



NTU's Development Strategy of Digital Collections – A Progress Report

**Jieh Hsiang (項潔)
University Librarian
Professor of Computer Science
National Taiwan University**

17 October, 2007



- **NTUL started digitizing its holdings in 1996**
- **Significantly increased the scale since 2002**
- **Made major revision of strategy in 2005**



■ Purposes of digitization

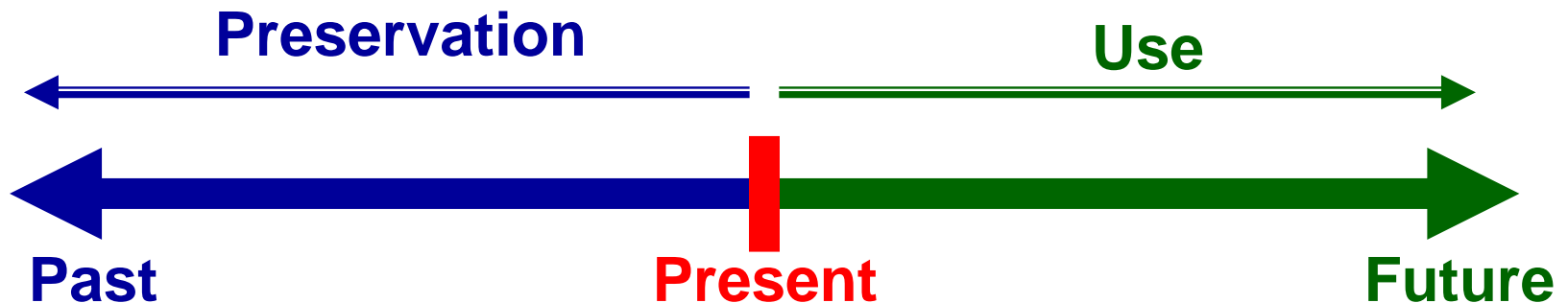
- ◆ Preservation of cultural materials
- ◆ Provide effective environment for using the mateterials

■ Who are the users?

- ◆ Faculty, students and researchers of **present** and **future**

■ For what purpose?

- ◆ Research and education





What to digitize – priority

Priority

- **Significance**
- **Uniqueness**
- **Timeliness**
- **Fragility**

Coverage

- **From within to without**
- **Starts from what are available in the university to the society at large**



What is involved in digitization

- **Metadata**
- **Images (mostly scanned)**
- **Full text** (keyed in if necessary)
- **Building systems and tools** for better use



Classification of digital collections

- **Digitization of library's collections**
 - ◆ **Digitalization of library's books**
 - ◆ **Digitalization of special collections (digital archives)**
- **Reincarnated digital collections from existing materials**
- **Digitization of external collections**
- **Contemporary materials**
 - ◆ **Institutional Repository**
 - ◆ **Born digital materials (such as web archiving)**



Digitization of Books

- **Only rare books, mostly in Chinese but some in western languages and in Japanese**
- **No project on massively digitizing common books such as done in the Million Book Project**



