

## 01 Metadata 需求確認表單

<b>主題計畫名稱 (中/英文) :</b> 台灣南島語數位典藏 The Formosan Language Digital Archive	<b>計畫主持人：(姓名、電話、電子郵件)</b> 齊莉莎 26523191 <a href="mailto:hsez@ccvax.sinica.edu.tw">hsez@ccvax.sinica.edu.tw</a>
<b>主題計畫所屬單位：中央研究院語言學研究所籌備處</b>	
<b>參與 Metadata 作業的主題計畫同仁：(姓名、電話、電子郵件)</b> 余清華(27863300*362 / <a href="mailto:harryu@gate.sinica.edu.tw">harryu@gate.sinica.edu.tw</a> ) 工程師 華加婧(27863300*362 / <a href="mailto:taligu@gate.sinica.edu.tw">taligu@gate.sinica.edu.tw</a> ) 助理	
<b>Q1、是否希望將 Metadata 需求納入計畫內？</b> <input checked="" type="checkbox"/> 是                      否 (原因：_____ )	
<b>Q2、是否希望中研院「Metadata 工作組」協助規劃與分析典藏品 Metadata？</b> <input checked="" type="checkbox"/> 是 (請至 Q3)                      否 (原因：_____ )	
<b>Q3、預計時程與期限？</b> 期望在 10 月完成 Metadata 分析，在 12 月有系統雛形可進行著錄測試	
<b>Q4、預期的需求重點與成果？(GIS system? Exhibition Management.....)</b> 1. 南島語 metadata 規格及需求 2. 完成語料的基本資料庫雛形 3. 線上單筆資料更新及檔案匯入功能	
<b>Q5、藏品資料類型與數量(若不同資料類型需不同的 Metadata 著錄格式，請註明，並標明計畫進行的優先順序)</b> 資料類型：魯凱語/雅美語/鄒語/賽夏語/泰雅語/排灣語/布農語/阿美語/卑南語 數量(指方言)：魯凱語: 6 雅美語: 1 鄒語: 3 賽夏語: 2 泰雅語: 2 排灣語: 2 布農語: 5 阿美語: 5 卑南語: 2	
<b>Q6、目前典藏品目錄資訊的狀態？</b> <b>6-1 使用的系統</b> 系統名稱、鍵檔數量、資料庫系統相關文件(規格書, Schema...等) <b>6-2 是否使用任何 metadata 標準？</b> <b>6-3 是否有任何既有之著錄表單？(如：文物基本資料表、登記表...等)</b> <b>6-4 是否有公開網站，可供查詢使用？</b> <b>6-5 是否有其他相關參考資訊(如：其他單位網站、著錄規範...等)</b> 6-1 使用系統：使用 VB 語言撰寫的系統轉換工具，將語料轉換成資料庫檔案 鍵檔數量：魯凱語 萬山方言已收集 764 句 相關文件：RDBMS, XML, HTML 6-2 OLAC 6-3 無 6-4 <a href="http://www.ling.sinica.edu.tw/formosan">http://www.ling.sinica.edu.tw/formosan</a> 6-5 參考國外 OLAC <a href="http://www.language-archives.org">http://www.language-archives.org</a> SIL <a href="http://www.sil.org">http://www.sil.org</a> LACITO <a href="http://lacito.archivage.vjf.cnrs.fr/">http://lacito.archivage.vjf.cnrs.fr/</a>	
<b>Q7、藏品數位化規格需求？(解析度、檔案類型、開放程度...等)</b> 典藏文件多含 IPA 符號，故使用者系統須安裝支援 Unicode 的字型檔： 字型檔：Lucida Sans Unicode, Arial Unicode MS, Code2000 語音檔為 MP3 格式	
<b>其他說明與建議：</b>	

填表者：\_\_\_\_\_

## 02 Metadata 藏品元素需求表單

(包含三部分：一、文本；二、人名資料庫；三、語言資料庫)

### 一、文本

#### 1. 文本層級

元素中文名稱		元素英文名稱		定義	著錄規範	
Element	Subelement	Element	Subelement			
文本號		ID		給定的唯一識別號		
標題	原名	Title	Original title	文本的原語言標題		
	中文		Chinese title	文本的中文標題		
	英文		English title	文本的英文標題		
語言	名稱	Language	Name	文本紀錄的語言方言名稱	下拉式選單。連結語言資料庫「方言 - 名稱 - 拼寫」欄位	
	漢名		Chinese name	語言方言的漢語名稱	下拉式選單	
文類		Genre		文章內容的類別	下拉式選單	
發音人	類別	Informant	Type	發音人對本文本負責的工作	下拉式選單	
	姓名		Name	填寫發音人的漢名	連結人名資料庫	
	語族		Language group	發音人所屬語族	連結人名資料庫	
參與者	類別	Participant	Type	參與語料收集、翻譯、分析、校對人員的類別	下拉式選單	
	姓名		Name	填寫參與者的原名	連結人名資料庫	
日期	類別	Date	Type	有關語料生命週期的類別	下拉式選單	
	起		Year	Begin	有關語料生命週期的起訖時間	年代以西元年著錄
			Month			
			Date			
	迄		Year	End		
			Month			
Date						
經費	類別	Funding	Type	經費來源的類型	下拉式選單	
	名稱		Title	採集文本時，經費來源的名稱		
	主持人		Director	計畫執行的負責者		
	單位		Sponsor	贊助計畫的單位		
	描述		Description	對於經費使用範圍的說明		
獲獎記錄	名稱	Award	Title	本文本曾經得獎的紀錄		
	得獎者		Awardee	得到獎勵的人		

