Plan Overview

A Data Management Plan created using HKUL DMPTool

Title: Which Way for Hong Kong Higher Education? Exploring Ideas of Universities in Hong Kong: Cultural Routes and Future Opportunities

Creator: Yuqi Yang

Affiliation: The University of Hong Kong

Template: HKU Template

Project abstract:

Hong Kong's publicly funded universities occupy a liminal space: they remain influenced by British cultural norms and common-law charters while being increasingly embedded in a resurgent Sinitic moral-political horizon, producing a daily collision of institutional languages and loyalties. This thesis contends that these institutions are best understood as constellational hybrids in which Western ideals of autonomous inquiry and Chinese ethics of scholarly service are continually renegotiated rather than neatly synthesised. Two overarching questions guide the study: (1) how are Western and Chinese ideas of the university enacted in everyday practice across human development, knowledge production, intellect formation, social contribution and organisational form, and (2) what patterns of tension, accommodation or synthesis emerge when academics, managers and policy actors weave those ideas together.

The inquiry is grounded in a relational-constructivist ontology and an interpretivist epistemology animated by iterative abductive looping between data and theory. Twenty-four semi-structured interviews drawn from two University Grants Committee institutions – stratified across senior leadership, mid-level managers and rank-and-file academics – provide a sector-wide, cross-sectional portrait of hybrid meaning-making during the 2024/25 academic year. Reflexive thematic analysis, executed in NVivo 14 and refined until thematic saturation, secures analytic transparency and rigour.

Early analysis discloses a patterned yet contested repertoire. Lehrfreiheit claims rooted in the Humboldtian tradition inform research strategy, yet promotion dossiers simultaneously invoke

Confucian junzi ethics of civic virtue; global ranking metrics are embraced for competitive advantage but are morally legitimised through narratives of harmony and national rejuvenation; and newly mandated National-Security committees create fresh fault-lines in which tactical compliance co-exists with quiet resistance expressed through hybrid vocabularies and mundane work-arounds.

The thesis advances the field on four fronts. Theoretically, it reconceptualises hybridity as enacted, layered and historically entangled rather than a simple convergence–divergence dichotomy. Empirically, it offers the first systematic, interview-based map of Hong Kong academics' hybrid imaginaries at a moment of acute political flux. Methodologically, it demonstrates the utility of abductive looping and reflexive thematic analysis for capturing situated organisational meanings. Practically, it furnishes policy makers and university leaders with an empirically-grounded vocabulary for negotiating autonomy, civic duty and global competitiveness.

By foregrounding the moral and epistemic grammars through which institutional actors navigate competing imperatives, the study illuminates both the possibilities and the limits of the future of Hong Kong's universities, as well as other hybrid higher-education systems, whilst concurrently offering a transferable lens for other post-colonial, globally-connected university sectors wrestling with similarly layered legacies.

Start date: 06-18-2025

End date: 09-30-2027

Last modified: 05-28-2025

Copyright information:

Which Way for Hong Kong Higher Education? Exploring Ideas of Universities in Hong Kong: Cultural Routes and Future Opportunities

Data Collection

What data will you collect or create?

The project primarily involves qualitative field research data, consisting of:

- Transcripts of semi-structured interviews: These are collected in text format (preferably Word or plain text) and stored electronically. Each transcript includes the participant's responses, along with researcher annotations and contextual notes. The volume depends on the number of interviews conducted, estimated to be a moderate volume manageable within institutional storage limits.
- Field notes and memos: These are kept as text documents linked to transcripts within qualitative analysis software.
- Policy texts and related documents: Supplementary materials gathered from public sources or institutional repositories, formatted as PDFs or Word documents.
- Researcher analysis data: Coding schemes, memos, and thematic annotations stored within qualitative analysis software.

Software and analysis tools: NVivo (version 14 or later) is required to read, organize, and analyze the qualitative data.

Data reuse and sharing: The text formats used (Word, PDF, NVivo project files) support sharing and ensure long-term access, especially when stored using open or widely adopted standards and repositories.

Existing data reuse: Existing publicly available policy texts and academic literature will be incorporated, which are in the public domain and can be reused without restrictions.