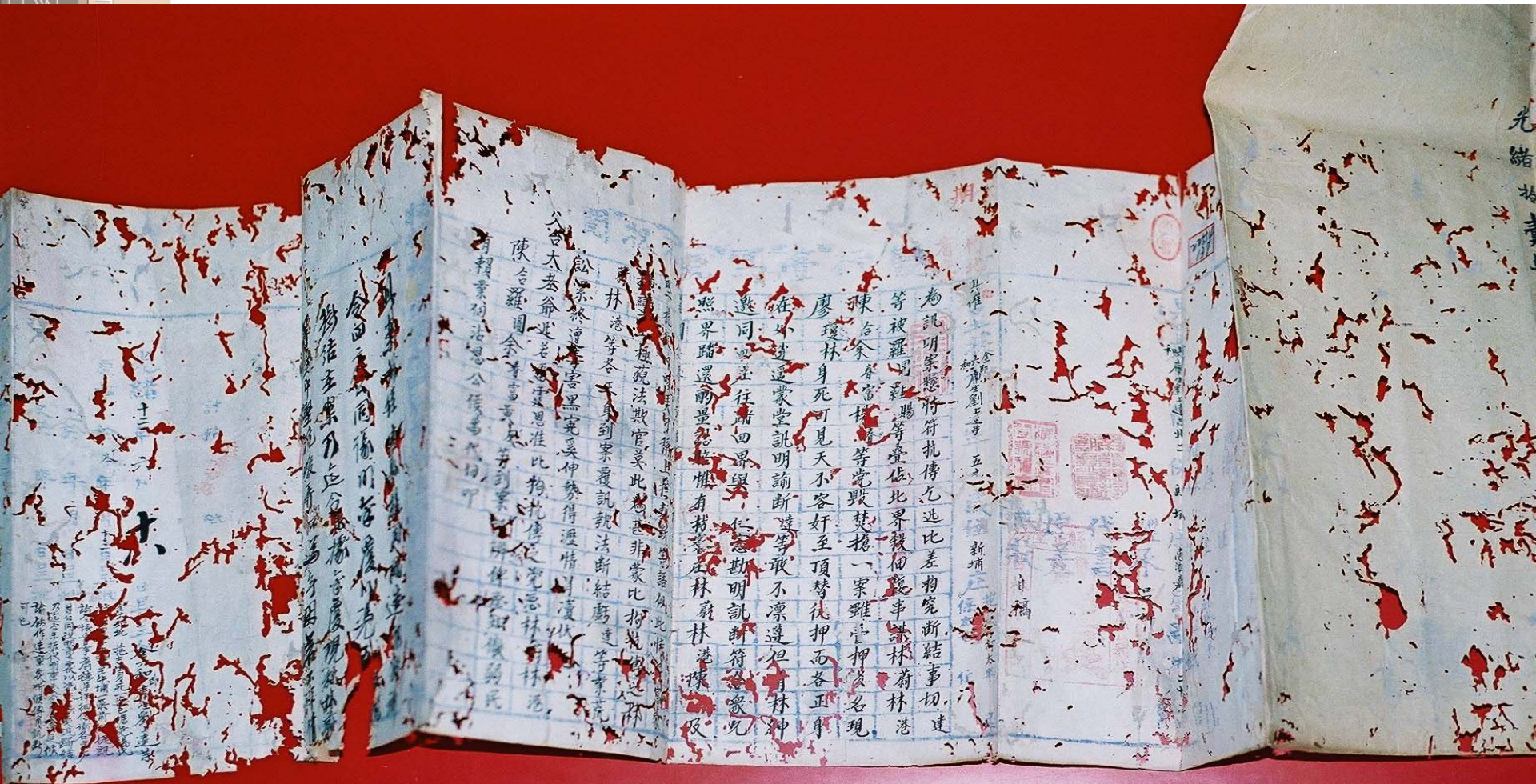
Digitalization of special collections

- All major historical archives are either completed, in progress, or in the pipeline
- **Completed:** Include Dan-Xin Archives (淡新檔案), An-Li Dashe Archives (岸裡大社文書), Ino Archives, the Taiwan Stone Rubbings Collection, The Local Religious Relicts Collection, etc.
- **E.g. Dan-Xin Archives**
 - ◆ Administrative and judicial archives during Qing Dynasty
 - ◆ Including documents of some 1,600 legal cases, containing as many as 19,152 items
 - ◆ Dating from 1776 to 1895
 - ◆ One of the only two judicial archives in China at the prefecture archive
 - ◆ An important first-hand resource for the study of Taiwan economics, society structure, and law
- **What we have done**
 - ◆ Image, metadata, full text, associated tools

A document from Dan-Xin – Archives before repairing



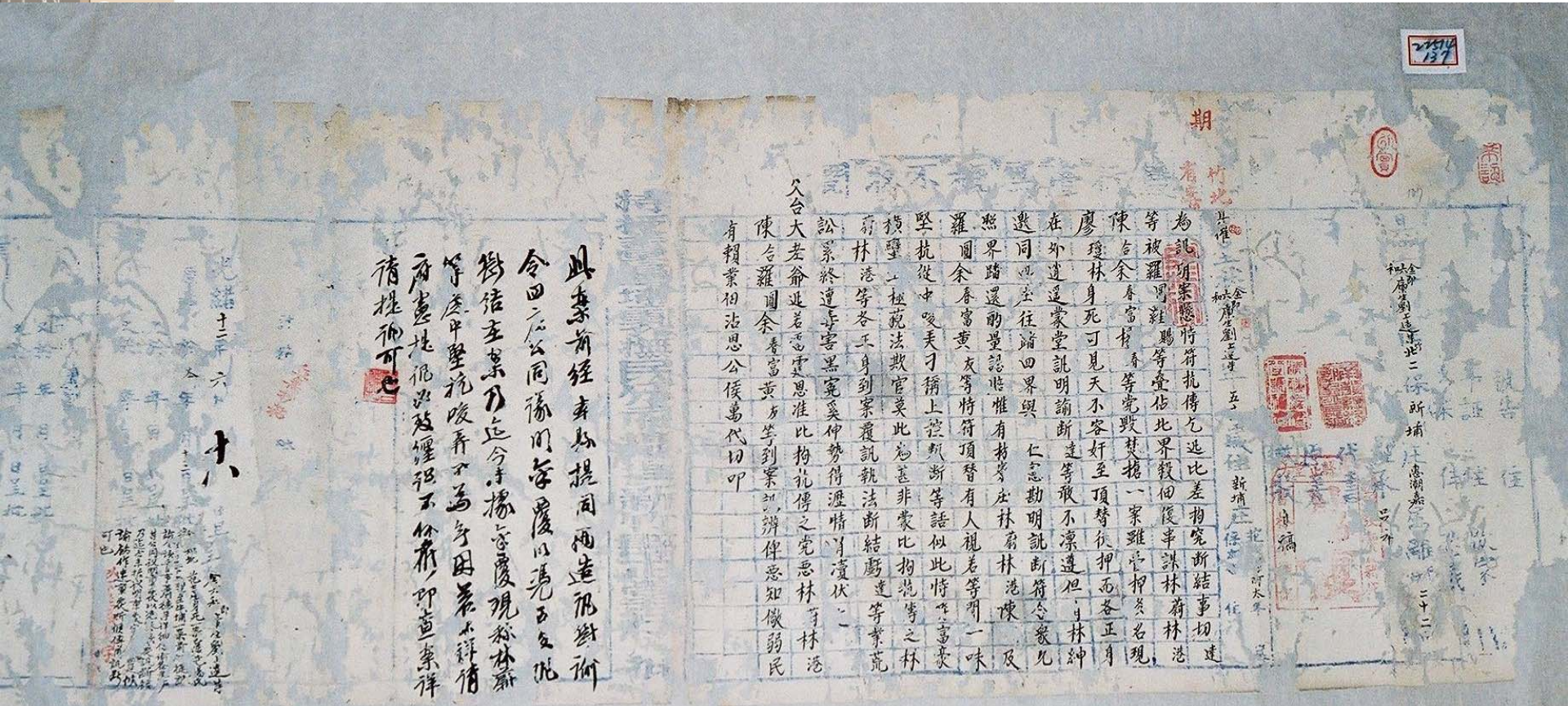
NAT



A document from Dan-Xin – Archives after repairing



NATI



Reincarnated digital collections from existing materials



■ Methodology

- ◆ Identify a specific topic, extract material from existing holdings, and create a new collection
- ◆ The digital collection is created from the physical collection, but there is no one-to-one correspondence between the two

