DIGITALLY ARCHIVING ADVERTISING EPHEMERA:
THE CASE OF THE CHINESE COMMERCIAL ADVERTISEMENT ARCHIVE (CCAA) (1880-1940)

Jing Chen, Nanjing University & Rice University

2014 International Conference of Digital Archives and Digital Humanities
近代中國報紙廣告

在華文報紙的早期，“大多數報紙的廣告都是大型外國海運公司、保險公司和貿易公司發布的告白。報紙廣告收入比訂閱費的獲利要少。中國商業尚未開始轉型，還不需要在報紙上刊登廣告”
——白瑞華著，王海譯。2011。中國報紙（1800-1912）。广州：暨南大学出版社。

据学者统计，《申报》、《新闻报》、《沪报》、《时报》、《大公报》、《益世报》、《中国报》、《忠言报》、《时闻报》等报纸广告版面逐渐超过了50%以上
——苏士梅。2006。中国近现代商业广告史。郑州：河南大学出版社。

The earlier newspapers which, like the Shun Pao, were designed to earn profits did solicit advertisements but got modest returns. For years, the mainstays of the daily advertising columns were the large foreign businesses, the shipping and insurance companies at first, and later the oil, tobacco, liquor and patent medicine companies. Advertising rates long remained a meatier of bargaining. ....... Medical advertising, covering quack doctors and generalists, philters and patent medicine, became the largest classification. Advertising agencies began to develop some years after the revolution, and a degree of uniformity and system gradually entered. By 1930 patent medicine advertisements had begun to lose the lead.

我国广告事业，年有进展，自为可喜之现象。如《申报》、《新闻报》、《益世报》之经济充裕，不可谓非广告之赐。然就上列各表观察，则外货居十之六七，国货仅十之二三。
——戈公振，《中国报学史》，1927
近代中國報紙廣告

这些“洋货”广告中涉及门类众多，包括药品、日化用品、乳制品、谷物、糖果、食物、化肥及杀虫剂、机动车、烟草、化工产品、纺织品、饮品、橡胶制品、机械、摄影器材、电子设备、钟表、乐器、纸制品和办公设备等三十多个类别。商品生产公司则主要来自英国、美国、加拿大、法国、日本、德国、印度、意大利、荷兰、瑞士、瑞士、新加坡等国。

（註：此分类标准是依据了“中国商业广告档案库”的分类，而非民国时期的商品分类。）
• For example, researcher investigates the cost of advertising of BAT (British American Tobacco) on 47 Chinese and English newspapers and magazines in 1924 and finds that BAT spend 100486.32 yuan on ads on Chinese newspapers in Shanghai, that is, 76.7% of the total cost, 131009.35 yuan. And in 1922 and 1923, the cost of advertising of BAT respectively is 2,146,351.14 yuan, 2.66% of the gross sales and 2,015,536.70 yuan, 2.10% of the gross sales.

——秦其文。2010。中国近代企业广告研究。北京：知识产权出版社
CCAA之理論基礎

- Tani Barlow 教授探讨了近现代中国广告中反复出现的“摩登女性”形象及以广告为代表的“蜉蝣”材料的研究意义，她将广告、中国现代商品社会的出现和本土社会学（借助中国启蒙社会思想来解释日常生活、阐释精英化社会关系与社会实践等的思想或著作）放置于特定的历史语境中进行考察，以包括Cutex（蔻丹）、BAT（英美烟草公司）、双美人牌化妆品、安生生电灯泡、卜内门化肥等商品的一系列广告为例说明商品在现代中国社会是如何与社会进化思想联系起来的，本土社会学又是如何与性感的摩登女性的图像建立关系的。广告又是如何将这种对现代性的另类的幻化图像自然化的。进一步的，她还指出女性主体在现代的出现及其同时期内呈现出的不同形式是独特性和多样性的统一，抽象化的女性不仅仅出现在政治理论，比如中国女性主义思想，也出现在具体的广告及社会学著作中，新女性的视觉呈现与理论（比如社会学）的“女性”范畴之间实为互补。Barlow教授在强调广告不是现实的反映或再现的同时，也指出广告，这种商业艺术，是一种体验的自我指涉（self-referentiality），“摩登女性”提供了一种从视觉艺术的角度去思考什么引起了女性的事件（the event of women），广告的背后是资本（普适价值）及其在中国本土的转化。她继而提出，商品－女性关系具象化为都市广告文化，而广告文化也为女性如何在“使用价值的另类景观”得到了生动呈现，这一角度为殖民现代性在亚洲的探讨供了另一种可能
• 出现在中国近现代的、基于纸质印刷品的广告是一种被称为“蜉蝣”的东西，是短暂即逝的，是文字和图像的综合体，是某一社会、历史时期内商业、文化活动的记录
  as an ephemera, advertisements embedded into the print paper medium is “the minor transient documents of everyday life” as well as mixture of text and images jointly recording activities in specific period and society

