

# **1. Business Requirements**

## **1.1. Background**

Каждый день в мире регистрируются тысячи патентов на изобретения, и их количество продолжает расти. В связи с этим появилась необходимость в автоматизированной обработке, кластеризации и поиска схожих патентов. Данный продукт будет решать представленные выше задачи в области химических патентов.

## **1.2. Business Opportunity**

В настоящее время представлено множество средств автоматизированной обработки патентов. В основном они не специализируются на какой-либо конкретной области научных исследований, имеют неэффективные инструменты поиска схожих патентов. Данный продукт специализируется в области химии и производит сравнение химических патентов, их кластеризацию и выявление патентных трендов с учётом упоминаемых в них химических формул.

## **1.3. Business Objectives and Success Criteria**

Основной бизнес-целью является повышение эффективности обработки патентов в патентном ведомстве. Сюда можно включить повышение скорости обработки химических патентов сотрудниками патентного ведомства, а также более быстрый поиск схожих патентов и автоматизированное выявление патентных трендов.

Успех этого проекта будет определяться по двум критериям:

- уменьшение временных затрат на обработку патентов сотрудниками патентного ведомства;
- уменьшение количества ошибок, совершаемых сотрудниками патентного ведомства, при обработке патентов.

#### **1.4. Customer of Market Needs**

Патентные ведомства сильно нуждаются в автоматизации работы с патентами. Ручная обработка патентов требует больших временных и физических затрат. Также в виду человеческого фактора возможны ошибки при обработке.

#### **1.5. Business Risk**

Непринятие продукта потенциальным пользователем ввиду непрозрачности его назначения и узкой специализированности может представляться серьезным риском. Также имеется большой риск того, что системные требования для этого продукта окажутся слишком высокими, либо невыполнимыми для большинства его потенциальных пользователей.

### **2. Vision of the Solution**

#### **2.1. Vision Statement**

Данный продукт должен представлять собой комплексное решение, внедряемое в вычислительный комплекс патентного ведомства. Он должен упростить поиск схожих химических патентов, автоматизировать кластеризацию химических патентов по выбранным признакам и выявление текущих трендов в данной области.

#### **2.2. Major Features**

- 1) Нацеленность продукта на химические патенты;
- 2) Поиск схожих патентов с учётом используемых химических формул;
- 3) Кластеризация химических патентов по выбранным критериям (время, тип исследуемых соединений, назначение исследуемых соединений);
- 4) Выявление патентных трендов (какие типы химических соединений становятся наиболее исследуемыми с течением времени).

### **2.3. Assumptions and Dependencies**

Предполагается, что к определённому этапу реализации проекта, в частности, к началу тестирования отдельных модулей, будет доступен массив патентных архивов USPTO.

Предполагается, что пользователь знаком с основами работы с персональным компьютером и сопутствующей ему периферией, а также умеет работать с интернет-браузером.

Так как данная система будет представлять собой веб-приложение, будет необходимо использовать интернет-браузер. Предполагается, что пользователи будут иметь бесперебойное подключение к локальной сети патентного ведомства.

## **3. Scope and Limitations**

### **3.1. Scope of Initial Release**

В первый релиз продукта будет включен следующий функционал:

- Распаковка патентных архивов;
- Извлечение информации о патенте (дата, название, авторы, заявитель, категория по Международной Патентной Классификации);
- Извлечение информации о химическом соединении, содержащемся в патенте (графическое представление и различные текстовые представления – SMILES, InChi и т.д.);
- Поиск схожих химических соединений;
- Поиск схожих патентов на основе схожести используемых в них химических соединений;
- Просмотр информации о патентах и схожих патентах, химических соединениях и схожих химических соединениях пользователем.

### **3.2. Scope of Subsequent Releases**

В последующем планируется дополнить функционал продукта:

- 1) Выполнение кластеризации патентов по различным критериям;
- 2) Выявление патентных трендов.

### 3.3. Limitations and Exclusions

Продукт не будет работать вне локальной сети патентного ведомства.