Implications: Using standard textual formats facilitates compatibility across analysis platforms and long-term preservation. Data volumes are expected to be manageable, with regular backups stored securely on university servers. Restricted access controls and encryption ensure data security, especially given the politically sensitive nature of interview content.

How will the data be collected or created?

The data collection and creation process will adhere to established qualitative research standards, emphasizing transparency, consistency, and rigor. The key elements are outlined below:

Standards and Methodologies

- Qualitative Data Standards: The project follows best practices in qualitative research, aligning with community standards such as those recommended by the Digital Curation Centre (DCC) and the Qualitative Data Repository (QDR). These include systematic documentation, participant validation, and anonymization procedures.
- Data Collection Methodology: Data will be gathered through semi-structured interviews, which are recorded, transcribed verbatim, and supplemented with field notes. The process emphasizes standardized interview protocols, piloting, and refining interview guides to ensure consistency across participants.

Data Organization During the Project

- **Folder and File Structure:**
  - Root Folder: Project_Name/
  - Data/ — contains raw and processed data
  - Audio/ — original interview recordings (.mp3)
  - Transcripts/ — verbatim transcripts (.docx)
  - Revised_Transcripts/ — corrected or participant-validated transcripts
  - Field_Notes/ — contextual notes and observations
  - Analysis/ — coding schemes, memos, themes
  - Documentation/ — consent forms, interview guides, protocols, audit logs
  - Backups/ — encrypted backups kept separately for data security
- **Naming Conventions:**
  - Participants: [Institution][Role]_[Number], e.g., HKU_Senior_001
  - Audio files: [ParticipantID]_Audio_[Date].mp3
  - Transcripts: [ParticipantID]_Transcript_[Date].docx
  - Revision files: Append _rev with date or version number, e.g., HKU_Senior_001_Transcript_20240215_rev.docx
  - Memos/analysis: [AnalysisTopic]_[Date].nvivo or .docx

Version Control

- **Handling Versions:**
  - Each major document (transcripts, memos, analysis files) will include version numbers or dates in filenames to track updates.
  - Record changes in a Change_Log.txt or in the file metadata, detailing what was updated and why.
  - Use NVivo's internal project versioning and regular backups to maintain a history of analytical stages.

Quality Assurance Processes

- **Consistency and Standardization:**
  - Interview guides will be piloted and iteratively refined to ensure questions align with research aims.
  - All interviews will be conducted using a standardized protocol, with interviewers trained to ensure consistency.
  - Audio recordings will be of high quality, verified before transcription.
  - Transcripts will be subject to participant validation within ten days—participants will review and correct transcripts to ensure accuracy.
- **Data Recording & Validation:**
  - Transcriptions will be double-checked for accuracy, with bilingual translations verified by professional translators.
  - Field notes capturing tone, non-verbal cues, and context will complement transcripts.
  - Coding will follow a predefined manual or coding scheme, with iterative refinement documented through memos.
  - Discrepancies during coding (e.g., variations between coder and peer) will be logged and adjudicated to ensure interpretive consistency.
- **Documentation and Audit Trails:**
  - All analytical decisions, coding schemes, and procedural steps will be documented in audit logs.
  - Each step of data collection and processing will be recorded systematically to facilitate reproducibility and transparency.
- **Community Data Standards and Metadata**
  - Data will be anonymized in accordance with community standards, replacing identifiable information with pseudonyms or generic descriptors.
  - Metadata for each data point (participant ID, date, role, institution, language, etc.) will be systematically recorded.
  - Data will be stored in open, interoperable formats (.docx, .mp3, NVivo project files), following standards such as those recommended by ICPSR for qualitative data sharing.

In summary, the thesis employs rigorous qualitative standards, with organized folder structures, clear naming conventions, thorough version control, and systematic quality assurance measures, including participant validation, double coding, and detailed documentation. This ensures the integrity, reproducibility, and ethical handling of the data throughout the research process.

Documentation and Metadata

What documentation and metadata will accompany the data?

