



Revisiting CUHK 60th Anniversary Oral History Project

LAM Hoi Sing, Colin (MPhil HIST/1)
HUANG Xiadong, Cory (PhD HIST/2)

Interactive Map and LDA Modelling

重訪《香港中文大學六十周年口述歷史項目》

Introduction

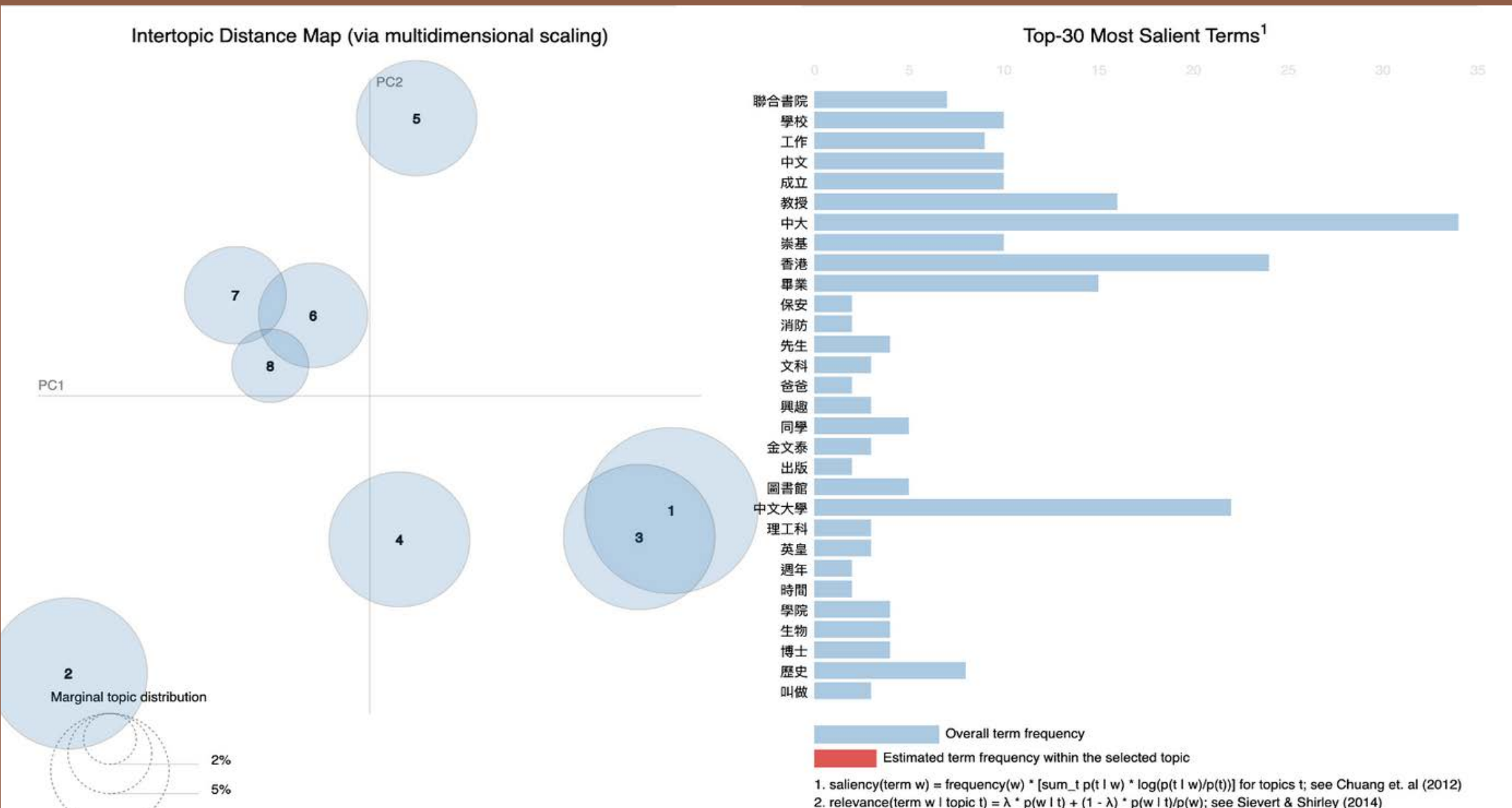
This project aims to explore how people remember places associated with CUHK between the interviewees in the **CUHK 60th Anniversary Oral History Project**. The 60 interviews in the Collection were transcribed to Cantonese texts using Whisper AI. **Latent Dirichlet allocation (LDA)** modelling has been used in analysing their speeches while the differences in memories of CUHK across generations and colleges are illustrated with an interactive map using **ArcGIS Online**.