Before: item → one book / document

After: item → one image / paragraph

Collection of Old Photographs of Taiwan – an example



- **There was no known systematic collections of pre-1945 photographs of Taiwan**
- **Most of the pre-1945 photos of the Central News Agency, the largest holding agency of photographs, are of China, not Taiwan**
- **The lack of photo images became a problem for people who wanted to study the Taiwanese society during Japanese occupation**



Old Photographs of Taiwan

- **Our librarians went through over 5,000 old Japanese books about Taiwan and selected about 40,000 photos**
- **Scanned the images and created metadata records**
- **The selection process took a month. The digitization work took another two**

Old Photographs of Taiwan

Project – outcome



- **The most comprehensive collection of photographs of its kind in the world**
- **Covers a wide spectrum of areas including society, education, commerce, politics, transportation, industry, people, culture, religion, architecture, landscape, plants, animals, etc**
- **Easier to use than many photo collections because of metadata extracted from books**

Old Photographs of Taiwan



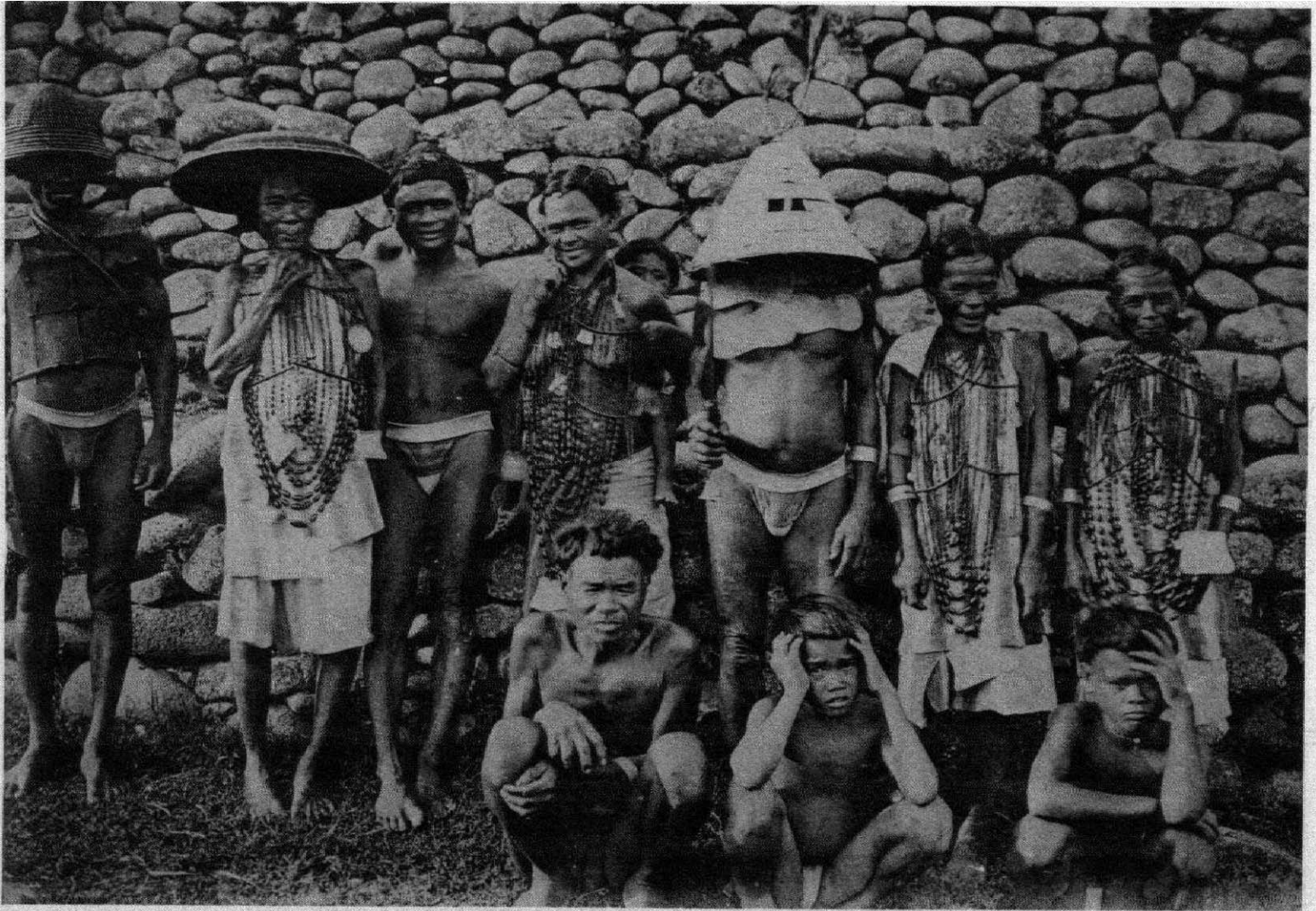
NATIONAL TAIWAN UNIVERSITY LIBRARY



2007/10/17
PRDLA2007

NTU's Development Strategy of Digital Collections: A Progress Report (Jieh Hsiang)

Old Photographs of Taiwan



Old Photographs of Taiwan



NATIONAL TAIWAN UNIVERSITY LIBRARY



200
PRDLA2007

Old Photographs of Taiwan



NATIONAL TAIWAN UNIVERSITY LIBRARY





Old Photographs of Taiwan



Reincarnation from static materials – other collections



- A full-text digital library of primary documents that cover more than 80% of all court material concerning Taiwan during Ming and Qing Dynasties (pre-1895) – over 35,000,000 words and about 40,000 items (a paper to appear in ICADL available upon request)
- A full-text database of the Chinese version of Taiwan Riri Xinbau (台灣日日新報) from 1905 to 1911
- ...