The documentation accompanying the data will be designed to ensure clarity, transparency, and usability for secondary users. It will include the following types of documentation:

1. ReadMe File

Content: Will provide an overview of the dataset, including the dataset's title, purpose, and scope. It will record the name and contact information of the creator(s), the date of data creation, and version details. It will also specify the formats of data files (preferably open formats like .txt, .csv, .xml, .json, .docx), conditions for access including any restrictions, licensing information, and citation instructions.
Method of capture: Created concurrently with data collection, stored in a dedicated folder, and kept updated.

2. Codebook / Data Dictionary

Content: Will define all variables, codes, and themes used in transcripts and analysis. It will explain coding schemes, thematic categories, and any abbreviations or terminology used. It will specify units of measurement and clarify assumptions made during transcription, translation, or coding.
Method of capture: Developed during and after coding sessions, possibly embedded within analytical software (e.g., NVivo), and exported as open format documents.

3. Methodology and Procedural Documentation

Content: Will detail data collection procedures, interview protocols, translation processes, validation steps, and coding strategies. It will include any changes made during the research process.
Method of capture: Drafted during research, stored as a separate document, and linked to the dataset in the ReadMe.

4. Metadata Records

Content: Will describe the dataset’s title, creator/contributor details, date of data collection, description, file formats, access conditions, methodological details, and any relevant vocabularies or code standards.
Standards: Will employ community standards such as Dublin Core or Digital Data Documentation

Initiative (DDI) to ensure interoperability and discoverability.
Method of capture: Metadata will be entered into a standardized metadata template, saved as separate structured files (e.g., JSON or XML), and linked with the dataset.

5. Storage and Recording: All documentation will be stored within an organized folder structure in the project directory, with filenames indicating content and version. The documentation will be updated regularly to reflect methodological changes, ensuring that all information remains current and comprehensive.

6. Rationale for Standards: Using established standards such as Dublin Core and DDI allows for consistent, interoperable metadata that supports long-term accessibility and reuse. Open formats are preferred to maximize accessibility and minimize barriers for future users.

Ethics and Legal Compliance

How will you manage any ethical issues?

1. Gaining Consent for Data Preservation and Sharing

The researcher will obtain informed consent from all participants through bilingual Participant Information Sheets and Consent Forms detailing the study's aims, data handling procedures, and potential future sharing or reuse of data.
Participants will be explicitly informed that they can decline to answer questions, withdraw their data up to two weeks after receiving their verified transcripts, and specify their preferences for anonymization.
Transcripts will be verified with participants within ten days to ensure accuracy and consent regarding data use.

2. Protecting Participant Identity through Anonymization

To safeguard identities, the research will substitute personal identifiers with generic descriptors (e.g., “a research-intensive institution”).
For sensitive remarks, quotations will be anonymized, and the secondary anonymization check by supervisory team will ensure no inadvertent identifiers remain before dissemination.
Participants will have the option to remain anonymous, and this preference will be explicitly recorded on the Consent Form.

3. Secure Storage and Transfer of Sensitive Data

Data—including audio recordings and transcripts—will be stored on encrypted (AES-256) university servers, with access restricted via password-protection and two-factor authentication.

Raw data will not be uploaded to commercial cloud services, and identifiers will be replaced by alphanumeric codes with the key stored securely on encrypted USB devices kept in locked cabinets.
Data retention will be in compliance with GDPR and Hong Kong's Personal Data (Privacy) Ordinance, with data retained for five years before secure destruction.
During collection, interviews may be conducted in neutral settings or via encrypted platforms (e.g., Zoom with data routing disabled), ensuring participant safety and confidentiality, especially considering political sensitivities.

4. Ethical Oversight and Approval

The researcher will seek ongoing approval from the relevant Institutional Review Board (e.g., the Faculty Research Ethics Sub-Committee or IRB at the University of Hong Kong).
The protocols, including consent procedures, anonymization processes, and data security measures, are designed to meet ethical standards and legal requirements.

5. Additional Safeguards

Clear procedures for data access, transfer, and destruction are in place to prevent unauthorized use or breaches.
Transparent communication with participants about their rights and data handling ensures respect for autonomy and confidentiality.

