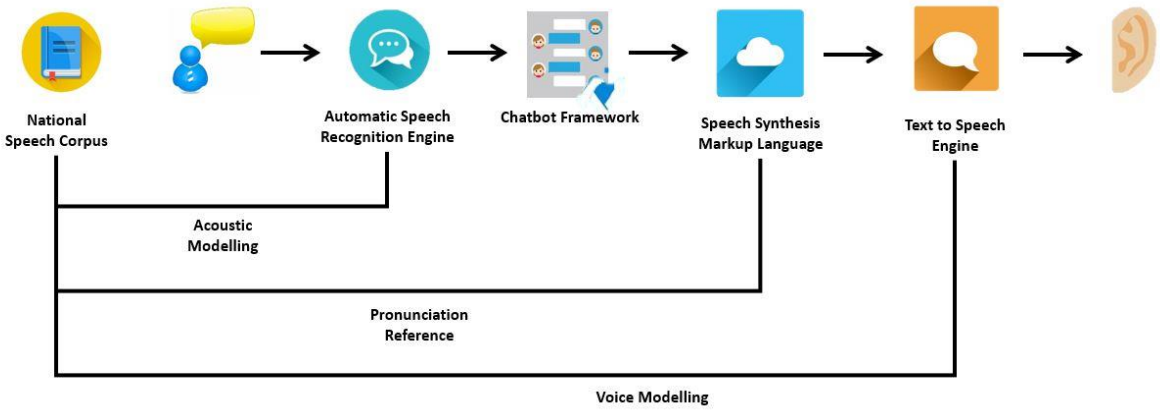
Introduction

The Food Ordering Speech AI Assistant showcases local food being recognized and pronounced by the AI.

This document explains the Methodology of development of the Chatbot.

Rationale

- Users can relate to calling a waiter at a busy restaurant
- Users will have a physical menu which they browse at their leisure
- Users will order from the AI assistant when they are ready
- Showcases local food being recognised and pronounced by the AI



Methodology - Creating a Domain Taxonomy of the Demo

- Methodology for developing NLP Taxonomy
    - Understanding the environment:
        - Typical operation of target business
            - Identify main customer interactions to consider as intents
            - Using 80/20 rule, focus on 20% of the use cases that occurs 80% of the time
    - Understanding the main customer interaction flow:
        - Such as: Greetings, Buying, Confirm and Payment
    - Perform site survey to bridge understanding and real-world scenario:
        - Confirm that the theoretical understanding developed matches the real world
    - Capturing taxonomy and data based on identified intents
        - Such as using existing brochures and menu

- Methodology applied for restaurant environment
    - Understanding the environment
        - Typical operation of a sit-in restaurant
            - Customer places an order ➞ Waiter takes order ➞ Confirm Order
            - Process: Greeting, Order (Remove and Change), Confirm Orders
    - Common utterances from waiter and customers
        - Restaurant taxonomy for Waiters

