

Final Project

Generation of Fairy Tales Scripts

Introduction of Artificial Intelligence and Literature

Team 3

107065501 蘇攷如

107065527 段凱文

Agenda

- Tools
- Extraction
- Scene
- Entity Expansion
- Case Results
- Competition Results



Part I - Tools



The Stanford Natural Language Processing Group

Stanford NLP

- 自然語言處理套件，支援多國語言
- 已訓練好的繁體中文模型
 - Stanford CoreNLP
 - StanfordNLP Python Library
 - Stanford Parser
 - **Stanford POS Tagger**
 - Stanford Named Entity Recognizer
 - Stanford Word Segmente
 - **Stanford Dependency Parser**
- 使用 Java models 運行至 Python



Jieba-zh_TW

- 進行句子有效斷詞
- 採用和原始結巴相同的演算法，替換其詞庫及HMM機率表
製做出針對台灣繁體的結巴斷詞器

['從前', '有', '個', '可愛', '的', '小姑娘', '，', '誰', '見', '了', '都', '喜
歡', '，', '但', '最', '喜歡', '她', '的', '是', '她', '的', '奶奶', '，', '簡
直', '是', '她', '要', '什麼', '就', '給', '她', '什麼', '。']

POS Tagger 標注詞性

- 查找詞性
- 抓住「人物」、「地點」



[(從前#AD), (有#VE), (個#NN), (可愛#VV), (的#DEC), (小姑娘#NN), (， #PU), (誰#NN), (見#VV), (了#AS), (都#AD), (喜歡#VV), (， #PU), (但#AD), (最#AD), (喜歡#VV), (她#PN), (的#DEC), (是#VC), (她#PN), (的#DEG), (奶奶#NN), (， #PU), (簡直#NR), (是#VC), (她#PN), (要#VV), (什麼#VV), (就#AD), (給#VV), (她#PN), (什麼#VV), (。 #PU)]

CNA Dataset

- 中文 word embedding 資料集
- 相似詞獲取

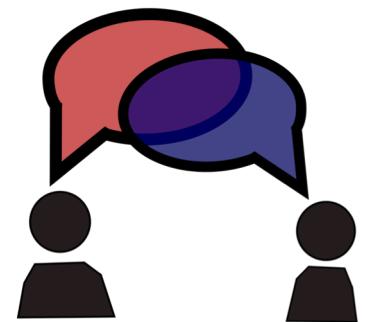
```
[4]: ch.most_similar('國王')

[4]: [('王儲', 0.8000344634056091),
      ('親王', 0.750373363494873),
      ('新王', 0.7259018421173096),
      ('女王', 0.7144579887390137),
      ('王后', 0.7024247646331787),
      ('哈山', 0.6718887686729431),
      ('包杜恩', 0.6705120801925659),
      ('阿布杜拉', 0.6595524549484253),
      ('賈南德拉', 0.6540836095809937),
      ('畢蘭德拉', 0.6476490497589111)]
```