Cantonese Transcriptions

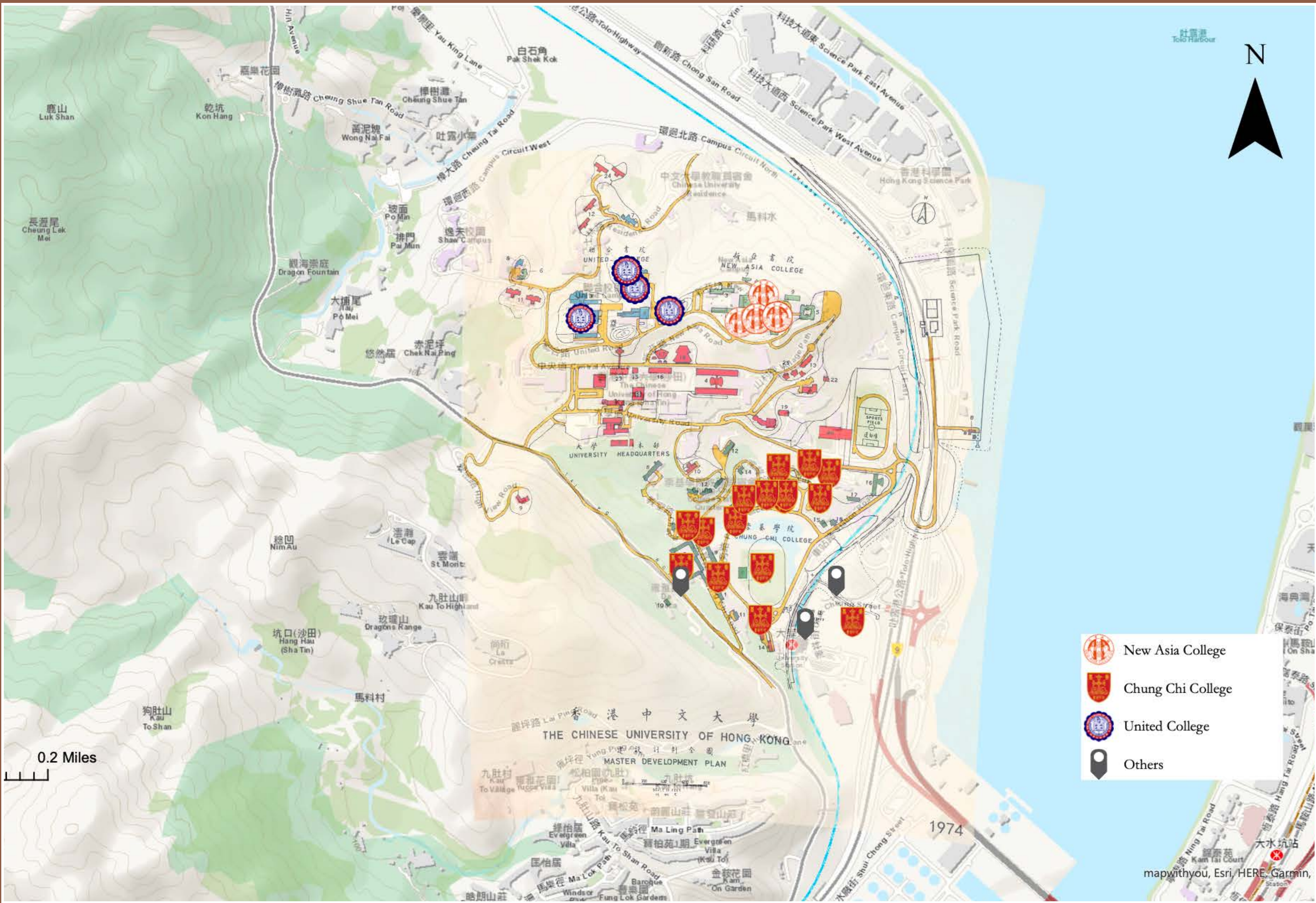
The 'alvanlii' pre-trained model in **Whisper AI**, designed for translating Cantonese speech, was used to create the initial transcriptions of the audio interview clips. Gemini from Poe was adopted in fine-tuning the transcription results with large language model (LLM).

LDA

This project used the **Jieba** library and its dictionary for segmentation. Stop word removal was conducted to eliminate Cantonese redundancy. However, the results were unsatisfactory since the interview content had no major differences.



Interactive Map



The map includes different layers for New Asia College, Chung Chi College, United College, and the Chinese University of Hong Kong. Additional layers categorize individuals by their affiliation - enrolled or worked 1) before 1963, 2) between 1963 and 1976, and 3) between 1976 and 1980.

