SMOEx Interview Analysis

Key Findings and Implications

Data-Driven Insights for Student Medical Organization Exchange Camp Planning

Analysis by: **นศพ.พุทธรักษ์ โสดถานา**

Chiang Mai University

Introduction & Scope

Analysis Goal

Discover common themes in qualitative feedback from 28 Student Medical Organisations (SMOs) across multiple institutions to inform exchange camp planning and identify best practices.

Methodology

- Modern Natural Language Processing (NLP) techniques
- Multilingual text embedding and semantic analysis
- Density-based clustering for natural grouping discovery
- Automated topic modeling with interpretable labels

Analysis Scope

- **Initial:** SWOT section only (structural insights)
- **Updated:** All narrative fields (comprehensive view)

⚠ **Important Note**

The choice of data columns significantly impacts the depth and breadth of analysis results.

SWOT only: Reveals structural and internal management issues (teaching, communication, handover).

Full narrative: Captures broader priorities including budget, student engagement, and community needs.

Initial Analysis: Methodology (SWOT Only)

1 **Pre-processing**

Text fields are combined into a single document per institution. The multilingual e5 transformer converts each document into a dense vector, capturing semantics across both Thai and English languages.

↓

2 **Dimensionality Reduction**

UMAP (Uniform Manifold Approximation and Projection) projects high-dimensional embeddings down to five dimensions, preserving local structure essential for accurate clustering.

↓

3 **Clustering**

HDBSCAN (Hierarchical Density-Based Clustering) discovers clusters of varying shapes and automatically identifies outliers, avoiding forced clustering of dissimilar documents.

↓

4 **Topic Modeling**

BERTopic leverages HDBSCAN clusters and generates bag-of-words representations using class-based TF-IDF, extracting representative keywords and assigning human-readable topic names.

Outliers and unassigned documents are manually reviewed for unique insights.

Five Key Topics: Workflow, Governance & Communications

TOPIC 0

BM, BUU, CICM, MFU, RSU

Clinic Workflow & Coordination

Pre-clinical/clinical stage tensions, split campuses, language barriers

RECOMMENDED ACTIONS

- Appoint liaisons between cohorts
- Hold regular retrospectives
- Share documentation standards

TOPIC 1

CMU, PSU, SU, WU

Governance & Evaluation

Weak communication, slow approvals, missing post-event evaluations

RECOMMENDED ACTIONS

- Define roles explicitly
- Streamline communication channels
- Create after-action templates

