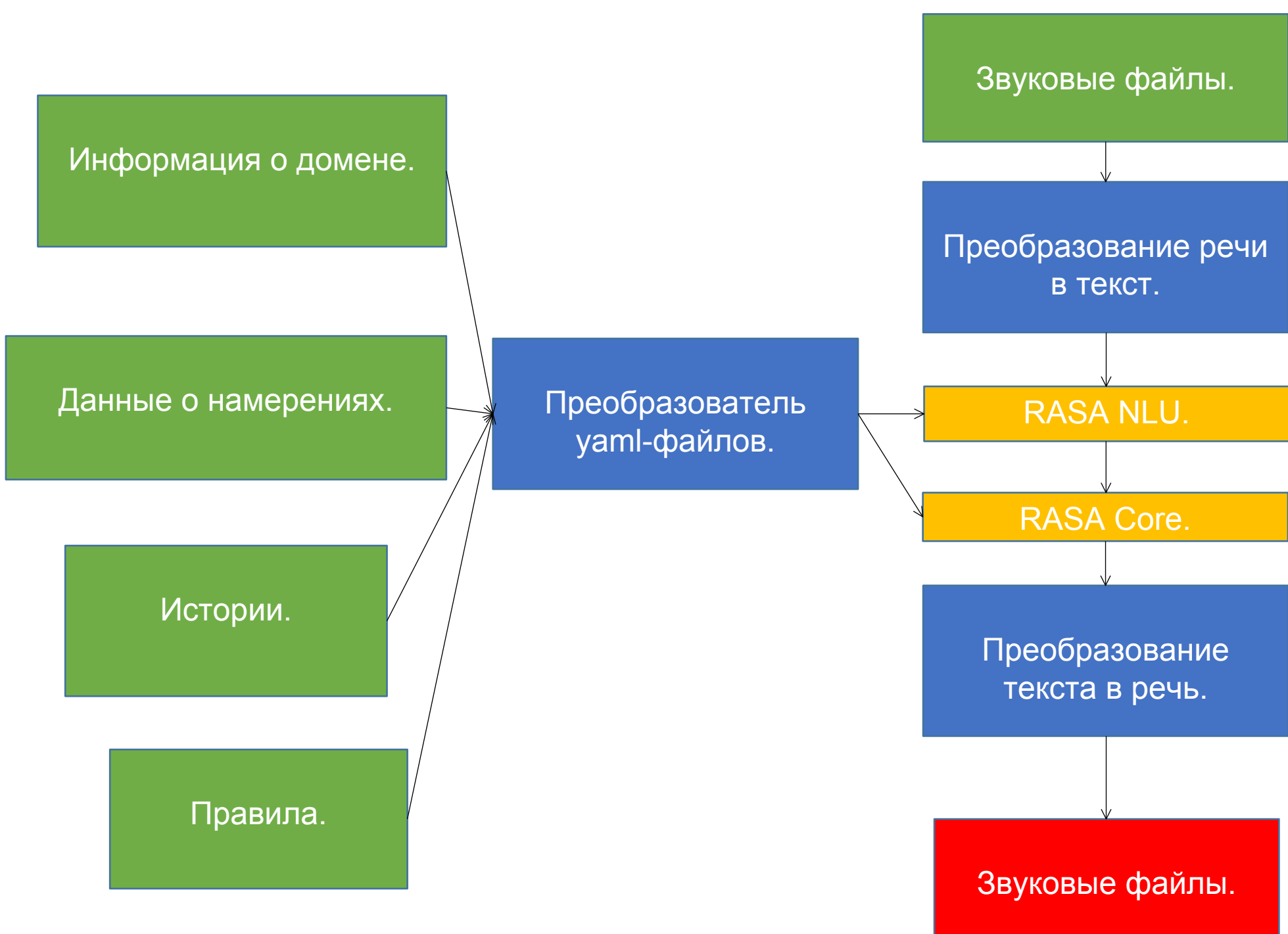


Решаемые задачи.

- Ответ на звонки (с синтетической генерацией голоса) с распознаванием интенгов пользователей и простой неиерархичной структурой ответов
- Автообзвон компаний с информационными звонками
- С точки зрения модели разница лишь в том, что в первом случае инициатива у пользователя, а во втором случае у нас.



Входные звуковые файлы.