Dependency Parser

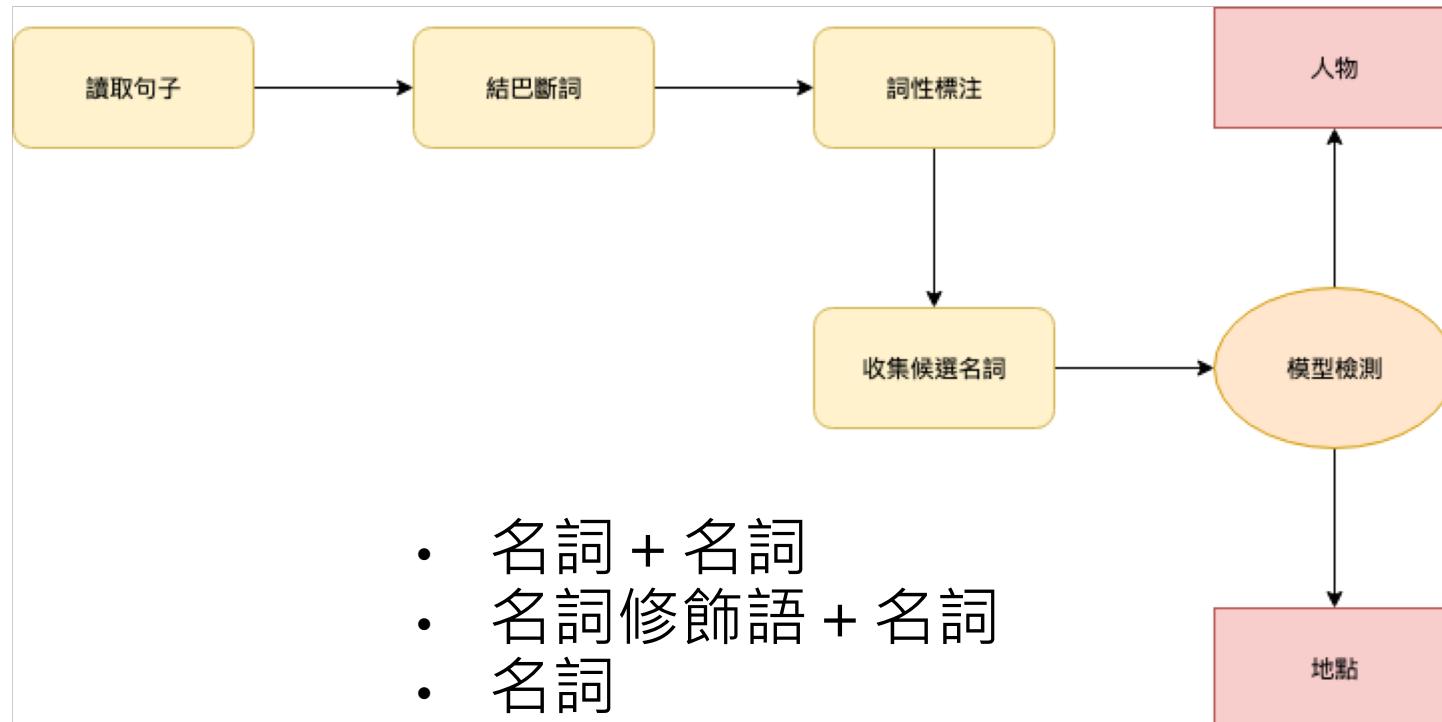
(('有', 'VE'), 'dep', ('從前', 'NN'))
((('有', 'VE'), 'dobj', ('小姑娘', 'NN'))))
((('小姑娘', 'NN'), 'nmod:assmod', ('可愛', 'NN'))))
((('可愛', 'NN'), 'compound:nn', ('個', 'NN'))))
((('可愛', 'NN'), 'case', ('的', 'DEG'))))
((('有', 'VE'), 'conj', ('見', 'VV'))))
((('見', 'VV'), 'nsubj', ('誰', 'NN'))))
((('見', 'VV'), 'aux:asp', ('了', 'AS'))))
((('見', 'VV'), 'conj', ('喜歡', 'VV'))))
((('喜歡', 'VV'), 'advmod', ('都', 'AD'))))
((('有', 'VE'), 'conj', ('奶奶', 'NN'))))
((('奶奶', 'NN'), 'advmod', ('但', 'AD'))))
((('奶奶', 'NN'), 'nsubj', ('喜歡', 'VV'))))
((('喜歡', 'VV'), 'advmod', ('最', 'AD'))))
((('喜歡', 'VV'), 'dobj', ('她', 'PN'))))
((('喜歡', 'VV'), 'mark', ('的', 'DEC'))))
((('奶奶', 'NN'), 'cop', ('是', 'VC'))))