My research plan exhibits awareness of ethical issues by securing informed consent, implementing anonymization procedures, and ensuring secure storage and transfer of data. These measures align with ethical guidelines and institutional requirements to protect participants and facilitate responsible data sharing and reuse.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

The researcher will retain ownership of the data and will license it under appropriate open licenses for reuse, respecting any restrictions on third-party data. Any restrictions needed—such as delays for publication, confidentiality, or patent considerations—will be clearly documented and adhered to, ensuring compliance with institutional policies, funder requirements, and legal frameworks related to copyright and IPR.

Ownership of Data:

The primary researcher will hold the copyright and IPR for the data generated through interviews, transcripts, coding, and analysis, unless institutional policies specify otherwise.
Given that the data originate from participants' narratives, participants will have had their informed consent regarding how their data may be used, shared, or kept confidential.

The researcher will ensure compliance with any institutional, departmental, or funder policies concerning IPR, which may involve institutional ownership or restrictions.

Licensing for Reuse:

The data will be licensed under an open but controlled license, such as Creative Commons Attribution (CC BY), allowing re-users to share and adapt the data while providing appropriate credit.
If restrictions are necessary—for example, to protect sensitive or identifiable information—more restrictive licenses (e.g., CC BY-NC or CC BY-NC-ND) may be used.
The licensing terms will be clearly documented in the accompanying metadata and documentation to specify permissible uses.

Restrictions on Third-Party Data:

Any third-party data integrated into the research (e.g., external documents, policies, or previous publications) will be properly cited and permissions obtained where necessary.
Reuse restrictions, such as embargoes or limited access, will be respected according to the licenses or agreements governing those third-party materials.
The researcher will ensure that sharing or republishing do not infringe on third-party copyrights or confidentiality.

Data Sharing and Post-Research Restrictions:

Data sharing will be planned in alignment with ethical approvals and participant agreements.
Sensitive data or confidential information (e.g., politically charged remarks) will be anonymized and potentially restricted from open access if necessary.
Data sharing may be postponed or restricted—for example, to allow publication or patent processes —by applying embargo periods or access controls.
The researcher will specify in the data management plan when and under what conditions data will be shared, including any restrictions for publication or intellectual property considerations.

Storage and Backup

How will the data be stored and backed up during the research? i. e. until stored in the final location (e.g. on your password protected laptop)?

Initial Data Storage:

Raw data, including audio recordings and transcripts, will be stored securely on the researcher’s password-protected university devices, such as a laptop or desktop computer, in accordance with

institutional guidelines.
During collection, data will be stored temporarily in encrypted, secure locations, with sensitive notes kept in password-protected NVivo project folders, ensuring only authorized access.

Backup Strategy:

Regular backups will be made to prevent data loss. The primary backup will occur daily or after each significant data collection session.
Backup copies will be created on encrypted, university-approved external storage devices (such as encrypted USB drives or external hard drives).
Additionally, a secondary backup will be stored on secure, institutionally approved cloud services that comply with data protection policies—preferably within Hong Kong or jurisdictions with equivalent protections—to facilitate recovery in case of hardware failure.

Responsibility and Data Recovery:

The researcher (or designated data manager, if applicable) will be responsible for performing backups and ensuring their proper management.
In the event of data loss, recovery procedures will involve restoring data from the latest secure backup stored on protected local or cloud locations.
Regular checks will be performed to confirm backup integrity and accessibility.

Frequency of Backups:

Data will be backed up at least once daily during active data collection phases, with increments after each data entry or significant event.
Final versions, after verification and anonymization, will be stored in multiple secure locations, including institutional repositories, to ensure availability and redundancy.

How will you manage access and security?

Risks to Data Security:

Potential risks include unauthorized access, data loss, or leakage during collection, storage, transfer, and analysis phases.
To mitigate these risks, all digital files—such as audio recordings, transcripts, and coding data—will be stored on encrypted devices and folders, for example, encrypted USB drives, password-protected computers, and secure NVivo projects.
Sensitive information will be segregated in separate, restricted folders to prevent accidental access or disclosure.

Control of Access:

Access to data will be restricted via password protections, with strong, unique passwords for all devices and software used.
Two-factor authentication (2FA) will be employed for access to cloud storage and databases.
Participants' identifiable data will be anonymized or pseudonymized, with a separate, encrypted key file stored securely offline, controlling who can link identities back to anonymized data.
Access rights will be limited to the researcher and/or designated supervisory personnel, with strict controls over permissions.

