factor [h]

Руководство пользователя и администратора

Factor H 1.0

Наименование программы	Factor H
Версия документа	1.0
Дата	18.10.2026

Аннотация *3*
1. Порядок работы с системой *3*
1.1. Вход в систему **3**
1.2. Роли пользователей **3**
1.3. Администрирование системных пользователей **4**
2. Настройка маркировки *4*
2.1. Реестр получателей **4**
2.2. Серверные параметры (администратор системы) **5**
3. Работа с документами *5*
3.1. Управление получателями **5**
3.2. Создание копий **6**
4. Загрузка найденного образца *7*
4.1. Файл или изображение **7**
4.2. Вставленный текст (copy-paste) **8**
5. Расследование *8*
5.1. Ход анализа **9**
5.2. Как опознаётся копия **9**
6. Просмотр результата *10*
6.1. Заключение **10**
6.2. Доказательная база **11**
7. Типовые ошибки и действия пользователя *11*

Аннотация

Настоящий документ представляет собой руководство пользователя (оператора) и администратора системы невидимой маркировки документов и трассировки источника утечки Factor H в составе передаваемой версии 1.0. Документ описывает роли пользователей, вход в систему, администрирование системных пользователей, настройку маркировки, работу с документами, загрузку найденного образца, проведение расследования, просмотр результата и типовые ошибки.

1. Порядок работы с системой

1.1. Вход в систему

Откройте в браузере адрес стенда (http://<адрес-хоста>:<порт>, по умолчанию порт 8000). Откроется форма входа, в которой нужно ввести логин и пароль. После входа система направит вас на рабочую страницу согласно вашей роли. Сеанс сохраняется между запросами; срок его жизни задаётся в параметрах стенда (см. руководство по установке). Для завершения сеанса используйте кнопку «Выйти».

Первые учётные записи (администратор и оператор) создаются автоматически при первом запуске системы в соответствии с заданными конфигурационными параметрами (см. руководство по установке).

1.2. Роли пользователей

В системе две роли, каждой доступен свой набор разделов:

Роль	**Доступные разделы и возможности**
Оператор	Создание документов (копий), Расследование, Управление получателями. Ведение собственного изолированного реестра получателей, выпуск копий и расследование происходит только в пределах этого реестра.
Администратор	Системные пользователи: заведение и удаление операторов и администраторов. Функции выпуска копий и расследований выполняются под учётной записью оператора.

Изоляция. У каждого оператора свой реестр получателей; операторы не видят данные друг друга. Администратор управляет только списком системных пользователей.

1.3. Администрирование системных пользователей

Войдите в систему администратором, откроется раздел «Системные пользователи». Для того чтобы добавить пользователя, в форме внизу укажите Логин, Пароль, выберите роль (оператор / админ) и нажмите **«+ Добавить»**. Для начала работы заведите хотя бы одного пользователя с ролью «**оператор**».

