

Интеграция МИС с Voice2Med по FormAPI

Введение

Руководство является выжимкой из [спецификации](#) и будет полезна разработчикам МИС, перед которыми стоит задача по созданию прототипа интеграции с программным комплексом Voice2Med по FormAPI предназначенным для заполнения шаблонных форм.

Общие требования

Для взаимодействия Voice2Med и МИС, FormAPI использует протокол полнодуплексной связи WebSocket. В связи с тем что часть операций больше соответствует синхронной модели запросов, предусмотрена также альтернатива для таких запросов в виде REST API.

Данная интеграция подразумевает подключение к клиентской части Voice2Med (Рисунок 1). В этом случае захват звука, его передача на сервер и некоторые другие операции берёт на себя клиент Voice2Med.

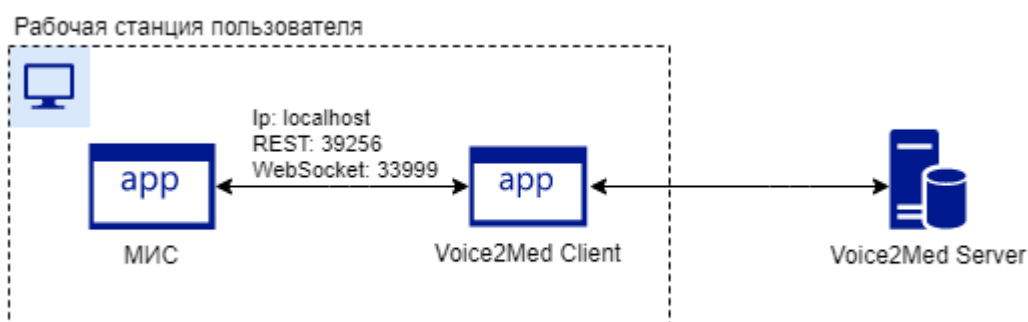


Рисунок 1 – Принципиальная схема деплоя

Функциональные требования

Реализованный прототип (МИС) должен обладать следующими функциональными возможностями:

1. [Как определить что V2M клиент установлен](#)
2. [Подключение к Voice2Med при запуске МИС \(или логине пользователя\)](#)
3. [Отключение МИС от клиентской части Voice2Med](#)
4. [Отслеживание состояния сервиса распознавания речи](#)
5. [Настройка синонимов полей на форме](#)
6. [Настройка постфиксов полей на форме](#)
7. [Настройка синонимов для перечислений](#)
8. [Запуск распознавания](#)
9. [Остановка распознавания](#)
10. [Навигация по элементам формы и её заполнение](#)
11. [Корректировка ранее заполненных полей](#)

Use-Cases

Как определить что V2M клиент установлен

В первую очередь следует объяснить зачем это нужно. Кейс достаточно прост: даже если форма поддерживает голосовое заполнение, без установленного V2M клиента, оно будет недоступно. Следовательно пользователю должен как-то об этом узнать. Как вариант можно не показывать элементы интерфейса, отвечающие за голосовое заполнение или показать соответствующее сообщение об ошибке и кнопку подключения.

Сценарий:

1. Пользователь запускает МИС и логиниться
2. МИС проверяет - установлен ли V2M

Примечание

Windows:

Следует проверить вот эти два раздела в реестре:

1) V2M клиент:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Installer\UpgradeCodes\22120261CAA61E74888C217F55ABCADB
```

2) V2M Standalone:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Installer\UpgradeCodes\56244D91C8F612F4287504B65E43BD1F
```

Если хоть один из них в наличии, значит V2M установлен.

Linux:

Следует проверить наличие файлов

1) /usr/bin/voice2med/DictationClient

2) /usr/local/bin/voice2med/DictationClient

Если хоть один из них в наличии, значит V2M установлен.