TOPIC 2

MSU, NU, SUT, TU

Communications & Participation

Uneven reach, timetable clashes, vague handover procedures

RECOMMENDED ACTIONS

- Publish central event calendar

Five Key Topics: Resources & Bureaucracy

TOPIC 3

KKU, PNU, PSMD

Resources & Engagement

Low participation, limited budgets/space, travel difficulties

RECOMMENDED ACTIONS
- Develop sponsor packages
- Offer hybrid participation
- Emphasize personal benefits
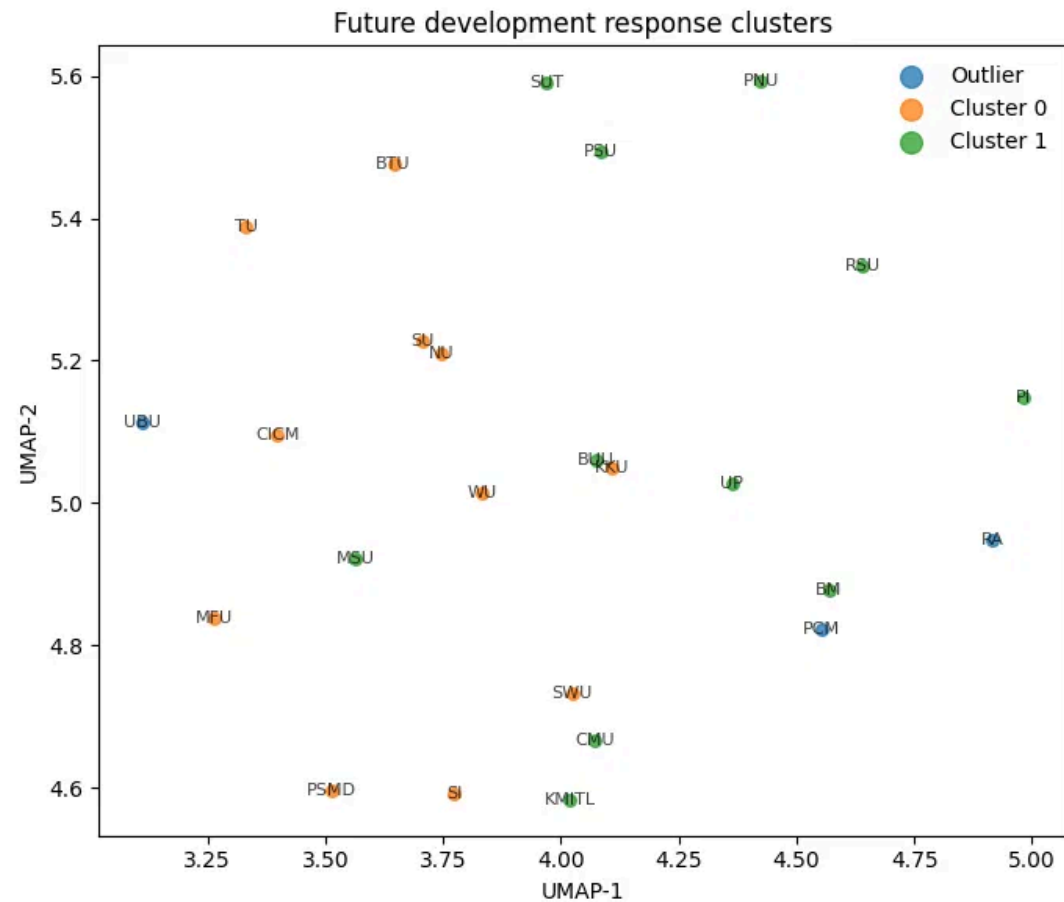
TOPIC 4

PCM, RA, SI

Bureaucracy & Handover

Complex approval chains, weak continuity during leadership changes

RECOMMENDED ACTIONS
- Adopt standardized request forms
- Map approval flows clearly
- Implement handover checklists

Future Development Response Clusters

Institutional clusters based on UMAP-reduced future development responses

What the Clusters Represent

Points close together represent clubs with similar future directions (e.g., academic support, community outreach, fundraising, internal restructuring).

Practical Application

- Form working groups based on cluster membership
- Group titles: "Engagement & Mental Health", "Funding & Sustainability", "Academic Innovation", "External Partnerships"
- Prioritize clusters with many institutions but weak current practice (high leverage)
- Identify cross-region clusters for national-level templates

Combine with Mentor Pairing

Strategy: Pick mentors from clubs whose current strengths match the dominant needs inside each "future plan" cluster.

Result: Targeted mentoring aligned with institutional future directions and strategic priorities.

Student Number Cluster Analysis

Methodology

Unsupervised clustering on total student numbers at each institution using DBSCAN, K-means, and agglomerative clustering algorithms with normalized data.

Key Finding

K-means and agglomerative clustering show highest silhouette score at three clusters, revealing natural institutional size groupings optimal for peer-to-peer matching.

Cluster Category	Student Range	Institutions
Very Large	1,500–1,868 students	CMU, CU, SI
Large	840–1,230 students	KKU, NU, PSU, RA, RSU, SWU, TU
Small	Fewer than ~600 students	BM, BUU, CICM, CMU, MFU, MSU, PCM, PI, PNU, PSMD, SI, SU, SUT, TU, WU, BTU

Four-Cluster Solution

A distinct mid-sized group (BM, PCM, PI, SUT) emerges between very large and large clusters, providing more granular peer-matching options.

Exchange Camp Application

Pair institutions with complementary strengths and needs based on cluster membership to maximize knowledge transfer and peer learning.

Updated Analysis: Full Narrative Scope

INITIAL APPROACH

SWOT Section Only

Focused on structural and internal management issues such as teaching administration, communication systems, and handover procedures. Limited to predefined SWOT framework.

UPDATED APPROACH

All Narrative Fields

Encompasses all qualitative data including Concern, Sharing, Expectation, and SWOT. Captures broader institutional priorities beyond structural insights.

