высоконагруженного сервиса агрегации идей из различных источников с использованием современных технологий.

Для серверной части продукта был разработан динамический менеджер задач приложения, который помог уменьшить нагрузку на веб часть приложения.

Функциональность приложения была расширена динамическими настройками, которые помогут пользователю создавать наиболее подходящую структуру каждой компании.

Также разработан специализированный модуль приложения, который упростил коммуникацию с пользователями сервиса. Данный модуль представляет собой настраиваемую систему внутренних триггеров с гибкой настройкой.

По окончании разработки было произведено тестирование приложения, которое показало хорошие результаты как с нагрузочными тестами, так и на тесты проникновения в систему.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Статьи в рецензируемых журналах

1. Кудин, Н.И. Автоматизированная система краудсорсинга данных / Н.И. Кудин, // Молодой ученый. – 2020. – № 21(311). – С. 54-56.
2. Кудин, Н.И. Автоматизированная система создания и редактирования этикеток для химической продукции / Д.И. Ярошевич, М.А. Сабадаш, Н.И. Кудин // Modern science: problems and innovations. Abstracts of the 3rd International scientific and practical conference. SSPG Publish. Stockholm, Sweden. – 2020. – с. 254-257.

Тезисы конференций

1. Кудин, Н.И. Автоматизированная система краудсорсинга данных / Н.И. Кудин, М.А. Сабадаш // материалы 56-ой науч. конф. аспирантов, магистрантов и студентов «Проектирование информационно-компьютерных систем», Минск, Респ. Беларусь, 18–20 мая 2020 г. / УО «БГУИР». – Минск, 2020. – с.149–150.