



Ontology Design for Supporting Digital Humanities on Historical Personages: The Taiwan Biographical Database (TBDB) Case

Hao-Ren Ke*, Shun-Hong Sie, Su-Bing Chang

National Taiwan Normal University

Email: clavenke@ntnu.edu.tw

Outline

- ▶ Introduction
- ▶ Ontology development
- ▶ System Demonstration
- ▶ Conclusions

INTRODUCTION

Overview

- ▶ Digital humanities: the use of digital materials and digital tools to transform scholarship in the humanities and social sciences
- ▶ Personage is important in history study
 - ✿ Individual's family background, academic achievement, professions, social status, kinship and social relationships, marriage networks and political networks
- ▶ China Biographical Database (CBDB)
 - ✿ A cooperative project among Harvard University, Academia Sinica and Peking University
 - ✿ Collect about 491,000 Chinese figures (2021/05) in 7-19 centuries

Taiwan Biographical Database (TBDB)

- ▶ A DH-oriented system that assists historians in discovering information on Taiwan historical personages and exploring research topics
- ▶ Collects biographies and offers historians software services including relational database search, fulltext search, text mining, social network analysis (SNA), information visualization, and geographic information system (GIS) functions
 - ✿ 1911 personages from 16th century, especially 19th-20th centuries
- ▶ <http://tbdb.ntnu.edu.tw>

ONTOLOGY DEVELOPMENT



What is an Ontology?

- ▶ An ontology models the vocabulary and meaning of domains of interests in a computer-usable form
→ **make knowledge reusable**
- ▶ An ontology defines the common words and concepts (the meaning) used to describe and represent an area of knowledge
 - ✿ Classes (general things) in the domain of interest
 - ✿ The property (and property values) of those things
 - ✿ Instances (particular things)
 - ✿ The relationships among those things

Ontology Development 101

- ▶ Determine the domain and scope of the ontology
- ▶ Consider reusing existing ontologies
- ▶ Enumerate important terms in the ontology
- ▶ Define the classes and the class hierarchy
- ▶ Define the properties of classes—slots
- ▶ Define the facets of the slots
- ▶ Create instances

Person Properties

CIDOC CRM Entity	Property Name	Data Type	Description
	additionalName	Text	別名、字、號
E74Group	affiliation	Organization	參與社團、職業、組織
E74Group	alumniOf	Organization	畢業學校/校友
E67Birth	birthDate	Text	誕生日
E53Place	birthPlace	Place	出生地
E21Person	children	Person	子女
E69Death	deathDate	Text	死亡日
E53Place	deathPlace	Place	死亡地點
P131isIdentified by	familyName	Text	姓氏
E21Person	follows	Person	追隨者/徒弟
	gender	Text	性別
P131isIdentified by	givenName	Text	名
	hasOccupation	Text	曾任職業、職稱
	homeLocation	Place	居住地
	honorificPrefix	Text	榮譽事蹟或頭銜
	jobTitle	Text	職位名稱
E21Person	knows	Person	相關人物,和其有關連人物
E74Group	memberOf	Organization	會員/組織成員
	nationality	Text	國別
E21Person	parent	Person	父母
	performerIn	Event	參與事件
	relatedTo	Event	其它相關事件

Organization Properties

CIDOC CRM Entity	Property Name	Data Type	Description
E53PPlace	address	Place	地址/地點/所在地點
E21Person	alumni	Person	成員
E53Place	areaServed	Place	主要活動地點
	award	Text	曾獲獎項
	brand	Organization	持有品牌/產品
	department	Organization	部門
E52Time-Span	dissolutionDate	Text	解散日期
E21Preson	employee	Person	員工
E5Event	event	Event	重大事件
E21Person	founder	Person	贊助商/創辦人
E52Time-Span	foundingDate	Text	創辦日期
E53Place	foundingLocation	Place	創辦地點
	legalName	Text	正式註冊名稱
E53Place	location	Place	地點
E74Group	relMember	Organization	相關組織
E74Group	memberOf	Organization	所屬組織
E74Group	parentOrganization	Organization	上層單位/組織
E74Group	subOrganization	Organization	子組織
E74Group	changeOrg	Organization	改名/重組
E74Group	disjoinOrg	Organization	分割組織來源
E74Group	mergeOrg	Organization	合併組織來源