Digitization of external collections

Approach

- **Through collaboration with other archive holders, digitize important collections and make them available to researchers**
- **Convince our staff and the university that one can consider a digital collection as special collection even if one doesn't own the physical material**



Digitization of external collections

- E.g. The archives of the Nationalist Party
- On-going project. The “Special Archives” (特種檔案) during the first year, 200,000 images

KMT Archives



NATLIC



中國國民黨史料資料庫
KMT Digital Archives Database

首頁 瀏覽 查詢 內容簡介

您的位置 首頁 > 查詢 > 查詢結果

- 查詢
- 一般查詢
- 進階查詢
- 分群

查詢結果

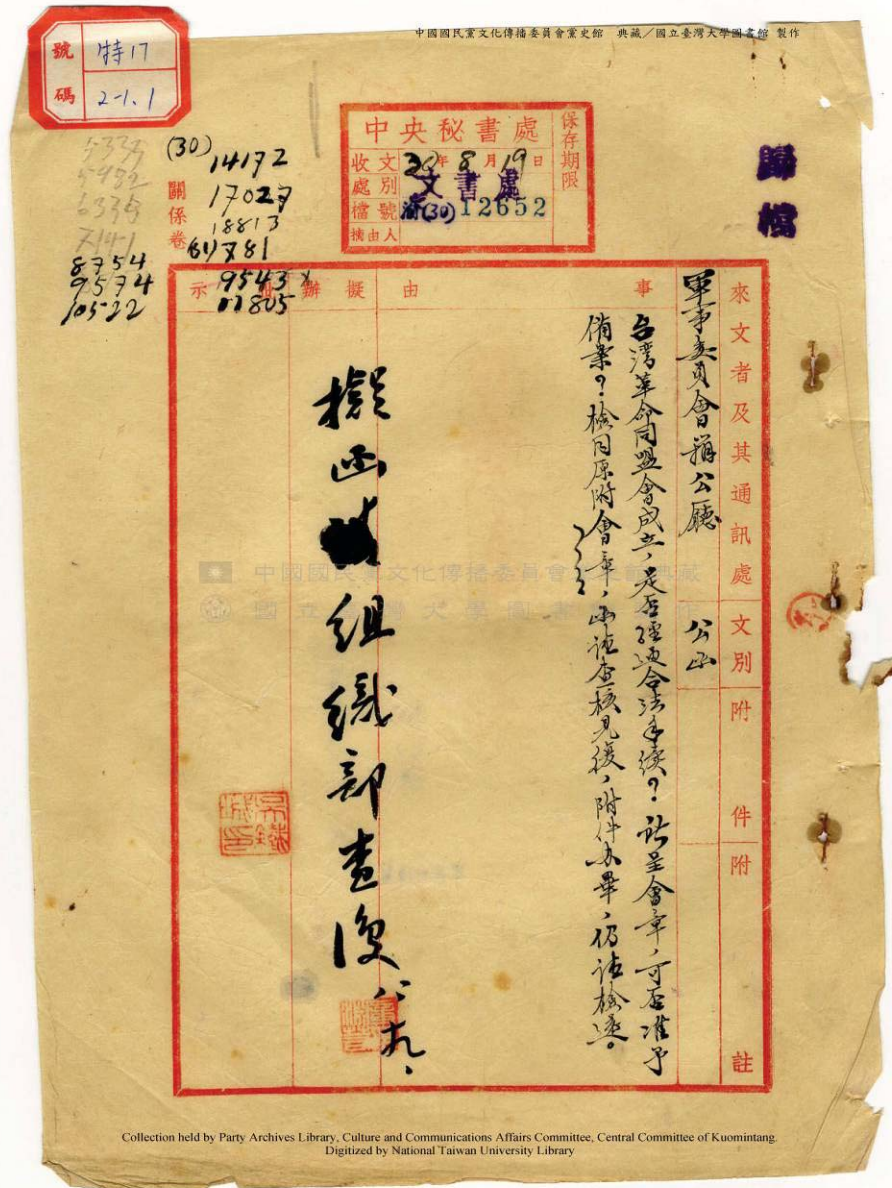
每頁顯示資料筆數 筆 共 39筆資料

序號	檔號	題名	作成日期/出版日期
21	特172.1	台灣革命同盟會成立案	民國30年8月16日(1941-08-16)
22	特172.2	軍委會辦公廳致中秘處函	民國31年1月15日(1942-01-15)
23	特172.5	台灣革命同盟會上蔣總裁呈	民國31年4月10日(1942-04-10)
24	特172.7	台灣革命同盟會會章及職員表	民國31年(1942)
25	特172.8	台灣革命同盟會上中央黨部呈文	民國31年6月23日(1942-06-23)
26	特172.9	孫科致吳鐵城函	民國31年11月4日(1942-11-04)
27	特172.16	台灣革命同盟會工作報告書	民國32年6月23日(1943-06-23)
28	特172.22	台灣革命同盟會第二屆大會宣言	民國31年4月5日(1942-04-05)
29	特172.30	台灣革命同盟會黨務李友邦等上吳鐵城呈	民國32年6月23日(1943-06-23)
30	特172.31	張壽賢上吳鐵城呈	民國32年10月22日(1943-10-22)

查詢：台灣革命同盟會
結果：39筆

2007/10/17
PRDLA2007

台灣革命同盟會成立案



Collection held by Party Archives Library, Culture and Communications Affairs Committee, Central Committee of Kuomintang. Digitized by National Taiwan University Library



