

題目：標點符號 -v2020-extend-to- 法規文字

隊名：fwy3s408k 金魚腦老頭

人員：郭 ○ ○

日期：2020 年 4 月 19 日

你想解決什麼問題？

- 法規文字解讀不易。



提案類型：

• 10. 消彌不平等.

一般社會印象		
法律人	有錢人	一般人
求學時，有教授專業指導，	可以 - 花費金錢，	每天，忙生活，忙家庭，忙三餐， 一旦，遇到法律事件，... ..，
工作時，有豐富實戰訓練。	購買 - 法律服務。	法律，只保障懂法律的人， 法條都有公布，你為何不讀勒？ 鳴～我就是看不懂啊！

這個問題為何很重要？

- 一般人， 非法律相關之從業人員，
或多或少，都曾感覺過：

法規文字，
反覆堆疊纏繞，
相當拗口難以解讀。



這個問題為何很重要？

- 法律條文：
為解決某些社會問題，所制定的規則，
- 法規文字：
經過：許多人＋許多年的反覆精煉，
融合社會多方面的生活經驗與解決方案
- 實難兼顧：語義簡化而不失周全。



這個問題為何很重要？

- 程式語言：

指揮電腦計算機，為人類規劃安排眾多生活事務，

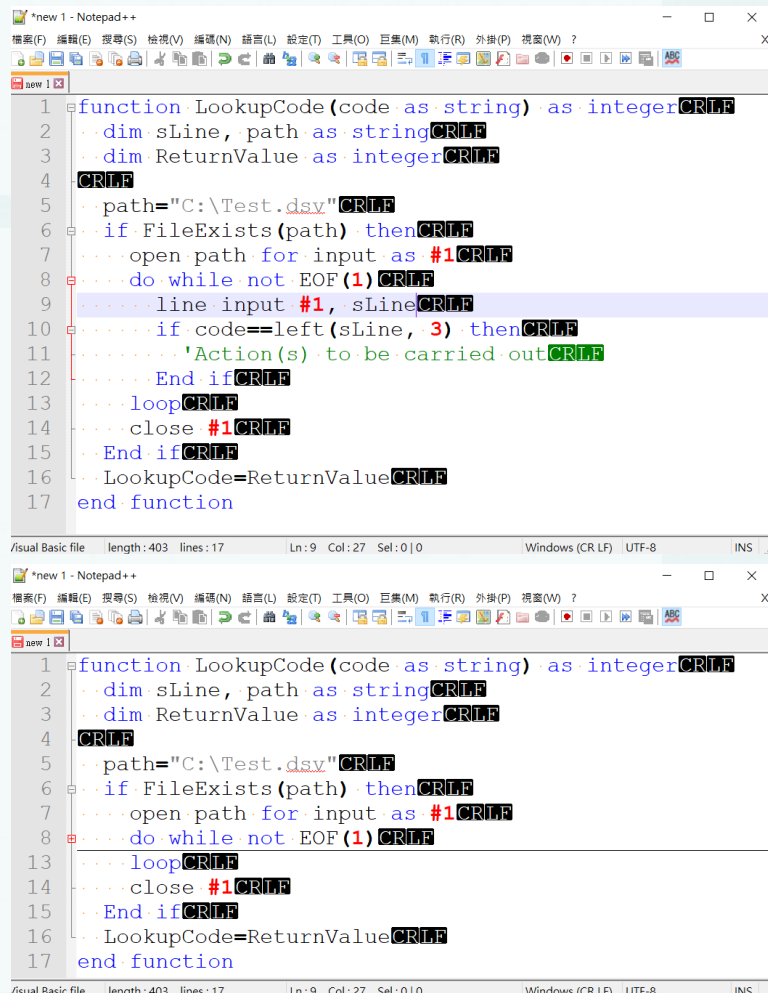
- 其敘事之完整性，易解讀，易編修，等性質，經過～許多人＋許多年～的驗證，無庸置疑。