Updated Analysis Dataset & Results

Documents Analyzed: 28 institutions (one document per institution)

Fields Used: All narrative columns including "โครงสร้างสโมสร", "สโมสร Concern", "คาดหวังอะไรจากโครงการค่ายแลกเปลี่ยนสโมสร"

Topics Detected: 4 distinct topics (labelled 0–3)

Distribution: Topic 0 (8 institutions), Topic 1 (7), Topic 2 (7), Topic 3 (6). **No documents left unclustered.**

Four Key Topics: Programme Design & Engagement

TOPIC 0

Programme Design & Event Consolidation

Institutions report overlapping activities and seek streamlined events with meaningful scenario-based sharing instead of passive presentations.

KEY THEMES

- Consolidate similar events to reduce resource strain
- Use scenario-driven sharing rather than formal presentations
- Align session timing with committee turnover
- Provide clear objectives and pre-work for participants

TOPIC 1

Student Engagement, Well-being & Communication

Low student participation and limited resources are compounded by weak mental-health support and inadequate PR channels for club visibility.

KEY THEMES

- Increase student participation and recruit first-year members
- Establish systematic mental-health support and monitoring
- Strengthen public relations and social media presence
- Share detailed job roles across institutions

Four Key Topics: Budget & Operations

Budget Flexibility, Resilience & Balanced Programming

Respondents focus on financial management challenges. They seek more flexible budgets, improved cash-flow collaboration with administrative staff, and a shift from generic resilience lectures to practical capacity-building workshops. There is also emphasis on balancing intensive discussions with lighter activities and enhancing feedback systems for committee continuity.

KEY THEMES

- Improve budget fluidity and cash-flow approvals
- Shift from generic resilience lectures to practical skill-building
- Balance intensive discussions with recreational bonding
- Enhance feedback pipelines to aid handovers

Operational Logistics & Staffing

This cluster emphasizes administrative pain points. Remote campuses create attendance barriers, document submissions require longer lead times, and staff shortages limit project management capacity. Respondents request online meeting options, better sharing of working cultures between teachers and students, and robust recruitment systems to maintain adequate committee membership.

KEY THEMES

- Facilitate online meetings to overcome travel distance
- Provide explicit timelines for document submissions
- Share working cultures between teachers and students
- Implement robust recruitment systems for committee staffing

Mentor-Learner Pairing Strategy

Mentor→Learner Complementarity (Strengths×Weaknesses)
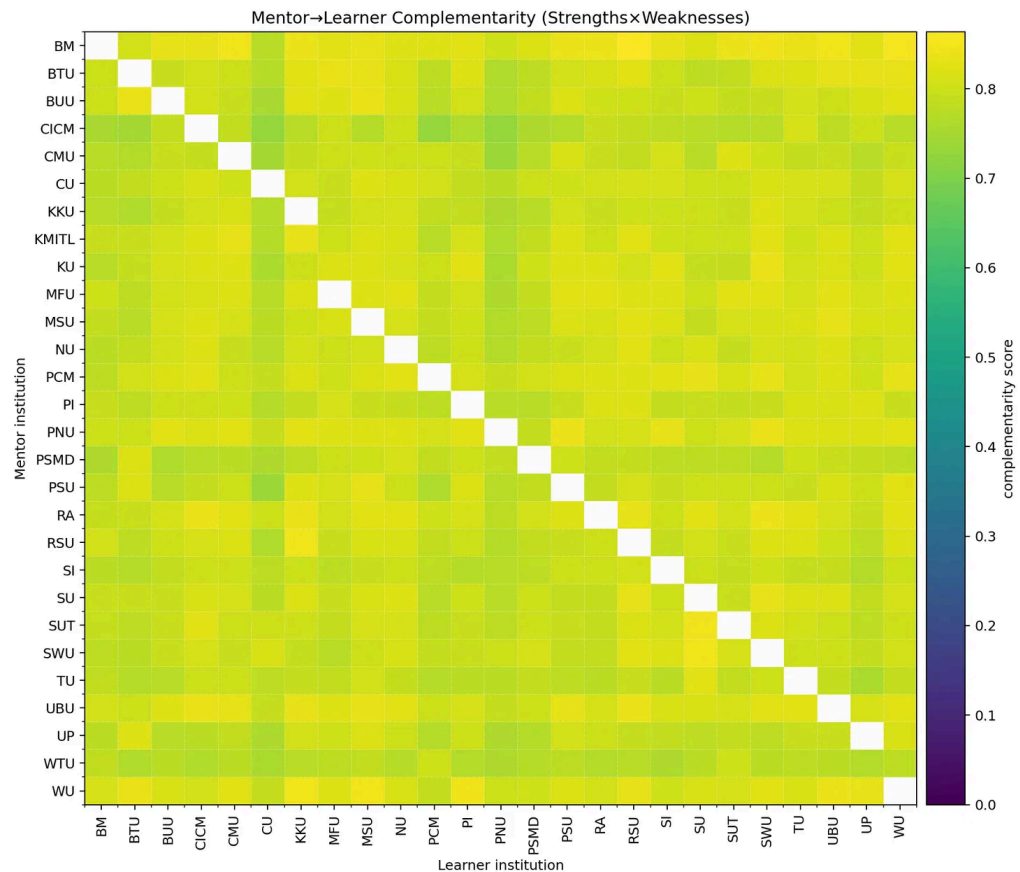
Complementarity Matrix: Rows = Mentors | Columns = Learners | Color Intensity = Strength-Need Alignment

