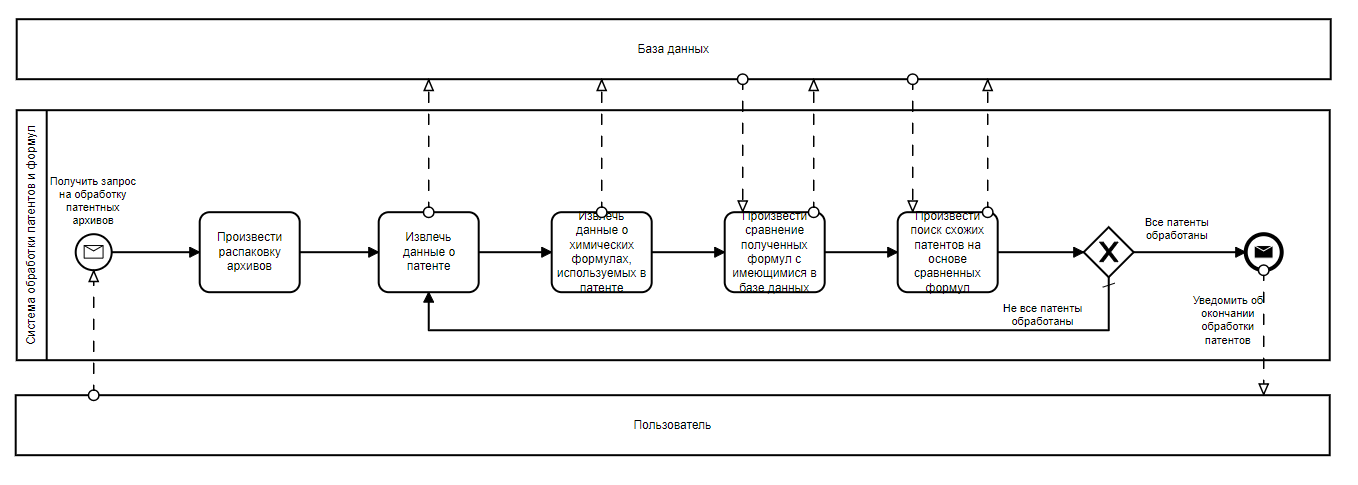
|  |  |  |
| --- | --- | --- |
| Проектирование АСОиУ  *Лабораторная работа №1*  Диаграммы BPMN | Студент | Вайнгольц Н.А. |
| Группа | САПР-1.1н |
| Преподаватель |  |

Бизнес-процесс «Модуля анализа химических формул» имеет следующий вид:



Модуль анализа химических формул производит поиск схожих химических патентов, основываясь на результатах сравнения содержащихся в них химических формул.

1. *Получить запрос на обработку патентных архивов.*

Сперва от пользователя поступает запрос на распаковку патентных архивов (архивы с тысячами патентов из разных научных отраслей).

1. *Произвести распаковку архивов.*

Производится распаковка патентных .tar и .zip архивов. По завершении распаковки производится выбор химических и «отсеивание» остальных патентов.

1. *Извлечь данные о патенте.*

Затем производится извлечение данных о патенте: регистрационный номер, дата подачи заявления, заявитель, авторы, классификация по МПК. Полученные данные записываются в базу данных.

1. *Извлечь данные о химических формулах, содержащихся в патенте.*

После производится извлечение информации о химических формулах, используемых в патенте: имена по номенклатуре SMILES, InChi, структурная формула, молекулярный отпечаток. Полученные данные также записываются в базу.

1. *Произвести сравнение полученных формул с имеющимися в базе данных.*

По молекулярным отпечаткам производится сравнение полученных формул с уже имеющимися в базе данных. Результаты сравнения отправляются также в БД.

1. *Произвести поиск схожих патентов на основе сравненных формул.*

По результатам сравнения химических формул из двух патентов (только что извлеченного и полученного из БД в цикле) делается вывод о схожести этих патентов. Полученный коэффициент схожести записывается в базу данных.

1. *Уведомить об окончании обработки патентов.*

Если все патенты из патентного архива извлечены и обработаны, то пользователю присылается сообщение об окончании работы.