

Wisers Information Portal

WiseSearch

Ming Pao Instant News

News Sources

Document ID

DOCUMENT ID: 201809123221525

港聞 | 2018-11-01

Basic Info.

港灣道爆水管路面下陷 西行線封閉

Title of the water pipe burst news

港灣道25號一地下食水管爆裂，導致路面下陷，事件中2人受傷。

Main context

提取到的数据

Extracted news articles

#	字段1 Basic Info.	字段2 Title	字段3 Main context	字段4 Document ID	字段5 News Sources
1	A02 港聞 澳門...	澳門口岸又出醜 旅檢...	【靜態組報道】被指...	DOCUMENT ID: 201...	Oriental Daily News
2	E01 文化 2018...	【港情講趣】Sor9y...	【港情講趣】香港人...	DOCUMENT ID: 201...	Apple Daily
3	A06 港聞 2018...	灣仔爆水管 大家樂...	【本報訊】灣仔港灣...	DOCUMENT ID: 201...	Apple Daily
4	P24 港聞 2018...	無妄之災 貨車司機跟...	灣仔華潤大廈對開路...	DOCUMENT ID: 201...	Headline Daily
5	A16 港聞 2018...	爆水管貨車輪陷 膠籃...	【本報訊】一輛隸屬...	DOCUMENT ID: 201...	Oriental Daily News
6	港聞 2018-11-01	港灣道爆水管路面下...	港灣道25號一地下食...	DOCUMENT ID: 201...	Ming Pao Instant Ne...

Web-crawler for news collection

- 21 pieces of news per minute shown in Fig. S2(a)
- Around 134 mins for collecting 2828 news in this study

Performance of the web-crawler

Collected news: 105 pieces (1 repetition); The time used: 4 Mins 48 Seconds; Collection speed: 21 pieces of news per minute

已采集:105条(1条重复) 已用时: 4分钟48秒 平均速度:21条/分钟

(a) Time consumed by web-crawler for collecting news articles

PythonExercise > infrastructure.py

Recognition of infrastructure entities

- 197 secs for extracting 7064 infrastructure entities shown in Fig. S2(b)

Codes for recording time

```
endtime = datetime.datetime.now()
print(endtime)
print((endtime-starttime).seconds)
```

Named entity recognition of infrastructure entities

Start time

2020-03-13 21:28:45.028323

Finish time

2020-03-13 21:31:54.091137

189

189 seconds consumed for recognizing infrastructure entities

(b) Time consumed by recognizing infrastructure entities

Recognition of stakeholder entities

- 208 secs for extracting 9150 stakeholder entities shown in Fig. S2(c)

Codes for recording time

```

import jieba
import numpy as np
import xlswriter
from opencc import OpenCC
import datetime

starttime = datetime.datetime.now()
print(starttime)

workbook = xlswriter.Workbook('C:\\Users\\zhou\\Desktop\\interdependence\\InterdependentStakeholders2020.xlsx')
WhetherTheRelationshipIsExistedWorksheet = workbook.add_worksheet('WhetherTheRelationshipIsExisted')
RMWorksheet = workbook.add_worksheet('RelationshipMatrix')
RMWorksheetAll = workbook.add_worksheet('RelationshipMatrixAll')
ColumnofText = 0
  
```

Named entity recognition of stakeholder entities

Run: stakeholder

Start time: 2020-03-13 21:45:28.417713

Building prefix dict from the default dictionary ...

Loading model from cache C:\\Users\\zhou\\AppData\\Local\\Temp\\jieba.cache

Loading model cost 1.381 seconds.

Prefix dict has been built successfully.

Finish time: 2020-03-13 21:48:56.432611

208 seconds consumed for recognizing stakeholder entities

Process finished with exit code 0

(c) Time consumed by recognizing infrastructure entities

Association rules learning

- 7 secs for learning 11943 potential rules shown in Fig. S2(d)

association rule learning

Run: AprioriRealization

Start time: 2020-03-13 16:58:30.612911

The Search 1 is conduction...

The Search 2 is conduction...

The Search 3 is conduction...

The Search 4 is conduction...

The Search 5 is conduction...

The Search 6 is conduction...

The Search 7 is conduction...

The Search 8 is conduction...

Finish time: 2020-03-13 16:58:38.254348

7 seconds consumed for association rule learning

Process finished with exit code 0

(d) Time consumed by association rule learning