

ООО «Преферентум»

Поисково-аналитическая система  
«АйТи-Поиск»

---

Описание проекта внедрения  
«Интеллектуальная обработка информации»

Листов 12

Москва

2022

## Оглавление

<b>1. Вводная информация .....</b>	<b>3</b>
<b>2. Термины и определения .....</b>	<b>3</b>
<b>3. Описание решения.....</b>	<b>7</b>
<b>3.1. Реализованный функционал .....</b>	<b>7</b>
<b>3.2. Описание работы системы.....</b>	<b>8</b>
<b>4. Сценарии внедрения .....</b>	<b>9</b>
<b>4.1. Типовые сценарии внедрения.....</b>	<b>9</b>
<b>4.2. Пример реализации проекта на базе АйТи-Поиск .....</b>	<b>9</b>
<b>4.2.1. Описание решаемой проблемы.....</b>	<b>9</b>
<b>4.2.2. Описание реализации проекта.....</b>	<b>11</b>
<b>4.2.2.1. Обследование и анализ.....</b>	<b>11</b>
<b>4.2.2.2. Разработка.....</b>	<b>11</b>
<b>4.2.2.3. Опытная эксплуатация.....</b>	<b>12</b>
<b>4.2.2.4. Гарантийное сопровождение.....</b>	<b>12</b>
<b>4.3. Выводы .....</b>	<b>12</b>

## 1. Термины и определения

В таблице ниже представлены термины и определения, а также сущности, с которыми работает Поисково-аналитическая система «АйТи-Поиск».

Термин	Описание
<b>Алгоритмическое правило</b>	Жесткие алгоритмы, настройка выполняется разработчиком или администратором
<b>Аннотация</b>	<p>Информационное сообщение определенного формата и структуры, сформированное в виде примечания к тексту документа.</p> <p>Аннотация:</p> <ul style="list-style-type: none"> <li>• Указывает на место в документе, где сработало (должно было сработать) правило проверки текста либо правило извлечения именованной сущности (факта).</li> <li>• Несет в себе информацию о том, как сработало (должно было сработать) правило или какой факт был извлечен (должен был быть извлечен).</li> <li>• В документе будет применяться терминология в рамках нижеперечисленных видов классификации аннотаций.</li> </ul> <p>По природе создания аннотации можно разделить:</p> <ul style="list-style-type: none"> <li>• На автоматические, созданные ПО при отправке документа на проверку.</li> <li>• На созданные вручную экспертом при проверке результатов автоматической обработки документа.</li> </ul> <p>По отношению к типу документа аннотации можно разделить:</p> <ul style="list-style-type: none"> <li>• На аннотации извлечения именованных сущностей (фактов).</li> <li>• На аннотации, созданные по результатам срабатывания интеллектуальных правил проверки.</li> </ul> <p>По отношению к проверяемому тексту аннотации можно разделить:</p> <ul style="list-style-type: none"> <li>• На информационные – аннотации, которые не производят корректировку текста проверяемого документа.</li> <li>• На корректирующие – аннотации, которые могут автоматически изменить текст проверяемого документа.</li> </ul>
<b>Аннотирование</b>	Подтверждение или опровержение Экспертом результатов срабатывания интеллектуальных правил
<b>АРМ</b>	Автоматизированное рабочее место
<b>АС</b>	Автоматизированная система
<b>БД</b>	База данных
<b>ГОСТ</b>	Государственный стандарт
<b>ИС</b>	Информационная система
<b>Интеллектуальные правила</b>	Жесткие алгоритмы, настройки индуцируются системой на основе примеров, задаваемых экспертом
<b>ЛВС</b>	Локальная вычислительная сеть
<b>ЛНА</b>	Локальные нормативные акты
<b>Набор правил</b>	Несколько Правил, объединенных в процесс, исполняемый в модуле «АйТи-Поиск». Набором правил можно управлять (конфигурировать Набор). Например, включать или выключать работу отдельных Правил Набора или устанавливать значения параметров.

