

Описание процессов, обеспечивающих  
поддержание жизненного цикла  
программного обеспечения  
ИИ «Агент Энерго»

Описание процессов, обеспечивающих поддержание жизненного цикла  
программного обеспечения ИИ «Агент Энерго»

## 1. Оглавление

|                                                                        |    |
|------------------------------------------------------------------------|----|
| 2. Введение и общие положения .....                                    | 3  |
| 2.1. Назначение .....                                                  | 3  |
| 2.2. Область применения .....                                          | 3  |
| 2.3. Термины, определения и сокращения .....                           | 3  |
| 3. Порядок сопровождения программного обеспечения .....                | 4  |
| 3.1. Порядок непрерывного мониторинга компонентов ПО .....             | 4  |
| 3.2. Метрики контроля работоспособности .....                          | 5  |
| 3.3. Регламент профилактического контроля .....                        | 5  |
| 4. Порядок устранения неисправностей и регламент инцидентов .....      | 6  |
| 4.1. Прием и регистрация обращений .....                               | 6  |
| 4.2. Матрица критичности инцидентов и сроки реагирования (SLA) .....   | 6  |
| 4.3. Регламент действий персонала при возникновении сбоев .....        | 7  |
| 5. Порядок совершенствования и модернизации ПО .....                   | 8  |
| 5.1. Порядок сбора и анализа предложений по модернизации .....         | 8  |
| 5.2. Регламент разработки и тестирования обновлений .....              | 9  |
| 5.3. Порядок дообучения ИИ-моделей и обновления базы знаний .....      | 10 |
| 5.4. Требования к квалификации и составу персонала .....               | 10 |
| 5.5. Условия размещения персонала и территориальная принадлежность ... | 11 |
| 5.6. Инфраструктура обеспечения ЖЦ .....                               | 11 |

## 2. Введение и общие положения

### 2.1. Назначение

Настоящий документ регламентирует совокупность организационно-технических процессов, процедур и мероприятий, обеспечивающих поддержание жизненного цикла программного обеспечения ИИ «Агент Энерго» (далее — ПО, Система). Документ определяет порядок сопровождения, устранения неисправностей, модернизации и технической поддержки Системы в процессе ее коммерческой эксплуатации.

### 2.2. Область применения

Положения настоящего регламента являются обязательными для исполнения всеми штатными сотрудниками правообладателя (инженерами технической поддержки, ML-инженерами, системными администраторами), задействованными в процессах обслуживания и развития ПО.

Цель – обеспечить высокое качество обслуживания, минимизировать время простоя и быстрое реагирование на возможные сбои, а также повысить эффективность эксплуатации системы.

### 2.3. Термины, определения и сокращения

| Термин | Определение                                                                                   |
|--------|-----------------------------------------------------------------------------------------------|
| API    | Application Programming Interface – интерфейс программирования приложений                     |
| LLM    | Large Language Model – большая языковая модель                                                |
| RAG    | Retrieval-Augmented Generation – метод генерации ответов с использованием внешней базы знаний |

Описание процессов, обеспечивающих поддержание жизненного цикла  
программного обеспечения ИИ «Агент Энерго»

|                |                                                                                                                                                                                                                                                                      |
|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SLA            | Service Level Agreement – Соглашение об уровне оказания услуг – регламентный документ правообладателя, устанавливающий целевые показатели доступности ПО, временные рамки регистрации, локализации и полного устранения возникших сбоев силами технической поддержки |
| ЖЦ             | Жизненный цикл                                                                                                                                                                                                                                                       |
| ИИ             | Искусственный интеллект                                                                                                                                                                                                                                              |
| Инцидент, сбой | Любое событие, повлекшее частичную или полную недоступность функций ПО для конечных пользователей                                                                                                                                                                    |
| ПО / Система   | Программное обеспечение ИИ «Агент Энерго»                                                                                                                                                                                                                            |

### 3. Порядок сопровождения программного обеспечения

#### 3.1. Порядок непрерывного мониторинга компонентов ПО

Для обеспечения стабильного функционирования Системы применяется распределенный автоматизированный стек мониторинга на базе компонентов *Prometheus* (сбор метрик), *Grafana* (визуализация), *Grafana Loki* (централизованный сбор логов) и *Prometheus AlertManager* (интеллектуальное оповещение). Автоматизированная система в режиме реального времени осуществляет непрерывный контроль доступности и производительности изолированных контейнеров ключевых микросервисов:

- **crm-ai**: микросервис ядра, отвечающий за базовую бизнес-логику интеграции и обработку входящих запросов

