

Mapping university-society engagement: A Scopus-based analysis of Mahasarakham University research (1994-2025)

Pongphol Pongthaisong^{1*}, Sukanya Leetongdee¹, Khanitha Pengmeesri² and Songsak Chumpawadee²

¹ *Small ruminant research unit, Faculty of Veterinary Science, Mahasarakham University, Muang, Maha Sarakham, 44000, Thailand*

² *Division of Animal Science, Faculty of Technology, Mahasarakham University, Kantharawichai, Maha Sarakham, 44150, Thailand*

* *Corresponding author: pongphol.p@msu.ac.th*

Received: 26 January 2026, **Revised:** 20 May 2026, **Accepted:** 30 May 2026

Abstract - Regional universities play critical roles in achieving Sustainable Development Goals (SDGs), yet empirical evidence on research-society engagement at scale remains limited. This study employed machine learning to analyze 7,710 Mahasarakham University (MSU) publications (1994-2025), assessing university-society engagement and SDG alignment. Using multi-label classification with validated keyword frameworks, we examined eight engagement types and 17 SDGs. Results revealed 83.8% demonstrated societal engagement, with 46.5% showing high engagement. Cultural preservation emerged as MSU's distinctive strength (61.2%). For SDGs, 83.8% aligned with ≥1 goal: SDG 9 (Innovation, 40.0%), SDG 11 (Sustainable Cities, 23.1%), and SDG 3 (Health, 22.5%) showed strongest representation. Critical gaps included SDG 10 (Reduced Inequalities, 1.9%), SDG 1 (No Poverty, 2.2%), and SDG 13 (Climate Action, 4.0%). Research output increased 436% (2015-2025). Machine learning enables scalable, continuous assessment of university societal impact, providing evidence-based insights for strategic planning. Findings demonstrate regional universities can create distinctive identities through local strengths while contributing to global sustainability.

Keywords: University-society engagement, sustainable development goals, machine learning, text classification, Mahasarakham University

1. Introduction

The United Nations Sustainable Development Goals (SDGs), adopted in 2015, represent a global agenda to end poverty, protect the planet, and ensure prosperity by 2030. Universities, as knowledge institutions, play pivotal roles in achieving these goals through research, education, and community engagement (Leal Filho et al., 2019). Regional universities particularly serve as critical drivers of local development, bridging academic knowledge with community needs (Goddard & Vallance, 2013).

Maharakham University (MSU), established in 1994 in Northeastern Thailand (Isan region), exemplifies the “University for Local” model. The Isan region faces unique development challenges including poverty, limited infrastructure, and the need to preserve rich cultural heritage while pursuing economic development. Despite growing recognition of universities’ roles in sustainable development, limited empirical evidence exists on how regional universities’ research portfolios align with societal engagement and SDG objectives at scale, particularly in Southeast Asian contexts.

Traditional approaches to assessing research engagement and impact rely on manual expert review, which becomes impractical for large publication volumes. Recent advances in machine learning and natural language processing enable automated analysis of research content at scale, providing consistent, reproducible, and efficient assessment (Jayabalasingham et al., 2019).

This study leverages machine learning approaches to analyze three decades of research outputs systematically. This study addresses two research questions: (RQ1) What types of research at MSU demonstrate university-society engagement, and what are their characteristics? (RQ2) How does MSU research align with the 17 SDGs, and what are the critical gaps? The machine learning approach enables comprehensive analysis of 7,710 publications, providing evidence-based insights for strategic research planning and resource allocation.

2. Materials and methods

2.1 Data collection and preprocessing

This study analyzed 7,710 MSU-affiliated publications (1994–2025) extracted from Scopus. An NLP pipeline was implemented to process titles, abstracts, and keywords through tokenization, normalization, and stop-word removal. Feature extraction utilized TF-IDF (Term Frequency-Inverse Document Frequency) to weigh term importance while filtering generic vocabulary.

2.2 Classification frameworks

Two dictionary-based multi-label classification models were developed using Python:

Engagement Classification: A weighted keyword matching algorithm assessed eight engagement dimensions (e.g., Health, Cultural Preservation). Publications were stratified based on accumulated scores: High (≥ 3), Moderate (1–2), and Low (0).