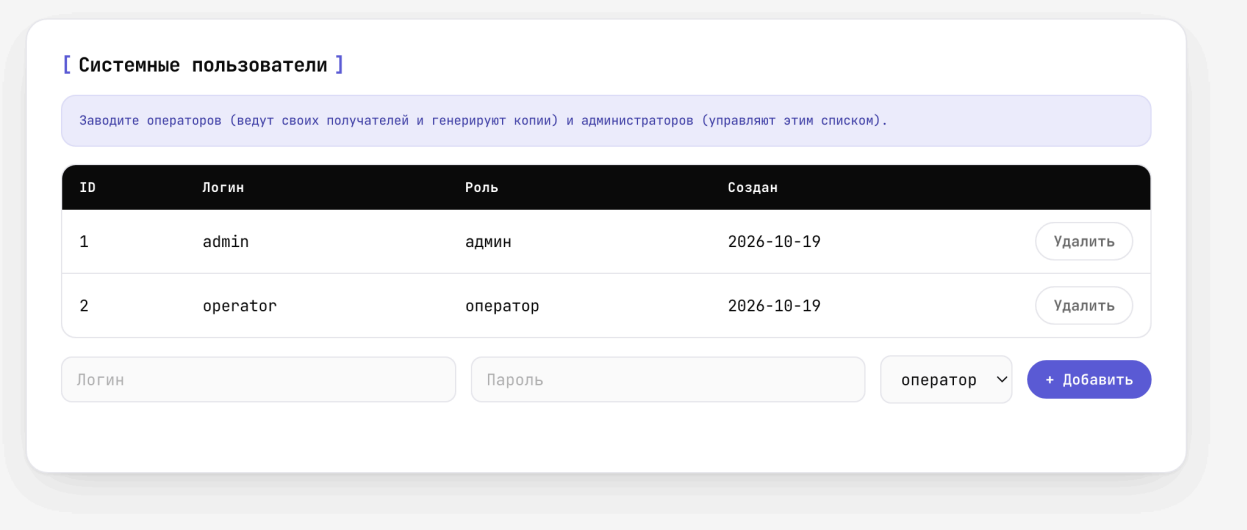


Рисунок 1 Добавление системного пользователя

2. Настройка маркировки

Маркировка в Factor H автоматическая и многослойная. В каждую копию всегда встраиваются все доступные для выбранного формата файла независимые слои маркировки. Возможность изменения параметров маркировки через интерфейс отсутствует, что снижает вероятность ошибки оператора. Настройка системы включает две составляющие: ведение реестра получателей (выполняется оператором) и задание серверных параметров (выполняется администратором при развёртывании системы)

2.1. Реестр получателей

Основной прикладной настройкой является реестр получателей (раздел «Управление получателями», см. Рисунок 2). Атрибут получателя служит единственным ключом маркировки: он определяет индивидуальный код получателя, встраиваемый в каждую копию, и область поиска при расследовании, которая ограничивается реестром текущего оператора. В связи с этим «атрибут получателя» обязателен и должен быть корректным.

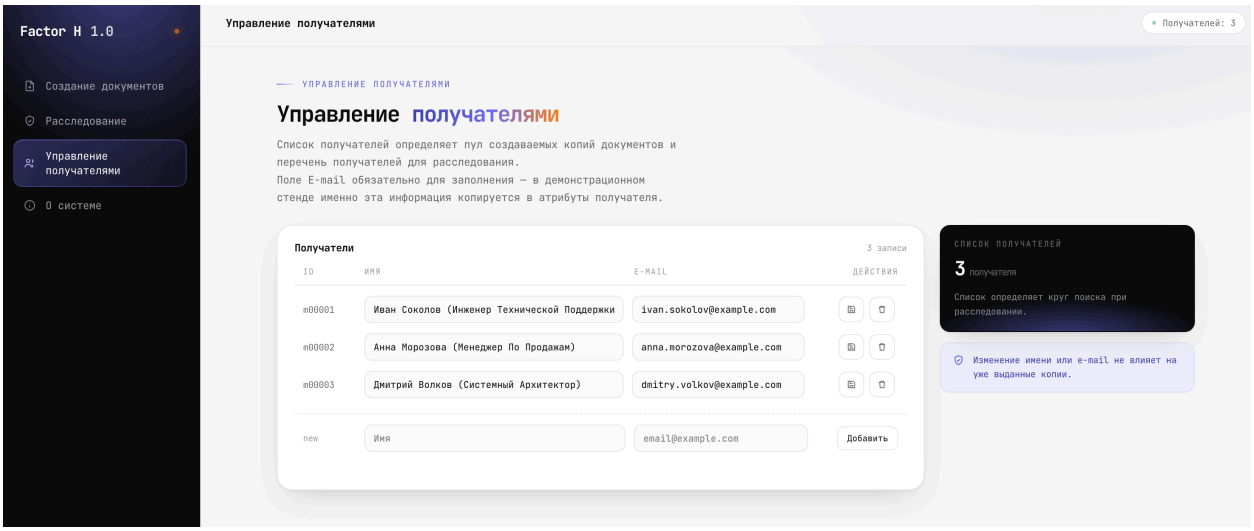


Рисунок 2 Реестр получателей

2.2. Серверные параметры (администратор системы)

Серверные параметры системы (секреты, поддерживаемые языки для оптического распознавания символов, срок жизни сеанса, и другие) задаются в файле окружения «.env» до первого запуска и описаны в руководстве по установке.

