



# DEDU—佛教對讀文本編輯平台的建置與發展

The Establishment and Development of DEDU – The Buddhist Parallel Text Editing Platform

洪振洲 Jen-Jou Hung  
Professor, Dharma Drum Institute of Liberal Arts

# Parallel Reading

- 同時閱讀(比較)不同文本內的相關內容，是正確理解佛陀教理的重要方法
- Place the corresponding sections of multiple texts together to read content and compare texts.
- An important way to understand the teaching of Buddha.

**Text-cluster:**  
別譯雜阿含: bza001    BZA English: 別譯雜阿含英譯    雜阿含: za1062    SN: SN,II,278 (Sujāto)

**bza001**  
   
如是我聞：一時佛在彌緋羅國菴婆羅園。  
  
爾時尊者善生初始出家，剃除鬚髮，來詣佛所，頂禮佛足，在一面坐。佛告諸比丘：  
  
「此族姓子善生有二種端嚴：一、容貌\*瑣偉，天姿挺特。二、能剃除鬚髮，身服法衣，深信家法會歸無常，出家學道；盡諸煩惱，具足無漏，心得解脫，慧得解脫；身證無為，生死永盡，梵行已立，不受後有。」  
  
佛說是已，即說偈言：  
  
「比丘常寂定 除欲離生死  
住最後邊身 能破於魔軍

**Sujāta is praised**  
Translation of BZA 001. First version published in Buddhist Studies Review vol. 23-1 (2006).  
   
  
Thus have I heard, once, the Buddha stayed in a mango grove in the country of Mithilā.  
  
At this time the Venerable Sujāta had just gone forth and cut off his hair. He came to the Buddha, payed homage by touching the Buddha's feet and sat down to one side. The Buddha told the other monks: 'This son of a good family, Sujāta, is splendid in two ways: First, his appearance is of extraordinary beauty, his features exceptionally handsome. Second, he was able to cut off his hair and cover his body with the robe. Feeling deeply that family life was all too ephemeral, he went forth

**tn0099-1062**  
   
如是我聞：一時佛住舍衛國祇樹給孤獨園。  
  
時有尊者善生，新剃鬚髮，著袈裟衣，正信非家，出家學道，來詣佛所，稽首佛足，退坐一面。爾時世尊告諸比丘：  
  
「諸比丘！當知此善生善男子有二處端嚴：一者、剃除鬚髮，著袈裟衣，正信非家，出家學道。二者、盡諸有漏，無漏，心解脫，慧解脫，現法自知作證：我生已盡，梵行已立，所作已作，自知不受後有。」  
  
爾時世尊即說偈言：  
  
「寂靜盡諸漏 比丘莊嚴好

# Parallel Reading in Buddhist Studies

- Scenario 1

- 佛學研究者，通常藉由比對不同語言的經典紀錄，找出佛陀教法的真正義理，或找出翻譯上的差異。
- Researchers often find **differences** between documents by **comparing the contents of multilingual documents**, thereby restoring and understanding the true meaning of Buddhist scriptures.
- 1. Buddha's teachings mainly spread among disciples through **oral transmission** when Buddhism emerged in India.
- 2. These **oral contents** were later **written down** in varieties.
- 3. **Different translators** may make the translation divorced from the original meaning.

# Different Versions of The Same Scripture

T0224

東漢支婁迦讖譯《道行般若經》  
C.E. 179

T0225

吳支謙譯《大明度經》  
C.E. 222–257

T0226

曇摩婢及竺佛念譯《摩訶般若鈔經》  
Late 4<sup>th</sup> Century

Lk.429a10f.

《摩訶般若波羅蜜·難問品》  
第一

爾時，釋提桓因與四萬天子相隨俱來，共會，坐。四天王與天上二萬天子相隨來，共會，坐。梵迦夷天與萬天子相隨來，共會，坐。首(←梵)多會天與五千天子相隨來，共會，坐。諸天子宿命有德，光明巍巍。持佛威神，持佛力，諸天子光明徹照。

釋提桓因白須菩提言：「賢者須菩提！是若干千萬天子大會，欲聽須菩提說般若波羅蜜。云何菩薩於般若波羅蜜中住？」

須菩提語釋提桓因言：

ZQ.482b1f.

《大明度經》卷第二《天帝釋問品》第二

爾時，帝釋與四萬天子、  
四天王與二萬天子、  
梵眾  
天與萬天子、  
梵輔天與  
五千天子俱皆來會  
坐。諸天子宿命功德，光  
耀巍巍。持佛神  
力，明徹照。

釋問善業言：「是諸天子大會，欲聽說智度無極。云何闍士[·大士]於大明中立乎？」

善業曰：

Zfn.511c14f.

《摩訶般若波羅蜜·問品》  
第二

爾時，釋提桓因與四萬天人俱來，共會。四天王與天上二萬天人俱來，共會。梵迦夷天與[一]萬天人俱來，共會。首陀譚天與五千天人俱來，共會。諸天宿命有德，光明巍巍。得佛威神力，諸天光明悉不復見。