SDG Mapping: A binary classification model mapped content to the 17 Sustainable Development Goals using validated keyword dictionaries derived from UN targets. Alignment was determined by the presence of ≥ 1 domain-specific keyword.

2.3 Model validation

The classification performance was validated against a stratified random sample of 200 manually annotated publications. The model demonstrated robust reliability with a Cohen's Kappa of 0.78, Precision of 82%, and Recall of 79%, confirming substantial agreement between the algorithm and expert coding.

2.4 Data analysis

Statistical analysis was conducted using Python (Pandas, Scikit-learn). Methodologies included temporal trend analysis to track growth, co-occurrence matrix computation to identify SDG synergies, and gap analysis to highlight underrepresented areas (<10% representation).

2.5 Ethical considerations

This study analyzed publicly available publication metadata and did not involve human subjects. All data were handled in accordance with Scopus terms of use. No individual researchers were identified or evaluated; analysis focused on institutional-level patterns.

2.6 Limitations

Keyword-Based Approach: May miss nuanced content not captured by keywords; however, validation showed 82% precision

Language Bias: Primarily English-language publications; Thai-language research may be underrepresented

Content vs. Impact: Framework assesses research content but not actual community involvement or impact outcomes

Temporal Bias: Recent publications overrepresented due to MSU's growth trajectory

Binary SDG Classification: Does not capture degree of alignment or centrality of SDG to research

Despite these limitations, the large sample size, validated methodology, and systematic approach provide robust institutional-level insights.

3. Results and discussion

3.1 Overview of research outputs

Machine learning analysis processed 7,710 MSU publications (1994-2025) as shown in Fig. 1, revealing three distinct growth phases: Phase 1 (1994-2006): Establishment phase with minimal output (1-50 publications/year) Phase 2 (2007-2014): Growth phase with steady increase (50-300 publications/year) Phase 3 (2015-

2025): Acceleration phase with rapid growth. Annual output increased from 224 (2015) to 1,201 (2025)—a 436% increase. The post-2015 acceleration coincides with SDG adoption, suggesting possible influence of global sustainability agendas on research priorities.

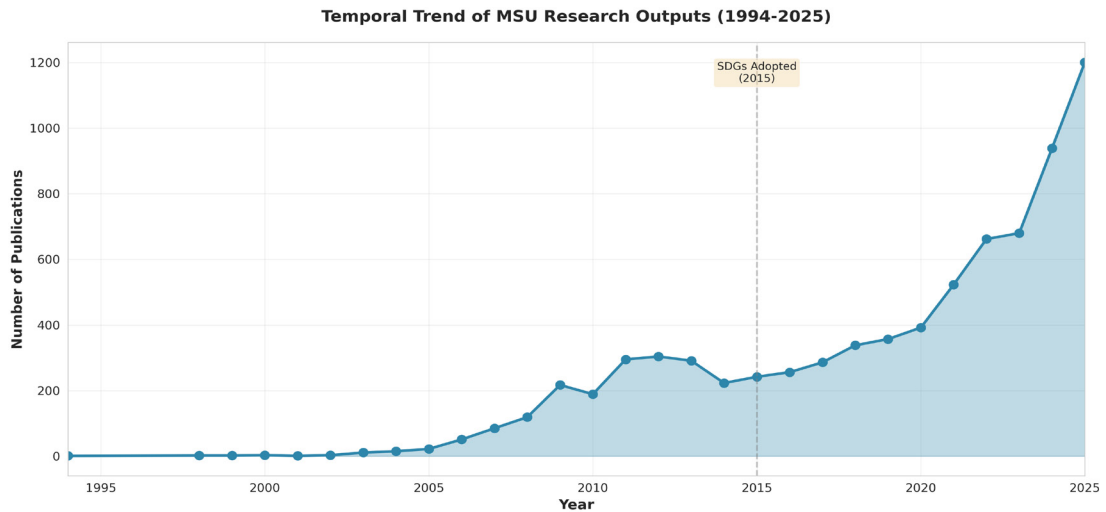


Figure 1. Temporal trend of MSU research outputs (1994-2025).