3. Работа с документами

Под «документами» понимаются исходные документы, на основе которых система выпускает персональные копии для получателей. Порядок работы: сначала задается реестр получателей, затем по нему выпускаются копии.

3.1. Управление получателями

Раздел «Управление получателями». В таблице отображаются столбцы ID, Имя, E-mail, Действия. Для добавления получателя в нижней строке укажите Имя и E-mail (e-mail обязателен) и нажмите «Добавить». Записи можно редактировать и удалять (см. Рисунок 2).

Реестр получателей одновременно:

- определяет состав выпускаемых копий, при выпуске создаётся по одной копии на каждого получателя реестра;
- определяет область поиска при расследовании (в пределах реестра текущего оператора).

Изменение имени или e-mail не влияет на уже выданные копии: копия несёт метку, сформированную в момент выпуска.

3.2. Создание копий

Работа выполняется в разделе «Создание документов»:

1. Перетащите исходный документ в зону «Оригинал документа» или выберите файл на диске (см. Рисунок 3). Поддерживаются форматы PDF, DOCX, XLSX, PPTX. При отсутствии выбранного файла используется демонстрационный документ.

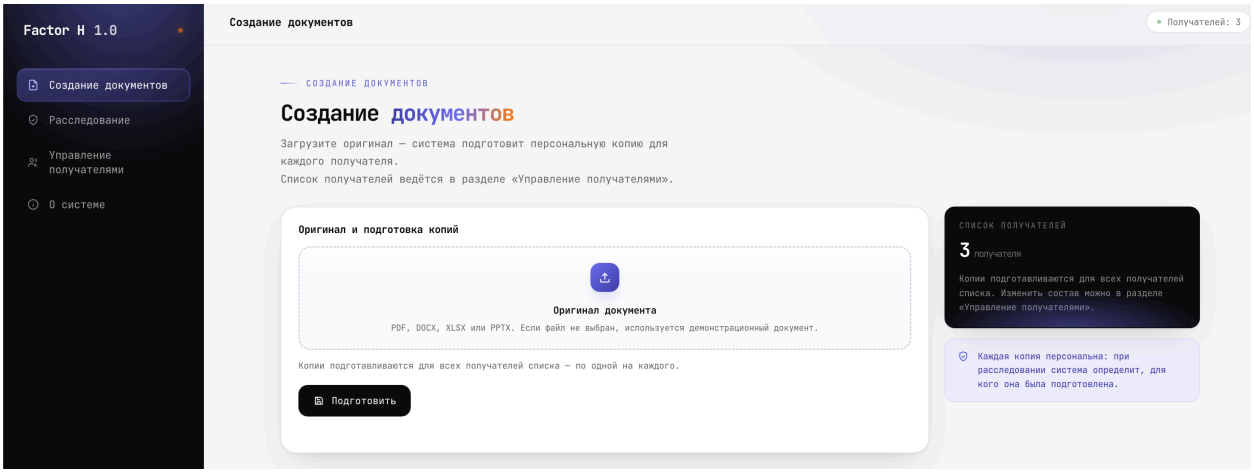


Рисунок 3 Загрузка исходного документа

2. Нажмите «Подготовить» — система выпустит по одной копии на каждого получателя из реестра и отобразит таблицу с данными получателей (см. Рисунок 4).