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
	給獎者		Donator	給予獎勵的人或單位	
	描述		Description	對於得獎內容的敘述	
文本內容	原文	Text	Original Text	文本全文	
	中文翻譯	Content	Chinese Translation	文本中文翻譯	
	英文翻譯		English Translation	文本英文翻譯	
	聲音檔路徑		Sound Path	聲音檔的完整路徑	
版權	版權宣告	Access	Availability		系統自動產生
	所有者		Owner		
	出版者		Publisher		系統自動產生
編目記錄	填表日期	Cataloging record	Cataloging date		系統自動產生
	更新日期		Updated date		系統自動產生
	填表者		Cataloger		系統自動產生

## 2. 段落層級

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
段落號		Paragraph ID		段落的識別編號	
文本號		Text ID		所屬文本的編號	
段落		Paragraph		段落內容	
翻譯	中文	Translation	Chinese	段落的中英文翻譯	
	英文		English		
聲音檔路徑		Sound path		聲音檔的完整路徑	

## 3. 句子層級

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
句子號		Sentence ID		句子的識別編號	
文本號		Text ID		所屬文本的編號	
段落號		Paragraph ID		所屬段落的編號	
句子		Sentence		句子拼寫	
翻譯	中文	Translation	Chinese	句子的中英文翻譯	
	英文		English		
聲音檔路徑		Sound path		聲音檔的完整路徑	

## 4. 詞層級

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
詞編號		Word ID		表示詞彙在句子中位置的編號	
句子編號		Sentence ID		表示詞彙所屬的句子編號	
段落號		Paragraph ID		表示詞彙所屬的段落編號	

元素中文名稱		元素英文名稱		定義	著錄規範	
Element	Subelement	Element	Subelement			
文本號		Text ID		表示詞彙所屬的文本編號		
詞	書寫	Word	Orthography	詞的拼寫		
	註解		中文	Gloss	詞彙的中英文語意註解	
				英文		

5. 詞素層級

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
詞素編號		Morpheme ID		表示詞素在詞中位置的編號	
詞編號		Word ID		表示詞彙在句子中位置的編號	
句子編號		Sentence ID		表示所屬詞彙的句子編號	
段落號		Paragraph ID		表示詞彙所屬的段落編號	
文本號		Text ID		表示詞彙所屬的文本編號	
詞素	書寫	Morpheme	Orthography	詞素的拼寫	
	形式		Form	說明詞素形式(bound/free)	下拉式選單
	類型		Category	說明詞素類型(詞根/詞幹)	下拉式選單
	功能標記		Functional tag	語法標記的英文縮寫	
	註解		中文	Gloss	Chinese
英文					

二、人名資料庫

元素中文名稱		元素英文名稱			定義	著錄規範	
Element	Subelement	Element	Subelement				
識別號		ID			唯一識別號	系統自動產生	
類別		Type			此人在此文本負責的工作	關連類別下拉式選單	
姓名	原名	Informant	Name	Original name	填寫原名		
	漢名			Chinese name	填寫漢名		
	代稱			Anonymous	當不願公開姓名時所使用的代稱		
職業	現職	類別	Occupation	Incumbent	Type	說明發音人或參與者的職業類型	下拉式選單
					Name	記錄名稱	
					Affiliation	記錄職業所屬單位	
	經歷			Experience	發音人過去做過的職業		
聯絡方式	電話	Contact		Telephone	說明發音人的電話或參與者的聯絡方式		
	Email			Email			

元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
語族			Language group	發音人所屬的語族	下拉式選單
性別			Sex	發音人的性別	下拉式選單
出生年月日	年	Birth date	Year	發音人的出生西元年	
	月		Month	發音人的出生月份	
	日		Date	發音人的出生日期	
年齡			Age	發音人錄音當時的年齡	
存歿			Live/Dead	說明發音人現在的存歿狀況	下拉式選單
籍貫			Nationality	發音人的籍貫	
母語	類別	First language	Type	說明發音人、其配偶、父母的母語	下拉式選單
	語言		Language	母語的語言別	下拉式選單
語言能力	語言	Language ability	language	發音人說某種語言的能力	下拉式選單
	流利度		Fluency	發音人說某種語言的流利程度	下拉式選單
出生地			Place of birth	發音人的出生地點	
成長地			Place of growth	發音人的成長地點	
居住地	現址	Residence	Address	發音人現居地的地址	不開放
	居住年數		Year	發音人在現居地的居住年數	
教育程度			Education	發音人的教育程度	下拉式選單
填表記錄	填表者	Cataloging record	Cataloging date		系統自動產生
	填表日期		Updated date		系統自動產生
	更新日期		Cataloger		系統自動產生

### 三、語言資料庫

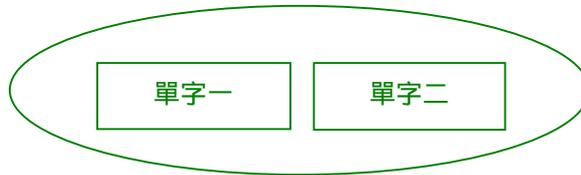
元素中文名稱		元素英文名稱		定義	著錄規範
Element	Subelement	Element	Subelement		
識別號		ID		給予語言的唯一識別號	系統自動產生

