A Comparative Analysis of University Strategic Plans and World Ranking Performance

1st Pallavit Aggarwal
Trinity College Dublin
Dublin, Ireland
aggarwpa@tcd.ie

2nd Arnav Bhattacharya
Trinity College Dublin
Dublin, Ireland
bhattaar@tcd.ie

3rd Haoxian Liu
Trinity College Dublin
Dublin, Ireland
liuha@tcd.ie

4th Kaize Shen
Trinity College Dublin
Dublin, Ireland
kshen@tcd.ie

***Abstract*—There has been growing interest in the relationship between university strategic plans and global rankings. Our study will investigate the correlation between the strategic plans of top-ranked universities and their performance in the QS World University Rankings.**

We plan to employ different techniques to analyze the similarity of university strategic plans like, Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA) and Hierarchical Clustering. These techniques have different strengths and weaknesses, and their suitability will depend on the data we scrape from University websites and the approach for data augmentation that we employ. This will further be extended by a correlation analysis to measure the strength of the association between the similarity of strategic plans and the world rankings of universities. The different ways to measure this can be Kendall's tau and Spearman's rank correlation or Pearson correlation, Point-biserial correlation, Biserial correlation etc. This will be best observed on the distribution of the variables being analyzed.

The study will be limited by the availability and quality of the strategic plans and statement of purposes which will be scraped from the websites, which may vary across institutions. Nonetheless, we hope that our findings provide valuable insights for university leaders, policymakers, and researchers seeking to improve institutional performance and reputation.

***Index Terms*—LSA, LDA, Document similarity matrix, Spearman's rank correlation, Text Analysis**

I. Introduction

In recent years, higher education institutions have increasingly emphasized the importance of developing mission and vision statements as part of their strategic planning process. These statements provide a framework for guiding the institution's goals and actions and can help to align the efforts of faculty, staff, and students towards a common purpose. However, there is limited research on how effectively mission and vision statements are being used in practice, and whether they are actually contributing to improved outcomes for higher education institutions.

To address this gap in the literature, this study employs a mixed-methods approach that includes document and rank similarity, as well as content and theme analysis, to examine the mission and vision statements of a sample of higher education institutions. Specifically, the study analyzes the language used in these statements, the themes and values that are emphasized, and the extent to which the statements align with the institution's strategic goals and objectives. By exploring these factors, the study aims to provide insights into how mission and vision statements can be better used to support the success of higher education institutions.

To accomplish this, the study draws on a range of existing research on mission and vision statements in higher education, as well as several recent studies that have used content and theme analysis to analyze these statements. The study also discusses the potential implications of its findings for higher education practitioners, policymakers, and researchers. Ultimately, the study seeks to contribute to a better understanding of how mission and vision statements can be leveraged to support the success of higher education institutions in a rapidly changing educational landscape.

II. Literature Review

A. Does Strategic Planning Matter?

According to the examined literature, strategic planning seems to be an essential component of the success of higher education institutions [3]. There is a need for more clarity and consistency in their content and execution [8] since mission and vision statements may offer a beneficial foundation for directing the strategic direction of these organizations [7].

Research has also demonstrated that the success of strategic planning is impacted by characteristics such as organizational culture, leadership, and stakeholder participation [4][5]. Specifically, extensive stakeholder participation in the planning process may guarantee that the ultimate strategic plan represents the different viewpoints and requirements of the institution's community [4].

Literature indicates that strategic planning is important for the success of higher education institutions. Strategic planning may assist organizations in aligning their resources and activities with their objectives by establishing a clear and shared vision for the future [6]. Yet, it is essential to note that the performance of strategic planning relies on a variety of elements outside the plan's content, and that constant monitoring and review are essential to ensure that the plan stays relevant and effective over time [5].

B. Challenges in University Management

University management is a difficult issue and relying on strategic plans alone to manage a university can often have many limitations. Galbraith [1] has explained exactly what to do when strategic plans are inadequate and why they are inadequate in university management.

Strategic management is effective in many situations, such as corporate management or certain business decisions, and using such policies for management is often effective. However, universities are in many cases different in nature from the institutions mentioned above, and their management cannot rely solely on the various strategies of the institution. There are many uncontrollable factors in the academic research process and it is not possible to simply set various assessment indicators and follow this system rigidly. This may affect some areas of research, such as those that require long periods of time and high costs, but where it is difficult to see significant research progress in a short period of time. At such times, if strategic plans are used for management, such as having an internal competition between university departments for the number of papers produced, etc. This type of management may result in such research directions not being able to compete with other departments within the university, thus losing adequate research funding, etc. The authors suggest that the university should be managed as an ecology, with many hidden influences and links between these influences, in addition to the usual evaluation indicators.