標點符號 -v2020-extend-to- 法規文字

這個問題為何很重要？

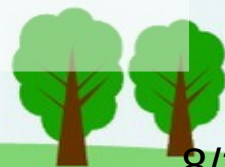
- 法規文字之排版方式，
或可向~程式語言~借鏡若干，
 - 包括：
 - 多層縮排 - 巢狀文字區塊
 - 語法突顯 - 關鍵字高亮度
 - 語法折疊 - 隱藏內層區塊
 - 單鍵切換 - 新型 / 傳統顯示
- 等手法。



```
1 function LookupCode(code as string) as integer
2     dim sLine, path as string
3     dim ReturnValue as integer
4
5     path = "C:\Test.dsv"
6     if FileExists(path) then
7         open path for input as #1
8         do while not EOF(1)
9             line input #1, sLine
10            if code == left(sLine, 3) then
11                'Action(s) to be carried out
12            End if
13        loop
14        close #1
15    End if
16    LookupCode = ReturnValue
17 end function
```

這個問題為何很重要？

- 法規文字之呈現方式，在
現代網頁瀏覽器（Firefox, Chrome, MsIE, MsEdge, …）
豐富動態顯示功能（javascript, css, sql, xml, …）
的強大支援下，有沒有更多想像？



這個問題為何很重要？

- 總算聊到－標點符號－了
- 中文古代文書（文言文）
 - －一般不加或不常使用標點符號

－ [https://zh.wikipedia.org/wiki/ 標點符號](https://zh.wikipedia.org/wiki/標點符號)

- 如增加－標點符號，可便利－文字解讀。
- 下雨天留客天留我不留，有 7 種解讀
- 喜歡上一個人，有 4 種解讀



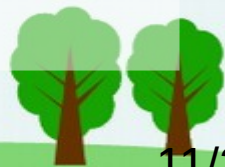
這個問題為何很重要？

- 現有標點符號 - 不好嗎？
- 嗯，一般日常生活使用時，並沒有不好，
- 但，某些領域（公式，法規，...）使用時，希望可以更好。
- 以今觀古，後見之明，現有標點符號，
- 只有，將長條文字串，剪切分段的功能，
- 缺乏，現代標記語法（html, xml, ...），巢狀區塊等能力。



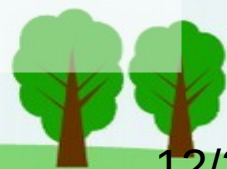
這個問題為何很重要？

- - 法規文字，拗口難解，眾人皆知
- - 把多層語意樹，展開拉直成，長條文字串，閱讀時，不會辛苦嗎？
- - 只用標點符號，僅可剪切分段，這樣足夠嗎？
- - 可以要求更多嗎？巢狀區塊？縮排？凸顯？摺疊？
- - 除了§- 某法 - 第 n 條 - 第 i 項 - 第 j 款之外，還可以看到什麼構造嗎？
平行（不是：舉1反3？而是：舉3反1！、，及，或，與，而，同，但，...）？
串列（主詞-動詞-受詞，於x時，依x規定，除x外，因x致y，以x為y，將x給y，...）？
指代（其，此，該，這，那，前揭，上揭，旨揭，...）？



這個問題為何很重要？

- -法規文字的解讀 => 解 + 讀 => 解析 + 閱讀 => => 解析 + 標記 + 動態顯示 + 閱讀
- -雖然，閱讀的部分，我暫時幫不了你。
但是，解析的部分，你願意幫助別人嗎？
- -如果有人願意預先解析 + 標記，
你會比較容易閱讀嗎？
- -如果先前有1人~解析時~多花1.5x1倍力氣~進行標記，
可否讓後續n人~閱讀時~節省0.5xn倍力氣？~搭配動態顯示。



這個問題為何很重要？

- 改善法規文字解讀性.