Термин	Описание
<b>НПД</b>	Нормативно-правовые документы
<b>Обратная связь</b>	Информация, внесенная экспертами по результатам проверки задания: решения по срабатыванию правил и предложения новых правил
<b>ОЗУ</b>	Оперативное запоминающее устройство
<b>ООО «Преферентум»</b>	Вендор решения «Поисково-аналитическая система «АйТи-Поиск»»
<b>ОС</b>	Операционная система
<b>ПО</b>	Программное обеспечение
<b>ПО «АйТи-Поиск»</b>	Поисково-аналитическая система «АйТи-Поиск»
<b>Правило</b>	Реализованный в виде программного кода интеллектуальный алгоритм, осуществляющий проверку текста на соответствие заданным критериям. Примерами критерия могут быть: наличие или отсутствие определенной информации (Фактов) в документе, соответствие или несоответствие проверяемого документа определенным требованиям: типовой форме или структуре, корректность употребления терминов и выражений, отсутствие ссылок на утратившие силу ГОСТы. Если в результате применения правила выявлены искомые факты или обнаружены ошибки в документе, правило формирует аннотации.
<b>Рейтингование правил</b>	Изменение рейтинга правила для целей проверки допустимости применения правила
<b>СУБД</b>	Система управления базами данных
<b>ТЗ</b>	Документ типа «Техническое задание»
<b>ТКП</b>	Документ типа «Технико-коммерческое предложение»
<b>Факт</b>	Смысловое понятие, выделяемое в тексте документа. Примерами фактов могут быть: наименование, телефон, e-мейл организации, ФИО руководителя и т.д. Синонимы – атрибут, реквизит, сущность.
<b>AD</b>	Active Directory, службы каталогов корпорации Microsoft для операционных систем семейства Windows Server
<b>API</b>	Программный интерфейс приложения (от англ. application programming interface)
<b>GUI, UI</b>	Графический пользовательский интерфейс (англ. graphical user interface, GUI)
<b>ИIS</b>	Internet Information Services, проприетарный набор серверов для нескольких служб Интернета от компании Microsoft
<b>IP-адрес</b>	Уникальный сетевой адрес узла в компьютерной сети, построенной на основе стека протоколов TCP/IP
<b>RabbitMQ</b>	Платформа, реализующая систему обмена сообщениями между компонентами программной системы на основе стандарта AMQP.
<b>SMB</b>	Сетевой протокол прикладного уровня для удалённого доступа к файлам и другим сетевым ресурсам (от англ. Server Message Block)
<b>SofA</b>	Source of Analysis, источник анализа. Объект данных сложной структуры, содержащий проверяемые набором правил документы в универсальном структурном представлении и набор аннотаций и реквизитов, выделяемых в процессе обработки документов набором правил.
<b>URL</b>	Единообразный определитель местонахождения ресурса (от англ.

<b>Термин</b>	<b>Описание</b>
	Uniform Resource Locator)
<b>Web</b>	Программное обеспечение или оборудование, обеспечивающее взаимодействие информационных систем посредством интернет-протоколов
<b>XML</b>	В данном документе термин используется для обозначения формата файлов обмена данными между информационными системами.

## 2. Вводная информация

Программное обеспечение «Поисково-аналитическая система АйТи-Поиск» (далее - АйТи-Поиск или ПО АйТи-Поиск) принадлежит компании ООО «Преферентум» (регистрационный №2009611815 в федеральной службе по интеллектуальной собственности (РОСПАТЕНТ), программа для ЭВМ).

ПО АйТи-Поиск реализует функции семантического анализа, извлечения сущностей из документов и проверки документов на соответствие реализованным в системе правилам.

ПО разрабатывается и поддерживается с учетом возможности работы как на платформе MS Windows, так и с использованием только импортозамещающего ПО.