3. Если МИС установлен, то работаем в соответствии с кейсом "[Подключение к Voice2Med](#)"
4. Если МИС не установлен, то подключаться к нему нет смысла, равно как и пытаться постоянно переподключиться. Все формы поддерживающие голосовое заполнение в этом случае будут выглядеть как на [рисунке 4](#). За тем исключением что сообщение может быть другим. Например: "Voice2Med не установлен. Обратитесь к системному администратору".
5. При нажатии на кнопку "Подключиться", процедура подключение повторится
 - a. Проверка установлен ли V2M
 - b. Подключение к V2M если он установлен

Подключение к Voice2Med при логине пользователя

1. Пользователь запускает Voice2Med
2. Пользователь выполняет логин в МИС
3. МИС проверяет установлен ли Voice2Med (см. [“Как определить что V2M клиент установлен”](#))
4. МИС вычитывает настройки подключения к Voice2Med

Примечание

Windows:

Пользовательские настройки подключения к Voice2Med находятся в файле

1) V2M клиент: %localappdata%\Speech Technology Center\DictationApp.Client\Settings\FormApiServer.Properties.Settings.config

2) V2M Standalone: %localappdata%\Speech Technology Center\DictationApp\Settings\FormApiServer.Properties.Settings.config

Дефолтные значения:

WebSocket порт: 33999

Linux:

Настройки подключения расположены в файле `~/.config/Speech Technology Center/DictationClient/settings.ini` в секции FormApiServer.

Пример секции:

```
[FormApiServer]
Port=33999
RestPort=39256
```

5. МИС пытается подключиться к клиентской части Voice2Med посредством WebSocket. При подключении МИС должен передать информацию о МИС и пользователе которым выполнен логин (см. раздел спецификации [“Подключение к Voice2Med”](#)).

Если форма, поддерживающая заполнение голосом, была открыта до момента подключения к V2M, то в окне МИС должна отображаться соответствующая индикация об попытке подключения к V2M, а кнопка активации распознавания - задизейблена. (Рисунок 2).



Рисунок 2 – индикация процесса подключения МИС к Voice2Med

6. Откроется форма с кнопкой запуска распознавания (Рисунок 3).

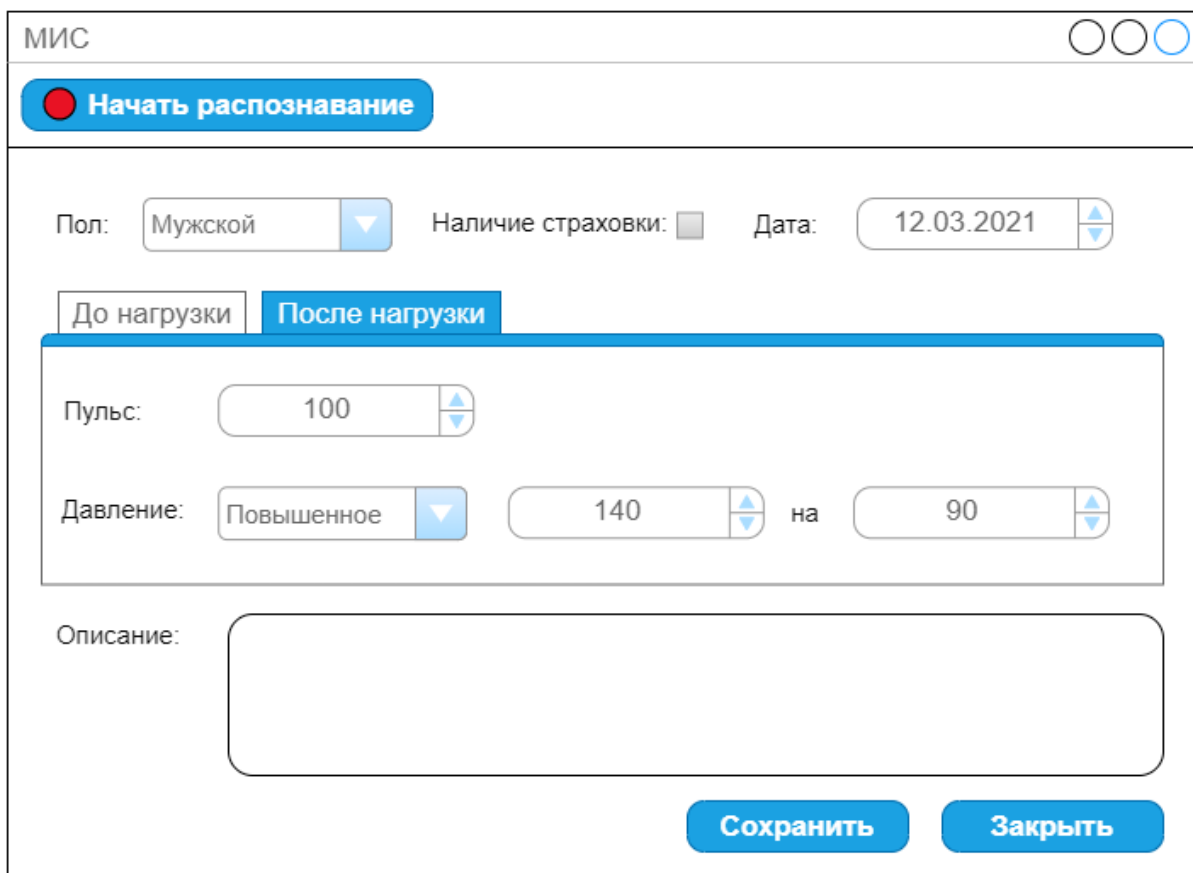
The image shows a window titled 'МИС' with three window control buttons in the top right corner. Below the title bar, there is a prominent blue button labeled 'Начать распознавание' with a red circular icon, indicating it is active. Below this button, there is a form with several fields: 'Пол:' with a dropdown menu showing 'Мужской'; 'Наличие страховки:' with a greyed-out checkbox; 'Дата:' with a date picker showing '12.03.2021'. Below these fields, there are two tabs: 'До нагрузки' and 'После нагрузки', with the latter being selected and highlighted in blue. Under the 'После нагрузки' tab, there are three input fields: 'Пuls:' with a value of '100'; 'Давление:' with a dropdown menu showing 'Повышенное', followed by two numeric input fields with values '140' and '90' separated by the word 'на'. At the bottom of the form, there is a large empty text area labeled 'Описание:'. At the bottom right of the window, there are two blue buttons: 'Сохранить' and 'Закрыть'.

