**ИСХОДНЫЕ ДАННЫЕ ДЛЯ ФОРМИРОВАНИЯ ТЕХНИЧЕСКОГО ЗАДАНИЯ НА ВЫПОЛНЕНИЕ РАБОТ**

**Мероприятия по развитию Системы автоматизированного делопроизводства и документооборота Государственной Думы**

**Оглавление**

[**Общие требования: 1**](#_Toc137253759)

[**Определения, обозначения и сокращения: 2**](#_Toc137253760)

[**Требования к Системе: 2**](#_Toc137253761)

[**1. Учет при подготовке информационно-аналитических и статистических материалов всех входящих документов (обращений) от граждан, зарегистрированных в САДД 2**](#_Toc137253762)

[**2. Добавление новых атрибутов, отображение их в таблицах базы данных 3**](#_Toc137253763)

[**3. Внедрение обязательного протокола работы с карточкой обращения и ответа на него 4**](#_Toc137253764)

[**4. Внедрение модуля анализа текстов обращений на основе применения современных информационных технологий, в том числе технологий искусственного интеллекта 5**](#_Toc137253765)

[**4.1 Подсистема интеллектуального анализа и структурирования обращений: 6**](#_Toc137253766)

[**4.2 Подсистема «Классификация обращений»: 7**](#_Toc137253767)

[**4.4 Подсистема «Подготовка проектов ответов»: 8**](#_Toc137253768)

[**4.5 Подсистема «Статистика и отчеты»: 8**](#_Toc137253769)

[**5. Внедрение модуля обработки массовых обращений 9**](#_Toc137253770)

**Общие требования:**

Развитие Системы автоматизированного делопроизводства и документооборота Государственной Думы необходимо для обеспечения единого учета обращений граждан, поступающих в Государственную Думу, учета расширенного перечня анализируемых показателей и дополнительных маркеров обращений в целях создания программно-аппаратного комплекса анализа обращений граждан (далее – Сервис анализа обращений).

Сервис анализа обращений создается в целях обеспечения сбора, агрегирования, анализа и представления в структурированном виде, в том числе посредством визуализаций *(аналитических выкладок и информационных панелей)*, детализированной информации из поступающих в Государственную Думу обращений.

# **Определения, обозначения и сокращения:**

**Термины и определения**

|  |  |
| --- | --- |
| **Термин** | **Определение** |
| ЕСИА | Единая система идентификации и аутентификации |
| Заказчик | Аппарат Государственной Думы Федерального Собрания Российской Федерации |
| САДД | Система автоматизированного делопроизводства и документооборота Государственной Думы |
| Сервис анализа обращений | Программно-аппаратный комплекс анализа обращений граждан |
| СОЗД ГАС «Законотворчество» | Система обеспечения законодательной деятельности Государственной автоматизированной системы «Законотворчество» |
| СООГ | Система обработки и анализа обращений граждан ГАС «Законотворчество» |

# **Требования к Системе:**

1. **Учет при подготовке информационно-аналитических и статистических материалов всех входящих документов (обращений) от граждан, зарегистрированных в САДД**

**Анализ действующей технологии:**

В настоящее время информационно-аналитические и статистические материалы по обращениям граждан подготавливаются на основе обращений, зарегистрированных в картотеке Управления по работе с обращениями граждан.

Картотеки, не учитываемые при формировании информационно-аналитических и статистических материалов по обращениям граждан:

* картотеки депутатов Государственной Думы;
* картотеки фракций в Государственной Думе (аппаратов фракций);
* картотеки комитетов Государственной Думы (аппаратов комитетов);
* картотека Управления документационного обеспечения;
* картотеки иных структурных подразделений Аппарата Государственной Думы.

**Предложения по развитию:**

Исполнителю необходимо осуществить функциональное обследование САДД и разработку документации в целях обеспечения единого и полного учета всех входящих документов от граждан и ответов на них при подготовке информационно-аналитических и статистических материалов по обращениям граждан.

Исполнителю необходимо обеспечить формирование набора данных и его импорт в отдельный программно-аппаратный комплекс анализа обращений граждан.

Набор данных должен быть нормализован до третьей нормальной формы.

До разработки отдельного программно-аппаратного комплекса анализа обращений граждан необходимо обеспечить учет обращений граждан, зарегистрированных во всех картотеках САДД, при подготовке Управлением по работе с обращениями граждан отчетов с использованием следующих выходных статистических (аналитических) форм:

* Обзор обращений граждан, поступивших (зарегистрированных) в ГД ФС РФ;
* Справка по ОГ. Рассмотрение адресатом;
* Анализ работы депутатов;
* Обращения, направленные по компетенции по отдельным группам (узлам) корреспондентов-адресатов, за определенный период;
* Статистические данные по ответам на обращения гражданам.

