



Guía d' anotació de slots

[versión 3.0]



Descripció de la tasca

Per a poder dur endavant la seva tasca, els xatbots no només necessiten identificar la intenció de l'usuari (*intents*) sinó que també han de reconèixer els segments de text que li aporten la informació concreta sobre la tasca (*slots*) i ser capaços d'identificar de quin tipus d'informació es tracta (*tags*).

“Bon dia, necessito saber els horaris per anar demà al matí de Barcelona a Mataró.”

intent: request_info
slots: info_type: horaris
location: Barcelona
destination: Mataró
date: demà
time: matí

Se us presentaran una sèrie de frases escrites per un usuari, l'*intent* que se'ls ha assignat i la llista de possibles *tags*, segons l'*intent* concret.

En un primer lloc haureu de valorar si l'*intent* assignat us sembla correcte. En cas que l'*intent* assignat no us semblés correcte, ho podeu assenyalar amb la bandera per a revisar-ho més tard.

Si l'*intent* assignat us sembla correcte, busqueu a la frase els segments que aporten les informacions indicades a la llista de *tags* i marqueu-los.

Les frases poden tenir diversos nombres d'*slots*, com també no tenir-ne cap. En alguns casos, pot haver-hi més d'un slot anotat amb el mateix tag.

Exemple: “Els meus fills es diuen Marta i Martí”

name: Marta
name: Martí

Quan l'usuari es corregeix o corregeix al bot, s'anota el nou slot amb el mateix tag que l'slot original.

Exemple: “Voldria un bitllet per anar a Banyoles”

destination: Banyoles
“Ai, no: Volia dir a Mata”
destination: Mata

Si trobeu que caldria marcar algun *slot* amb informació no prevista a la llista de *tags*, podeu indicar-ho.

Llista de possibles tags

Nom	Descripció
accommodation ¹	tipus d'allotjament
address ²	adreça concreta (via i número).

1 Accommodation: Pot ser "casa rural", "hotel", etc.

2 Address: Correspon només a adreces concretes (via i número). Per a la resta d'ubicacions (municipi, país, nom d'un negoci, barri, etc.) usem **location**. En cas que una adreça contingui el nom del municipi tot seguit (per exemple, “carrer dels

age	edat
asset ³	en el domini assegurances, l'item que es vol assegurar.
birth_date ⁴	data de naixement
booking_num	número de reserva
brand	marca del producte
card_exp	data de caducitat de la targeta de crèdit
card_num	número targeta de crèdit
cvv	número de control de la targeta de crèdit
date ⁵	dia del que es parla
destination ⁶	destinació per la què es pregunta
dni_num	DNI de l'usuari
driver_card_num	número del carnet de conduir
duration ⁷	durada d'una activitat

Blanquers, 24, Banyoles”), l'incloem dins el mateix slot, anotat com a **address**. Si el municipi no es menciona tot seguit (per exemple “carrer dels Blanquers, 24. I el municipi és Banyoles”), el carrer i el número s'anoten com a **address**, i el municipi com a **location**.

3 date: Pot ser que sigui el dia del que es demana informació, dia que ha passat alguna cosa, etc. Si parlem d'un període de temps, farem servir **from_date** i **to_date**. En el cas d'una data d'aniversari, farem servir el tag **birth_date**. Cal anotar la màxima informació possible (per exemple, dia, mes i any).

Exemple: Vindrem el 16 de juliol.

4 destination: S'utilitza només per parlar de trajectes. En el cas que lloguem un vehicle i el retornem en un punt diferent d'on l'hem recollit, també l'anotarem com a **destination**. En la resta de casos, fem servir **location**.

5 health_condition: Anotem com a **health_condition** les condicions de salut que es comenten durant la conversa i que poden ser rellevants, però que no són el motiu principal de la conversa. Poden ser més o menys problemàtiques, com ara presència de pigues, pell molt blanca, alçada i pes, tenir la regla, etc.

