

## 詞的分類

=====

- \* 實詞：名詞、動詞、形容詞、狀態詞、區別詞、數詞、量詞、代詞
- \* 虛詞：副詞、介詞、連詞、助詞、擬聲詞、嘆詞。

## ICTPOS3.0詞性標記集

=====

### n 名詞

#### nr 人名

- nr1 漢語姓氏
- nr2 漢語名字
- nrj 日語人名
- nrf 音譯人名

#### ns 地名

##### nsf 音譯地名

- nt 機構團體名
- nz 其它專名
- nl 名詞性慣用語
- ng 名詞性語素

### t 時間詞

#### tg 時間詞性語素

### s 處所詞

### f 方位詞

### v 動詞

- vd 副動詞
- vn 名動詞
- vshi 動詞“是”
- vyou 動詞“有”
- vf 趨向動詞
- vx 形式動詞
- vi 不及物動詞（內動詞）
- vl 動詞性慣用語
- vg 動詞性語素

### a 形容詞

- ad 副形詞
- an 名形詞
- ag 形容詞性語素
- al 形容詞性慣用語

### b 區別詞

bl 區別詞性慣用語  
 z 狀態詞  
 r 代詞  
     rr 人稱代詞  
     rz 指示代詞  
         rzt 時間指示代詞  
         rzs 處所指示代詞  
         rzv 謂詞性指示代詞  
     ry 疑問代詞  
         ryt 時間疑問代詞  
         rys 處所疑問代詞  
         ryv 謂詞性疑問代詞  
     rg 代詞性語素  
 m 數詞  
     mq 數量詞  
 q 量詞  
     qv 動量詞  
     qt 時量詞

## 虛詞

=====

d 副詞  
 p 介詞  
     pba 介詞“把”  
     pbei 介詞“被”  
 c 連詞  
     cc 並列連詞  
 u 助詞  
     Uzhe 著  
     ule 了嘍  
     uguo 過  
     ude1 的底  
     ude2 地  
     ude3 得  
     usuo 所  
     udeng 等等等云云  
     uyy 一樣一般似的般  
     udh 的話  
     uls 來講來說而言說來  
     uzhi 之  
     ulian 連 (“連小學生都會”)

e	嘆詞
y	語氣詞(delete yg)
o	擬聲詞
h	前綴
k	後綴
x	字符串
	xx 非語素字
	xu 網址URL
w	標點符號
	wkz 左括號，全角：（ [ [ { 《 【 【 〈 半角：( [ { <
	wky 右括號，全角：) ] ] } 》 】 】 〉 半角：) ] { >
	wyz 左引號，全角：“ ’ 『
	wyy 右引號，全角：” ’ 』
	wj 句號，全角：。
	ww 問號，全角：？半角：?
	wt 嘆號，全角：！半角：!
	wd 逗號，全角：， 半角：,
	wf 分號，全角：； 半角：;
	wn 頓號，全角：、
	wm 冒號，全角：： 半角：:
	ws 省略號，全角：…… …
	wp 破折號，全角：—— — — 半角：--- ----
	wb 百分號千分號，全角：% ‰ 半角：%
	wh 單位符號，全角：¥ \$ £ ° ℃ 半角：\$

## ICTCLAS 漢語詞性標註集

=====

代碼	名稱	幫助記憶的詮釋
----	----	---------

\_\_\_\_\_ | \_\_\_\_\_ | \_\_\_\_\_

Aq | 形語素 | 形容詞性語素。形容詞代碼為a，語素代碼q前面置以A。

a | 形容詞 | 取英語形容詞adjective的第1個字母。

ad | 副形詞 | 直接作狀語的形容詞。形容詞代碼a和副詞代碼d並在一起。

an | 名形詞 | 具有名詞功能的形容詞。形容詞代碼a和名詞代碼n並在一起。

b | 區別詞 | 取漢字“別”的聲母。

c | 連詞 | 取英語連詞conjunction的第1個字母。

Dg | 副語素 | 副詞性語素。副詞代碼為d，語素代碼g前面置以D。

d | 副詞 | 取adverb的第2個字母，因其第1個字母已用於形容詞。

e | 嘆詞 | 取英語嘆詞exclamation的第1個字母。

f | 方位詞 | 取漢字“方”的聲母。

語素 | 絕大多數語素都能作為合成詞的“詞根”，取漢字“根”的聲母。

h	前接成分	取英語head的第1個字母。
---	------	----------------

i | 成語 | 取英語成語idiom的第1個字母。

j | 簡稱略語 | 取漢字“簡”的聲母。

k | 後接成分 |  
 l | 習用語 | 習用語尚未成為成語，有點“臨時性”，取“臨”的聲母。  
 m | 數詞 | 取英語numeral的第3個字母，n，u已有他用。  
 Ng | 名語素 | 名詞性語素。名詞代碼為n，語素代碼g前面置以N。  
 n | 名詞 | 取英語名詞noun的第1個字母。  
 nr | 人名 | 名詞代碼n和“人(ren)”的聲母並在一起。  
 ns | 地名 | 名詞代碼n和處所詞代碼s並在一起。  
 nt | 機構團體 | “團”的聲母為t，名詞代碼n和t並在一起。  
 nz | 其他專名 | “專”的聲母的第1個字母為z，名詞代碼n和z並在一起。|  
 o | 擬聲詞 | 取英語擬聲詞onomatopoeia的第1個字母。  
 p | 介詞 | 取英語介詞prepositional的第1個字母。  
 q | 量詞 | 取英語quantity的第1個字母。  
 r | 代詞 | 取英語代詞pronoun的第2個字母，因p已用於介詞。  
 s | 處所詞 | 取英語space的第1個字母。  
 Tg | 時語素 | 時間詞性語素。時間詞代碼為t，在語素的代碼g前面置以T。  
 t | 時間詞 | 取英語time的第1個字母。  
 u | 助詞 | 取英語助詞auxiliary 的第2個字母，因a已用於形容詞。  
 Vg | 動語素 | 動詞性語素。動詞代碼為v。在語素的代碼g前面置以V。  
 v | 動詞 | 取英語動詞verb的第一個字母。  
 vd | 副動詞 | 直接作狀語的動詞。動詞和副詞的代碼並在一起。  
 vn | 名動詞 | 指具有名詞功能的動詞。動詞和名詞的代碼並在一起。  
 w | 標點符號 |  
 x | 非語素字 | 非語素字只是一個符號，字母x通常用於代表未知數、符號。  
 y | 語氣詞 | 取漢字“語”的聲母。  
 z | 狀態詞 | 取漢字“狀”的聲母的前一個字母。

jieba 字典中出現的類型

=====

一個

至

在

一個

b

C

d

DF

DG

和

F

G

H

一世

Ĵ

至  
升  
米

毫克  
平方米。

ñ

的  
號  
nrfg  
NRT  
NS  
例如，  
新西蘭

該  
p  
q  
[R

RG  
RR  
政府

小號  
ř

TG

在

出  
和  
新  
該  
紫外線  
上

v

CEO  
VG  
我們  
VN  
VQ

x  
和  
從

ZG

simhash 過濾的掉的詞彙  
====

C  
和  
H  
至

該  
p  
在

出  
和  
新  
該  
紫外線  
上

x  
和