*На этапе функционального обследования Исполнителю необходимо обратить внимание, что в настоящее время статистическая отчетность формируется с учетом экземплярности обращения и данный атрибут может использовать в других картотеках не в числовом, а в ином значении.*

* 1. **Добавление новых атрибутов, отображение их в таблицах базы данных**

В соответствии с планируемыми мероприятиями по развитию СООГ и сбора дополнительных атрибутов, необходимо предусмотреть учет этих атрибутов в базе данных САДД. В том числе, будут включены атрибуты:

* ID пользователя (уникальный идентификационный номер пользователя на портале Госуслуг);
* Объект вопроса;
* № законопроекта;
* Оценка отношения заявителя к законопроекту *(за / против / внесение предложений / особое мнение)*.
* Признак повторности и пр.
* Типы обращений (выделение благодарности, без сути)
* Типы обращений с детализацией (***Предложение****: в федеральное законодательство, в подзаконные акты, региональное законодательство, предложение, не касающееся законотворческой деятельности.* ***Жалобы****: на депутатов Государственной Думы (действие, бездействие), на деятельность Государственной Думы / Аппарата Государственной Думы, на деятельность ФГО, региональных органов власти, органов МСУ, юридических лиц, физических лиц (действие / бездействие).*
* Предмет ведения (федеральный / региональный/ местный)

Перечень атрибутов может быть уточнен на этапе технического проектирования.

* 1. **Внедрение обязательного протокола работы с карточкой обращения и ответа на него**

Требуется разработка по согласованию с Заказчиком последовательного сценария заполнения карточки с обязательным отображением конфликта сценариев *(возврат к предыдущему шагу и / или возникновение всплывающего окна)*.

Выстраиваемый порядок работы с обращениями должен обеспечить создание прозрачного контролируемого движения документа с отображением всей информации о движении документа в Системе.

В рамках внедрения обязательного протокола работы необходимо реализовать функциональные возможности, позволяющие:

* обеспечить частичную автоматизацию процессов, связанных с обработкой обращений *(например: автоматическая смена статуса вопроса при регистрации связанного документа в адрес заявителя; автоматизация функционала направления уведомлений гражданам о переадресации их обращений по компетенции с возможностью редактировать текст уведомления; упрощение процедуры выгрузки на ССТУ путем учета связки, а не совпадения статусов вопроса в карточке обращения и ответа на него)*;
* осуществлять мониторинг действий пользователей *(в целях повышения уровня исполнительской дисциплины)*, включая сбор информации для осуществления внутреннего контроля порядка рассмотрения обращений, в том числе в части организации их учета и регистрации, соблюдения сроков рассмотрения.

Исполнителю необходимо:

* совместно с Заказчиком провести UX-исследование (анализ пользовательского опыта) в целях протоколирования и документирования единого сценария работы с обращением с момента регистрации (внесение данных о заявителе, общая информация об обращении, включая внесение данных о тематике вопросов, объекте вопроса, заполнение краткого содержания, сканирование обращения при поступлении на бумажном носителе) до завершения работы с обращением и списанием в архив;
* проанализировать совместно с Заказчиком порядок регистрации обращений граждан, направленных с сопроводительным письмом от организаций, разработать методологию, позволяющую исключить раздельную регистрацию сопроводительных писем *(в настоящее время осуществляется Управлением документационного обеспечения)* и приложенных к ним обращений граждан *(в настоящее время осуществляется Управлением по работе с обращениями граждан)*;
* разработать маршрут движения документа с целью визуализации данных о текущем этапе работы с документом (в том числе разработка и визуализация древовидной схемы движения документа со сроками и исполнителями на отдельной вкладке «История» - с возможностью углубленной детализации);
* согласовать и внедрить последовательный сценарий заполнения карточки обращения и ответа на него;
* выполнить работы по тестированию и отладке технологии.
  1. **Внедрение модуля анализа текстов обращений на основе применения современных информационных технологий, в том числе технологий искусственного интеллекта**

Модуль анализа текстов обращений на основе применения технологий автоматизированного анализа документов, использования методов интеллектуальной обработки неформализованной информации на основе машинного обучения должен обеспечить решение следующих задач:

* повышение скорости обработки поступающих обращений, автоматизация процесса предобработки обращений, поступивших в электронном виде *(определение тематики поднимаемых в обращении вопросов, проверка выбранной заявителем классификации обращения как предложения/ заявления/ жалобы, проверка адресата обращения, исходя из определенной тематики, выявление отношения заявителя к поднимаемым вопросам (определение тональности текста для оперативного реагирования на особо критичные обращения), определение, является обращение индивидуальным, коллективным либо направлено в интересах 3-х лиц, реагирование на упоминание безрезультатности его обращений в иные органы власти и организации).*
* сокращение трудозатрат на предобработку обращений *(автоматическое заполнение полей: Рубрика, Краткое содержание, Проект резолюции с возможностью редактирования данных Исполнителем)*;
* обеспечение полноты данных по обращениям для подготовки информационно-аналитических и статистических материалов;
* повышение качества формирования текстовых и графических отчетов и для проведения оперативного комплексного мониторинга и анализа актуальных предложений и проблем, существующих на местах.