釋提桓因白須菩提：「若干萬千天人大會，欲聽須菩提說般若波羅蜜。云何菩薩[·摩訶薩]於般若波羅蜜中住？」

須菩提語釋提桓因：

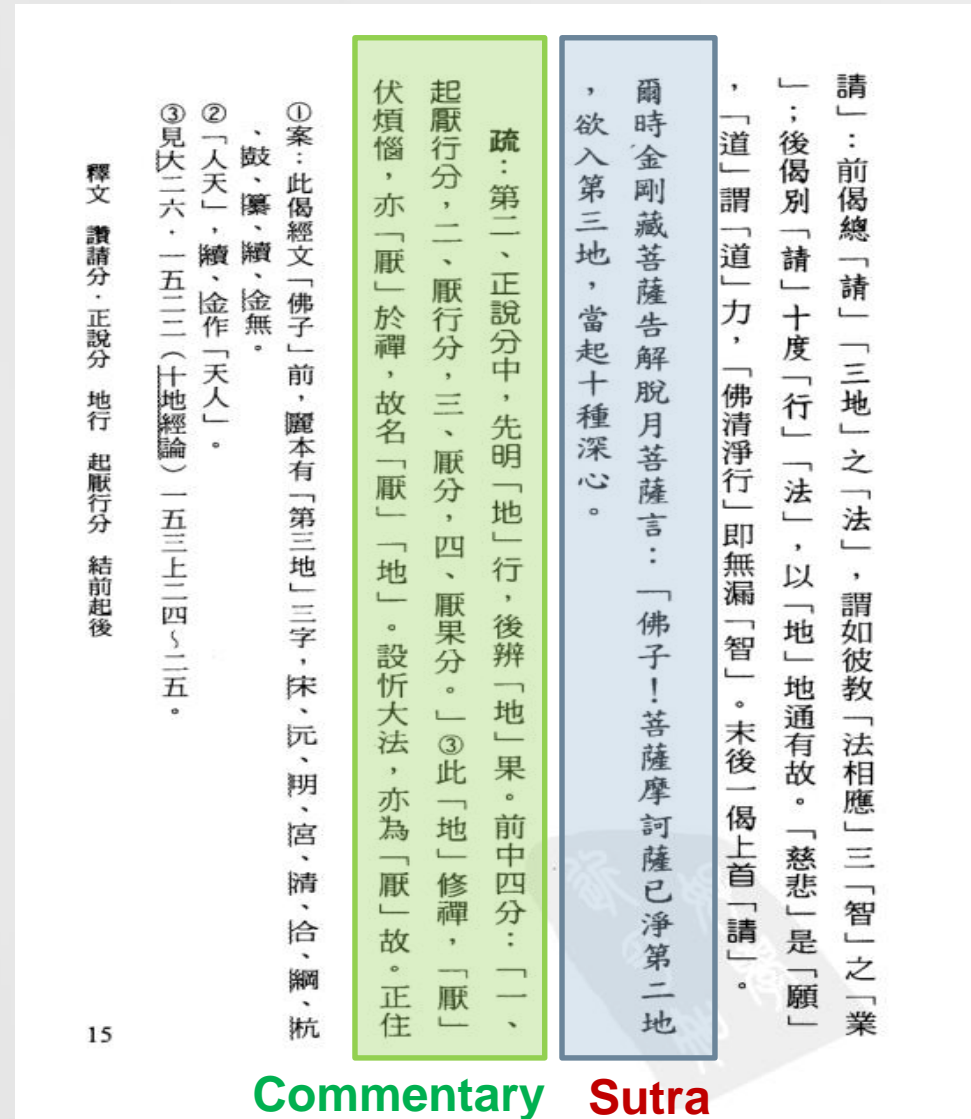
# Parallel Reading in Buddhist Studies

- Scenario 2

- 另外**注疏**(**解釋經文**的內容)的內容，也是通常必須跟經文內容一起閱讀，或相互比對
- Many **commentaries** are used to explain the complicated contents of Buddhist texts.
- Those commentaries are also possibly written in **different languages** like **Sanskrit** and **Tibetan**. When there are many different translations of the Buddhist scriptures and their commentaries, it will be more complicated to compare and read.

# Sutra Text and Commentary

- 右圖是「新脩華嚴經疏抄」的實際範例。
- Picture in right hand side shows an example of a printed edition
- The Avatamsaka Sutra, accompanied by commentaries..





# Buddhist Parallel Corpora

- The SAT Daizōkyō Text Database
  - English and Chinese

BDK Paralell corpus

Sentence Corresponding

1. ————告諸比丘。是德藏菩薩。次  
-> Addressing the monks, he said:...

2. ———爾時佛告諸比丘。———  
-> Then the Buddha addressed the ...

3. ————佛告諸比丘我最後弟子而今已  
-> The Buddha said to the bhikṣus...

4. 於天人大眾中告諸比丘。———  
-> "After the Buddha had perfecte...

5. 一佛遙見女來告諸比丘衆此女極端正能留  
-> When the Buddha saw the lady c...

6. ————告諸比丘我以佛眼見是迦葉  
-> I tell you, O monks, That I s...

7. 子心之所念。告諸比丘。———  
-> Then the

Up-and-Down  
Arrangement

神通。威力弘大。盡知過去無數劫事。以能  
善解法性故知。亦以諸天來語故知。佛時  
頌曰

|       |                      |
|-------|----------------------|
| 比丘集法堂 | 講說賢聖論                |
| 如來處靜室 | 天耳盡聞知                |
| 佛日光普照 | 分別法界義                |
| 亦知過去事 | 三佛般泥洹                |
| 名號姓種族 | 受生分亦知                |
| 隨彼之處所 | 淨眼皆記之                |
| 諸天大威力 | 容貌甚端嚴                |
| 亦來啓告我 | 三佛般 <sup>18</sup> 泥洹 |
| 記生名號姓 | 哀 <sup>19</sup> 鸞音盡知 |
| 無上天人尊 | 記於過去佛                |

又告諸比丘。汝等欲聞如來識宿命智知  
於過去諸佛因緣不。我當說之。時諸比丘  
白<sup>20</sup>言。世尊。今正是時。願樂欲聞。善哉世尊。  
以時講說。當奉行之。佛告諸比丘。諦聽諦  
聽。善思念之。吾當爲汝分別解說。時諸比  
丘受教而聽