How to Read the Heatmap

Each row represents a potential mentor institution; each column represents a potential learner institution. Darker/warmer cells indicate higher complementarity.

Color Scale:

High (0.8–1.0)   Medium (0.4–0.6)   Low (0.0–0.2)

Finding Best Mentors

For any given learner (column), the best candidate mentors are the institutions with the darkest cells. Enforce constraints (region, school size) by selecting top-k cells that satisfy rules.

Practical Application

- Start from topic-based grouping (programme, engagement, budget, logistics)
- Ensure each group has mixed topics
- Use matrix to assign mentors within groups

Session Allocation

High-Complementarity: Small clinics (6–8 participants)

Medium-Complementarity: Round-table discussions

Low-Complementarity: Networking only

Implications for Exchange Camp Planning

Governance & Paperwork (Topics 1 & 4)

MANDATORY SESSIONS

- Role charters and responsibility mapping
- Decision-making frameworks and approval workflows
- Standardised request forms and documentation
- End-of-term handover checklists and procedures

Communication & Engagement (Topics 0 & 2)

INTERACTIVE WORKSHOPS

- Liaison roles bridging pre-clinical and clinical cohorts
- Event scheduling and conflict resolution strategies
- Promotion and publicity best practices
- Scenario-driven sharing and peer discussion formats

Resources & Constraints (Topic 3)

ROUND-TABLE DISCUSSIONS

- Sponsorship models and targeted sponsor packages
- Hybrid event formats for remote participation
- Personal benefits messaging for recruitment
- Budget optimization and resource sharing strategies

Maximize Knowledge Transfer Through Peer Pairing

Topic-Based Pairing: Match institutions with complementary strengths and needs based on their topic profiles to facilitate targeted learning.

Size-Based Pairing: Use the Student Number Cluster Analysis to facilitate peer-to-peer learning among similar-sized institutions (Very Large, Large, Small).

Conclusion & Next Steps

Analysis Summary

HDBSCAN and BERTopic successfully uncovered underlying themes in qualitative SMO feedback without predefining cluster numbers. The dual-approach analysis—combining SWOT-focused and full-narrative perspectives—reveals both structural challenges and broader institutional priorities, providing a comprehensive roadmap for targeted interventions and effective exchange camp design.

Key Findings

SWOT ANALYSIS (5 TOPICS)

- Clinic workflow & coordination
- Governance & programme evaluation
- Communications & participation
- Resources & engagement constraints
- Bureaucracy, paperwork & handover

FULL NARRATIVE ANALYSIS (4 TOPICS)

- Programme design & event consolidation
- Student engagement & well-being
- Budget flexibility & balanced programming
- Operational logistics & staffing

Future Work & Recommendations

Model Refinement

Tune clustering parameters, manually review outliers (BTU and PI), and incorporate additional data sources for enhanced model accuracy and coverage.

Exchange Camp Design

Implement targeted curriculum based on topic profiles, pair institutions with complementary strengths, and use cluster analysis for peer-to-peer matching.