Shannon Chance [2] also points out in her article that current university management tends to be a linear management model, i.e. a simple cause-and-effect accountability system. Simple linear and mechanical thinking in university management is detrimental to the university itself.

Non-linear planning can be modeled on the teaching approach already used in studio-based curricula at universities. The planning strategies used in design programs, including architecture, involve iterative, synthesizing processes that do not follow a strict linear path. The studio format, which is used in the earliest classes, requires high-order thinking, making it an excellent model for educational planners to learn from in developing and implementing responsive, well-synthesized plans. By adopting the design studio example, universities can improve their strategic planning processes and promote critical thinking and adaptive learning among students, faculty, and administrators across all disciplines.

III. Research Methods

A. LDA

Latent Dirichlet Allocation is a probabilistic model used in natural language processing to find latent topics in a collection of documents. LDA stands for Latent Dirichlet Allocation. Topic modeling of this kind assumes that every document in a corpus is a mix of several topics and that each topic is a word distribution. By examining the document-to-document co-occurrence patterns of words, LDA then tries to infer these topic distributions.

Each document is assumed to be generated in the following manner by the LDA model:

From a Dirichlet distribution, select a topic distribution for the document. For each of the document's words:

- *Select a topic from the document's topic distribution*
- *Select a word from the topic's word distribution*

Based on the observed words in the corpus, the objective of LDA is to infer the topic and word distributions for each topic for each document. Bayesian inference methods like variational inference and Markov Chain Monte Carlo (MCMC) can be used to accomplish this.

In natural language processing, LDA can be used for sentiment analysis, document clustering, and information retrieval. It is a well-liked method for analyzing large text data sets, such as social media posts, news articles, and scientific papers.

B. LSA

LSA Latent Semantic Analysis is a natural language processing technique for analyzing and extracting a collection of documents' underlying semantic structure. A type of dimensionality reduction technique, it depicts the meaning of documents and words in a vector space with lower dimensions.

The distributional hypothesis, which states that words that appear in similar contexts typically have similar meanings, serves as the foundation for LSA. A word co-occurrence count matrix is created from the document collection in order to apply LSA. After that, singular value decomposition (SVD) is used to factorize the matrix, which preserves the most crucial information while simultaneously reducing the matrix's dimensionality. In the lower-dimensional semantic space, where semantically similar words and documents are closer together, the resulting vectors represent the words and documents.

Natural language processing tasks like document classification, information retrieval, and question answering can all be accomplished with LSA. It has also been used to investigate human language processing and semantic representation in cognitive psychology and computational linguistics. Nonetheless, LSA has a few restrictions, for example, its powerlessness to catch a few subtleties of significance and its dependence on the co-event measurements of words in the record assortment.

C. Spearman's Rank Correlation

Spearman's rank correlation is a statistical measure that assesses the strength and direction of the association between two ranked variables. It is a non-parametric test that does not assume any particular distribution for the variables.

The Spearman's rank correlation coefficient, denoted as "r_s" or "rho," ranges from -1 to 1, where:

r_s = 1 indicates a perfect positive correlation (i.e., both variables increase or decrease together).

r_s = -1 indicates a perfect negative correlation (i.e., one variable increases while the other decreases).

r_s = 0 indicates no correlation between the variables.

Spearman's rank correlation coefficient is calculated by first ranking the observations of each variable and then calculating

the correlation between the ranked variables. Specifically, it measures the extent to which the ranks of one variable are associated with the ranks of the other variable.

Spearman's rank correlation coefficient is often used when the variables being compared are ordinal or ranked data, but it can also be used for interval or ratio data. It is commonly used in social sciences, business, and other fields to analyze the relationship between variables that may not have a linear association.

D. Kendall's tau Correlation

A correlation coefficient called Kendall's tau measures how strongly two ranked variables are linked. Because it is a non-parametric statistic, it does not rely on any particular data distribution.

In statistical analysis, Kendall's tau is frequently utilized to evaluate the agreement or similarity between two rankings. It can be used to compare how two algorithms rank search engine results or how products are ranked based on reviews from customers.

Kendall's tau has a value between -1 and 1, with -1 representing a perfect negative correlation, 0 representing no correlation, and 1 representing a perfect positive correlation. The context of the analysis and the magnitude of the correlation influence how the value of Kendall's tau is interpreted.

Counting the number of concordant and discordant pairs between the two rankings is necessary for the calculation of Kendall's tau. If the rankings agree on the relative order of the two items, a pair is said to be concordant, and if they disagree, it is said to be discordant. After that, the total number of pairs divided by the difference between the number of concordant and discordant pairs determines the value of Kendall's tau.

Outliers and data non-normality have no effect on Kendall's tau, which is a robust statistic. However, compared to other correlation measures like Pearson's correlation coefficient, it is less sensitive to changes in the data's central tendency.