Relation Properties

CIDOC CRM Entity	Property Name	Data Type	Description
	title	Text	關係名稱
E55TYPE	type	Kinship/Social Relationship	關係類型
E21Person	actor	Person	主要參與者
E21Person	attendee	Person	參與者
E11 Modification	startTime	Text	起始時間
	order	Text	次序
E21Person	upperKin	Text	上層關係

SYSTEM DEMONSTRATION

Home Page and Search Results

The screenshot shows the homepage of the Taiwan Biographical Database (TBDB). At the top left is the logo "TBDB". On the right are buttons for "登入/註冊" (Login/Register) and a menu icon. The central graphic features vertical Chinese text: "臺灣歷史" (Taiwan History) and "人物傳記資料庫" (Biographical Database). Below the text is a silhouette of the island of Taiwan. In the background, there's a faint image of the Taipei 101 building and the Taiwan Strait Bridge.

傳主姓名 出生地 內文關鍵字

Name **Birth Place** **Full text** **搜尋**
[完整查詢功能](#)

Home Page and Search Results

Facet Analysis	總筆數：11筆						查看 Full text
	Name	Birth/Active Place	Type	姓名	出生地	事蹟發生地	
全部列出				柯謀	伸港	臺中	經濟人物
出生地				柯批	伸港	伸港	經濟人物
伸港 (5)				柯福	北斗	北斗	經濟人物
北斗 (2)				柯扱	伸港	伸港	政治人物
線西 (1)				柯子彰	臺北市		教育學術
臺北市 (1)				柯秋潔	士林		教育學術
士林 (1)				柯慶瑞	北斗	北斗	經濟人物
大甲 (1)				柯煥章	伸港	和美	文化人物
事蹟發生地				柯子元	伸港	伸港	社會人物
伸港 (3)				柯永昌	線西	線西	社會人物
北斗 (2)				柯莊庭	大甲		婦女類
和美 (1)							
線西 (1)							
臺中 (1)							
傳文出處							
新修彰化縣志 (8)							
續修臺北市志 (2)							
續修臺中縣志 (1)							

Home Page and Search Results

傳主／

柯謀

出生地/事蹟發生地／

伸港/臺中

生卒年／

1893-1964

撰稿者／

顧雅文

柯謀（臺灣官紳年鑑）線西新港人（今伸港鄉大同村）。父錦雞，母阮燦，為長男。自幼修習漢學，造詣深厚，且立志於實業界發展。[\[註1\]](#)

明治42年（1909），16歲進入臺中林金水支店工作，努力學習商貿業務，逐漸熟悉經營的秘訣。大正8年（1919）自立門戶，於臺中市榮町開設柯廣川商店，主要經營和洋雜貨販賣，最初資本僅有百圓，但至1930年代已成知名百貨店，每年有十數萬的營業額，經常與各銀行及公司行號交易，信用良好。後為拓展販賣通路，於臺中老松町設立分店，被評為手腕敏潤、經驗豐富。[\[註2\]](#)

與臺中市內居住的許、柯、蔣、巫四大姓家族有密切往來，自大正14年（1925）同許克綏（另有傳）、許炳雲、蔣良、柯金水等人發起四姓懇親會，廣邀四姓宗親80餘人團聚聯歡，並計畫收受會費，每年懇親甚或樹立團體，以彰四姓協力一致之心，並表宗親之誼。[\[註3\]](#)

此外，擔任過臺中商工協會評議員、[\[註4\]](#)庶民組合評定委員等名譽職。對於地方公共事業也竭盡心力，經常捐助大筆金額，如於昭和4年（1929）與地方巡查以及曾灶、黃合親（另有傳）等人，共同發起在線西庄新港造橋，並與地方有志之士96人捐資5千圓，是為善舉。[\[註5\]](#)

二弟柯尊、四弟克靖（克清），於臺中市經營雜貨商；六弟安然。

妻黃撻養，育有女呈淑、欽、雀、秋梨、素晴，養子水枝。妾何雲月，生女阿昭，男煥榮、建助。

[顧雅文]

**Biography
Full Text**

Automatic Named Entity Recognition (NER)