Secure Data Transfer:

During field collection, data will be transferred securely from devices to the main storage, utilizing encrypted transfer methods, such as encrypted file transfer protocols or secure, institution-approved cloud platforms.
When transferring data from field devices to the primary storage location, all data will be encrypted to prevent interception.
If data is collected in the field via portable devices, those devices will be secured physically and encrypted, and files will be transferred promptly to secure systems with encryption.

Secure Collaboration:

If collaborating with colleagues, access to sensitive data will be managed through secure, institutional collaboration platforms that enforce data encryption and access controls.
External collaborators will be granted access only to de-identified data or specific datasets necessary for their role, with permissions closely monitored.

Security Standards and Compliance:

All data handling procedures will comply with relevant standards such as the EU's GDPR (where applicable) and Hong Kong's Personal Data (Privacy) Ordinance.
Encryption standards (AES-256) will be used for data at rest and in transit.
Regular audits and security checks will be performed to ensure that data protection measures are functioning effectively.
Physical documents or sensitive notes will be stored in locked cabinets with limited access.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Data to be Retained and Their Rationale:

Audio Recordings and Transcripts:

Reason: These form the primary data sources, essential for verifying coding and interpretations.
Duration: Retained for at least five years post-project completion, aligning with data protection policies and institutional guidelines.
Format: Stored securely in encrypted formats, with transcripts anonymized to protect participant identities.
Annotated Notes and Memos in NVivo:
Reason: These support ongoing analysis, coding decisions, and methodological transparency.
Duration: Same retention period as raw data, with potential for reuse in publications or methodological pedagogies.
Participant Consent and Data Management Documentation:
Reason: Necessary for contextual understanding and compliance with ethical requirements.
Duration: Retained alongside data for audit purposes.

Legal, Ethical, and Contractual Considerations:

Data will be retained in accordance with institutional policies, data protection laws (e.g., GDPR where applicable, Hong Kong's Personal Data (Privacy) Ordinance), and funder requirements.
Any personal or sensitive data not directly necessary for future use will be securely destroyed after the retention period, unless participants have explicitly consented to longer retention.

Research Use and Reuse Potential:

The main data (audio and transcripts) have high potential for reuse to validate findings, support secondary analyses, or develop new studies related to ethics, governance, or institutional identities in Hong Kong's higher education context.
Anonymized transcripts could be shared with other researchers or included in repositories for educational or methodological examples, provided that confidentiality is maintained and participants' anonymity is preserved.

Retention Duration and Preservation:

Data will be retained for at least five years after the project concludes, consistent with institutional policies and best practices for qualitative research.
The data will be stored in formats conducive to long-term accessibility (e.g., encrypted PDFs, standard audio formats).
Preparing data for sharing involves documenting metadata, anonymizing, and possibly converting files to open formats, which entails additional effort. Therefore, only data with demonstrable long-term value and reuse potential will be preserved for the extended period.

What is the long-term preservation plan for the dataset?

Preservation Location:

Primary Repository: The dataset—including anonymized transcripts, audio recordings, analysis notes, and related documentation—will be stored in the HKU DataHub, which is an institutional repository that supports secure long-term data preservation and sharing.

Alternative or supplementary archiving: If necessary, supplementary datasets or sensitive data may be stored in a secure, encrypted institutional server with controlled access, but the primary long-term storage will be in the HKU DataHub.

URL: https://data.hku.hk (or the specific link provided by HKU DataHub upon deposit).

Cost Considerations:

Repository Charges: The HKU DataHub generally does not charge individual researchers for data storage or deposit. However, if external repositories are used, review of the specific's repository fee schedule will be necessary; currently, no additional costs are anticipated for the planned deposit.

Preparation and Documentation Efforts:

The dataset will be systematically prepared for sharing, including:

Anonymization: Ensuring all personal identifiers and sensitive information are removed or pseudonymized.

Metadata Creation: Detailed data documentation will be developed, including file descriptions, variable definitions, codebooks, contextual information, and methodological notes to facilitate reuse.

File Formatting: Data will be converted into open, sustainable formats (e.g., PDF for transcripts, WAV for audio, CSV or TXT for notes) to enhance longevity and accessibility.