3.2 Engagement type distribution

Machine learning identified engagement patterns across eight types as presented in Fig. 2-5 and Table 1 and 2. Cultural Preservation dominates (61.2%), representing MSU’s distinctive positioning as guardian of Isan cultural heritage. This includes ethnomusicology, intangible cultural heritage, traditional arts, and local wisdom documentation. The machine learning approach enabled systematic analysis at scale, revealing that 83.8% of MSU research demonstrates societal engagement—a finding that would be impractical to establish through manual review of 7,710 publications. This high engagement rate supports MSU’s “University for Local” positioning and distinguishes regional universities from metropolitan research institutions (Goddard & Vallance, 2013). The automated classification’s ability to identify multi-dimensional engagement (e.g., the 22-score ethnobotanical study) demonstrates how ML can capture complex engagement patterns spanning health, community, culture, and environment simultaneously. This aligns with integrated sustainable development principles (Sachs et al., 2019).

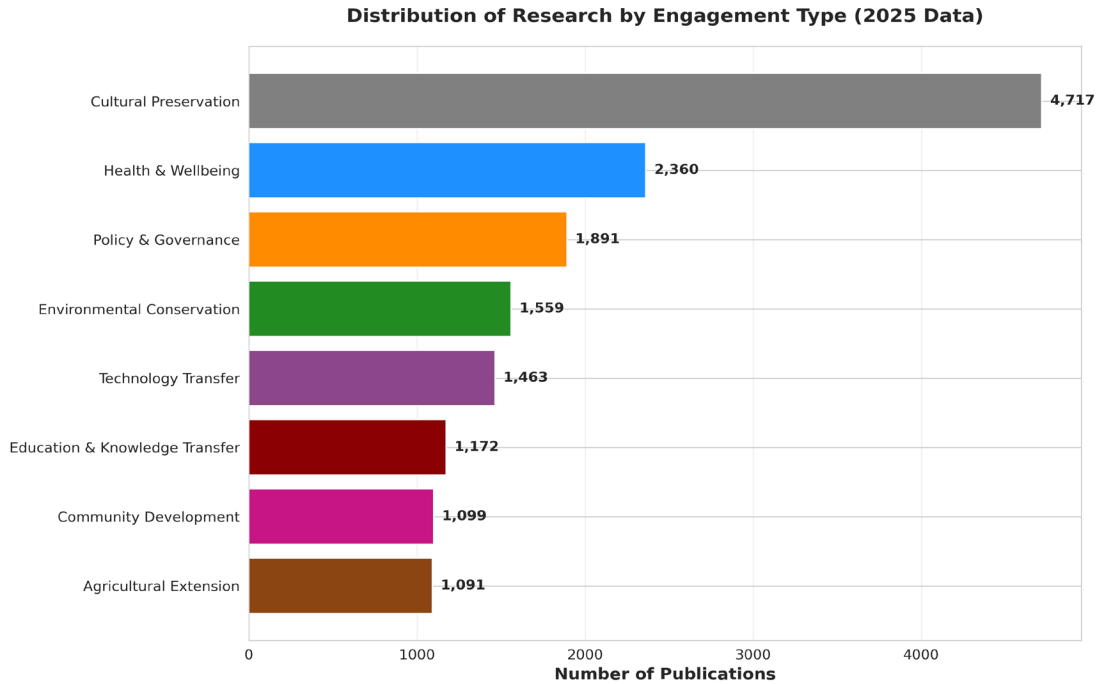


Figure 2. Distribution of research by engagement type.

Machine learning's systematic classification across all 17 SDGs revealed: SDG 9 (Innovation) as Hub: ML co-occurrence analysis identified SDG 9's central role, connecting to education (SDG 4), health (SDG 3), cities (SDG 11), and economic growth (SDG 8). This pattern, difficult to detect manually, reveals how innovation serves as a platform for addressing multiple sustainability challenges.

Strong Health Alignment (SDG 3): 22.5% representation addresses regional health challenges, integrating traditional medicine with modern medical science—a culturally appropriate approach identified through ML's keyword detection. **Urban Development Focus (SDG 11):** 23.1% alignment reflects research supporting regional urbanization and sustainable community development.