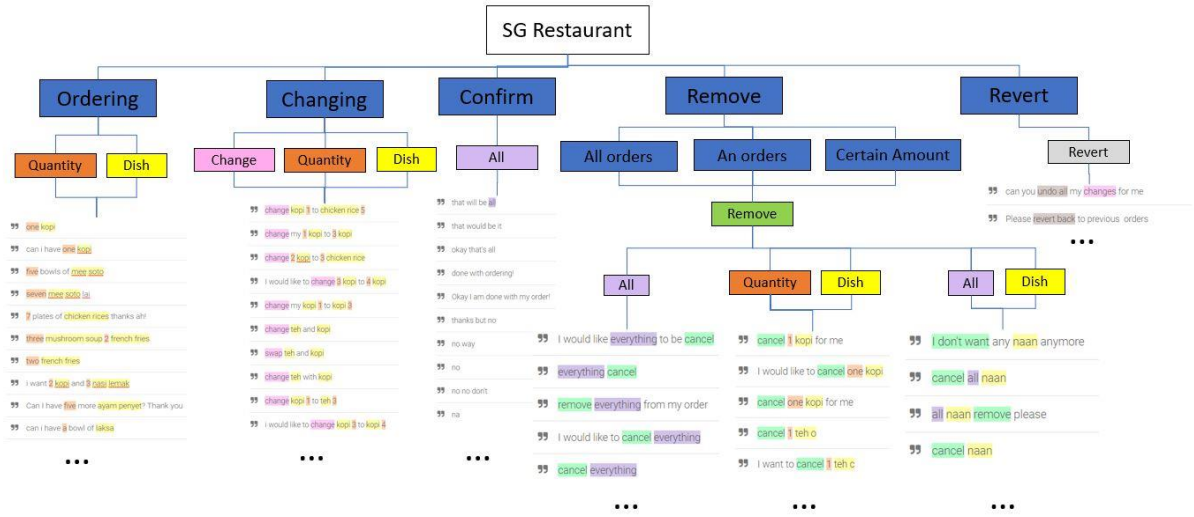
Processes	Utterances	Sentence structure
Welcome	"Hi Welcome to _____. How may I help you?"	[Greetings] [Restaurant Name] [Query next actions]
	"Hi, How may I help you?"	
	"Hi Welcome to _____. What would you like?"	
	"Good Evening! How may I help you?"	
Ordering	"Okay can!"	[Affirmative Reply] [Query next actions]
	"Alright 1 kopi ah!"	
	"Can, is that all?"	
	"Sure, anything more?"	
Removing Orders	"Okay kopi ah"	[Affirmative Reply] [Order name]
	"Alright kopi"	
	"Kopi noted"	
	"Done!"	
Changing Orders	"okay changed your order"	[Affirmative Reply]
	"Alright! Changed already!"	
Confirming Orders	"Okay! 1 kopi, 2 Teh Peng, Total is $4.00"	[Affirmative Reply] [All Dishes] [Total Amount]
	"1 Fillet-O-Fish set meal , Total is $6.50"	

        - Restaurant taxonomy for Customers

Processes	Utterances	Sentence structure
Ordering	"Can I have 1 kopi, please!"	[Quantity] [Order name(s)]
	"1 kopi, 2 teh peng, 3 Teh C"	
	"umm.... I need ..... 1 Fillet-O-Fish set"	
	"Give me 1 Fillet-O-Fish set meal"	
	"1 Kopi lorh"	
	"1 kopi peng thanks"	
	"1 chicken rice ahhh!"	
Removing Orders	"remove 1 chicken rice for me leh"	[Quantity] [Order name]
	"Sorry, remove kopi for me"	
Changing Orders	"Change 1 kopi to 2 teh please"	[Quantity] [Order name] to another [Quantity] [Order name]
	"Sorry order wrongly, change all my kopi to teh"	
Confirming Orders	"okay, that's all"	[Affirmative Reply]
	"yep correct!"	
	"yes"	
	"absolutely"	

- Understanding the food menu
    - A selected few numbers of local food was chosen for this demo to showcase the recognising and pronunciation of local food (e.g. Mee Soto, Soon Kueh)
- Site Survey
    - Conducted at the restaurant to study the interactions between the customer and the waiter
    - Example Location: PS. Café, Morgan Field

Chatbot Logic and Intents



- Ordering
    - Quantity and Dish
        - Example: one kopi, can I have three mee soto
- Changing
    - Change, Quantity and Dish
        - Example: change one kopi to two chicken rice
- Confirm
    - All
        - Example: that would be all, I am done
- Remove
    - All
        - Example: Cancel Everything
    - Quantity and Dish
        - Example: Cancel one Teh O, I want to cancel one Teh C
    - All Dish
        - Example: Cancel all naan
- Revert
    - Example: Undo/revert to previous order

Automatic Speech Recognition (ASR)

