https://tw.news.yahoo.com/www.businessweekly.com.tw--finance/archive/1.html



* 奇摩這邊可以篩選範圍:

所有時間(目標2014/01/01起)

如果頁面內有2014/2013,可以選擇用判斷的方式,讓他自動剔除,若是太麻煩,請直接一起抓下來,後端再手動剔除

* 可以挑選多家報紙

各家報紙來源會在網址列變動,所以請參考下方我所提供的

**內文要抓取的部分**



<cite class="byline vcard">作者： <span class="fn">文／林苑卿</span> | <span class="provider org">財訊</span>&nbsp;–&nbsp;<abbr title="2015-03-26T01:28:56Z">2015年3月26日 上午9:28</abbr></cite>

**原來出處和日期時間都在這邊 (作者不要),其他要**

<div class="bd"><!-- google\_ad\_section\_start --><p class="first">

聯發科董事長前特助袁帝文，農曆春節期間擔任展訊資深副總裁的消息傳出後，業界譁然。

<p>

<!-- google\_ad\_section\_end --></div>

**內文都在這部分的p標籤裡面,雅虎這邊排版目前沒有換頁的困擾,他們習慣在一個頁面內**

<h1 class="headline">展訊挖走核心戰將 爭中國ＩＣ設計霸主 聯發科眼中的叛將 袁帝文的沉痛復仇</h1> **標題**



<div class="pluginCountButton pluginCountNum"><span class="pluginCountConnected" id="u\_0\_1"><span class="pluginCountTextConnected">338 萬</span><span class="pluginCountTextDisconnected">338 萬</span></span></div>

**按贊數**

PAGE

新聞來源

**財訊 https://tw.news.yahoo.com/monthly.wealth.com.tw--finance/archive/3.html**

**請抓取到第4頁**

**法新社 https://tw.news.yahoo.com/afp.cnanews.gov.tw--finance/archive/1.html**

**請抓取到第40頁**

 2014年11月18日

**風傳媒 https://tw.news.yahoo.com/afp.cnanews.gov.tw--finance/archive/1.html**

**請抓取到第20頁**

2014年10月29日

 2014年11月18日

**中時電子報 https://tw.news.yahoo.com/chinatimes.com.tw--finance/archive/40.html**

**請抓取到第40頁**

2015年4月6日

**中央社 https://tw.news.yahoo.com/cna.com.tw--finance/archive/11.html**

**請抓取到第40頁**

2015年3月23日