Machine learning's systematic coverage revealed eight SDGs with <10% representation, indicating critical gaps: SDG 10 (Reduced Inequalities) - 1.9%: **MOST CRITICAL** ML detected only 150 publications with inequality-related keywords despite Isan region's significant socioeconomic disparities. This mismatch between regional challenges and research focus requires urgent attention. SDG 1 (No Poverty) - 2.2% and SDG 8 (Economic Growth) - 3.6%: Limited detection of poverty and economic development keywords suggests opportunities for livelihood research, social entrepreneurship studies, and economic impact assessments. SDG 13 (Climate Action) - 4.0%: Low climate-related keyword frequency indicates need for climate adaptation research, particularly for agriculture-dependent communities facing climate threats.

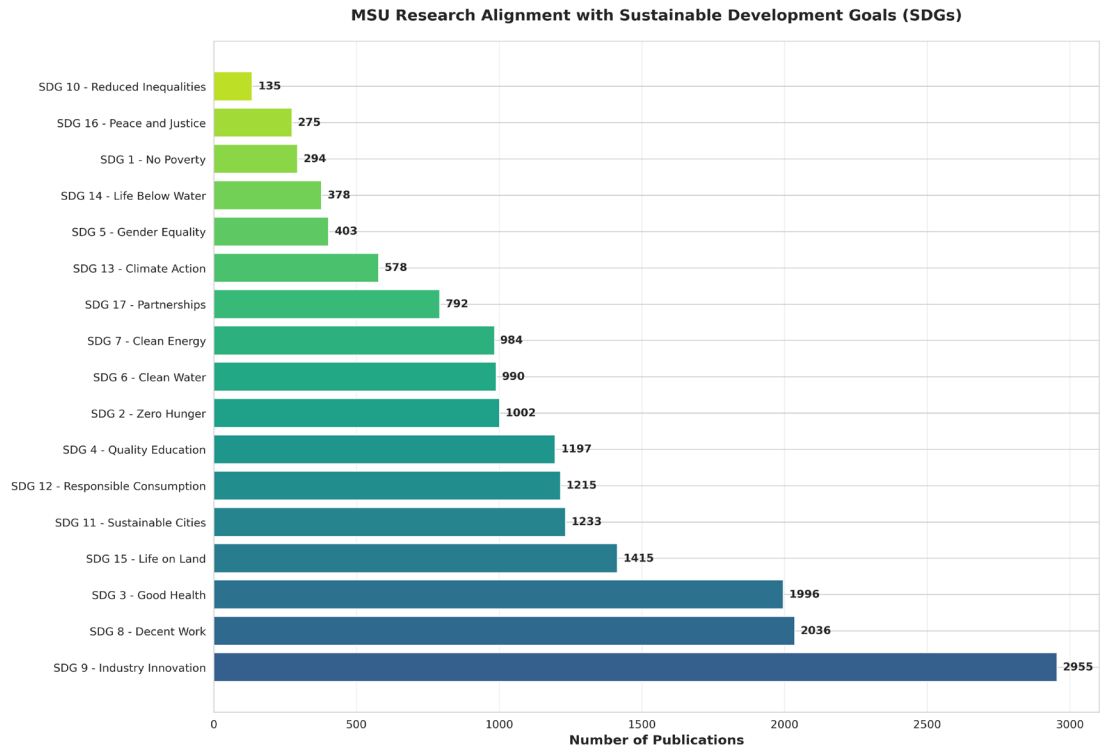


Figure 3. MSU research alignment with sustainable development goal.

