

# CWordTM (0.7.6) Toolkit Usage on the Holy Bible (CUV)

This Jupyter notebook demonstrates how to use the package "CWordTM" on the Holy Bible (Chinese Union Version - Traditional Chinese):

1. Utility Features
2. Text Visualization - Word Cloud
3. Text Summarization
4. Pivot Table
5. OT Quotes
6. Topic Modeling - LDA, BERTopic and NMF

**CWordTM Toolkit's Documentation:** <https://cwordtm.readthedocs.io>

```
In [1]: import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: # Import the Package CWordTM
import cwordtm
from cwordtm import *
```

## 1. Utility Features

```
In [3]: # Load the whole Bible (Chinese Union Version)
bible = "cuv.csv"
cdf = util.load_word(bible, info=True)
```

Loading file 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\cuv.csv' ...

Dataset Information:  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 31102 entries, 0 to 31101  
Data columns (total 9 columns):  
 # Column Non-Null Count Dtype  
---  
 0 book 31102 non-null object  
 1 book\_no 31102 non-null int64  
 2 chapter 31102 non-null int64  
 3 verse 31102 non-null int64  
 4 text 31102 non-null object  
 5 testament 31102 non-null int64  
 6 category 31102 non-null object  
 7 cat 31102 non-null object  
 8 cat\_no 31102 non-null int64  
dtypes: int64(5), object(4)  
memory usage: 2.1+ MB

## Extract Partial Scripture

```
In [4]: # Extract OT Scripture
otdf = util.extract(cdf, testament=0)
otdf.head()
```

	book	book_no	chapter	verse	text	testament	category	cat	cat_no
0	Gen	1	1	1	起初 神創造天地。	0	Torah	tor	0
1	Gen	1	1	2	地是空虛混沌·淵面黑暗· 神的靈運行在水面上。	0	Torah	tor	0
2	Gen	1	1	3	神說、要有光、就有了光。	0	Torah	tor	0
3	Gen	1	1	4	神看光是好的、就把光暗分開了。	0	Torah	tor	0
4	Gen	1	1	5	神稱光為晝、稱暗為夜· 有晚上、有早晨、這是頭一日。	0	Torah	tor	0

```
In [5]: # Extract Gospels (The first four book in NT)
gos = util.extract(cdf, category='gos')
gos.head()
```

						text	testament	category	cat	cat_no	
Out[5]:	book	book_no	chapter	verse							
23145	Mat	40	1	1	亞伯拉罕的後裔、大衛的子孫、耶穌基督的家譜・〔後裔子孫原文都作兒子下同〕			1	Gospel	gos	5
23146	Mat	40	1	2	亞伯拉罕生以撒・以撒生雅各・雅各生猶大和他的弟兄・			1	Gospel	gos	5
23147	Mat	40	1	3	猶大從他瑪氏生法勒斯和謝拉・法勒斯生希斯崙・希斯崙生亞蘭・			1	Gospel	gos	5
23148	Mat	40	1	4	亞蘭生亞米拿達・亞米拿達生拿順・拿順生撒門・			1	Gospel	gos	5
23149	Mat	40	1	5	撒門從喇合氏生波阿斯・波阿斯從路得氏生俄備得・俄備得生耶西・			1	Gospel	gos	5

```
In [6]: # Extract the Book of Romans
rom = util.extract(cdf, book='Rom')
rom.head()
```

Out[6]:	book	book_no	chapter	verse		text	testament	category	cat	cat_no	
27930	Rom	45	1	1	耶穌基督的僕人保羅、奉召為使徒、特派傳神的福音・			1	Pauline Epistles	pau	7
27931	Rom	45	1	2	這福音是神從前藉眾先知、在聖經上所應許的、			1	Pauline Epistles	pau	7
27932	Rom	45	1	3	論到他兒子我主耶穌基督・按肉體說、是從大衛後裔生的、			1	Pauline Epistles	pau	7
27933	Rom	45	1	4	按聖善的靈說、因從死裡復活、以大能顯明是神的兒子。			1	Pauline Epistles	pau	7
27934	Rom	45	1	5	我們從他受了恩惠、並使徒的職分、在萬國之中叫人為他的名信服真道・			1	Pauline Epistles	pau	7

```
In [7]: # Extract Verse 28 of Chapter 8 of Romans
rom8_28 = util.extract(cdf, book='Rom', chapter=8, verse=28)
rom8_28.text
```

Out[7]: 28143 我們曉得萬事都互相效力、叫愛神的人得益處、就是按他旨意被召的人。  
Name: text, dtype: object

```
In [8]: # Another way to extract Rom 8:28
rom8_28 = util.extract2(cdf, filter='Rom 8:28')
rom8_28.text
```

Out[8]: 28143 我們曉得萬事都互相效力、叫愛神的人得益處、就是按他旨意被召的人。  
Name: text, dtype: object

```
In [9]: # Extract Rom 8:1~10
rom8_1_10 = util.extract2(cdf, filter='Rom 8:1-10')
list(rom8_1_10.text)
```