- 對個人

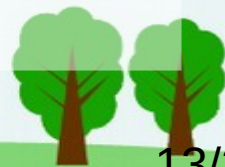
個人，每個人都會老化，大腦記憶退化，語言文字解讀變慢
若能預先解析+留下標記，可否利於法規文字之後續閱讀？

- 對組織

組織，大量工商契約，作業手冊規範，工作標準流程，盤根錯節，
若能改善法規文字公文書類的解讀性，可否利於工作流程與經驗傳承？

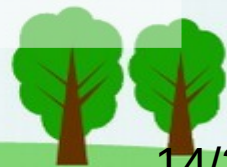
- 對社會

社會，若能改善~法規文字~的呈現方式，可否利於民眾瞭解法規內容？



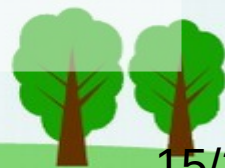
這個問題為何很重要？

- - 摺疊 = 簡化
- - 法規文字，畫成語意樹狀圖 = 巢狀區塊縮排。
- - 保留主幹，修剪分枝 = 摺疊內層區塊。
- - 雖可，簡化主題，但也，丟失細節。
- - 近看_{可得}細節 ~VS~ 遠觀_{得見}主題。
- - 以上，路途還很遙遠，目前只是想像。



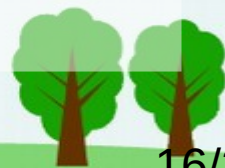
這個問題為何很重要？

- - 哪些分枝要保留？哪些分枝要修剪？
- - 如果：把：法規文字 ~ 畫成：資料結構的圖 (Graph),
 - 則可：與我的 ~ 狀況及需求 ~ 有關的部分, ~ 保留！
與我的 ~ 狀況及需求 ~ 無關的部分, ~ 修剪！
- - 我的 ~ 狀況及需求 \neq 你的 ~ 狀況及需求,
 - 你的 ~ 修剪或保留 \neq 我的 ~ 修剪或保留.
- - 以上，路途還很遙遠，目前只是想像.



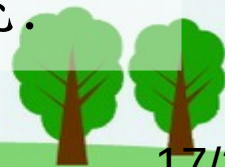
這個問題為何很重要？

- - 對這標記系統的構想
- - 在法規文字的解讀過程，有部分力氣，花費在詞語的掃描－解析－辨認－層級－範圍等上面．
- - 如果，可以預先～解析＋標記，能否便利後續～閱讀．
- - 是否可以，把你的理解，變成我的理解，變成眾人的理解．



這個問題為何很重要？

- 對這標記系統的構想
- 手工標記對象，以法規文字為之，
相較於裁判書，法規文字，數量相對較少，且為眾人所需閱讀熟悉。
- 以手工標記方式為主，的標記系統，
但不排除他日可以局部自動化。
- 可視覺檢視 + 手工編修，的標記系統。
- A君的標記結果，B君認為局部不ok，也可手工潤飾，的標記系統。

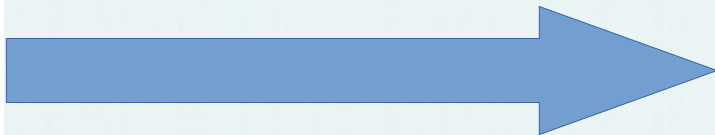


標點符號 -v2020-extend-to- 法規文字

這個問題為何很重要？



3D



2D



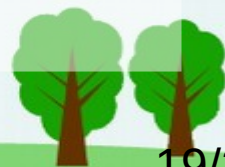
1D

畫面來源：均來自網路，如不慎冒犯，懇請告知，願意刪除



這個問題為何很重要？

- - 講這麼多，我想幹什麼？
- - 我想建構一個標記語法，for 法規文字
- - 標記語法，需要實兵演練。
- - 我想發動多人協作，實戰標記既有法規文字。



