

Spring 3-10-2022

# COMPARING TOP INDUSTRY DEMANDS FOR TALENT TO GENERAL STUDIES CURRICULUM FOR COMPETITIVE JOB ACQUISITION

Gregory Sansone

Follow this and additional works at: <https://aquila.usm.edu/dissertations>



Part of the [Curriculum and Instruction Commons](#), and the [Training and Development Commons](#)

---

## Recommended Citation

Sansone, Gregory, "COMPARING TOP INDUSTRY DEMANDS FOR TALENT TO GENERAL STUDIES CURRICULUM FOR COMPETITIVE JOB ACQUISITION" (2022). *Dissertations*. 1979.  
<https://aquila.usm.edu/dissertations/1979>

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact [Joshua.Cromwell@usm.edu](mailto:Joshua.Cromwell@usm.edu).



COMPARING TOP INDUSTRY DEMANDS FOR TALENT TO GENERAL STUDIES  
CURRICULUM FOR COMPETITIVE JOB ACQUISITION

by

Gregory Sansone

A Dissertation  
Submitted to the Graduate School,  
the College of Business and Economic Development  
and the School of Leadership  
at The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

Approved by:

Dr. H. Quincy Brown, Committee Chair  
Dr. Heather M. Annulis  
Dr. Dale L. Lunsford  
Dr. John J. Kmiec  
Dr. Jonathan Beedle

May 2022



COPYRIGHT BY

Gregory Sansone

2022

*Published by the Graduate School*





## ABSTRACT

University General Studies degrees are rapidly increasing in enrollment, driven by an increased focus on college completion. Yet, non-traditional, non-science, technology, engineering, and math (STEM) degrees are perceived as less competitive in the job market upon graduation. This study finds that students predominantly acquire skills in three areas, analysis, critical thinking, and communication, regardless of courses taken. Variability in salary suggests that the degree was not a factor in the graduates' ability for wage earning. Through a combination of skills outcomes for external job postings, natural language processing on course syllabi outcomes, General Studies graduates' course records, and a survey to General Studies graduates, this study finds that there is a limited number of skills a student can acquire in school and that only a few of those skills match the demand requested in entry-level jobs, however the General Studies degree did not inhibit wage earning potential.

Keywords: General Studies, curriculum alignment, talent demands, skills assessment



## ACKNOWLEDGMENTS

The completion of a dissertation requires a team, this page serves to acknowledge their dedication and commitment to facilitating this outcome. My chair, Dr. H. Quincy Brown who believed and encouraged this project, and spent hours invested in virtual meetings, responding to spontaneous MS Teams messages, and sending texts to ensure I was still alive. I owe you a debt of gratitude and will continue to seek your guidance as I traverse the next stage of my professoriate journey.

I would also like to thank my other committee members. Dr. Heather Annulis, without you and your investment in scholarship related to Human Capital Development (HCD), I would not have attended USM. You have been a constant, in a sea of change, and I am grateful for your words of encouragement and outcome-driven approach to this process. Dr. Dale Lunsford, our time together was brief, but as the resident statistician and as a guiding standard for quantitative research, this manuscript would not have been possible. Dr. Jon Beedle, thanks for the excitement. Since the time of presenting the proposal, you inspired me to keep investing in this complex methodology and saw its potential for making improvements to the academy. That inspiration was infectious, and I appreciate all your help in getting here. To Karen Daas, while not on the committee, without your leadership of the Multidisciplinary Studies (MDS) department at The University of Texas at San Antonio (UTSA) and your investment in me and our MDS students, this project would never have existed.

The staff at USM, especially within Human Capital Development, deserve no less thanks. Robin Johnson, thank you for always making sure I had what I needed. It was not a rare occurrence to receive a late night or early morning email with all the details along



with alternative to ensure I could keep going and pursue this focus. Joyce Powell, thank you for the smiles and ensuring we were always better than good. You encourage all of us to get to where we want to be and you believe in this program.

Finally, I would like to thank my fellow students in the program, especially those that helped me in this venture directly. Tammy Rutland, thank you for being our cohort's Ms. Congeniality, always rooting us on and being available to help. You were my most prodigious tagger, exceeding other taggers in the number of labels, but also thank you for the texts and the phone calls to just check on how we were all doing. Keiasha Hypolyte, thank you for investing in my success. Cory Wicker, my battle buddy, without your support, your brotherhood, your investment, this would not be possible. We did this together and it will be cemented as such in our personal histories.

On a personal note, I thank my parents for always encouraging their children to pursue their goals and providing unwavering support, especially when it was the thousandth time hearing the material contained within these pages. I also thank the two companies where I was employed, Ally and USAA, who funded a large portion of the tuition for this program. Educational advocacy is real, and I thank employers who recognize this, regardless of the degree in pursuance. And lastly to all those friends, family members, coworkers, and other special people who invested in me too numerous to name. I appreciate all your time, energy, and belief that led to this completion. I hope to repay the debt you have so graciously paid forward.



## TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS .....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES .....	x
LIST OF ILLUSTRATIONS .....	xii
LIST OF ABBREVIATIONS.....	xiii
CHAPTER I – INTRODUCTION.....	1
Background of the Study .....	2
Statement of Problem.....	4
Purpose of the Study .....	5
Research Questions .....	5
Research Objectives.....	5
Conceptual Framework.....	6
Human Capital Theory and Signaling Theory .....	9
Curriculum Alignment and Skills Weights.....	10
Significance of the Study .....	10
Limitations .....	11
Delimitations.....	12
Assumptions.....	12



Definition of Terms.....	12
Organization of the Study .....	13
Summary .....	14
CHAPTER II – LITERATURE REVIEW .....	15
Theoretical Framework.....	16
Human Capital and Education .....	16
Signaling Theory.....	18
Skills Weight Theory .....	20
Power of Three.....	22
Higher Education and Performance Funding.....	22
The Focus on College Completion .....	24
Criticisms against Performance Funding .....	25
Evolution of General Studies .....	26
The Flexible Degree.....	27
Interdisciplinary, Multidisciplinary, or Something Else.....	28
Vocational Demand, Missing Skills, and Curriculum Alignment .....	30
Natural Language Processing .....	36
Tokenization .....	36
Parts of Speech Tagging .....	39
Phrase Matching.....	43



Latent Dirichlet Allocation and Amazon Comprehend .....	47
Summary .....	52
CHAPTER III – METHODOLOGY .....	54
Research Design.....	54
Population .....	56
Sampling and Sampling Procedures .....	57
Informed Consent and IRB .....	58
Instruments and Operationalization of Constructs.....	58
Demographic Questions.....	59
Employment Acquisition Questions .....	59
Job Inquiry Questions .....	60
Summary .....	60
Data Collection .....	62
Data Collection Phase I – Entry-Level Skills .....	62
Data Collection Phase II – Course Outcomes.....	63
Data Collection Phase III – Student Data .....	67
Data Collection Phase IV – Survey .....	67
Summary and Data Collection Timeline.....	69
Data Analysis Plan.....	70
RO1 – General Studies Student Characteristics.....	70



RO2 – Employer Skills .....	70
RO3 and RO4 – Course Outcome and Skills Alignment.....	71
RO5 – Students Enrollment .....	73
RO6 – Competitive Position Relationship.....	74
Additional Survey Questions .....	76
Summary and Research Objective/Analysis Alignment.....	77
Threats to Validity .....	78
Ethical Concerns .....	79
Summary .....	80
CHAPTER IV – RESULTS OF THE ANALYSIS .....	81
Demographic Data .....	81
Entry-Level Skills .....	83
Course Outcomes / Skills Alignment.....	85
Student Enrollment .....	95
Competitive Position Correlation .....	106
Summary .....	111
CHAPTER V – DISCUSSIONS AND CONCLUSIONS .....	114
Findings, Conclusions, and Recommendations .....	114
Finding 1 .....	114
Conclusion 1 .....	115



Recommendation 1 .....	116
Finding 2 .....	117
Conclusion 2 .....	117
Recommendation 2 .....	118
Finding 3 .....	118
Conclusion 3 .....	118
Recommendation 3 .....	119
Limitations .....	120
Discussion .....	121
Recommendations for Future Research .....	123
Summary .....	124
APPENDIX A – APPROVAL EMAIL .....	126
APPENDIX B – IRB Approval Letter .....	127
APPENDIX C – SURVEY .....	128
APPENDIX D – LIGHTTAG TRAINING.....	134
REFERENCES .....	135



## LIST OF TABLES

Table 1 <i>Survey Questions and Sources</i> .....	61
<b>Table 1</b> (continued) .....	62
<b>Table 2</b> Data Collection Timeline .....	69
<b>Table 3</b> Survey Map .....	77
<b>Table 4</b> Research Objective/Analysis Alignment .....	78
<b>Table 5</b> Frequency Table for Variable Gender.....	82
<b>Table 6</b> Frequency Table for Ethnicity .....	82
<b>Table 7</b> Frequency Table for Enrollment .....	83
<b>Table 8</b> Salary Bands in Surveyed Population .....	83
<b>Table 9</b> Skills Frequency Table.....	84
<b>Table 10</b> Labels Frequency Table .....	86
<b>Table 11</b> Confusion Matrix .....	88
<b>Table 12</b> Precision, Recall, F1-Score Values for the Multi-label Custom Classification Model .....	90
<b>Table 13</b> Skills-ranking Based on Syllabi Count .....	91
<b>Table 14</b> Top Programs by Skills.....	92
<b>Table 15</b> Total Skills as Average of Syllabi in Corpus .....	93
<b>Table 16</b> Skill Frequency for Analytical Skills.....	96
<b>Table 17</b> Skills Frequency for Communication Skills .....	98
<b>Table 18</b> Skills Frequency for Critical Thinking Skills .....	100
<b>Table 19</b> Skills Frequency for Mathematics .....	101
<b>Table 20</b> Skills Frequency for Problem Solving .....	102



<b>Table 21</b> Skills Frequency for Research .....	103
<b>Table 22</b> Skills Frequency for Teamwork.....	103
<b>Table 23</b> Skills Frequency for Presentation Skills .....	104
<b>Table 24</b> Skills Frequency for Listener .....	105
<b>Table 25</b> Skills Frequency for Creativity .....	105
<b>Table 26</b> Skills Frequency for Building Effective Relationships.....	106
<b>Table 27</b> Results of Initial Variable Influence Factor Analysis .....	107
<b>Table 28</b> Results of Second Variable Influence Factor Analysis.....	108
<b>Table 29</b> Odds Ratio Results for Income .....	109



## LIST OF ILLUSTRATIONS

<b>Figure 1</b> Conceptual Framework.....	8
<b>Figure 2</b> Research Design .....	56
<b>Figure 3</b> Mapped Data.....	65
<b>Figure 4</b> Tagging Screen Example.....	134