Рисунок 3 – МИС подключен к Voice2Med и готов к работе

Отключение МИС от клиентской части Voice2Med

1. Пользователь выполняет логат или завершает работу МИС подключенному к Voice2Med.
2. МИС разрывает WebSocket соединение.

Отслеживание состояния сервиса распознавания речи

На форму, которая может быть заполнена голосом, добавляется панель управления голосовым заполнением протокола. На этой панели располагаются:

- кнопка включения и выключения распознавания речи
- строка с информационным сообщением

Именно эта панель отображает состояние сервиса распознавания речи. Возможные статусы сервиса распознавания:

Отсутствует подключение к сервису распознавания речи

Сценарий:

1. пропало соединение по WebSocket-у между МИС и V2M клиентом
2. МИС отображает сообщение об отсутствии подключения к серверу распознавания, а кнопка активации распознавания заменена на кнопку подключения (Рисунок 4)
3. МИС автоматически раз в 30 секунд пробует подключиться к Voice2Med, пока открыта форма, поддерживающая заполнение голосом

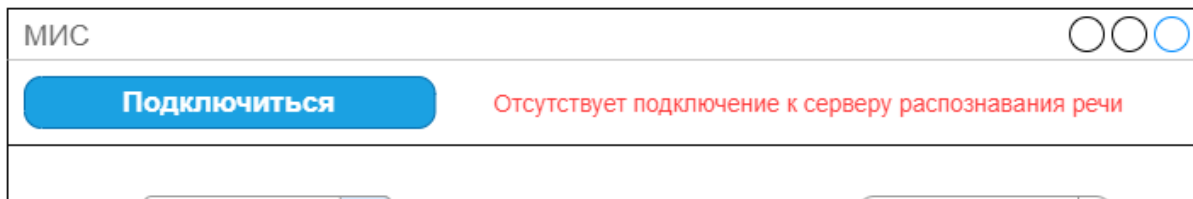


Рисунок 4 – Отсутствие подключения к Voice2Med

Подключение к сервису распознавания речи успешно установлено

Сценарий:

1. пользователь нажал кнопку "Подключиться" (рисунок 4)
1. соединение по WebSocket-у между МИС и V2M клиентом успешно установлено
2. на МИС нет сообщений об ошибках, и есть кнопка старта распознавания (рисунок 5)