Out[9]: ['如今那些在基督耶穌裡的、就不定罪了。',  
'因為賜生命聖靈的律、在基督耶穌裡釋放了我、使我脫離罪和死的律了。',  
'律法既因肉體軟弱、有所不能行的、\u3000神就差遣自己的兒子、成為罪身的形狀、作了贖罪祭、在肉體中定了罪案。',  
'使律法的義、成就在我們這不隨從肉體、只隨從聖靈的人身上。',  
'因為隨從肉體的人、體貼肉體的事、隨從聖靈的人、體貼聖靈的事。',  
'體貼肉體的就是死、體貼聖靈的乃是生命平安。',  
'原來體貼肉體的、就是與\u3000神為仇、因為不服\u3000神的律法、也是不能服。',  
'而且屬肉體的人、不能得\u3000神的喜歡。',  
'如果\u3000神的靈住在你們心裡、你們就不屬肉體、乃屬聖靈了、人若沒有基督的靈、就不是屬基督的。',  
'基督若在你們心裡、身體就因罪而死、心靈卻因義而活。']

## 2. Text Visualization - Word Cloud

```
In [10]: # Extract the NT Scripture for Word Cloud
text_list = util.get_text_list(cdf[cdf.testament==1]) # Load New Testament Scripture
text_list[:10]
```

Out[10]: ['亞伯拉罕的後裔、大衛的子孫、耶穌基督的家譜・〔後裔子孫原文都作兒子下同〕',  
'亞伯拉罕生以撒・以撒生雅各・雅各生猶大和他的弟兄・',  
'猶大從他瑪氏生法勒斯和謝拉・法勒斯生希斯崙・希斯崙生亞蘭・',  
'亞蘭生亞米拿達・亞米拿達生拿順・拿順生撒門・',  
'撒門從喇合氏生波阿斯・波阿斯從路得氏生俄備得・俄備得生耶西・',  
'耶西生大衛王。大衛從烏利亞的妻子生所羅門・',  
'所羅門生羅波安・羅波安生亞比雅・亞比雅生亞撒・',  
'亞撒生約沙法・約沙法生約蘭・約蘭生烏西亞・',  
'烏西亞生約坦・約坦生亞哈斯・亞哈斯生希西家・',  
'希西家生瑪拿西・瑪拿西生亞們・亞們生約西亞・']

```
In [11]: # White background with no image mask
viz.chi_wordcloud(text_list)
```

```
Building prefix dict from the default dictionary ...
Loading model from cache C:\Users\User\AppData\Local\Temp\jieba.cache
Loading Chinese vocabulary 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
Loading model cost 1.662 seconds.
Prefix dict has been built successfully.
Building prefix dict from C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
Dumping model to file cache C:\Users\User\AppData\Local\Temp\jieba.ufaf52121053d30f6b6740fa1422773b4.cache
Loading model cost 3.446 seconds.
Prefix dict has been built successfully.
C:\Dev\Anaconda3\envs\aiml\lib\site-packages\wordcloud\wordcloud.py:106: MatplotlibDeprecationWarning: The get_cmap
function was deprecated in Matplotlib 3.7 and will be removed two minor releases later. Use ``matplotlib.colormaps[n
ame]`` or ``matplotlib.colormaps.get_cmap(obj)`` instead.
    self.colormap = plt.cm.get_cmap(colormap)
```



```
In [12]: # Use internal image mask  
viz.chi_wordcloud(text_list, bg='black', image=1)
```

```
Building prefix dict from C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
Loading model from cache C:\Users\User\AppData\Local\Temp\jieba.ufaf52121053d30f6b6740fa1422773b4.cache
Loading Chinese vocabulary 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
Loading model cost 2.850 seconds.
Prefix dict has been built successfully.
C:\Dev\Anaconda3\envs\aiml\lib\site-packages\wordcloud\wordcloud.py:106: MatplotlibDeprecationWarning: The get_cmap
function was deprecated in Matplotlib 3.7 and will be removed two minor releases later. Use ``matplotlib.colormaps[n
ame]`` or ``matplotlib.colormaps.get_cmap(obj)`` instead.
    self.colormap = plt.cm.get_cmap(colormap)
```



### 3. Text Summarization

```
In [13]: # Extract and summarize Chapter 8 of Romans
rom8 = util.extract2(cdf, 'Rom 8')
ta.summary_chi(rom8)
```

```
Building prefix dict from C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
Loading model from cache C:\Users\User\AppData\Local\Temp\jieba.ufaf52121053d30f6b6740fa1422773b4.cache
Loading Chinese vocabulary 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
```

Out[13]: ['因為隨從肉體的人、體貼肉體的事、隨從聖靈的人、體貼聖靈的事，  
而且屬肉體的人、不能得神的喜樂'，  
'如果神的靈住在你們心裡、你們就不屬肉體、乃屬聖靈了：人若沒有基督的靈、就不是屬基督的']

