Утверждено Приказом
Генерального директора ООО «Нейродеск»
№ 01/02 от «02» февраля 2026 г.

Приложение № 1
к Политике конфиденциальности ООО «Нейродеск»
(утв. Приказом Генерального директора № 04/12 от «02» декабря 2025 г.)

ПОЛОЖЕНИЕ ОБ ОБЕЗЛИЧИВАНИИ ПЕРСОНАЛЬНЫХ ДАННЫХ
и механизм работы сервиса обезличивания

1. Выдержка из Политики конфиденциальности

Настоящее Приложение развивает положения Политики конфиденциальности ООО «Нейродеск» (далее - Политика) в части обезличивания персональных данных и описывает технический механизм их реализации в Сервисе.

В Политике зафиксировано:

1.1. Действие Политики распространяется на все процессы обработки персональных данных, включая обезличивание.

1.2. Обезличивание - действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту.

1.3. Обрабатываемые персональные данные подлежат уничтожению либо обезличиванию по достижении целей обработки или в случае утраты необходимости в достижении этих целей, если иное не предусмотрено федеральным законом.

1.4. Персональные данные собираются, хранятся и обрабатываются на территории Российской Федерации. Трансграничная передача персональных данных не осуществляется.

2. Назначение механизма

В текстах диалогов поддержки, заявок и документов могут содержаться персональные данные: ФИО, телефоны, номера документов, адреса и иные категории, указанные в разделе 4.

Чтобы такие данные не покидали внутренний контур системы при обращении к внешнему LLM, Сервис автоматически находит персональные данные в тексте, заменяет их плейсхолдерами и после обработки точно восстанавливает исходные значения в ответе, предназначенном пользователю.

Маскирование включено всегда. Это не опция заказчика и не режим «по запросу»: в штатной эксплуатации во внешний LLM направляется только замаскированный текст - токены и плейсхолдеры. Реальные персональные данные остаются во внутреннем

контуре на стороне Оператора. В этом состоит назначение микросервиса маскирования.

3. Принцип работы: три шага

Шаг 1. Маскирование. Сервис анализирует текст и находит персональные данные (список - в разделе 4). Каждое найденное значение заменяется плейсхолдером вида <ТИП_N_КОД>, например ФИО - <PERSON_1_a7b2c3>, телефон - <PHONE_NUMBER_1_a7b2c3>. Одновременно создаётся ключ восстановления: таблица «плейсхолдер → исходное значение».

Шаг 2. Обработка. Во внешний LLM уходит только обезличенный текст с плейсхолдерами. Исходных персональных данных и ключа восстановления в нём нет.

Шаг 3. Восстановление. Когда обработанный текст возвращается, сервис по ключу подставляет исходные значения обратно. Пользователь видит обычный читаемый текст.

Пример:
исходный текст - «Меня зовут Иван Петров, тел. +7 999 123-45-67»;
во внешний LLM уходит - «Меня зовут <PERSON_1_a7b2c3>, тел. <PHONE_NUMBER_1_a7b2c3>»;
после восстановления пользователь видит готовый ответ с исходными значениями.

4. Какие персональные данные распознаются

Распознавание настроено на российские форматы документов и данных. Ниже - соответствие категории данных типу плейсхолдера сервиса маскирования.

Категория	Тип Плейсхолдера	Примеры
ФИО	PERSON	Иван Петров; Петров Сергей Александрович
Телефон	PHONE_NUMBER	+7 999 123-45-67; 8-915-555-00-11
email	EMAIL_ADDRESS	ivanov@example.com
ИНН	RU_INN	10 и 12 цифр
Адрес регистрации/проживания (почтовый адрес)	ADDRESS	г. Москва, ул. Ленина, д. 5, кв. 10 (индекс, регион, корпус, строение)
Паспортные данные	PASSPORT	4514 123456
Дата рождения	BIRTH_DATE	15.08.1992

Банковские реквизиты (в т.ч. номера карт)	CREDIT_CARD	4276 8500 1234 5678
Удостоверение личности	RU_DRIVER_LICENSE (частично)	водительское удостоверение 77 12 123456
СНИЛС (страховой номер в Социальном фонде)	RU_SNILS	123-456-789 01
ОГРН (регистрационный номер юрлица в налоговой)	RU_OGRN	1027700132195

СНИЛС и ОГРН маскируются дополнительно: это устойчивые идентификаторы, которые могут встречаться в переписке и документах наряду с перечисленными выше категориями.

Не детектируются (не входят в текущий набор распознавания):

- пол - практически не извлекается из свободного текста;
- место рождения - редкий кейс для диалогов поддержки.