# Buddhist Parallel Corpora

- SuttaCentral
  - Pali, Sanskrit, English, and Chinese

| The Prime Net — Bhikkhu Sujato  |  | 🔍 | ⋮ |
|---|--|---|---|
| resentful, bitter, and exasperated.   | saṅghassa vā avaṇṇaṃ bhāseyyuṃ, tatra tumhehi na āghāto na appaccayo na cetaso anabhiraddhi karaṇīyā.  |   |   |
| You'll get angry and upset, which would be an obstacle for you alone.   | Mamaṃ vā, bhikkhave, pare avaṇṇaṃ bhāseyyuṃ, dhammassa vā avaṇṇaṃ bhāseyyuṃ, saṅghassa vā avaṇṇaṃ bhāseyyuṃ, tatra ce tumhe assatha kupitā vā anattamanā vā, tumhaṃ yevassa tena antarāyo.   |   |   |
| If others were to criticize me, the teaching, or the Saṅgha, and you got angry and upset, would you be able to understand whether they spoke well or poorly?"   | Mamaṃ vā, bhikkhave, pare avaṇṇaṃ bhāseyyuṃ, dhammassa vā avaṇṇaṃ bhāseyyuṃ, saṅghassa vā avaṇṇaṃ bhāseyyuṃ, tatra ce tumhe assatha kupitā vā anattamanā vā, api nu tumhe paresaṃ subhāsitaṃ dubbhāsitaṃ ājāneyyāthā"ti?                                   |   |   |
| "No, sir."  | "No hetam, bhante".  |   |   |
| "If others criticize me, the teaching, or the Saṅgha, you should explain that what is untrue is in fact untrue:<br>'This is why that's untrue, this is why that's false. There's no such thing in us, it's not found among us.' | "Mamaṃ vā, bhikkhave, pare avaṇṇaṃ bhāseyyuṃ, dhammassa vā avaṇṇaṃ bhāseyyuṃ, saṅghassa vā avaṇṇaṃ bhāseyyuṃ, tatra tumhehi abhūtaṃ abhūtaṃ nibbēthetabbaṃ: 'itipetaṃ abhūtaṃ, itipetaṃ atacchaṃ, natthi cetam amhesu, na ca panetaṃ amhesu saṃvijjati'ti. |   |   |
| If others praise me, the teaching, or the Saṅgha, don't make yourselves thrilled, elated, and excited.  | Mamaṃ vā, bhikkhave, pare vaṇṇaṃ bhāseyyuṃ, dhammassa vā vaṇṇaṃ bhāseyyuṃ, saṃghassa vā vaṇṇaṃ bhāseyyuṃ, tatra tumhehi na ānando na somanassaṃ na cetaso uppilāvitattaṃ karaṇīyaṃ.  |   |   |
| You'll get thrilled, elated, and excited, which would be an obstacle for you alone.   | Mamaṃ vā, bhikkhave, pare vaṇṇaṃ bhāseyyuṃ, dhammassa vā vaṇṇaṃ bhāseyyuṃ, saṃghassa vā vaṇṇaṃ bhāseyyuṃ, tatra ce tumhe assatha ānandino sumanā uppilāvitā tumhaṃ yevassa tena antarāyo.  |   |   |

Sentence Corresponding

Side-by-Side Arrangement



# Buddhist Parallel Corpora

- A Digital Comparative Edition and Translation of the Shorter Chinese **Samyukta Āgama** (T.100)
  - Sanskrit, Chinese, English, Tibetan, and Pali

[HOME](#) [SEARCH](#) [CLUSTER PAGE](#)  001 [next](#) [TWO COLUMN VIEW](#)

**bza001 -- Sujāta is praised 善生二種端嚴**

Text-cluster:

**bza001**  
   
如是我聞：一時佛在彌締羅國 菴婆羅園。  
爾時尊者善生初始出家，剃除鬚髮，來詣佛所，頂禮佛足，在一面坐。佛告諸比丘：  
「此族姓子善生有二種端嚴：一、容貌 \* 瑣偉，天姿挺特。二、能剃除鬚髮，身服法衣，深信家法會歸無常，出家學道；盡諸煩惱，具足無漏，心得解脫，慧得解脫；身證無為，生死永盡，梵行已立，不受後有。」  
佛說是已，即說偈言：  
「比丘常寂定 除欲離生死  
住最後邊身 能破於魔軍」

**Sujāta is praised**  
Translation of BZA 001. First version published in Buddhist Studies Review vol. 23-1 (2006).  
   
Thus have I heard, once, the Buddha stayed in a mango grove in the country of Mithilā.  
At this time the Venerable Sujāta had just gone forth and cut off his hair. He came to the Buddha, payed homage by touching the Buddha's feet and sat down to one side. The Buddha told the other monks: 'This son of a good family, Sujāta, is splendid in two ways: First, his appearance is of extraordinary beauty, his features exceptionally handsome. Second, he was able to cut off his hair feeling deeply ... he went forth

**tn0099-1062**  
   
如是我聞：一時佛住舍衛國 祇樹給孤獨園。  
時有尊者善生，新剃鬚髮，著袈裟衣，正信非家，出家學道，來詣佛所，稽首佛足，退坐一面。爾時世尊告諸比丘：  
「諸比丘！當知此善生善男子有二處端嚴：一者、剃除鬚髮，著袈裟衣，正信非家，出家學道。二者、盡諸有漏，無漏，心解脫，慧解脫，現法自知作證：我生已盡，梵行已立，所作已作，自知不受後有。」  
爾時世尊即說偈言：  
「寂靜盡諸漏 比丘莊嚴好」

Side-by-Side Arrangement

Discourse Corresponding

# For A Good Parallel Reading Platform

- 好的佛教對讀系統，必須顯示多語，並妥善呈現資料間的對應。
- Able to handle **multilingual characters** correctly.
- Display **whole texts**, **corresponding blocks**, and **corresponding words** at once.
- Provide a function to switch from **side-by-side** arrangement to **up-and-down** display mode of parallel texts.
- Provide a **content index**.
- Provide a function for **selecting texts for display**.

# The Need of Parallel Corpora Creation

- 但資料對讀需從未停止，現有結果永遠難以滿足使用者。
- These databases are rich in content and convenient for public use.
- However, they cannot meet all the different needs of researchers.
- Researchers may have **their own texts** to compare.



# DEDU Editor

Create, Edit and Share. All in one Site

Create parallel text in the most intuitive way,  
and share your work with the world immediately.

START TO EDIT

DEDU 工作坊錄影

There are currently 842 registered users. A total of 1212 reading projects  
have been produced. Click [here](#) to view the public projects.



DEDU Editor  
DEDU 線上對讀文本編輯平台

<https://DEDU.dila.edu.tw>

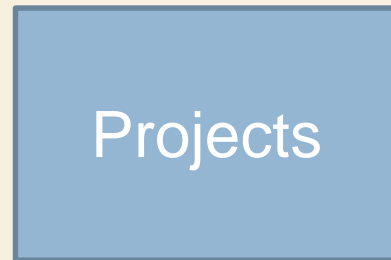
Google : 「DEDU」

# DEDU Editor Workflow



**Login**  
**DEDU**

Free For Everyone



+ Add New Project

Edit Parallel  
Document

Collaboration &  
Sharing

Online Display

# Create New Project

編輯 - 藏文彌陀經對讀專案 [Sukhavativyuha Sutra Comparison]

檔案 編輯 外部資料來源 設定 說明

+ 新增 ✎ 編輯 ✕ 移除

藏文彌陀經對讀專案

對應區塊1

對應區塊2

Content  
Index  
Structure

文章標題 ✎

對應區塊1 ✎

Doc 1  
Tibetan

對應區塊2 ✎

文章標題 ✎

對應區塊1 ✎

Doc 2  
Tibetan

對應區塊2 ✎

→  
English

文章標題 ✎

對應區塊1 ✎

Doc 3  
Chinese

對應區塊2 ✎

文章標題 ✎

對應區塊1 ✎

Doc 4  
Chinese

對應區塊2 ✎

→  
English

← 回管理介面 預覽成果



# Block Editing

編輯 - 藏文彌陀經對讀專案 [Sukhavativyuha Sutra Comparison]