Effort Budget: Adequate time will be allocated (estimated at approximately 20–30 hours) for processing data, creating metadata, and ensuring compliance with data standards and repository requirements.

Long-term Preservation and Curation Strategies:

Curatorial Policies: The HKU DataHub provides ongoing data curation, backup, and version control.

Ongoing Management: A designated data steward or researcher will maintain the dataset, monitor access requests, and update documentation if necessary.

Data Use Policy: Clear licensing (e.g., CC BY or similar license) will be applied to facilitate appropriate reuse while respecting participant confidentiality.

Future Accessibility: The data will be stored securely with persistent identifiers (e.g., DOI) assigned to support discoverability and citation.

Data Sharing

How will you share the data?

Mechanisms for Data Sharing:

Primary access point: Data will be shared primarily via the HKU DataHub, an institutional repository providing controlled access, persistent identifiers, and discoverability.

Complementary mechanisms: For genuine case-by-case requests—such as for sensitive or detailed transcripts—access may be granted upon formal request, subject to ethical review and participant consent restrictions. Access requests will be handled through a formal data access procedure documented in the data repository's policies.

Data Files to be Shared:

Anonymized transcripts of interviews (text data).

Audio recordings (if ethically permissible and with appropriate de-identification).

Metadata and codebooks describing data collection procedures and variable definitions.

Analysis notes and memos to facilitate understanding of the coding and interpretation process.

Timing of Data Availability:

Embargo/Restriction: Data will be made publicly available after the completion of the project and upon publication of main findings, anticipated approximately 6–12 months post-project.

Restrictions: Sensitive data containing potentially identifiable information will be subject to a restricted access period of 2 years or until ethical approval allows broader sharing, whichever is sooner. The public version of the dataset will exclude any sensitive or personally identifiable information.

Access Conditions and Licensing:

Conditions: Data will be shared under a Creative Commons Attribution Non-Commercial (CC BY-NC) license, allowing reuse for non-commercial purposes with appropriate attribution.

Restrictions: Direct access will be granted only to bona fide researchers who agree to the license terms and abide by confidentiality and data protection guidelines.

Persistent Identifier:

The dataset will be assigned a persistent identifier (such as a DOI) upon deposit in the HKU DataHub.

Existing DOI: If a DOI has already been obtained (e.g., during deposit), it will be provided here; otherwise, DOI registration will be completed at the time of data publication.

Long-term Accessibility and Acknowledgment:

Data will be curated to ensure it remains accessible beyond the project's duration, facilitating future research and teaching.

Proper acknowledgment guidelines will be included in the metadata and any data reuse licenses, encouraging users to cite the dataset in publications, e.g., referencing the DOI, author, and publication year.

Track Record and Best Practices:
Prior research projects employing similar data sharing strategies have effectively facilitated secondary analysis, replication, and scholarly acknowledgment, exemplifying the value of transparent, open data practices.

Are any restrictions on data sharing? If yes, Why?

Nature of Restrictions:
Confidentiality and Anonymity: Some interview transcripts and recordings may contain sensitive information that could identify participants or disclose politically sensitive opinions, especially given the impact of the National Security Law and the political climate.
Lack of Explicit Consent for Broad Sharing: Participants may not have consented to open sharing of raw data, especially audio recordings or detailed transcripts that could be potentially identifiable.
Legal and Institutional Policies: Data must comply with data protection regulations like Hong Kong's Personal Data (Privacy) Ordinance and GDPR, which impose restrictions on data dissemination and secondary use.
Intellectual Property Rights (IPR): Some data, particularly if derived from institutional documents or proprietary content, could be subject to IPR restrictions.

Actions to Mitigate Restrictions:
Data Anonymization: Rigorous anonymization and de-identification procedures will be applied to data files to prevent identification of participants and mitigate politicized or sensitive content.
Restrict Access: Sensitive datasets will be placed under controlled access conditions, requiring approval for access via a data access committee or formal request process.
Informed Consent and Clear Communication: Participants will be informed about data sharing plans, and consent forms will explicitly include provisions about limited sharing, ensuring ethical compliance.
Time Restrictions: Data containing sensitive content will be shared after a 3–5 year embargo or once sensitivity diminishes, allowing for an initial exclusive use period necessary for primary analysis.