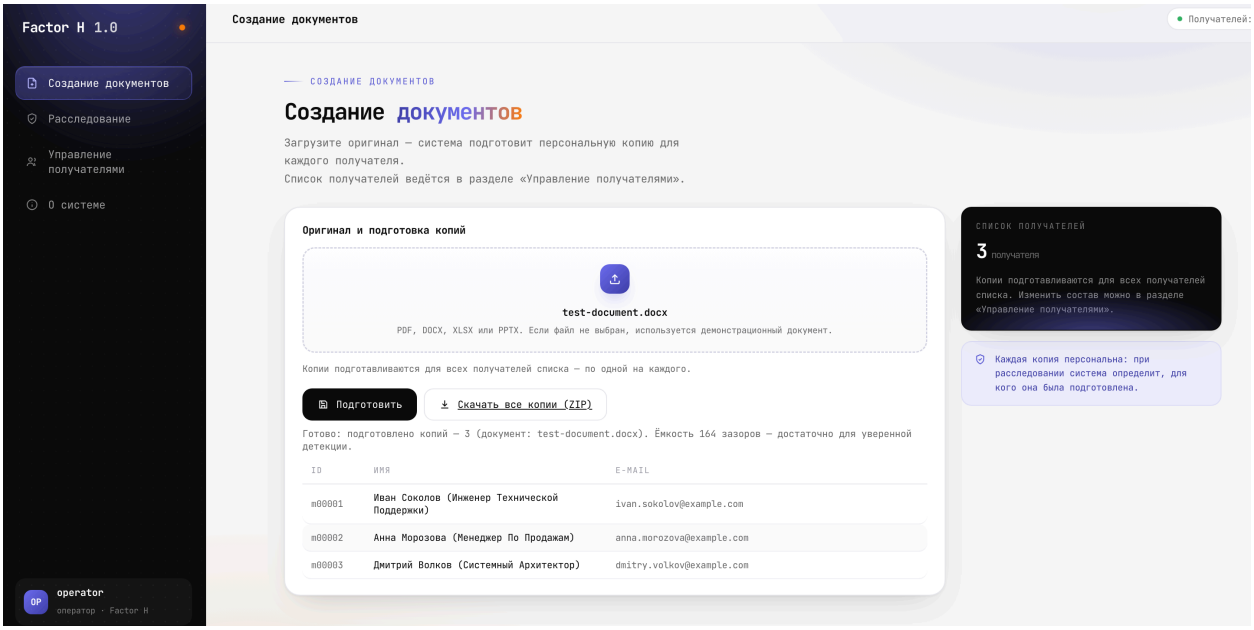


Рисунок 4 Подготовка копий с встроенной маркировкой

3. Нажмите «Скачать все копии (ZIP)». В архиве содержатся файлы, имена которых имеют вид copy_<id>.<расширение> (по одному на каждого получателя).
4. Выдайте каждому получателю соответствующую копию. Видимый текст, состав, содержимое и смысл документа не изменяются; маркировка невидима.

Состав встраиваемых в копию слоёв защиты зависит от формата исходного документа.

4. Загрузка найденного образца

Найденный образец загружается в разделе **«Расследование»** одним из двух способов.

4.1. Файл или изображение

Перетащите файл в зону «Перетащите файл сюда» или нажмите, чтобы выбрать на диске (см. Рисунок 5). Поддерживаются форматы файлов:

Файлы-копии:

PDF, DOCX, XLSX, PPTX

Изображения (скриншот / фото / скан):

- PNG, JPG/JPEG, TIFF/TIF, BMP

Ограничение размера файла до 50 МБ; формат определяется по сигнатуре содержимого, а не по расширению. Загрузка запускает расследование автоматически.

- Достаточно фрагмента — одной страницы с текстом; загружать весь документ не обязательно.
- Текст должен быть читаемым; допустимы сжатие, кадрирование и небольшой наклон, если текст разборчив.
- Скриншот средствами устройства надёжнее фотографии экрана камерой; чем больше текста и чем он чётче тем надёжнее опознание.

