

## **ИСПОЛЬЗОВАНИЕ АНАЛИТИЧЕСКИХ ПЛАТФОРМ DEDUCTOR И LOGINOM ПРИ ПОДГОТОВКЕ БАКАЛАВРОВ НАПРАВЛЕНИЯ «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ»**

*Чернышева Кира Владимировна, доцент кафедры прикладной информатики, ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева*

*Афанасьева Светлана Ильинична, доцент кафедры прикладной информатики, ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева*

*Карпузова Надежда Васильевна, доцент кафедры прикладной информатики, ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева*

***Аннотация.** В статье рассматриваются вопросы практического применения отечественных аналитических платформ Deductor и Loginom при подготовке бакалавров направления «Информационные системы и технологии».*

***Ключевые слова:** аналитические платформы, OLAP, KDD, Data Mining, ETL, ELT, хранилища данных, визуализаторы данных.*

В настоящее время согласно «Стратегии развития отрасли информационных технологий в Российской Федерации на 2014-2020 годы и на перспективу до 2025 года» одной из основных точек роста сегмента разработки программного обеспечения становятся технологии обработки больших массивов данных [1]. Это обуславливает необходимость использования современных аналитических платформ при подготовке бакалавров и магистров направлений «Экономика», «Менеджмент», а также «Информационные системы и технологии».

В РГАУ-МСХА имени К.А. Тимирязева осуществляется подготовка бакалавров направления «Информационные системы и технологии» направленности «Информационные технологии анализа данных». Согласно требованиям ФГОС ВО по направлению подготовки 09.03.02 «Информационные системы и технологии» и учебного плана по данному направлению предусмотрено изучение дисциплины «Информационные технологии». Данная дисциплина включена в обязательную часть учебного плана и задействована в формировании у обучающихся следующих общепрофессиональных компетенций ОПК-2 и ОПК-7. Компетенция ОПК 2 предусматривает знание, умение и владение студентом современными информационными технологиями, в том числе отечественного производства, при решении задач профессиональной деятельности. В перечень профессиональных стандартов, соответствующих профессиональной деятельности выпускников, освоивших программу бакалавриата по направлению подготовки 09.03.2 «Информационные системы и технологии», входит и профессиональный стандарт «Специалист по информационным

системам»[2]. Одной из указанных в стандарте групп занятий является «Разработчики и аналитики информационных систем». Кроме того, в данном стандарте имеется обобщенная трудовая функция «Выполнение работ и управление работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес-процессы», возможной должностью по которой является «Бизнес-аналитик». Необходимыми знания являются знания предметной области автоматизации, в т.ч. источников информации, знание систем хранения и анализа баз данных, программных средств и платформ информационной инфраструктуры организации, методик описания и моделирования бизнес-процессов, средств моделирования бизнес-процессов, инструментов и методов моделирования бизнес-процессов организации и пр.

Исходя из вышесказанного, при изучении дисциплины «Информационные технологии» студенты изучают теоретические основы и приобретают практические навыки работы с отечественными аналитическими платформами Deductor Academic и Loginom. В данных программных продуктах реализованы технологии: ETL (Extract, Transform and Load), ELT, OLAP (Online Analytical Processing) , Data Mining, KDD (Knowledge Discovery in Databases) для работы с хранилищем данных и оперативным складом данных.

Студенты выполняют проектирование различных сценариев обработки данных. Это анализ сельскохозяйственного производства в разрезе районов Московской области за два года, прогнозирование временных рядов, построение деревьев решений, проектирование и анализ данных хранилища организации по продажам [3].

Важнейшим процессом, реализуемым в любом аналитическом проекте, является извлечение данных из различных источников, их преобразование к единому формату и модели данных, очистка от дубликатов, противоречий и других факторов, расчет каких-либо показателей, загрузка в единый интегрированный источник. Для реализации данного процесса предназначена специальная технология преобразования данных, называемая ETL (англ.: *Extract, Transform, Load* – *извлечение, преобразование, загрузка*). При проектировании сценариев обработки данных по анализу сельскохозяйственного производства в разрезе районов Московской области за два года источником данных является база данных по Московской области, представляющая собой файл MS Excel и содержащая информацию по сельскохозяйственным организациям. ETL- технологии подготовки данных могут выполняться двумя способами. Первый способ заключается в выполнении процессов преобразования данных в MS Excel. Это расчет промежуточных итогов по районам Московской области, организация расчетных полей и пр. Второй способ предусматривает выполнение процессов преобразования данных непосредственно в аналитической платформе. Применяются такие методы преобразования данных как настройка набора данных (параметры полей), замена данных, группировка, сортировка, фильтры строк, вычисляемые значения (калькулятор),

соединение, объединение наборов данных. Данный способ является наиболее приближенным к реальным аналитическим проектам, использующим «сырые», неподготовленные данные.

Непосредственно в аналитических платформах производится обработка подготовленных данных с использованием OLAP и Data Mining технологий. Выполняется кластеризация районов Московской области с визуализацией OLAP- куб и профили кластеров.

При выполнении задания по прогнозированию значений временного ряда исходные данные загружаются в аналитические платформы с официального сайта Федеральной службы государственной статистики [www.gks.ru](http://www.gks.ru). ETL- технологии выполняются в аналитических платформах и включают транспонирование данных, преобразование упорядоченных данных (скользящее окно). Прогнозирование выполняется по модели линейной регрессии или декомпозиции временного ряда (ARIMAX), учитывающей сезонную и трендовую составляющие.

Важным для реализации заявленной компетенции является выполнение задания по проектированию хранилища данных. Реляционное хранилище данных (ROLAP) типа «Снежинка» проектируется в аналитической платформе Deductor Academic, наполняется из источников данных, подготовленных на основе демобазы «Борей». Анализ данных хранилища с применением OLAP и Data Mining технологий возможен как в аналитической платформе Deductor Academic, так и в Loginom [4].

Таким образом, использование в учебном процессе отечественных аналитических платформ Deductor Academic и Loginom позволяет реализовать заявленные образовательным стандартом и учебным планом подготовки бакалавров направления «Информационные системы и технологии» компетенции, а также подготовить обучающихся к будущей профессиональной деятельности.

### **Библиографический список**

1. Стратегия развития отрасли информационных технологий в Российской Федерации на 2014 - 2020 годы и на перспективу до 2025 года.
2. Профессиональный стандарт «Специалист по информационным системам» : утв. М-вом труда и социальной защиты Рос. Федерации от 18.11.2014 г. N 896н.
3. Информационные системы и технологии в менеджменте АПК : учебное пособие / В. И. Карпузова, Э. Н. Скрипченко, К. В. Чернышева, Н. В. Карпузова. – Москва : БИБКМ, ТРАНСЛОГ, 2016. - 462 с.
4. Карпузова, В. И. Формирование и использование хранилищ данных в АПК : материалы IX научно-практической конференции / В. И. Карпузова, К. В. Чернышева, Н. В. Карпузова. – Москва : ФГБНУ «Росинформагротех», 2017. – С. 395-399.