檔案 編輯 外部資料來源 設定 說明

← 回管理介面 預覽成果

+ 新增 編輯 × 移除

藏文彌陀經對讀專案

對應區塊1  
對應區塊2

| 文章標題  | 文章標題  | 文章標題  | 文章標題  |
|-------|-------|-------|-------|
| 對應區塊1 | 對應區塊1 | 對應區塊1 | 對應區塊1 |
| 對應區塊2 | 對應區塊2 | 對應區塊2 | 對應區塊2 |

Start Edit  
This Block

# Block Editing

## 編輯 - 藏文彌陀經對讀專案 [Sukhavativyuha Sutra Comparison]

檔案 編輯 外部資料來源 設定 說明

← 回管理介面 預覽成果

+ 新增 ✎ 編輯 ✕ 移除

藏文彌陀經對讀專案

對應區塊1

對應區塊2

| བོད་སྐད་ཀྱི་   | Translation  | 中文                            | Translation  |
|--|--|-------------------------------|--|
| 對應區塊1  | 對應區塊1  | 對應區塊1                         | 對應區塊1  |
| <p>  ཁྱེ་ལ་སྐད་དུ། ལྷ་མྱ་ལ་བ་ཏི་<br/>སྐྱུ་ཏུ་རྒྱ་མ་མ་རྒྱ་ན་སྐྱུ་ཏ།</p> <p>བོད་སྐད་དུ་འཕགས་པ་བདེ་བ་<br/>ཅན་གྱི་བོད་པ་ཞེས་བྱ་བ་ཐེག་<br/>པ་ཆེན་པོའི་མདོ། སངས་རྒྱལ་<br/>དང་བྱང་ཆུབ་སེམས་དཔའ་<br/>ཐམས་ཅད་ལ་བྱུག་འཆལ་ལོ།</p> | <p>In the language of India:<br/><i>Āryasukhāvatīvyūha-<br/>nāmamahāyānasūtra</i></p> <p>In the language of Tibet:<br/>The Noble Mahāyāna Sūtra<br/>“The Display of the Pure<br/>Land of Sukhāvati”</p> <p>Homage to all the buddhas<br/>and bodhisattvas.</p> | <p>佛说阿弥陀经<br/>姚秦三藏法师鸠摩罗什译</p> | <p>The Buddha Speaks of<br/>Amitabha Sūtra</p> <p>Translated into Chinese<br/>during the Yao-Qin<br/>Dynasty by Tripiṭaka<br/>Master Kumārajīva of<br/>Kucha</p> |

# Side-by-Side Table Display

藏文彌陀經對讀專案



對讀結構

選擇文本

匯出

取得分享連結

● 表格式 ● 條列式

● 中文 ● English

● 專案說明

| བོད་སྐད།   | Translation   | 中文                             | Translation  |
|--|---|--------------------------------|--|
| 對應區塊1<br>  ཁྱེ་གར་སྐད་དུ། ལྷ་ཅུ་སྤུ་ལ་བཏི་བྱུ་ཏ་རྩ་མ་<br>མ་རྩ་ལྷ་ན་སྤྱུ་ཏ།<br>བོད་སྐད་དུ་འཕགས་པ་བདེ་བ་ཅན་གྱི་བཀོད་<br>པ་ཞེས་བྱ་བ་ཐེག་པ་ཆེན་པོའི་མདོ། སངས་<br>རྒྱུ་དང་བྱང་ཆུབ་སེམས་དཔའ་ཐམས་ཅད་<br>ལ་ཕུག་འཆལ་ལོ། | 對應區塊1<br>In the language of India:<br>Āryasukhāvativyūha-<br>nāmamahāyānasūtra<br>In the language of Tibet:<br>The Noble Mahāyāna Sūtra<br>“The Display of the Pure Land<br>of Sukhāvati”<br>Homage to all the buddhas<br>and bodhisattvas. | 對應區塊1<br>佛说阿弥陀经<br>姚秦三藏法师鸠摩罗什译 | 對應區塊1<br>The Buddha Speaks of<br>Amitabha Sūtra<br><br>Translated into Chinese during<br>the Yao-Qin Dynasty by<br>Tripiṭaka Master Kumārajīva of<br>Kucha |
| 對應區塊2  | 對應區塊2   | 對應區塊2                          | 對應區塊2  |

# Up-and-Down List Display

藏文彌陀經對讀專案



對讀結構

選擇文本

匯出

取得分享連結

表格式

條列式

中文 English

專案說明

## 對應區塊3

བོད་སྐད་།

དེ་ནས་བཙམ་ལྟན་འདས་ཀྱིས་ཆོ་དང་ལྟན་པ་ཤུ་རིའི་བྱ་ལ་བཀའ་རྒྱལ་པ། ཤུ་རིའི་བྱ། སངས་རྒྱུས་ཀྱི་ཞིང་འདི་ནས་ཐུབ་ཕྱོགས་ལོགས་སུ་སངས་རྒྱུས་ཀྱི་ཞིང་བྱེ་བ་ཁག་ཁྲིག་འབྲུམ་འདས་པ་  
ན་འཇིག་རྟེན་གྱི་ཁམས་བདེ་བ་ཅན་ཞེས་བྱ་བ་ཡོད་དེ། དེ་ན་དེ་བཞིན་གཤེགས་པ་དགྲ་བཙམ་པ་ཡང་དག་པར་རྫོགས་པའི་སངས་རྒྱུས་ཆོ་དཔག་མེད་ཅེས་བྱ་བ་བཞུགས་ཏེ་འཆོ་ཞིང་ག་ཞེས་  
ལ་ཆོས་ཀྱང་སྟོན་ཏོ།

### Translation

2. At that time, the Buddha said to Elder Śāriputra, "To the west from here, beyond a hundred thousand million Buddha lands, there is a world called Great Bliss. In that land there is a Buddha named Amitābha. He is there now, teaching the Dharma.

### 中文

尔时，佛告长老舍利弗：“从是西方，过十万亿佛土，有世界名曰极乐，其土有佛，号阿弥陀，今现在说法。

# Application of DEDU

- 教學、理解、閱讀比較上，DEDU 的成果都可以提供協助
- Teaching and Researching
  - DEDU is a flexible tool to navigate and compare the scripture and the commentary together for teaching.