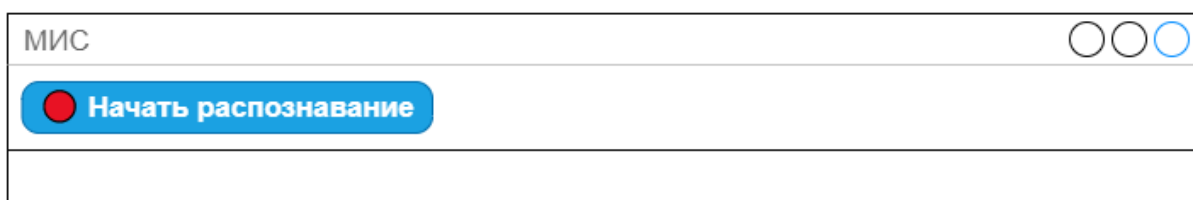


Рисунок 5 – Подключение к Voice2Med успешно установлено. Распознавание выключено

Распознавание в процессе запуска

Сценарий:

1. пользователь нажал кнопку “Запустить распознавание”
2. пришло сообщение о том что распознавание в процессе запуска:

```
{
  "MessageType": 0,
  "Data":
  {
    "Mode": 3,
    "State": 0,
    "Info": "... "
  }
}
```

Примечание

Активные режимы распознавания: Mode >= 2

Состояние этого режима: State == 0 (Starting)

Более подробно [здесь](#)

3. МИС заменяет кнопку “Начать распознавание” на кнопку “Остановить распознавание” (рисунок 6)

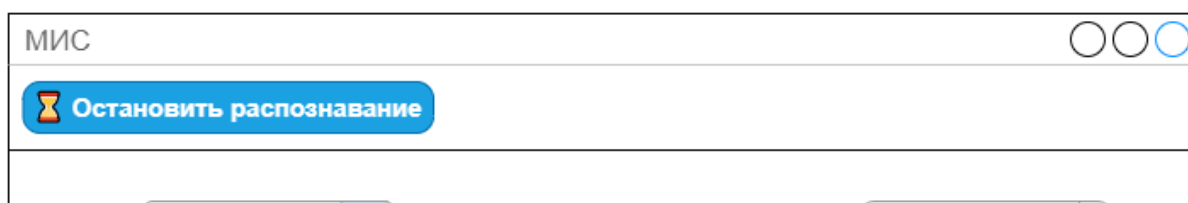


Рисунок 6 – Процесс запуска распознавания

Распознавание запущено и работает

Сценарий:

1. пришло сообщение о том что распознавание успешно запущено:

```
{
  "MessageType": 0,
  "Data":
  {
    "Mode": 3,
    "State": 1,
    "Info": "...
  }
}
```

Примечание

Активные режимы распознавания: Mode >= 2

Состояние этого режима: State == 1 (Started)

Более подробно [здесь](#)

2. МИС отображает кнопку “Остановить распознавание” (рисунок 7)

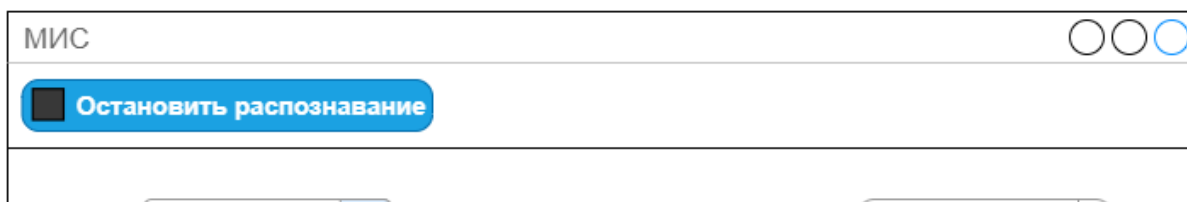


Рисунок 7 – Распознавание запущено и работает

Распознавание остановлено

Сценарий:

1. пользователь нажал кнопку “Остановить распознавание”
2. пришло сообщение о том что распознавание остановлено:

```
{
  "MessageType": 0,
  "Data":
  {
    "Mode": 0,
    "State": 1,
    "Info": "...
  }
}
```