((('奶奶', 'NN'), 'nmod:assmod', ('她', 'PN'))))
((('她', 'PN'), 'case', ('的', 'DEG'))))
((('奶奶', 'NN'), 'conj', ('要', 'VV'))))
((('要', 'VV'), 'advmod', ('簡直', 'AD'))))
((('要', 'VV'), 'cop', ('是', 'VC'))))
((('要', 'VV'), 'nsubj', ('她', 'PN'))))
((('要', 'VV'), 'ccomp', ('給', 'VV'))))
((('給', 'VV'), 'nsubj', ('什麼', 'NN'))))
((('給', 'VV'), 'advmod', ('就', 'AD'))))
((('給', 'VV'), 'dobj', ('什麼', 'NN'))))
((('什麼', 'NN'), 'compound:nn', ('她', 'PN'))))

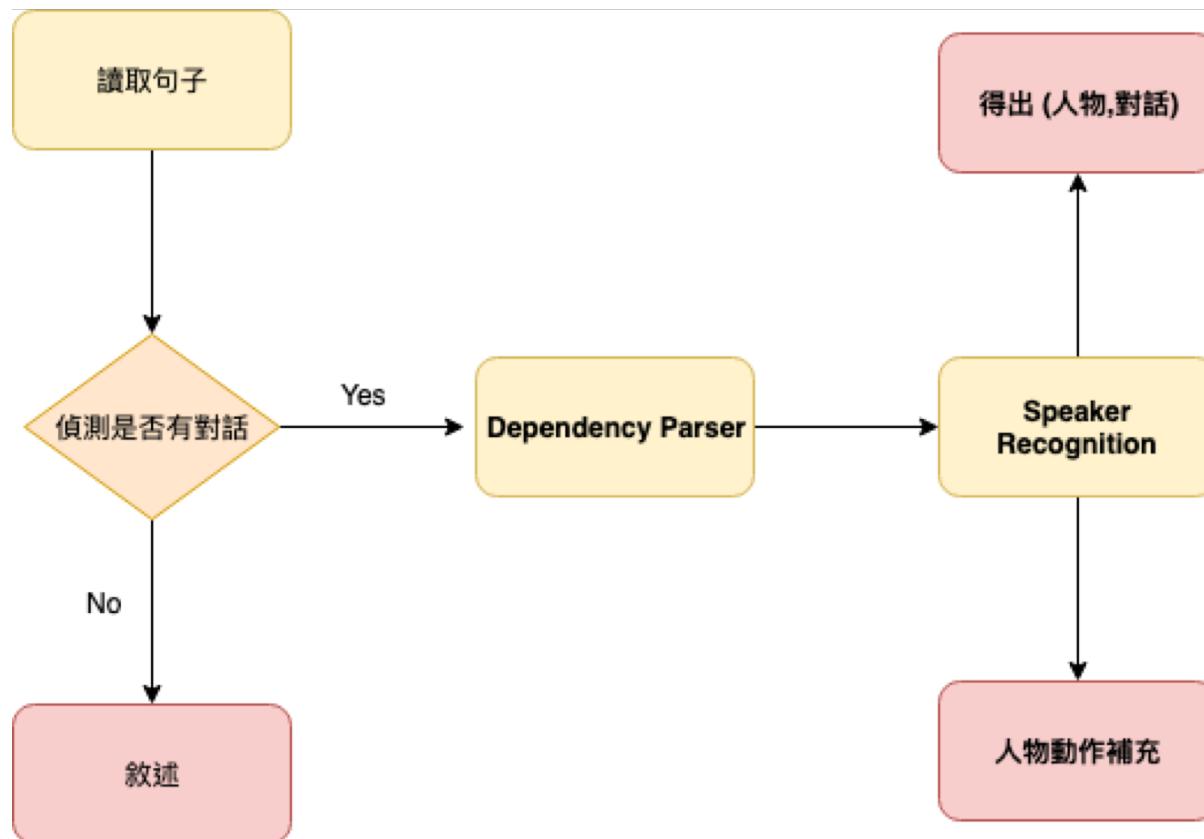


Part II - Extraction

Character & Location



Dialogue



Part III - Scene

場景判斷

- 依照「**時**」、「**空**」分景
 - 時間：清晨、白天、黃昏、晚上
 - 空間：Location
- 每句話都會紀錄其時空
- 只要跟前期狀態不同，尤其是對於地點特別敏感，立即切換場景
- 地點在文本中所屬位置會影響分景正確度