• 广告作为一种“使用价值的另类景观”，具象化了商品－女性关系，生动展示了女性是如何成为消费文化的主体性，通过幻化出的美好生活图景刺激中国的消费者以消费商品的方式来满足其对现代生活的幻想
  as “other scene of use value”, advertisement in Modern China concretizes the commodity-female relationship, vividly exhibits how female (modern girl) became the subjectivity of consumer culture and satisfies the consumer’s dream of modern life by fantasize the beautiful illusion of life to simulate consumer to buy commodity

• 广告并不是真实社会的反映或再现，也不是欧洲或者西方现代主义影响／殖民中国的证据，而是中国现代社会经验（同时面向过去和未来）的图像化
  advertisement is not the reflection/representation of real society or evidence of influence of European or Western modernism on China but the imagination of experiences of Chinese modern society(facing pass and future)

• 广告却是全球商业资本、现代跨国公司和本土社会学思想在特定时空中接合而产生出的真实之物，是得以探寻中国现代殖民性消费市场、跨过资本和消费主义的。
  advertisement is the real thing produced by the articulation of global capital, Modern international company and vernacular sociology in specific period
從元數據到概念網絡/From Metadata to Concept Network

(A graph of Database Structure)
## Chinese Commercial Advertisement Archive (1880-1940)

<table>
<thead>
<tr>
<th>NEWSPAPER</th>
<th>NUMBER</th>
<th>Resource</th>
<th>Original Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hankow Times</td>
<td>12,961</td>
<td>Shanghai Library; Peking University Library</td>
<td>Microfilm</td>
</tr>
<tr>
<td>Yuehua Bao</td>
<td>4,719</td>
<td>Sun Yat-Sen Library of Guangdong Province</td>
<td>Microfilm</td>
</tr>
<tr>
<td>Shengjing Shibao</td>
<td>54,804</td>
<td>The National Library of China</td>
<td>Microfilm</td>
</tr>
<tr>
<td>Dagong Bao</td>
<td>60,000</td>
<td>Library of University of Washington, Seattle</td>
<td>Microfilm</td>
</tr>
<tr>
<td>Shen Bao</td>
<td>60,000</td>
<td>Shanghai Library</td>
<td>Microfilm</td>
</tr>
</tbody>
</table>
Chinese Commercial Advertisement Archive (1880-1940)
Chinese Commercial Advertisement Archive (1880-1940)

• Theoretical objective
  ➢ to create a means of drawing general pictures of social-economic change in major Chinese treaty ports and their marketing catchment areas, between the late 19th and mid-20th centuries.

• The first set of goals
  ➢ collecting and digitizing advertisements collected from the five representative newspapers
  ➢ designing and laying the foundations for a searchable, online archive/database
  ➢ placing all images and their metadata in archive, free for use and publicly accessible online.
選擇標準 / Selection Criteria

- 首先，廣告中必須有明确的商品名称或生产公司，商业服务广告或者公共信息广告则不包括在内；其次，是否为外国商品，则需要找到明确的历史信息，比如外国公司历史记录。因此广告之外的历史档案和研究资料也非常的重要；第三，时间范围上，考虑到广告对市场和时空的依赖性，CCAA选择了从这五份报纸创刊时间到分别因为战乱而停刊或迁移的时间。
- Shanghai, Shen-pao (申报), Tianjin’s Ta Kung Pao (大公报), Shenyang’s Sheng-ching Shih-pao (盛京时报), Hankou’s Hankow Times (汉口中西报) and the Yuet Wa Po (越华报) in Guangzhou
視覺廣告與知識生產 / Visual Advertisement and Knowledge Production

• “Pictorial Turn”, W. J. T. Mitchell, 1992
• “Dian Shi Zhai” (点石斋), Chen Pingyuan (陈平原)
• “Visual Modernity”, “Chinese Women's Magazines in the Late Qing and Early Republican Period”, Barbara Mittler
• “Virtual Shanghai”, Christian Henriot
• “China Made”, Karl Gerth
• The “Modern Girl”, Tani Barlow
• “Visual Knowledge Production”, Johanna Drucker
已有的广告数据库现状 / Why we want to archive ads

• 一方面是因为中国近现代报纸的数量非常庞大，其保存地也相对比较分散，这就造成了学者在研究的时候，获取相应的资源并不是非常的方便；另一方面也是因为尽管中国及国外很多机构，比如图书馆和一些商业公司都对报纸进行了商业化，但这些数据库大部分都是要收费的，而且费用不菲，并且这些数据库的建设主要针对的是报刊上的新闻及评论文章，所以对广告的内容加工和信息提炼并不是很充分，大部分都只有广告中的一行字，但是没有对图像进行进一步的分析。

• Firstly, not all newspapers published during the Modern China period have been digitized, which means researchers still need to visit those archiving location to get access to newspaper’s physical material in the form of microfilm or photo copy.

• Secondly, even the newspapers have been digitized but not all of them could be access online.