Пользователи также не смогут расширить доступ к данному продукту.

Приложение будет работать только с патентами USPTO.

## 4. Business Context

### 4.1. Stakeholder Profiles

<b>Stakeholder</b>	<b>Major Value</b>	<b>Attitudes</b>	<b>Major Interests</b>	<b>Constraints</b>
работник патентного ведомства	меньшее время работы	видит продукт как способ уменьшения затрат собственных усилий при работе с патентами	меньше времени затрачивается на поиск схожих патентов, анализ патентных трендов; простота использования	клиент должен работать на низкопроизводительном аппаратном обеспечении
разработчик	меньшее количество ошибок, требующих исправление в будущем	воспринимает продукт как возможность улучшить свои навыки работы с большими объемами данных	«умная» среда разработки: автоматическое исправление ошибок, автодополнение, подсказки о параметрах и пр.	тестирование необходимо проводить на высокопроизводительном аппаратном обеспечении
начальник патентного ведомства	получение большего объема обработанных патентов за меньшее время	может не пойти на компромисс, если потребуются дополнительная закупка аппаратного обеспечения	возможность обработки больших массивов патентных данных, отсутствие необходимости в долгом обучении работников	ограничен бюджет на доработки

### 4.2 Project Priorities

<b>Dimension</b>	<b>Driver (state objective)</b>	<b>Constraint (state limits)</b>	<b>Degree of Freedom (state allowable range)</b>
<i>Schedule</i>	<i>release 1.0 должен быть готов к 01.02.2020, release 1.1 - к 01.04.2020</i>	<i>Любая непредвиденная проблема приведет к задержке релиза</i>	<i>Во время тестирования продукта сервер будет запущен всё время.</i>
<i>Features</i>	<p>1) Распаковка патентных архивов;</p> <p>2) Извлечение информации о патенте</p> <p>3) Извлечение информации о химическом соединении, содержащемся в патенте</p> <p>4) Поиск схожих химических соединений;</p> <p>5) Поиск схожих патентов на основе схожести используемых в них химических соединений;</p> <p>6) Просмотр информации о патентах и схожих патентах, химических соединениях и схожих химических соединениях пользователем;</p> <p>7) Выполнение кластеризации патентов по различным критериям;</p> <p>8)Выявление патентных трендов.</p>	<i>Ошибки программного кода.</i>	<i>70-80% приоритетного функционала должны быть включены в release 1.0</i>
<i>Quality</i>	<i>Система повысит эффективность обработки химических патентов.</i>	<i>На низкопроизводительных вычислительных системах использование продукта может быть затруднительным.</i>	<i>90-95% приемочных испытаний должны проходить к release 1.0, 95-98% - к release 1.1</i>
<i>Staff</i>	<i>Разработчик базы данных, Front-end разработчик, Back-end разработчик.</i>	<i>Проектом занимается 3 разработчика. Их количество не может быть увеличено.</i>	<i>Знания персонала покрывают большую часть требований, необходимых для реализации этого проекта.</i>

Cost	500 000 рублей	Изменение курса рубля может повлиять на конечную стоимость.	Приемлемо превышение бюджета в 15%.
------	----------------	---	-------------------------------------

### 4.3 Operating Environment

*Данная система должна использоваться внутри единого подразделения патентного ведомства, на территории одного предприятия.*

*Данные, полученные при анализе химических патентов и соединений, используемых в них, должны храниться в распределенной базе данных ( NoSQL) на внутренних хранилищах вычислительного центра патентного ведомства.*

*Одновременно может выполняться только один процесс обработки патентных архивов. В то же время просматривать содержимое базы данных и результаты уже совершённой обработки (кластеризации патентов и патентные тренды) имеет возможность множество пользователей.*

*Так как система будет запускаться на собственной вычислительной системе, доступной только по локальной сети, нет необходимости в создании системы авторизации и аутентификации, SSL-сертификатов и пр. Единственная аутентификация будет происходить при подключении очередного клиента к базе данных.*