語言	類型		Language	Type		記錄語言的族稱或他稱	下拉式選單	
	書寫			Orthography		記錄語言名稱的拼音	下拉式選單；或由著錄者輸入	
漢名			Chinese name			記錄語言的漢語名稱	下拉式選單	
代碼			Code			根據 SIL Code 給予的語言代碼	下拉式選單	
分佈			Distribution			描述語言在台灣的地理分布		
人口			Population			以本語言為母語的人口總數		
現況			Existential condition			語言保存的現況描述		
鄰近語言			Neighboring languages			地理位置上和本語言相鄰的其他語言		
方言	類型		Dialect	Type		指出語言為方言或方言變體	下拉式選單	
	名稱	類型		Language	Type		記錄語言的族稱或他稱	下拉式選單
		拼寫			Orthography	Type		記錄語言名稱的拼寫
	漢名			Chinese name		方言或方言變體的漢名	類別關聯下拉式選單；或由著錄者輸入	
	代碼			Code		記錄方言的代碼	下拉式選單	
	分佈			Distribution		描述語言在台灣的地理分布		
	人口			Population		以本語言為母語的人口總數		
	現況			Existential condition		語言保存的現況描述		
	鄰近語言			Neighboring languages		地理位置上和本語言相鄰的其他語言		

填表者：\_\_\_\_\_

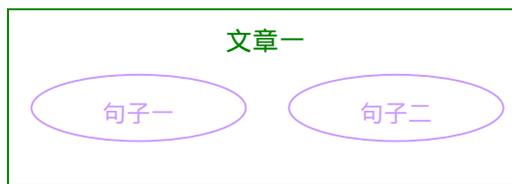
### 03 藏品單元(unit)層級關係圖

#### 群組關係一：句子



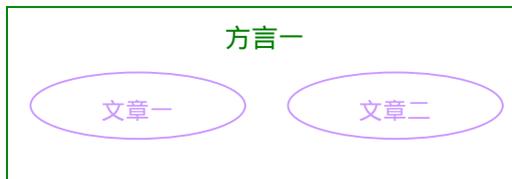
說明：最小之藏品單元(unit)為單字，兩單字或更多會組合為有意義之一組單字組(即為句子)，最小之藏品單元與句子皆各需著錄一筆紀錄。

