

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ РАДИОФИЗИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ
Кафедра интеллектуальных систем**

Аннотация к дипломной работе

**АВТОМАТИЗАЦИЯ СБОРА И ОБРАБОТКИ ДАННЫХ В
КОРПОРАТИВНЫХ ХРАНИЛИЩАХ ДАННЫХ**

Лобков Дмитрий Сергеевич

Научный руководитель: старший преподаватель Д.С. Штукатер

Минск, 2025

РЕФЕРАТ

Дипломная работа содержит 88 с., 10 рис., 6 табл., 65 источн., 2 прил.

АВТОМАТИЗАЦИЯ, АНАЛИТИЧЕСКАЯ ОТЧЕТНОСТЬ, СБОР ДАННЫХ, ОБРАБОТКА ДАННЫХ, КОРПОРАТИВНОЕ ХРАНИЛИЩЕ ДАННЫХ, ВИТРИНЫ ДАННЫХ, ВИЗУАЛИЗАЦИЯ ДАННЫХ

Объект исследования – процесс автоматизации сбора и обработки данных в существующей системе корпоративного хранилища данных Национального банка Республики Беларусь для подготовки информации к аналитической визуализации.

Цель работы – разработка и описание процесса автоматизации сбора, обработки и подготовки данных в корпоративном хранилище данных для последующей визуализации аналитической отчетности, а также демонстрация его применения на практических кейсах в рамках существующей ИТ-инфраструктуры.

Методы исследования – компьютерное моделирование, эксперимент.

В работе проведен анализ теоретических основ автоматизации аналитической отчетности в Национальном банке Республики Беларусь. Выявлены недостатки существующих ручных процессов сбора и обработки данных, обосновывается необходимость разработки автоматизированного процесса для подготовки информации к визуализации.

Приводится описание разработанного процесса автоматизации сбора и обработки данных в рамках существующей архитектуры корпоративного хранилища данных Национального банка с использованием стека технологий Apache NiFi, Greenplum и Apache Superset. Описана практическая реализация процесса «Расчетные величины стандартного риска», включая формирование аналитических витрин и создание интерактивных дашбордов.

Показано, что предложенный автоматизированный процесс позволяет значительно сократить трудозатраты, повысить оперативность и качество аналитической информации, минимизировать ошибки, связанные с человеческим фактором.

Разработанный процесс и его элементы внедрены в практические решения для Национального банка Республики Беларусь, что подтверждено актом внедрения. Результаты могут быть использованы для автоматизации широкого спектра аналитической отчетности в организациях с аналогичной ИТ-инфраструктурой.

РЭФЕРАТ

Дыпломная праца змяшчае 88 ст., 10 мал., 6 табл., 65 крыніц, 2 дадатка.

АЎТАМАТЫЗАЦЫЯ, АНАЛІТЫЧНАЯ СПРАВАЗДАЧНАСЦЬ, ЗБОР ДАДЗЕННЫХ, АПРАЦОЎКА ДАДЗЕННЫХ, КАРПАРАТЫЎНАЕ СХОВІШЧА ДАДЗЕННЫХ, ВІТРЫНЫ ДАДЗЕННЫХ, ВІЗУАЛІЗАЦЫЯ ДАДЗЕННЫХ

Мэта працы – распрацоўка і апісанне працэсу аўтаматызацыі збору, апрацоўкі і падрыхтоўкі дадзеных у карпаратыўным сховішчы дадзеных для наступнай візуалізацыі аналітычнай справаздачнасці, а таксама дэманстрацыя яго прымянення на практычных кейсах у рамках існуючай ІТ-інфраструктуры. Аб'ектам даследавання з'яўляецца працэс аўтаматызацыі збору і апрацоўкі дадзеных у існуючай сістэме карпаратыўнага сховішча дадзеных Нацыянальнага банка Рэспублікі Беларусь для падрыхтоўкі інфармацыі да аналітычнай візуалізацыі.

У працы праведзены аналіз тэарэтычных асноў аўтаматызацыі аналітычнай справаздачнасці і даследавана прадметная вобласць Нацыянальнага банка Рэспублікі Беларусь. Выяўлены недахопы існуючых ручных працэсаў збору і апрацоўкі дадзеных, што абгрунтавала неабходнасць распрацоўкі аўтаматызаванага працэсу для падрыхтоўкі інфармацыі да візуалізацыі.

Распрацаваны і апісаны дэталізаваны працэс аўтаматызацыі збору і апрацоўкі дадзеных у рамках існуючай архітэктury карпаратыўнага сховішча дадзеных Нацыянальнага банка з выкарыстаннем стэка тэхналогій Apache NiFi, Greenplum і Apache Superset. Практычная рэалізацыя працэсу прадэманстравана на кейсе «Разліковыя велічыні стандартнай рызыкi», уключаючы фарміраванне аналітычных вітрын і стварэнне інтэрактыўных дашбордаў.

Паказана, што прапанаваны аўтаматызаваны працэс дазваляе значна скараціць працавыдаткі, павысіць аператыўнасць і якасць аналітычнай інфармацыі, мінімізаваць памылкі, звязаныя з чалавечым фактарам.

Распрацаваны працэс і яго элементы ўкаранёны ў практычныя рашэнні для Нацыянальнага банка Рэспублікі Беларусь, што пацверджана актам укаранення. Вынікі могуць быць выкарыстаны для аўтаматызацыі шырокага спектра аналітычнай справаздачнасці ў арганізацыях з аналагічнай ІТ-інфраструктурай.

ABSTRACT

Diploma thesis contains 88 p, 10 fig., 6 tables, 65 sources, 2 appendices.

AUTOMATION, ANALYTICAL REPORTING, DATA COLLECTION, DATA PROCESSING, CORPORATE DATA WAREHOUSE, DATA MARTS, DATA VISUALIZATION

The aim of the work is to develop and describe the process of automating data collection, processing, and preparation in a corporate data warehouse for the subsequent visualization of analytical reports, as well as to demonstrate its application through practical case studies within the existing IT infrastructure. The object of research is the process of automating data collection and processing within the existing corporate data warehouse system of the National Bank of the Republic of Belarus to prepare information for analytical visualization.

The thesis analyzes the theoretical foundations of analytical reporting automation and investigates the subject area of the National Bank of the Republic of Belarus. Shortcomings of existing manual data collection and processing procedures were identified, justifying the need to develop an automated process for preparing information for visualization.

A detailed process for automating data collection and processing within the existing corporate data warehouse architecture of the National Bank was developed and described, utilizing a technology stack including Apache NiFi, Greenplum, and Apache Superset. The practical implementation of the process was demonstrated through the case study "Calculated Values of Standard Risk" including the formation of analytical data marts and the creation of interactive dashboards.

It has been shown that the proposed automated process significantly reduces labor costs, increases the timeliness and quality of analytical information, and minimizes errors associated with human factors.

The developed process and its elements have been implemented into practical solutions for the National Bank of the Republic of Belarus, as confirmed by an implementation act. The results can be used to automate a wide range of analytical reports in organizations with similar IT infrastructure.