Описание процессов, обеспечивающих поддержание жизненного цикла программного обеспечения ИИ «Агент Энерго»

- **ml-auto-reply:** микросервис генеративного контура, отвечающий за автоматическое формирование текстовых ответов

### 3.2. Метрики контроля работоспособности

Система автоматического мониторинга фиксирует и анализирует технические параметры с периодичностью каждые 5 минут. Контролю подлежат:

- **Параметры API-эндпоинтов:** время ответа и HTTP-коды ответов по адресам /extract-entities, /classify, /answer-rag. Критическим отклонением признается превышение времени ответа более 30 секунд для эндпоинта /answer-rag или возврат кодов ошибок HTTP 500, 503.
- **Статус контейнеризации:** непрерывная проверка активности виртуальных контейнеров среды выполнения (docker compose ps).
- **Доступность внешних и внутренних технологических зависимостей:** контроль работоспособности векторной базы знаний OpenSearch (через системный запрос /\_cluster/health) и локального сервера инференса LLM (через системный запрос /v1/models).

### 3.3. Регламент профилактического контроля

Служба сопровождения правообладателя выполняет обязательные плановые процедуры проверки:

**Ежедневно:** ручная и автоматизированная верификация активного статуса всех контейнеров ПО («Up») и проверка отсутствия неотработанных предупреждений в панели мониторинга.

**Еженедельно:** проведение сквозного функционального тестирования софта путем отправки контрольного сообщения и валидации полного цикла его обработки (включая автоматическое создание черновика ответа и генерацию вложений).

Описание процессов, обеспечивающих поддержание жизненного цикла программного обеспечения ИИ «Агент Энерго»

**Контроль бэкапов:** ежедневная автоматическая проверка наличия свежего снимка (snapshot) базы знаний OpenSearch со сроком давности не более 24 часов.

## 4. Порядок устранения неисправностей и регламент инцидентов

### 4.1. Прием и регистрация обращений

Пользователи (операторы сбытовых и энергокомпаний) направляют информацию о сбоях в работе ИИ «Агент Энерго» по официальным каналам связи:

- Электронная почта технической поддержки: support@integrysys.ru

Все обращения регистрируются в корпоративной системе учета заявок с присвоением уникального номера.

### 4.2. Матрица критичности инцидентов и сроки реагирования (SLA)

SLA определяется договором на сопровождение и поддержку ПО. Базовые сроки реагирования:

| У р о в е н ь<br>критичности | Описание ситуации                                      | М а к с .<br>в р е м я<br>ответа | Макс. время<br>исправления |
|------------------------------|--------------------------------------------------------|----------------------------------|----------------------------|
| Авария                       | Полная недоступность API, система не генерирует ответы | 1 час                            | 2 часа                     |

Описание процессов, обеспечивающих поддержание жизненного цикла программного обеспечения ИИ «Агент Энерго»

|         |                                                                                                                           |         |                           |
|---------|---------------------------------------------------------------------------------------------------------------------------|---------|---------------------------|
| Высокий | Ошибки в работе одного из ИИ-модулей, некорректная классификация типов обращений при сохранении общей доступности системы | 4 часа  | 10 часов                  |
| Низкий  | Вопросы по интерфейсу, необходимость добавления новых документов в базу знаний, ложные срабатывания на редких запросах    | 8 часов | В рамках планового релиза |

#### 4.3. Регламент действий персонала при возникновении сбоев

При фиксации инцидента (системой автоматического мониторинга или по заявке пользователя) дежурный инженер службы поддержки правообладателя осуществляет анализ журналов логов (docker compose logs) и действует по следующей матрице:

| Сценарий сбоя                            | Регламентированные действия персонала правообладателя                                                                                                                                 |
|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| П а д е н и е микросервиса crm-ai        | Иницируется процедура перезапуска контейнера сервиса. При повторном падении выполняется экстренный откат на предыдущую стабильную версию Docker-образа из репозитория правообладателя |
| П а д е н и е микросервиса ml-auto-reply | Выполняется принудительный перезапуск контейнера службы средствами среды оркестрации контейнеров                                                                                      |