論曰：世間、聖教說有我、法，但由假立，非實有性。<sup>[注]</sup>我謂主宰，法謂軌持。<sup>[注]</sup>

## [述記]

述曰。我如主宰者。如國之主有自在故。及如輔宰能割斷故。有自在力。及割斷力義同我故。或主是我體。宰是我所。或主如我體。宰如我用。法謂軌持。軌謂軌範可生物解。持謂住持不捨自相。一體有無對。二自性差別對。三有為無為對。四先陳後說對。前唯有體。後亦通無。瑜伽論五十二。說意不壞法現前。無亦名法。今者相分必有。似無名無。前是共相。後是自相。前唯有為。後通無為。前唱者名持。後唱者名軌。合有四對名為軌持。

## [概要]

「我謂主宰，法謂軌持」：「主」是俱生我，無分別故；「宰」是分別我，有割斷故。「主」是第七我；「宰」是第六我。「主」是世間我，能作、受故；「宰」是聖教我，依用辨故。並疏為五解。聖教法名「軌」，依用辨故；世間法名「持」，軌實自體能自持故。並疏為五。此中皆依增上義說，四解通依世間、聖教，第五別配。

## [了義燈]

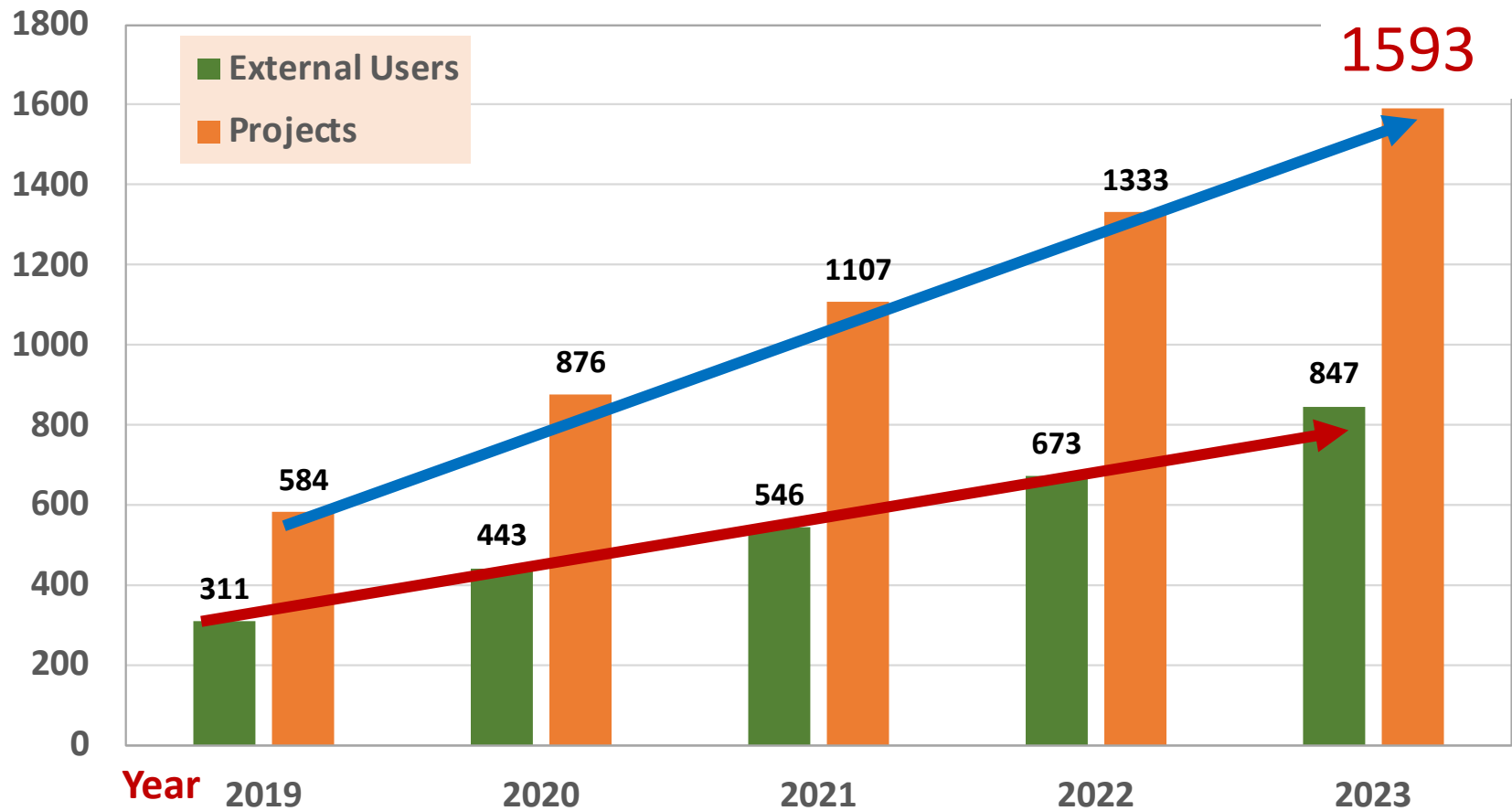
法謂軌持。本疏釋云：軌為軌範，可生物解。持謂任持，不捨自相。此意能有屈曲軌範，表彰差別生物，解者隨增名軌。不為屈曲表彰，但直生解，即名為持。持自性故，緣無為等，生解之時，通得名法。不是定說，無及真如，不能生解，故名為持。今於總一法名，分為軌持。不爾，軌持二言何別？又以軌持，疏有四對，不是如次配前軌持，但隨所應，為其四對，准義配取，并《樞要》中，聖教世間對。《樞要》解云：「此中皆依增上義說。前四解通世間聖教，第五別配。」若解此釋，西明引立，無亦名軌，如《因明論》立我為無，亦得名軌，徒設劬勞。有解不得本疏主意，救此義，云：以我體無為疎所緣緣，以無體故，不能生解。親緣有體，生解名軌，亦是徒施。有云：復引《瑜伽論》，軌通有法、<sup>[動]</sup>無法、有為、無為、若有、若無俱名論軌，亦應自息。《要集》復云：有說四對，辨持、軌異、無為、無法，唯持非軌，義即難尋。復引《楞伽》真智緣如。乃至云後得智，緣遠從本質，生解名論軌，亦無有失者。此不異前，何煩廣說？又引《瑜伽》五十二云：「由二種緣，諸識得生。乃至意、法。」答：由能軌持諸五識身所不行義故。佛世尊假說名法，五識不行，通為、無為，既欲證無及如名軌。復引通語，如無軌言。不解本釋，徒為證據。應尋前解，經論無違。