Need for Data Sharing Agreements:
Yes, a data sharing agreement or data access request process will be required for access to sensitive or restricted datasets. This agreement will specify the scope of use, confidentiality obligations, and data destruction or return procedures after use.
Purpose: To ensure legal and ethical conduct, protect participant confidentiality, and prevent misuse.

Expected Difficulties and Measures:
Difficulty: Ensuring full anonymization without losing contextual richness.
Mitigation: Employ advanced anonymization techniques and provide detailed metadata to contextualize data without revealing identities.

Difficulty: Obtaining sufficient participant consent for broad sharing.
Mitigation: Clearly communicate sharing plans during consent and offer participants the option to restrict certain content.
Difficulty: Political sensitivity limiting open dissemination.
Mitigation: Share only processed, anonymized, and non-sensitive summaries in open-access repositories, reserving sensitive raw data for restricted access with proper agreements.
Difficulty: Institutional policies or intellectual property restrictions.
Mitigation: Engage early with institutional legal teams and rights holders to clarify usage rights and negotiate licensing terms.

Responsibilities and Resources

Who will be responsible for data management?

The primary responsibility for implementing and overseeing the Data Management Plan (DMP) lies with the Principal Investigator (PI), who will coordinate all data management activities throughout the research project.

Specific roles include:

Data Capture: The researcher will conduct interviews, record data securely, and ensure compliance with informed consent procedures.
Metadata Production and Data Quality: The researcher will produce detailed metadata for all datasets, ensuring data accuracy, consistency, and completeness.
Data Storage and Backup: The researcher will manage the secure storage of all data on encrypted, password-protected institutional servers, establishing routine backup procedures to prevent data loss.
Data Archiving: Post-project, the researcher or designated institutional data steward will archive the data securely for the required retention period, adhering to institutional policies and legal requirements.
Data Sharing and Dissemination: The researcher will oversee responsible data sharing, ensuring anonymization, compliance with participant consent, and adherence to relevant policies and legal restrictions.
Policy Compliance and Oversight: The Principal Investigator will ensure that all data management activities conform to institutional policies, ethical standards, and legal regulations, consulting the Institutional Ethics Committee when necessary.

What resources will you require to deliver your plan?

1. Hardware Resources

Secure Data Storage Devices:

Access to institutional encrypted servers for storing sensitive interview transcripts, audio recordings, and metadata.

Backup drives (external or network-based) for redundancy and disaster recovery.

Computing Equipment:

A reliable computer with sufficient processing power and storage capacity for audio transcription, data coding, and qualitative analysis.

2. Software Resources

Qualitative Data Analysis Software:

NVivo (latest version): Already mentioned as part of the research workflow for coding and thematic analysis.

Justification: Critical for organizing, coding, and analyzing qualitative interview data systematically.

Transcription Software (Optional):

Software such as Express Scribe, Otter.ai, or Temi could expedite transcription but is optional if transcription is done manually.

Secure Communication Tools:

Encrypted video conferencing platforms (e.g., Zoom with end-to-end encryption, Microsoft Teams) for remote interviews.

Data anonymization tools:

Software or scripts (if needed) to support anonymization of sensitive data before sharing or archiving.

3. Specialist Expertise / Training

Data Security and Ethical Compliance Training:

Training sessions for the researcher on data security protocols, handling sensitive data, and ensuring compliance with institutional and legal requirements.

Qualitative Data Analysis Training:

Additional training on using NVivo effectively, if necessary, to ensure proper use of coding and memo functions.

Language Translation:

Professional translation support for bilingual transcripts to maintain semantic fidelity (already included in plan).

Legal and Ethical Guidance:

Consultation with institutional data protection officers or legal advisors for compliance with privacy legislation, especially considering Hong Kong’s legislation.

4. Charges and External Services

Translation/Transcription Services:

If external services are employed for transcription or translation, costs should be justified.

5. Additional Considerations

Technical Support:

IT support for setting up secure data storage, troubleshooting encryption, and maintaining data security.

Ongoing Maintenance:

Resources allocated for maintaining backups, performing data integrity checks, and updating software.