

萬讀

**Mandoku:**

**An incubator for  
premodern Chinese**

Christian Wittern, Kyoto



一 炁生炁 炁生巽 巽生艮 艮生離 此相生也

一 夬尅革 革尅焱 焱尅器 器尅蠱 蠱尅巽 巽尅艮 艮尅坤

一 艮尅蠱 蠱尅夬 夬尅巽 巽尅離 離尅坤 坤尅艮 艮尅坤

一 艮尅困 困尅艮 艮尅巽 巽尅離 離尅坤 坤尅艮 此相尅也

輪範屬先天之學

一 夬炁一道炁尅夬者言地可以勝天而人更可勝地蓋地近而天遠而人更近於天地也他則尅爲凶

一 夬生陽道生天生五行生星斗生金玉生龍馬生日月生男女一夬炁雖分爲二而生養萬物培植二氣則一也

一 炁生陰道生地生山川生草木生金玉生飛禽生雲霧生走獸

炁生炁 炁生巽 巽生艮 艮生離 此相生也

夬尅革 革尅焱 焱尅器 器尅蠱 蠱尅巽 巽尅艮 艮尅坤

艮尅蠱 蠱尅夬 夬尅巽 巽尅離 離尅坤 坤尅艮 艮尅坤

艮尅困 困尅艮 艮尅巽 巽尅離 離尅坤 坤尅艮 此相尅也

輪範屬先天之學

夬炁一道炁尅夬者言地可以勝天而人更可勝地蓋地近而天遠而人更近於天地也他則尅爲凶

夬生陽道生天生五行生星斗生金玉生龍馬生日月生男女一夬炁雖分爲二而生養萬物培植二氣則一也

炁生陰道生地生山川生草木生金玉生飛禽生雲霧生走獸



# The problems

- Premodern Chinese texts do not use modern standardized orthography
- They do not contain punctuation and only minimal hints to the structure
- A modern view of the text (content based OHCO) is difficult to establish



# Stages in the life of a text

- We consider the life of a digital text in different stages, with quite different requirements:
  - (1) Birth: First cut transcription (outsourced?)
  - (2) Incubator: Establishment of the character content and hierarchical structures
  - (3) Education: Exploration of the content, translation, annotation etc.
  - (4) Adulthood: Publication, public collaboration



# Birth of a text

- Provisional transcription of the characters, according to one source
- Placeholders are used for not transcribable characters to mark the position of these characters
- line breaks are retained and page numbers indicated



# In the incubator

- The sequential position of a character in the text is the main reference point
- This allows to assign different characters (original/regularized) forms to the same position
- A character database helps in establishing and maintaining such relationships both on the level of the individual text instance, text collections and generally valid relations between characters
- Collation with other redactions of the text



# Growing up

- Transform from the position and page/line oriented model of the incubator to the TEI text model
- Punctuation and / or a more fine grained analysis structural analysis is added
- Adding of additional information about textual features, named entities etc



# Publication

- Side by side view, similar to the incubator
- Additional views with a layout corresponding to the different purposes of the view
- Support for translating and research on the text
- Collaborative framework



# from the Incubator

<text><body><div><p><milestone unit="dummy"/>  
<lb n="JY006\_01p001a\_01"/><milestone unit="h1"/>爕元<app><rdg wit="#CK-KZ">贊</rdg><rdg wit="#YP-Z">贊</rdg></app>運孚佑帝君重刻大洞玉經敘  
<lb n="JY006\_01p001a\_02"/><milestone unit="p"/>道可得而言乎不可得而言也道可得而秘乎不可得而秘也  
<lb n="JY006\_01p001a\_03"/>道不容終秘又難以言傳二者之所不居而經典之所由<g>[宣-旦+且]</g>示  
<lb n="JY006\_01p001a\_04"/>也經之旨有在於言字之內者由淺以<choice><sic>厯</sic><corr>歷</corr></choice>深有在於糸微之際  
<lb n="JY006\_01p001a\_05"/>者由表以窺裏得其解者挾經之心司道之契即經中之語言  
<lb n="JY006\_01p001a\_06"/>文字符籙咒訣均可不存得言忘象筌蹄可棄然知之者鮮矣  
<lb n="JY006\_01p001a\_07"/>故古來高真上聖有不得不由解識而入則經之爲予以津梁  
<lb n="JY006\_01p001a\_08"/>使登彼岸其度濟之益莫可名言至於消災救劫拔厄延生乃  
<lb n="JY006\_01p001a\_09"/>靈通所蘊萬古不磨其神妙固出於言思擬議之外者也茲編  
<lb n="JY006\_01p001a\_10"/>乃  
<lb n="JY006\_01p001b\_01"/><milestone unit="p"/>元始天王爲神禹治水鎮伏諸瀆而傳厥後一授於天台再授  
<lb n="JY006\_01p001b\_02"/>於七曲奉而行之皆證妙果不特爲救護小劫也今夏南水盛  
<lb n="JY006\_01p001b\_03"/>漲侍壇超子曾爲河帥故欲其積誠感召以應天和幸覺子等  
<lb n="JY006\_01p001b\_04"/>同志贊勸發愿刊<!--Norm 刊-->刻大洞玉經以弭患於未然志良善也夫此  
<lb n="JY006\_01p001b\_05"/>經之足以消弭水厄已著於  
<lb n="JY006\_01p001b\_06"/><milestone unit="p"/>文帝前敘徵信已久自不必更述故爲闡傳道之初旨更爲述  
<lb n="JY006\_01p001b\_07"/>復刊<!--Norm 刊-->之由以付諸子尙望讀是經者窺測<choice><sic>玄</sic><corr>玄</corr></choice>奧其大義不外以  
<lb n="JY006\_01p001b\_08"/>人合天其究竟歸於以天從人學道者知人所固有之天復無  
<lb n="JY006\_01p001b\_09"/>始以來自然之天天外之天既<!--Norm 既-->得而本心之天全矣  
<lb n="JY006\_01p002a\_01"/><milestone unit="h1"/>九天開化文昌帝君自敘  
<lb n="JY006\_01p002a\_02"/><milestone unit="p"/>余之居事畎澮忽於鋤下得一金像詢之父老曰乃元始天王  
<lb n="JY006\_01p002a\_03"/>相昔夏禹治水冶金爲像以鎮方岳余香火奉事未嘗少<choice><sic><g>[↑\*鮮]</g></sic><corr>懈</corr></choice><!--Add Table variant-->因