#### 群組關係二：文章



說明：最小之藏品單元(unit)—單字組成一個句子，若干句子進而組成一篇文章。

#### 群組關係三：方言



說明：一個方言將包含若干文章。

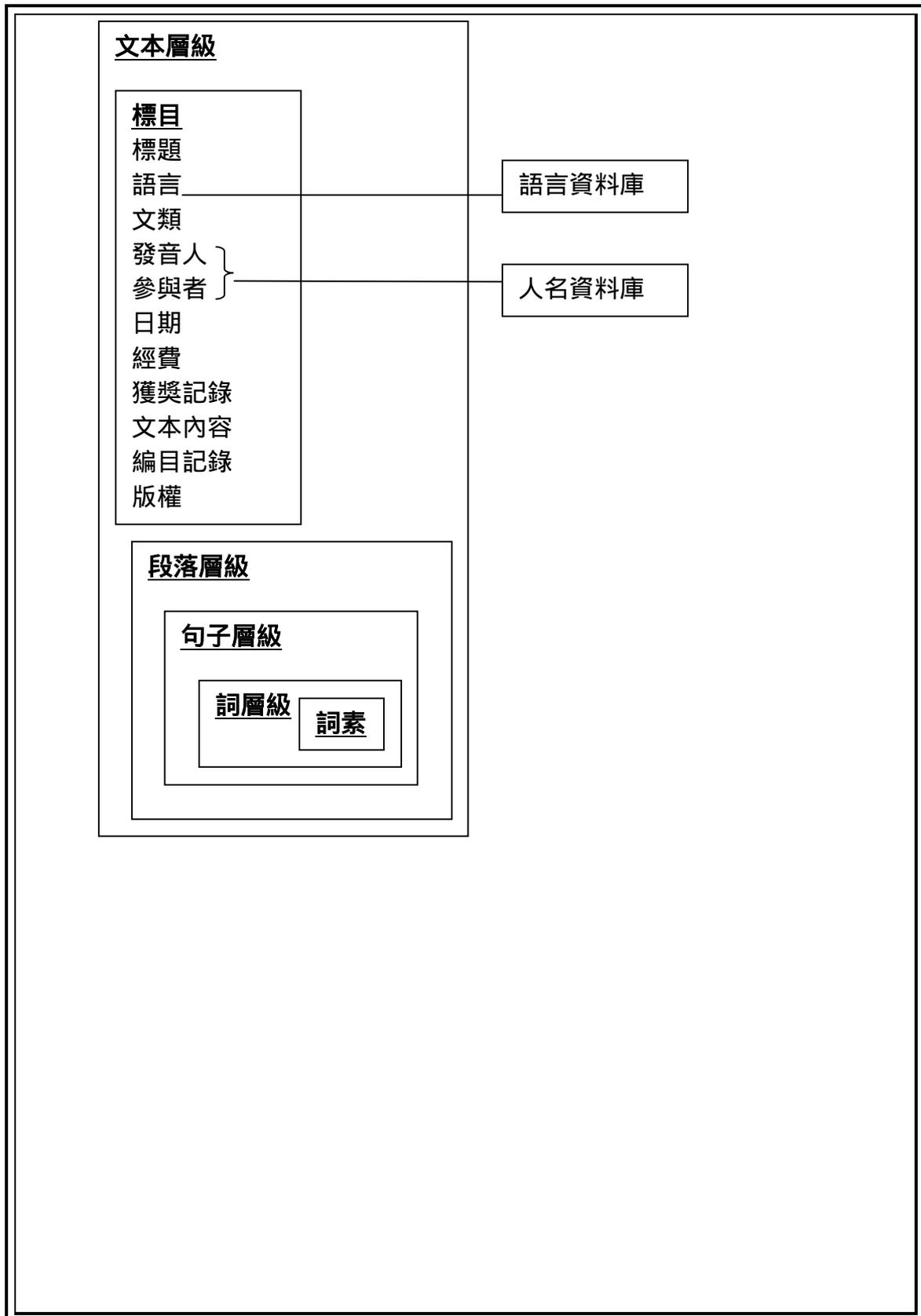
#### 群組關係四：語言



說明：一個語族將包含若干方言。

填表者：\_\_\_\_\_

## 04 元素關係結構圖



填表者：\_\_\_\_\_

## 05 Metadata 元素代碼表單

### 一、文本

#### 1. 文本層級

元素中文名稱		代碼	
Element	Subelement		
語言	名稱	Atayal	Squliq   C'uli'
		Saisiyat	Taai   Tungho
		Pazeh	
		Bunun	Takituduh   Takibakha   Takbanuaz   Takivatan   Isbukun
		Yami	
		Tsou	Duhtu   Tfuya   Tapangu
		Kanakanavu	
		Saaroa	
		Rukai	Mantauran   Maga   Tona   Budai   Tanan   Labuan
		Paiwan	t-dialect   tj-dialect
		Puyuma	Nanwang   Kapitul
		Amis	Sakizaya   Northern   Tavalong-Vata'an   Central   Southern
		Seediq	Paran   Toda   Taroko
		Kavalan	
	漢名	泰雅	賽考利克   澤敖利
		賽夏	大隘   東河
		巴澤	
		布農	卓社群   卡社群   巒社群   丹社群   郡社群
		雅美	
		鄒	久美   特富野   達邦
		卡那卡那富	
		沙阿魯阿	
		魯凱	萬山   茂林   多納   霧台   大南   大武
		排灣	西北部   東南部
		卑南	南王   知本
		阿美	沙奇萊亞   北部(南勢群)   馬太安   中部(海岸群、秀姑巒群)   南部(卑南群、恆春群)
		賽德克	霧社   春陽   太魯閣
噶瑪蘭			
文類		傳說 口述 對談 歌謠	
發音人	類型	口述者   口譯者   修改者	
參與者	類別	紀錄者 翻譯者 分析者   校對者	
日期	類別	語料收集 校對編輯 分析 標註翻譯	

元素中文名稱		代碼
Element	Subelement	
經費	類別	計畫   贊助

5. 詞素層級

元素中文名稱		代碼
Element	Subelement	
詞素	形式	附加詞素   自由詞素
	類型	詞根   詞幹

二、人名資料庫

元素中文名稱			代碼
Element	Subelement		
類型	發音人		口述者   口譯者   修改者
	參與者		紀錄者 翻譯者 分析者   校對者
職業	現職	類別	語言學界   非語言學界
語族	魯凱族 雅美族 鄒族 賽夏族 泰雅族 排灣族 布農族 阿美族 卑南族		
性別	男 女		
存歿	存   歿		
母語	類別	本人 父親 母親 配偶	
	語言	魯凱語 雅美語 鄒語 賽夏語 泰雅語 排灣語 布農語 阿美語 卑南語 國語 閩南語 客家語 日語 其他	
語言能力	名稱	母語 國語 閩南語 客家語 日語 其他	
	流利度	很流利 可與人溝通 勉強可溝通 聽得懂 完全不懂	
教育程度	自修 幼稚園 小學肆 小學畢 國中肆 國中畢 高中(職)肆 高中(職)畢 大學以上肆 大學以上畢		

## 三、語言資料庫

元素中文名稱		代碼		
Element	Subelement			
語言	類型	族稱   他稱		
	書寫	Atayal   Saisiyat   Pazeh   Bunun   Yami   Tsou   Kanakanavu   Saaroa   Rukai   Paiwan   Puyuma   Amis   Seediq   Kavalan		
漢名		泰雅   賽夏   巴澤   布農   雅美   鄒   卡那卡那富   沙阿魯阿   魯凱   排灣   卑南   阿美   賽德克   噶瑪蘭		
代碼		TAY   SAI   PZH   BNN   YMI   TSY   QNB   SXR   DRU   PWN   PYU   ALV   SDQ   CKV		
方言	類型	方言   方言變體		
	名稱	類型	族稱   他稱	
		書寫	Atayal	Squliq   C'uli'
		Saisiyat	Taai   Tungho	
		Pazeh		
		Bunun	Takituduh   Takibakha   Takbanuaz   Takivatan   Isbukun	
		Yami		
		Tsou	Duhtu   Tfuya   Tapangu	
		Kanakanavu		
		Saaroa		
		Rukai	Mantauran   Maga   Tona   Budai   Tanan   Labuan	
		Paiwan	t-dialect   tj-dialect	
		Puyuma	Nanwang   Kapitul	
		Amis	Sakizaya   Northern   Tavalong-Vata'an   Central   Southern	
		Seediq	Paran   Toda   Taroko	
	Kavalan			
漢名	泰雅	賽考利克   澤敖利		
	賽夏	大隘   東河		
	巴澤			
	布農	卓社群   卡社群   巒社群   丹社群   郡社群		
	雅美			

元素中文名稱		代碼	
Element	Subelement		
		鄒	久美   特富野   達邦
		卡那卡那富	
		沙阿魯阿	
		魯凱	萬山   茂林   多納   霧台   大南   大武
		排灣	西北部   東南部
		卑南	南王   知本
		阿美	沙奇萊亞   北部(南勢群)   馬太安   中部(海岸群、秀姑巒群)   南部(卑南群、恆春群)
		賽德克	霧社   春陽   太魯閣
		噶瑪蘭	
	代碼	TAY	Sq   Cu
		SAI	Ta   Tu
		PZH	
		BNN	Td   Th   Tb   Tn   Is
		YMI	
		TSY	Du   Tf   Ta
		QNB	
		SXR	
		DRU	Mn   Mg   To   Bu   Ta   La
		PWN	Td   Tj
		PYU	Na   Ka
		ALV	Sa   No   Ta   Ce   So
		SDQ	Pa   To   Ta
		CKV	