The Chi-Lin foundation Archives

- **430 titles, 7,000 volumes of Taiwanese magazines during 1957-2004, many of which banned during the Marshall Law period. NTU doesn't have 293 of the titles (neither do other libraries in Taiwan).**
- **Newspaper clippings of social movements during 1951-2006**
- **Total 400,000 pages**

臺灣年代

1

目錄

慈林教育基金會 — 台灣社運史料



慈林教育基金會 典藏 / 國立臺灣大學圖書館 自由談



台灣獨立 是追求民主化的 唯一活路

對蔣氏王朝結束後台灣政局的思考

■陳台篤

國會全面改選，事實上已是台灣人民追求政治民主化最後的請求，除此之外，地方自治選則遲遲未結，省長、院轄市市長不能民選以及人民沒有行使自決的權利，也是台灣民主化的兩大障礙。現在的問題是，什麼是台灣人民追求民主化的有效手段？什麼才是解決台灣國際地位問題，避免國民黨與共產黨聯手出賣台灣人的最佳途徑？

蔣經國非死一週以來，台灣表面上看來好像平靜無事，但事實上却是暗潮洶湧，尤其是國民黨，在蔣氏亡故後及一週內，內部傳出強派權力鬥爭的傳言，我敢說人說謊，他惡一黨內會變成不可收拾的社、動亂，惟此之際，台灣應加速民主化，已成為一項國人必須正視的嚴重課題。

國民黨長年一黨獨統下的台灣，橫梗在眼前的民主障礙至少有以下三項，第一、現存三個由大多數無黨派基礎的民意代表組成的民意機構，非代表人民行使權。第二、地方自治無法制度化，省市(安撫)官不能選，第三則是人民決定自選國家前途的權利被剝奪。

就中國共產黨提出自己的中國國民黨，選擇台灣之後，為了其統治上方便，四十來一連串持三不從不改的國會，代表人民行使權，其間雖然為了安撫台灣人民對的政改願，而局部提出部份多額，供令人爭取，但基本上，仍由大對壟台的「一兩派」老法統」把持多數，讓強派自衛民意，在這種結構下，台灣人追求民主或為一種「事實上」的不可能。雖是之故，打破現有國會結構，重新進行全面改選，這是台灣走向民主化，所必須跨越的第一道關。

目前強派台灣最大反對黨一民運

Collection held by Chin Foundation
Digitized by National Taiwan University Library

66

公論時代週刊

2007/10/17
PRDLA2007

慈林教育基金會 典藏 / 國立臺灣大學圖書館 自由談

■鄭南榕

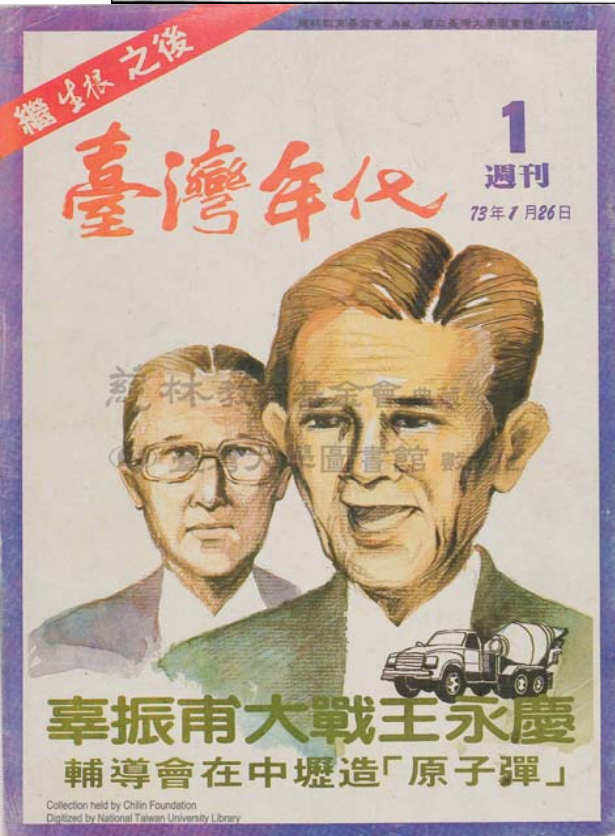
揚棄威權統治文化 尋求台灣政治活路

蔣經國雖然逝世之後，全島的擁擠擁擠氣氛，一律改以「黑白片」出現，然而這位國家最後主人的去世所引發的權力鬥爭，在一陣應殺的氣氛中，却時時透出濃濃的色彩。

在電視、報紙的新聞片斷裏，我們看到伏地跪拜的鏡頭，也看到有人神傷喪，哀哀痛哭。蔣經國死，是有些人感到如喪所依，卻他死了，他們失去期待，中心無主，不覺悲從中來。另外我們也看到大多數人的反應，他們在一陣陣得之後，雖然不安心惶惶，却也不再表示什麼，他們在查看報、看電視、聽廣播、自語、自語、自語……

1 Collection held by Chin Foundation
Digitized by National Taiwan University Library

NTU's Development Strategy of Digital Collections: A Progress Report (Jieh Hsiang)



王慶大戰南振甫 輔導會在中壢造「原子彈」

Collection held by Chin Foundation
Digitized by National Taiwan University Library

封面故事——王慶大戰南振甫

- 從「錢力」火拼到「權力」大戰——王慶、水泥大戰「青景」分析
- 拆穿崇德、水龍大戰的幕後真相
- 山不青、水不流、鳥不語的太魯閣？

李桓 4
曉林 6
江靜惠 13

2

台灣年代週刊

25

Judicial court records during Japanese period (1895-1945)