## LIST OF ABBREVIATIONS

<i>BG</i>	Burning Glass
<i>UTSA</i>	The University of Texas at San Antonio
<i>HCT</i>	Human Capital Theory
<i>CTE</i>	Career and Technical Education
<i>IT</i>	Information Technology
<i>NLP</i>	Natural Language Processing
<i>TTS</i>	Text to Speech
<i>POST</i>	Parts of Speech Tagging
<i>BLSTM</i>	Bidirectional Long Short-Term Memory
<i>DNN</i>	Deep Neural Networks
<i>CBHG</i>	1-D convolution bank + highway network + bidirectional gated recurrent units
<i>LDA</i>	Latent Dirichlet Allocation
<i>LS</i>	LimeSurvey
<i>PII</i>	Personally Identifiable Information
<i>IRB</i>	Institutional Review Board
<i>SLO</i>	Student Learning Outcomes
<i>TN</i>	True Negative
<i>FP</i>	False Positive
<i>FN</i>	False Negative
<i>TP</i>	True Positive



## CHAPTER I – INTRODUCTION

The top 50% of high school graduates earn the same amount or more than the bottom 50% of college graduates several years after graduation (Harris, 2020). A growing reality exists in today's labor market where recent college graduates find themselves underemployed based on oversaturation, economic instability, increased automation, and misalignment of necessary skills (Pennington & Stanford, 2019). Business leaders worsen this issue by focusing solely on specific majors for entry-level positions (Leighton & Speer, 2020). These leaders attest their focus on these college majors, specifically those in science, technology, engineering, and math (STEM), better identifies talent with the skills required to meet their industry demands (Weise, 2019). Yet, every year, US colleges and universities graduate 50% more students from STEM programs than entry-level STEM jobs (Mills, 2021), suggesting that there may be a greater focus on this major than necessary.

The potential exclusion of non-STEM majors for entry-level jobs worsened with the pressure on universities to maintain graduation rates. Since 2009, the federal government and most state governments have encouraged higher graduation rates using performance funding mandates (Gándara & Rutherford, 2018). These mandates help ensure students don't incur high levels of student debt by supporting 4-year graduations and theoretically assisting in meeting job market demands (Gandara et al., 2017). These mandates have spurred the growth of non-traditional degrees such as General Studies (Hill et al., 2019). A degree that allows flexibility in class selection and more focus on hour accumulation than the specificity of study (Cadwell Bazata, 2020) and intends to assist in helping students complete a degree (Hill et al., 2019). This research study



focuses on understanding whether General Studies provides students a curriculum framework that prepares students to compete in the market and identify specific classes or roadmaps that students can take in the degree program to increase their relevancy in the job marketplace.

This first chapter outlines the foundation of research performed to include a background on the challenges college graduates face in the entry-level job market. The chapter continues with an outline of the research objectives and the questions the research presented. The chapter concludes with a highlight of the theories that frame this study, along with the significance of helping to address issues graduates face in being competitive.

### Background of the Study

Surveys of incoming first-year students' perceptions continue to confirm the belief that having a degree, regardless of major, prepares them to do well financially (Lock & Kelly, 2020). The finding validates a human capital theory principle that investment and attaining a degree will lead to higher wages (Kunz & Staub, 2020; Schweri & Hartog, 2017). Several studies conclude that bachelor's degree holders earn significantly more than those who have earned only a high school diploma over an entire work-life, with special consideration given to academic program selection (Cellini et al., 2019; Lobo & Burke-Smalley, 2018). This academic major selection is more than a special consideration. As research suggests, the difference in lifetime earning potential between the highest-earning and lowest-earning majors averaged \$3 million (Steele et al., 2020). College major selection also affects total employment, as non-STEM, non-business



graduates experience higher unemployment than those with technical or perceived occupationally aligned degrees (Leighton & Speer, 2020).

According to National Center for Education Statistics, 41% of bachelor's degrees conferred in 2019 were non-Business, non-STEM degrees (Hussar et al., 2020). Based on the research that employers are focused primarily on these majors (Leighton & Speer, 2020), these graduates may struggle with competitiveness in the job market. Many graduates fill positions that do not require a degree, ultimately weakening employee engagement, and increasing overall attrition for employers (Adeola, 2017). For other graduates, the disconnect between employer expectations and their degree can mean unemployment (Leighton & Speer, 2020). However, business leaders lament that sourcing talent with the desired capabilities remains challenging even with a surplus of graduates (Weise, 2019). In over a third of external job postings, most skills requested by employers are foundational, such as analysis, critical thinking, or research (Jackson, 2020). Therefore, the issue might not be overeducation, as some research suggests (Ralston, 2021), but a lack of career readiness (Burke et al., 2020).

Career readiness presents a set of unique challenges for both universities and their students, the greatest of which is determining industry requirements and the university's curriculum (Pennington & Stanford, 2019). Saulnier (2017) proposes several strategies for overcoming these challenges. Strategies include using industry partnerships to help forecast future skills and employment opportunities, leveraging requisite skills examination against job postings (Verma et al., 2019), deploying employer and former student surveys (Thouin et al., 2018), and holding focus groups (Handali et al., 2020). All of which can help develop an alignment between industry and institutions of higher



learning. Most of these methods, however, are used once and not continuously (Saulnier, 2017). Another challenge is that these methods align to only a single industry or a structured degree program such as education, engineering, or information technology (IT), limiting applicability to all degree programs (Handali et al., 2020; Thouin et al., 2018; Verma et al., 2019).

### Statement of Problem

Academic major specificity is a significant factor in employability and positive industry outcomes. This factor is a concern when recognizing that one of the fastest-growing majors in the United States is not a STEM degree. General Studies grew 79% in 10 years and averaged over 50,000 graduates each year (Wright, 2016). Much of the General Studies' degree growth was not organic, but the universities attempting to address a growing challenge in meeting graduation rates. Over 30 states have performance funding, whereby public institutions' graduation rates affect state contributions and support (Laderman & Weeden, 2020). The impetus for funding changes results from an attempt by lawmakers to reduce total college expenditure by reducing the time an individual student spends earning a degree and increasing the number of students completing a bachelor's degree (Li, 2020). By preceding a traditional degree plan, universities allow students to leverage many courses to facilitate the minimum credit hours for graduation (Wright, 2016). The General Studies degree, therefore, is advantageous in helping students remain interested in school and complete a bachelor's degree (Gill, 2018). However, the lack of skill specificity magnifies an existing challenge present in traditional degrees, ensuring more opportunities for under- and unemployment



for degree holders (Hora, 2020), driving hiring managers' struggle to identify positions for General Studies graduates (Deming & Noray, 2019).

### Purpose of the Study

This study's primary aim was to understand the talent demands for the entry-level job market related to outcomes in courses available to General Studies degree holders and their relation to competitive job acquisition. Improving the understanding of the influence of course selection in employability for General Studies degree holders may help advise future students in course selection. The understanding may also assist employers in identifying General Studies students as hireable assets and maximize financial returns for General Studies graduates.

### Research Questions

The following questions satisfy the purpose of comparing industry talent demands concerning course outcomes for courses available and completed by General Studies degree holders and the potential for competitive job acquisition:

1. Do General Studies degree holders meet the talent demands of industry upon graduation?
2. Does the curriculum available to General Studies students align with the talent demands of the job market?
3. Do General Studies students enroll in curriculum which aligns to industry talent demands?

### Research Objectives

Research objectives are This study aimed to understand course alignment related to talent demands of industry, specifically for students who earn a General Studies



degree. The research assessed this alignment through several independent assessments, culminating in a final evaluation of competitive job acquisition.

*RO1* - Describe the demographics of General Studies graduates participating in this study.

*RO2* - Describe the skills employers seek for competitive entry-level positions that do not require a specific college major.

*RO3* - Describe the skills, as defined by course outcomes, for courses available for enrollment by General Studies degree-seeking students.

*RO4* - Determine which courses align to competitive entry-level positions.

*RO5* - Determine General Studies students' enrollment in courses that align with competitive entry-level positions.

*RO6* - Determine the relationship between students who took competitive courses and those who achieved competitive entry-level positions.

### Conceptual Framework

A conceptual framework visualizes the research questions and their alignment (Kivunja, 2018). Figure 1 illustrates the conceptual model for this study. The first research objective supports or refutes the literature on the makeup of General Studies students. Literature suggests that these programs no longer attract highly motivated students and instead attract students from more underrepresented groups such as adults, ethnic and racial minorities, and military personnel seeking degree completion over knowledge acquisition (Hoyt & Allred, 2008). The research assesses this phenomenon through university institutional research data.



RO2 was satisfied by analyzing a 90-day sample of entry-level job data collected by the employment analytics firm Burning Glass Technologies (BG) using the skills-bundle analysis. BG is an analytical software company that aggregates job posting information from over 40,000 sources, providing a list of critical skills for each job and role (Burning Glass Technologies, 2021). This assessment assumed that individual skills are general and that the number of unique skills required, if different when aggregated than the labor market, makes a job unique. The more specific the skill set needed, the supposition is that more competitive the level of wage offered (Chan et al., 2017; Deming & Noray, 2019; Eggenberger et al., 2018; Heisig, 2018; Jackson, 2018; Leighton & Speer, 2020; Light & Schreiner, 2019).

An automated process extracted the course outcomes for General Studies courses students can enroll in from syllabi stored in a university system. The research derived an exhaustive set of skills using a language processing algorithm and the dictionary established in RO2. Recognizing that the skills were not exclusive for each course, this approach highlighted each course's potential for matching against the skills required for entry-level jobs to satisfy RO3 and R04.

RO5 and RO6 required a survey of General Studies degree graduates. The research involved analyzing all completed courses per student before students' survey receipt. The students received an individual survey to verify their employment, industry, job class and help provide a perception of their salaried potential. In tandem with the associative skills, these results ran through an ordinal regression to determine the likelihood of a positive impact on acquiring a competitive entry-level job.