填表者：\_\_\_\_\_

## 06 Metadata 著錄範列表單

### 一、文本

#### 1. 文本層級

元素中文名稱		範例	
Element	Subelement		
文本號		1	
標題	原名		
	中文	我們的話	
	英文	Our language	
語言	名稱	Rukai-Mantauran	
	漢名	魯凱 - 萬山	
文類		傳說	
發音人	類型	口述者	
	姓名	呂玉枝	
	語族	魯凱族	
參與者	類別	紀錄者	
	姓名	Elizabeth Zeitoun	
參與者	類別	記錄者	
	姓名	林惠娟	
日期	類別	語料收集	
	起	年	1992
		月	
		日	
	迄	年	1992
		月	
日			
日期	類別	語料收集	
	起	年	1997
		月	
		日	
	迄	年	1999
		月	
日			
日期	類別	校對編輯	
	起	年	1999
		月	
		日	
	迄	年	2001
		月	
日			
日期	類別	標註翻譯	
	起	年	1992
		月	
日			

元素中文名稱		範例
Element	Subelement	
	迄	2001
	年	
	月	
	日	
經費	類別	
	名稱	
	主持人	
	單位	
	描述	
得獎記錄	名稱	
	得獎者	
	給獎者	
	描述	
文本內容	原文	<p>onaʔi ʔaamaʔalaə-nai ta-piʔa-aə-na-ʔa po-a acə ʔoponoho m-ia. ʔonaʔi a-kaavaʔi-ŋa-ʔa ʔaomo mani a a ʔicaʔocaʔo ʔina vaha-nai ʔoponoho la tali-lao-ləoʔo. mani poa a-pa-rama-ramao kapa-ka-ʔao-ʔao ta-ʔini ao ʔi tali-ʔəkai ʔina vaha-nai. mani a -inamə patoʔo maʔ-itololo-nomi m-ia. mani lihoʔo-nai ʔina ʔoponoho ka-iʔa-ta aanai m-ia-nai. ʔonaʔi a-ki-kaavaʔi-na-ʔa ʔaomo o-lihoʔo-ka-nai maʔ-itololo-mita ʔəkai m-ia. kasa-ta ʔoponoho vaha-ta om-ia-nai. ʔinaʔi vaha-nai ʔi ʔoponoho toramoro ka ma-kociŋai. cəla ʔinaʔi la-a-lakə-nai ta-ovalisi-ŋa o-lihoʔo-ka-i ka pa-loloʔo ʔinaʔi ta-ia-ə-nai ʔao-vao-vaha ta-ka-a-roʔaŋə. cəla ʔi va ava ə lo m-a a-nai siri-ri ka mati-lihoʔo-ka-i m-iiʔa vaha-nai. olo ʔao-vao-vaha-nai ka-ʔoponoh-aə ocao ʔinaʔi vaha-nai ta-lihoʔo to-siʔi olo vaha-nai ʔako-ʔoponoh-aə-nai om-iʔa-ka-i ka ma-rama-ramao. toʔa-toloraə-li vaha-nai ʔoponoho toramoro ma-kociŋai. toʔa-toloraə-li ʔinaʔi vaha-nai ʔoponoho alaka-i ʔinaʔi ta-ovalisi-ŋa a-o olai ka o-lihoʔo-ka-i ka ʔoponoho vaha. a-vasaə-nai vaha-nai amo-ka-olipi ʔinaʔi vaha-nai ʔoponoho. mani ʔatoloro-la-ilinə piʔa ʔinaʔi la-lini ka-amai pa-solatə la-lini vaaʔi patoʔo-toʔo ʔinaʔi a-o olai la-ni lihoʔo la-ni ki-ka-o o vaha-nai ʔoponoho.</p>

元素中文名稱		範例
Element	Subelement	
	中文翻譯	我們的祖先萬山（自己是）萬山人。日本人來了以後就（開始）學我們萬山和屏東縣的話。他們比較這兩種語言後，（就發現）我們的話有許多（和）魯凱族是一樣的。然後他們告訴我們說：「你們是同宗（的）。」我們萬山人（才）知道「原來咱們（屬於魯凱族）。」日本人還沒來之前我們不知道我們和魯凱族同宗。我們只（認為）咱們的語言（就）是萬山話。我們萬山話很難（學）。看看現在年輕一輩的小孩：他們（已經）不會隨著我們老人家說話的方法。（再）看看那些我們從外村娶來的人：他們（也不）完全懂我們的話。（而且）我們這些當地的萬山（人和那些）會一點（萬山話）的人，所說的萬山話（也）不完全一樣。原因（就在於）我們萬山話（真的）很難。我（參與編輯這套書的）原因，（是）因為現在的小孩（幾乎都）不會講萬山話。很可惜，我們萬山話即將消失。所以我（參與）這件事，希望他們（兩位作者）可以學會（萬山話），（再）教會（我們的）小孩，如果小孩（學）會（萬山話），我們的語言就不會消失。
	英文翻譯	Our ancestors used to call (themselves) Mantauran. When the Japanese came, they learnt our language as well as (the languages spoken by the people living in) Pingtung county. They compared (our) languages and (discovered that) there were many resemblances between Rukai and our (own) language. Then they told us: "You share the same ancestry." (That's how) we, Mantauran, learnt that actually we were Rukai. Before the Japanese came, we did not know we were related to the Rukai. We just knew that our language was Mantauran. Our language is very difficult (to learn). Look at the young generation of children: (they) do not (even) know (how to) repeat (what) the old people say. Look at those outsiders: (even) if they inter-marry with the Mantauran, they do not understand our language very well. (If you listen to) us authentic Mantauran or (to those who) only know a little the language, (you'll find that our speech) varies (from one person to another). That's because our language is very difficult. The reason for my (doing) this (is to preserve) our language, because the young generation does not know how to speak Mantauran. It is (such) a pity that our language might become extinct. I am therefore (collaborating on) this (book) so that they [the authors] might teach themselves (our language) to teach (it in turn to our) children; (if) the children learn (our language), it will not disappear.
	聲音檔路徑	<a href="http://www.ling.sinica.edu.tw/formosan/sound/mantauran/text/01.mp3">http://www.ling.sinica.edu.tw/formosan/sound/mantauran/text/01.mp3</a>