- Push-to-Talk as a method to activate the ASR
    - Removes the need for speech detection
        - Voice detection can be unreliable (especially when the detection of end of speech fails)
        - Noisy environment
    - Greater user control
        - Enables users to be in control of when the speech detection starts and stops
    - Light and Robust operation
        - As compared to constantly listening for a keyword

- ASR Language Model
    - Data used is small and concise
    - Domain-specific use case
        - Restaurant/Food Ordering domain
        - Not meant for transcription work
    - Performed admirably
    - Manually fix common sounding words
        - “two two” > “to two”

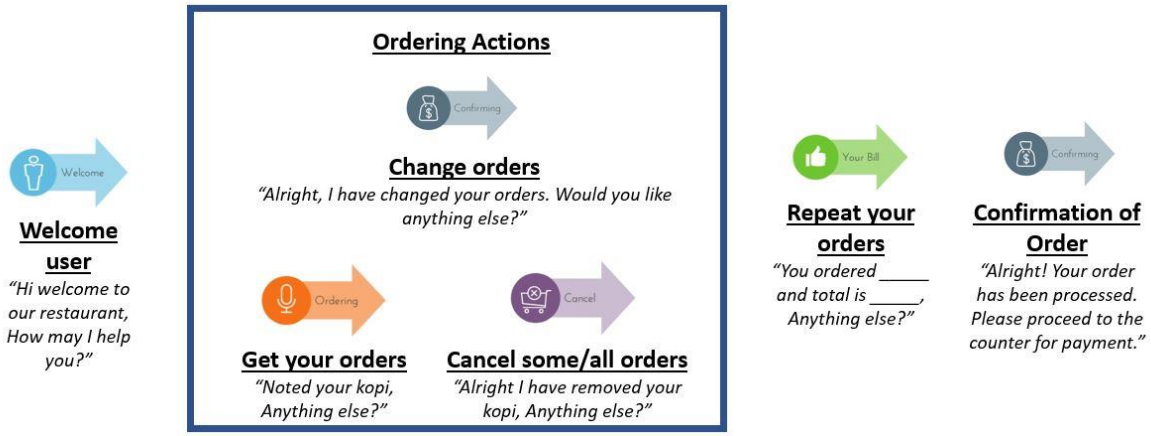
- Language Model Sample

```
A
A BEEF BOLOGNAISE
A BLACK TEA
A BOWL
A CHEESY GARLIC PRAWN PASTA
A CHICKEN CHOP
A CHICKEN CHOP BURGER
A CHICKEN CHOP WITH FRIES
A CHICKEN CURRY NOODLES
A CHICKEN CURRY RICE
A CHICKEN KARAAGE
A CHICKEN KARAAGE WITH RICE
A CHICKEN NUGGET WITH FRIES
A COFFEE
A CRAB SALAD
A CREAMY CHEESE GARLIC TIGER PRAWNS
A EARL GREY TEA
A FRIED POTATO SKINS
A FRIED SQUID ARMS
A GARDEN SALAD
A GRAIN FED RIBEYE
A GRILLED SABA FISH
A GRILLED SQUID
A GYOZA
A HONEY MUSTARD CRISPY CHICKEN WINGS
A HORLICKS
A HOT CHOCOLATE
A JAPANESE POTATO SALAD
A KOPI
A KOPI C
A KOPI C KOSONG
A KOPI C PENG
A KOPI C SIEW DAI
A KOPI KOSONG
A KOPI O
A KOPI O KOSONG
CHICKEN CHOP WITH FRIES
CHICKEN CURRY NOODLES
CHICKEN CURRY RICE
CHICKEN KARAAGE
CHICKEN KARAAGE WITH RICE
CHICKEN NUGGET WITH FRIES
CHICKEN RICE
CLEAR
COFFEE
CONFIRM
COST
COULD YOU PLEASE CANCEL
CRAB SALAD
CREAMY CHEESE GARLIC TIGER PRAWNS
CRISPY FRIED CHICKEN
CUP
CUPS
DISCARD MY CHANGES
DISCARD MY ORDER
DOES
EACH
EARL GREY TEA
EIGHT
EIGHTEEN
EIGHTY
EIGHTY-EIGHT
EIGHTY-FIVE
EIGHTY-FOUR
EIGHTY-NINE
EIGHTY-ONE
EIGHTY-SEVEN
EIGHTY-SIX
EIGHTY-THREE
EIGHTY-TWO
ELEVEN
ELIMINATE
I WOULD LIKE TO SWITCH MY
IS
IS IT
IS IT POSSIBLE TO CANCEL
IS IT POSSIBLE TO CHANGE
IS IT POSSIBLE TO DELETE
IS IT POSSIBLE TO DISCARD
IS IT POSSIBLE TO ELIMINATE
IS IT POSSIBLE TO EXCHANGE
IS IT POSSIBLE TO GET RID
IS IT POSSIBLE TO MODIFY
IS IT POSSIBLE TO REMOVE
IS IT POSSIBLE TO SWAP
IS IT POSSIBLE TO SWITCH
JAPANESE POTATO SALAD
KOPI
KOPI C
KOPI C KOSONG
KOPI C PENG
KOPI C SIEW DAI
KOPI KOSONG
KOPI O
KOPI O KOSONG
KOPI O PENG
KOPI O SIEW DAI
KOPI PENG
KOPI SIEW DAI
KUEH LAPIS
LAKSA
LAMB SHOULDER
LATTE
LOR MEE
MAY I GET
MAY I HAVE
MAY I MODIFY
MAY I ORDER
```