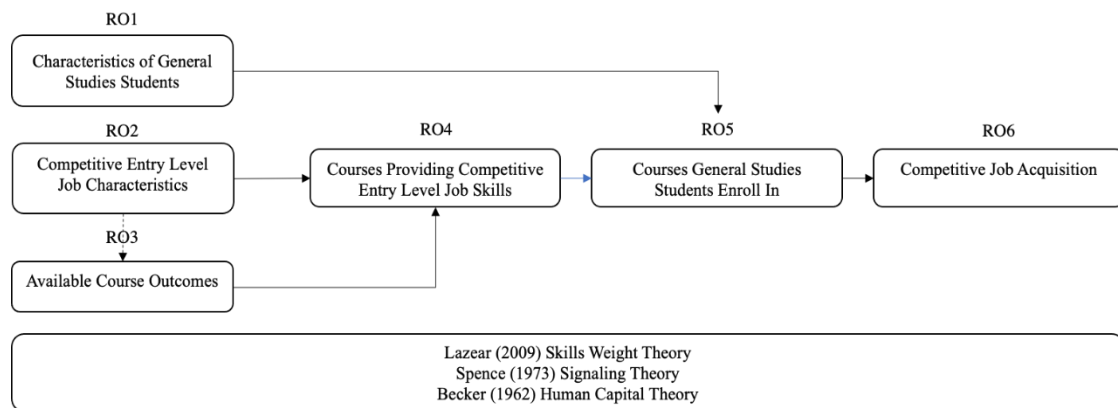


The five research objectives outlined provided a perspective of the specific skills required for entry-level positions not requiring a specific major and identified if General Studies degree holders are competitive upon graduation based on courses' availability. The first step recognized that a student has a choice to take these courses; therefore, the research sought to understand if it was possible to be competitive based on different decisions. Influencing this outcome was the results from RO4. It intended to identify the most competitive students based on the skills bundle and their effectiveness at gaining employment competitively.

Each of the research objectives was supported fundamentally by Becker's (1962) human capital theory, Spence's (1973) signaling theory, and Lazear's (2009) skills weight theory which provide the framework of investigating skills adoption. Each of the researchers focused on investments people make to increase their competitive potential.

**Figure 1**

*Conceptual Framework*





### *Human Capital Theory and Signaling Theory*

As Becker (1962) outlined, the human capital theory suggests that education positively increases ones' human capital, allowing effective competition in the labor market, translating into higher earnings. Education provides general capital, widely transferable skills between industry and organizations, and specific capital useful for only one organization or industry. Traditionally, the theory suggests that direct investments in only specific capital are riskier investments because they limit individuals' ability to adapt to industry changes, resulting in a higher potential for wage loss or unemployment (Gibbons & Waldman, 2004). However, investments in specific capital generate higher returns, particularly theories on task-specific (Kambourov & Manovskii, 2009), occupational-specific (Neal, 1995), and industry-specific (Spence, 1973) human capital because the skills align to occupational requirements. As Jackson (2020) suggested, a large percentage of competitive entry-level job postings seek primarily general capital outside of specific college majors. Therefore, employers may leverage a principle from signaling theory (Heisig, 2018).

Signaling theory stipulates the overvalue a specific degree or specific skills over the many general skills they require. Because of the overwhelming number of candidates applying for each position, employers use methods to limit the number of individuals requiring in-depth review, often settling on elements of a resume, specifically identified skills, as indicators of productivity as opposed to indicators of actual capabilities (Chan et al., 2017; Pennington & Stanford, 2019). Leveraging these two elements in tandem provides a foundational lens on the capital required to meet industry demands and potential opportunities for generalist degrees to meet specific skill demands.



### *Curriculum Alignment and Skills Weights*

Several studies have offered guidance on how universities can better align with local labor markets (Light & Schreiner, 2019; Silos & Smith, 2015). Researchers focused on labor market alignment to the curriculum have isolated their investigation to specific programs instead of an entire university catalog (Ntiri et al., 2004). However, several researchers have provided insight into grouping subjects' and courses and aligning them into categories, most of whom use an adopted framework from Lazear's (2009) skills weights approach. This approach assumes that all skills are general capital when viewed individually. The combination and weighting of these skills in a defined capacity make a skill bundle or a form of specific capital. In this way, one can compare the bundle and weights of single skills provided within a curriculum against a set of required skills and weights in the labor market (Lazear, 2009).

### *Significance of the Study*

Previous research has long considered higher education attainment a method for achieving high-income growth (McMahon, 2018). These realities are changing, and the focus on skills is more relevant today. Several studies have investigated degree specificity or curriculum as a predictor of labor performance (Pennington & Stanford, 2019); however, they have not focused on General Studies degree programs. Using skill bundles (Lazear, 2009), this study can provide a quantitative perspective on the job prospects for General Studies graduates and provide a view on the course alignment most congruent with positive labor performance. With the growing number of graduates with this degree, this study could benefit academic advisors in assisting students, not just those in General Studies, curate courses that have the most significant marketability for future



employment, potentially increasing the university's brand. Employers and industry leaders will benefit by being able to draw from a qualified workforce. Aided by using skills alignments, employers can evaluate more effectively candidates, especially those from non-traditional degrees, reducing underemployment for current degree holders and future graduates and meeting the competitive demands of industry.

### Limitations

A limitation in a study design is the systematic bias that a researcher could not or did not control and could affect the result (Brutus et al., 2013). There are several limitations to this study. The primary will leverage BG data and will only use a year's worth of job vacancy data. This decision is because of the availability of data and the extraction of associative skills. As extracting and translating the skills from course outcomes was already great, leveraging an already listed skills inventory aligned to the current market ensured a more straightforward translation.

Another limitation and one that researchers have frequently identified (Föll & Thiesse, 2017; Goldring, 2017; Gromov et al., 2019) is those course outcomes may not accurately define the course's competencies. With data mining techniques, a concrete definition set should be attainable. However, these definitions may not accurately associate with job requirements directly, rendering some jobs unassociated. However, with the size of the skills leveraged, a macro association should suffice.

The final limitation is that this study potentially lacks external validity, or the ability to generalize to a large population (Field, 2018). The competencies defined should be universal and, with little work, be associated with other university curricula. However, course availability differs between universities and degree requirements, and the results



did not match directly with other institutions, reducing the generalizability of these results.

### Delimitations

The boundaries or choices a researcher makes within a study are called delimitations (Brutus et al., 2013). This study exclusively looked at General Studies graduates from a single university for the past 3 years. The use of a larger population may have increased survey participation, but the 3-year window ensured alignment with the available course catalog for syllabi at the university. The variability of curriculum available to students and program requirements at other universities guided the focus on a single university program.

### Assumptions

An assumption in a study are items or features that are out of the control of the researcher, but if they were to disappear the study would become irrelevant (Brutus et al., 2013). There are several assumptions concerning this study. The first of which is that all degree holders will answer all questions in the survey. The second is that when answering questions, they provide an accurate perception of their realities. A third assumption is that BG's data accurately represent the skills associated with entry-level jobs, and there are no data quality issues. A fourth assumption is that all syllabi contain course outcomes.

### Definition of Terms

The following definitions of terms are pertinent terminology to this study and used throughout the manuscript.



1. *Competitive Job* - A competitive job will be one that, without requiring a specific college major, has a starting salary in congruence with averages projected by the National Association of Colleges and Employers > \$58K (Muafi, 2019).
2. *Entry-Level Job* - A position that does not require advanced skills or previous experience (Muafi, 2019) .
3. *General Capital* - Skills that widely apply in all settings or organizations (Becker, 1962).
4. *General Studies* - Sometimes referred to as interdisciplinary or multidisciplinary, the degree allows students to focus on multiple concentrations and does not include the rigor of a traditional college major (Cadwell Bazata, 2020).
5. *Non-technical Degree* - College majors that do not align directly to an industry or offer "technical skills" as defined as advanced math, science, healthcare, or engineering. These degrees include business generalists, liberal arts, social sciences (Becker, 1962).
6. *Specific Capital* - Skills that only apply to a specific firm or industry (Becker, 1964).

### Organization of the Study

This dissertation proposal explores the association between human capital and competitive acquisition related to students' curriculum in a General Studies degree program. The research design model includes degree holder surveys, examining course outcomes, and analyzing entry-level job requirements. Chapter I introduces the concept of skills association to entry-level job alignment. The chapter related the aforementioned problem to a lack of specificity concerning skills acquisition and theoretical model lenses



to guide the analysis. Chapter II explores theoretical models that influence entry-level job acquisition decisions and a theory that explains the importance of human capital formation related to career development. Introducing previous research in the space of labor alignment, curriculum development, employer expectations, and major specificity illustrate the nuanced challenges students face in general curriculum education. Chapter III presents the data collection method, the sample, and the analysis strategy. Chapter IV provides the results of the analysis. Chapter V concludes the study with findings, discussions, and recommendations for future research.

### Summary

This chapter provided an overview of General Studies degrees, the challenges graduates face in a competitive job market, and the purpose of this study focused on investigating methods to improve graduates' outcomes. In summary, several factors draw attention to the value of a college degree. These factors include employment competition, a demand for higher skills within the global labor market and institutional accountability for occupational preparedness. By focusing on the specificity of skills required for competitive entry-level positions, General Studies degree programs might be able to enhance their effectiveness in terms of positive career outcomes by providing outlines of curriculum that meet current and future industry needs. The following chapter, Chapter II, focuses on the relevant literature and provides a foundation for the body of this study's investigation.



## CHAPTER II – LITERATURE REVIEW

The ability to get and maintain employment relies on one's possession of skills, a set of skills deemed helpful by the labor market. One's employability or the continuous process of acquiring experience, knowledge, and relevant skills to remain marketable in shifting labor markets (Pologeorgis, 2019) begins often and continues into formal education. Based on this, labor force participants believe education, regardless of which degree program, will help lead to a better job (Lobo & Burke-Smalley, 2018), a value proposition critical to validate for the future of higher education.

This chapter summarizes the literature related to gaining a General Studies degree and meeting industry talent prerequisites. The chapter begins with a theoretical perspective on skills acquisition, precisely its function as a competitive advantage, as defined in the realm of human capital (Becker, 1962) through later focused efforts to understand its relevancy in the modern labor market. The chapter will then explore the role of federal funding, the pressures to maximize graduation, to ultimately the creation of degrees like General Studies to meet these demands. The chapter expands on the General Studies design at several universities, settling in on its unique implementation at a single institution in the Southwest. In conclusion to curriculum, the chapter reviews research and findings on methods previously performed related to industry talent expectations and curriculum alignment, providing the fundamental support for the research questions. The chapter concludes with a focus on analysis methods, specifically the use of natural language processing and the foundational context on its utilization and advancement, lending credence to its selection as a method expounded in chapter three.



## Theoretical Framework

Human capital theory (HCT) and its evolution frame the foundation of this research, specifically how education as human capital influences earning potential for workers and serves as a positive investment for employers. This literature review does not neglect the criticisms but leverages them as additional elements to evaluate.

### *Human Capital and Education*

Defined as the traits (knowledge, skills, aptitudes, etc.) that humans used to contribute to production (Goode, 1959), human capital is a combination of early abilities (whether innate or gained) and skills attained through formal training and education (Abuselidze & Beridze, 2019). Human capital differs from other types of capital, as it only yields a return concerning an individual's labor supply. Human capital is one's stock of capital that determines their earnings, or the value of the skills leveraged and length of time those skills remain relevant (Klees, 2016).