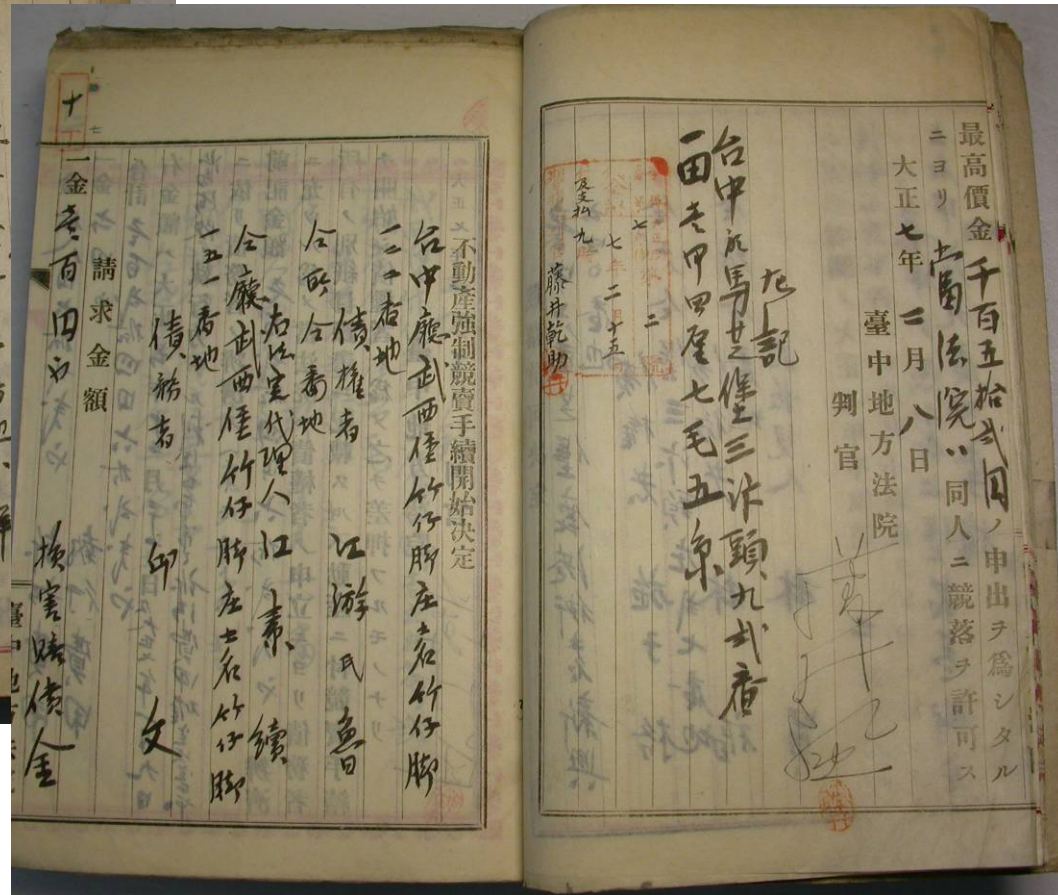


- **The court records of the judicial courts of Hsin-Chu, Taichung, Chia-Yi, Taipei, and the Superior Court**
- **5,460 volumes, almost destroyed due to lack of care**
- **Connects the link between Dan-Xin Archives (Qing court records) and the post-war court records**
- **Digitized using digital camera, already over 4 Terabytes**

Judicial court records during Japanese period (1895-1945)