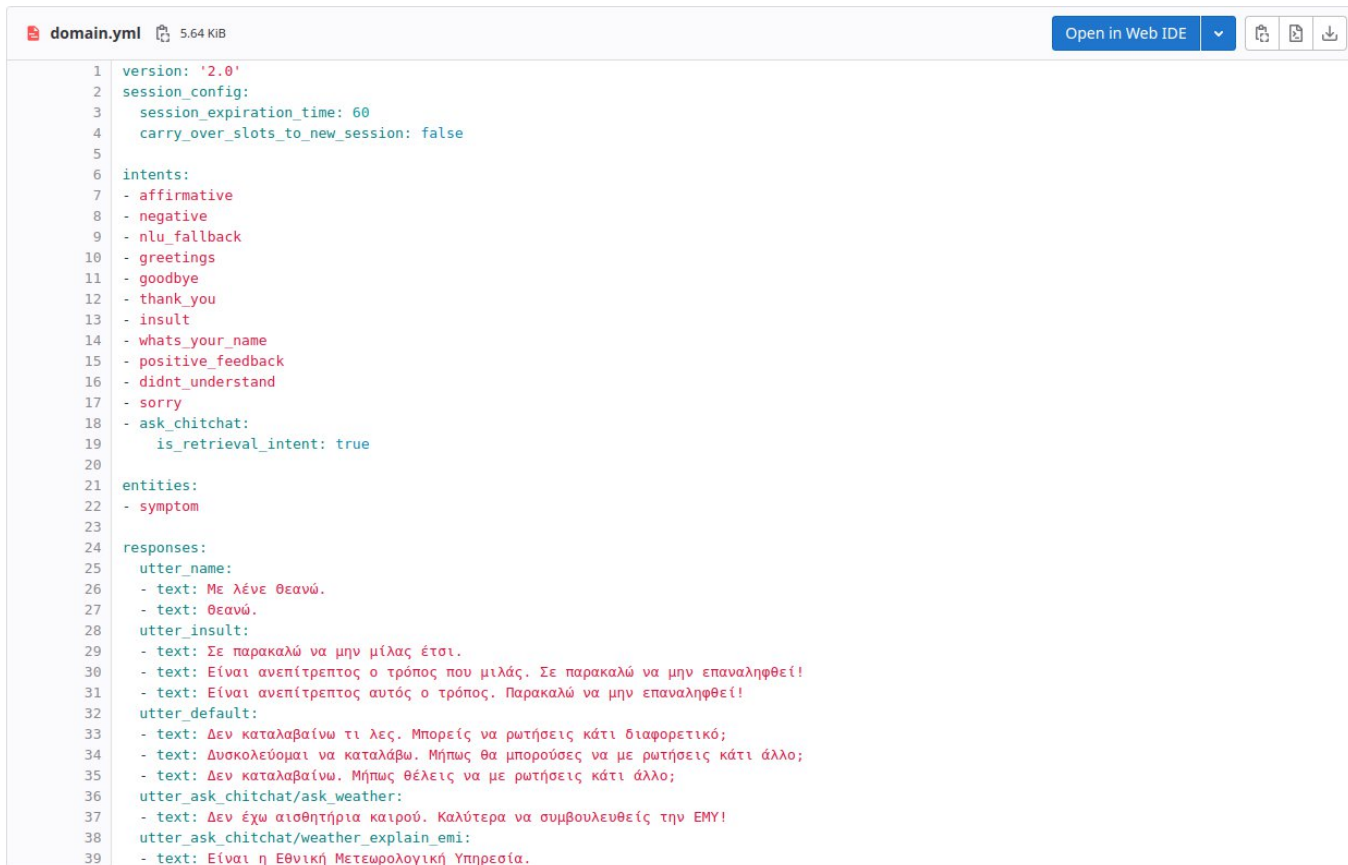
- Предполагается самостоятельно записывать небольшое количество входных звуковых файлов для проверки модели.
- Скорее всего, файлы придётся преобразовывать, чтобы с ними могла работать предобученная модель.

Преобразование речи в текст.

- Предполагается использовать предобученную модель.
- Один из кандидатов ESPnet2:
https://zenodo.org/record/4541727#.Y_YdXtJBzmg
- Второй кандидат kaldi:
<https://github.com/SergeyShk/Speech-to-Text-Russian>
- Третий кандидат:
<https://huggingface.co/jonatasgrozman/wav2vec2-large-xlsr-53-russian>

Информация о домене.