Becker (1962) surmised that human capital development was "all activities that influence future actual income through the embedding of resources in people" (p. 9). Therefore, any expenditure on formal education or training, health services, and labor mobility is an investment in, not the consumption of, human capital (Weisbrod, 1966). These investments are costs undertaken to gain a higher return rate regarding education, which is most interesting to this study. These costs can come as tuition expenditure or lost or reduced earnings during training (Becker, 1964). Therefore, like investments in physical capital, individuals will only spend the resources if the expected return is more significant than that achieved with the current level of capital, i.e., skills, held (Becker, 1964).



As the theory relates to investments in education, the original model for human capital distinguishes between general and specific capital (Becker, 1964). Capital useful to a current employer or industry and potentially beneficial for other employers or industries is general. Alternatively, if the capital only increases productivity in a current role or industry, the capital is specific. This distinction is vital to understanding the accumulation of human capital as this only occurs through three methods. These methods are formal schooling (individual devotes most of the time to learning), off-the-job training (post formal education provided by proprietary institutions), and on-the-job training (supplied by current employer; Becker, 1992). In Becker's opinion (1962), education is the single most significant investment a human can make; because education migrates with them regardless of employer, cautioning too much investment in specific capital restricts the individual's ability to move.

However, several challenges arise when evaluating an investment in education as positive capital. The first of which is the definition of education. Schultz (1963) suggested that education has several connotations and depends on the culture and community education occurs. Education is the process by which learning, and teaching happen (Schultz, 1963). The development of moral and mental aspects of a person to assist in rational decision making and training for vocational acumen in both skills and cultural suitability are direct objectives of education (Schultz, 1963). Yet, as is the same about the cultural connotations of education, the outcomes achieved also exhibit variability, especially with value. As with any form of capital, depreciation is a certainty. Different forms of human capital reach obsolescence faster than others. For instance, education in Assembly, a computer programming language, while valuable to understand



theoretically, is all but obsolete to newer languages such as Python or Java. This obsolescence risk requires evaluations by the education provider and the human wishing to expand their investment.

Another challenge facing education as a positive human capital investment is that education is challenging to measure like all human capital investments (Sweetland, 1996). Therefore, most of the 'value' achieved from education is indirect. The method pioneered by Schultz (1963) and further advanced by Becker (1964) was to look at differences in financial outcomes for people who pursued higher education. They found in both cases that the average rate of return, or rather the amount of money spent on education post-high school vs. the financial compensation the 'investor' received, was an average of ten to thirteen percent (Becker, 1964; Schultz, 1963). This finding has served as a framework for many studies related to higher education value and has held relatively constant for several decades (Sweetland, 1996).

### *Signaling Theory*

A confirmation that investment in education results in a perceived positive outcome for the investor is the increase of the population earning a college degree. The percentage of people with a bachelor's degree or higher in the United States increased by 16% between 2009 and 2019 (McElrath & Martin, 2021). However, the ubiquity of college degrees among potential job applicants may have a negative effect. Spence (1973) suggested that any attribute indistinguishable between candidates is no longer valuable as an indicator of productive ability. Productive ability is a crucial outcome for investing in human capital (Becker, 1962).



In his seminal work, Spence (1973) theorized that people, to employers, are a combination of two types of attributes, fixed attributes, such as gender or race, which he calls Indices, and alterable attributes, such as education level, which he calls Signals. Potential employees manipulate their signals to communicate to future employers their productive ability (Spence, 1973). If enough applicants are successful, the result has the potential of creating signaling equilibrium, whereby the employers' beliefs (associated wages, level of desired acumen, future hiring decisions) shift to meet the recent signals (Spence, 1973). This reality is both positive and negative. With hiring entry-level employees, an employer may believe that a specific signal, i.e., a bachelor's degree regardless of major, is a valuable indicator of productive ability and pay accordingly. However, as that employer hires continuously over time, they may identify that particular majors within a bachelor's degree are most valuable and hire for that, only shifting the signals required for job applicants to send. The obverse, the employer may hire with a particular signal in mind and not find enough signals in the market within the cost range and therefore choose to abandon that signal or increase associated wages. Thus, according to Spence (1973), a signal is only a positive one in that ultimately it sets the signaler apart from others.

This theory is not without its weaknesses, primarily that the theory doesn't account for signals communicated during the hiring process, nor does the theory show multi-dimensional scales when many signals are transmitted (Karasek & Bryant, 2012). Several research efforts have attempted to address this by offering cues where employers evaluate individual signals (Forbes, 1987; Herriot & Pemberton, 1996; Iellatchitch et al., 2003). However, a combination of signals, including years of experience, unique



positions held, certifications gained, etc., had an impact, not a specific unique signal. This weakness is essential, as it suggests that individual signals may be more impactful early in one's career when one has less differentiation than in later years.

One item of Spence's (1973) theory is the importance of indices. While important to employers for candidates to have powerful signals that differentiate them, their unalterable indices, such as race or sex, may still play a role. Instead of education as a signal being a single group, candidates become multiple groups Latinos with education vs. African Americans with education, etc. According to Spence (1973), these signals may still influence an employer's beliefs, which, while prejudiced, may create an internal signaling equilibrium. For a Latino to communicate more remarkable productive ability, they must have stronger signals than someone who is White, i.e., be Latino and not just a bachelor's degree, but a bachelor's degree in statistics.

### *Skills Weight Theory*

Yet, another challenge with the human capital theory is that it allows someone employment only with specific skills; for instance, organizations that deal with computer programming will only hire individuals with specializations in that field (Lazear, 2009). It has several implications that wage profiles and the split of human capital cost are dependent only on the market thickness. When individuals change their jobs involuntarily, they are likely to get less income on their following jobs (Lazear, 2009). Firms only pay for what appears to be general training (Lazear, 2009). Skill weight theory, however, allows for all skills to be available. The skill weight approach implies that workers are likely to lose earnings associated with their previous job possessions in an exogenous job change. When the exogenous probabilities are low, involuntary



changes are higher, and employers require workers to distinguish themselves from others based on their weights from the initial firm. The more one showcases his distinction, the more unbalanced the incomes are.

In skill weight theory, there are no skills that need to be firm-specific, or rather that skill is only valuable to a single firm. Conversely, the individual's assets may appear to be general because other firms may need them. Day et al. (2001) extended research to support the skill weight idea that individuals can quickly learn new skills, increasing skill retention and skills transfer. However, an individual's knowledge structure influences the cognitive ability and skill-based outcomes in their new fields after training (Day et al., 2001).

The skill-weights method developed by Lazear offers a model for defining specificity at the skill level, assuming that all skills are general but that the specificity of combinations of single skills differs (Lazear, 2009). In the study by Day et al. (2001), they identified specific crucial skills for an occupational performance degree.

Simultaneously, the more precise an occupation's skill requirements are, the lower the likelihood of an employee changing jobs over their entire career (Day et al., 2001). Detail varies on the general contents of bundled training within an apprenticeship training.

Therefore, firms reduce customer mobility by offering training to their employees to increase their specificity of skills combinations (Day et al., 2001). It differs from other types of capital, as it only yields a return concerning an individual's labor supply.

Therefore, one's stock of skill assets determines their earnings, or the value of the skills leveraged and the time those skills remain relevant (Klees, 2016).



### *Power of Three*

Each of the three theoretical approaches is vital in understanding higher education's role in creating economic returns for individuals. Rising tuition costs and mounting student debt have recently led some to question a traditional bachelor's degree (Fishman et al., 2019). Most researchers' arguments focus on the direct monetary value ratio, primarily that a college degree does not guarantee a high-paying profession upon graduation (Bird & Smith, 2005). However, evidence (which varies by degree type, geography, and ethnicity) still shows that college graduates earn more over their lifetime than non-degree holders (Abel & Deitz, 2018). This evidence is consistent with human capital theory, which states that the accumulation of capital combined with time spent on application creates a positive return. Yet, as shown through signaling theory and skills weight theory, as degree attainment becomes more ubiquitous, differentiators are necessary to narrow the entry-level scope of candidates, creating large wage variability upon exit from the institution (Hora, 2020). It is this variability that drives the questions about higher education.

### Higher Education and Performance Funding

The modern university in the U.S. has developed considerably since the founding of Harvard almost four centuries ago. Harvard's initial role was to train clergy, doubling as a source of validation of prestige for families with multiple sons, sending all but the oldest to "learn life skills" (Dorn, 2017). This model existed until the founding of the first publicly funded institution, The University of Virginia, in the early 1800s. Championed by Thomas Jefferson, The University of Virginia controversially shifted away from the traditional religious doctrine and specializations in medicine or law (which previously



were the only three focus areas). Instead, they opted to provide affordability and invest in beneficial economic pursuits such as astronomy, architecture, philosophy to advance the public (Dorn, 2017). This perspective extended further when, in 1862, President Lincoln signed the Morrill Land Grants Act. This act allowed states to apply for free land from the government to enact public universities with a caveat. They promoted agriculture, engineering, and mechanic arts, deemed practical to the burgeoning workforce (Dorn, 2017). Introducing the GI Bill, the U.S. government advocated the necessity of a 4-year college degree to transition in life, further promoting the mainstream attitude and expansion of traditional college education (Dorn, 2017). By the 1950s, a college degree became the differentiator for gaining access to the middle class. With little downside financially, a college degree quickly became a requirement for competitively priced jobs (Wierschem & Mediavilla, 2018).

In the 1970s, however, federal government funding subsided as partnerships with universities, especially around military projects, lost favor for private industry (Dorn, 2017). The loss of funding led to increased tuition costs and an increased dependency on state provisions (McConnell, 2017). This increased financial dependency on state provisions and growing student tuition led many state lawmakers to require the adoption of performance funding (Kelderman, 2019). First implemented in 1979 in Tennessee, performance-based funding has waned in its popularity from being almost abandoned in the late 1980s (Snyder et al., 2018) to growing in popularity because of pressures of state budgets, whereby the mid-2000s, over 30 states had a funding model (Gándara & Rutherford, 2018). This increased interest, motivated beyond the impact of depleted state coffers, is due to increased interest in college completion rates (Kelderman, 2019).



### *The Focus on College Completion*

National attention had been growing around degree attainment rates in the U.S. (Gandara et al., 2017), specifically concerning our economic competitors in Canada, Japan, and the U.K. (Gándara & Rutherford, 2018). However, degree attainment became a primary national agenda item with the election of President Obama. After being elected, one of his first addresses stated that "By 2020, America will once again have the highest proportion of college graduates in the world" (Obama, 2009, para. 66). To achieve this goal, according to the U.S. Dept of Education, 60% of people aged between 25-34 would have to have completed at least an associate degree by 2020 (OECD, 2020). Politicians and employers alike agree as they face predicted workforce shortages of 55 million jobs in 2020 (Dougherty & Natow, 2015).