場景空間的判斷

- 偵測對話框內的地點 v.s. 偵測對話框外的地點
- 對話地點判斷器
- 以三隻小豬為例，以對話內的地點為偵測目標
- 以小紅帽為例，以對話外的地點為偵測目標

Part IV Entity Expansion

人物、地點標註

- 我們先人工標註一些童話故事裡的名詞為「人物」跟「地點」
- 用 cosine similarity 把「人物」跟「地點」的相似詞加入詞組中
- 人工清理新生成的詞組

分類器

- 用 word vector 為 input 訓練一個 logistic regression 的分類器
• 把名詞分成「人物」、「地點」跟其他
- 當遇到不在現有資料中的詞彙時，用分類器來猜測



Part V - Case Results



小紅帽

人物	[‘奶奶’, ‘媽媽’, ‘小紅帽’, ‘狼’, ‘獵人’]
地點	[‘奶奶家’, ‘家’, ‘森林’]
幕數	3
切景方式	對話外
劇本精準度	97%

小紅帽 - 劇本

=====

Scene: 時/ 白天 景/ 家

從前有個可愛的小姑娘，誰見了都喜歡，但最喜歡她的是她的奶奶，簡直是她要什麼就給她什麼。一次，奶奶送給小姑娘一頂用絲絨做的小紅帽，戴在她的頭上正好合適。從此，姑娘再也不願意戴任何別的帽子，於是大家便叫她'小紅帽'。

媽媽: 來，小紅帽，這裡有一塊蛋糕和一瓶葡萄酒，快給奶奶送去，奶奶生病了，身子很虛弱，吃了這些就會好一些的。趁著現在天還沒有熱，趕緊動身吧。在路上要好好走，不要跑，也不要離開大路，否則你會摔跤的，那樣奶奶就什麼也吃不上了。到奶奶家的時候，別忘了說'早上好'，也不要一進屋就東瞧西瞅。

媽媽: 我會小心的。

=====

=====

Scene: 時/ 白天 景/ 森林

奶奶住在村子外面的森林里，離小紅帽家有很長一段路。小紅帽剛走進森林就碰到了一條狼。小紅帽不知道狼是壞家伙，所以一點也不怕它。

狼: 你好，小紅帽

小紅帽: 謝謝你，狼先生。

狼: 小紅帽，這麼早要到哪裡去呀？

小紅帽: 我要到奶奶家去。

狼: 你那圍裙下面有什麼呀？

小紅帽: 蛋糕和葡萄酒。昨天我們家烤了一些蛋糕，可憐的奶奶生了病，要吃一些好東西才能恢復過來。

狼: 你奶奶住在哪裡呀，小紅帽？

小紅帽: 進了林子還有一段路呢。她的房子就在三棵大橡樹下，低處圍著核桃樹籬笆。你一定知道的。

狼: 這小東西細皮嫩肉的，味道肯定比那老太婆要好。我要講究一下策略，讓她倆都逃不出我的手心。

狼: 小紅帽，你看周圍這些花多麼美麗啊！幹嗎不回頭看一看呢？還有這些小鳥，它們唱得多麼動聽啊！你大概根本沒有聽到吧？林子里的一切多麼美好啊，而你卻只管往前走，就像是去上學一樣。