Результаты работ должны обеспечить создание и внедрение комплекса программно-технологических средств интеллектуальной информационно-технологической среды управления обращениями граждан.

**Исполнителем должны быть выполнены следующие работы:**

* Доработка функциональных возможностей Системы в связи с изменением в СООГ алгоритма подачи гражданами обращений в электронном виде, разработкой схем маршрутов подачи обращения и сбора дополнительных атрибутов, учет которых необходимо осуществить в Системе *(требуется на этапе подготовки к внедрению технологий искусственного интеллекта с целью формирования для обучающих выборок качественных данных)*;
* Формирование обучающих выборок по видам обращений *(при участии Заказчика)*;
* Разработка проектных решений по автоматизированной технологии дообучения;
* Разработка программного обеспечения, апробация подсистемы интеллектуального анализа и структурирования обращений на реальном массиве данных на тестовом контуре Заказчика;
* Выполнение пусконаладочных работ, включая: развертывание, настройку и апробацию технологии в тестовом режиме, ее интеграцию в действующие информационные системы.
* Проведение предварительных испытаний, опытной эксплуатации технологии, сбор и устранение замечаний, включая доработку технологии по замечаниям.
* Проведение приемо-сдаточных испытаний, передача для ввода в эксплуатацию

**Требования к архитектуре:**

**В состав технологии должны входить следующие подсистемы** *(требуется уточнение на этапе проектирования)***:**

* 1. **Подсистема интеллектуального анализа и структурирования обращений:**

*Позволяет осуществлять прием и централизованную интеллектуальную обработку, структурирование и диспетчеризацию потоков входящих в электронном виде обращений, в т.ч.:*

* *классификацию документов на основе методов машинного обучения и искусственного интеллекта;*
* *проверку документов с помощью настраиваемых лексических шаблонов и правил;*
* *формирование, ведение и актуализацию тематического рубрикатора обращений с использованием технологий машинного обучения;*
* *автоматическую диспетчеризацию индексированных обращений в зависимости от вида обращения (мониторинг и****диспетчеризация****позволяет оперативно, точно, адекватно оценивать и анализировать текущую ситуацию, принимать обоснованные решения);*
* *ведение и актуализацию справочника заявителей с расширенными атрибутивным составом c функцией синхронизации и дедупликации (устранение избыточных копий информации) справочников информационных систем;*
* *формирование, ведение и обогащение базы знаний с применением технологий машинного обучения, актуальными сведениями из поступающих по электронным каналам обращений.*
  1. **Подсистема «Классификация обращений»:**

*Предназначена для решения задачи распределения неструктурированной текстовой информации по заранее заданным тематическим рубрикам (классам).*

*Необходимо поддерживать возможность интеграции с другими подсистемами через REST API или путём непосредственного вызова функций программных библиотек.*

*Подсистема должна поддерживать возможность экспорта результатов автоматической классификации в виде CSV-файла с целью последующего анализа качества обучающей выборки. Для устранения из процесса обучения классификатора некорректных данных, вызванных ошибками оператора и повышения качества классификации, подсистема «Классификация обращений» должна поддерживать возможность автоматического обнаружения таких данных и устранения их из обучающей выборки.*

*Подсистема должна поддерживать возможность тонкой настройки алгоритмов классификации для установления порога неуверенной классификации (близких значений определяемых вероятностей первой и последующей рубрики) и передачи неуверенно классифицируемых текстов на ручную обработку. Подсистема должна поддерживать возможность раздельной настройки порогов неуверенной классификации по каждой отдельной рубрике.*

*Подсистема должна поддерживать возможность расширения используемого лексического словаря для учета специфики деятельности Государственной Думы: терминологии, аббревиатур, сленговых выражений, названий структурных подразделений и т.п.*

*Должна обеспечиваться возможность ручной корректировки весовых коэффициентов в обученном индексе, в том числе с защитой от их изменения при последующем дополнительном обучении.*

* 1. **Подсистема «Анализ подготовленных ответов»:**

*Предназначена для получения и анализа текстов исходящих ответных документов, подготовленных исполнителем по обращению.*

*Подсистема должна обеспечивать автоматизированное выявление наиболее типичных тематических кластеров в массивах подготовленных ответов.*

*Результаты анализа подготовленных ответов должны использоваться для пополнения базы ответов на типовые вопросы заявителей.*