Scholars Partnering together, several initiatives accomplished the President's goal. The first began with establishing the American Graduation Initiative (AGI) that promised to offer \$12 billion to fund community colleges (Dougherty & Natow, 2015). The second initiative was increasing Pell Grant funds and the American Opportunity Tax Credit development, which helped families save money on and for college (Dougherty & Natow, 2015). Several private organizations also offered to take part, most prolific of the Bill and Melinda Gates Foundation, who, as of 2006, has spent over \$500 million directly on American Higher Education initiatives (Complete College America, 2021). However, their most influential work in this area is the funding of the Complete College America Initiative, whose mission is to be "a bold national advocate for dramatically increasing college completion rates and closing institutional performance gaps" (Complete College America, 2021, para. 6). Using this program, they have encouraged several states to



expand their performance-based or outcomes-based funding model, even suggesting states' focus on credit attainment instead of completing a degree to disrupt capital to colleges and universities (Hillman, 2016). As the second-largest funding source for a college after enrollment (Dougherty & Natow, 2015), this new level of accountability did not come without its critics.

### *Criticisms against Performance Funding*

Performance-based funding policies, while helping leaders of colleges and universities prioritize assisting students' complete school, several studies have shown that this type of funding has not been effective at increasing college completion (Dougherty & Natow, 2015). One reason for this is the complexity of these organizations (Cutler White, 2019); student success isn't simply a matter of focus on the classroom; it takes place from major selection to academic advising to housing to individual evaluations of student performance. Not every resource understands its role or who receives performance-based funding (Rios, 2019). Another reason is that many colleges do not have the resources, such as technology, staff, or ability to improve their completion rates (Umbricht et al., 2017). If those same under-resourced schools receive a decrease in funding, it further exacerbates their ability to meet the demand (Bell et al., 2018).

These institutions' efforts to meet the high requirements often involve detrimental outcomes to already underserved populations. Their efforts include increasing enrollment to provide a larger pool of potential graduates to complete credits (Rios, 2019) to reducing requirements for remedial education, forcing academically underqualified students through programs (Bell et al., 2018). The evolution of General Studies programs section discusses this reality in greater detail.



Suppose the colleges do not provide ample resources to enable students to complete school. In that case, the outcome will affect the job market, of which, even now, in 2020, only 48% of people have completed college (Sublett & Tovar, 2021). Researchers project that the workforce will face a shortage of over 85.2 million workers by 2030 based on current college attainment levels (Sublett & Tovar, 2021).

### Evolution of General Studies

The first incarnation of General Studies degrees, sometimes called interdisciplinary or multidisciplinary studies, began as early as 1917 (Erickson & Winburne, 1972). The degree allowed academically motivated students to focus on areas of interest, build their curriculum with faculty help, and investigate a core topic. Yet, by the 1980s, the degree shifted towards fulfilling the demands for a career-oriented curriculum (Cohen & Kisker, 2010) and offer flexibility to the nontraditional student whose work or family commitments prevented them from participating in the standard academy (Green et al., 2007). The demand for these degrees often came from adult learners, who emphasized the practical application of knowledge and acquisition of a degree solely on postindustrial economic pressures (Ntiri et al., 2004). However, as tensions mounted against universities to increase college completion rates, the flexibility of these degrees made them helpful to expand enrollment from nontraditional students to all students at risk of not completing their degrees (Busby, 2012). The degree once



reserved for a few select students has become the thirteenth most popular major in the United States (Busby, 2012).

### *The Flexible Degree*

Initially, the General Studies degree program, named differently across institutions, followed a similar set of principles identified by Erickson and Winburne (1972). The first principle must not be a free elective system, whereby students can take any course to fulfill their intellectual curiosity. Nor can it be a rigid major, whereby many require prescribed introductory courses or work within a distinct discipline (Erickson & Winburne, 1972). The second principle is that the degree must not have many institutional requirements (Erickson & Winburne, 1972). The third principle is an expectation of the required number of credit hours (Erickson & Winburne, 1972). The fourth principle is that the students within the major need greater attention from academic advisors, and the last principle is that students express academic superiority and motivation (Erickson & Winburne, 1972).

However, by the mid-1980s, these programs only aligned with the first three principles (Green et al., 2007). The commonality between these programs became a requirement for a specific number of credit hours and a focus on two or more disciplines (Green et al., 2007). These disciplines align with courses within a major family, i.e., business, economics, communication, kinesiology, etc. The General Studies degree programs also focused on concentrations that offered more elasticity in schedule, night and weekend courses, distance learning, and independent studies (Green et al., 2007). The demographics in these programs also changed from highly motivated students focused on expanding their education to students from more underrepresented groups



such as adults, ethnic and racial minorities, and military personnel looking for degree completion over knowledge acquisition (Hoyt & Allred, 2008).

As an example, according to the student catalog at The University of Texas at San Antonio, to complete a General Studies degree (called multidisciplinary studies), students must complete one hundred and twenty semester credit hours. Forty-two of these hours must be from the university's core curriculum requirements (UTSA, 2021). An additional six semester credit hours must come from foundation courses (one technology course and one communications course). The student must complete forty-eight hours split between three focus areas of their choosing. The student must also complete three semester hours of Intro to Multidisciplinary Studies, three Semester hours in a senior seminar, and eighteen hours of free electives (UTSA, 2021).

#### *Interdisciplinary, Multidisciplinary, or Something Else*

Using two or three concentrations may lead scholars to view General Studies like other interdisciplinary programs, programs that traverse deliberately across traditional disciplinary boundaries, according to Moran (2010). Interdisciplinary studies programs, such as women's studies or African American studies, allow broader inquiries and skills development to address problems (Lyon, 1992). For example, the issues affecting women are not solely sociological but are political, economic, and health-related. By intertwining disciplines, interdisciplinary students can engage in various interests and have exposure to different perspectives to similar problems (Davies & Devlin, 2010).

However, while allowing enrollment in courses from different disciplines, interdisciplinary studies programs have a core curriculum and only allow registration in outside domains aligned to specific topics (Davies & Devlin, 2010). This model, while



arguably does not build maturity in applying a particular discipline as one might in psychology or philosophy (Moran, 2010), it likens itself closer to a major in that specific faculty align directly to the program, specifically developed courses. The topical material expands the intellectual discourse (Davies & Devlin, 2010).

Yet even this definition of interdisciplinary studies researchers challenged. Researchers contend that the interdisciplinary studies model established at universities is multidisciplinary studies, not interdisciplinary. They argue that for interdisciplinary studies to exist, there needs to be a "common approach to common problems distinct from those of traditional disciplines" (Miller & McCartan, 1990, p. 2). Therefore, according to them, no interdisciplinary studies program is integrative; instead, it is multidisciplinary because it exposes students to multiple approaches used to solve problems. However, regardless of the debate between multidisciplinary interdisciplinary studies, these programs more closely align with the traditional General Studies degree defined by Erickson and Winburne (1972) than today's program. They are often more costly and require more advising than the approach seen in modern General Studies.

General Studies degrees are like other interdisciplinary studies programs as it allows students to take courses in different disciplines, but the courses do not have to align topically (McCray, 2011). A similar argument against interdisciplinary programs that students do not gain maturity in a single discipline (Moran, 2010) applies to General Studies, along with the risk of zero maturity within a specific topic or problem. Therefore, General Studies are neither multidisciplinary nor interdisciplinary. This lack of definitions has led researchers to question the quality of a General Studies degree and its impact on the quality of other degrees because of its undefined boundaries (Green et



al., 2007). Yet, based on the limited research, there is little to suggest that General Studies lack a definition, fares students differently in the job market or graduate school than graduates from interdisciplinary studies programs.

#### Vocational Demand, Missing Skills, and Curriculum Alignment

Industry voices a different opinion about graduates in the job market. Based on a study from the U.S. National Association of Colleges and Employers, college students' preparation is inadequate for the workforce on skills and application of those skills (Hora, 2018). Only 40% of skills students learn in school apply to industry or directly attribute to the student's success after graduation (Garner et al., 2019). The changes in the work landscape compound this skills inadequacy, with automation and the need for transferrable skills not part of the typical academic curriculum (Woodside, 2018).

Business leaders have harshly criticized Universities for this negative result, citing that they can no longer find talent with critical thinking, communication, or collaborations skills (Weise, 2019). These requested skills are vital to the modern workplace, one in which has replaced its manual processes with automation and shifting its needs to the "human" skillset (Weise et al., 2018). The inability to find and acquire relevant technical skills and translate and apply these human skills has created long-term negative impacts (Weise, 2019).

To fill this gap, modern employers of size, companies like Google or Facebook, choose to hire a varied group of new graduates, some from liberal arts and others from technical arenas, in hopes of cross-pollinating their skills (Gerstein & Friedman, 2016). The challenge they face is that it still requires a significant amount of time and investment to realize the potential of this model entirely. As indicated by a recent patent



from Google, leading companies look for alternative sources for job candidates (Foss et al., 2017).

Traditionally, vocational education or workforce development often has had a negative stigma among university leadership associated with lower-level students and relegated to community colleges. Until recently, universities and four-year colleges never believed they could or should provide vocational acumen or sometimes referred to as career and technical education (CTE; Ntiri et al., 2004). This type of education was the sole impetus for creating the community or junior college in the early 1900s (Whitcomb et al., 2016). One of these first colleges was the Joliet Junior College in Chicago, designed to reduce the vocational focus of the University of Chicago (Ntiri et al., 2004). However, due to changing opinions around workforce skills and the perception that universities are not filling this gap, community colleges have developed four-year applied degree programs to meet the demand (Whitcomb et al., 2016). This transition is not without criticism for both the community college on the quality of the education and universities for not offering these degrees directly (Ntiri et al., 2004). Some universities have considered alternative methods to maintain relevancy. These methods include using competency-based education models (Gerstein & Friedman, 2016) and certificate programs (Craven & DuHamel, 2003).

Unlike traditional education models that focus on fixed schedules and routines based on standardized processes, competency-based models focus on demonstrating learning that reflects students' skills, abilities, and knowledge (Halibas et al., 2020). This model attempts to translate what a student will learn in a program at a granular level and how it correlates into the modern workforce. It attempts this by measuring what is known



instead of traditional time requirements. This shift also allows individual students to focus on what they want and often at their own pace (Gerstein & Friedman, 2016).

The competency-based model, shifting from knowledge acquisition to application and evaluation, provides students with fundamental skills. These skills allow them to be dynamic and assess context, making them more valuable to modern employers (Gruppen et al., 2016). The model is not unique to other forms of traditional curriculum, but it is the primary emphasis on this specific feature.