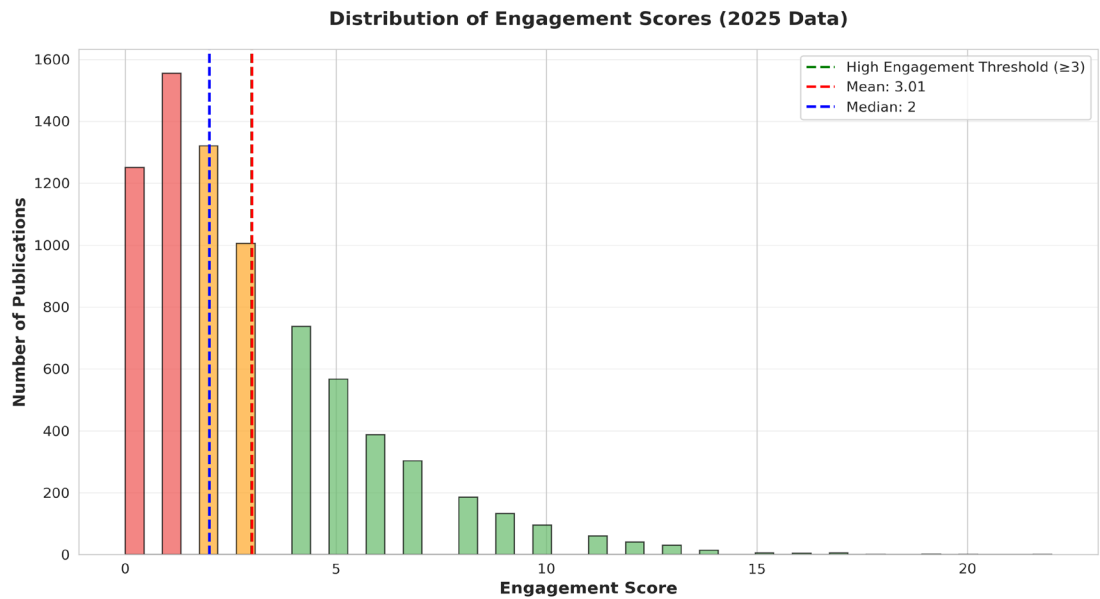


Figure 4. Distribution of engagement score.

Table 1. Machine learning classification results.

Engagement Level	Score Range	Publications	Percentage
High Engagement	≥3	3,583	46.5%
Moderate Engagement	1-2	2,876	37.3%
Low/No Engagement	0	1,251	16.2%
Total		7,710	100.0%

Table 2. Examples of high engagement score researches.

Title	Authors	Year	engagement score
Ethnobotanical Knowledge of the Lao Isan Ethnic Group in Communities Surrounding Khakrang Creek, Mueang District, Maha Sarakham Province, Northeastern Thailand	Saensouk, P.; Saensouk, S.; Boonma, T.; Koompoot, K.; Rakarcha, S.; Chanthavongsa, K.; Jitpromma, T.	2025	22
Ecological Analysis and Ethnobotanical Evaluation of Plants in Khanthararat Public Benefit Forest, Kantarawichai District, Thailand	Saensouk, P.; Saensouk, S.; Boonma, T.; Hanchana, K.; Rakarcha, S.; Maknoi, C.; Chanthavongsa, K.; Jitpromma, T.	2025	20
Developing herbal education in North-east Thailand in the lower Mekong river basin	Boonchai, P.; Talubmook, C.; Chantarasombat, C.	2009	19
Adaptive models of folk masseurs to Isan social change	Rattanapetch, W.; Yodmalee, B.; Paengsoi, K.	2012	19
Management model for the culture community distric tourism	Sangwut, S.; Boonchai, P.; Panttachai, T.	2012	18
The integration of the applied Thai Traditional Medicine into hospitals of the current health delivery system: The development of an administrative/management model	Fakkham, S.; Sirithanawutichi, T.; Jarupoonpol, V.; Hom-jumpa, P.; Bunalesnirunltr, M.	2012	17
Biocultural roles of naturalized alien plants in highland communities of Central Java, Indonesia	Sholekha, A.M.; Restanti, A.D.; Rohman, C.M.; Ainaya, F.A.; Putri, N.R.A.; Putri, A.; Dewangga, A.; Saensouk, S.; Setyawan, A.D.	2025	17