## ② 釋第二句

彼二俱有種種相轉：<sup>[注]</sup>「我種種相」謂有情、命者等，<sup>[注]</sup>預流、一來等。<sup>[注]</sup>「法種種相」謂實、德、業等；<sup>[注]</sup>蘊、處、界等。<sup>[注]</sup>「轉」謂隨緣施設有異。<sup>[注]</sup>

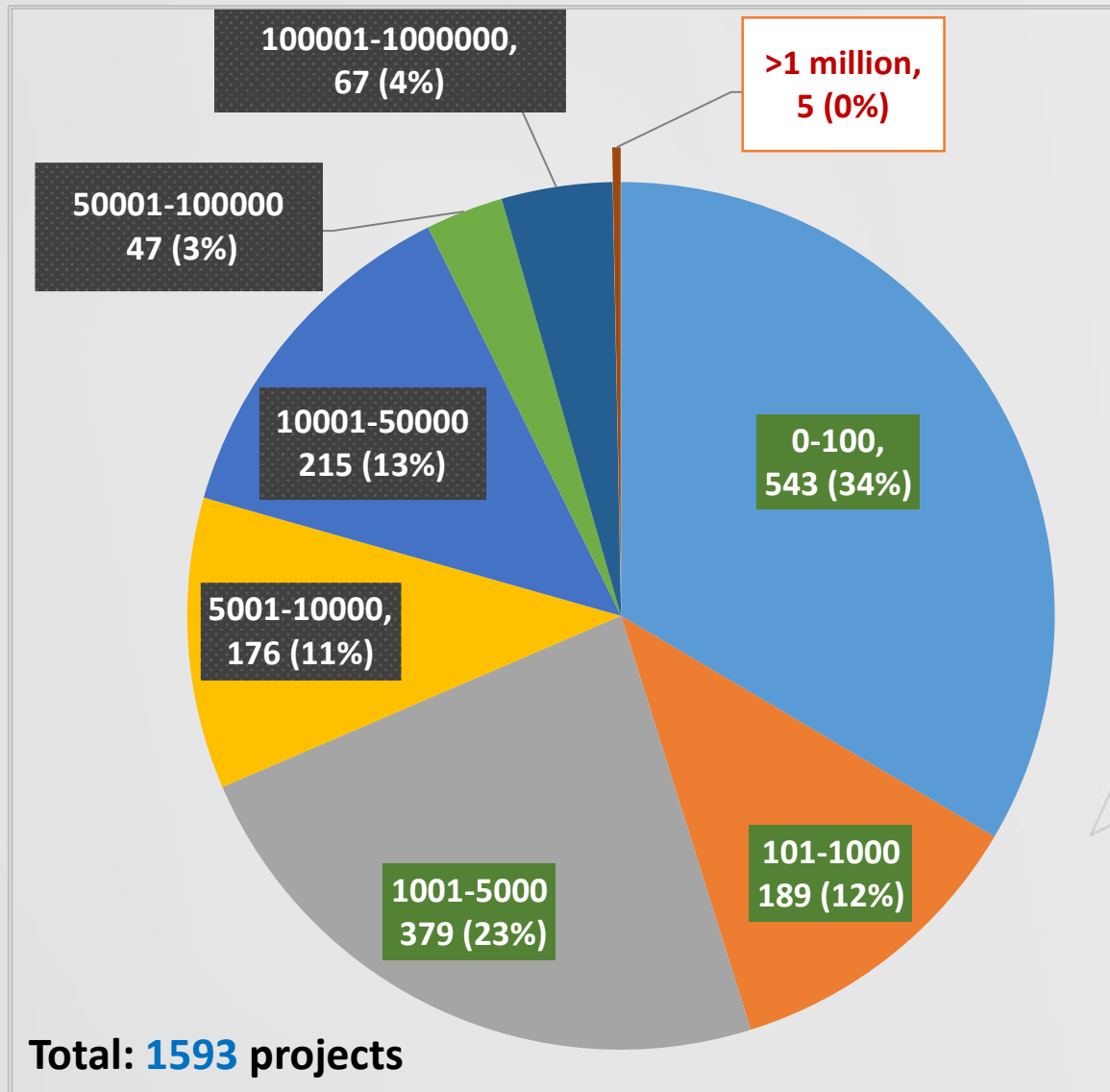
# DEDU Usage

Total Number of External Users and Projects (By Year)





# Count By # of Characters



- 69 % of Projects contain less than 5,000 characters
- 120 projects contain more than 50,000 characters
- 5 projects contain more than 1 M characters

# Language and Category

Count from 103 Open Projects

| Language        | # of Projects |
|-----------------|---------------|
| <b>Chinese</b>  | <b>90</b>     |
| <b>Sanskrit</b> | <b>47</b>     |
| <b>Tibetan</b>  | <b>47</b>     |
| <b>English</b>  | <b>41</b>     |
| <b>Pali</b>     | <b>10</b>     |
| <b>Japanese</b> | <b>10</b>     |
| <b>French</b>   | <b>3</b>      |
| <b>Others</b>   | <b>7</b>      |

| Category     | # of Projects |
|--------------|---------------|
| Buddhism     | 99            |
| Non-Buddhism | 4             |

| Category under Buddhism      | # of Projects |
|------------------------------|---------------|
| <b>Yogācāra</b>              | 15            |
| <b>Āgama</b>                 | 12            |
| <b>Ratna-kūṭa</b><br>(寶積部類)  | 11            |
| <b>Sutta-nipāta</b><br>(經集部) | 11            |
| <b>Pure Land</b>             | 8             |

# Top 10 popular projects

| Project's Title               | Views  | Characters |
|-------------------------------|--------|------------|
| 宗義寶鬘(四部宗義)                    | 42,914 | 13,078     |
| 八識規矩頌                         | 6,274  | 30,941     |
| 大乘百法明門論                       | 4,634  | 18,250     |
| 桂林古本《傷寒雜病論》理法條文               | 4,425  | 158,953    |
| 《辯中邊論・辯相品》諸譯本及梵本對照<br>( 測試版 ) | 3,981  | 84,731     |
| 百法明門論                         | 2,467  | 18,250     |
| 《攝大乘論・應知勝相品》真諦與玄奘版本比對         | 1,475  | 13,644     |
| 《攝大乘論 ( 釋 ) 》五種文本對讀           | 1,397  | 141,799    |
| 《解深密經・一切法相品》真諦與玄奘版本比對         | 1,220  | 2,543      |
| 佛說阿彌陀經 ( 含科判 )                | 1,148  | 6,399      |

# Recent Development

專注研發更容易製作對讀專案的方式

## 1. New Parallel Corpus Creating Method

- Import from **spreadsheet** format. (ex: EXCEL)
- **Fork and Enhance**: A Function to Copy an Existing Project for Easy Creation of New Parallel Corpora.

