

```

1  # -*- coding:utf-8 -*-
2  import re, datetime, csv , jaconv, spacy
3  class Ext:
4      def __init__(self,rawText):
5          self.rawText = rawText
6          self.datas = {"dates": "", "times": "", "places": ""}
7          self.nlp = spacy.load("ja_ginza")
8
9      def NotationVariability(self,rawText):
10         dt_now = datetime.datetime.now()
11         text = jaconv.z2h(rawText, kana=False, digit=True, ascii=True)
12         with open("dict/NotationVariability.csv",encoding="UTF-8") as f:
13             for row in csv.reader(f):
14                 for i in range(1, len(row)):
15                     text = text.replace(row[i], row[0])
16         text = re.sub("明日|あした|あす",f"{dt_now.month}/{dt_now.day+1}",text)
17         text = re.sub("明後日|あさって|みょうごにち", f"{dt_now.month}/{dt_now.day+2}",
text)
18         return text
19
20     def DateCheck(self, text):
21         pattern1 = r"([0-9]{1,2}月[0-9]{1,2}日)|([0-9]{1,2}/[0-9]{1,2})|([0-9]{1,2}
日)"
22         pattern2 = r"[0-9]{1,2}時[0-9]{0,2}分?"
23         dates = re.findall(pattern1, text)
24         times = re.findall(pattern2, text)
25         return dates,times
26
27     def PlaceCheck(self, text):
28         places = []
29         doc = self.nlp(text)
30         for ent in doc.ents:
31             if ent.label_ == "City" or ent.label_ == "Province" or ent.label_ ==
"Road":
32                 places.append(ent.text)
33                 #print(ent.text,ent.label_)
34         return places
35
36     def Extract(self):
37         textNotate = self.NotationVariability(self.rawText)
38         dates, times = self.DateCheck(textNotate)
39         places = self.PlaceCheck(textNotate)
40         return dates,times,places

```