• Thirdly, some important newspapers haven been digitized repeatedly but the accesses are limited by the high access fee.

• Except for the financial condition, for the already-digitized newspaper, advertisements are not considered as important as news so for users it is not easy to get information beside title and date, which becomes obstacle for researcher who is willing to use such information into their researches, not to mention the systemically analysis.

• Furthermore, we think archive advertisement and extract information from images to set up a database could provide new way of looking though and analyze newspaper advertisements for researchers.
CCAA Metadata

• Metadating Images
  - inventing of new systems of image description and categorization
  - inventing new interfaces to image collections
  - inventing new kinds of images and approaching the new “super-human” scale of visual data available as a creative opportunity

• CCAA Metadata
  - Descriptive data
  - Contextual information data
  - Bibliographical data
  - Technical data
**What metadata can tell us**

- Samples of statistical analysis of frequency we made based on the incomplete metadata of advertisement in Hankou Times 1907-1922.

<table>
<thead>
<tr>
<th>行标签</th>
<th>计数项：Text Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>第威德补肾丸[Dr WITT'S kidney and bladder pill]</td>
<td>7</td>
</tr>
<tr>
<td>韦廉士医生红色补丸</td>
<td>6</td>
</tr>
<tr>
<td>如意膏[SHE-KO]</td>
<td>5</td>
</tr>
<tr>
<td>第威德润肠丸[Dr WITT'S]</td>
<td>5</td>
</tr>
<tr>
<td>大炮台香烟</td>
<td>4</td>
</tr>
<tr>
<td>六零六</td>
<td>4</td>
</tr>
<tr>
<td>第威德痔疮膏[Dr WITT'S MAN ZAN FOR PILES]</td>
<td>4</td>
</tr>
<tr>
<td>五华牌香烟[MAY BLOSSOM]</td>
<td>4</td>
</tr>
<tr>
<td>哈德门香烟[HATAMEN]</td>
<td>4</td>
</tr>
<tr>
<td>别克汽车[Motor Vehicle][Buick]</td>
<td>3</td>
</tr>
</tbody>
</table>

A chart of frequency of brands
A chart of the frequency of categories of commodities

A chart of frequency of the first five companies
A graph of “Product”: the most frequent advertised brands respectively are 如意膏, 韦廉士吸入止咳片, 韦廉士医生红色补丸.
Statistical analysis of metadata of 2,554 images in Hankow Times in 1934.

The sequence of pages of newspapers advertisement published mostly is “第二张第八版,” “第二张第六版” and “第二张第五版”
The first three brands in "张" are "红", "红", and "导".

The first three brands in "张" are "红", "红", and "导".

"2014D"
The result shows that company “韦廉士” did publish on two pages, the reason is “如意膏” is also a brand of “韦廉士” but the brand name doesn’t indicate it.
it is very interest that ad of “如意膏” on “第二张第八版” was very small and the information of brand and company is not very distinct comparing to other ads on the same page and the ad of other products of “韦廉士” on “第二张第六版.”
Another interesting result is from the analysis of keywords of “subject”. I ran two keywords, “情景插图” and “女性” and the result shows that brands having those two keywords are overlapped, which indicates that image in ad that describes a scene (情景插图) not appearance of product, usually has a female figure (女性).
Text mining is another way we look into the content of advertisements. When CCAA completed, our Stage II database will offer statistical mechanisms enabling scholars to analyze advertisements for their own research projects.

A word list is not only for visualize the frequency of words used by ads but also a preparation for the next step of text mining and concept network. Chinese language is lack of word boundaries and punctuation. Different from English, the word boundaries in Chinese are invisible; even worse, many advertisement texts (such as slogans) have no punctuation marks. Thus, the raw data is just a long sequence of Chinese characters (the basic units of Chinese) without breaks. The uncertainty in the segmentation of Chinese texts makes Chinese text mining a very challenging job, compared to mining English texts.
• With text mining, we basically want to achieve two things
  ➢ discover potential associations among features and technical terms from advertisements
  ➢ build links between these features/terms and our knowledge about ideological trend and social dynamics of China in that period to achieve quantitative analysis or create new knowledge
從元數據到概念網絡 / From Metadata to Concept Network

Visual interface of Concept Network
• Thanks to all researchers of CCAA, such as Professor Tani Barlow at Rice University, Professor Ke Deng at Tsing Hua University, Mr. He Lin at Shanghai Library, Ying Jin at Fondren Library of Rice University and Mr. Minhua Guo at Easy for You Company; and all collaborators from University of Washington, Seattle, Shanghai Jiao Tong University, Nanjing University, Xiamen University, the National Library of China and Sun Yat-Sen Library of Guangdong Province.

• Thanks also to Dr. Haejin Koh at CCAS, Rice University and intern students at Rice University, University of Washington, Seattle, Shanghai Jiao Tong University, Peking University, Huadong Normal University, Nanjin University and Nanjing Normal University. All of them have contributed a lot of time and efforts into this project and only because of them, this project would be possible!