

ё

Программа для ЭВМ

«Атом.Поиск: Базовый релиз»

**Описание функциональных характеристик**

На 8 листах

Дата: 04.03.2024

Версия: 1.0

ЛИСТ ИЗМЕНЕНИЙ

| Дата | Версия | Краткое описание изменений | ФИО |
| --- | --- | --- | --- |
| 04.03.2024 | 1.0 | Документ создан | Бочарова О.В |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

СОДЕРЖАНИЕ

[ГЛОССАРИЙ 4](#_Toc174359486)

[1 Общие сведения 5](#_Toc174359487)

[1.1 Наименование программы 5](#_Toc174359488)

[1.2 Область применения программы 5](#_Toc174359489)

[2 Программа для ЭВМ «Атом.Поиск: Базовый релиз», описание функциональных характеристик 6](#_Toc174359490)

[2.1 Функциональные возможности 6](#_Toc174359491)

[2.2 Описание функциональных возможностей 6](#_Toc174359492)

ГЛОССАРИЙ

| Термины/сокращения | Определение |
| --- | --- |
| Аутентификация | Процедура проверки подлинности пользователя путем сравнения введенных им логина / пароля с соответствующими данными, сохраненными в системе |
| Программа | Программа для ЭВМ «Атом.Поиск: Базовый релиз» |
| Сущности | Объекты определённого типа, чаще всего составные. К ним относятся, например, наименование адреса, организации, контактные данные, даты, ФИО. |

1. Общие сведения
	1. Наименование

Полное наименование: Программа для ЭВМ «Атом.Поиск: Базовый релиз».

* 1. Область применения

 предназначена для поиска информации в различных корпоративных информационных системах по документам различных форматов в рамках единого поискового окна.

Единое поисковое окно позволяет пользователям сократить время на поиск необходимой информации и повышает эффективность сотрудников.

1. Программа для ЭВМ «Атом.Поиск: Базовый релиз», описание функциональных характеристик
	1. Функциональные возможности

Программа обеспечивает выполнение следующих функциональных возможностей:

* четкий поиск – поиск точного совпадения;
* нечеткий поиск – поиск с учетом морфологии (окончания, падежи, опечатки, неправильная раскладка клавиатуры);
* семантический поиск – поиск по смыслу;
* ранжирование поисковой выдачи;
* сбор статистики поисковых запросов;
* поиск по содержимому документов;
* поиск по разным форматам документов;
* извлечение сущностей из документов;
* поддержка пользовательских словарей синонимов;
* поиск похожих по смыслу документов;
* единое окно поиска во всех источниках (аутентификация, фильтры, настройки);
* администрирование через web-интерфейс.
	1. Описание функциональных возможностей
		1. Четкий поиск

Программа осуществляет поиск точному совпадению поискового запроса (дословный поиск).

При поиске точного совпадения результаты должны соответствовать поисковому запросу ключевого слова или фразы.

* + 1. Нечеткий поиск

Программа осуществляет поиск, с учетом различных вариантов окончаний и падежей поискового запроса, а также опечатки, неправильной раскладки клавиатуры.

Нечеткие запросы носят размытый характер и имеют вариативность в выдаче.

* + 1. Семантический поиск

Семантический поиск – это способ и технология поиска информации, основанная на использовании контекстного (смыслового) значения запрашиваемых фраз вместо словарных значений отдельных слов или выражений при поисковом запросе.

Цель семантического поиска – определять особенности пользователя и предоставлять ему наиболее релевантные результаты.

* + 1. Ранжирование поисковой выдачи

Ранжирование поисковой выдачи – процесс упорядочивания найденных результатов по их релевантности, другими словами, по соответствию поискового запроса содержанию документов.

В программе используется функция ранжирования bm25 для упорядочивания документов по их релевантности данному поисковому запросу

Именно от ранжирования зависит качество поиска – то, насколько поисковая система умеет показать пользователю нужный и ожидаемый результат.

* + 1. Сбор статистики поисковых запросов

Для просмотра метрик работы программы используются дашборды.

* + 1. Поиск по содержимому документов

Программа позволяет осуществлять поиск по содержимому файлов.

* + 1. Поиск по разным форматам документов

Программа загружает, обрабатывает и осуществляет поиск по следующим форматам файлов: docx, txt, xlsx, msg, pdf, jpg, png, odp, odt, ods, pptx.

* + 1. Извлечение сущностей из документов

Программа выполняет выделение именованных сущностей из содержимого детальной информации. В каждой выделенной сущности отображается тип именованной сущности.

Извлечение именованных сущностей – это извлечение конкретной информации из текстовых данных.

* + 1. Поддержка пользовательских словарей синонимов

Администратор создает текстовые файлы (набор синонимов и стоп-слов) в определенном формате и копирует в созданные директории в контейнере.

* + 1. Поиск похожих по смыслу документов

Программа осуществляет поиск похожих по смыслу документов.

* + 1. Единое окно поиска во всех источниках (аутентификация, фильтры, настройки)

В рамках единого окна пользователь может пройти процедуру аутентификации, произвести поиск во всех доступных источниках и при необходимости выполнить фильтрацию и настройку отображения результатов поиска.

* + 1. Администрирование через web-интерфейс

Программа имеет системные настройки. Через web-интерфейс есть возможность изменять системные настройки.