## 4. Pivot Table

## Show Bible Scripture Statistics through a Pivot Table

```
In [14]: util.set_rows()  
pivot.stat(cdf, chi=True)
```

Out[14]:

		chapter	verse	text
category	book_no	book		
Torah	1	Gen	50	1533
	2	Exo	40	1213
	3	Lev	27	859
	4	Num	36	1288
	5	Deu	34	959
<b>Sub-Total</b>			187	5852
OT History	6	Jos	24	658
	7	Jug	21	618
	8	Rut	4	85
	9	1Sa	31	810
	10	2Sa	24	695
	11	1Ki	22	816
	12	2Ki	25	719
	13	1Ch	29	942
	14	2Ch	36	822
	15	Ezr	10	280
	16	Neh	13	406
	17	Est	10	167
<b>Sub-Total</b>			249	7018
Ketuvim	18	Job	42	1070
	19	Psm	150	2461
	20	Pro	31	915
	21	Ecc	12	222
	22	Son	8	117
<b>Sub-Total</b>			243	4785
Major Prophets	23	Isa	66	1292
	24	Jer	52	1364
	25	Lam	5	154
	26	Eze	48	1273
	27	Dan	12	357
<b>Sub-Total</b>			183	4440
Minor Prophets	28	Hos	14	197
	29	Joe	3	73
	30	Amo	9	146
	31	Oba	1	21
	32	Jon	4	48
	33	Mic	7	105
	34	Nah	3	47
	35	Hab	3	56
	36	Zep	3	53
	37	Hag	2	38
	38	Zec	14	211
	39	Mal	4	55
<b>Sub-Total</b>			67	1050
Gospel	40	Mat	28	1071
	41	Mak	16	678
	42	Luk	24	1151

category	book_no	book	chapter	verse	text
			43	Jhn	21 878 27378
<b>Sub-Total</b>			89	3778	116160
<b>NT History</b>	44	Act	28	1007	33762
<b>Sub-Total</b>			28	1007	33762
<b>Pauline Epistles</b>	45	Rom	16	433	14365
	46	1Co	16	437	14154
	47	2Co	13	257	9260
	48	Gal	6	149	4817
	49	Eph	6	155	4591
	50	Phl	4	104	3408
	51	Col	4	95	3312
	52	1Ts	5	89	2983
	53	2Ts	3	47	1624
	54	1Ti	6	113	3579
	55	2Ti	4	83	2639
	56	Tit	3	46	1544
	57	Phm	1	25	717
<b>Sub-Total</b>			87	2033	66993
<b>General Epistles</b>	58	Heb	13	303	10428
	59	Jas	5	108	3534
	60	1Pe	5	105	3905
	61	2Pe	3	61	2343
	62	1Jn	5	105	3779
	63	2Jn	1	13	463
	64	3Jn	1	15	478
	65	Jud	1	25	1030
<b>Sub-Total</b>			34	735	25960
<b>Apocalypse</b>	66	Rev	22	404	15606
<b>Sub-Total</b>			22	404	15606
<b>Total</b>			1189	31102	1062163

In [15]: `util.bible_cat_info()`

	category	cat	book_list	nbooks
0	Torah	tor	Gen Exo Lev Num Deu	5
1	OT History	oth	Jos Jug Rut 1Sa 2Sa 1Ki 2Ki 1Ch 2Ch Ezr Neh Est	12
2	Ketuvim	ket	Job Psm Pro Ecc Son	5
3	Major Prophets	map	Isa Jer Lam Eze Dan	5
4	Minor Prophets	mip	Hos Joe Amo Oba Jon Mic Nah Hab Zep Hag Zec Mal	12
5	Gospel	gos	Mat Mak Luk Jhn	4
6	NT History	nth	Act	1
7	Pauline Epistles	pau	Rom 1Co 2Co Gal Eph Phl Col 1Ts 2Ts 1Ti 2Ti Ti...	13
8	General Epistles	epi	Heb Jas 1Pe 2Pe 1Jn 2Jn 3Jn Jud	8
9	Apocalypse	apo	Rev	1