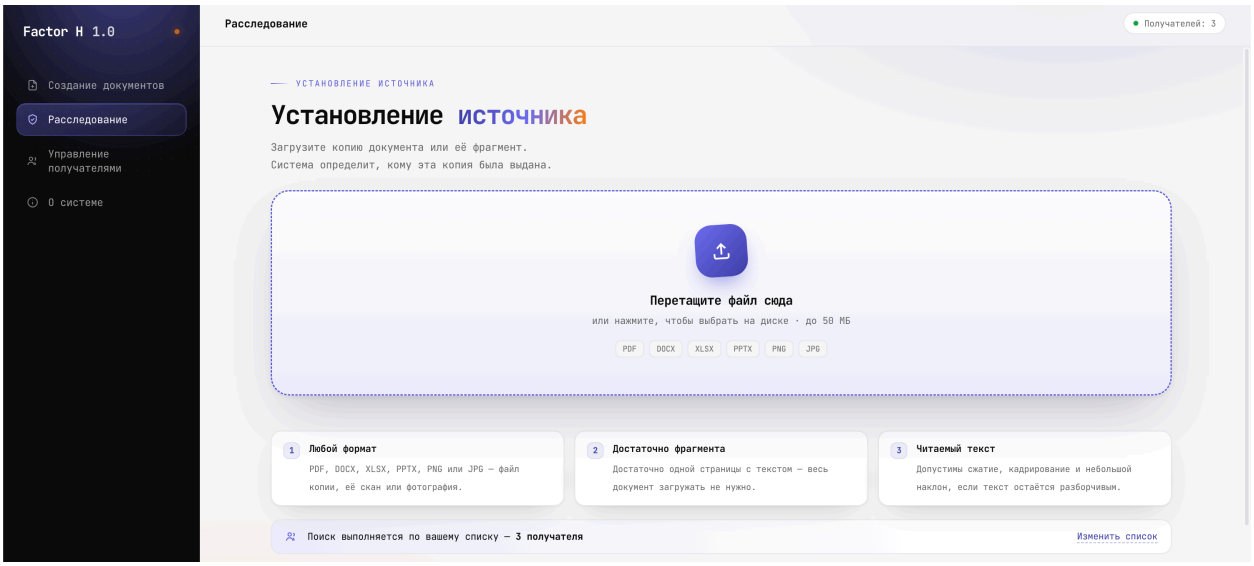


Рисунок 5 Расследование по файлу

4.2. Вставленный текст (copy-paste)

Если доступен только скопированный текст, вставьте его в поле **«...или вставьте текст из копии»** и нажмите **«Проверить текст»** — проверка пройдёт без загрузки файла. Работает для текста из копий **DOCX и XLSX** (см. Рисунок 6).

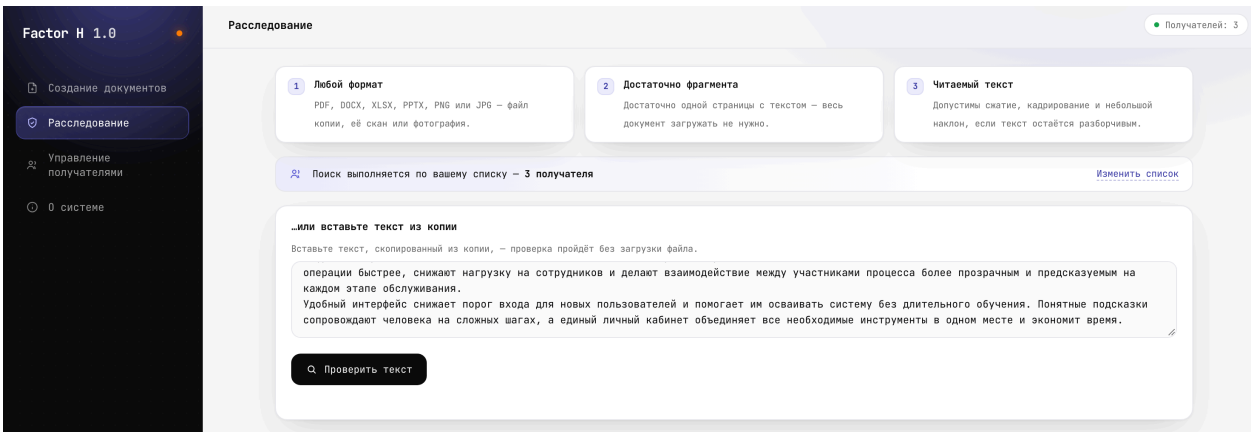


Рисунок 6 Расследование по фрагменту текста

5. Расследование

В ходе расследования по загруженному образцу устанавливается получатель (источник). Область поиска ограничивается реестром получателей текущего оператора: в шапке отображается счётчик «Получателей: N», список можно изменить по ссылке «Изменить список»

5.1. Ход анализа

На панели **«Анализ образца»** отображаются этапы обработки:

- Приём образца (Принимает загруженный файл и определяет его формат по сигнатуре содержимого)
- Подготовка изображения (Если это скриншот, фотография или скан, то улучшает изображение (масштаб, освещённость), чтобы надёжно прочитать текст)
- Анализ содержимого (Находит и считывает невидимую маркировку)
- Определение получателя (По считанной маркировке вычисляет, какому получателю соответствует образец)
- Сверка со списком получателей (строит рейтинг)
- Оценка достоверности(Оценивает, насколько уверенно можно назвать источник, и либо подтверждает получателя с показателем достоверности, либо сообщает, что «источник не определён»)

По завершении панель сворачивается и показывается заключение (см. Рисунок 7).

