

Rasa-Dokumentation

Vergleich RASA mit Alexa:

Intent-Erkennung in RASA: Rasa überprüft die eingegebenen Eingaben mit den vorhandenen Examples der jeweiligen Intents und bewertet die Ergebnisse nach Passgenauigkeit. Die jeweiligen Intents werden dabei mit einer id, dem namen und der „confidence“, also der Wahrscheinlichkeit, dass dieser Intent gemeint war ausgegeben. Der wahrscheinlichste Intent wird hierbei am weitesten oben abgebildet und dementsprechend verwendet (Siehe snippet)

```
Next message:
Inzidenz Landkreis Worms
{
  "text": "Inzidenz Landkreis Worms",
  "intent": {
    "id": -3219611831882340642,
    "name": "LandkreisInzidenzIntent",
    "confidence": 0.9842750430107117
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": -3219611831882340642,
      "name": "LandkreisInzidenzIntent",
      "confidence": 0.9842750430107117
    }
  ],
}
```

Intent-Erkennung in Alexa: Alexa überprüft die bereits definierten Utterances. Wenn der eingegebene Satz / die eingegebene Aufforderung auf einen der Intents zutrifft, so wird dieser ausgeführt. Sollte es zu Abweichungen / Nicht übereintreffungen kommen, wird der benötigte Intent nicht ausgeführt.

Fazit: Rasa bietet dem Nutzer eine bessere Intent-Erkennung, da die Utterances nicht zu 100% übereinstimmen müssen um dennoch in den Intent zu springen. Vorrausgesetzt der Benutzer / der Skill hat nicht zuvor alle möglichen Variablen als Utterances gespeichert, ist es möglich dass der Benutzer des Skills bei Alexa eine Formulierung wählen könnte, die der Skill nicht kennt. Dies ist eher umgehbar in Rasa.

Utterance-Erkennung:

Die Utterance-Erkennung bei Rasa funktioniert solange eine ähnliche Formulierung bereits als Utterance in dem jeweiligen Intent vorhanden ist. Sollten hierbei jedoch die Keywords (z.B. bei Landkreis-Intents das Wort „Landkreis“) fehlen, so ist eine Abweichung des Intent möglich.

Wenn der Nutzer also unter anderem das Keyword „Landkreis“ und „Inzidenz“ eingibt, landet man im „LandkreisInzidenzIntent“:

```
}
Next message:
inzidenz im Landkreis Worms
{
  "text": "inzidenz im Landkreis Worms",
  "intent": {
    "id": 2080917895408328301,
    "name": "LandkreisInzidenzIntent",
    "confidence": 0.989575207233429
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": 2080917895408328301,
      "name": "LandkreisInzidenzIntent",
      "confidence": 0.989575207233429
    },
    {
      "id": 5390951073538412984,
      "name": "DeutschlandInzidenzIntent",
      "confidence": 0.004956534598022699
    }
  ],
}
```

Hierbei reicht es sogar, nur die beiden Keywords einzugeben:

```
Next message:
Inzidenz Landkreis
{
  "text": "Inzidenz Landkreis",
  "intent": {
    "id": 2080917895408328301,
    "name": "LandkreisInzidenzIntent",
    "confidence": 0.9605277180671692
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": 2080917895408328301,
      "name": "LandkreisInzidenzIntent",
      "confidence": 0.9605277180671692
    },
    {
      "id": 5390951073538412984,
      "name": "DeutschlandInzidenzIntent",
      "confidence": 0.0304340161383152
    }
  ],
}
```

Sollte wiederum nur eines der Keywords eingegeben werden, ist es schnell möglich im falschen Intent zu landen. So landet man bei „Inzidenz Worms“ in dem DeutschlandInzidenzIntent, obwohl eindeutig nach Worms gefragt wurde

```
Next message:
inzidenz worms
{
  "text": "inzidenz worms",
  "intent": {
    "id": 5390951073538412984,
    "name": "DeutschlandInzidenzIntent",
    "confidence": 0.9991226196289062
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": 5390951073538412984,
      "name": "DeutschlandInzidenzIntent",
      "confidence": 0.9991226196289062
    },
    {
      "id": 2080917895408328301,
      "name": "LandkreisInzidenzIntent",
      "confidence": 0.00018056468979921192
    },
  ],
}
```