Примечание

Неактивные режимы распознавания: Mode < 2
(Т.е. режимы ожидания: None и Voice Activation)

Состояние этого режима: State == 1 (Started)

Более подробно [здесь](#)

3. МИС вернулся в состояние готовности запустить распознавание (рисунок 8)

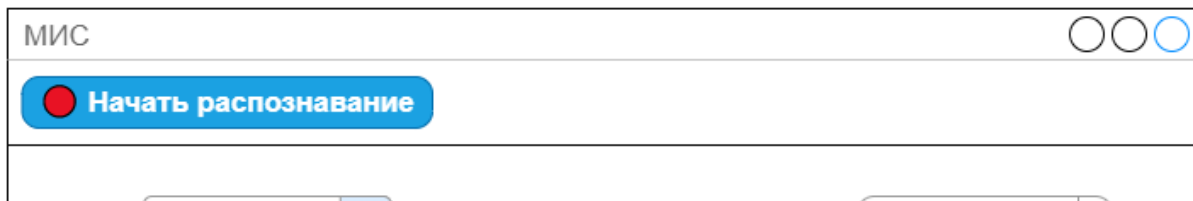


Рисунок 8 – Распознавание выключено

Отображение ошибки, возникшей в процессе работы

Сценарий:

1. от V2M пришло сообщение о том что произошла ошибка:

```
{
  "MessageType": 5,
  "Data":
  {
    "Message": "Некорректный формат загружаемых метаданных"
  }
}
```

2. МИС отображает эту ошибку (рисунок 9)

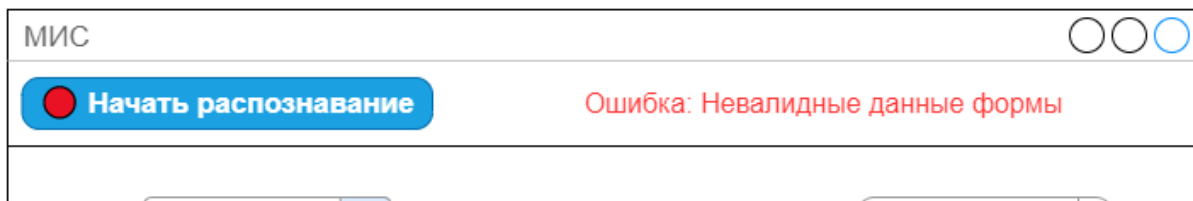


Рисунок 9 – Отображение ошибки

3. сбросить ошибку имеет смысл:

- a. либо по таймеру
- b. либо при нажатии на кнопку старта распознавания

Настройка синонимов полей на форме

При старте распознавания формы, МИС передаёт в V2M метаданные этой формы. Подробнее о формате этих метаданных рассказано в спецификации. В рамках данного кейса важно понимать что у любого поля на форме, должно быть “имя” или “название”. Именно по этому имени пользователь будет обращаться к полю и заполнять его. Например если на форме есть целочисленное поле с именем “Пульс”, то чтобы его заполнить пользователь должен будет произнести фразу “Пульс сто двадцать”, и тогда поле “Пульс” должно быть заполнено значением 120.

Такое “имя” у поля далеко не всегда совпадает с меткой этого поля. Например на форме может быть отображена метка в виде аббревиатуры: “ККР”, или фразы целиком: “кранио-каудальный размер”. При этом врач захочет произносить его в виде транскрипции “Кэ Ка Эр”, которую уже не уместно располагать в метке поля.

Более того, имён у одного и того же поля может быть несколько. Например на форме будет поле “КВР”. Часть врачей будет произносить: “Ка Вэ Эр”, часть “Косой вертикальный размер”, ещё часть просто “Вертикальный размер”. Поэтому в API предусмотрена возможность задавать набор синонимов для каждого поля.

Таким образом администратор МИС во время настройки формы, поддерживающей голосовое заполнение, должен иметь возможность задать для каждого поля на форме набор таких синонимов.

