

РАЗРАБОТКА МОДУЛЯ СИНХРОНИЗАЦИИ ДАННЫХ НАУКОМЕТРИЧЕСКИХ ПОКАЗАТЕЛЕЙ С БИБЛИОГРАФИЧЕСКИХ БАЗ ДАННЫХ

*А.О.Савельев, к.т.н., доц. ОИТ ИШИТР,
Д.С.Меденцева, студентка, гр8ИИМ92,
А.А.Сутягина, студентка, гр8ИИМ92,
Томский политехнический университет
E-mail: sava@tpu.ru*

Введение

В последнее время все большее значение уделяется вопросам повышения эффективности научной деятельности, разрабатываются критерии оценки и показатели, демонстрирующие, как работают отдельные ученые, группы ученых, институты и ВУЗы.

Для проведения оценки эффективности научной деятельности отдельных авторов или целых научных сообществ на данный момент существует огромное число ключевых показателей. Используя библиографические базы данных можно получить данные о различных журналах, публикациях и авторах. Это могут быть как количественные показатели, например, количество цитирований публикации или индекс Хирша автора, так и качественные, например, ключевые слова публикации или области интереса автора. Основываясь на данных библиографических баз можно построить такие аналитические инструменты, которые помогут экспертам принимать управленческие решения наиболее быстро и результативно. Подобные инструменты должны систематизировать данные со всех библиографических баз так, чтобы конечный пользователь для вынесения экспертной оценки мог полагаться на результаты работы интеллектуальных аналитических инструментов и карт науки.

Это дает возможность применения метода принятия управленческих решений на основе анализа больших данных для поиска наиболее эффективного пути развития научной деятельности. Такой подход обеспечивает обоснованность и объективность принятых решений и, как правило, качественно улучшает управленческую деятельность.

Описание команд модуля

В данной статье приведено описание модуля интеграции с различными библиографическими базами данных.

Для сбора данных система включает в себя следующие команды:

- Команда синхронизации интересующих научных сообществ по наименованию сообществ;
- Команда синхронизации публикаций выгруженных сообществ с возможностью указать интересующие годы публикационной активности сообщества;
- Команда синхронизации авторов выгруженных публикаций (в том числе и соавторов из других научных сообществ);
- Команда синхронизации источников выгруженных публикаций.

На рисунках 1 и 2 представлены схемы работы команд синхронизации научных сообществ и публикаций, остальные команды работают по аналогичным схемам.

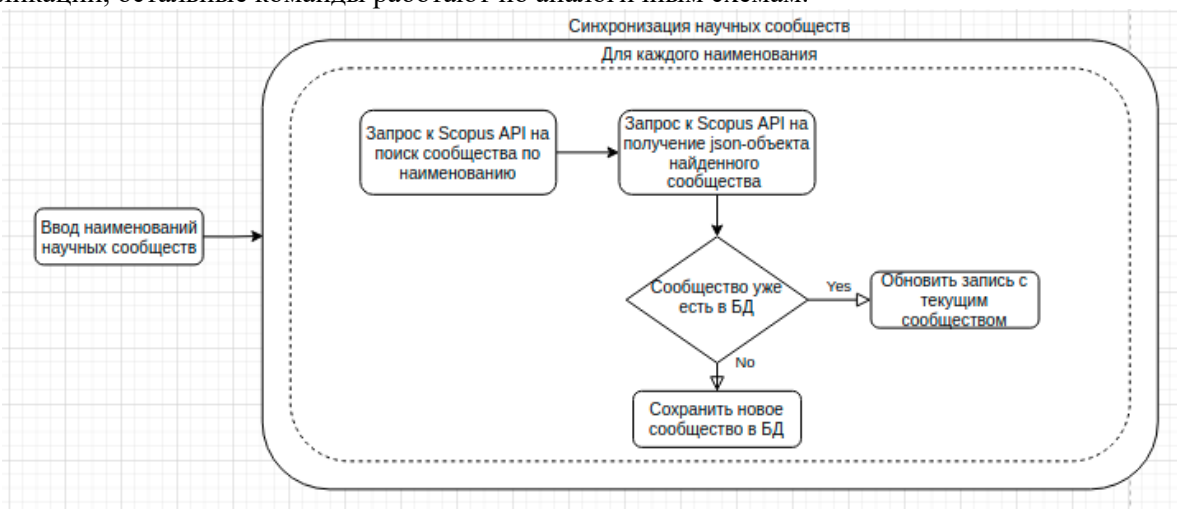


Рис. 1. Схема работы команды синхронизации научных сообществ.