Dies liegt daran, dass bei DeutschlandInzidenzIntent bereits ausreicht, wenn nur „Inzidenz“ als Keyword eingegeben wurde über die Examples

```
- intent: DeutschlandInzidenzIntent
examples: |
  - wie sieht die infektionslage aus
  - wie sieht das infektionsgeschehen aus
  - welchen inzidenzwert haben wir
  - inzidenzwert
  - wie hoch ist die sieben tage inzidenz
  - sieben tage inzidenz deutschland
  - sieben tage inzidenz
  - wie hoch ist unsere inzidenz
  - wie hoch ist die inzidenz
  - inzidenz
  - Wie hoch ist die Inzidenz in Deutschland
  - Was für eine Inzidenz haben wir in Deutschland
  - Wie lautet der Wert für die Inzidenz in Deutschland
  - Was beträgt der Inzidenz Wert in Deutschland
  - Deutschland Inzidenz
  - Inzidenz Deutschland
  - Was ist die Inzidenz in Deutschland
  - Was haben wir für eine Inzidenz in Deutschland
  - Was für einen Wert hat die Inzidenz in Deutschland
  - Wie hoch ist in Deutschland die Inzidenz
```

Experiment:

Neuer Intent: „InzidenzVergleichIntent“

Dieser Intent benötigt zwei Landkreise -> Landkreis (Landkreis) Landkreis (Landkreis)

```
Next message:
landkreis landkreis
{
  "text": "landkreis landkreis",
  "intent": {
    "id": 1480221285900846212,
    "name": "LandkreisInzidenzIntent",
    "confidence": 0.3314915597438812
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": 1480221285900846212,
      "name": "LandkreisInzidenzIntent",
      "confidence": 0.3314915597438812
    },
    {
      "id": -936305207244450605,
      "name": "LandkreisNeuGeneseneIntent",
      "confidence": 0.20496170222759247
    },
    {
      "id": -6568468075998926874,
      "name": "LandkreisNeuinfektionenIntent",
      "confidence": 0.17523130774497986
    },
    {
      "id": 4212774624262977500,
      "name": "LandkreisNeueTodesfaelleIntent",
      "confidence": 0.15912549197673798
    },
    {
      "id": -808641800000785572,
      "name": "InzidenzVergleichIntent",
      "confidence": 0.05430031195282936
    }
  ],
}
```

LandkreisInzidenzIntent beinhaltet „Inzidenz“ und „Landkreis“ als Keywords -> abgefangen

➔ Keyword „Vergleich“ hinzufügen -> „Vergleich Inzidenz Landkreis Landkreis“

```
Next message:
Vergleich Inzidenz Landkreis Landkreis
{
  "text": "Vergleich Inzidenz Landkreis Landkreis",
  "intent": {
    "id": -874524456063881116,
    "name": "InzidenzVergleichIntent",
    "confidence": 0.6664578914642334
  },
  "entities": [],
  "intent_ranking": [
    {
      "id": -874524456063881116,
      "name": "InzidenzVergleichIntent",
      "confidence": 0.6664578914642334
    },
    {
      "id": 678619672317008342,
      "name": "DeutschlandReproduktionszahlIntent",
      "confidence": 0.13023625314235687
    }
  ],
}
```

⇒ **Fazit:**

RASA arbeitet wie bereits gedacht stark mit Keywords. Wenn es diese Schlüsselwörter in der Nachricht wiederfindet, gleicht sie diese mit in utterance vorhandenen Wörtern ab. Hierbei werden natürlich bei kürzeren Utterances eher diese ausgewählt, weswegen es bei einem Vergleich zum Beispiel einem weiteren Keyword benötigt, und demnach die Formulierungsvariationen deutlich geringer sind.

Je spezifischer die Anforderung bzw. der Intent, desto relevanter ist die gewählte Wortwahl.

```
- intent: InzidenzVergleichIntent
  examples: |
    - Vergleich Inzidenz Landkreis (landkreis) und Landkreis (landkreis)
    - Vergleich Inzidenz Landkreis Landkreis
```

Letztenendes bleibt festzuhalten, dass bei Intents, welche die gleichen Formulierungen verwenden, einzelne Wörter sehr relevant sein können wenn es darum geht, welcher Intent ausgewählt wird.

Daraus ergeben sich folgende Handlungsempfehlungen:

- Die Utterances dürfen niemals identisch sein
- Es gibt wenig Formulierungsspielraum bei verschiedenen Intents wenn die Utterances ähnlich formuliert worden sind
- Auf für verschiedene für die Intents relevanten Keywords ist zu achten (z.B. Vergleich beim Vergleichs-Intent das Keyword „Vergleich“)
- Je mehr „Keywords“ desto genauer die Erfassung