創立「**菱香**吟社」，為賦風雅，是日治時期中部地區活躍的詩社之一.....他被推選為**菱香**吟社首任社長，擅長傳統古詩.....亦多次於北斗**螺溪**吟社、員林**興賢**吟社、二林**香草**吟社擔任課題詞宗，是促使**菱香**吟社茁壯有成的靈魂人物。

$n=2, l=4, W=\text{吟社} \rightarrow \{\text{菱香吟社}, \text{螺溪吟社}, \text{興賢吟社}, \text{香草吟社}\}$

精於書法，詩品俊逸。昭和5年（1930）桃園人周石輝發行《詩報》，主要刊載各地漢語文言詩人作品與全臺**各地**詩社吟會、徵詩等，戰爭末期，《詩報》重心轉為漢詩與書道，開始大量刊載書法的推廣事項，他被「詩書報書道壇發行委員會」任命為參事及臨時組審查委員，可見其書法在當時頗受肯定。

$n=2, l=4, W=\text{詩社} \rightarrow \{\text{各地詩社}\}$

Automatically recognize persons, places, events, and organizations (like poetry clubs) in the full text of biographies

Interface for Determining Kinship and Social Relationship

A

人物關係定義

李成章

李成章（中興大學農業暨自然資源學院）花壇中庄人（今花壇鄉中庄人）。父萬國，母黃金花。其父畢業於臺中師範學院，曾在中西庄組西洋樂隊，開花壇鄉西樂學習風氣之先，其後擔任彰化中級職業學校（今秀水高工）創校校長。|民國49年（1960）臺灣省立農學院（今中興大學）農藝學系畢業，民國51年在臺中區農業改良場服務，53年獲洛克斐勒獎學金前往菲律賓國際稻米研究所進修，55年取得農業植物學碩士學位。翌年返國，回中興大學任教，擔任農藝學系講師、副教授。民國61年獲國科會補助赴美國留學，64年取得加州大學戴維斯分校遺傳學博士學位。|其後，再回中興大學任教，翌年升任教授。民國72年（1983）借調沙烏地阿拉伯農業技術團，教導當地農政人員相關農業技術。74年返中興大學兼任農藝系所主任，設立當時全國第一個電腦教室，以改進生物統計教學實習設備。79年出任農學院長，任內致力提升研究水準及改善教學品質。民國86年10月，當選中興大學校長。|在教學研究方面，提出教學特優獎、青年教師研究獎、研究績優獎等三項獎勵辦法，鼓勵教師從事教學與研究。又將大學部課程重新規劃為語言課程、通識課程、專業課程三大類，在每一大類下分若干領域與學群開設。在系所擴充方面，先後增設4個學系、6個碩士班及7個博士班；民國88年（1999）8月成立獸醫學院。又為配合臺北校區改制歸併為臺北大學，特成立社會科學暨管理學院。此外，在改善教學研究設備、整修改建學生宿舍、椰林大道的美化、排水設施整建等均投注心血和大筆經費。為提高行政效率，更引進外部評估模式，推行標準作業程序方案，簡化採購流程；為因應網路時代，推動教務電腦化、選課網路化、期刊電子化，使資訊取得更為便利。而民國88年9月發生九二一大地震，中興大學受創嚴重，他帶領全校投入賑災及復健工作，受到當時總統李登輝及教育部的高度肯定。|在學術方面，對於作物育種、遺傳及生理學之研究不遺餘力，發表百餘篇學術論文，並獲得國科會甲種及優良獎，出版遺傳學詞彙及英文之水稻遺傳育種專書。民國80年（1991）獲加州大學戴維斯分校傑出校友，84年在日本獲選為第三屆亞洲作物學會（ACSA）會長，86年獲頒中華農藝學會傑出研究獎。|妻洪素娟，育有一女一男，分別為安莉、紹祺。|〔李毓嵐〕|

B

李萬國 父	黃金花 母	李登輝 社會關係
洪素娟 妻子	李安莉 長女	李紹祺 長子；第一子
1 新增姓名 請選擇	Kinship First-Born Daughter	
<input type="button" value="增加"/>	<input type="button" value="確認"/>	