元素中文名稱		範例
Element	Subelement	
版權	版權宣告	中研院擁有「台灣南島語數位典藏」的相關智慧財產權，包括介面程式、語料組成方式、後設資料欄位、詞類標記、詞綴分析及詞彙庫等。使用者僅能使用網頁介面查詢所需資料以及利用查詢結果進行研究，但是不得隨意擷取、修改、出版查詢結果。語料的版權仍歸原著作人或提供單位，他人不得轉載、抄襲或侵犯任何語料的智慧財產權。
	所有者	中央研究院語言學研究所籌備處
	出版者	中央研究院語言學研究所籌備處
編目記錄	填表日期	2002/8/23
	更新日期	2002/11/12
	填表者	翁翠霞

## 2. 段落層級

元素中文名稱		實例
Element	Subelement	
段落號		001
文本號		1
段落		onaʔi ʔaamaðalaə-nai ta-piʔa-aə-na-ða po-a acə ʔoponoho m-ia. ðonaʔi a-kaavaʔi-ŋa-ða ʔaomo mani a a ʔicaʔocaʔo ʔina vaha-nai ʔoponoho la tali- ao- aoðo. mani poa a-pa-rama-ramao kapa-ka-ðao-ðao ta-ʔini ao ʔi tali-dəkai ʔina vaha-nai. mani a -inamə patoʔo maʔ-itoloto o-nomi m-ia. mani  ihoʔo-nai ʔina ʔoponoho ka-iʔa-ta aanai m-ia-nai. ðonaʔi a-ki-kaavaʔi-na-ða ʔaomo o- ihoʔo-ka-nai maʔ-itoloto o-mita dəkai m-ia. kasa-ta ʔoponoho vaha-ta om-ia-nai.
翻譯	中文	我們的祖先萬山（自己是）萬山人。日本人來了以後就（開始）學我們萬山和屏東縣的話。他們比較這兩種語言後，（就發現）我們的話有許多（和）魯凱族是一樣的。然後他們告訴我們說：「你們是同宗（的）。」我們萬山人（才）知道「原來咱們（屬於魯凱族）。」日本人還沒來之前我們不知道我們和魯凱族同宗。我們只（認為）咱們的語言（就）是萬山話。
	英文	
聲音檔路徑		<a href="http://www.ling.sinica.edu.tw/formosan/sound/mantauran/para/001.mp3">http://www.ling.sinica.edu.tw/formosan/sound/mantauran/para/001.mp3</a>

## 3. 句子層級

元素中文名稱	實例
--------	----

Element	Subelement	
句子號		a
段落號		001
文本號		1
句子		onaʔi ʔaamaðalaə-nai ta-piʔa-aə-na-ða po-aɭacə ʔoponoho m-ia.
翻譯	中文	我們的祖先萬山(自己是)萬山人
	英文	Our ancestors used to call (themselves) Mantauran.
聲音檔路徑		http://www.ling.sinica.edu.tw/formosan/sound/mantauran/001a.mp3

4. 詞層級

元素中文名稱		著錄規範	
Element	Subelement		
詞號		2	
句子號		a	
段落號		001	
文本號		1	
詞	拼寫	ʔaamaðalaə-nai	
	註解	中文	我們祖先
		英文	our ancestor

5. 詞素層級 (1)

元素中文名稱		著錄規範	
Element	Subelement		
詞素編號		1	
詞編號		2	
句子編號		a	
段落號		001	
文本號		1	
詞素	書寫	ʔaamaðalaə	
	形式	自由詞素	
	類型	詞幹	
	功能標記		
	註解	中文	祖先
英文		ancestor	

5. 詞素層級 (2)

元素中文名稱		著錄規範	
Element	Subelement		
詞素編號		2	
詞編號		2	
句子編號		a	
段落號		001	

元素中文名稱		著錄規範
Element	Subelement	
文本號		1
詞素	書寫	nai
	形式	附加詞素
	類型	詞幹
	功能標記	
	註解	中文 英文

## 二、人名資料庫

元素中文名稱		範例
Element	Subelement	
識別號		1
類別		發音人 口述者
姓名	原名	
	漢名	呂玉枝
	代稱	
職業	現職	類別
		名稱
		所屬機構
	經歷	
聯絡方式	電話	
	Email	
語族		魯凱族
性別		女
出生年月日	年	1924
	月	7
	日	24
年齡		
存歿		歿
籍貫		台灣省
母語	類別	本人
	語言	魯凱語
	類別	父親
	語言	魯凱語
	類別	母親
	語言	魯凱語
語言能力	語言	母語
	流利度	很流利
出生地		高雄縣舊萬山
成長地		高雄縣舊萬山
居住地	現址	高雄新萬山