小紅帽: 也許我該摘一把鮮花給奶奶，讓她高興高興。現在天色還早，我不會去遲的。

她於是離開大路，走進林子去採花。她每採下一朵花，總覺得前面還有更美麗的花朵，便又向前走去，結果一直走到了林子深處。

=====

=====

Scene: 時/ 白天 景/ 奶奶家

就在此時，狼卻直接跑到奶奶家，敲了敲門。

奶奶: 是誰呀？

狼: 是小紅帽，我給你送蛋糕和葡萄酒來了。快開門哪。

奶奶: 我身上沒有力氣，起不來。

狼剛拉起門栓，那門就開了。狼二話沒說就衝到奶奶的床前，把奶奶吞進了肚子。然後她穿上奶奶的衣服，戴上她的帽子，躺在床上，還拉上了簾子。

可這時小紅帽還在跑來跑去地採花。直到採了許多許多，她都拿不了啦，她才想起奶奶，重新上路去奶奶家。看到奶奶家的屋門敞開著，小紅帽感到很奇怪。

小紅帽: 天哪！平常我那麼喜歡來奶奶家，今天怎麼這樣害怕？

小紅帽: 早上好！

小紅帽: 哎，奶奶，你的耳朵怎麼這樣大呀？

狼: 為了更好地聽你說話呀，乖乖。

小紅帽: 可是奶奶，你的眼睛怎麼這樣大呀？

狼: 為了更清楚地看你呀，乖乖。

小紅帽: 奶奶，你的手怎麼這樣大呀？

狼: 可以更好地抱著你呀。

小紅帽: 奶奶，你的嘴巴怎麼大得很嚇人呀？

狼: 可以一口把你吃掉呀！

狼剛把話說完，就從床上跳起來，把小紅帽吞進了肚子，狼滿足了食欲之後便重新躺到床上睡覺，而且鼾聲震天。

獵人：這老太太鼾打得好響啊！我要進去看看她是不是出什麼事了。

獵人：你這老壞蛋，我找了你這麼久，真沒想到在這裡找到你！

他正準備向狼開槍，突然又想到，這狼很可能把奶奶吞進了肚子，奶奶也許還活著。

獵人就沒有開槍，而是操起一把剪刀，動手把呼呼大睡的狼的肚子剪了開來。

他剛剪了兩下，就看到了紅色的小帽子。

小紅帽：真把我嚇壞了！狼肚子里黑漆漆的。

接著，奶奶也活著出來了，只是有點喘不過氣來。小紅帽趕緊跑去搬來幾塊大石頭，塞進狼的肚子。狼醒來之後想逃走，可是那些石頭太重了，它剛站起來就跌到在地，摔死了。

三個人高興極了。獵人剝下狼皮，回家去了；奶奶吃了小紅帽帶來的蛋糕和葡萄酒，精神好多了。

===== The End =====

三隻小豬

人物	[‘媽媽’, ‘大哥’, ‘二哥’, ‘小弟’, ‘野狼’]
地點	[‘稻草屋’, ‘家’, ‘木屋子’, ‘磚屋’]
幕數	5
切景方式	對話內
劇本精準度	97%

三隻小豬 - 劇本

從前，有一隻胖胖的豬媽媽，她生了三隻小豬。最大的小豬，豬大哥很貪睡，很懶惰。一天到晚都在打瞌睡。第二個小豬，豬二哥很愛吃，他也很懶惰。幸好最小的小豬，豬小弟是個勤勞的好孩子。常常努力的工作。

Scene: 時/ 白天 景/ 家

媽媽: 你們都長大了，應該自己蓋房子，自己住，自己種田，自己生活。我要你們自己照顧自己
豬大哥噘起嘴，懶洋洋走出家門。

Scene: 時/ 白天 景/ 稻草屋

大哥: 呀呀．．．．．好討厭哦！

大哥: 我的房子蓋好了，很漂亮吧～，你們也快一點蓋吧！

二哥: 哇，果然很漂亮，大哥，你真了不起啊！

小弟: 哥，你不擔心稻草屋會被風吹倒嗎？

二哥: 稻草屋會被風吹倒，那我就用木頭來蓋好了，木屋較重，不怕風吹

二哥: 你們看，這麼漂亮的房子！而且釘得很牢固，不會被風吹倒，我真的好棒呀！

小弟: 木屋雖然不會被風吹倒，可是用力打，木頭會被打斷，房子就垮了。

二哥: 你以為你最聰明？看你搬磚頭搬一整天了，房子還沒蓋出來。笨蛋～～

小弟: 我蓋房子雖然比你們慢，但是我要蓋的房子不怕風吹，也不怕敲打，紅磚房子最牢固了。

豬小弟不理會哥哥的嘲笑，他搬好了磚塊，又搬水泥，他把水泥和好了，開始堆砌磚塊，一塊一塊將抹上水泥的磚頭堆砌起來。

小弟: 你們不要笑我, 等我蓋好了, 你們就知道了, 我的房子比什麼都堅固, 野狼來了我也不怕呢!
豬小弟繼續加油工作, 他趕呀趕, 趕到天黑, 月亮掛在天上了, 他的紅磚房子才好不容易蓋好了。豬小弟正想休息, 却聽到了大野狼的吼叫聲。