Another advantage of competency-based education is the focus away from fixed time. In traditional academics, students' complete courses in specific units of time, most often in semesters. They are usually required to complete a minimum number of studies as well. In competency-based education, the students simply need to have their competencies assessed. Therefore, some may require a series of courses to attain it, while others may require only an assessment. This form of education allows students to enter and exit the program (Gruppen et al., 2016).

An argument provided by Gerstein and Friedman (2016) is that all degrees benefit from a holistic competency-based education. Leveraging several national associations, they identify four critical skills that need to be incorporated and assessed: communication, collaboration, creativity, and critical thinking. The authors provide examples of effective higher education programs (e.g., MIT, Augustana College) that have evolved their programs and seen higher than expected returns in the form of employment for their students upon graduation (Gerstein & Friedman, 2016).

Competency-based education isn't without its proponents that suggest there are still translation problems for students and employers. Although national associations



identify competencies, it can still be problematic when universities and their programs interpret the outcomes of their programs independently (Weise, 2019). This challenge may address why modern employers seek certifications and more easily tangible definitions (Foss et al., 2017) rather than relying on a traditional degree program.

Unlike competency-based education, which attempts to provide an overt assessment of skills learned regardless of occupational fit, certificate programs try to directly meet the demands of practitioners (Friedman & Friedman, 2018). There are two variants within the University structure, academic certificates and professional certificates (Craven & DuHamel, 2003). Academic certificates are developed by the university directly, often through requirements from local employers (Craven & DuHamel, 2003), whereby professional certificates are often developed directly by companies themselves or through third-party partnerships (Goldring, 2017).

The use of certificates can offer several advantages to students and future employers. The first is like competency-based education in that it offers levels of flexibility (Craven & DuHamel, 2003; Guthrie & Callahan, 2016). Most competency-based education approaches focus on degree completion, whereby certification programs, by their nature, are focused on the certificate itself. This focus can serve as simply a skills refresher or a method of earning academic credit towards a degree (Craven & DuHamel, 2003). Another benefit of certificates is that it is often hyper-focused (Goldring, 2017), degree programs tend to focus on holistic material, whereby a certificate may have a student pursue deeply into a single topic. This focus is attractive to modern business leaders by ensuring employees remain relevant. The third and most appealing benefit to employers and universities is that certificates can be self-taught. These courses are often



delivered virtually and reach more students than a classroom-focused teaching methodology (Goldring, 2017).

An area seen where certificate programs meet industry demands is in leadership competencies. As identified by Guthrie and Callahan (2016), leveraging a leadership certificate offered at Florida State, they argue that focus needs to be made on five competencies to make better citizens, leaders, and future employees: critical thinking, communication, cross-cultural understanding, ethical capacity, and civic engagement. They based these conclusions on surveys and the review of vocational literature on the skills employers find to be in most demand. When a student completes both a certificate and a degree, skills gaps are closed, improving industry.

Several Universities have considered ingesting industry certifications directly into their existing curriculum (Goldring, 2017). These certifications offered by companies like Google or Udemy provide industry validation to students. These certificates provide incentives to future employers and students who might otherwise not attempt them and instead rely solely on their degrees. In all cases, the certifications were technical, supplementing already existing STEM or business degrees (Goldring, 2017).

The certificate program quickly replaces vocational learning, a long-valued proposition within community colleges (Traver, 2016). Several employers can partner to offer their exact requirements and integrate them into this format (Bravenboer & Lester, 2016). This method is excellent at helping first-time graduates procure a position and meet employers' immediate demands. An argument against this approach is that students do not learn the critical skills that will help them maintain relevancy after the technical skill sets are no longer relevant (Gerstein & Friedman, 2016).



The most valuable aspect of a certificate program is skills adoption. Yet, as noted, skills adoption may be of finite value. More easily interpreted than traditional degrees by employers (Foss et al., 2017), recipients depend on continue earning certifications to remain relevant. This dependency, in part, is due to the technical nature of certifications, but also, if not paired with a degree, may not teach foundational competencies that expedite learning (Yu, 2019).

Universities have attempted to address this by advertising computer science, business, STEM programs, even focusing on individual certifications. None of these programs are proper workforce development. While offering skills outcomes, like a CTE program, there is no direct alignment to specific job tasks or functions (Ntiri et al., 2004). This alignment challenge has led many scholars to identify that this lack of translation between the job market and course outcomes negatively affects degree satisfaction and interpretation of a traditional bachelor's degree value (Ntiri et al., 2004).

For universities, especially tier-one institutions, this degree alignment to industry demands may not be their most relevant challenge. Historically, as suggested above, funding sources fueled universities' identity. And while, arguably, degree satisfaction matters if it harms overall enrollment, it has not yet, even with students in recent years opting for sabbaticals (Whitcomb et al., 2016), significant effect on school funding.

Several studies have attempted to understand whether a specific academic curriculum aligns with industry demands (Föll & Thiesse, 2017; Kendricks et al., 2019; Kitto et al., 2020). The most common approach is to pull a sample of job descriptions and review the required qualifications qualitatively (Woolridge & Parks, 2016). These studies focus on specialized degree programs such as Information Systems or Computer Science.



These researchers benefit from isolating job descriptions based on degree requests and making it easier to understand alignment with the curriculum's skills offerings. When applying to General Studies, the challenge with this method is that there is no uniform set of courses taken, nor do employers request a General Studies degree. Therefore, all jobs not explicitly requesting a specific degree are available, and the ability to review becomes too many in methods used in these previous studies. However, it does not negate the need to understand the specific and general skills required for competitive jobs.

### Natural Language Processing

This research effort uses modern algorithms to identify, categorize, and define words used within course syllabi. Therefore, as these are potentially unfamiliar techniques to the human capital development discipline, the following sections outline these techniques and their background in detail.

#### *Tokenization*

Natural Language Processing (NLP) explores the usage of computers to analyze the natural languages of human beings (Chowdhury, 2003). Using the NLP's primary aims, the scholars gather knowledge on human beings' understanding and use languages to develop computers. To perform specific tasks, they can analyze and understand natural human language (Chowdhury, 2003). NLP encompasses a wide range of disciplines, including linguistics, information science, and mathematics. Programmers use NLP for text analysis and summarization in natural language, multilingual user interfaces, and cross-language information retrieval (Chowdhury, 2003). It also applies to speech recognition, artificial intelligence, and expert systems.



Tokenization is a task that involves chopping the defined programs into units called tokens. Programmers refer to units as tokens when they group programs as a semantic unit for text processing (Rai & Borah, 2021). The tokenization process removes characters, such as the remaining spelling errors not released in the previous steps. In this step, the texts are broken into single words called tokens, used to collect observations in the classifier related to the respective text to represent the data (Rai & Borah, 2021). Each of the words connects to the reader. This thesis considers degree courses a column consisting of defined program units (Rai & Borah, 2021). We now take the program units, break them into pieces, and consider them as tokens. These pieces of tokens represent their respective job qualifications.

According to Aso et al. (2020), one must have professional knowledge of the targeted text language to synthesize the prosody, statistical parametric text-to-speech (TTS). Consequently, languages limited to only rich-resource languages are suitable for TTS synthesis (Aso et al., 2020). Acoustic model-based sub-word tokenization does not need supervision to extract prosodic contents to achieve TTS synthesis without understanding the language. Both methods guarantee prediction accuracy without any prior knowledge of the language. The practices, in effect, chop tokenized texts into sub-terms that are very similar to language-specific units. They also assure for synthesis speech qualities and the resultant works are empirically stable.

Scholars are moving from conventional tokenization methods to contemporary ones (Aso et al., 2020). For instance, when used, the bidirectional long short-term memory (BLSTM)-based context extraction requires intensive training after acoustic model-based sub-word tokenization (Aso et al., 2020). While on the other hand, the



proposed contemporary method, the deep neural networks (DNN), is more straightforward and does not require any training after sub-word tokenization. There are several presented advantages of modern techniques. For example, the DNNs can simplify the extraction process compared to the BLSTM (Aso et al., 2020).

Moreover, the DNNs ensure that the perpetual speech synthesized is of high quality. Finally, the DNNs compared to the BLSTMs are significantly more robust hyperparameter situations (Aso et al., 2020). Therefore, today, the tokenization process that combines the new models rather than the conventional models (end-to-end methods) will provide a much closer desirable performance.

Xiao et al. (2020) have also proposed new model-based tokenization suitable for large-volume audios; the new models integrate phone- and prosody-related linguistic properties. That combines with a 1-D convolution bank + highway network + bidirectional gated recurrent units (CBHG) -like encoder (LinFeat); the second one is the bidirectional encoder representations from Transformers (BERT) -conditional random field (CRF), available in the BertBr package, model that enriches inputs with predicted breaks. The third is the standard BERT model, available in the BertEmb package, which uses up-sampled character embedding to concatenate phoneme inputs (Xiao et al., 2020). Thus, the three new models improve the quality of the text to speech in the targeted domain, and the most effective one is the BertEmb package (Xiao et al., 2020).

However, Zhou et al. (2019) have opposed the idea of using the multilingual BERT. They claim that the model, rather than using a common, shared interlingua space, reveals the model partition representations (Zhou et al., 2019). Therefore, the model does not abstract semantic languages progressively as it disregards languages. Further, they



dispute the BERT model that provides a more substantial bias towards word-level tokenization and such characters (Zhou et al., 2019). Thus, it isn't objective in discovering the evolutionary and linguistic relationship between languages. Finally, the BERT model does not embed languages together into a shared space (Zhou et al., 2019). Conversely, they praise the idea of unsupervised usage of the BERT model enabling the use of several languages.

### *Parts of Speech Tagging*

Employers can consider focusing on talent or General Studies curriculum to identify competitiveness for job acquisition in the contemporary competitive job environment. In this process, parts of speech tagging (POST) have emerged as a method deployed in competitive job acquisition. POST consists of reading sentences and identifying the words that serve as verbs, nouns, pronouns, and adverbs. Malhotra and Godayal (2018) report that identifying POST is increasingly complex than mapping words to their POST. The POS tagging is never generic. In most cases, employers should be aware that a single word can have various POS tags in varied sentences depending on the different contexts of the word. It follows why having a generic mapping for the POS tags is an impossibility.

Additionally, the POS tags make the automation of text processing tools possible to consider the POS tag in every word. In this way, employers could use linguistic criteria in combination with statistics in determining recruitment competitiveness. Two forms of POS tagging can be used, manual annotation and automatic POS annotation (Malhotra & Godayal, 2018). Under manual annotation, human annotators tag the data. However, this technique is highly labor-intensive and rarely used in the modern business and



recruitment environment. Significantly, attaining quality and accurate results using this approach requires the employer to use more than one annotator. The process often receives support from specialized annotation software that checks for inconsistencies between annotators, but it remains laborious (Malhotra & Godayal, 2018). Suppose the software recognizes a token or word with a varying POS tag from every annotator. In that case, the human annotators must identify a suitable solution to annotating the word or expand the tagset to ensure it accommodates the new condition (Malhotra & Godayal, 2018).