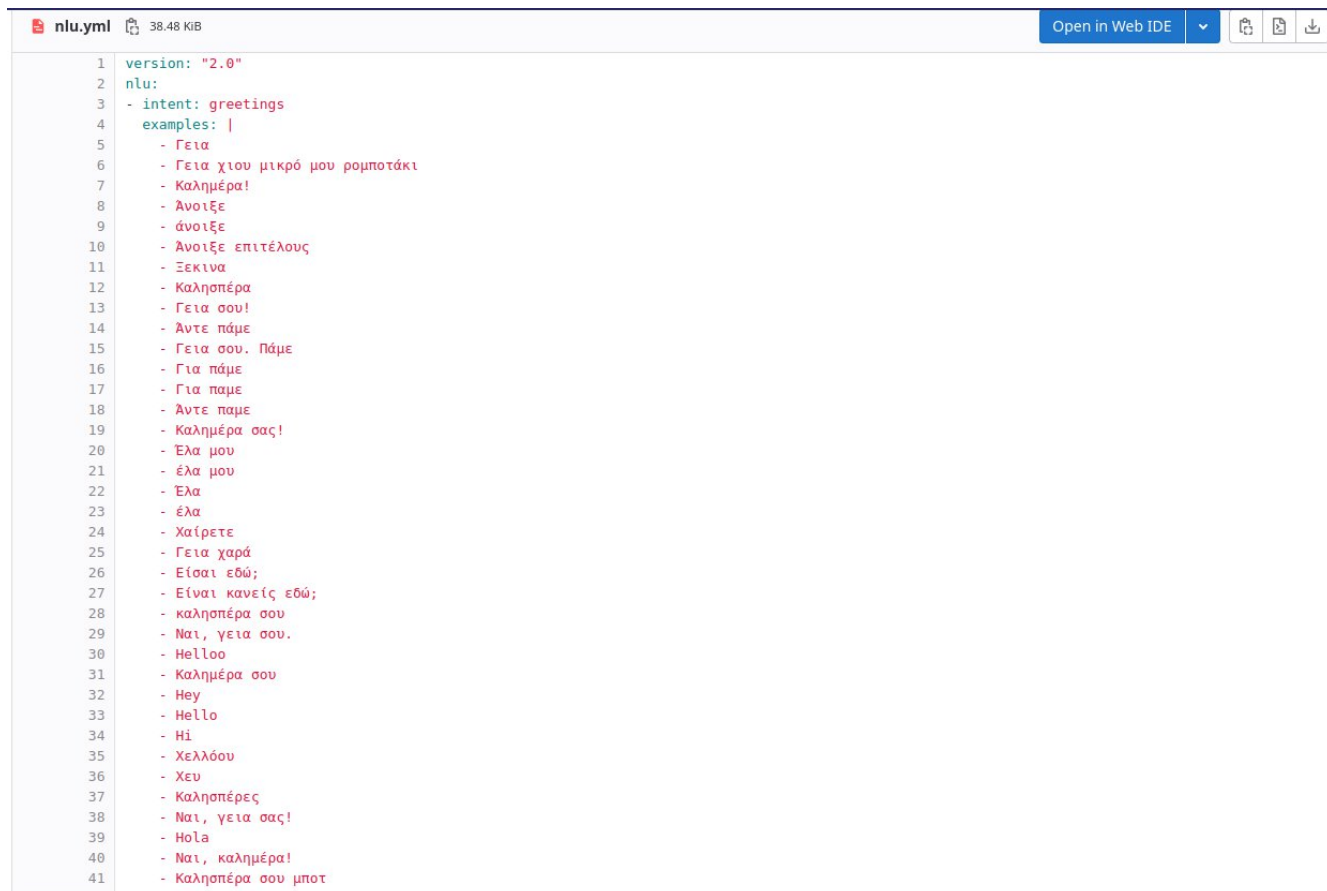
- Домен — задачи, который предстоит решать боту.



```
domain.yml 5.64 KiB
1 version: '2.0'
2 session_config:
3   session_expiration_time: 60
4   carry_over_slots_to_new_session: false
5
6 intents:
7   - affirmative
8   - negative
9   - nlu_fallback
10  - greetings
11  - goodbye
12  - thank_you
13  - insult
14  - whats_your_name
15  - positive_feedback
16  - didnt_understand
17  - sorry
18  - ask_chitchat:
19    is_retrieval_intent: true
20
21 entities:
22  - symptom
23
24 responses:
25   utter_name:
26     - text: Με λένε Θεανώ.
27     - text: Θεανώ.
28   utter_insult:
29     - text: Σε παρακαλώ να μην μίλας έτσι.
30     - text: Είναι ανεπίτρεπτος ο τρόπος που μιλάς. Σε παρακαλώ να μην επαναληφθεί!
31     - text: Είναι ανεπίτρεπτος αυτός ο τρόπος. Παρακαλώ να μην επαναληφθεί!
32   utter_default:
33     - text: Δεν καταλαβαίνω τι λες. Μπορείς να ρωτήσεις κάτι διαφορετικό;
34     - text: Δυσκολεύομαι να καταλάβω. Μήπως θα μπορούσες να με ρωτήσεις κάτι άλλο;
35     - text: Δεν καταλαβαίνω. Μήπως θέλεις να με ρωτήσεις κάτι άλλο;
36   utter_ask_chitchat/ask_weather:
37     - text: Δεν έχω αισθητήρια καιρού. Καλύτερα να συμβουλευθείς την ΕΜΥ!
38   utter_ask_chitchat/weather_explain_emi:
39     - text: Είναι η Εθνική Μετεωρολογική Υπηρεσία.
```

Информация о намерениях.

