

**А. В. Барановский**  
Науч. рук. **Н. Б. Осипенко**,  
канд. физ.-мат. наук, доцент

## МЕТАЯЗЫК ФОРМИРОВАНИЯ ОТЧЕТА НА ОСНОВЕ ДАННЫХ ИЗ ОБЛАЧНЫХ ХРАНИЛИЩ

Сегодня тема облачных вычислений получила активное развитие. Традиционно используемый для задач формирования отчета на основе данных из облачных хранилищ Apache Spark имеет ряд недостатков, например, код при его использовании малопрозрачен. Работа посвящена проектированию и разработке библиотеки метаязыка трансформации данных различного формата из облачных хранилищ, базирующегося на языке Scala. Библиотека включает в себя несколько модулей. Схема работы метаязыка с Apache Spark для трансформации данных различного формата из облачных хранилищ приведена на рисунке 1. Модули библиотеки принимают данные и возвращают их в измененном виде.

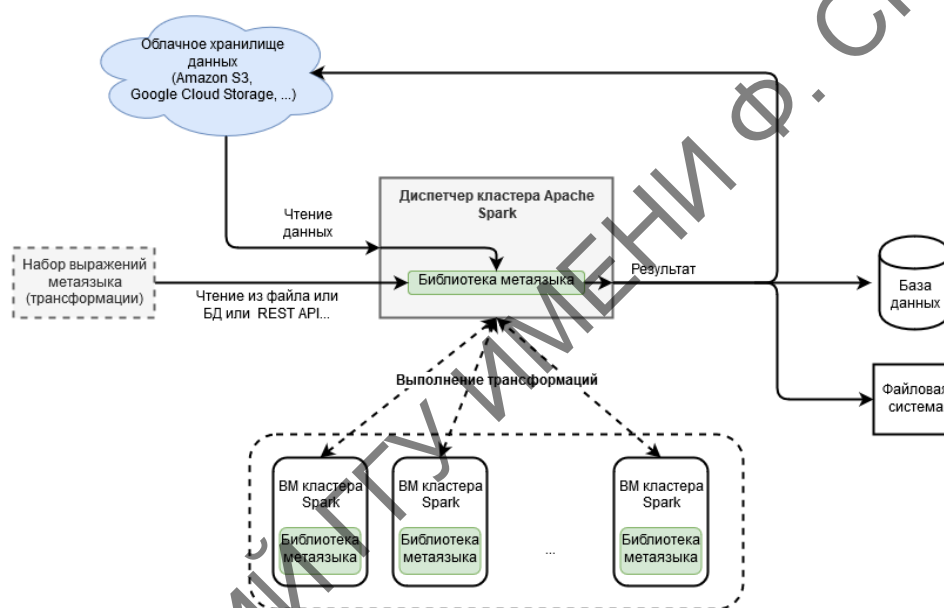


Рисунок 1 – Работа метаязыка совместно с Apache Spark

Модули чтения трансформируют данные во внутренний формат, преобразуя их в набор строк, а затем – в объекты внутренних классов библиотек. При этом модуль, отвечающий за преобразование правил трансформации, выполняет их компиляцию с использованием встроенных утилит языка Scala. На выходе получается функция, запуск которой означает выполнение трансформаций. При работе с данными главный модуль применяет правила трансформации над данными. Для обращения к БД специально создан модуль. БД могут использовать разные языки запросов: SQL, ELK и т. д. На входе модуль получает не конкретный запрос, а его структуру в некотором внутреннем формате, определенном библиотекой. В зависимости от используемой БД внутренний формат преобразуется в конкретный язык запросов, а внутренние объекты библиотеки преобразовываются в формат, соответствующий входному, т. е. если входные данные были в формате json, то и результат библиотека будет возвращать в формате json. Итак, модули разработанного метаязыка связаны друг с другом, формируя каркас библиотеки, определяя ее функционал. Применение разработанного метаязыка увеличивает читаемость запроса, уменьшает время на ознакомление с трансформациями и время на поиск ошибок.