* 1. **Подсистема «Подготовка проектов ответов»:**

*Предназначена для поддержки исполнителей по обращению при подготовке проектов ответов на обращения, в том числе:*

* *семантический анализ обращений заявителя, выявление ранее поступивших «схожих» обращений;*
* *выборка ответов на обращения по семантически «близким» обращениям;*
* *проверка поступления аналогичных обращений с указанием иного вида доставки (дублирование);*
* *выявление противоречивых ответов на семантически идентичные обращения;*
* *выделение ссылок на нормативные правовые акты, присутствующих в обращении и в подготовленных ранее официальных ответах на похожие обращения;*
* *выборка текстов нормативных правовых актов и их фрагментов для принятия обоснованных решений по обращению;*
* *применение сервисов автоматизированной проверки обращений на требования законодательства;*
* *генерация проекта ответа на обращение на основе автоматизированного типового шаблона ответа, соответствующего типовому обращению (запросу);*
* *ретроспективный анализ всех обращений заявителя за выбранный период.*
  1. **Подсистема «Статистика и отчеты»:**

*Предназначена для формирования аналитической и статистической отчетности по следующим направлениям:*

* *по типам и классам обращений;*
* *по статистике правильности подготовленных проектов ответов.*

*Состав и перечень параметров отчетов определяется на этапе проектирования.*

Применяемое программное обеспечение должно предусматривать использование в технологических ландшафтах, основанных как на MS Windows, так и на Linux системах.

Модуль анализа текстов обращений должен обеспечивать интеграцию с электронными каналами поступления обращений для получения и обработки данных в подсистеме интеллектуального анализа и структурирования обращения.

Модуль в рамках выполнения требуемых функций должен быть интегрирован с СООГ.

**Требования к информационному обеспечению:**

Должно обеспечиваться систематизированное хранение поступающих обращений и иных сообщений в виде неизменных файлов.

Должно обеспечиваться хранение получаемых из обращений, сообщений формализованных данных в базе (базах) данных.

Должно обеспечиваться систематизированное хранение результатов отнесения обращений и иных сообщений к заданным категориям, предпочтительно – в базе (базах) данных.

Должно обеспечиваться формирование и ведение тематического реестра категорий и групп категорий обращений и иных сообщений с использованием технологий машинного обучения, в том числе хранение списка используемых категорий и групп категорий с историей изменений в базе (базах) данных.

Должно обеспечиваться формирование и ведение базы знаний, используемой для отнесения обращений и иных сообщений к заданным категориям; допускается общая техническая реализация тематического реестра категорий и групп категорий и базы знаний; в ходе разработки Заказчиком обеспечивается предоставление поступивших за предшествующие периоды обращений и иных сообщений в виде файлов и полученных из них формализованных данных в объеме, достаточном для формирования первичной базы знаний либо выполняется формирование первичной базы знаний по поступившим в предшествующие периоды обращениям и иных сообщениям с помощью предоставленных Исполнителем инструментов.

1. **Внедрение модуля обработки массовых обращений**

**Модуль обработки массовых обращений должен быть реализован на базе САДД и должен предусматривать:**

* тиражирование ответа на однотипные обращения *(решение об однотипности обращений принимается исполнителем на основе анализа текста обращений, с учетом совпадения тематики обращений заявителей)* и обращения из веерной рассылки;
* тиражирование резолюции на исполнителя с одновременным направлением уведомления *(в случае перенаправления обращения на рассмотрение)* или ответа *(в случае направления исполнителю обращения к сведению).*

**Тиражирование предусматривает:**

1. выделение части однотипных обращений заявителей или части обращений из веерной рассылки одним заявителем;
2. обработка одного обращения из выделенного списка по определенному алгоритму с подготовкой ответа *(регистрация связанного документа, заполнение обязательных реквизитов в части: тематики, статуса вопроса, создание шаблона ответа заявителю с возможностью отправки с использованием функционала САДД)* либо с формированием поручения по обращению *(посредством ввода резолюции)*, созданием шаблона уведомления/ ответа заявителю;
3. применение алгоритма ко всем выделенным обращениям граждан;
4. автоматическое направление ответа на адрес электронной почты заявителя;
5. внутренняя проверка системы на корректность работы модуля.

Схема обработки обращения модулем может быть изменена на этапе технического проектирования.

При тиражировании ответа на однотипные обращения и обращения из веерной рассылки необходимо создание для каждого обращения уникального исходящего документа *(обеспечение связи: на 1 входящий документ от заявителя ≥ 1 исходящего документа в адрес заявителя).*

Действия в рамках работы Модуля обработки массовых обращений должны храниться в Системе и быть доступны по поиску как обращения, к которым применен модуль в целях сбора статистической отчетности по массовым обращениям.