

УДК 004

ИНФОРМАЦИОННАЯ СИСТЕМА ИССЛЕДОВАНИЯ РЕЙТИНГОВ
УНИВЕРСИТЕТА

Н. А. ЛАПИН, Д. Ю. МАЖУГОВ
Научный руководитель И. Г. ПЛИСКО
Белорусско-Российский университет
Могилев, Беларусь

Публикация рейтингов является мощным и эффективным инструментом для запуска и консолидации процессов изменения в научных кругах, увеличения приверженности ученых и создания необходимых долгосрочных стратегий в развитии научного потенциала вуза.

Следить одновременно за несколькими сайтами-рейтингами является достаточно сложной задачей. У каждого рейтинга своя периодичность обновления данных. В результате сбор данных вручную занимает большое количество времени, особенно если анализ происходит по университетам из нескольких стран. Собранные данные необходимо анализировать, чтобы получить возможность прогнозирования и управления рейтингом университета.

Одним из параметров рейтинга, например, является Преподавание (Teaching) (40 %), который может быть представлен следующими оценками: Количество студентов на одного преподавателя (8 %); Количество преподавателей, имеющих степень бакалавриата (8 %); Количество докторских степеней, присужденных преподавателям (8 %); Количество докторских степеней, присуждаемых на одну степень бакалавра (8 %); Мировая репутация преподавателей (8 %).

Для автоматизации сбора и обработки информации о рейтингах университета поставлены и решены следующие задачи: разработка алгоритмов сбора данных с каждого сайта-рейтинга; экспорт данных в программу для работы с электронными таблицами Excel и последующим анализом; разработка подсистемы первичной обработки собранных данных с целью обработки некорректной информации и ошибок работы алгоритма сбора; разработка системы анализа собранных данных с помощью алгоритмов факторного анализа, который позволит получить коэффициенты влияния учитываемых параметров на конечный результат.

Для выполнения поставленных задач используются следующие технологии:

– сбор данных – библиотека React.js; библиотека axios для выполнения запросов; библиотека cheerio для анализа и манипулирования результирующей структурой данных; библиотека.xlsx для работы с файлами Excel;

– анализ данных – язык программирования Python; библиотека scikit; библиотека pandas.