E. Hierarchical Clustering

Hierarchical clustering is a technique used in data analysis to group data points into clusters based on their similarities or dissimilarities. It is a type of unsupervised learning algorithm, meaning it does not require labeled data to train a model.

The basic idea of hierarchical clustering is to create a hierarchy of clusters, where each cluster contains a set of data points that are similar to each other. There are two main types of hierarchical clustering: agglomerative and divisive.

Agglomerative hierarchical clustering starts with each data point as its own cluster and then iteratively merges the closest pairs of clusters until a single cluster containing all data points is obtained. Divisive hierarchical clustering starts with all data points in a single cluster and then iteratively divides the cluster into smaller clusters until each data point is its own cluster.

The similarity or dissimilarity between data points is measured using a distance metric, such as Euclidean distance or cosine similarity. The choice of distance metric can have a significant impact on the resulting clusters.

Hierarchical clustering can be visualized using a dendrogram, which is a tree-like diagram that shows the relationships between clusters. The height of each branch in the dendrogram represents the distance between clusters, with the longest branches indicating the largest differences in similarity or dissimilarity.

Applications of hierarchical clustering include market segmentation, image segmentation, and gene expression analysis. It is a powerful technique for exploring and understanding complex data sets and can provide insights into patterns and relationships that may be difficult to detect using other methods.

REFERENCES

[1] Galbraith, P. L. (1998). When strategic plans are not enough: Challenges in university management. System Dynamics: An International Journal of Policy Modelling, 10(1/2), 55-84.

[2] Chance, S. (2010) Strategic by Design Iterative Approaches to Educational Planning , Planning for Higher Education, Vol. 38, no. 2, Jan - Mar , 2010. doi:10.21427/D7GK72

[3] Akgun, A. E., Kocabas, I., and Byrne, J. C. (2014). Analyzing strategic plans of higher education institutions using document similarity analysis. Journal of Applied Research in Higher Education, 6(2), 194-209.

[4] Bryson, J. M. (2018). Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement (5th ed.). John Wiley and Sons.

[5] DeCelles, K. A., Leclerc, M. J., and Caza, A. (2016). Exploring the factors that enhance or detract from the effectiveness of nonprofit strategic planning. Journal of Nonprofit and Public Sector Marketing, 28(1), 1-23.

[6] Mullins, M. E., and Komives, S. R. (2014). Mission and vision in student affairs: Mapping the future. Journal of College and Character, 15(1), 1-10.

[7] Oplatka, I. (2009). Educational leaders' perspectives on the mission and vision of their schools. Journal of Educational Administration, 47(3), 352-371.

[8] Ozdem, G. (2011). An analysis of the mission and vision statements on the strategic plans of higher education institutions. Educational Sciences: Theory and Practice, 11(4), 1887-1894.

Statement of Contribution

Although everyone contributed to every section, all authors contributed equally to the work.

Arnav Bhattacharya (Chair) wrote the Introduction and the first part of the Literature Review(Does Strategic Planning Matter?). He actively engaged in discussions and provided constructive feedback to assist the group in determining the most effective methods to implement. Their ability to listen carefully to the viewpoints of others and to convey their own thoughts effectively aided in maintaining the group's focus and forward momentum.

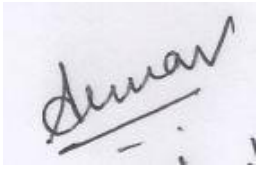
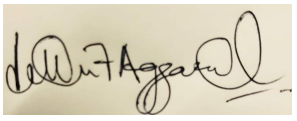
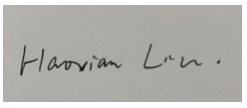
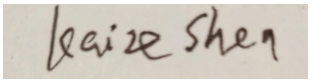
Pallavit Aggarwal (Recorder) wrote the abstract and shared summary of research papers with the team to give more structure to the literature review and research methods. He also provided definition to the approach of the project as per professor's feedback and helped in organising the next steps.

Haoxian Liu (Accountant) wrote the second part of the Literature Review (Challenges in University Management). He actively participated in every meeting, offered constructive advice and conducted thorough research on the subject. He kept track of team members' contributions and effectively promoted the team's work.

Kaize Shen (Verifier) wrote the research methods part and shared his idea of the research paper. He attended every meeting, verifying that the weekly responsibilities of the chair, recorder and accountant are met and the responsibilities of each group member in proceeding with research integrity.

Arnav Bhattacharya, Pallavit Aggarwal, Haoxian Liu, Kaize Shen (Ambassador) took part in other group meetings and exchanged ideas with them.

Arnav Bhattacharya, Pallavit Aggarwal, Haoxian Liu, Kaize Shen

_____ , _____ , _____ , _____