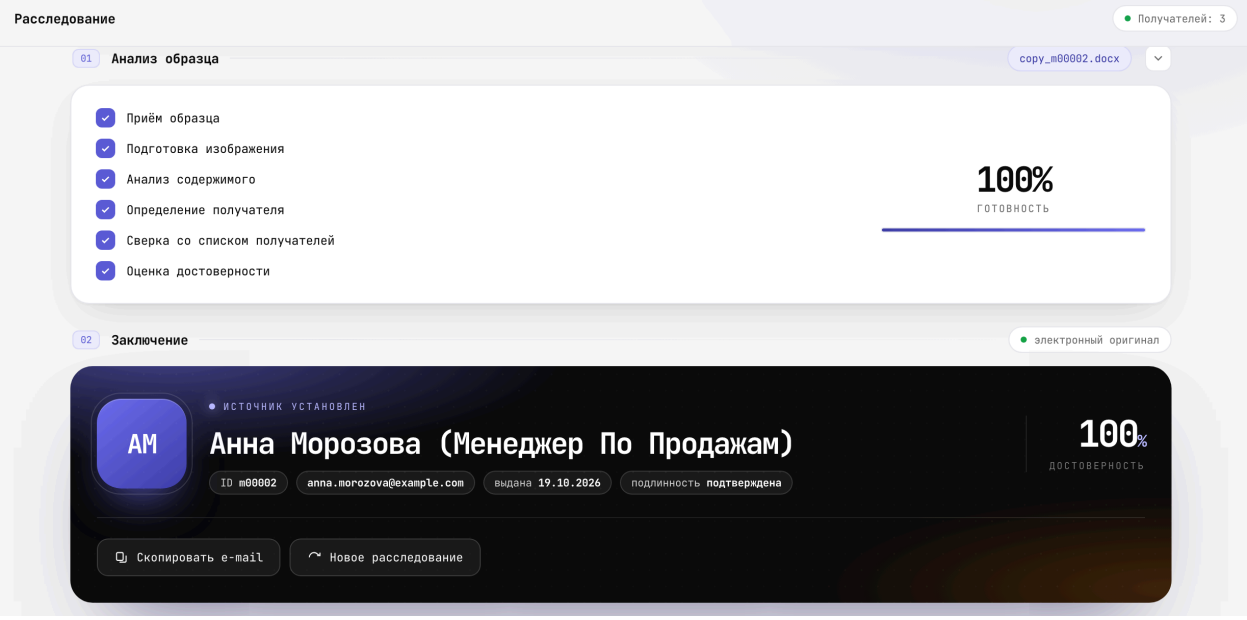


Рисунок 7 Анализ образца

5.2. Как опознаётся копия

Система определяет получателя по невидимой метке, автоматически выбирая самый надёжный доступный способ для конкретного образца:

- если предоставлен файл копии, то маркировка считывается из файла напрямую. Это наиболее быстрый способ с наивысшей достоверностью;
- если предоставлены скриншот, фотография или скан, то метка восстанавливается по изображению и сопоставляется со списком получателей;
- если предоставлен только скопированный текст, то получатель определяется по вставленному тексту без загрузки файла.

Чем полнее и качественнее образец, тем выше достоверность вывода. Подробнее о свойствах метки и слоях защиты копий см. «Функциональные характеристики Factor H».