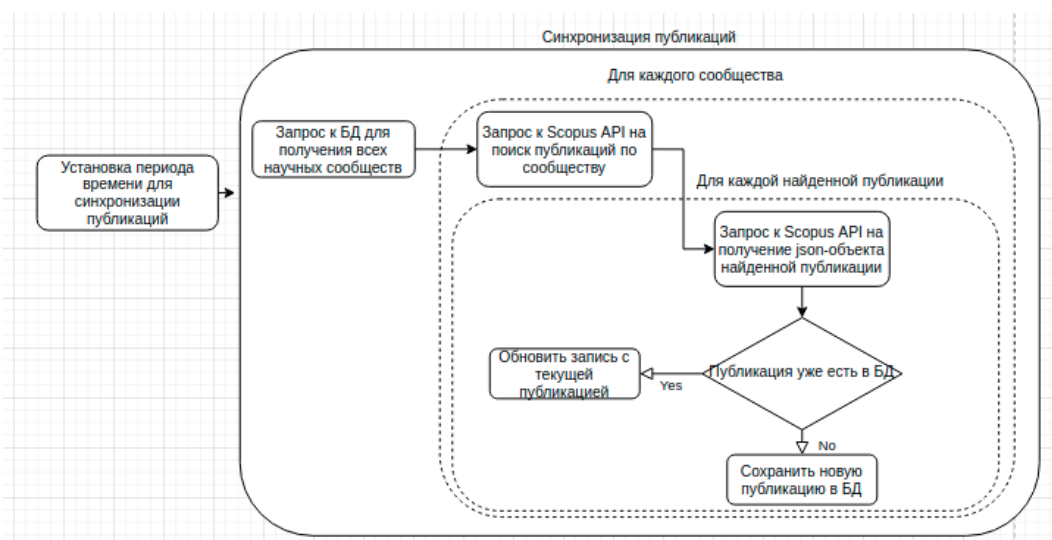


Рис. 2. Схема работы команды синхронизации публикаций.

В результате работы вышеперечисленных команд заполняются соответствующие таблицы в базе данных, данные для заполнения получают посредством отправки необходимых запросов к API Scopus. Для того чтобы избежать дублирования записей, каждая таблица хранит идентификаторы Scopus. Если в ответе на запрос приходит запись с идентификатором, уже имеющимся в базе данных системы, новая запись не создается, вместо этого происходит обновление полей уже имеющейся записи.

Описание выгружаемых данных

В ответе на запросы к API Scopus приходят json-объекты, описывающие научные сообщества, публикации, авторов и источники. При выгрузке в БД отдельно сохраняются необходимые на текущий момент поля json-объекта и сам json-объект.

По каждому научному сообществу сохраняется json-объект с информацией о сообществе из Scopus и наименование.

По каждой публикации сохраняется json-объект с информацией о публикации из Scopus, название публикации, год публикации, количество цитирований и тип публикации.

По каждому автору сохраняется json-объект с информацией об авторе из Scopus и ФИО. По выгруженным публикациям рассчитывается количество публикаций, количество цитирований автора и количество документов, в которых цитируется автор, на текущий день. Данную информацию можно получать непосредственно из json-объекта автора, однако, в таком случае она будет актуальна только в рамках одной библиографической базы данных.

По каждому источнику сохраняются json-объект с информацией об источнике из Scopus, название источника и тип источника.

Заключение

Описанный в данной статье подход к реализации сбора данных удобно использовать для реализации систем принятия управленческих решений в области развития научной деятельности. При подключении дополнительной библиографической базы данных необходимо каждой сущности добавить атрибут типа json для сохранения json-объектов из подключаемой библиографической базы данных.

Список использованных источников

1. API Interface Specification [Электронный ресурс] // Elsevier Developers - URL: https://dev.elsevier.com/api_docs.html
2. Трубников В.С., Туральчук К.А. Проектирование системы сбора, анализа и визуализации наукометрических данных [Электронный ресурс] // CyberLeninka - 2015. -URL: <https://cyberleninka.ru/article/n/proektirovanie-sistemy-sbora-analiza-i-vizualizatsii-naukometricheskikh-dannyh>
3. Бондарь М.А. Методы принятия управленческих решений [Электронный ресурс] // CyberLeninka - 2014. -URL: <https://cyberleninka.ru/article/n/metody-prinyatiya-upravlencheskih-resheniy-1>