Данный документ содержит описание процессов, обеспечивающих поддержание жизненного цикла программного обеспечения ПО АйТи-Поиск, в том числе устранение неисправностей, выявленных в ходе эксплуатации ПО, совершенствование ПО, а также информацию о персонале, необходимом для обеспечения такой поддержки.

ПО разработано и развивается сотрудниками ООО «Преферентум».

Исключительные права на программное обеспечение АйТи-Поиск принадлежат ООО «Преферентум».

### 3. Описание решения

#### 3.1. Реализованный функционал

Система обеспечивает следующие функциональные возможности:

- Возможность выделения информационных объектов. Например, организаций, персон, должностных лиц, дат, денежных сумм, валют, условий договора, условия поставки/отгрузки/оплат, названия ТМЦ и Услуг, единиц измерения, нормативной документации (ГОСТ, ТУ), географических объектов, почтовых адресов, ссылок на нормативные правовые акты и пр.;
- Возможность автоматической классификации информации (документов) с использованием методов машинного обучения;
- Возможность автоматического формирования в тексте проверяемого документа гипертекстовых ссылок на обнаруженные упоминания нормативных правовых актов и их составных частей в ЛНА (Локальных нормативных актов);
- Получение заданий на проверку документов пользователем или из внешней системы;
- Возможность администрирования системы, в том числе возможность добавления новых правил обработки документов;
- Возможность настройки получения и обработки документов из подключаемых источников (например, база ЛНА) и передачу документов в подсистему интеллектуальной обработки информации;
- Возможность сохранения в БД:
  - данных о выделенных фактах и сработавших правилах, в том числе рейтинги срабатывания правил;
  - фрагментов текстов, аннотации и комментарии Экспертов (в объеме, необходимом для настройки имеющихся или создания новых правил).

### 3.2. Описание работы системы

ПО «АйТи-Поиск» предоставляет возможность с помощью внутренних алгоритмов выделять из неструктурированного текста так называемые «именованные сущности» (информационные объекты), которые позже используются правилами для обработки документа.

Правила являются основным средством системы для обработки документов. Правила кодируются разработчиками на языке С#, загружаются в виде .DLL-файлов и при обработке могут использовать выделенные ранее информационные объекты.

После завершения работы правила, для документа создаются аннотации по отработанным правилам, в БД записывается список отработавших правил, выделенные сущности и фрагменты текста с аннотациями.

В системе реализована возможность автоматического формирования в тексте проверяемого документа гипертекстовых ссылок на обнаруженные упоминания нормативных правовых актов и их составных частей в ЛНА. Данную функциональность предоставляет специализированный компонент системы – «Документорий», в который заранее загружаются все необходимые документы.



## 4. Сценарии внедрения

### 4.1. Типовые сценарии внедрения

Архитектура и функционал ПО АйТи-Поиск предоставляют возможности нескольких сценариев внедрения решения или его компонент.

Основными сценариями проектов реализации информационных систем построенных с использованием ПО АйТи-Поиск можно назвать следующие сценарии:

- Использование базовой версии ПО АйТи-Поиск в качестве подсистемы обработки текстовой информации.
- Использования компонент ПО АйТи-Поиск для реализации ИС Заказчика использующие данные текстовой аналитики, как отдельные модули ИС.
- Встраивание компонент ПО АйТи-Поиск как WEB-сервис в ИС на стороне Заказчика.

Базовая версия ПО АйТи-Поиск может быть настроена по конкретные требования Заказчика. Некоторые требования могут привести к необходимости доработки базового функционала ПО АйТи-Поиск. Подобные работы выполняются на основе договорных отношений с ООО «Преферентум». Ниже приведено описание реализации проекта внедрения АИС «Интеллектуальная обработка информации», построенная на платформе ПО АйТи-Поиск.