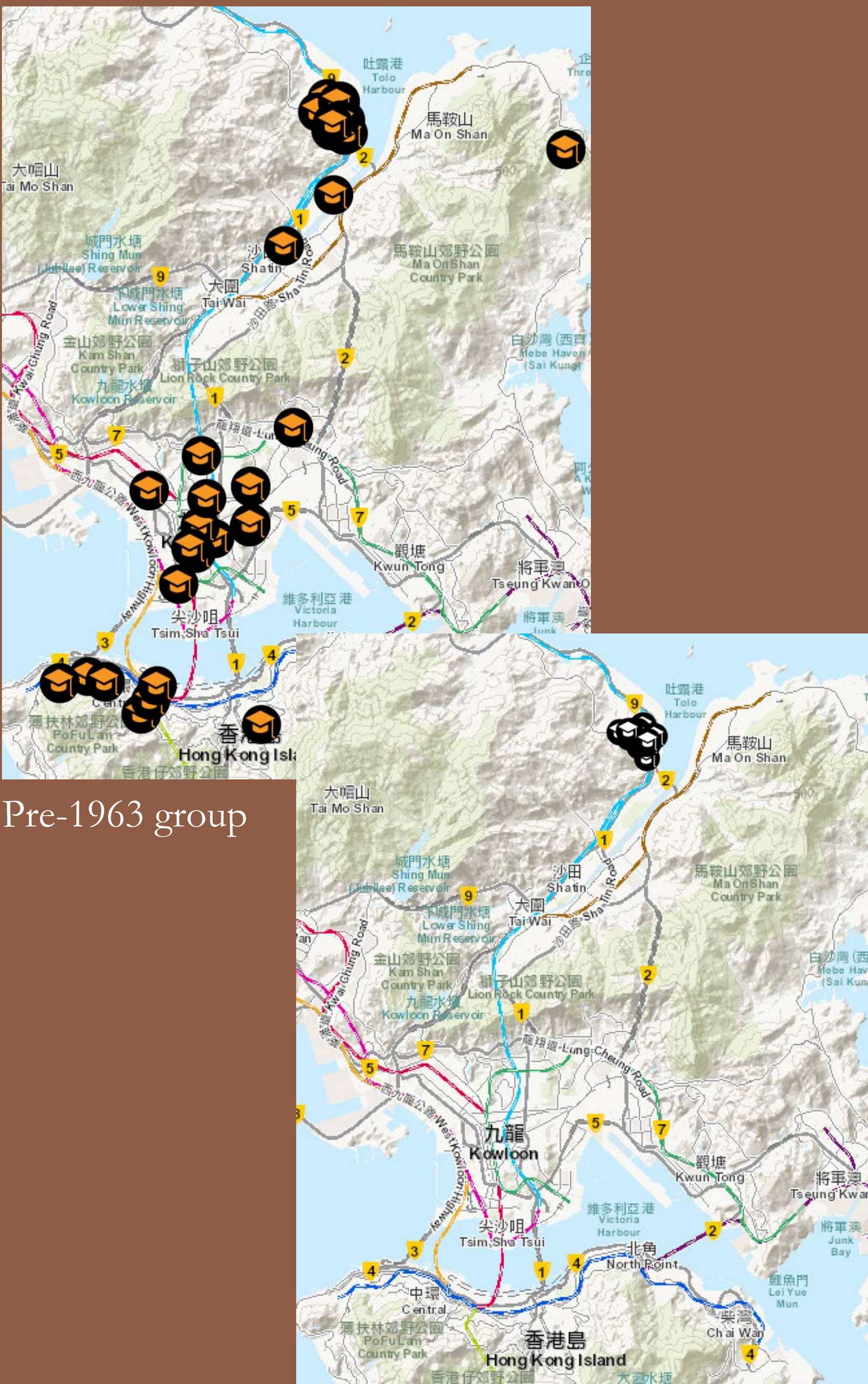
Findings

Data

82 place names and 519 quotes were generated from the 60 interviews.

The Diverse Lives of Different Generations

This project found that there are significant differences in the scope of activities between the pre-1963 group and the 1976-1980 cohort. This is likely because the three founding colleges were located in different areas of Hong Kong before 1963. Subsequently, people concentrated their activities on the CUHK campus.



More about this project

<https://dsprojects.lib.cuhk.edu.hk/en/projects/cuhk-60th-oralhistory/introduction/>
CUHK 60th Anniversary Oral History Project:
<https://repository.lib.cuhk.edu.hk/en/collection/oralhist>

If you would like to reuse the data, please contact data@cuhk.edu.hk.



Many thanks to

- Ms. Kitty Siu, CUHK Library
- Ms. Sintra Tsang, CUHK Library
- Prof. Puk Wing Kin, Department of History, CUHK
- Digital Scholarship Team, CUHK Library