Должна быть возможность настроить синонимы не только для поле, но и для групп (закладки, контейнеры итд)

Сценарий:

1. Администратор создает в конструкторе новый шаблон\форму
2. Администратор имеет возможность указать каким-то образом, что форма будет поддерживать заполнение голосом (например добавив на форму некий специальный компонент “Голосовое заполнение”)
3. Администратор добавляет новый элемент на форму (поле или группу)
4. При попытке сохранить поле без синонимов (и наличии компонента “Голосовое заполнение”), администратор увидит предупреждение о том что синонимы не заданы и заполнение голосом может работать некорректно
5. Администратор может нажать правой кнопкой на это поле и в контекстном меню выбрать “Задать синонимы”
6. Откроется окно со списком фраз-синонимов (имен) этого поля
7. Администратор имеет возможность добавить, изменить или удалить любой из синонимов (синонимы могут содержать только символы кириллицы)

Следует понимать что синонимы - это фактически транскрипция названия поля. И если ввести туда что-то вроде “ьюъ^” - то это вряд ли будет корректно распознано.

Настройка постфиксов полей на форме

Постфиксы чем-то похожи на синонимы по реализации (см. предыдущую главу), но имеют несколько иное смысловое назначение. Так если на форме мы расположим два числовых поля “Рост” и “Пульс”, то пользователь сможет произносить заполнять их при помощи голосовых команд “Рост 170” и “Пульс 120” соответственно. Однако на практике пользователю может быть привычнее произносить фразу “Рост 170 сантиметров” и “Пульс 120 ударов в минуту” или даже “Пульс 120 ударов”. В данном примере значения “сантиметров”, “ударов в минуту” и “ударов” это и есть постфиксы. Они не влияют на итоговое значение которое должно быть записано в поле (170 и 120), но влияют на то какими фразами допустимо заполнять эти поля. Так без постфиксов поле “Рост” может быть заполнено только фразой “Рост 170” (т.е. название поля + цифра), а если задать постфиксы, то допустимо и короткое произношение “Рост 170” и полное “Рост 170 сантиметров”. Обратите внимание что следует предусмотреть возможность ввода нескольких вариантов постфиксов, как минимум из-за склонений слов. С тем же полем “Рост” это будут: “сантиметра”, “сантиметров”.

Сценарий:

1. Администратор создает в конструкторе новый шаблон\форму
2. Администратор имеет возможность указать каким-то образом, что форма будет поддерживать заполнение голосом (например добавив на форму некий специальный компонент “Голосовое заполнение”)
3. Администратор добавляет новый элемент на форму (поле или группу) и задаёт ему синонимы как в кейсе [“Настройка синонимов”](#)
4. Администратор может нажать правой кнопкой на это поле и в контекстном меню выбрать “Задать постфиксы”
5. Откроется окно со списком постфиксов для этого поля
6. Администратор имеет возможность добавить, изменить или удалить любой из постфиксов (постфиксы могут содержать только символы кириллицы)

Следует понимать что постфиксы, как и синонимы - это фактически транскрипции. И если ввести туда что-то вроде “ьюъ^” - то это вряд ли можно будет корректно распознать.

Настройка синонимов для перечислений

Для полей с типом Link в метаданных передается массив пар Key + Value, где Key - идентификатор записи в справочнике, а Value - фраза, которую должен произнести пользователь для выбора этого значения из справочника. Например на форме есть поле "Пол", с вариантами возможных значений:

- { Key="0", Value="Мужской" }
- { Key="1", Value="Женский" }

И для заполнения такого поля, пользователь должен произнести: "Пол женский" или "Пол мужской".

Однако не все перечисления и справочники настолько просты. В некоторых справочниках отображаемое выражение зачастую содержит спецсимволы или сокращения. Такие значения просто не смогут быть распознаны системой. В таком случае полезно иметь возможность задать синонимы для конкретного значения из справочника (по аналогии с синонимами полей).

Сценарий:

1. Администратор редактирует справочник и видит список всех значений справочника
2. Администратор нажимает правой кнопкой мыши на запись в справочнике и выбирает пункт меню "задать синонимы"
3. Открывается окно со списком синонимов для этой записи справочника
4. Администратор имеет возможность добавить, изменить или удалить любой из синонимов (синонимы могут содержать только символы кириллицы)

Следует понимать что синонимы - это фактически транскрипци. И если ввести туда что-то вроде "ьюъ^" - то это вряд ли будет корректно распознано.