6 id_num: S'utilitza només quan no hi ha un identificador més concret (p.ex. **dni_num**, **booking_num**, etc.). S'utilitza també per a NIEs, números de pòlissa i números de matrícula dels vehicles. En cas que s'especifiqui en algun lloc de quin tipus d'identificador es tracta, aquesta informació s'anotarà com a **info_type**.

Exemple: El número de matrícula és 4563BCE

info_type: matrícula

id_num: 4563BCE

7 incidence: Per informar de problemes, en general. A complaint, serà el motiu de la queixa, a report_incidence, el problema que hi ha hagut, etc. En el domini de les assegurances, és el problema que hi hagi hagut (per exemple “m'han robat la cartera”). S'engloben coses com haver perdut un objecte al tren, no haver trobat l'habitació com s'esperava, haver perdut un vol, haver rebut un producte en mal estat, etc.

Els problemes de salut, en el context de la clínica, es marquen com a **health_issue**.

email	email
flight_num	número de vol
from_date ⁸	primer dia del període del que es parla
gender	gènere de l'usuari
health_condition ⁹	característiques que poden influir en l'estat de salut però no són el motiu de la interacció amb el bot
health_issue ¹⁰	qüestió de salut que motiva la conversa
iban_num	número iban
id_num ¹¹	qualsevol identificador no previst en aquesta llista
incidence ¹²	incidència
info ¹³	informació rellevant no especificada en altres tags. Sovint és una

En canvi, en els altres contextos, sí que es marquen com a **incidence**.

8 **info**: Exemple: “El participant és major o menor d’edat? –Menor d’edat”.

Veure també **info_type**.

9 **location**: Tant pot referir-se a un país, com un poble, un barri o un carrer. També pot referir-se a la seu d’un servei (per exemple “La clínica del carrer urgell”), el nom d’un establiment (“Estan fent una festa al Bar Celó”) o a l’expressió “a casa” en contextos com “m’han entrat a robar a casa”.

Per a adreces concretes (via i número) fem servir el tag **address**.

10 **service**: Aquí ens referim a servei en el sentit més ample possible. En transports, els abonaments es consideren **service**. A la clínica, les especialitats (dermatologia, dentista, etc.) i les proves mèdiques, també. A allotjament, les habitacions i altres allotjaments, també. A assegurances, és el tipus d’assegurança (per exemple, “assegurança a tot risc”). A lloguer de vehicles és el tipus de vehicle (ex: “necessito una furgoneta per fer un trasllat”). En cas que hi hagi una quantitat, s’anotaran junt amb el servei (ex: “voldria dues habitacions individuals i una de doble”).

11 **person**: es refereix sempre a persones concretes. Pot ser un nom (la Joana Raspall), un cognom (el Dr. Baular), un càrrec (l’editora de la revista) o una relació (el meu fill). Les persones indeterminades (“un agent”) no s’anoten.

12 **product**: Anotarem cada un dels productes per separat.

En el domini del menjar a domicili, anotem la quantitat amb el producte.

Ex. “Voldria una pizza de tonyina i pinya, una barbacoa i dues cerveses”

product: “una pizza de tonyina i pinya”

product: “una barbacoa”

product: “dues cerveses”.

En altres dominis, no cal anotar la quantitat.

Ex. “Voldria un iphone”

product: “iphone”

13 **options**: En el domini de lloguer de vehicles marcarem com a **options**: cadireta per a nens, tipus de tapisseria, GPS, assegurança, etc.

En el domini d’allotjament, les característiques de l’habitació (amb balcó, llits twin, etc.) o de la reserva (mitja pensió, arribar passada una hora, etc.)

	resposta o la informació a la que es refereix info_type
info_type ¹⁴	en preguntes, és el tipus d'informació que es pregunta. En altres frases, qualsevol tipus d'informació no previst a la llista de tags
item ¹⁵	a how_to_fulfill, la documentació que es vol presentar
location ¹⁶	Lloc: Barri, zona, municipi, seu d'un negoci, etc.
method	mètode de pagament
model ¹⁷	model del producte
modification ¹⁸	tipus de canvi que es vol fer
motive ¹⁹	raó d'una baixa, devolució, etc.

14 **route**: Qualsevol identificador de línia: "línia 5", "línia blava", "la r3", "la línia de l'Hospitalet a la Tor de Querol".

