ORACLE

The Challenge of Entity Extraction

Program agenda

1   "Real World" entity extraction challenges

2   What to do about it?

Program agenda

1. **"Real World" entity extraction challenges**
2. What to do about it?

"Real World" entity extraction challenges

> Humans have a natural ability to understand a conversation and extract all relevant information from it.

How can a conversational AI platform match such capabilities?

How to address ambiguity?

How to apply any implicit rules we may need?

And many other rules that are natural to human conversation!

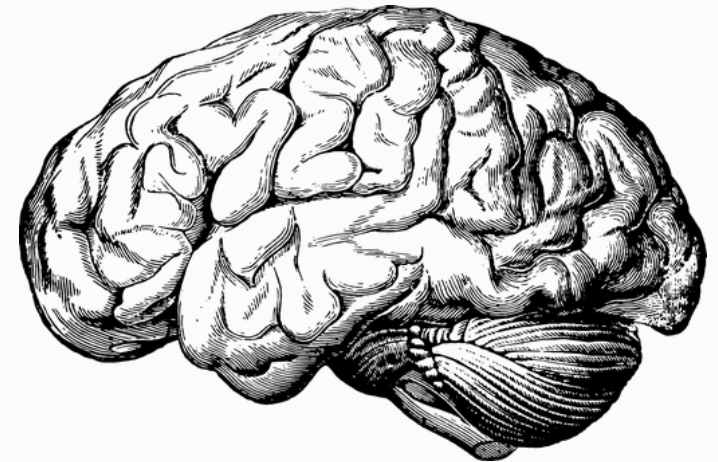# "Real World" entity extraction challenges

Ambiguity

- A message is not clear in what it means
  - E.g. "I'd like to arrange dinner for 6 tonight"
    - Is it 6 people of 6 O'clock?
  - E.g. "Book vacations for June"
    - June the month or June the person?
  - E.g. "It was me and 2 others"
    - For a human it's obviously 3 people, but is it that obvious for a machine?

Information provided unexpectedly

- E.g. "Make it large please"
  - When being prompted for the pizza delivery time
- E.g. "Actually, I want the shoes in blue instead"
  - When being asked for shipping address

NO WAY TO BEAT THIS?

Image courtesy of pixabay.com

"Real World" entity extraction challenges

When context matters
- Assistant: "The Daly City store is open from 8 – 5"
- User: "Can I pick my meds there?"
 - "There" is where? Information from the previous message is needed to establish context

Order in which information is provided
- What is the source, what is the target?
  - "I want to change 100 USD to EUR"
  - "For 100 EUR how many USD do I need?

Date ambiguity
- For humans it's easy to understand if the Date is past or future
  - "I want to book a flight for Monday"
  - "I want to create an expense for a meal I had Monday"

NO WAY TO BEAT THIS?
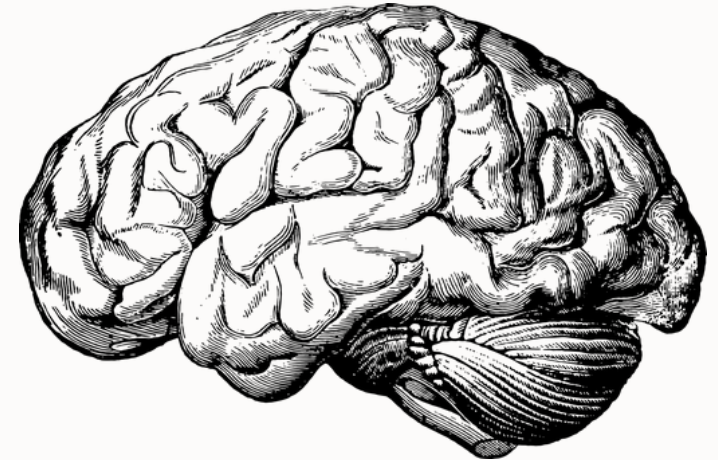
"Real World" entity extraction challenges

Usually, multiple entities are used when resolving an intent
- Each entity captures a single piece of information

