ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ «ДИДЖИТАЛ ВАНДЕР»
ИНН 7203542555 · КПП 720301001 · ОГРН 1227200013655
625000, Тюменская обл., г. Тюмень, ул. Дзержинского, д. 15, помещ. 1, этаж 3 · info@dwon.ru · +7 (3452) 91-11-08

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ «PD-AI» ДОКУМЕНТАЦИЯ, НЕОБХОДИМАЯ ДЛЯ УСТАНОВКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

1. Доступ к системе

«PD-AI» — интернет-сервис; установка программного обеспечения на стороне пользователя не требуется. Доступ предоставляется после регистрации на https://app.pd-ai.ru:

1. Откройте https://app.pd-ai.ru — откроется страница «Вход в PD-AI».
2. Перейдите к регистрации, укажите адрес электронной почты и пароль, отметьте согласие с условиями Договора-оферты и на обработку персональных данных.
3. Подтвердите адрес электронной почты по ссылке из письма.
4. После входа откроется раздел «Чат». Для пользователя автоматически создаётся организация на тарифе «Бесплатный».

2. Требования к рабочему месту

Параметр	Требование
Операционная система	Любая с поддерживаемым браузером: Astra Linux, РЕД ОС, Альт, Windows 10 и выше, macOS 12 и выше, Android 10 и выше, iOS 15 и выше
Браузер	Яндекс Браузер, Google Chrome, Microsoft Edge, Mozilla Firefox, Safari — актуальные версии, с включёнными JavaScript и cookie
Сеть	Доступ по HTTPS (порт 443) к pd-ai.ru, app.pd-ai.ru, auth.app.pd-ai.ru
Голосовой ввод (необязательно)	Микрофон и разрешение браузера на его использование

3. Развёртывание экземпляра (для инженеров правообладателя)

Экземпляр PD-AI разворачивается правообладателем в управляемом кластере Kubernetes. Порядок приведён для воспроизведения промышленного контура.

3.1. Программное окружение

Компонент	Версия
Kubernetes	1.35 (k0s, Timeweb Cloud), сеть flannel, отказоустойчивая плоскость управления
Ingress / TLS	ingress-nginx, cert-manager
Управление релизами	Helm 3, kubectl
Инфраструктура как код	Terraform 1.6 и выше
Хранилища	PostgreSQL с pgvector, Redis, S3-совместимое хранилище
Платформенные сервисы	Authentik (OIDC), Lago и billing-microservice, LiteLLM, конвертация файлов, OCR, эмбеддинги, зеркало исходного кода Forgejo; распознавание речи — модель GigaAM на GPU-серверах правообладателя (г. Тюмень, ул. Дзержинского, д. 15)

3.2. Ресурсы приложения

Компонент	Запрос CPU / ОЗУ	Лимит CPU / ОЗУ
pdai-ui	0,1 / 256 МиБ	1 / 768 МиБ
pdai-chat-backend	0,1 / 256 МиБ	1 / 1 ГиБ
document-service: служебный интерфейс	0,25 / 512 МиБ	1 / 1 ГиБ
document-service: обработчик	0,25 / 512 МиБ	2 / 2 ГиБ
pdai-search-agent (агент поиска)	0,1 / 256 МиБ	2 / 1 ГиБ
pdai-doc-gen-agent (2 экземпляра)	0,05 / 64 МиБ каждый	0,5 / 256 МиБ каждый
pdai-site-check (агент проверки сайта, браузер Chromium)	0,1 / 384 МиБ	1 / 1,5 ГиБ
pd-ai.ru	0,025 / 32 МиБ	0,5 / 128 МиБ

Сумма лимитов — около 9,5 vCPU и 7,9 ГиБ ОЗУ. Рекомендуемая конфигурация с учётом платформенных сервисов — не менее четырёх рабочих узлов по 2 vCPU.

3.3. Порядок развёртывания

1. Заполнить файлы переменных окружения по образцам environments/<env>/*.tfvars.example (секреты в систему контроля версий не помещаются).
2. Создать хранилище состояния Terraform (слой 00-state).
3. make ENV=pd-ai-prod foundation-apply — сеть, кластер, реестр образов, DNS, ingress-nginx, cert-manager.
4. make ENV=pd-ai-prod platform-apply — PostgreSQL, Redis, Authentik, Lago,

billing-microservice, LiteLLM, зеркало исходного кода Forgejo.

5. Применить слои 03-applications: pd-ai, pd-ai-chat-backend, pd-ai-ui, pd-ai-search-agent, pd-ai-agent-doc-gen, pd-ai-agent-doc-gen-kii, pd-ai-agent-site-check, pd-ai-ru (пространство имён, базы данных, бакеты, приложение в Authentik, секреты).
6. В репозитории каждого компонента запустить конвейер выкатки: сборка образа на сервере сборки правообладателя с тегом коммита, публикация в Timeweb Cloud Container Registry, helm upgrade --install --wait --timeout 10m. Порядок: document-service, search-agent, doc-gen-agent, site-check, chat-backend, ui, сайт.
7. Миграции баз данных выполняются автоматически (Prisma — при старте оркестратора и отдельным заданием document-service; Alembic — заданием биллинга).
8. Загрузить корпус нормативных актов через служебный интерфейс document-service.

3.4. Проверка работоспособности

Адрес	Компонент	Проверка
pd-ai.ru	Сайт	GET /healthz
app.pd-ai.ru	Веб-приложение	GET /api/health → {"status":"ok"}
chat-backend.app.pd-ai.ru	Оркестратор	GET /api/v1/health
doc-gen-agent.app.pd-ai.ru, kii-doc-agent.app.pd-ai.ru	Агенты документов	GET /api/v1/health
pdai-site-check.pd-ai.svc.cluster.local:8080 (внутренний адрес)	Агент проверки сайта	GET /api/v1/health
document-service.app.pd-ai.ru	Служебный интерфейс корпуса	страница входа

Сквозная проверка: вход в app.pd-ai.ru, вопрос по 152-ФЗ, ответ со ссылкой на статью, открытие панели первоисточника.

Откат: helm rollback <релиз> <ревизия> -n <пространство имён> или повторная выкатка предыдущего коммита.