Распознавание учитывает живой язык переписки: сокращения («г.», «ул.», «кв.») и полные слова, слитное написание, строчные буквы и отсутствие пунктуации в поддерживаемых форматах. Для ряда числовых данных (паспорт, ИНН, дата рождения, водительское удостоверение, ОГРН, СНИЛС) учитываются контекстные слова рядом со значением: это снижает риск принять номер заказа или дату документа за персональные данные.

5. Как устроено распознавание

Распознавание выполняется гибридным методом: сочетаются правила по форматам (телефоны, карты, ИНН, СНИЛС, паспорт и иные строгие шаблоны) и контекстный анализ (ФИО, адреса, живые формулировки переписки). Конкретные программные фреймворки и библиотеки в настоящем Приложении не раскрываются.

6. Плейсхолдеры и ключ восстановления

Плейсхолдер - метка-заменитель вида <ТИП_N_КОД>, например <PHONE_NUMBER_1_a7b2c3>. В ней указаны тип данных, порядковый номер и уникальный код, который генерируется заново при каждом запросе. Плейсхолдеры из разных запросов не пересекаются, даже если фрагменты объединяются в одном диалоге.

Ключ восстановления - таблица соответствий «плейсхолдер → исходное значение». Пример: <PERSON_1_a7b2c3> → Иван Петров; <PHONE_NUMBER_1_a7b2c3> → +7 999 123-45-67.

Ключ хранится только во внутреннем контуре системы, на стороне компонента, который вызвал маскирование. Внешний LLM и промежуточный маршрутизатор ключ не получают ни на одном этапе. Ключ может сохраняться во внутреннем состоянии диалога на срок жизни этого состояния и удаляется вместе с ним.

Без ключа плейсхолдер бесполезен: по <PERSON_1_a7b2c3> невозможно вычислить «Иван Петров». Это не шифрование с ключом расшифровки, а замена на случайную метку. Узнать исходное значение можно, только если ключ уже есть.

Восстановление - точная обратная подстановка по ключу, буква в букву, включая пробелы и пунктуацию.

Технически механизм является обратимым (псевдонимизация): без ключа восстановления определить субъекта невозможно, при этом ключ не покидает внутренний контур. Это соответствует определению обезличивания в Политике (невозможность определить субъекта без дополнительной информации).

7. Куда направляется обезличенный текст

В текущей конфигурации обезличенный текст направляется из внутреннего контура системы во внешний облачный LLM-сервис через OpenRouter. OpenRouter маршрутизирует запрос выбранному модельному провайдеру.

Это не российский контур и не on-prem. Исходные персональные данные и ключ восстановления в OpenRouter и далее модельному провайдеру не отправляются и остаются во внутреннем контуре.

Оператор расценивает это так: во внешний LLM не передаются персональные данные, а обезличенный текст, по которому без ключа восстановления невозможно определить субъекта. Исходные персональные данные хранятся и обрабатываются во внутреннем контуре.

Оператор предусматривает переход на размещение LLM на собственной инфраструктуре (самостоятельный хостинг). После такого перехода обезличенный текст будет обрабатываться внутри собственной инфраструктуры без передачи во внешний LLM.

8. Меры безопасности

Исходные персональные данные и ключ восстановления не покидают внутренний контур. Наружу уходит только текст с плейсхолдерами.

PII-сервис не хранит тексты и ключи на диске: обработка идёт в памяти в момент запроса. Ключ держит вызывающая система во внутреннем контуре на срок жизни состояния диалога.

В служебные логи PII-сервиса пишутся только технические метрики (длина текста, количество найденных значений). Сами значения, тексты и ключи в логи не попадают.

Запросы к сервису принимаются только при предъявлении ключа доступа (токена).

9. Надёжность и ограничения

Качество измеряется автоматическими тестами текущего набора распознавания (в том числе 33 проверки адресов: поиск, отсутствие ложных срабатываний, точность восстановления) и проверками на реальных диалогах поддержки.

На проверенных данных: ложные срабатывания не зафиксированы; точность восстановления - 100%; покрытие почтовых адресов - около 93%; одно и то же значение в одном тексте маскируется одной меткой; несколько персональных данных в одной фразе распознаются одновременно.

Сервис настроен так, чтобы номера заказов, суммы, номера актов и даты документов не принимались за персональные данные.

Ограничения: редкие форматы адресов (улица без слова «улица» и без сокращения, полностью строчный текст без знаков препинания) могут не распознаться; имена с грубыми опечатками могут быть пропущены; пол и место рождения в текущем наборе не детектируются. Удостоверения личности, отличные от поддерживаемых форматов (в т.ч. вне RU_DRIVER_LICENSE), могут быть покрыты неполностью. Остальные категории (телефоны, email, ИНН, паспорт, СНИЛС, ОГРН, карты) распознаются по строгим форматам и стабильно надёжны. Полнота покрытия на данных и сценариях конкретного заказчика проверяется при внедрении.