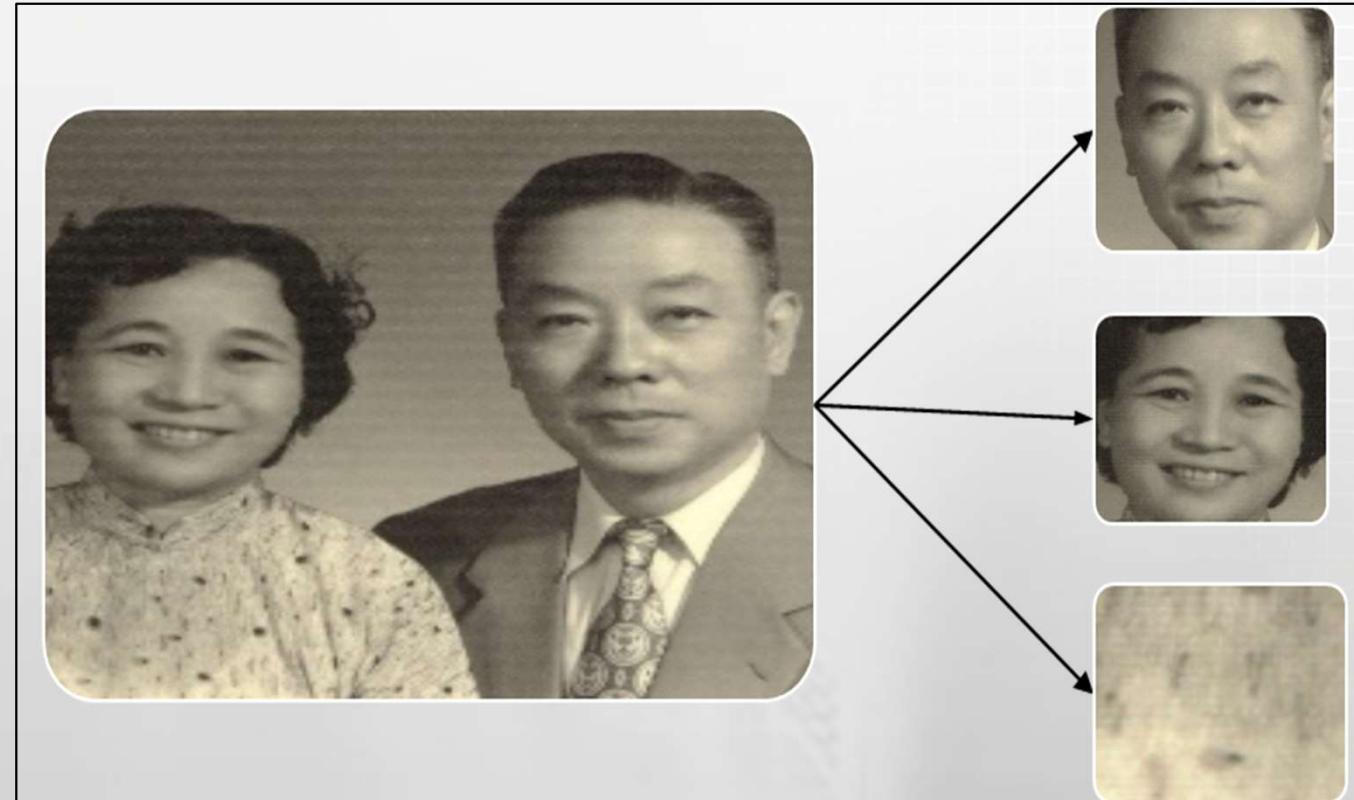
C

Person Name

Biography Full Text

Social Relationship

Face Detection from Historical Photographs



- ◆ OpenCV (<https://opencv.org/>) is used for face detection
- ◆ Precision: 32.38%, Recall: 98.16%

Face Detection and Social Network Construction

TBDB



Uploaded
Photograph

著者:

系統管理者

相片名稱:

1931年後攝於台中。葉榮鐘（後排左）與友人合影。

圖片說明:

1931年後攝於台中。葉榮鐘（後排左）與友人合影。

出版者:

系統管理者

貢獻者:

系統管理者

日期:

2021-01-18

(YYYY-mm-dd)

資料來源:

作者不詳（年代不詳）。[副系列名：友人照片]。《數位典藏與數位學習聯合目錄》。<http://catalog.digitalarchives.t>

DC-based
Metadata

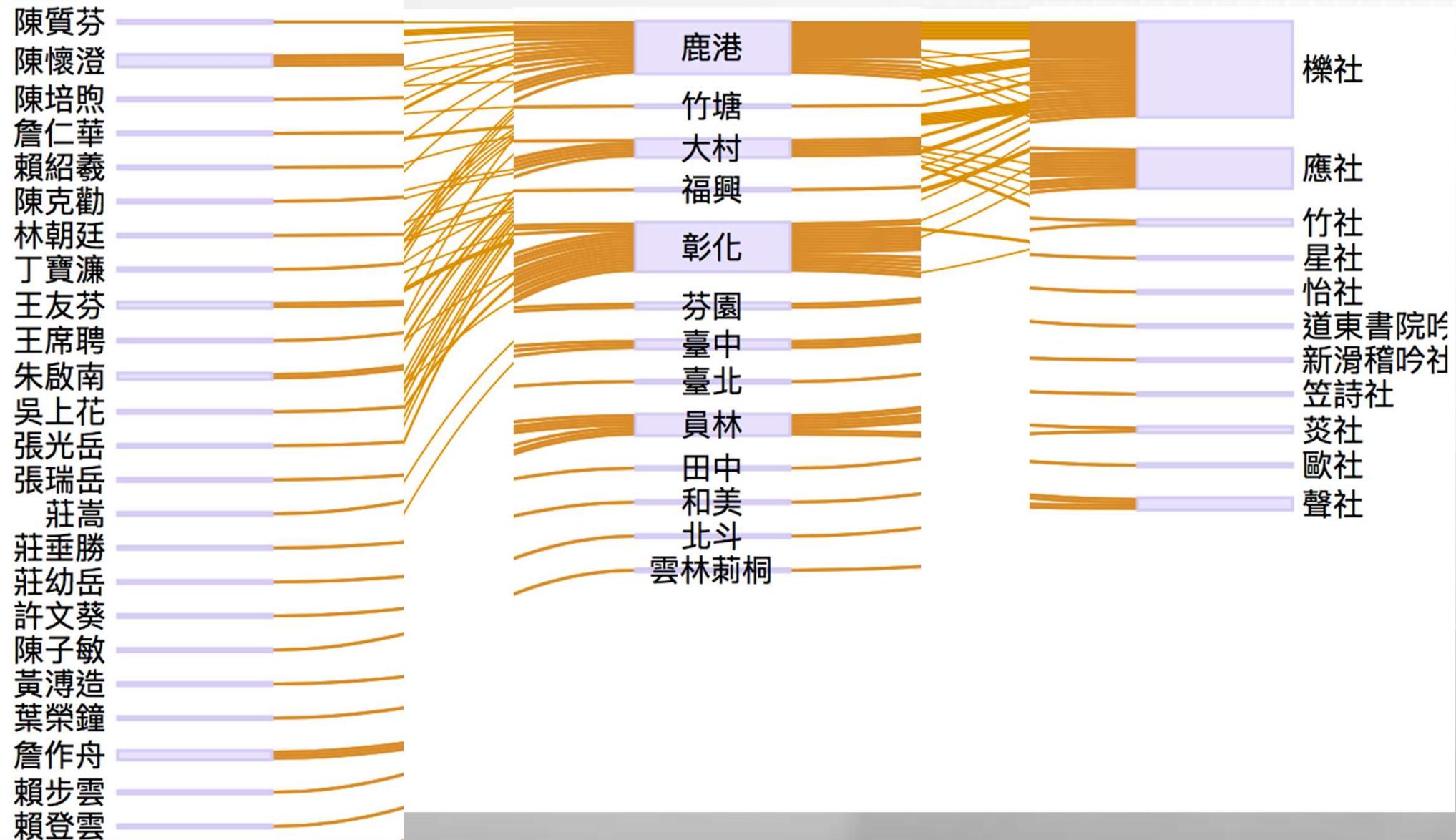
Correct faces	
False-positive faces	



Social Network Analysis (SNA)