## 2. Automatic Alignment

- Use **AI approaches** to automatically divide documents into segments and then align the segments for further editing.
- **BertAlign** and **LLM**

# EXCEL Template

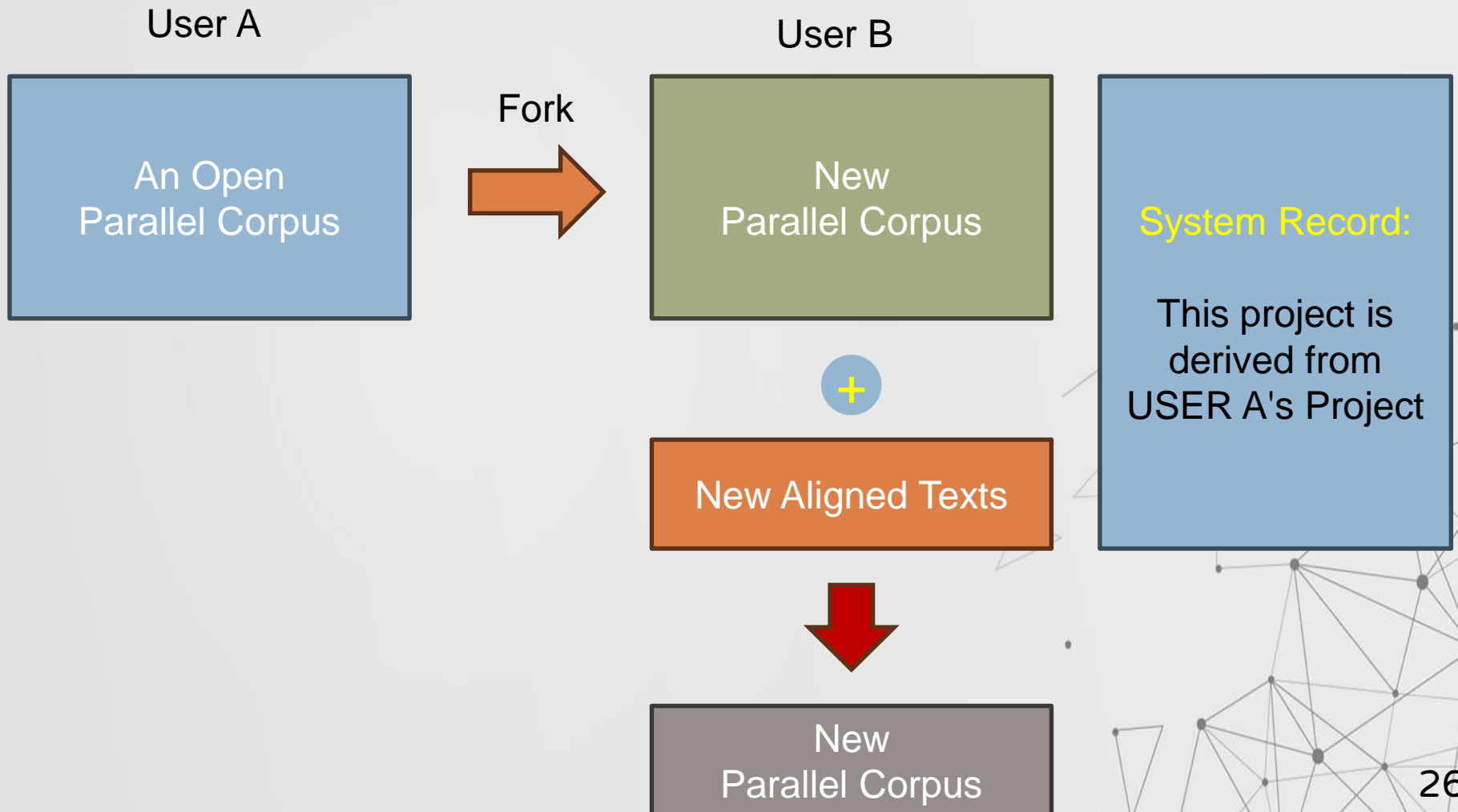
Text-1  
Title and Content

Title of Text-2  
Title and Content



|   | A          | B         |                        |   |
|---|------------|-----------|------------------------|---|
| 1 | Node Level | Node Name | 成唯識論 (唐 玄奘)            | 成唯識論 現代英文譯本   |
| 2 | #          | 節點1       | 此經不說異色心等有實自性，<br>為證不成。 | this scripture does not say that they have a<br>real nature apart from form, mind, etc., so<br>your proof is inconclusive.  |
| 3 | ##         | 節點1.1     | 非第六聲便表異體，色心之體<br>即色心故。 | The grammatical genitive relationship<br>between characteristics and dharmas does not<br>signify a different substance, because the<br>substance of form and mind is identical with<br>form and mind.                                       |
| 4 | ##         | 節點1.2     | 非能相體定異所相，勿堅相等<br>異地等故。 | You cannot clearly distinguish the substance<br>of that which can characterize from that<br>which is characterized, because the<br>characteristics of conditioned dharmas, such<br>as solidity, etc., are not different from earth,<br>etc. |