6. Просмотр результата

6.1. Заключение

Блок «Заключение» отображает один из исходов:

- **«Источник установлен»** — указываются получатель (имя, e-mail, технический идентификатор), достоверность (в процентах) и, при наличии в метке, дата выдачи; статус «подтверждено». Доступны кнопки «Скопировать e-mail» и «Новое расследование» (см. Рисунок 8)
- **«Источник не определён»** — сигнала недостаточно для достоверного вывода. Получатель не указывается; система не формирует предположительный вывод при отсутствии достаточного сигнала. Рекомендации приведены в разделе 7.

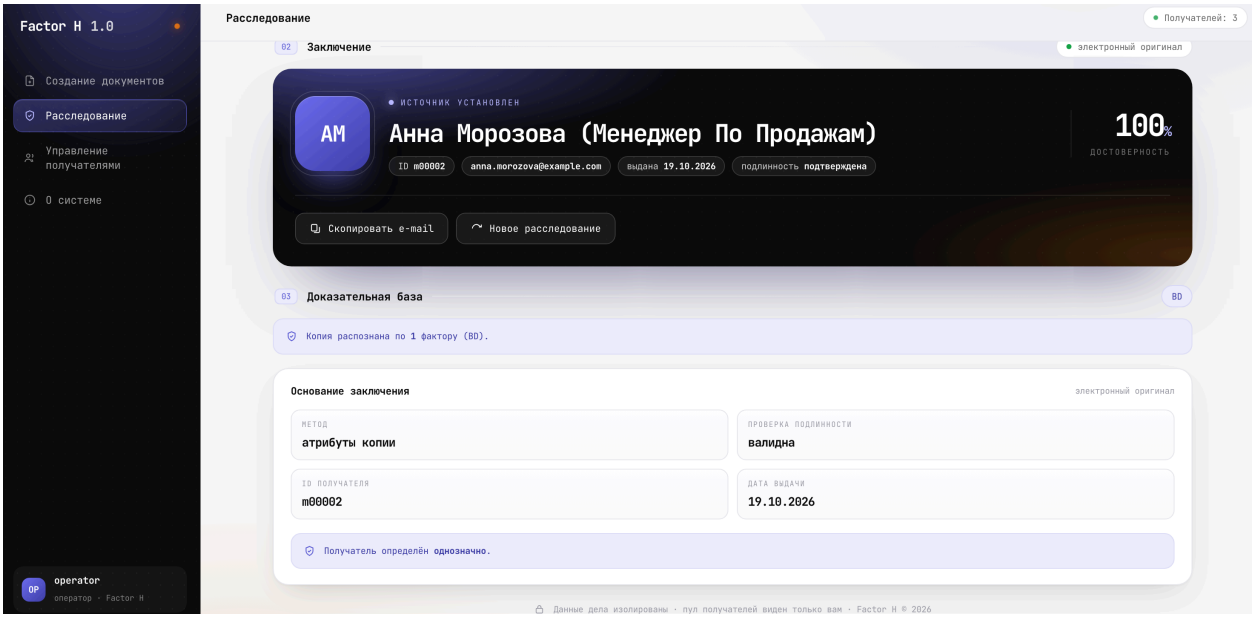


Рисунок 8 Блок "Заключение" по успешному расследованию

6.2. Доказательная база

Блок «Доказательная база» показывает, на чём основано заключение. Состав материалов зависит от того, что было предоставлено:

- **по файлу копии** — подтверждённые сведения о копии: получатель (его идентификатор) и дата выдачи, а также отметка о подтверждении метки. Если получатель отсутствует в текущем реестре, выводится предупреждение о том, что сведения получены из самой копии;
- **по изображению или тексту** — наглядные материалы сопоставления:
  - «Соответствие по получателям» — у истинного получателя соответствие существенно выше, чем у остальных (топ-20);
  - «Детализация совпадения» — насколько образец совпал с меткой названного получателя, а также показатели «Объём образца», «Сигнал / порог», «Совпадение, %»;
  - «Рейтинг получателей» — упорядоченный список кандидатов (ID, получатель, e-mail, оценка), топ-10.

Пример доказательной базы приведён на Рисунок 9.

