

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

УДК 005.336.6:004.732.6

Степанов Антон
Александрович

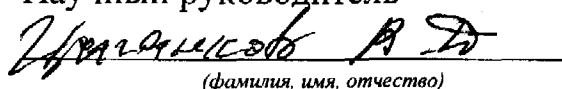
Автоматизированная система обработки данных в проектной деятельности

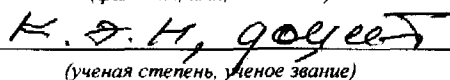
АВТОРЕФЕРАТ

на соискание степени магистра информационных технологий
по специальности 1-40 81 05 «Информационно-коммуникационные
технологии в экономике»


(подпись магистранта)

Научный руководитель


(фамилия, имя, отчество)


(ученая степень, ученое звание)


(подпись научного руководителя)

Минск 2019

Краткое введение

Сегодня новое время, время информатизации и компьютеризации. Информация – ценнейший ресурс наравне с такими традиционными видами ресурсов, как нефть, газ, полезные ископаемые и др., а значит, процесс ее обработки по аналогии с процессами обработки материальных ресурсов можно воспринимать как технологию. Процесс обработки информации с каждым годом становится все быстрее благодаря новым методам обработки информации и растущей вычислительной мощностью серверов. В тоже время растет и количество производимой информации. С каждым годом все больше видов деятельности становятся перегруженными информацией. Управление проектами – одна из самых перегруженных информацией область деятельности.

Из-за растущего объёма информации построение рациональных и оптимальных моделей управления проектной деятельностью является достаточно сложной задачей, однако с решением данной задачи хорошо справляются современные технологии и подходы сбора, обработки и визуализации больших объёмов данных.

Первым этап подразумевает сбор данных из разных источников и их предварительную подготовку. Предварительная обработка данных является важным шагом в процессе анализа больших объёмов данных. *Garbage In, Garbage Out*, «Мусор на входе — мусор на выходе» – термин означающий, что при неверных входящих данных будут получены неверные результаты, даже если сам по себе алгоритм правилен. Таким образом, предварительно обрабатывая данные на этапе сбора данных, позволяют улучшить точность конечного результата обработки.

Этап обработки данных – ядро процесса анализа данных, на данном этапе новые знания или свойства получают путем сложения разных факторов в один.

Визуализации данных является важной составной частью качественных систем анализа данных, особенно ориентированных на обработку больших объёмов информации.

Таким образом цель данной работы – предложить оптимальный набор методологий и инструментов для обработки и визуализации данных проектной деятельности.

Для достижения поставленной цели требуется решить следующие задачи:

- изучить основные методологии управления проектами;

- изучить и описать процесс управления IT-проектами на предприятии ИООО «ЭПАМ Системз»;
- провести сравнительный анализ современных подходов сбора и обработки данных
- изучить, с последующим анализом, основных методы, подходы и технологии обработки больших объемов данных.

Общая характеристика работы

Проект знакомит с основными управления проектами, поднимает проблемы, с которыми сталкиваются компании, занятые в сфере разработки программных продуктов, отражает методы решения этих проблем, а также описывает способы, которые позволяют более эффективно управлять проектной деятельностью.

В данной работе изучены основные способы решения проблемы обработки данных проектной в деятельности, для этих целей детально описаны способы сбора данных, методы обработки данных и их визуализацию. В результате выполненной работы были определен оптимальный набор методов, технологии и подходов для решения такой сложной задачи как обработка данных в проектной деятельности.

Краткое содержание работы

В первой главе описываются теоретические аспекты управления проектами. Для этого были изучены основные проблемы, с которыми сталкиваются менеджеры проектов и методологии созданные с целью решить эти проблемы и максимально упростить процесс управления проектами. Детально были изучены такие методологии как scrum и agile.

Вторая глава посвящена описанию процесса управления проектами в компании ИООО «ЭПАМ Системз». В начале второй главы дается краткое описание что представляет из себя компания ЕПАМ в конце 2018 года. После чего дается подробное описание процесса управления проектной деятельности по разработке программного обеспечения на предприятии. Заканчивается вторая глава небольшим отчетом от том каким образом компания использует программное решения Jira для упрощения процесса управления проектной деятельностью.

Третья глава просвещена технической стороне решения проблемы обработки данных в проектной деятельности. И включает в себя 3 основных пункта:

- сбор и обработка данных;

- поиск и фильтрация данных
- визуализация данных;

В подглаве посвященной сбору и обработке данных кратко описаны и два главных подхода сбора и обработки больших объемов данных: Extract Load Transform и Extract Transform Load. В следующей подглаве посвященной фильтрации и поиску произведен сравнительных анализ основным средств полнотекстового поиска. Также во второй главе кратко описан процесс поиска в большом объеме данных. Процесс визуализации подробно расписан в третьей главе. Приведены основные виды визуализации и примеры использования. Каждая из визуализация сопровождается иллюстрацией. В последней подглаве «Пример использования системы обработки данных проектной деятельности» приведен пример использования программного продукта, созданного на основе рекомендаций приведенных в первых трех подглавах.

Заключение

Изучив результаты работы, можно сделать вывод, что все поставленные задачи были выполнены в полном объеме. В результате написания работы были изучены основные методологии управления проектами, выявлены основные преимущества и недостатки.

Более тщательно были изучены и описаны методологии, наиболее часто используемые в IT-сфере. Во время выявления основных этапов и ключевых особенности управления IT-проектами на примере предприятия ИООО «ЭПАМ Систем» было изучено и использовано на практике программное средство управления IT-проектами – Atlassian JIRA.

Далее были изучены основные технологии и методологии сбора, подготовки и обработки данных больших объемов. Были исследованы и описаны такие подходы как Business Intelligence и Big Data, дополнительно был описан процесс MapReduce.

Так же были изученные программные продукты и решения, используемые в ИООО «ЭПАМ Систем» для обработки, хранения и визуализации данных проектной деятельности.

Подробно был изучен процесс трансформации больших объёмов данных. Был произведен сравнительный анализ двух основных подходов обработки данных: «Extract Transform Load» и «Extract Load Transform».

Изучены основные средства хранения данных. Наиболее подробно были рассмотрены особенности базы данных NoSQL, описаны основные

особенности и отличия от традиционных баз данных. В качестве примера использовалась MongoDB.

Подробно описан процесс фильтрации больших объемов данных и полнотекстовый поиск. Были изучены основные системы полнотекстового поиска: Lucene, Elasticsearch, Sphinx.

Были описаны наиболее часто применяемые виды визуализации с описанием их назначения и примерами. Также был приведен пример использования системы анализ проектной деятельности, используемой в ИООО «ЭПАМ Систем», приведены скриншоты работы программы для.

Список опубликованных работ:

– Степанов, А. А. Программная поддержка управления ИТ-проектами с возможностью рассылки уведомлений / А. А. Степанов // Проблемы экономики и информационных технологий: материалы 53-й научной конференции аспирантов, магистрантов и студентов (Минск, 02 – 06 мая 2017 г.). – Минск: БГУИР, 2017. – С. 53 – 54.