Запуск распознавания

1. [МИС подключился к Voice2Med](#)
2. Пользователь нажал на кнопку активации распознавания
3. МИС по WebSocket соединению передает сообщение на V2M содержащее полные метаданные формы в формате JSON

Примечание

Сам формат метаданных описан в спецификации. Прежде всего рекомендуем изучить [“Короткий пример работы с формами”](#).

При необходимости подробности и детали можно найти в спецификации: [“Запуск распознавания формы”](#), [“Описание классов, их полей и особых требований к ним”](#).

4. МИС ожидает сообщения об успешном старте распознавания:

```
{
  "MessageType": 0,
  "Data":
  {
    "Mode": 3,
    "State": 1,
    "Info": "... "
  }
}
```

Важно

Помимо этого сообщения до старта распознавания будут приходить и другие сообщения (о прогрессе старта или ошибке при старте распознавания). Реакция на них описывается [чуть выше](#).

5. На МИС кнопка активации распознавания изменяется на кнопку “Остановить распознавание” (Рисунок 10). Пользователь может приступить к заполнению формы.



Рисунок 10 – Индикация что МИС в активном сеансе распознавания

Остановка распознавания

1. Активирован сеанс распознавания
2. Пользователь нажимает кнопку “Закончить распознавание”
3. МИС отправляет [сообщение в Voice2Med об окончании сеанса голосового заполнения формы](#)
4. Voice2Med отправляет в ответ [уведомления об изменении режима распознавания](#). Ожидаемая последовательность сообщений:
 - 4.1. текущий режим “шаблонное распознавание” остановлен;
 - 4.2. новый режим “распознавание отключено” начал процесс запуска;
 - 4.3. режим “распознавание отключено” запустился;

Примечание

Реакция МИС на эти сообщения [описывается выше](#)

Навигация по элементам формы и её заполнение

Элементы, расположенные на форме, могут и как правило должны быть различным образом сгруппированы (закладки на форме, GroupBox, Expander, и прочие контейнеры). Возможна ситуация, когда в двух разных группах будут содержаться одинаковые элементы.

Например на рисунке 3 в группах “До нагрузки” и “После нагрузки” есть одинаковое поле “Пульс”. Чтобы не возникало путаницы какое из полей на форме требуется заполнить когда пользователь произнесет “Пульс 78”, V2M хранит в памяти текущее активное поле. Подробнее описано [здесь](#).

Таким образом, при голосовом заполнении, важно чтобы и МИС, и V2M, и главное пользователь считали активным одно и то же поле.

Примечание

Под активным полем понимается поле, находящееся в фокусе

Для того чтобы обеспечить такую синхронизацию МИС должен поддерживать обработку следующих кейсов:

1. МИС уведомляет V2M о смене фокуса на форме

Сценарий:

1. Пользователь активировал распознавание на форме
2. Пользователь вручную перенес фокус на новое поле или закладку (при помощи мышки, клавиатуры или иным способом изменил фокус на стороне МИС)
3. МИС отправляет V2M [уведомление о смене фокуса](#).

Пример:

```
{
  "MessageType": 5,
  "Data": "ea0ac1b8-d91d-4b98-8d63-4f565048ab77"
}
```

2. V2M уведомляет о смене фокуса на поле

Сценарий:

1. Пользователь активировал распознавание на форме
2. Пользователь произнёс название какой-то группы или поля (например “После нагрузки”)
3. От V2M пришло уведомление о том что [поле/группа стала активной](#):

```
{
  "MessageType": 2,
  "Data":
  {
    "Key": "ea0ac1b8-d91d-4b98-8d63-4f565048ab77"
  }
}
```

4. МИС, получив это сообщение, устанавливает визуальный фокус на этом поле/группе и делает его/её активнойй.

Примечание

Важно также обеспечить видимость этого элемента для конечного пользователя. Например, если элемент находится внутри какой-то другой вкладки, которая была не активна, то она должна стать активной. Или если элемент, которому присваивается фокус, находится внутри ScrollBar-a, то нужно обеспечить автоскроллинг к этому элементу.

Примечание

Отправлять обратное уведомление о фокусе на V2M в этом случае нет необходимости.