## 5. OT Quotes

Identify Cited Sources in OT Scripture for some NT verses

```
In [16]: rom10 = util.extract2(cdf, 'Rom 10')
quot.show_quot(rom10, lang='chi')

Building prefix dict from C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\dictionary\dict.txt.big.txt ...
Loading Chinese vocabulary 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\bible_vocab.txt' ...
Loading file 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\cuv.csv' ...

*** Quote(s) from OT
Loading file 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\book_categories.csv' ...
( 1) 羅 10:5 摩西寫著說、『人若行那出於律法的義、就必因此活著。』
Loading model from cache C:\Users\User\AppData\Local\Temp\jieba.ufaf52121053d30f6b6740fa1422773b4.cache
Loading model cost 1.924 seconds.
Prefix dict has been built successfully.
( 2) 羅 10:6 惟有出於信心的義如此說、『你不要心裡說、誰要升到天上去呢、就是要領下基督來。
( 3) 羅 10:8 他到底怎麼說呢、他說、『這道離你不遠、正在你口裡、在你心裡。』就是我們所傳信主的道。
( 4) 羅 10:11 經上說、『凡信他的人、必不至於羞愧。』
( 5) 羅 10:13 因為『凡求告主名的、就必得救。』
( 6) 羅 10:15 若沒有奉差遣、怎能傳道呢、如經上所記、『報福音傳喜信的人、他們的腳蹤何等佳美。』
( 7) 羅 10:16 只是人沒有都聽從福音、因為以賽亞說、『主阿、我們所傳的有誰信呢。』
    -> 0.6723 賽 53:1 我們所傳的、〔或作所傳與我們的〕有誰信呢、耶和華的膀臂向誰顯露呢。
( 8) 羅 10:18 但我說、人沒有聽見麼、誠然聽見了、『他們的聲音傳遍天下、他們的言語傳到地極。』
( 9) 羅 10:19 我再說、以色列人不知道麼、先有摩西說、『我要用那不成子民的、惹動你們的憤恨、我要用那無知的民、觸動你們的怒氣。』
    -> 0.5403 申 32:21 他們以那不算為神的、觸動我的憤恨、以虛無的神、惹了我的怒氣、我也要以那不成子民的、觸動他們的憤恨、
以愚昧的國民、惹了他們的怒氣。
(10) 羅 10:20 又有以賽亞放膽說、『沒有尋找我的、我叫他們遇見、沒有訪問我的、我向他們顯現。』
    -> 0.6651 賽 65:1 素來沒有訪問我的、現在求問我、沒有尋找我的、我叫他們遇見、沒有稱為我名下的、我對他們說、我在這裡、
我在這裡。
(11) 羅 10:21 至於以色列人、他說、『我整天伸手招呼那悖逆頂嘴的百姓。』
    -> 0.6086 賽 65:2 我整天伸手招呼那悖逆的百姓、他們隨自己的意念行不善之道。
** End of Quotes **
```

## 6. Topic Modeling

```
In [17]: import warnings
warnings.filterwarnings('ignore')
```

### LDA Modeling

```
In [18]: # Build an LDA Model on the OT Scripture
lda = tm.lda_process(bible, cat=1, chi=True, eval=True, timing=True)
```

```
Loading Bible 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\cuv.csv' ...
Corpus loaded!
Text preprocessed!
Text trained!
You may execute the following commands to show the visualization:
    import pyLDAvis
    pyLDAvis.display(lda.vis_data)

Visualization prepared!

Topics from LDA Model:
[(0,
  '0.009*"約瑟" + 0.009*"肘" + 0.009*"祭司" + 0.008*"潔淨" + 0.006*"作" + 0.006*"法老" + '
  '0.005*"說" + 0.005*"沒有" + 0.005*"行" + 0.004*"長"' ),
(1,
  '0.018*"耶和華" + 0.015*"必" + 0.009*"神" + 0.006*"中" + 0.006*"嘆" + 0.005*"不可" + '
  '0.005*"惡人" + 0.005*"智慧" + 0.004*"說" + 0.004*"好像"' ),
(2,
  '0.007*"肘" + 0.006*"說" + 0.005*"以色列" + 0.005*"軍隊" + 0.005*"殿" + 0.004*"見" + '
  '0.004*"兒子" + 0.004*"共有" + 0.004*"被數" + 0.004*"基路伯"' ),
(3,
  '0.027*"說" + 0.026*"大衛" + 0.020*"掃羅" + 0.016*"耶和華" + 0.011*"以色列" + '
  '0.010*"非利士人" + 0.006*"兒子" + 0.005*"王" + 0.005*"沒有" + 0.005*"約拿單"' ),
(4,
  '0.016*"神" + 0.016*"耶和華" + 0.012*"求" + 0.008*"子孫" + 0.007*"必" + 0.007*"智慧" + '
  '0.006*"願" + 0.006*"中" + 0.006*"惡人" + 0.005*"說"' ),
(5,
  '0.038*"耶和華" + 0.014*"不可" + 0.014*"摩西" + 0.013*"以色列" + 0.011*"祭司" + '
  '0.010*"作" + 0.009*"亞倫" + 0.008*"吩咐" + 0.008*"說" + 0.007*"兒子"' ),
(6,
  '0.027*"耶和華" + 0.025*"說" + 0.018*"王" + 0.013*"以色列" + 0.009*"兒子" + 0.008*"必" + '
  '0.007*"猶大" + 0.007*"耶路撒冷" + 0.006*"神" + 0.005*"巴比倫"' ),
(7,
  '0.053*"兒子" + 0.012*"生" + 0.008*"子孫" + 0.007*"族長" + 0.007*"以色列" + 0.007*"中" + '
  '0.006*"弟兄" + 0.005*"族" + 0.004*"住" + 0.004*"支派"' ),
(8,
  '0.029*"說" + 0.016*"雅各" + 0.014*"耶和華" + 0.007*"兒子" + 0.007*"潔淨" + 0.007*"以撒" + '
  '0.007*"父親" + 0.006*"嘆" + 0.006*"神" + 0.006*"亞伯拉罕"' ),
(9,
  '0.041*"耶和華" + 0.037*"說" + 0.014*"以色列" + 0.012*"神" + 0.008*"必" + 0.006*"沒有" + '
  '0.005*"百姓" + 0.005*"面前" + 0.005*"中" + 0.004*"作"' )]
```

Model Evaluation Scores:

Coherence: 0.3602963255407491

Perplexity: -8.201667878702157

Topic diversity: 0.0013003907101656665

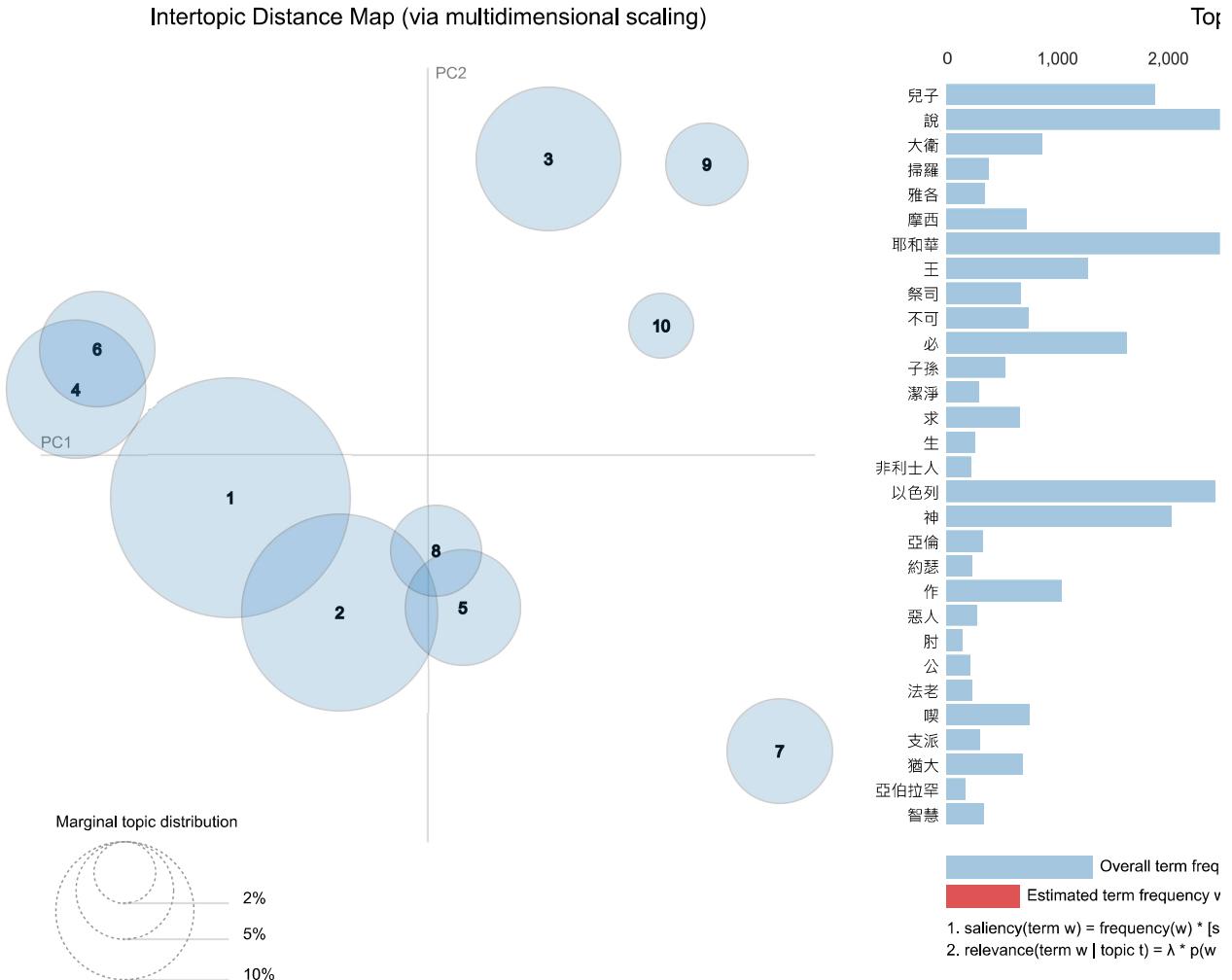
Topic size distribution: 0.0028653295128939827

Finished 'lda\_process' in 97.3219 secs

```
In [19]: # Show LDA Model Visualization
import pyLDAvis
pyLDAvis.display(lda.vis_data)
```

Out[19]: Selected Topic: 0 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric (2)

 $\lambda = 1$ 

## BERTopic Modeling

```
# Build a BERTopic Model on the OT Scripture
btm = tm.btm_process(bible, cat=1, chi=True, eval=True, timing=True)
```

```
Loading Bible 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\cuv.csv' ...
Corpus loaded!
Chinese text preprocessed!
Text trained!
```

Topics from BERTopic Model:

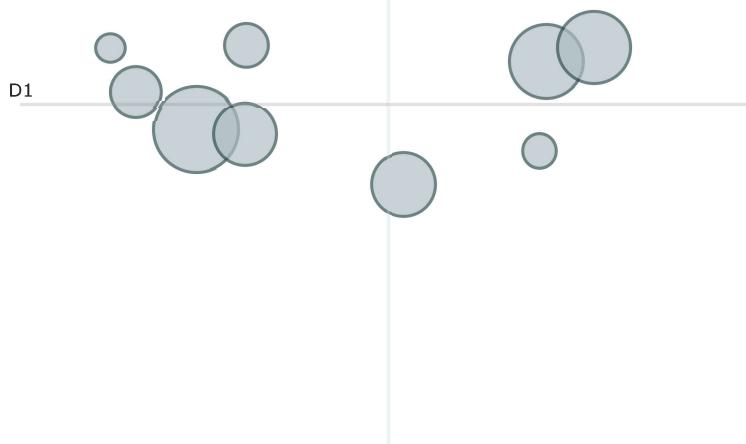
- Topic 0: 耶和華 | 不可 | 以色列 | 兒子 | 雅各 | 埃及 | 面前 | 沒有 | 地上 | 慈愛
- Topic 1: 兒子 | 耶和華 | 以色列 | 巴比倫 | 子孫 | 耶路撒冷 | 猶大 | 萬軍 | 大衛 | 耶利米
- Topic 2: 耶和華 | 以色列 | 摩西 | 支派 | 百姓 | 沒有 | 吩咐 | 兒子 | 屬城 | 面前
- Topic 3: 祭司 | 耶和華 | 兒子 | 以色列 | 潔淨 | 利未人 | 子孫 | 一隻 | 燭祭 | 亞倫
- Topic 4: 兒子 | 父親 | 大衛 | 僕人 | 一個 | 雅各 | 女兒 | 亞伯拉罕 | 以掃 | 沒有
- Topic 5: 耶和華 | 讚美 | 永遠 | 之詩 | 大衛 | 的詩 | 伶長 | 上行 | 交與 | 惡人
- Topic 6: 智慧 | 惡人 | 愚昧 | 義人 | 知識 | 正直 | 言語 | 日光 | 之下 | 虛空
- Topic 7: 耶和華 | 獻給 | 素祭 | 摩西 | 以色列 | 基路伯 | 羊羔 | 活物 | 百姓 | 不可
- Topic 8: 歌頌 | 我要 | 耶和華 | 歡呼 | 稱謝 | 詩歌 | 稱讚 | 萬民 | 救恩 | 細拉

Model Evaluation Scores:  
Coherence: 0.348970526130946

BERTopic Model Visualization:

## Intertopic Distance Map

D2



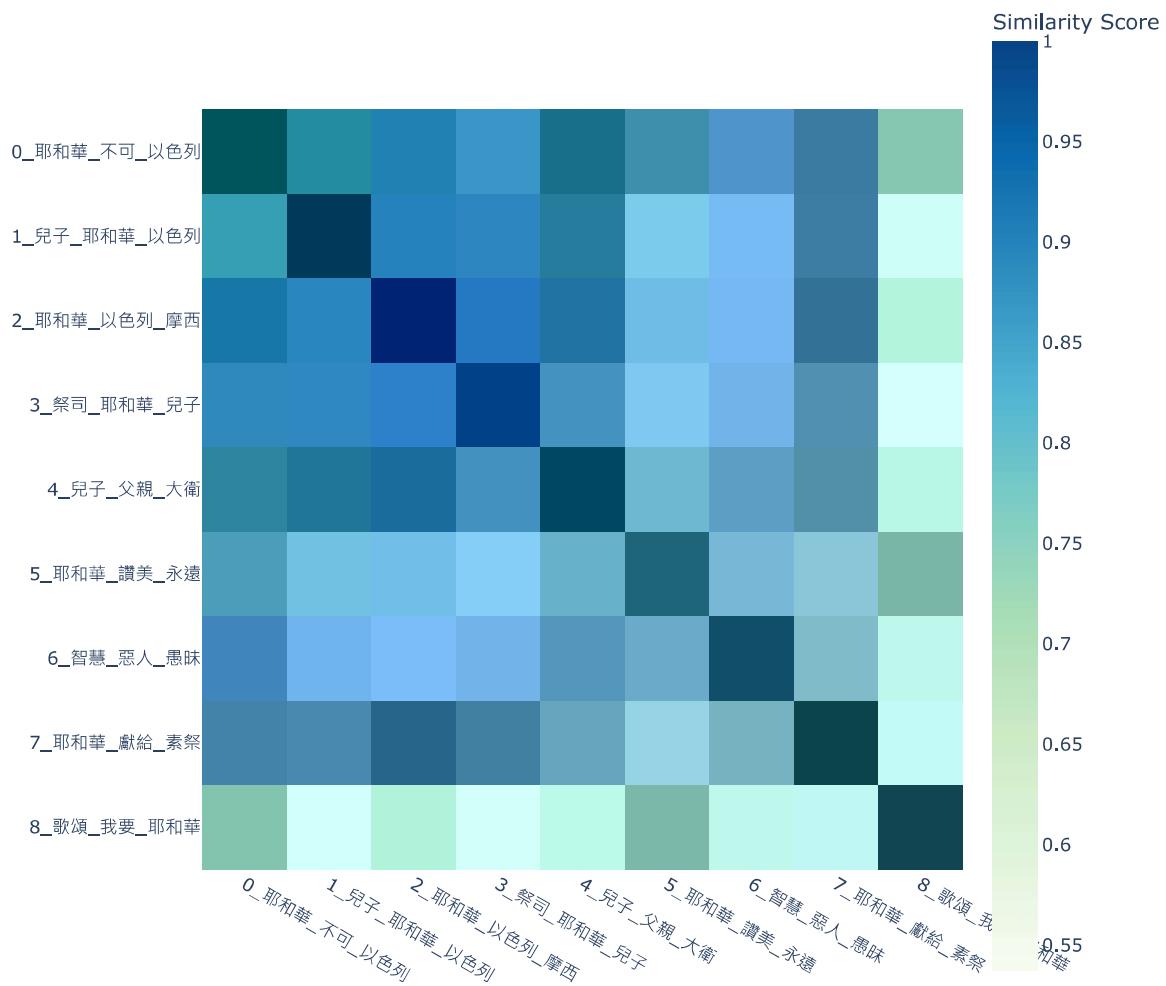
Topic 0



### Topic Word Scores



## Similarity Matrix



Finished 'btm\_process' in 232.2511 secs  
 <Figure size 1000x800 with 0 Axes>

## NMF Modeling

```
In [21]: nmf = tm.nmf_process(bible, num_topics=8, cat=1, chi=True, eval=True, timing=True)
```

Loading Bible 'C:\Dev\Anaconda3\envs\aiml\lib\site-packages\cwordtm\data\cuv.csv' ...

Corpus loaded!

Text preprocessed!

Text trained!

Topics-Words from NMF Model:

Topic 1:

兒子 (0.036530)  
以色列 (0.019223)  
子孫 (0.015938)  
族 (0.012946)  
作 (0.010833)  
屬 (0.009531)  
中 (0.008765)  
掃羅 (0.008661)  
支派 (0.008193)  
公 (0.006838)

Topic 2:

說 (0.027964)  
耶和華 (0.024642)  
必 (0.015475)  
神 (0.007957)  
王 (0.007388)  
不可 (0.006311)  
巴比倫 (0.006227)  
雅各 (0.005762)  
事 (0.005199)  
子孫 (0.004780)

Topic 3:

兒子 (0.016690)  
作 (0.007659)  
事 (0.007644)  
猶大 (0.007391)  
雅各 (0.006581)  
行 (0.005641)  
肘 (0.005616)  
神 (0.004874)  
求 (0.004418)  
耶路撒冷 (0.004362)

Topic 4:

神 (0.015200)  
王 (0.010441)  
說 (0.007960)  
求 (0.007413)  
面前 (0.005244)  
願 (0.004557)  
見 (0.004401)  
起來 (0.003956)  
不可 (0.003924)  
住 (0.003890)

Topic 5:

說 (0.079232)  
耶和華 (0.050291)  
沒有 (0.008814)  
喫 (0.008743)  
大衛 (0.008730)  
神 (0.006840)  
以色列 (0.006692)  
百姓 (0.006522)  
祭司 (0.006267)  
告訴 (0.006164)

Topic 6:

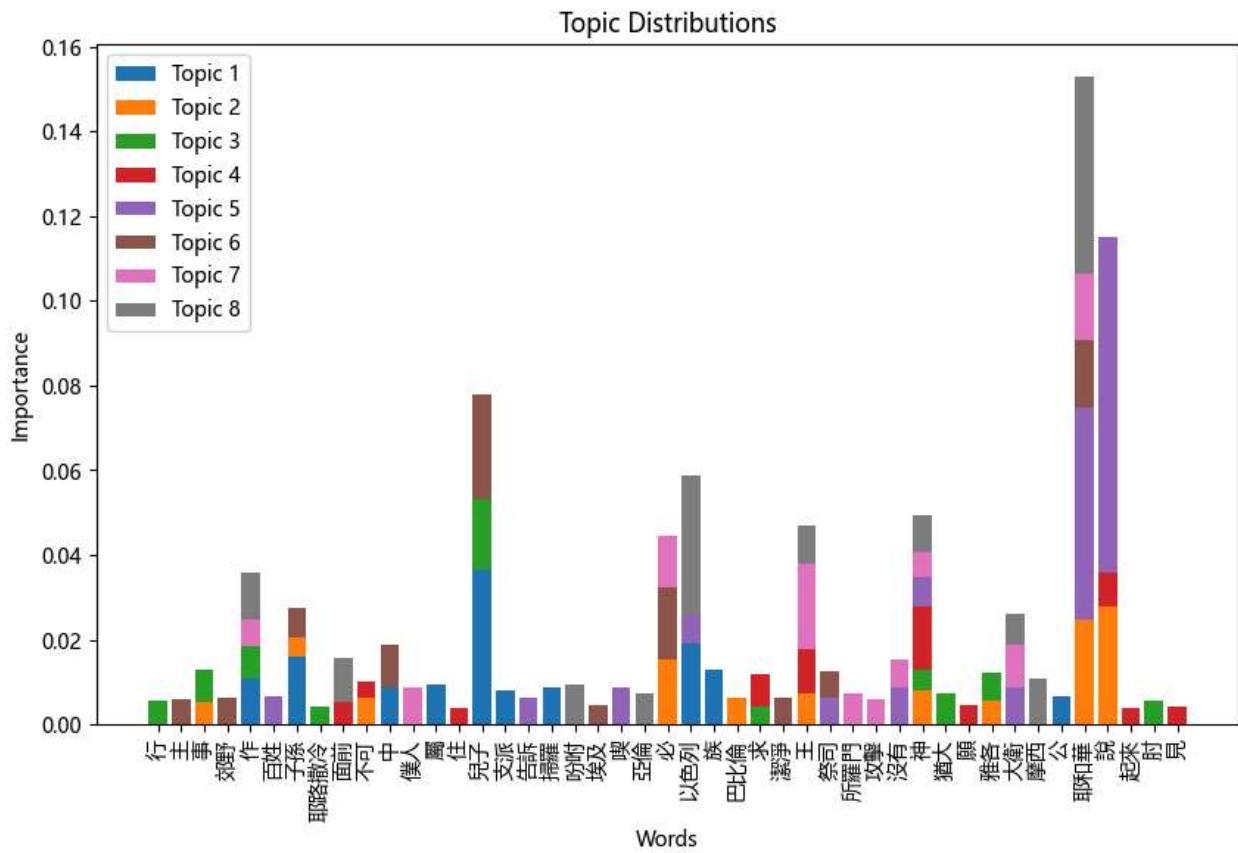
兒子 (0.024629)  
必 (0.016736)  
耶和華 (0.015751)  
中 (0.010130)  
子孫 (0.006968)  
潔淨 (0.006490)  
祭司 (0.006260)  
郊野 (0.006252)  
主 (0.005835)  
埃及 (0.004716)