元素中文名稱		範例
Element	Subelement	
	居住年數	41
教育程度		小學畢
填表記錄	填表者	翁翠霞
	填表日期	2002/8/23
	更新日期	2002/11/12

### 三、語言資料庫

元素中文名稱		範例	
Element	Subelement		
識別號		1	
語言	類型	族稱	
	書寫	Rukai	
漢名		魯凱語	
代碼		DRU	
分佈			
人口			
現況			
鄰近語言			
方言	類型	方言	
	名稱	類型	族稱
		拼寫	'Oponoho
	名稱	類型	他稱
		拼寫	Mantauran
	漢名		萬山
	代碼		Mn
	分佈		
	人口		
	現況		
	鄰近語言		

填表者：\_\_\_\_\_

## 07 Metadata 系統屬性功能需求表單

## 一、文本

## 1. 文本層級

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
文本號		1	2	Y					
標題	原名	2	255	Y					
	中文	3	100	Y					
	英文	4	100	Y					
語言	名稱	5	12	Y		代碼表			
	漢名	6	12	Y		代碼表			
文類		7	6	Y		代碼表			
發音人	類型	8	10		Y	代碼表			
	姓名	9	20		Y	人名資料庫			
	語族	10	12		Y	人名資料庫			
參與者	類別	11	10		Y	代碼表			
	姓名	12	20		Y	人名資料庫			
日期	類別	13	20		Y	代碼表			
	起	年	14	4		Y			
		月	15	2		Y			
		日	16	2		Y			
	迄	年	17	4		Y			
		月	18	2		Y			
日		19	2		Y				
經費	類別	20	20		Y	代碼表			
	名稱	21	20		Y				
	主持人	22	20		Y				
	單位	23	20		Y				
	描述	24	50		Y				
獲獎記錄	名稱	25	50		Y				
	得獎者	26	20		Y				
	給獎者	27	20		Y				
	描述	28	50		Y				
文本內容	原文	29	64k	Y					
	中文翻譯	30	64k	Y					
	英文翻譯	31	64k	Y					
	聲音檔路徑	32	50						

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
版權	版權宣告	33	100	Y					中研院擁有「台灣南島語數位典藏」的相關智慧財產權，包括介面程式、語料組成方式、後設資料欄位、詞類標記、詞綴分析及詞彙庫等。使用者僅能使用網頁介面查詢所需資料以及利用查詢結果進行研究，但是不得隨意擷取、修改、出版查詢結果。語料的版權仍歸原著作人或提供單位，他人不得轉載、抄襲或侵犯任何語料的智慧財產權。
	所有者	34	20		Y		中央研究院語言所籌備處		
	出版者	35	20		Y		中央研究院語言所籌備處		
編目記錄	填表日期	36	10	Y		系統自動產生			
	更新日期	37	10	Y		系統自動產生			
	填表者	38	20	Y		系統自動產生			

## 2. 段落層級

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
段落號		1	3	Y					
文本號		2	2	Y					
段落		3	32k	Y					
翻譯	中文	4	32k	Y					

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
	英文	5	32k	Y					
聲音檔路 徑		6	50						

## 3. 句子層級

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
句子號		1	1	Y					
文本號		2	2	Y					
段落號		3	3	Y					
句子		4	255	Y					
翻譯	中文	5	255	Y					
	英文	6	255	Y					
聲音檔路 徑		7	50						

## 4. 詞層級

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
詞編號		1	1	Y					
句子編號		2	1	Y					
段落號		3	3	Y					
文本號		4	2	Y					
詞	書寫	5	100	Y					
	註解	中文	6	100	Y				
		英文	7	100	Y				

## 5. 詞素層級

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
詞素編號		1	1	Y					
詞編號		2	1	Y					
句子編號		3	1	Y					
段落號		4	3	Y					
文本號		5	2	Y					
詞素	書寫	6	25	Y					
	形式	7	5	Y					
	類型	8	4	Y					
	功能標記	9	20	Y					
	註解	中文	10	50	Y				
		英文	11	50	Y				

## 二、人名資料庫

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
識別號		1	系統定義			系統自動 產生			
類別		2	10	Y		代碼表			
姓名	原名	3	20	Y					
	漢名	4	20	Y					
	代稱	5	20						
職業	現職	類別	6	20		代碼表			
		名稱	7	20					
		機構	8	50					
	經歷	9	50		Y				
聯絡方式	電話	10	20		Y				
	Email	11	20		Y				
語族		12	12	Y		代碼表			
性別		13	2	Y		代碼表			
出生年月 日	年	14	4						
	月	15	2						
	日	16	2						
年齡		17	3						
存歿		18	2			代碼表			
籍貫		19	10						
母語	類別	20	12	Y	Y	代碼表	本人		
	語言	21	12	Y	Y	代碼表			
語言能力	語言	22	10		Y	代碼表	母語		
	流利度	23	10		Y	代碼表			
出生地		24	20						
成長地		25	20						
居住地	現址	26	20						
	居住年數	27	3						
教育程度		28	10			代碼表			
填表記錄	填表者	29	20			系統自動 產生			
	填表日期	30	10			系統自動 產生			
	更新日期	31	10			系統自動 產生			