SNA – Poetry Clubs, Places, and Members

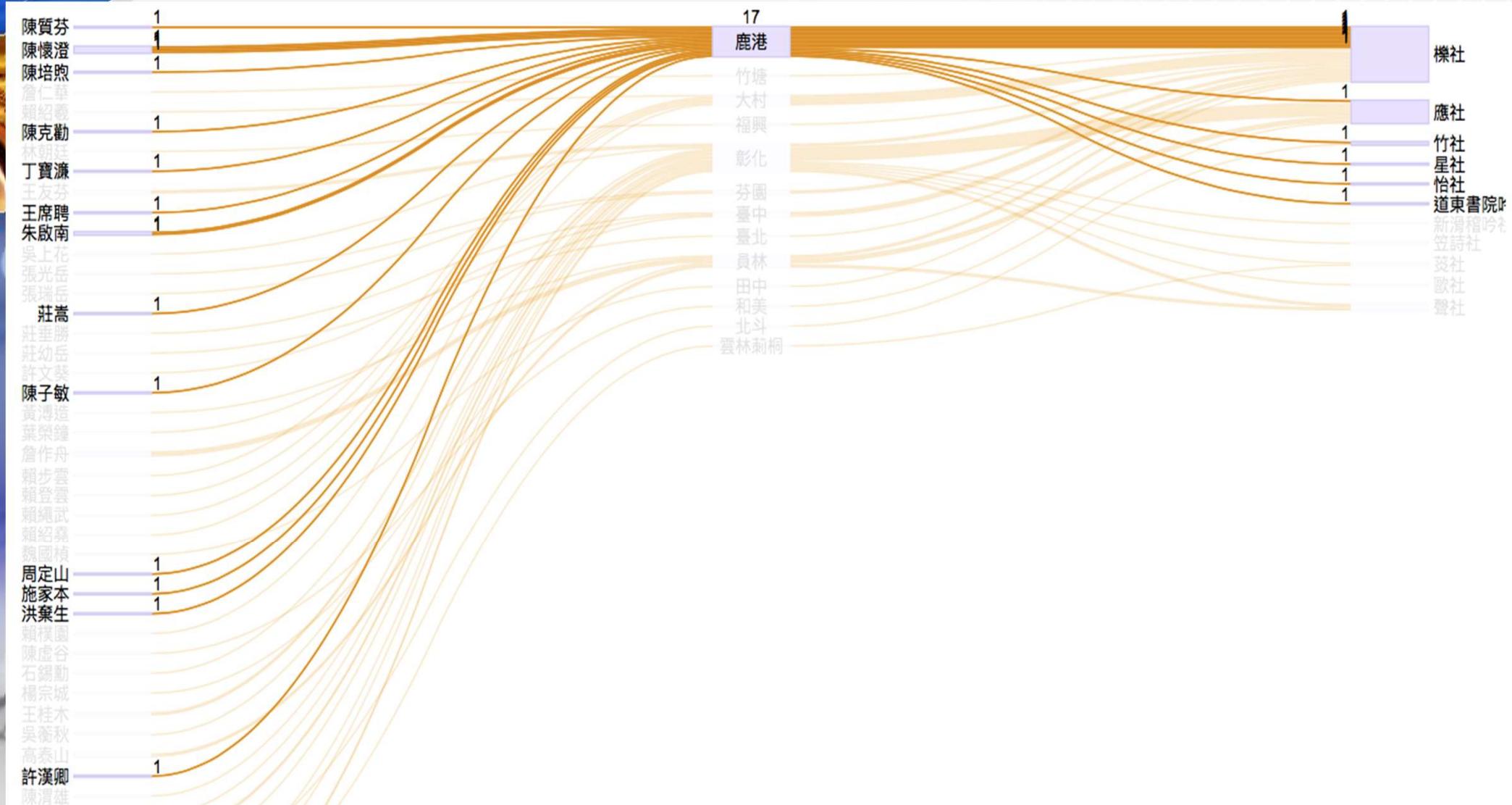


Members

Places

Clubs

SNA – Poetry Clubs, Places, and Members



Conclusions

- ▶ Follow the steps of Ontology Development 101 to create an ontology for historical persons
 - ✿ CIDOC CRM + Schema.org
 - ✿ Linked Data
- ▶ More instances to be created for each class/entity
- ▶ (Semi-)automatically populate the properties and relationships of classes/entities
- ▶ Refinement of the Ontology
- ▶ DC Abstract Model + DC Application Profiles



Thank you for Listening