15 **health_issue**: Problema de salut que motiva la conversa. Es pot confondre fàcilment amb **health_condition**.

Ex: "Bon dia, necessito una decepta pel xarop per l'anèmia. He d'anar especialment en compte perquè estic embarassada" -> la qüestió que motiva la interacció és l'anèmia, i l'embaràs és una qüestió secundària.

Health_issue: "anèmia"

Health_condition: "embarassada"

Ex: "Bon dia, he de fer-me l'ecografia dels 5 mesos d'embaràs. Estic preocupada perquè tinc anèmia" -> en aquest cas, en canvi, la qüestió que motiva la interacció és l'embaràs i l'anèmia és una informació secundària.

Health_issue: "embaràs"

Health_condition: "anèmia"

Ex: "M'he trobat malament tota la nit i he perdut l'avió" -> poden aparèixer també fora de l'àmbit mèdic.

Issue: "he perdut l'avió"

Health_issue: "trobat malament tota la nit"

16 **modification**: per exemple: "voldria canviar el nom del titular del meu compte", "voldria afegir el meu fill a l'assegurança del cotxe", etc.

17 **people_num**: anotem només el número ("voldria reservar taula per a tres persones"), a no ser que siguin tipus de persones diferents ("Serem dos adults i un infant")

18 **quantity**: anotem només el número. Pot referir-se també a diners (per exemple, "uns 300 euros") Recordeu que per a número de persones hi ha l'etiqueta **people_num**, i que en els dominis de menjar a domicili i reserva d'allotjament les quantitats s'anoten juntament amb el producte.

19 **info_type**: Vigilar no confondre **info** amb **info_type**. **info_type** es refereix al tipus d'informació, i **info**, a la informació concreta. **info_type** sovint és el tipus d'informació pel que es pregunta.

Ex. 1: "quina és l'edat mínima per demanar el carnet jove?" -> **info_type**

Ex. 2: "El meu fill és menor d'edat" -> info_type: "edat", info: "menor".

Ex. 3: "Fa tres anys que tinc el carnet" -> years_num: "tres", info_type: "tinc el carnet"

name	nom de l'usuari
options ²⁰	opcions no previstes per altres tags
order_num	número de comanda
person ²¹	persona concreta
people_num ²²	número de persones
phone_num	número de telèfon
product ²³	producte del que es parla
quantity ²⁴	quantitat
route ²⁵	línia de transport públic
service ²⁶	servei del que s'està parlant
surname	cognom de l'usuari
surname2	2n cognom de l'usuari
time	hora
to_date ²⁷	últim dia del període pel que es pregunta
via ²⁸	mitjà pel qual es transmeten o envien comunicacions
years_num	nombre d'anys (antiguitat, no edat)
zip_code	Codi postal

20 **motive**: S'anota només quan és rellevant. Per exemple "vull tornar la impressora perquè no funciona". En altres casos, com ara "vull un bitllet de tren per anar a Girona a la comunió de la meva neboda" no cal anotar-lo perquè no és rellevant per a la tasca.

21 **asset**: Pot referir-se a un bé material, a un animal, etc.

22 **model**: En cas que es mencioni la marca i el model tot seguit, s'anotarà tot l'slot com a model.

Ex: Vull un Seat Leon.

23 **from_date**: Per la delimitació de l'slot, seguir les indicacions que es donen a **date**.

24 **to_date**: Per la delimitació de l'slot, seguir les indicacions que es donen a **date**.

25 **birth_date**: Per la delimitació de l'slot, seguir les indicacions que es donen a **date**.

26 **item**: Document que s'ha de presentar, o tràmit del que es demana informació:

Ex. 1: "Què he de fer per demanar una beca?"

Ex. 2: "Com us puc fer arribar la nòmina?"

27 **via**: Pot ser per correu postal, per internet, etc.

Ex: Ho puc fer pel web?

28 **duration**: Per exemple: "Arribem el 16 de juliol i ens hi estarem tot el mes"

Exemple2: "Voldria llogar un cotxe per tres dies"