NAT

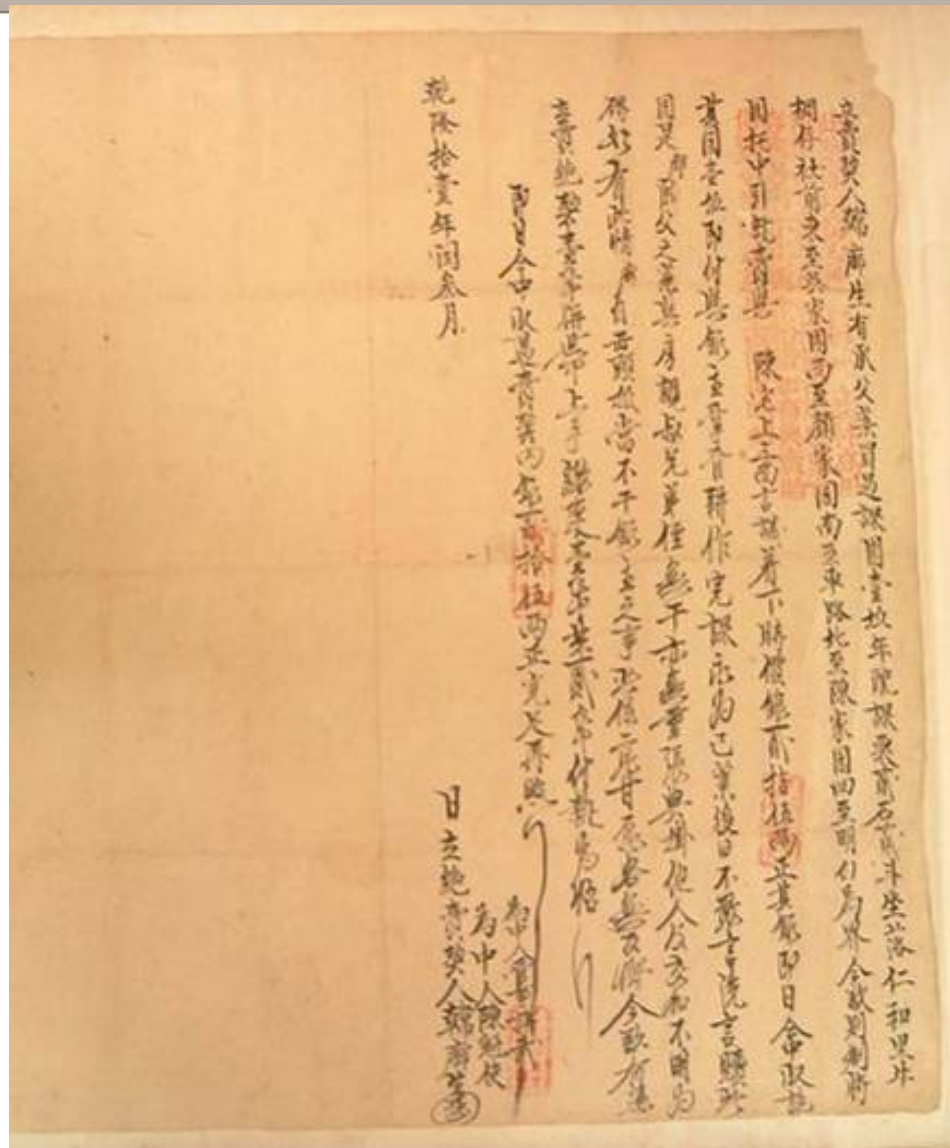




Several collections of old deeds

- **Numbered at about 1,000**
- **Metadata, images and full-text**
- **A full-text DL of 25,000 old deeds (not all done by us)**

Found a forgery





Digital collections at NTUL

- **Over 20 special collections**
- **Totally over 500,000 metadata records, 4,000,000 images, and 200,000,000 Chinese words in full text**
- **List available (in pdf) upon request**



Contemporary materials

- **Institutional Repository**
- **Web Archiving**



Institutional Repository

- **Based on DSpace 1.4, with significant enhancements**
- **National Taiwan University Repository (<http://ntur.lib.ntu.edu.tw/>)**
- **Over 33,000 metadata records**
- **About 16,000 in searchable pdf**
- **Just initiated the IR30 project, to help at least 30 universities in Taiwan to build their IRs**

搜尋範圍 全部NTUR

搜尋 進階搜尋

我的NTUR · 上傳 · 說明 · 關於NTUR · 管理

瀏覽全部

- 社群與類別
- 題名
- 作者
- 日期

相關新聞

校內推廣說明會
(9/27-10/25) **New**
 臺大機構典藏實施辦法
 臺大機構典藏系統文件
 臺灣機構典藏學術研討會
 (2007/08/31)

臺大機構典藏 NTUR > 主頁



+ NTUR中的社群 [15894/32543] [全文筆數/總筆數]

- + [文學院](#) [764/1368]
- + [理學院](#) [1831/4112]
- + [社會科學院](#) [353/726]
- + [醫學院](#) [3270/5882]
- + [醫學院附設醫院](#) [610/651]
- + [工學院](#) [2330/5446]
- + [生物資源暨農學院](#) [1335/2963]
- + [管理學院](#) [550/1005]
- + [公共衛生學院](#) [230/475]
- + [電機資訊學院](#) [4203/5929]
- + [法律學院](#) [155/323]
- + [生命科學院](#) [444/1143]
- + [其他校內單位](#) [429/698]

上傳排行

下載排行

社群

- + [電機工程學系](#) [2695/3693]
- + [資訊工程學系](#) [940/1262]
- + [物理學系](#) [637/998]
- + [機械工程學系](#) [511/1216]
- + [醫學系內科](#) [335/788]

作者

- + [陳良基](#) [181/210]
- + [郭華仁](#) [139/146]
- + [傅立成](#) [128/174]
- + [貝蕪章](#) [100/138]
- + [郭斯彥](#) [86/107]



Web Archiving

- **Born digital**
- **Websites are an important record of our time, but extremely fragile**



★ 關鍵字檢索

網頁 網站

送出

關於臺大圖書館網頁典藏庫：

隨著網際網路的流行，網路成為發表資訊的一個熱門途徑，它也成為資訊收集的一個重要來源，然而，網路資訊愈是流通快速，網路資訊的保存卻成為一個燙手的問題，根據調查，平均一個網站的生命週期大概只有44天，而重要的資訊，例如學者發佈的文獻資料、文化或科學資源這些寶貴的資訊，可能就因為時間的更迭而從此消逝掉了。

臺灣大學圖書館，身在臺灣第一學府~臺灣大學，自然負有保存文史資料的重大責任，因此成立了 NTU Web Archive網站，主要任務為負責「選擇」、「收集」、「保護」台灣大學、台灣政府機構、台灣政治團體、焦點人物、原住民族及重大事件之原始性網頁資料，並提供檢索服務以滿足當代與未來世代研究發展之需求。冀望這樣的開展能夠符合現實環境之需求，為台灣大學善盡保護網路資產之責任。

本館共典藏 2658 個網站，包含：

- 306 個台大相關網站 [More...](#)
- 690 個政府機構網站 [More...](#)
- 91 個事件議題網站 [More...](#)
- 86 個焦點人物網站 [More...](#)
- 417 個教育學術網站 [More...](#)
- 424 個藝文相關網站 [More...](#)
- 153 個政治相關網站 [More...](#)
- 277 個科技經貿網站 [More...](#)
- 250 個族群相關網站 [More...](#)
- 148 個非營利社團網站 [More...](#)

典藏縮影



讀者推薦

最新典藏 Latest Updated

中華郵政全球資訊網	2007-05-10
國立臺灣大學圖書館	2007-05-09
國立臺灣大學法律暨社...	2007-05-09
國立臺灣大學醫學院暨...	2007-05-09
國立臺灣大學附設醫院...	2007-05-09
臺大科法中心—科技倫...	2007-05-09
國立臺灣大學公共衛生...	2007-05-09
國立中正紀念堂管理處	2007-05-09
臺灣大學國際華語研習...	2007-05-09
臺灣大學文學院語文...	2007-05-09

