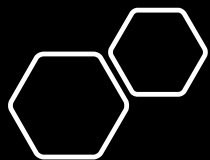


## Padrão para telefone (41) - 778778877

```
from spacy.matcher import Matcher
#inicializa
matcher = Matcher(nlp.vocab)
# Adiciona padrão para telefone (41) - 778778877
padrao = [{"ORTH": "(", "SHAPE": "dd"}, {"ORTH": ")"}, {"ORTH": "-", "OP": "?"}, {"IS_DIGIT": True}]
matcher.add("Telefone", [padrao])

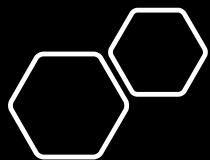
documento5 = nlp("Você pode ligar para (51) - 9964656570 ou (11) 12344988 ")

matches = matcher(documento5)
for match_id, start, end in matches:
    print(documento5[start:end])
```



# Atributos de Tokens

ATRIBUTO	DESCRIÇÃO
ORTH, TEXT	O texto literal exato de um token. Texto
LOWER	Forma minúscula do token text. Texto
LENGTH	O comprimento do token text. Inteiro
IS_ALPHA, IS_ASCII, IS_DIGIT, IS_LOWER, IS_UPPER, IS_TITLE, IS_PUNCT, IS_SPACE, IS_STOP	O texto do token consiste em caracteres alfabéticos, caracteres ASCII, dígitos etc. Booleano.
IS_SENT_START	O token é o início da frase. Booleano
LIKE_NUM, LIKE_URL, LIKE_EMAIL	O texto do token se assemelha a um número, URL, email. Booleano
SPACY	O token tem um espaço à direita.bool
POS, TAG, MORPH, DEP, LEMMA, SHAPE	Usar anotações linguísticas ou formatos, por exemplo, encontrar adjetivos.
ENT_TYPE	Entidades Nomeadas
–	Propriedades em atributos de extensão personalizados.
OP	Operador ou quantificador (!,?,+,* )



# Padrões estendidos

ATRIBUTO	DESCRIÇÃO
IN	O valor do atributo é membro de uma lista.
NOT_IN	O valor do atributo não é membro de uma lista.
IS_SUBSET	O valor do atributo é um subconjunto de uma lista.
IS_SUPERSET	O valor do atributo é um superconjunto de uma lista.
INTERSECTS	O valor do atributo tem uma interseção não vazia com uma lista.
==, >=, <=, >, <	O valor do atributo é igual, maior ou igual, menor ou igual, maior ou menor.

# Operadores

OP	DESCRIPTION
!	Negate the pattern, by requiring it to match exactly 0 times.
?	Make the pattern optional, by allowing it to match 0 or 1 times.
+	Require the pattern to match 1 or more times.
*	Allow the pattern to match zero or more times.