3. V2M уведомляет о заполнении поля

Сценарий:

1. Пользователь активировал распознавание на форме
2. Пользователь произнёс название и значение поля (т.е. фразу соответствующую корректному заполнению поля). Например: "Дата первое октября".
3. От V2M пришло уведомление о заполнении поля "Дата" соответствующим значением:

```
{
  "MessageType": 1,
  "Data":
  {
    "Type": 4,
    "Value": "01.10.2021",
    "Key": "0d2f0d70-be6a-4f02-b71f-b3e2adc80542"
  }
}
```

4. МИС, получив это сообщение, устанавливает визуальный фокус на этом поле, и заполняет соответствующим значением.

Примечание

[Примечания из предыдущего сценария](#) актуальны и здесь.

Примечание

Как правило при записи значений в соответствующие поля, предыдущие значения перезаписываются. Однако для текстовых полей имеет смысл использовать другую логику. Для прототипа достаточно [дописать текст в конец строки](#).

При заполнении формы обратите внимание:

Поля со свободным вводом (обычные текстовые поля) можно заполнять в двумя способами:

1) Произнеся название поля и его значение одной фразой. Например: "Комментарий: 1 2 3"

2) Установив фокус на поле и произнеся только его значение. Например: "Комментарий" пауза, после которой фокус переходит на это поле. Пользователь произносит "1 2 3" и текущее поле со свободным вводом заполняется значением "1 2 3"

Поля без свободного ввода могут быть заполнены только полным произношением одной фразой: "Толщина 45 миллиметров" (т.е. название поля + значение + необязательный постфикс).

Поддерживаемые FormApi типы полей можно посмотреть в соответствующем разделе спецификации "[Доступные для распознавания типы данных](#)".

Корректировка ранее заполненных полей

Поля которые заполняются голосом могут на момент заполнения уже содержать какие-то данные. В этой ситуации МИС должен принять решение - что делать с уже имеющимися данными (перезаписать новым значением, дописать к имеющемуся значению итд). Для прототипа предлагается реализовать следующую логику:

Тип поля	Действие МИС	Комментарий
Текстовое поле	Дописывать новые данные по месту расположения курсора. Если выделен блок текста, то заменять его на новый текст. При фокусе на текстовое поле следует устанавливать курсор в конец строки. При нахождении курсора в середине текста (в кейсах корректировки например), следует обеспечить устранение двойных пробелов.	в полной реализации рекомендуется предусмотреть более сложную логику
Поле Link с поддержкой свободного ввода	Перезаписывать значение	
Остальные	Перезаписывать значение	

Также при заполнении форм следует обеспечить обработку команд:

1. “Очистить поле” - очищает текущее поле, сбрасывая его значение к дефолтному или пустому, если это возможно
2. “Закончить распознавание” - завершает распознавание формы

Приложение

Примеры возможных ошибок и нештатных ситуаций

Ошибки которые могут возникнуть в процессе взаимодействия МИС и V2M как правило делятся на две группы:

1. потеряно сетевое соединение между МИС и V2M клиентом
2. ошибки возникшие в процессе работы с V2M (от V2M пришло сообщение с кодом 5)

Примечание

Поведение МИС в обоих случаях описано выше, в разделе [“Отслеживание состояния сервиса распознавания речи”](#)

Ниже рассмотрены типичные ошибки, относящиеся ко второй группе:

Ошибка	Описание	Что можно сделать
Ошибка при старте распознавания	МИС отправляет сообщение о старте распознавания формы и метаданные этой формы. Обычно причина в некорректном формате передаваемых метаданных.	<ol style="list-style-type: none">1) Попробовать локализовать проблему исключая из метаданных поля одно за другим.2) Сравнить структуру отправляемых метаданных с примерами из спецификации3) Убедиться что для всех полей заданы синонимы (хотя бы один)
Сообщение об отсутствии лицензии при запуске распознавания	Превышено число активных сессий распознавания, разрешённых вашей лицензией. Ну или лицензии нет вообще.	<ol style="list-style-type: none">1) Приобрести лицензию2) Работать и вести отладку с публичным стендом (инструкция по подключению на сайте voice2med.ru), не требующим

		лицензии
Сообщение об отсутствии лицензии при старте приложения V2M	Скорее всего истекла триальная лицензия	То же что и в предыдущем пункте