Setting of default values
- In some cases, we need to set up a value even if the user does not provide it
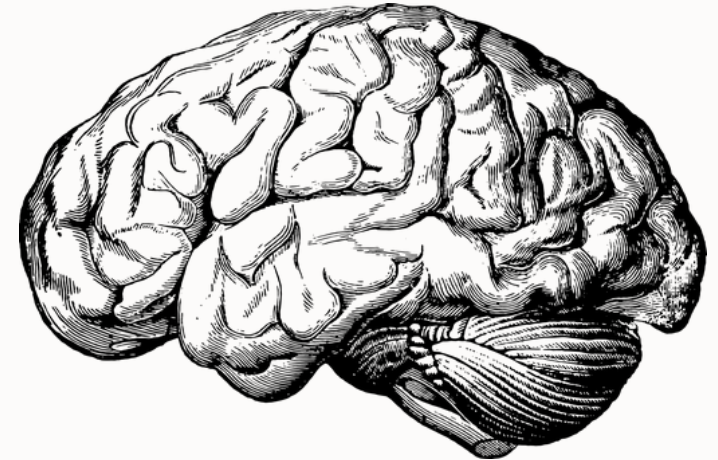
Conditional validation or prompts
- Prompt for expense receipt if expense is > 25$
- Suppress credit card prompt for payment type "bank transfer"

Business dependency
- E.g "Backend stock check"

NO WAY TO BEAT THIS?

Image courtesy of pixabay.com

"Real World" entity extraction challenges

Dynamic information

- We can have value list entity values that change often

When user locale matters

- 11/7 as a Date is November 7 in the USA but 11 July in other countries
- "It cost 35", what is the currency in this value?
- Numbers can be represented with commas or dots
  - 1.000.000 vs 1,000.000

How to deal with an infinite list of values?

- Not all information is known when a model get trained
  - Movie titles, merchant names

NO WAY TO BEAT THIS?


Image courtesy of pixabay.com
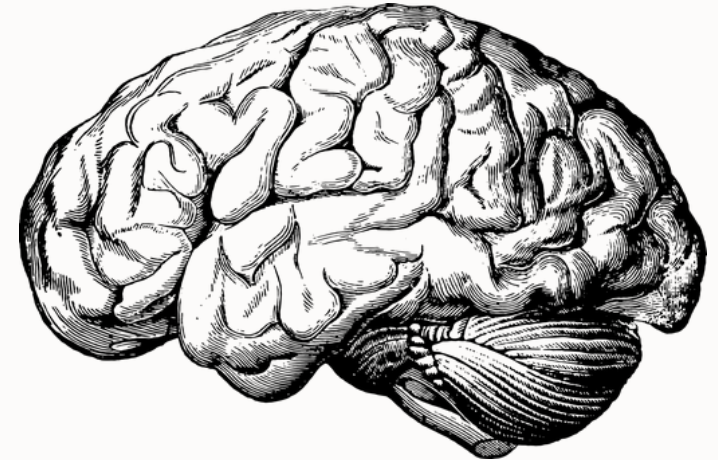
"Real World" entity extraction challenges

How to validate entities beyond their type

- Book a flight to SFO for June 12 2022
  - Dates in the past don't work for booking flights, though from an entity perspective it's a valid date
- I spent 1000 USD for a flight to SFO on June 12 2024
  - Again, this is a valid entity value, albeit incorrect from an expenses perspective if today is 2022
- I spent 29$ for breakfast
  - What of your expense is "per diem"?

NO WAY TO BEAT THIS?


Image courtesy of pixabay.com
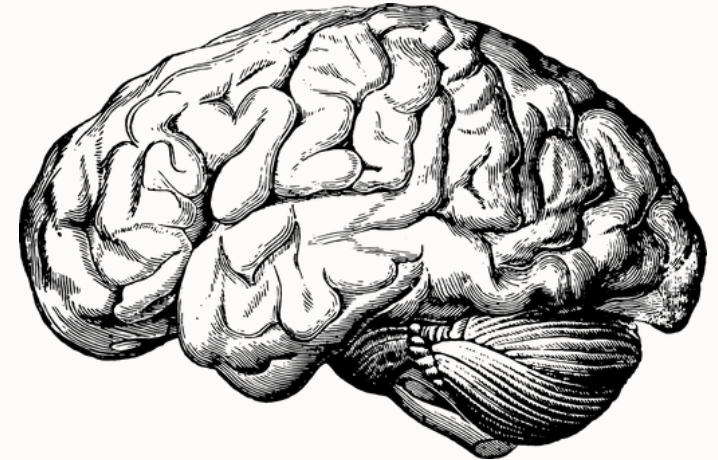
Program agenda