# to “real” TEI XML

```
22 <text><body>
24 <lb n="JY006_01p001a_01"/><div><head>燮元<app><rdg wit="#CK-KZ">贊</rdg><rdg wit="#YP-Z">贊</rdg></app>
24 運孚佑帝君重刻大洞玉經敘
25 </head><lb n="JY006_01p001a_02"/><p/><p>道可得而言乎不可得而言也道可得而秘乎不可得而秘也
26 <lb n="JY006_01p001a_03"/>道不容終秘又難以言傳二者之所不居而經典之所由<g>[宣-旦+且]</g>示
27 <lb n="JY006_01p001a_04"/>也經之旨有在於言字之內者由淺以<choice><sic>厯</sic><corr>歷</corr></choice>深
27 有在於糸微之際
28 <lb n="JY006_01p001a_05"/>者由表以窺裏得其解者挾經之心司道之契即經中之語言
28 <lb n="JY006_01p001a_06"/>文字符籙咒訣均可不存得言忘象筌蹄可棄然知之者鮮矣
29 <lb n="JY006_01p001a_07"/>故古來高真上聖有不得不由解識而入則經之爲予以津梁
30 <lb n="JY006_01p001a_08"/>使登彼岸其度濟之益莫可名言至於消災救劫拔厄延生乃
31 <lb n="JY006_01p001a_09"/>靈通所蘊萬古不磨其神妙固出於言思擬議之外者也茲編
33 <lb n="JY006_01p001a_10"/>乃
34 <lb n="JY006_01p001b_01"/></p><p>元始天王爲神禹治水鎮伏諸瀆而傳厥後一授於天台再授
35 <lb n="JY006_01p001b_02"/>於七曲奉而行之皆證妙果不特爲救護小劫也今夏南水盛
36 <lb n="JY006_01p001b_03"/>漲侍壇超子曾爲河帥故欲其積誠感召以應天和幸覺子等
37 <lb n="JY006_01p001b_04"/>同志贊勳發愿刊<!--Norm 刊-->刻大洞玉經以弭患於未然志良善也夫此
38 <lb n="JY006_01p001b_05"/>經之足以消弭水厄已著於
39 <lb n="JY006_01p001b_06"/></p><p>文帝前敘徵信已久自不必更述故爲闡傳道之初旨更爲述
40 <lb n="JY006_01p001b_07"/>復刊<!--Norm 刊-->之由以付諸子尙望讀是經者窺測<choice><sic>玄</sic><corr>玄
40 </corr></choice>奧其大義不外以
40 <lb n="JY006_01p001b_08"/>人合天其究竟歸於以天從人學道者知人所固有之天復無
41 <lb n="JY006_01p001b_09"/>始以來自然之天天外之天既<!--Norm 既-->得而本心之天全矣
42 <lb n="JY006_01p002a_01"/></p></div><div xmlns:tei="http://www.tei-c.org/ns/1.0"><head>九天開化文昌帝君自敘
44 </head><lb n="JY006_01p002a_02"/><p/><p>余之居事畎澮忽於鋤下得一金像詢之父老曰乃元始天王
45 <lb n="JY006_01p002a_03"/>相昔夏禹治水冶金爲像以鎮方岳余香火奉事未嘗少<choice><sic><g>[↑*鮮]
45 </g></sic><corr>懈</corr></choice><!--Add Table variant-->因
```



# Redundant textual representation

JY006\_01p001a\_01

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
JYE		變	元	贊	運	孚	佑	帝	君	重	刻	大	洞	玉	經	敘									
CK-KZ		變	元	贊	運	孚	佑	帝	君	重	刻	大	洞	玉	經	敘									
YP-Z		變	元	贊	運	孚	佑	帝	君	重	刻	大	洞	玉	經	敘									

JY006\_01p001a\_02

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
JYE		道	可	得	而	言	乎	不	可	得	而	言	也	道	可	得	而	秘	乎	不	可	得	而	秘	也
YP-Z		道	可	得	而	言	乎	不	可	得	而	言	也	道	可	得	而	秘	乎	不	可	得	而	秘	也
CK-KZ		道	可	得	而	言	乎	不	可	得	而	言	也	道	可	得	而	秘	乎	不	可	得	而	秘	也

JY006\_01p001a\_04

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
JYE		也	經	之	旨	有	在	於	言	字	之	內	者	由	淺	以	歷	深	有	在	於	系	微	之	際
CK-KZ		也	經	之	旨	有	在	於	言	字	之	內	者	由	淺	以	歷	深	有	在	於	系	微	之	際
YP-Z		也	經	之	旨	有	在	於	言	字	之	內	者	由	淺	以	歷	深	有	在	於	系	微	之	際



# Reduced representation of regularized and additional text witnesses

JY006\_01p001a\_01

[illegible]

JY006\_01p001a\_02

[illegible]

JY006 01p001a 04

[illegible]



# Interconnected textual layers