## 三、語言資料庫

元素中文名稱		欄位 順序	欄位最大 字元數	必填 欄位	多值 欄位	功能連結	預設值	固定值	唯一值欄位
Element	Subelement								
識別號		1	系統定義			系統自動 產生			
語言	類型	2	10	Y	Y	代碼表			
	書寫	3	20	Y	Y	代碼表			
漢名		4	12	Y		代碼表			
代碼		5	3	Y		代碼表			
分佈		6	20		Y				
人口		7	5						
現況		8	20						
鄰近語言		9	20		Y				
方言	類型	10	20	Y	Y	代碼表			
	名稱	類型	11	20	Y	Y	代碼表		
		拼寫	12	20	Y	Y	代碼表		
	漢名	13	12	Y	Y	代碼表			
	代碼	14	2	Y	Y	代碼表			
	分佈	15	20		Y				
	人口	16	5		Y				
	現況	17	20		Y				
	鄰近語言	18	20		Y				

填表者：\_\_\_\_\_

## 08 Metadata 系統查詢功能需求表單

一、文本

## 1. 文本層級

元素中文名稱		Element	Subelement	簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目	
Element	Subelement							
文本號								
標題	原名				Y	Y	Y	
	中文	Y		Y	Y	Y	Y	
	英文			Y	Y	Y	Y	
語言	名稱	Y		Y	Y	Y	Y	
	漢名	Y		Y	Y	Y	Y	
文類			Y				Y	
發音人	類別						Y	
	姓名	Y		Y			Y	
	語族			Y			Y	
參與者	類別						Y	
	姓名	Y		Y			Y	
日期	類別						Y	
	起	年			Y		Y	
		月						Y
		日						Y
	迄	年			Y			Y
		月						Y
日							Y	
經費	類別						Y	
	名稱						Y	
	主持人						Y	
	單位						Y	
	描述						Y	
獲獎記錄	名稱						Y	
	得獎者			Y			Y	
	給獎者						Y	
	描述						Y	
文本內容	原文				Y	Y	Y	
	中文翻譯	Y		Y			Y	
	英文翻譯			Y			Y	
	聲音檔路徑							
版權	版權宣告					Y	Y	
	所有者					Y	Y	
	出版者					Y	Y	
編目記錄	填表日期							
	更新日期							

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
	填表者				

## 2. 段落層級

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
段落號					
文本號					
段落			Y	Y	Y
翻譯	中文	Y	Y		Y
	英文		Y		Y
聲音檔路徑					

## 3. 句子層級

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
句子號					
文本號					
段落號					
句子			Y	Y	Y
翻譯	中文	Y	Y		Y
	英文		Y		Y
聲音檔路徑					

## 4. 詞層級

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
詞編號					
句子編號					
段落號					
文本號					
詞	書寫	Y	Y	Y	Y
	註解	中文	Y		Y
		英文		Y	

## 5. 詞素層級

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
詞素編號					
詞編號					
句子編號					
段落號					
文本號					
詞素	書寫	Y	Y	Y	Y
	形式	Y	Y		Y

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目	
Element	Subelement					
	類型	Y	Y		Y	
	功能標記		Y		Y	
	註解	中文	Y	Y	Y	Y
		英文		Y		Y

## 二、人名資料庫

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
識別號					
類別			Y	Y	Y
姓名	原名		Y	Y	Y
	漢名	Y	Y	Y	Y
	代稱				
職業	現職	類別		Y	Y
		名稱		Y	Y
		所屬機構	Y	Y	Y
	經歷				Y
聯絡方式	電話				
	Email				
語族		Y	Y	Y	Y
性別				Y	Y
出生年月日	年			Y	Y
	月				
	日				
年齡			Y	Y	Y
存歿				Y	Y
籍貫					
母語	類別		Y		Y
	語言	Y	Y		Y
語言能力	語言				Y
	流利度				Y
出生地			Y		Y
成長地					Y
居住地	現址		Y		
	居住年數				
教育程度				Y	Y
填表記錄	填表者				
	填表日期				
	更新日期				

## 三、語言資料庫

元素中文名稱		簡要查詢	進階查詢	簡要顯示款目	詳細顯示款目
Element	Subelement				
識別號					
語言	類型		Y	Y	Y
	書寫	Y	Y	Y	Y
漢名		Y	Y	Y	Y
代碼				Y	Y
分佈				Y	Y
人口				Y	Y
現況				Y	Y
鄰近語言					Y
方言	類型		Y	Y	Y
	名稱	類型		Y	Y
		拼寫	Y	Y	Y
	漢名	Y	Y	Y	Y
	代碼				Y
	分佈			Y	Y
	人口				Y
	現況				Y
鄰近語言				Y	

填表者：\_\_\_\_\_

## 09 Metadata 系統紀錄建檔流程說明

著錄人員要進行紀錄建檔時，進入著錄介面之前需經過身份的確證，著錄人員必須輸入自己的姓名與密碼，方可登入 Metadata 系統進行建檔，系統需記錄建檔人姓名與建檔時間。同樣的，當著錄人員需修改紀錄內容時，也需經過身份的確證，藉由確證的動作可知該筆記錄是經由誰人所修改，系統亦需記錄修改人姓名與修改時間。系統不需記錄每次的修改資訊，只需記錄最近一次的修改資訊即可。另外，著錄的過程應包含檔案的匯入功能。

## 10 Metadata 系統主鍵元素架構格式

系統可將「句子」作為主鍵的基礎，以其識別碼作為系統主鍵，其架構格式，如下所示：  
三個英文字母(族語別) + 兩個英文字母(方言別) + 兩個阿拉伯數字(文章代碼) + 三個阿拉伯數字(段落代碼) + 一個英文字母(段落中的句子代碼)

如：DRUMN01001A

說明：魯凱語(DRU)的萬山方言(MN)，第 01 個文本的第 001 段，其中第 A 句 (A 表 1，B 表 2，其餘類推)。

填表者：\_\_\_\_\_