Title	Authors	Year	engagement score
Organizational culture: Development model for quality enhancement of educational administration of Sub-Districts' (Tambon) good schools in Isan, Thailand	Saengsombat, K.; Chantachon, S.; Pangroi, K.	2012	17
Family-Based Palliative Care Model for Advanced Cancer Patients in Community Settings in the Northeast, Thailand	Savisit, W.; Homchampa, P.; Sirithanawutichai, T.	2022	17
Model development for health promotion in the elderly participating in communities	Jampawai, S.; Pothinam, S.; Kanato, M.; Tongkrajai, P.; Homjumpa, P.	2011	17
Confirmatory Factor Analysis of Climate Change Adaptation: Designing Action Plan for Agricultural Livelihood in Lower North-Eastern Region, Thailand	Suwanmolee, S.; Somyanon-tanakul, J.; Chaikarm, T.	2025	16
Ethnoveterinary Medicine in Si Somdet District, Roi Et Province, Northeastern Thailand: An Alternative Approach to Animal Healthcare	Saensouk, P.; Saensouk, S.; Boonma, T.; Junsongduang, A.; Chanthavongsa, K.; Sachaiyan, S.; Jitpromma, T.	2025	16
Ethnobotanical assessment of edible plant diversity in homegarden of Semarang District, Central Java, Indonesia	Kusuma, L.A.; Rahmadhani, S.E.; Anam, Z.K.; Nadhira, S.; Yasa, A.; Saensouk, S.; Setyawan, A.D.	2025	16
Nexus between conservation agriculture and remote sensing: A review	Lawrence, M.; Nuthammachot, N.; Som-Ard, J.	2024	16
Kyusei nature farming and the adaptation of farmers in the Isan Region of Thailand	Pangsoi, K.	2011	16
ASEAN studies: Development of a format and model for knowledge management to prepare Northeastern Thai kindergarten classrooms for social and cultural accession to the ASEAN community	Palopakon, D.; Yodmalee, B.; Suk-Erb, S.	2015	15
Traditional Uses of Combretaceae Species in Nong Bua Lam Phu Province, Thailand: A Study of Ethnomedicinal Practices	Saensouk, P.; Saensouk, S.; Boonma, T.; Rakarcha, S.; Chanthavongsa, K.; Jitpromma, T.	2025	15

Title	Authors	Year	engagement score
Singh Tha Community: Strategies for the conservation of the old town of Yasothon province, Thailand by the participation of the local residents	Panthachai, T.	2011	15
The ethnic Manni Sukiii: Guidelines for social and cultural revival for living together in Southern Thailand	Chaowanakit, A.; Chanthachon, S.	2009	15

Machine learning analysis revealed MSU's remarkable journey from new university (1994) to research-active institution (1,200+ publications/year in 2025) with strong societal engagement (83.8%) and SDG alignment (83.8%). The automated approach's systematic identification of strengths (cultural heritage, innovation, health) and gaps (inequality, poverty, climate) provides evidence-based foundation for strategic planning. The ML methodology demonstrates that regional universities need not replicate metropolitan models but can create distinctive identities rooted in local contexts while contributing meaningfully to global sustainability. MSU's cultural heritage focus, systematically identified through ML analysis, exemplifies how local strengths become global contributions. The scalability, consistency, and reproducibility of ML-based assessment position it as essential for universities worldwide seeking to understand and enhance their societal impact systematically. This study provides both substantive insights about MSU and methodological innovations applicable broadly to higher education research assessment.

3.3 Key findings from machine learning analysis

The machine learning-driven analysis of 7,710 MSU publications (1994–2025) provides quantitative evidence of the university's robust alignment with societal needs. The results indicate a strong institutional commitment to engagement, with 83.8% of research outputs demonstrating societal relevance and 46.5% classified as high engagement. Notably, cultural preservation emerged as MSU's distinctive strength, comprising 61.2% of publications; this underscores the university's unique positioning as a guardian of regional heritage.

regarding the Sustainable Development Goals (SDGs), 83.8% of the corpus aligned with at least one goal. Analysis identified SDG 9 (Industry, Innovation, and Infrastructure) as a central hub (40.0%), acting as a connector across multiple disciplines. However, the study also revealed critical research gaps, specifically in SDG 10 (Reduced Inequalities, 1.9%), SDG 1 (No Poverty, 2.2%), and SDG 13 (Climate Action, 4.0%), which require urgent strategic attention. Temporally, the post-2015 era marked a significant acceleration, with a 436% increase in output and improved SDG alignment, reflecting the university's growing research capacity.