The second and more common POST approach is automatic annotation (Sketch Engine, 2018). The new data or word setting exists in large volumes, making automatic annotation the best available tagging solution. Here, employers would employ the tagging tool with up to 98% accuracy and limited mistakes to high-interest phenomena, including occasional use, interjections, and misspelling (Sketch Engine, 2018). Additionally, automation takes care of the ambiguity problem. Sketch Engine (2018) reports that modern technological tools can accurately annotate most of the corpus, and even if mistakes might happen, they rarely lead to challenges during the corpus's use (Sketch Engine, 2018).

Different languages require varied taggers to help detect accurate skills compatible with the job requirements and roles. These taggers deploy various algorithms, techniques, configurations, and programming languages (Sketch Engine, 2018). Also, programmers have trained and adapted tools to annotate two or more languages. These tags are more convenient, especially in an increasingly diverse employment environment.



It is possible to find companies or employers attempting to understand candidates' skills from different cultural backgrounds.

According to Jurafsky and Martin (2020), the POST algorithms have increased accuracy. The author report that a study found a 97% accuracy level in 15 different languages from the Universal Dependency (U.D.) Treebank (Jurafsky & Martin, 2022). The percentage is mainly associated with human performance, making it a suitable tool for processing language to identify skills in a syllabus. While some words are simple to disambiguate, others are highly complex, with their various tags appearing unlikely. While ambiguous words only account for around 14-15% of the vocabulary, they are increasingly common and used in running texts (Jurafsky & Martin, 2022). In this article, the author also presents the concept of named entity tagging. This technique is usually the essential first phase in most of the natural language understanding jobs and tasks. For instance, the sentiment analysis can involve inquiring about the employment candidates feeling toward the company or a client's emotions toward a specific product. During question answering sessions or in connecting a text to structured information sources, entities are often helpful. The named entity tagging is at the core of language understanding tasks.

In his article, Manning (2011) examines the ability to improve the POST performance accuracy from 97.3% to 100% token accuracy. The author argues that it is possible to increase tagging performance by examining various improvements to the Stanford POST. Despite introducing machine learning or improved characteristics within the discriminative sequence classifier, errors remain challenging to reduce. Semi-supervised understanding also presents opportunities for advanced gain (Manning, 2011).



However, the article maintains that the most significant opportunity for growth is improving the taxonomic linguistic resource perspective for training the tagger. It is also fundamental to consider the word's status because some terms cannot adequately capture by assigning the words to a smaller number of groups. In such situations, deployed conversations enhance tagging consistency.

Additionally, Pham (2020) focuses on the rule-based form of POST. The author highlights the POST exists in different categories, including stochastic and rule-based approaches. Most natural language processing applications deploy the stochastic process in determining the parts of speech. The favor of stochastic methods over traditional rule-based systems originates from the simplicity and ease accompanying the necessary stochastic automated attainment. Further, the rule-based approaches are challenging to implement and are never strong. The author uses Yacc and Lex to develop rule-based POST for English (Pham, 2020). According to the author, the stochastic taggers have an increased accuracy range depending on the pure syntactic input analysis. They mainly rely on the Hidden Markov Model (HMM) that captures the contextual and linguistic data. From the research, the author concludes that rule-based tagging is a highly efficient and faster tagging technique than the stochastic approach. It is significantly more rapid in the execution of the tagging process using linguistic rules. Nevertheless, the method has a weakness in eliminating ambiguities and tagging unknown words. A rules-based approach can enhance the ability to tag unfamiliar words in rich morphology languages.

It is becoming increasingly significant within the modern workplace environment to process social media language to detect essential skills in a syllabus. According to Neunerdt et al. (2013), social media text has distinct nature from the standardized test,



which calls for the adaption of natural language processing techniques to achieve reliable Processing. The authors introduce a new social media text corpus and examine the various POS taggers that have undergone retraining in the corpus (Neunerdt et al., 2013). For social media texts, it is possible to improve the performances of state-of-the-art POS taggers. However, to achieve the improvement, it is fundamental to consider social media text data training. With the addition of the in-domain training data, there would be a more than 5% increase in performance accuracy (Neunerdt et al., 2013). For instance, the cross-validation in the TreeTagger results in the maximum average per word accuracy of around 93.72% (Neunerdt et al., 2013). This tagger leads to better results than the Stanford, Trigrams'n'Tags, and SupportVectorMachineTool for social media texts. This study illustrates the significance of selecting an appropriate tagger for language processing to produce highly accurate results. Accuracy is critical in identifying the relevant syllabus skills.

### *Phrase Matching*

A phrase is a building block of a track consisting of different bars. A phrase describes the word group that shows a concept and is a unit in a sentence in standard terms (Yin & Schütze, 2017). Phrases exist in eight forms: verb, noun, infinitive, appositive, gerund, absolute, participial, and preposition. While noun phrases usually entail a noun and its modifiers, a verb phrase consists of a verb and modifiers. Additionally, the gerund phrase is a noun phrase that has begins with a gerund, and the



infinitive phrase often entails a noun that starts with an infinitive verb (Yin & Schütze, 2017).

Further, the appositive phrase primarily restates and defines the noun and consists of one or more terms (Yin & Schütze, 2017). The other form of a phrase is the participial phrase that begins with a present or past participle. The prepositional phrases often start with a preposition and serve as an adjective, noun, or verb (Yin & Schütze, 2017). Finally, the absolute phrase consists of a subject but not an action verb, meaning that it cannot stand on its own as a complete sentence (Yin & Schütze, 2017). Understanding these forms of phrases is usually essential in natural language processing to identify skills within a syllabus. Gaining this detailed understanding of the concepts is also fundamental in understanding the definition and application of the phrase matching methodology.

Now, it is fundamental to expose the reader to the accurate definition of phrase matching. A language-dependent process where advanced search engines deploy different word sets as a cohesive unit when scanning within the search index to reveal the most relevant documents (Swifttype, 2021). For instance, an advanced search engine can contain the phrase "Dell computer" as a solitary query and only return those documents that feature the whole phrase. On the contrary, a simpler search engine will primarily return the documents with the term "Dell" and all those featuring the word "computer." For instance, Swifttype deploys phrase matching to return different findings for several search queries (Swifttype, 2021). Another phrase matching tool can be the Google search engine. Here, an individual can key in any phrase, and Google will return the most viable documents that match the search. Besides, it is fundamental to note that phrase matching plays a significant role when looking for voluminous content or when a researcher is



unsure of the most accurate search to use in a particular situation. In this perspective, one will type a phrase in a search engine and get different matching documents that they can use for their specific research or any other purpose. Phrase matching can also be essential when the search engine contains a built-in bias towards new information, such as recently published articles (Swifttype, 2021). For instance, looking for information in a newly published news article or website requires using emerging phrases that would lead to more accurate results from the search engine. However, the date bias usually has less significance and would return the older results first if they relate to the searched phrase (Swifttype, 2021). In essence, using phrase matching makes the results appear with the accurate wording as the keyword but with minor differences. It is also the same case when one adds more wordings before and after the keyword. The wording order is fundamental in the match, implying that it would not appear to an individual who adds any unrelated word at the middle of the keyword.

Phrase matching is another approach used in natural language processing. It is a robust tool for quick retrieval of results. Patterson et al. (2008) investigate document retrieval through proximity-based phrase searching. Phrase matching is an objective approach. Phrase matching can be used alongside phrase searching to identify and locate documents with complete text that they seek. Here, search engines, including Google, can be used to enable users to provide accurate phrases. The search needs to return the same document as it exists except for the addition and exemption of punctuation. Significantly, the authors report that unique phrases lead to more accurate the document (Patterson et al., 2008). This approach is highly effective and efficient in locating needed text since it



is highly likely to result in more desired results. This accuracy makes this method better in the natural language processing and subsequent identification of skills in a syllabus.

Other researchers, including Nguyen-Son et al. (2015), investigate phrase matching in detecting paraphrasing. The authors begin by highlighting that paraphrase detection has tremendous applications within natural language processing. To identify the presence of plagiarism, one can effectively deploy phrase matching to detect paraphrased phrases and words quickly. In their article, the authors deploy a heuristic algorithm in phrase matching. It is significant to note that the heuristic algorithm identifies the most extended duplicate phrase within every iteration. Like in the previous article, the authors reveal that phrase matching has increased accuracy levels, making it an appropriate and suitable method in natural language processing. In this study, similarity matching (SimMat) has improved efficiency in detecting paraphrases. Different translation methods, when combined, yield more significant results (Nguyen-Son et al., 2015). Combining the SimMat metrics with the re-implemented metrics delivers higher accuracy than when used alone. It illustrates that phrase matching is a viable and valuable technique in identifying paraphrasing.

Phrase matching is also fundamental within the medical field. In the modern healthcare environment, health facilities rely on the information system to store and retrieve medical information for diagnosis and treatment purposes. According to Liu et al. (2018), disease phrase matching facilitates reliable and efficient medical data processing. The disease phrase matching tasks are significant in this aspect. The method involves identifying whether they can interpret one another (Liu et al., 2018). Like any other professional environment, terminologies are increasingly complicated, overlapping, and



have similar syntax structures in the medical setting. These characteristics make medical terminologies unreliable for phrase matching. In identifying information, people often deploy supervised sentence or phrase matching, primarily because of the variability in human languages. This approach has also become significant in the entailment of text, web search, entity linking, and disease interference (Liu et al., 2018). It is essential to note that the modern developments in deep learning have drawn people's attention towards different tasks. The study shows that phrase matching is increasingly becoming necessary as a natural language processing method.

#### *Latent Dirichlet Allocation and Amazon Comprehend*

The latent Dirichlet allocation (LDA) is a common form of topic modeling. Topic modeling refers to the method deployed in unsupervised document classification (Kulshrestha, 2019). Under LDA, every document comprises different words, where each topic contains other words. This method focuses on locating or finding a document belonging to and based on the terms within its domain when presented with various documents containing varied word lists. Consider having five documents with different word lists.

Doc1: word2, word4, word5, word43, word13, word73, word88...

Doc2: word6, word73, word31, word92, word66, word11...

Doc3: word77, word83, word3, word91, word19, word20...

Doc4: word19, word53, word33, word61, word47, word18...

Doc5: word12, word41, word94, word22, word1, word17....