Topic 7:

王 (0.019971)  
耶和華 (0.015606)  
必 (0.012357)  
大衛 (0.010021)  
僕人 (0.008593)  
所羅門 (0.007420)  
沒有 (0.006531)

作 (0.006422)  
攻撃 (0.005982)  
神 (0.005733)

Topic 8:  
耶和華 (0.046554)  
以色列 (0.033024)  
作 (0.011025)  
摩西 (0.010887)  
面前 (0.010425)  
吩咐 (0.009510)  
王 (0.009022)  
神 (0.008915)  
亞倫 (0.007391)  
大衛 (0.007323)



Model Evaluation Scores:  
Coherence: 0.3049749209034524  
Topic diversity: 0.0011440087670163942  
Topic size distribution: 0.001848001848001848

Finished 'nmf process' in 50.4699 secs

## Save NMF Model

```
In [22]: nmf.save("models/nmf_cuv")
```

NMF model has been saved: 'models/nmf\_cuv\_model.gensim' and 'models/nmf\_cuv\_dictionary.gensim'

## Load NMF Model

```
In [23]: nmf2 = tm.NMF("", nmf.num_topics)
nmf2.model, nmf2.dictionary = nmf2.load("models/nmf_cuv")
nmf2.show_topics_words()
```

Topics-Words from NMF Model:

Topic 1:

兒子 (0.036530)  
以色列 (0.019223)  
子孫 (0.015938)  
族 (0.012946)  
作 (0.010833)  
屬 (0.009531)  
中 (0.008765)  
掃羅 (0.008661)  
支派 (0.008193)  
公 (0.006838)

Topic 2:

說 (0.027964)  
耶和華 (0.024642)  
必 (0.015475)  
神 (0.007957)  
王 (0.007388)  
不可 (0.006311)  
巴比倫 (0.006227)  
雅各 (0.005762)  
事 (0.005199)  
子孫 (0.004780)

Topic 3:

兒子 (0.016690)  
作 (0.007659)  
事 (0.007644)  
猶大 (0.007391)  
雅各 (0.006581)  
行 (0.005641)  
肘 (0.005616)  
神 (0.004874)  
求 (0.004418)  
耶路撒冷 (0.004362)

Topic 4:

神 (0.015200)  
王 (0.010441)  
說 (0.007960)  
求 (0.007413)  
面前 (0.005244)  
願 (0.004557)  
見 (0.004401)  
起來 (0.003956)  
不可 (0.003924)  
住 (0.003890)

Topic 5:

說 (0.079232)  
耶和華 (0.050291)  
沒有 (0.008814)  
喫 (0.008743)  
大衛 (0.008730)  
神 (0.006840)  
以色列 (0.006692)  
百姓 (0.006522)  
祭司 (0.006267)  
告訴 (0.006164)

Topic 6:

兒子 (0.024629)  
必 (0.016736)  
耶和華 (0.015751)  
中 (0.010130)  
子孫 (0.006968)  
潔淨 (0.006490)  
祭司 (0.006260)  
郊野 (0.006252)  
主 (0.005835)  
埃及 (0.004716)

Topic 7:

王 (0.019971)  
耶和華 (0.015606)  
必 (0.012357)  
大衛 (0.010021)  
僕人 (0.008593)  
所羅門 (0.007420)  
沒有 (0.006531)  
作 (0.006422)  
攻擊 (0.005982)  
神 (0.005733)

Topic 8:

耶和華 (0.046554)  
以色列 (0.033024)  
作 (0.011025)  
摩西 (0.010887)  
面前 (0.010425)  
吩咐 (0.009510)  
王 (0.009022)  
神 (0.008915)  
亞倫 (0.007391)  
大衛 (0.007323)