目前掌握了哪些相關內部或外部資料？

- 感謝法務部網站 - 全國法規資料庫

<https://law.moj.gov.tw/>

- 辛苦的，把法規文字，從印刷紙本，數化成html網頁，並仁慈的，慷慨的，無私的，

-----借用電影台詞：地心毀滅 TheCore, 辛斯基教授「借用」凱斯的研究成果

- 分享全部法規文字資料。



期待中的專案成果為何？

- 法規文字 的
標記語法 及
協作網頁



期待本次專案成果之影響或效益？

- 改善
法規文字
的解讀性。



結尾

---Finish---



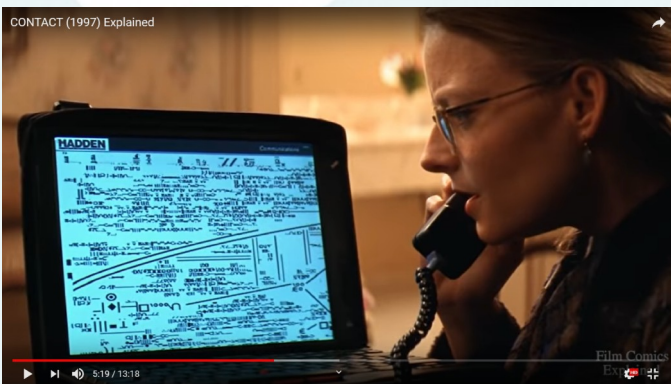
廢棄區

---Obsolete pages---



標點符號 -v2020-extend-to- 法規文字

這個問題為何很重要？



畫面來源：電影：接觸未來 (Contact, 1997), 及網站：Youtube

電影：接觸未來 (Contact, 1997)

蟲洞傳輸系統，外星人傳送1維的訊號串列，用德國奧運影像為固定間距切割，變成多幅的2維圖紙影像，再用規則及記號，經 ∞ 維拼圖，組合成鉅細靡遺 (BIM-LOD=max) 的設計圖。

（為何，會聊到工程設計圖的標記及拼接 (BIM-LOD=max)? 因為，這是我原來進攻的主線任務，而法規文字的標記-收合-修剪，是後來追加的支線任務，而且是刻意挑選~機器自動計算，於近年內-短期間，幫不上忙的，非人力不可，需要工讀生大軍~的任務）

法規文字，1維串列或1維長條魷魚絲，用標點符號之逗號，分號，句號，或烤肉攤剪刀，剪開為短詞，短句，或短魷魚絲，可以再組合回到2維-香烤魷魚片？或是組合成3維-新鮮現撈魷魚？



標點符號 -v2020-extend-to- 法規文字

這個問題為何很重要？



如果～

電影：異星入境 (Arrival, 2016)

電影裡的外星人～的語言形式～與思考模式，
是首尾相接的圓環

那麼～

身處資訊時代的我們，

適合閱讀的法規文字～的語言形式～與思考模式，

該是何種形狀？漢書 PE2 往復折返？多層巢狀區塊？樹狀語意？



畫面來源：電影：異星入境 (Arrival, 2016), 及網站：Youtube



這個問題為何很重要？

- 為何法規文字，有很多平行詞句？

• 舉 1 反 3	輸入 1 個參數 選用先驗函數 產生 3 或多個回傳值
使用知識	
舉 3 反 1	輸入 3 或多個參數 反推生成函數 產生 1 或多個回傳值
產生知識	



這個問題為何很重要？

- 為何不使用 NLP (如 LSTM, ELMo, GPT, BERT, XLNet, ... 等) 相關演算法, 自動計算, 完成~全部法規文字~之解讀+標記?
- 我承認, 在 NLP 領域, 我還未深入瞭解, 也未曾細讀論文, 更未動手實作例題。
- NLP 演算+自動標記? 我認為, 標記的動作, 是完成~先前未完成語意, 機器演算猜測, 或許可以代勞若干, 但, 終究仍需人類檢查潤飾,
- 譬如, 辦公室新手辦事員, 初擬公文簽稿, 尚需各層主管檢查修改潤飾, 使語意更趨完整, 何況機器的演算成果?
- 彭端淑-為學一首示子侄---蜀鄙二僧, 欲往南海,
貧僧, 手工解析+標記, 搖櫓小舟, 先行出發, 探崎嶇小徑, 如有人機合作的水上摩托車, 也是不錯的啦。
富僧, 自動 NLP 計算, 巨型戰艦, 寬闊河道, 若可後發先至, 亦是美事一樁



標點符號 -v2020-extend-to- 法規文字

廢棄區

---Finish---