Chatbot Framework – Dialogflow

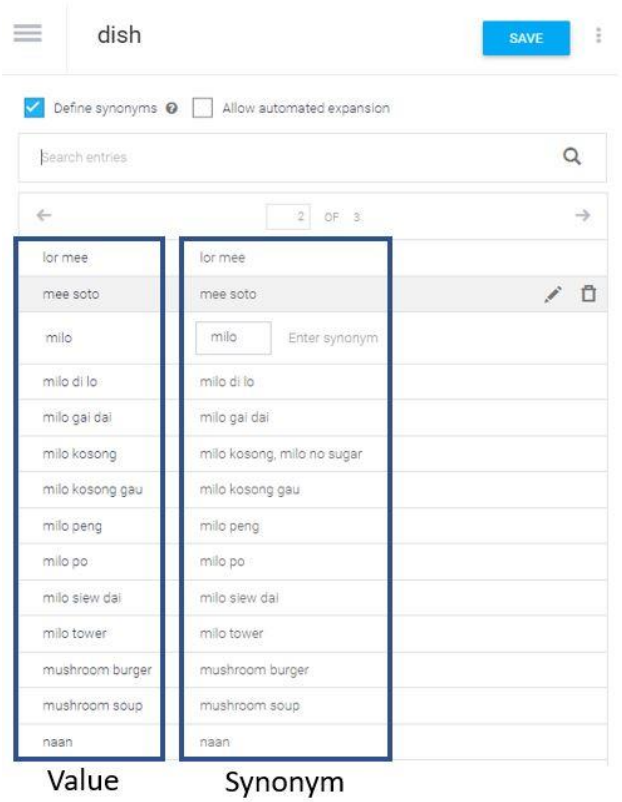
- Intents
  - An intent categorizes an end-user's intention for one conversation turn.
    For each Chabot domain, defining many intents are possible, where combined intents can handle a complete conversation.
    Dialogflow matches the end-user expression (user input) to the best intent in the agent.