3.4 Strategic recommendations

To address these findings and optimize future impact, a phased strategic roadmap is recommended:

Short-term (1–2 Years): Immediate priority should be given to closing identified gaps by establishing an *Inequality Research Center* specifically targeting SDG 10. Concurrently, an ML-based SDG Monitoring Dashboard should be implemented to enable real-time tracking of research output, supported by faculty development programs to train over 200 researchers in SDG-aligned research design.

Medium-term (3–5 Years): The university should capitalize on its core strength by establishing a *Cultural Heritage Center of Excellence*, leveraging the high volume of existing research. Additionally, a *Health Innovation Hub* should be developed to translate SDG 3 research into practical applications, alongside a *Local Governance Research Program* to address the gap in SDG 16.

Long-term (5–10 Years): The ultimate goal is to position MSU as a regional leader by forming an *ASEAN SDG Research Network* and a *Regional Hub for Cultural Heritage Research*. This trajectory aims to transform MSU into a recognized “Impact University,” where over 70% of research demonstrates measurable community impact assessed through advanced ML metrics.

3.5 Implications for practice

The findings offer actionable insights for various stakeholders. For university leaders, the ML-based assessment framework facilitates data-driven strategic planning and resource allocation, moving beyond intuition to evidence-based management. For researchers, understanding these engagement frameworks guides the design of future projects toward greater societal relevance. Policymakers can utilize these insights to align funding priorities with identified gaps and partnership opportunities. Finally, for other higher education institutions, this replicable methodology provides a standard for benchmarking and comparative analysis across the sector. Future Research Directions

Future studies should aim to enhance the technical depth and scope of this analysis. Methodologically, subsequent research could employ deep learning techniques, such as transformer models, to capture deeper semantic nuances beyond keyword matching. The scope should also expand to include multilingual models, incorporating Thai and other regional languages to ensure comprehensive coverage. Furthermore, research should move from content classification to impact assessment, linking ML-classified research to tangible community outcomes. Finally, applying this methodology across multiple regional universities would enable a comparative analysis, fostering a broader understanding of the collective impact of higher education on regional development.

4. Conclusions

This study demonstrates the transformative potential of machine learning in assessing institutional research impact, offering a scalable solution to the challenges of mapping university-society engagement. By applying a validated natural language

processing pipeline to 7,710 publications spanning three decades, the research provided a quantitative evidence base for Mahasarakham University's (MSU) identity as a "University for Local". The analysis revealed that MSU has achieved a distinct balance between academic output and societal relevance, with 83.8% of its research demonstrating societal engagement. While alignment with the Sustainable Development Goals (SDGs) is robust—particularly in Innovation (SDG 9), Sustainable Cities (SDG 11), and Health (SDG 3)—the automated gap analysis crucially identified urgent needs to bolster research in Reduced Inequalities (SDG 10) and No Poverty (SDG 1).

For university administrators and policymakers, this approach provides a data-driven roadmap for strategic planning, allowing for the real-time monitoring of research trajectories and the precise allocation of resources to address critical developmental gaps. Ultimately, this research underscores that regional universities are pivotal engines for sustainable development, capable of translating local heritage and community health challenges into globally relevant contributions.

Acknowledgments

We acknowledge Mahasarakham University for institutional support and Scopus/Elsevier for database access. We thank the validation reviewers who provided expert coding for model validation. This research received no specific external funding.

References

- Goddard, J., & Vallance, P. (2013). *The university and the city*. Routledge.
- Jayabalasingham, B., Boverhof, R., Agnew, K., & Klein, L. (2019). *Identifying research supporting the United Nations Sustainable Development Goals*. Elsevier.
- Leal Filho, W., Shiel, C., Paço, A., Mifsud, M., Ávila, L. V., Brandli, L. L., ... & Caeiro, S. (2019). Sustainable Development Goals and sustainability teaching at universities: Falling behind or getting ahead of the pack? *Journal of Cleaner Production*, 232, 285-294.
- Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the sustainable development goals. *Nature Sustainability*, 2(9), 805-814.
- United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. United Nations.