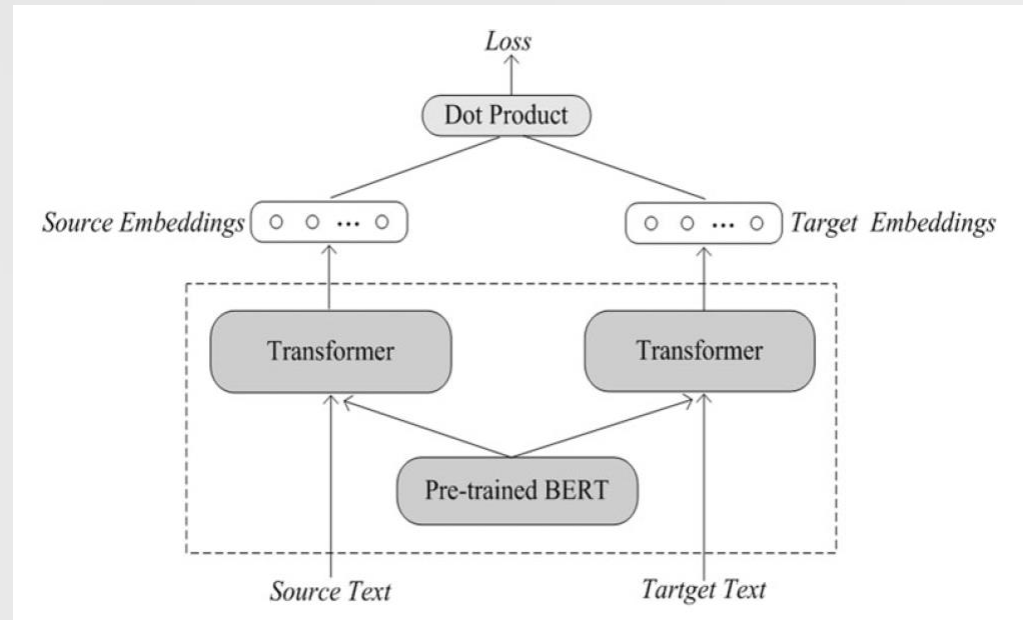
# Fork a Parallel Corpus





# Automatic Alignment (with BertAlign)

- BertAlign Algorithm
- Lei Liu, Min Zhu, Bertalign: Improved word embedding-based sentence alignment for Chinese–English parallel corpora of literary texts, *Digital Scholarship in the Humanities*, Volume 38, Issue 2, June 2023, Pages 621–634,
- Separate Texts into Sentences.
- Use LaBSE Model to Generate Sentence Embeddings.

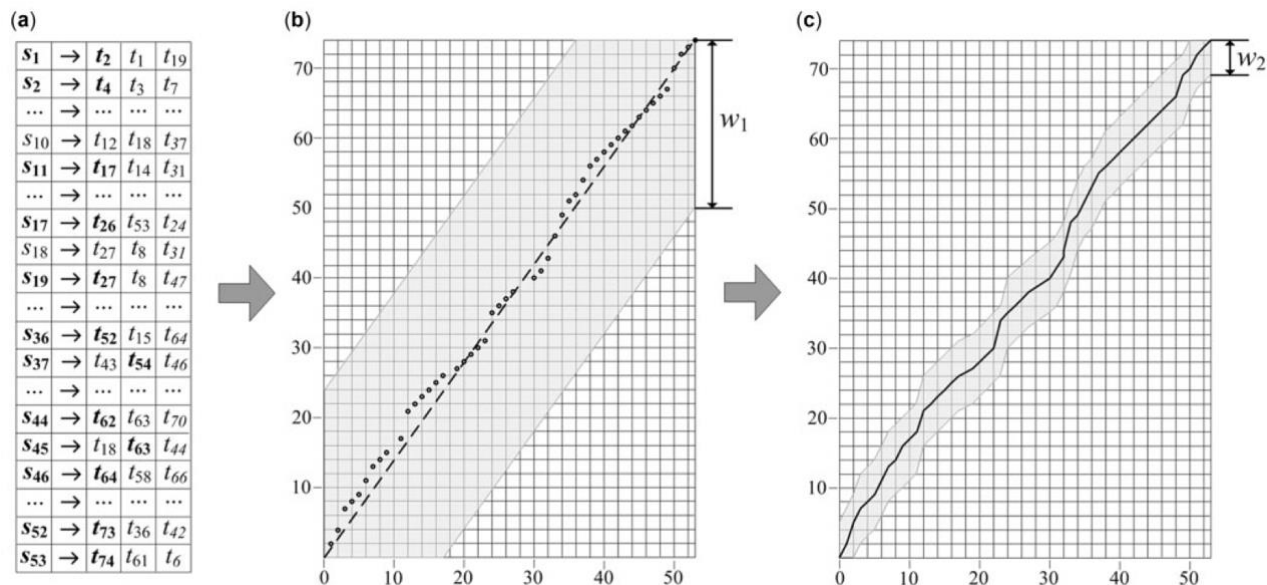


**LaBSE Model is a Dual Bert Model, specially trained for measuring the Semantical Similarity between two sentences**

# Two-step Alignment approach

Then use Dynamic Programming Algorithm to decides the best matches between sentences.

- Two-step Alignment approach (Reduce the time complexity of the DP )



Two-step alignment algorithm (a) top-k ( $k=3$ ) similarity matrix, (b) DP table for the first-step alignment, and (c) DP table for the second-step alignment

# 4 Experiments of Auto-Alignment

| Title  |   | Texts   | # of Original Sentences | # of Aligned Sentences | Accuracy    |
|--|---|---|-------------------------|------------------------|-------------|
| <b>Experiment 1</b><br>《三聚經》<br>(English Translation)    | 英 | Sutra 12 of Dirgha Agama:<br>Doctrines in Groups of Three | <b>86</b>               | <b>83</b>              | <b>100%</b> |
|  | 中 | 《長阿含經·三聚經》  | <b>90</b>               | <b>83</b>              |             |
| <b>Experiment 2</b><br>《佛說阿彌陀經》<br>(English Translation) | 英 | The Buddha Speaks of Amitābha<br>Sūtra                    | <b>31</b>               | <b>22</b>              | <b>100%</b> |
|  | 中 | 佛說阿彌陀經  | <b>22</b>               | <b>22</b>              |             |
| <b>Experiment 3</b><br>《百法明門論》<br>(Chinese Commentaries) | 中 | 《大乘百法明門論直解》玄奘譯  | <b>242</b>              | <b>201</b>             | <b>67%</b>  |
|  | 中 | 藕益沙門智旭 慈恩法師註解   | <b>372</b>              | <b>201</b>             |             |

# Future Work

- Based on our preliminary testing, AI model seems to have amazing performance for aligning original text and translation text, but it requires further verification.
- The performance of aligning text and commentaries needs improvement.
- For the DEDU user interface, we will focus on creating features that allow for quick alignment corrections, such as directing content editing, splitting one text block into two, or to merge two blocks into one.