  - Intents to Develop (Food Ordering)



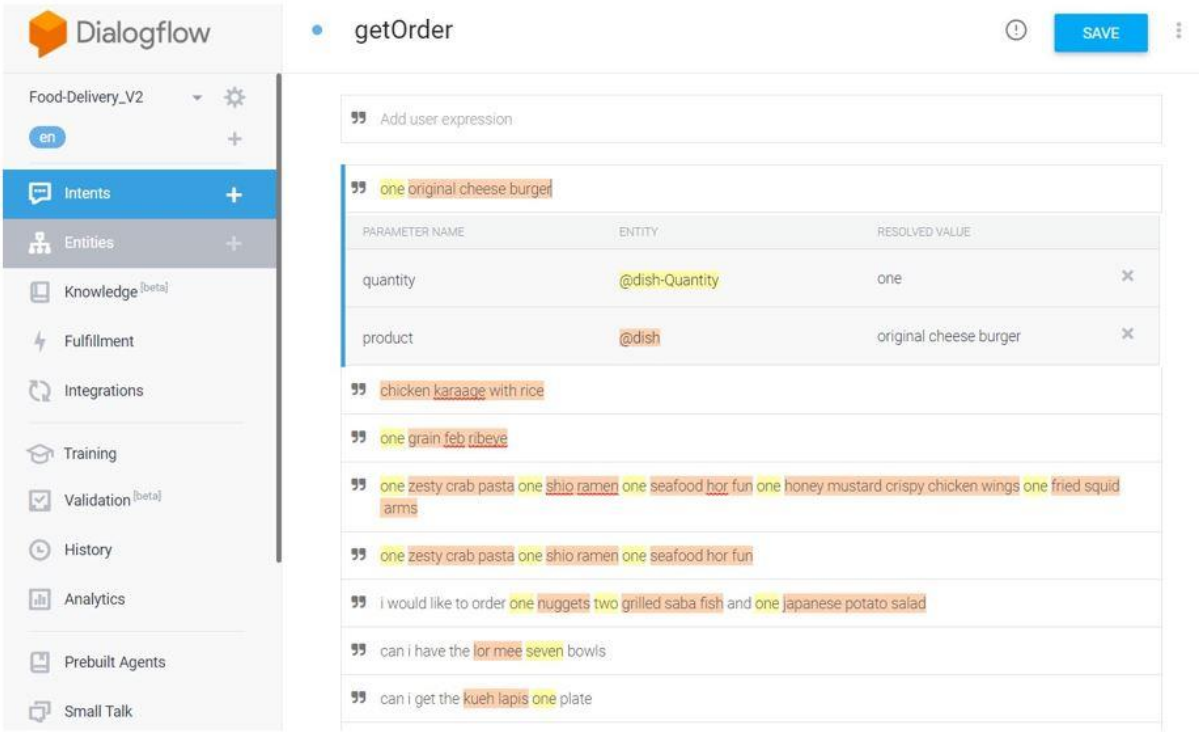
- Entities

  - Dictates exactly how the data from intents are extracted from an end-user expression.
    - Dish Names
    - Supports custom vocabulary
    - Provide Synonyms if needed
      - E.g. "Milo Kosong" can be "Milo without Sugar" or "Milo Kosong"
    - Can use menu codes: A1, A2, B3. etc.



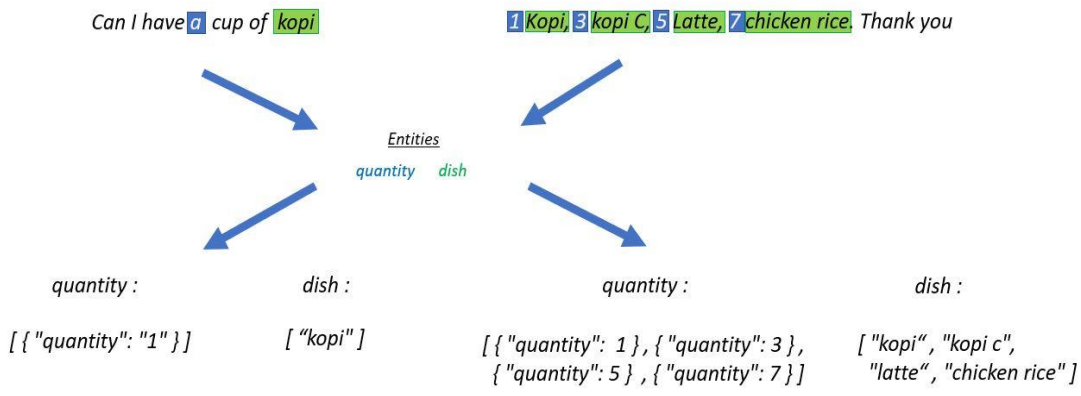
- Training Phrases in Intents
    - ‘Teaches’ Diagflow how to recognise each entity in an intent
        - E.g. One Original Cheese Burger