小弟: 幸好房子蓋好了, 我不怕大野狼。

這時兩個豬哥哥也聽到了野狼的吼叫聲, 怕得發抖, 他們怕野狼來, 所以整夜都不敢安心睡覺。

Scene: 時/ 白天 景/ 稻草屋

隔天, 豬媽媽要三隻小豬到田裡工作, 沒想到走到半路的時候, 他們被一隻可怕的大野狼發現了, 野狼一直跟蹤他們。

野狼決定先吃豬大哥。

豬大哥嚇得臉都白了, 他趕緊拿一把大鎖, 把門緊緊鎖了起來。

野狼: 你這個大笨蛋, 這種稻草屋, 我吹一口氣就倒了。

他說完話, 用力一吹, 果然把稻草屋吹倒了, 豬大哥嚇得直跑到豬二哥家。

Scene: 時/ 白天 景/ 木屋子

大哥: 救命呀!

野狼: 木頭屋子一樣擋不住我, 我一定要把你吃掉!

兩隻小豬怕得縮著身子擠在一起, 頭上直冒冷汗。

兩隻小豬逃呀逃呀, 逃到小弟弟的家中。

大哥: 弟弟, 弟弟, 小· · · · · 弟弟, 野狼在我們· · 身· · · 後· · ·

大哥: 快把門鎖緊, 不然我們會被吃掉的。

=====Scene: 時/ 白天 景/ 磚屋

小弟: 怕什麼？這麼堅固的磚屋，野狼進不來的。
很快的，野狼就追來了。

野狼: 死小豬，看我把你們的房子撞倒，你們就要被我吃掉了。
野狼說著說的開始撞牆了。

野狼: 呀！～～

野狼: 啟唷，痛死我了，痛死我了！

野狼: 萬歲！

從此，他們三個兄弟住在一起，每天一起吃飯睡覺，也一起工作，日子過得很快樂，而且野狼一直沒再出現呢！

=====The End=====

Part VI - Competition Results

白雪公主-分析

人物	['國王', '王后', '白雪公主', '魔鏡', '僕人', '老婆婆', '農婦', '王子']
地點	['森林', '王宮', '家', '山頂']
幕數	6
切景方式	對話外
劇本精準度	94%

白雪公主與七個小矮人 - 劇本

Scene: 時/ 白天 景/ 家

從前，在一個遙遠的國度裡，住著一個國王和王后。

他們一直想有一個孩子。

王后: 我多想要一個孩子啊，希望她的皮膚像這雪一樣白，臉蛋像蘋果一樣紅。

沒過多久，王后果然生下了一個可愛的小公主，小公主長得和她希望的一模一樣，國王和王后給她取名為白雪公主。

白雪公主出生不久，王后就去世了。

一年後，國王娶了新王后。

新王后長得非常美麗，但是她的嫉妒心十分強，她不能容忍任何人比她更漂亮。

王后: 魔鏡，魔鏡，請你告訴我，誰是天底下最漂亮的的女人?

魔鏡: 王后，這還用問嗎? 天底下最漂亮的的女人就是你。

王后聽完，滿意地笑起來。

白雪公主漸漸地長大了，在她七歲的時候，已經長成一個人見人愛的美少女。

王后: 魔鏡，魔鏡，請你告訴我，誰是天底下最漂亮的的女人?

魔鏡: 我的王后，你非常漂亮。可是白雪公主比你更漂亮!