熱門點選 Most Popular Links

國立臺灣大學	273
臺灣大學哲學系	204
中華民國醫用超音波學...	107

NTU web archive system

<http://webarchive.lib.ntu.edu.tw/>

Archiving over 3,000 websites

National Taiwan University (1997/07/15)



國立臺灣大學 / National Taiwan University - Mozilla Firefox
檔案 (F) 編輯 (E) 檢視 (V) 瀏覽 (G) 書籤 (B) 工具 (T) 說明 (H)

http://webarchive.lib.ntu.edu.tw/19970715/www.ntu.edu.tw/index.html



椰林記事 ● English
學術研究單位 ●
行政服務單位 ●
臺大新聞網 ●
臺灣學術網路 ●
虛擬世界之旅 ●
社會公益 ●

歡迎自八十五年九月九日起之本校第 888888 位貴賓！

[請您批評與指教](#) ● [常見問答集](#)



[[最新消息](#) | [椰林記事](#) | [學術研究單位](#) | [行政服務單位](#) | [台大新聞網](#) | [臺灣學術網路](#) | [虛擬世界之旅](#) | [社會公益](#)]

校總區地址：台北市大安區10617羅斯福路四段一號

總機：02-3630231

傳真：02-3627651

版權所有

國立臺灣大學“全球資訊系統”內以任何媒體呈現之所有內容—包括文字、音訊、視訊及版面設計—著作權屬“國立臺灣大學”。任何形式之轉載，請先與我們聯繫。

Copyright (c) 1997 National Taiwan University ALL RIGHTS RESERVED

National Taiwan University (2007/07/15)



國立臺灣大學 - Mozilla Firefox
檔案 (F) 編輯 (E) 檢視 (V) 瀏覽 (G) 書籤 (B) 工具 (T) 說明 (H)
http://webarchive.lib.ntu.edu.tw/20060911/www.ntu.edu.tw/chinese2006/index.htm

新生 · 學生 · 教職員 · 訪客 · 校友
舊版首頁 | 網站導覽 | 搜尋 Go!



認識臺大 學術單位 研究發展
行政組織 常見詢答 服務資源 English

圖書館 · 虛擬博物館 · 國際交流 · 課程 · 招生 · 推廣教育 · 行事曆 · INFO

>>> 活動快報

>>> 校園焦點

>>> 最新消息



>>計資中心資訊應用課程

>>電機系電子計算機訓練班

>>資訊系統訓練班

>> 第二屆亞洲與WTO政策與法律國際研討會



許倬雲院士(左1)、李嗣涇校長(左2)及多位專家學者參與人文社會高等研究院定期舉辦的腦力激盪(brain-storming)活動,藉著開放性的討論,尋找人文社會領域更多元的研究可能。

提昇人文社會研究 落實實際整合願景

為推動「邁向頂尖大學」計畫,本校積極整合並提升本校人文社會領域之創新研究,促進校內及跨校人文社會研究人才間之交流合作,臺灣大學成立「人文社會高等研究院」,預定於2006年9月29日舉行揭牌儀式,正式運作... >> [全文](#)

- ◆ 新生入學校園導覽人員培訓報名
- ◆ 9/3~11/5發現臺大實驗林之美-綠色寶藏展覽活動
- ◆ 化學系與日本岡山大學進行國際交流及學術研討
- ◆ 經典人文學程接受申請
- ◆ 亞洲固態線路研討會 9/7舉辦記者會
- ◆ 9/4 即時疾病偵測及通報成果發表
- ◆ 校內徵才

>>其它最新消息

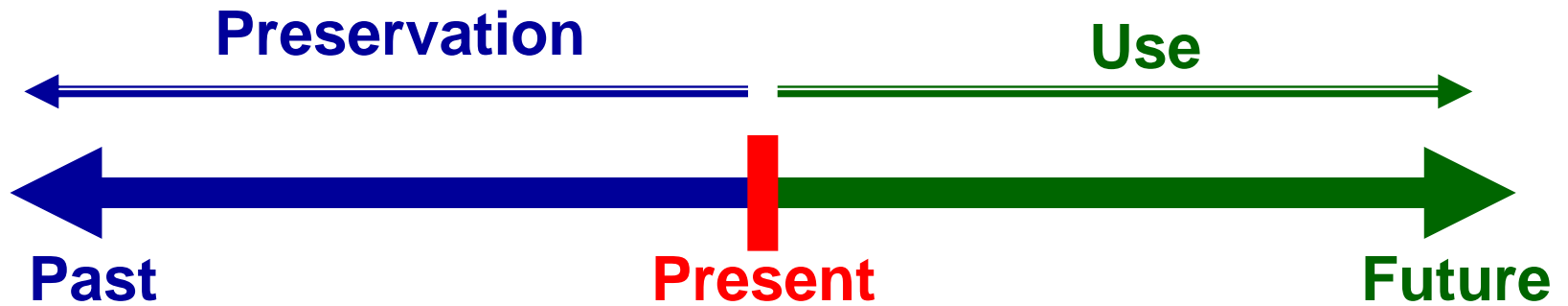
校園公告

邁向頂尖大學

統計年報

學雜費專區

用電資訊



- Digitization of special collections
- Reincarnated digital collections from existing materials
- Digitization of external collections

- **Born digital collections**
 - ◆ Institutional repository
 - ◆ Web archiving
 - ◆ Free academic web resources
 - ◆ Free e-books

- From OPAC to integrated content management system
- Pushing our content into gravitational hubs
- Incorporate services on the web into our systems



Thank you