Описание процессов, обеспечивающих поддержание жизненного цикла  
программного обеспечения ИИ «Агент Энерго»

| Сценарий сбоя                                              | Регламентированные действия персонала правообладателя                                                                                                                                                  |
|------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Недоступность компонента LLM                               | Инженер производит перезапуск службы инференса моделей и выполняет диагностику доступности графических вычислительных ускорителей сервера (GPU) с помощью низкоуровневых системных утилит (nvidia-smi) |
| Повреждение структуры или данных векторной базы OpenSearch | Запускается процедура аварийного восстановления данных из последнего суточного снимка (snapshot) либо инициируется процесс полной переиндексации на основе сохраненной сырой базы знаний               |
| Потеря или повреждение конфигурационных файлов             | Производится экстренное восстановление параметров среды (.env, docker-compose) из версионного удаленного хранилища бэкапов                                                                             |
| Полный отказ или уничтожение серверной инфраструктуры      | Развертывание системы осуществляется с нуля на резервных мощностях с последующим накатом конфигурационных файлов и восстановлением базы OpenSearch из архивной копии                                   |

## 5. Порядок совершенствования и модернизации ПО

### 5.1. Порядок сбора и анализа предложений по модернизации

Обратная связь от предприятий ЖКХ и энергетики (запросы на новые функции, изменения в логике классификации) аккумулируется в системе управления требованиями. Аналитики правообладателя ежемесячно

## Описание процессов, обеспечивающих поддержание жизненного цикла программного обеспечения ИИ «Агент Энерго»

анализируют данные, формируют бэклог задач и передают их в отдел разработки для планового улучшения алгоритмов.

### 5.2. Регламент разработки и тестирования обновлений

Все изменения в программном коде проходят обязательное код-ревью и автоматическое тестирование. Перед развертыванием в продуктивной среде (на серверах клиентов) компоненты проходят полный цикл функционального и нагрузочного тестирования на изолированном тестовом окружении (staging-стенде) правообладателя.

Технологический процесс установки обновлений компонентов

Все работы по плановому обновлению программных модулей осуществляются инженерами в периоды минимальной пользовательской нагрузки на систему (ночное время). Технологический процесс обновления включает следующие последовательные этапы, минимизирующие риски несовместимости модулей:

- Временная остановка обработки задач микросервисом ml-auto-reply
- Обновление и пересборка контейнера основного микросервиса crm-ai из утвержденного дистрибутива правообладателя (`docker compose up -d --build crm-ai`)
- Проведение автоматизированной проверки доступности и корректности ответов базовых интерфейсов API (проверка работоспособности функций `/classify`, `/extract-entities`, `/answer-rag`)
- Обновление исполняемых файлов и запуск заблокированного ранее микросервиса ml-auto-reply
- Отправка верификационного тестового сообщения для финальной сквозной проверки работоспособности всех компонентов обновленной

### 5.3. Порядок дообучения ИИ-моделей и обновления базы знаний

При изменении нормативно-правовых актов в сфере ЖКХ (например, Постановлений Правительства РФ № 354 или № 442) ML-инженеры правообладателя производят обновление базы знаний. Актуализированные текстовые корпуса загружаются в Систему, выполняется процедура безопасного обновления индексов базы знаний OpenSearch через механизм переиндексации (reindex) с обязательным предварительным созданием защитного снимка текущего индекса

Персонал и инфраструктура обеспечения жизненного цикла

### 5.4. Требования к квалификации и составу персонала

Для поддержания всех процессов ЖЦ Системы правообладатель выделяет штатную команду специалистов со следующими минимальными компетенциями:

**Инженер технической поддержки:** знание архитектуры Системы, опыт работы с логированием.

**Системный администратор / DevOps-инженер:** экспертные знания Docker Compose, опыт управления базами данных OpenSearch и PostgreSQL, настройка Prometheus.

**ML-инженер:** опыт работы со средами инференса больших языковых моделей, навыки векторизации и подготовки текстовых данных для RAG-систем.

### 5.5. Условия размещения персонала и территориальная принадлежность

Все штатные сотрудники службы технической поддержки, системного администрирования и модернизации софта являются налоговыми резидентами Российской Федерации. Выполнение любых работ по поддержке, администрированию, развертыванию обновлений и дообучению моделей ИИ «Агент Энерго» осуществляется сотрудниками исключительно с рабочих мест, физически расположенных на территории Российской Федерации.

### 5.6. Инфраструктура обеспечения ЖЦ

Вся вспомогательная и технологическая инфраструктура, обеспечивающая ЖЦ программы, полностью локализована в РФ:

- Серверы автоматизированной системы мониторинга (Prometheus, Grafana) размещены в российском облачном контуре правообладателя.
- Удаленное хранилище для резервных копий баз данных, снимков OpenSearch и файлов конфигурации развернуто на собственных физических серверах правообладателя на территории РФ.
- Доступ к исходному коду и репозиториям Docker-образов осуществляется через каналы связи внутри периметра РФ.