- Намерения — то, что может быть нужно пользователю.



```
1 version: "2.0"
2 nlu:
3   - intent: greetings
4     examples: |
5       - Γεια
6       - Γεια χιου μικρό μου ρομποτάκι
7       - Καλημέρα!
8       - Άνοιξε
9       - άνοιξε
10      - Άνοιξε επιτέλους
11      - Ξεκίνα
12      - Καλησπέρα
13      - Γεια σου!
14      - Άντε πάμε
15      - Γεια σου. Πάμε
16      - Για πάμε
17      - Για παμε
18      - Άντε παμε
19      - Καλημέρα σας!
20      - Έλα μου
21      - έλα μου
22      - Έλα
23      - έλα
24      - Χαίρετε
25      - Γεια χαρά
26      - Είσαι εδώ;
27      - Είμαι κανείς εδώ;
28      - καλησπέρα σου
29      - Ναι, γεια σου.
30      - Hello
31      - Καλημέρα σου
32      - Hey
33      - Hello
34      - Hi
35      - Χελλόου
36      - Χευ
37      - Καλησπéρες
38      - Ναι, γεια σας!
39      - Hola
40      - Ναι, καλημέρα!
41      - Καλησπέρα σου μποτ
```

Информация о историях.

- Истории — это сценарии развития диалога.

stories.yml 189 bytes

Open in Web IDE



```
1 version: "2.0"
2
3 stories:
4
5 - story: happy path
6   steps:
7     - intent: greetings
8     - action: action_utter_greetings
9
10 - story: happy sorry
11   steps:
12     - intent: sorry
13     - action: utter_never_mind
```

Информация о правилах.

- Правила — жёсткие сценарии ответов.



```
rules.yml 611 bytes
1 version: '2.0'
2 rules:
3   - rule: Rule to map `ask_chitchat` intent to `respond_ask_chitchat` (automatic conversion)
4     steps:
5       - intent: ask_chitchat
6       - action: utter_ask_chitchat
7
8   - rule: Rule to handle messages with low NLU confidence (automated conversion from 'FallbackPolicy')
9     steps:
10      - intent: nlu_fallback
11      - action: action_default_fallback
12
13   - rule: Rule to map `greetings` to `action_utter_greetings`
14     steps:
15       - intent: greetings
16       - action: action_utter_greetings
17     wait_for_user_input: false
18
19   - rule: insult
20     steps:
21       - intent: insult
22       - action: utter_insult
23
```


Преобразователь уатл-файлов.

- Мы будем использовать для построения чат-бота фреймворк RASA, open source версию. Он широко используется в индустрии для построения чатботов, заточенных под конкретные задачи.
- RASA принимает на вход уатл-файлы. Часть файлов будет создана нами, часть будет меняться преобразователем в соответствии с входными данными пользователя.

RASA NLU.

- Часть RASA, отвечающая за понимание текста.
- Внутри RASA мы будем использовать spaCy. Эта открытая библиотека широко используется в индустрии для решения разных задач NLP. Там поддерживается русский язык.
- В RASA есть целый спектр моделей, можно выбрать релевантные в yaml-файлах.

RASA Core.

- Это часть RASA, отвечающая за управление диалогом и за генерацию ответов.
- Надо предусмотреть возможность переспрашивать пользователя.
- Надо предусмотреть класс других интенгов для того, чтобы реагировать на нерелевантные ответы.
- В идеале было бы хорошо иметь два пути для довольного и недовольного пользователей.

Преобразование текста в речь.

- Речь идёт об использовании предобученной модели.
- Вот один из примеров (использование ESPnet2):
https://zenodo.org/record/5149485#.Y_YeTdJBzmg
- Можно рассмотреть варианты склейки маленьких звуковых файлов в соответствии с диалогом.

Выходные звуковые файлы.

- Предполагается сгенерить небольшое количество звуковых файлов, протестировать их в ручную.
- Предполагается в будущем регулярно проверять выходные файлы и улучшать модель в соответствии с реакциями пользователей.