王后生氣極了，她開始越來越討厭白雪公主。

Scene: 時/ 白天 景/ 森林

有一天，她實在忍不下去了，於是叫來一個僕人，命令他把白雪公主帶到森林裡去殺掉，並把她的心和肝帶回來作為憑證。

僕人: 公主，王后讓我來殺掉你，可我不能那麼做，你還是往森林裡逃走吧!

王后看了僕人帶回來的東西，以為白雪公主真的死了。

可憐的白雪公主一個人在森林中四處亂走，四周傳來野獸的號叫，她害怕極了。

屋子裡有一張小桌子，鋪著潔白的桌布，上面放著七個小碗、七把小勺子和七個小杯子，桌子的周圍放著七張小椅子；屋子的另一邊，放著七張小床。

屋子裡的一切實在太有趣了，白雪公主碰碰這兒，又碰碰那兒，還品嚐了放在盤子裡的點心，但是她實在是太累了，就倒在最後一張床上睡著了。

傍晚，七個小矮人回來了，他們是這間小屋的主人。

每天清晨，他們都會帶著工具到森林的深處去掏金子，一直工作到很晚才回家。

他們點亮七盞小油燈，發現屋子裡的東西被人動過了。

他們到處尋找，最後，在第七個小矮人的床上發現了白雪公主。

僕人: 好漂亮的姑娘呀!

Scene: 時/清晨 景/森林

白雪公主: 你叫什麼名字?是從哪裡來的呢?