### 4.2. Пример реализации проекта на базе АйТи-Поиск

#### 4.2.1. Описание решаемой проблемы

Целью описываемого внедрения АйТи-Поиск является повышение качества подготавливаемых документов, что выражается в следующем:

- Повышении оперативности подготовки документов.
- Снижении трудозатрат сотрудников.
- Снижении количества ошибок, допущенных в процессе разработки и обработки документов.

Основными задачами внедрения являлось обеспечение возможностей:

- Извлечение информационных сущностей (фактов) из технико-коммерческих предложений (ТКП), поступающих для анализа.
- Проверка технических заданий и проектов договоров, на предмет корректности согласно правилам (в т.ч. например, проверка корректности ссылок НПА).
- Классификация и проверка документов из пакетов документов, прилагаемых к ТКП, на предмет корректности согласно правилам и требованиям документации Заказчика;
- Обеспечение возможностей автоматической и ручной обработок обрабатываемой документации.
- Обеспечение возможности проверки экспертами результатов автоматической обработки документов, и фиксация результатов проверки в системе.
- Передача во внешнюю систему обработанных ПО документов.
- Настройка процессов и правил обработки документов.

Пользователями внедренной АИС являются сотрудники подразделений, связанных с поставками и закупками товаров, работ и услуг.

ИС была реализована с учетом возможностей масштабируемости:

- По функциональному объему.
- По организационному объему.

## 4.2.2. Описание реализации проекта

### 4.2.2.1. Обследование и анализ

По результатам проведенного обследования был сформирован пакет документов отражающие конфигурацию и настройки ПО АйТи-Поиск для решения поставленных задач, в частности:

- Проанализированы документы, подлежащие анализу.
- Сформированы количественные критерии качества обработки документов.
- Сформированы алгоритмы правил обработки документов.
- Сформированы требования классификации обрабатываемых документов.
- Уточнены параметры обработки документов, учитывающие работу как алгоритмических правил, так и использование машинного обучения (МО).
- Для реализации МО сформированы требования к тестовой выборке и требованиям к разметке тестовой выборки.

### 4.2.2.2. Разработка

На данном этапе реализации были обеспечены:

- Адаптация ПО АйТи-Поиск для решения задач проекта.
- Интеграционное взаимодействие со смежными ИС Заказчика.
- Размечен корпус тестовой выборки для проведения МО.
- По результатам реализации этапа и согласования был сформирован пакет проектной и исполнительной документации (руководства, презентации, обучающие материалы, документы обеспечивающие проведение испытаний перед переводом системы в опытную и промышленную эксплуатацию, регламенты проведения опытной и промышленной эксплуатаций).
- Проведен комплекс приемо-сдаточных испытаний перед переводом ИС в опытную эксплуатацию.

#### **4.2.2.3. Опытная эксплуатация**

На данном этапе были подтверждены заложенные решения, а также внесены уточняющие корректировки в настройки ПО на основании полученных от пользователей замечаний. Обеспечено сопровождение ИС и решение вопросов эксплуатации ПО. После проведения приемо-сдаточных испытаний, реализованная на базе ПО АйТи-Поиск, ИС передана в промышленную эксплуатацию. ИС была передана на сопровождение в ЦК со стороны Заказчика.

#### **4.2.2.4. Гарантийное сопровождение**

На этапе гарантийного сопровождения выпускались необходимые обновления ПО в соответствии с планами развития ПО АйТи-Поиск. Заказчику оказывалась необходимая консультационная поддержка, решались вопросы в рамках гарантийных обязательств.

### **4.3. Выводы**

По результатам реализации проекта все поставленные цели были достигнуты. Уточнены подходы к реализации типового внедрения и сценариев использования ПО АйТи-Поиск.

Подтверждены подходы к возможностям кастомизации (реализация функционала под специализированные задачи Заказчика) ПО АйТи-Поиск.

Решены задачи интеграционного взаимодействия с разными ИС Заказчика.