問題一: 股票更名

例如google在之前有一隻獨立的股票，但在15、16年左右更名成alphabet，所以在15、16年以前SEC的資料若寫的是GOOGLE INC那現在會在證劵行找不到這間公司的股票。

2016-09-30 GARDNER RUSSO & GARDNER LLCSEC資料



2016-06-30 GARDNER RUSSO & GARDNER LLCSEC資料



Google 搜尋google股票顯示結果

一張含有 螢幕擷取畫面 的圖片

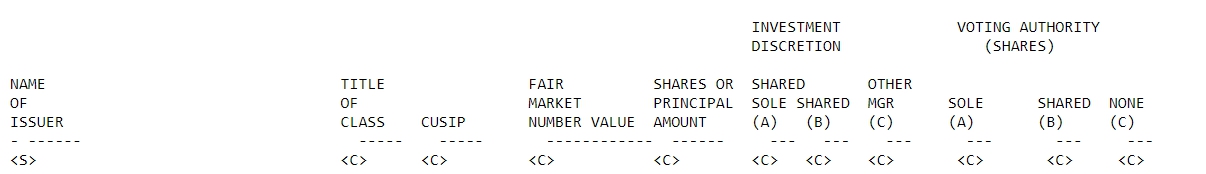
自動產生的描述

問題解決方案: 若能找到每年在那斯達克股票交易中心、紐約證券交易所、美國證券交易所中，所有有股票公司的ticker symbol資料，就可以配合年去對上symbol值，後面再做查詢股票的任何資訊時才不會有資訊缺漏的問題。

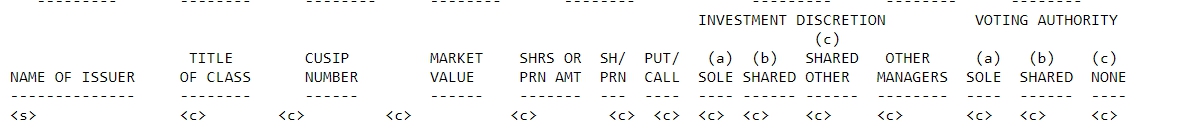
問題二: SEC填寫規範

在SEC的新報表有共同的格式畫規範但再2013年前的舊報表沒有，導致即使使同一間公司也以可能不同年間提交的方式不同。

2003-09-30 GARDNER RUSSO & GARDNER LLCSEC資料



2001-09-30 GARDNER RUSSO & GARDNER LLCSEC資料



2019-12-31 GARDNER RUSSO & GARDNER LLCSEC資料(新報表欄位)



問題解決方案: 直接只取前幾項可能會用到的共同欄位就好，後面的直接捨棄，可以一隻爬蟲廣泛應用到之後的投資人時會減少一些問題。

問題三: SEC命名規範

在SEC中，公司的命名方式我不是很清楚，例如下圖一，有些是Inc 有些是Corp，有些後面有C1有些沒有，A和B和C我查了一下好像是有些股票有分類，但對應的symbol會不同，下面圖二是我google了那個字眼但找不太到對應的資料。

2019-12-31 GARDNER RUSSO & GARDNER LLCSEC資料

一張含有 電腦 的圖片

自動產生的描述

2013-06-30 GARDNER RUSSO & GARDNER LLCSEC資料



問題解決方案: 這邊碰到的問題可能還是得找到更多更詳盡的ticker symbol對照表，以及名稱要如何切才能較好對應到該記錄的ticker symbol，所以我這邊目前想到的解法就是加入更多不同網站和時間symbol lookup，譬如有兩個網站記錄apple公司的寫法是apple APPL、apple Inc APPL，那不管在SEC中記法是apple或apple Inc都能順利的對應到APPL