白雪公主: 我叫白雪公主。

白雪公主: 你願意留下來嗎?沒事的時候，你可以幫我們把家裡收拾整潔。

白雪公主: 我會把家裡的事料理好的。

白雪公主每天都把小屋打掃得乾乾淨淨。

當小矮人從森林回來後，總能吃上可口的晚餐。

他們在一起生活得很快樂。

而王后呢?她一直以為白雪公主已經死了，認為自己是世界上最漂亮的女人的。

魔鏡: 我的王后，你非常漂亮。可是和七個小矮人住在一起的白雪公主比你更漂亮。

王后感到非常憤怒，她知道一定是僕人騙了她，不管怎樣一定要想辦法讓白雪公主從這個世界上消失。

於是，她裝扮成一個賣雜貨的老婆婆，翻過七座山，來到七個小矮人的屋外。

王后: 可愛的小姑娘，買絲帶嗎?這麼漂亮的絲帶只配繫在你雪白的脖子上。

善良的白雪公主打開門，把老婆婆讓了進去。她選了一條彩色的絲帶，並讓老婆婆幫她繫在脖子上。

老婆婆用絲帶緊緊地勒住白雪公主的脖子，白雪公主透不過氣來，很快便失去知覺倒在了地上，一動也不動。

王后: 哈哈，現在你再也不是最漂亮的的女人了。

傍晚，七個小矮人剛一進屋就看見白雪公主躺在地上，一動不動，臉色蒼白。

他們趕緊將她抱起來放到床上，把那條緊緊繫在她脖子上的絲帶鬆開。

過了一會兒，白雪公主漸漸有了呼吸，臉色也開始慢慢變紅。

王后: 那個老婆婆肯定是邪惡的王后假扮的，你要小心一點，以後，千萬不能讓陌生人進屋。

王后: 魔鏡，魔鏡，請你告訴我，誰是天底下最漂亮的女人？

魔鏡: 我的王后，你非常漂亮，可是和七個小矮人住在一起的白雪公主比你更漂亮。

王后氣得滿臉通紅。

這次，她做了一把有毒的梳子，扮成另一個老婆婆，翻過七座山，來到七個小矮人的屋外。

老婆婆: 可愛的小姑娘，買梳子嗎？這麼好的梳子只配梳你黑亮的頭髮。

白雪公主: 我才不會讓你進來呢。

老婆婆: 我不進屋，你看看梳子總可以吧？

白雪公主接過梳子，她實在太喜歡這個樣式了，她想用這把漂亮的梳子在頭髮上試一試。

可是，梳子剛一觸到頭髮，毒性就發作了，白雪公主立即失去了知覺，倒在地上，一動也不動。

王后: 這就是你漂亮的下場。

傍晚，七個小矮人進屋又看見白雪公主躺在地上，他們把她抱到床上，並在她的頭髮裡找到一把梳子。

他們知道這一定又是王后搞的鬼。

他們剛一抽出梳子，白雪公主就甦醒過來了。小矮人又一次囑咐她，千萬別輕信陌生人了。

王后: 我的王后，你非常漂亮。可是和七個小矮人住在一起的白雪公主比你更漂亮。

王后: 我一定要讓她永遠消失！

這次王后想出了一個更陰險的辦法。

她做了一個好看的蘋果，一半是紅色的，一半是綠色的，並在紅色的一邊塗上了毒藥，只要咬上一小口，就會喪命。

她扮成一個農婦，翻過七座山，來到七個小矮人的屋外。

有了前兩次的教訓，白雪公主再也不敢開門了。

農婦: 不用你開門，你看，我有一個可愛的蘋果，送給你當禮物吧。

白雪公主: 不，我什麼也不要。

農婦: 你怕蘋果有毒嗎？來！我吃一半，你吃一半。

她把紅色的那半蘋果從窗縫裡遞給白雪公主。

白雪公主見農婦吃了另一半蘋果，忍不住咬了一口。

王后: 白雪公主現在永遠從世界上消失了。

魔鏡: 是你啊，我的王后！天下最漂亮的女人就是你。

七個小矮人回到家中，看到白雪公主倒在地上。他們馬上把她抬到床上，但用盡一切辦法，也沒能讓白雪公主甦醒過來。

小矮人非常傷心，他們不願把白雪公主埋到黑漆漆的地下，所以訂做了一口水晶棺，這樣他們隔著棺材也能看見她。

Scene: 時/白天 景/山頂

他們將棺材放到小山頂上，每天都由一個小矮人看守。白雪公主在水晶棺裡面躺了很久，她的皮膚依然非常白皙，微微透出健康的紅潤，頭髮仍然像烏木一樣漆黑，她就像只是睡著了一樣。

=====

Scene: 時/ 白天 景/ 森林

有一天，一位王子走進森林，在山頂上見到了躺在水晶棺裡的白雪公主。

王子: 我會愛她勝過愛我自己，會永遠呵護她的。

就在王子叫人抬著水晶棺下山的時候，棺材被路旁的矮樹撞了一下，那塊有毒的蘋果從白雪公主的嘴中吐了出來，她很快醒了過來，從棺材裡坐了起來。

王子: 我怎麼和你在一起？

王子: 我想永遠照顧你，你能嫁給我嗎？

他們一起回到王子的國家，舉辦了隆重的婚禮。

婚禮舉行前，他們邀請了許多人，包括那個惡毒的王后。

王后: 魔鏡，魔鏡，請你告訴我，誰是天底下最漂亮的女人呢？

魔鏡: 我的王后，你非常漂亮，但是年輕的王后比你更漂亮。

王后聽了十分生氣，她想去看看年輕的王后到底是什麼樣子。

=====

Scene: 時/ 白天 景/ 王宮

她剛走進王宮，就一眼認出了白雪公主。她知道自己的陰謀會被揭穿的，嚇得連夜逃到森林深處再也不敢出來了。

從此，白雪公主和王子幸福地生活在一起.....

=====The End=====

補鞋匠的奇特經歷 - 分析

人物	['法官', '法蒂瑪', '馬魯夫']
地點	['家', '清真寺', '補鞋攤']
幕數	4
切景方式	對話外
劇本精準度	63%

結果分析

- 大部分的角色都有抓出來
- 地點提取得比較差
 - 切景受到影響
- 代名詞沒有處理
 - 對話會給錯誤的角色
- 故事的敘述沒有處理

Conclusion

Conclusion

- Stanford NLP 有助於元素抽取
- Entity Expansion 提高「人物」、「地點」判斷的精準度
- 場景切割對地點特別敏感，時間其次
 - 因為在故事中時間可能不像地點特別說明出來
- 生成忠於原作的劇本

Q & A