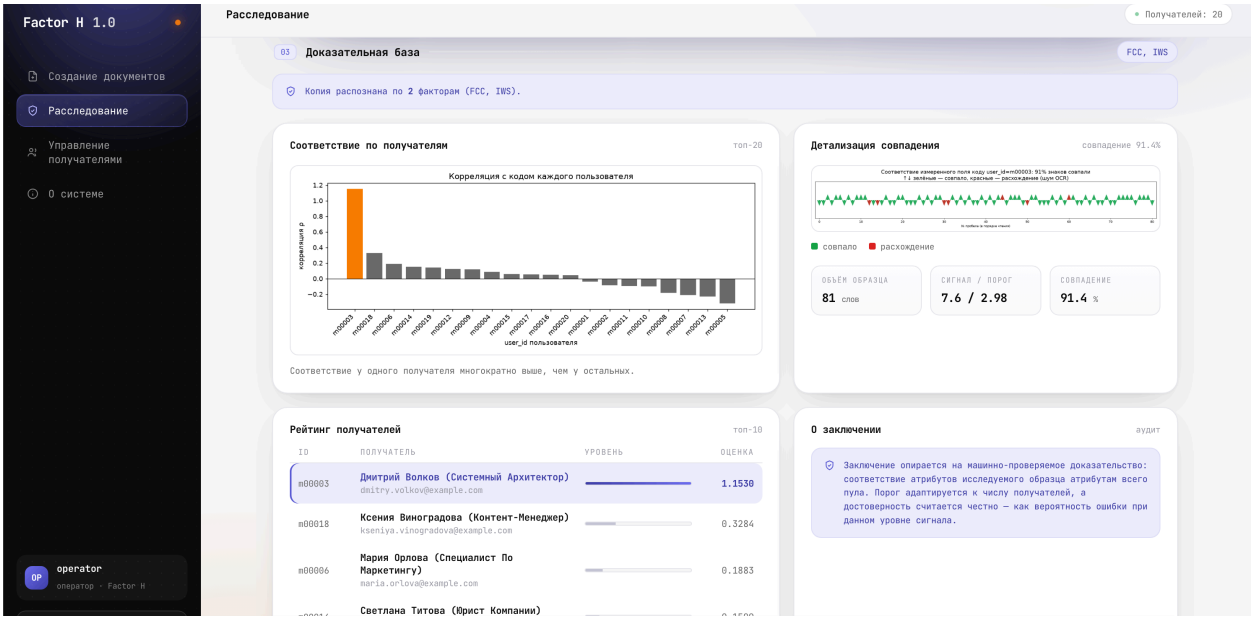


Рисунок 9 Доказательная база

7. Типовые ошибки и действия пользователя

Симптом	Причина	Действие
«Источник не определён» на скриншоте/фото	Мало текста или низкое качество образца	Загрузить фрагмент с бóльшим объёмом текста, крупнее и чётче; предпочесть скриншот фотографии

Симптом	Причина	Действие
«Источник не определён» на числовой таблице/выгрузке	В документе почти нет текста — несущую метить не на чем	Опознавать по файлу копии, если он доступен
В заключении — получатель не из вашего списка (предупреждение)	E-mail прочитан из метаданных копии, а в реестре его нет	Проверить, тот ли оператор/реестр; при необходимости добавить получателя в «Управление получателями»
Расследование идёт долго	Тяжёлый OCR на большом изображении/многостраничном скане	Загружать один разворот или страницу с текстом, а не весь документ
Формат не поддержан	Устаревшие .doc/.xls/.ppt	Пересохранить в DOCX/XLSX/PPTX или PDF
Кириллица не распознаётся по скриншоту	Не установлен словарь OCR для русского языка	Задать TESSERACT_LANG=rus+eng (в поставляемом образе)
Не работает вход под администратором	Пароли из «.env» применяются только при первом старте	Сменить пароль в веб-интерфейсе или пересоздать стенд (удалить том данных) с заданными паролями
Расследование по вставленному тексту не находит источник	Мало текста или текст из PPTX/PDF	Вставить больше текста (DOCX/XLSX); для PPTX/PDF опознавать по файлу/изображению
Скриншот justified-документа опознаётся ненадёжно	Выравнивание по ширине может влиять негативно на расследование по фото	В данной ситуации рекомендуется провести распознавание по файлу копии