1. "Real World" entity extraction challenges

2. **What to do about it?**

" All these challenges can be overcome using entity features of the Oracle Digital Assistant platform and, of course, conversation design
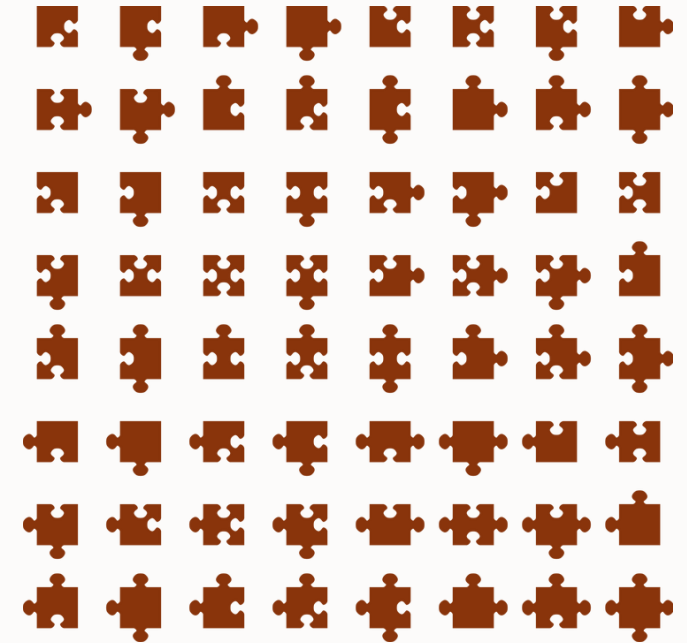
What to do about it?

- Most of these challenges can be addressed with Platform features, out of the box, in a non-programmatic way
  - No need to reinvent the wheel

- We can use system entities, like *Date* that allows us to specify future/nearest/past and format

- Composite bags allow to gain platform benefits such as understanding out of order during entity resolution, the ability to perform validation, to have multiple and conditional prompts, to resolve multiple entities at the same time, etc.

Image courtesy of pixabay.com
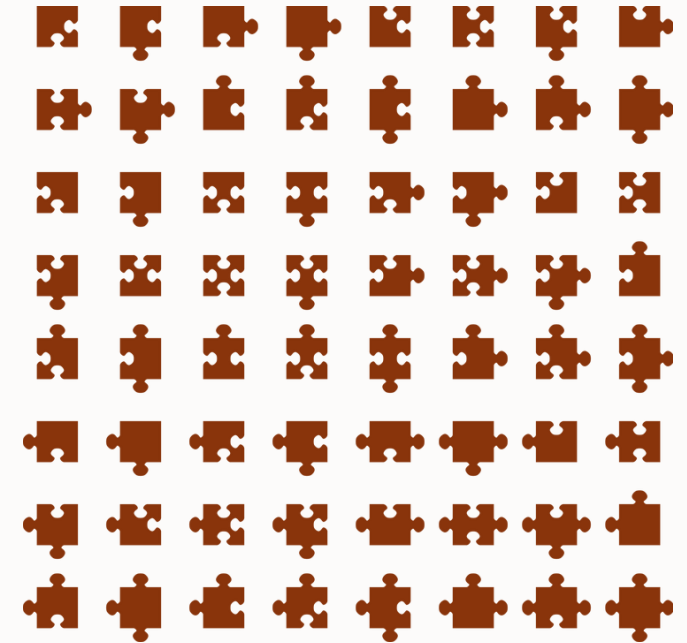
What to do about it?

- For those odd cases where more complex requirements need to be met, we can extend composite bags with a programmatic approach by using entity event handlers

- Sometimes we simply don't know and should use smart strategies to disambiguate
  - Clear wording
  - Good disambiguation messages
  - Guide the user to provide the desired message