Parameter Value	Entity	Resolved Value
Quantity	@dish-Quantity	one
Product	@dish	original cheese burger

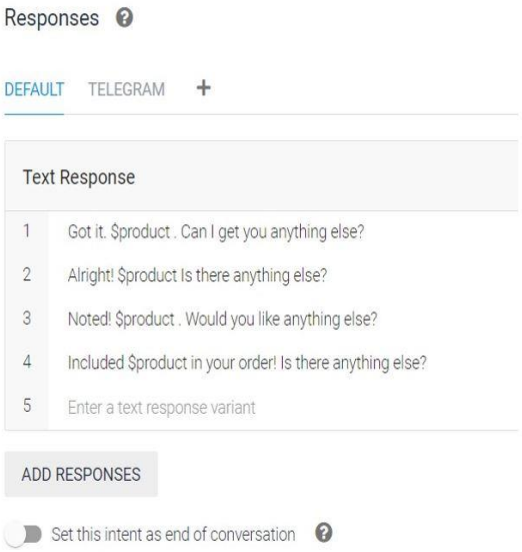


    - Developing Training Phases in Intents
        - Mapping out and tagging the association of different entities
            - E.g. Order Intent, whereby numbers are associated with the quantity entity and food names are associated with the dish entity

Order Intent

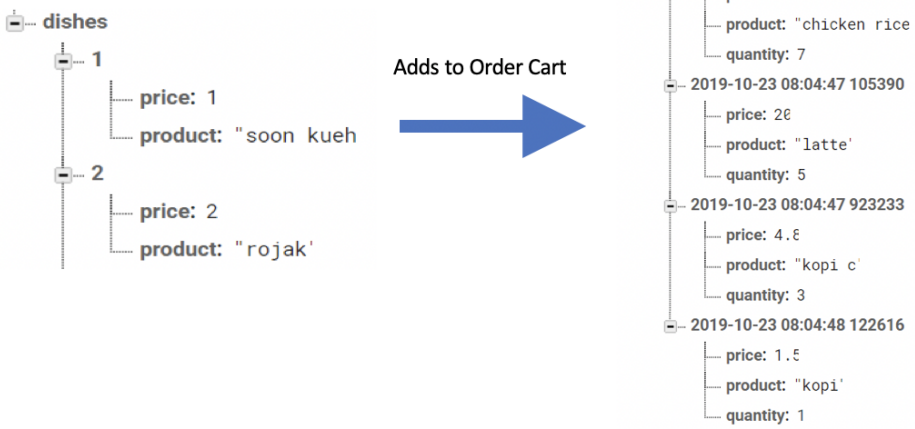


- Responses
  - Programme the output of the Chabot once it receives an intent request



- Ideally to keep the responses short
- Randomise the responses when appropriate
- Design the responses to carefully elicit the user's response
- Creation of a fallback response to cater for any unexpected user's response

- Fulfilment
  - Dialogflow fulfilment handles the parameters, quantity and product
  - In the Food Ordering demo
    - We *Add* or *Remove* orders to the *order_cart*
    - *Query* the database to retrieve the prices under *dishes* to process the order to the cart.



- Machine Learning Settings
    - Users can modify the ‘Match Mode’ which best fits their criteria
    - ‘Machine Learning Classification Threshold’
        - The closer the number is to 1, it becomes more rule-based and prediction by the Machine Learning Algorithm is lesser.