From the documents, it is significant to identify the words that exist in varied topics



The rows reflect a varied topic in the table, while every column contains a different term within the corpus. The cells have the probability of the word belonging to the topic. There are various assumptions to consider when using the LDA. For instance, the user should assume that every document is a collection of different words, where the model disregards word order and their grammatical responsibility. Terms like 'a', 'of,' and 'the,' never carry any important information for the topics of interest and can be eradicated from the documents during the preprocessing stage (Kulshrestha, 2019). Also, one should know the number of topics beforehand while all the topic segments are accurate, except for the current interest.

Kulshrestha (2019) provides sufficient information regarding the function of the natural language processing method. The LDA operates in two components, words within a document and words that belong to a topic that should be calculated (Kulshrestha, 2019). This approach primarily represents documents as a combination of different topics. Likewise, a subject is a combination of other words. Essentially, the LDA also deploys an algorithm to help undertake natural language processing to identify appropriate skills that would determine the competitiveness of a job.

The article by Sharma (2020) provides a basis for understanding the LDA. According to the author, LDA assumes the documents contain words that assist in determining the topics while also mapping the documents to a topic list by assigning every word within the document to distinct topics (Sharma, 2020). It is significant to note that the assignment occurring here happens based on conditional probability approximations.



It is fundamental to remember that the LDA ignores the word occurrence order and the syntactic data while treating documents as word collection. After estimating the probabilities, the user can find the word collection representing a specific topic by selecting the top possibilities of the terms or setting a threshold for the likelihood and selecting the terms with greater probabilities only. The author reports that the LDA algorithm assumes that a statistical generative procedure produced every document. That implies that each document is a combination of topics while every topic is a mixture of terms. Assume that a document consists of three different topics, including facilities, feedback, and tourism.

On the other hand, every topic is a combination of diverse word collections. In the document generation process, the initial step is to choose a subject from the distribution of documents and topics. After the topic selection, the user then picks a fitting word from the distribution of topics and terms.

The LDA performs the opposite of the process of generation during the topic identification stage from the documents. It is essential to remember that LDA commences with the random topic assignment to every word and the iterative improvement of the topic assigned to terms via the Gibbs sampling method (Sharma, 2020).

The other significant point to note in the LDA approach is its hyperparameters. Sharma (2020) reports that this method has three different hyperparameters, including the topic-word density, the number of topics considered, and the document-topic density factor (Sharma, 2020). In the document-topic density, there is significant control of the topic number expected within the document. A low value is used when the topic number is fewer in the desired mix, while a high value represents that an individual would expect



the document to possess an increased number of topics within the mix. Besides, the topic-word density hyperparameter controls and manages the word distribution per a specific topic. A lower value would indicate that the subject will contain fewer words and increased word count when the value is higher. The final hyperparameter specifies the number of topics expected within the document corpus.

In their research, Blei et al. (2003) investigated LDA. The authors define LDA as the generative probabilistic model deployed in allocating discrete data like the text corpora. According to Blei et al. (2003), this approach is a three-level hierarchy Bayesian model. Every item of a group is a finite combination over an existing topic set of probabilities. In the text modeling context, the topic probabilities provide a clear document representation. The authors highlight significant elements in the LDA, including its influence on interchangeability. Here, it is essential to note that a finite sequence consisting of random variables is infinitely exchangeable if each finite subsequence is exchangeable (Blei et al., 2003). Based on De Finetti's representation theorem, the joint distribution of the finitely exchangeable sequence of the random variables exist as if the random parameter comes from the various distribution" (Blei et al., 2003). According to the LDA, word generation is essential to occur by topics, and the topics have infinite exchangeability in a document.

The article concludes by highlighting the flexibility existing in the LDA model. The approach depends on the simple interchangeability assumption for both words and topics within a specific document. As a result, the authors maintain that the LDA is a straightforward use of De Finetti's theorem of representation (Blei et al., 2003). The article also presents the use of "any of a large suite of approximate inference algorithms



can be used for inference and parameter estimation within the LDA framework" (Blei et al., 2003, p. 1014). Some techniques for accomplishing the inference to yield a faster algorithm to lead to an excellent comparative performance include the higher-order variational methods, Laplace approximation, and Monte Carlo approaches (Blei et al., 2003). From the research, the authors report that the LDA is a simple model and is highly illustrative. The primary benefits of generating and using the LDA model include its extensibility and modularity. It is possible to embed the LDA into an increasingly complex model. This property makes this technique have high quality and flexibility, allowing the user to manipulate information and obtain valuable results. LDA also has various possible extensions, including continuous data and non-multinomial information (Blei et al., 2003). Developing a constant variant is straightforward, where the Gaussian observables are utilized rather than the multinomial. Permitting the combinations of the Dirichlet distributions in the place of a solitary Dirichlet of the LDA is another significant extension of the model. This extension allows the improved arrangement within the patent topic space and document clustering that is varied from the clustering achieved through shared topics (Blei et al., 2003). The final extension that is possible entails elaboration of the topic variables. The article shows that the LDA is an efficient and effective language processing tool that can detect accurate skills in a syllabus.

Moreover, different scholars have also investigated the use of Amazon Comprehend in natural language processing. The Amazon Comprehend is a new service introduced in 2017. This model is now available across the world and can process different languages. The Amazon Comprehend can detect language, entity, extract key phrases, analyze sentiment, and accomplish topic modeling on a massive document



collection (Simon, 2017). These features make the model increasingly applicable in most settings and varied languages. Simon (2017) reports that Amazon Comprehend can detect around 100 languages, making it highly relevant and valuable in different language environments. Within Amazon Comprehend exists the topic modeling concept that allows the user to automatically build a topic list from an expanded document collection and group documents based on the selected topics.

Mishra also investigates the deployment of the Amazon Comprehend in natural language processing. The author describes this model as an entirely managed service offering elaborate access to deep-learning-dependent natural language processing. It also contains a topic modeling engine that developers can deploy within their different projects to analyze the text document's content and implement features, including topic-oriented classification, customer sentiment examination, and content-based searches (Mishra, 2019). Amazon Comprehend uses pre-training machine deep learning models. Amazon also regularly maintains and improves through inputs from different real-world sources like Amazon product reviews.

### Summary

Unlike other assets, the value of an investment in human capital is difficult to assess, none more so than the value of an investment in education. The challenge is even more significant when the investment is not specific, as seen with the General Studies degree. As seen in the literature, the value of a human capital investment comes from its ability to signal the investor's productive ability to employers. For entry-level employers, this often comes in the form of a college major. Yet, that college major is simply an assumed obfuscation for a set of skills acquired by that candidate. General Studies degree



holders may have developed several different skills based on their areas of concentration and face a labor market that does not recognize this college major. The other challenge identified is that General Studies degree students are not from the traditional, highly motivated student groups but from groups traditionally underrepresented, presenting an additional bias in their hire ability based on their degree alone.

These challenges, however, might produce value to these students as the literature suggests employers prefer skills over traditional college degrees. If students can increase their acquired skills, specifically those in demand in industry, upon graduation and signal to employers, this is the case; the degree might hold more value than traditional degrees. The literature review outlined the process of leveraging natural language processing. NLP is limited in its use in the discipline of human capital development. These techniques are revolutionary in detecting and assigning definitions to words and topics used within documents. As suggested by the literature, if used properly, these techniques may assist in defining those skills for courses available to the General Studies degree candidate and setting the foundation for this research effort. The following chapter, Chapter III, presents the methodology for investigating the objectives addressed in Chapter I, and building upon the foundation provided in Chapter II.



## CHAPTER III – METHODOLOGY

The goal of this research is to understand the effectiveness of General Studies programs in meeting industry demands. As presented in previous chapters, understanding how the curriculum can be aligned to ensure students are more competitive in the economic market. The literature review identifies a significant effort in addressing college curriculum within specific programs and a drive to understand the value of a General Studies degree. Still, no study has applied both focuses. This study is an effort to examine how these two lenses, when joined together, can strengthen General Studies.

This chapter outlines, and in detail, describes the methodology for this research. This chapter includes the submitted research design, population and sample, the data collection methods, procedures, and analysis efforts.

### Research Design

This study examined the relationship between variables in a quantitative design. Specifically, the impact of the skills students acquired from their curriculum choices (IV) and their ability to earn a competitive entry-level job denoted by salary (DVs) after graduation. The design of this study, outlined in Figure 6, required several steps. The first step required establishing a baseline of industry talent demands or skills. These skills came from external job postings gathered by the 3<sup>rd</sup> party vendor, Burning Glass. The skills were rank ordered by number of postings and isolated to the top 25 most identified skills.

Step two required external taggers leveraging these identified skills as tags, or markers, for outcomes extracted from syllabi. The syllabi outcomes used for tagging came from a subset of total syllabi for the past three years from The University of Texas



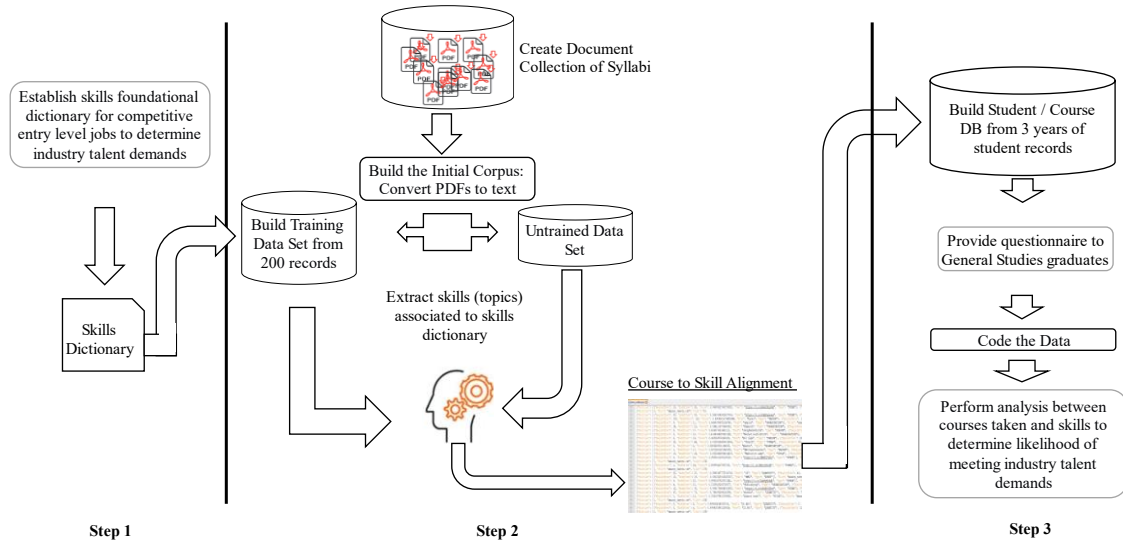
at San Antonio (UTSA). The researcher created the total body, or corpus, of syllabi by downloading them all systematically from <http://bluebook.utsa.edu>, converting them from Adobe PDF to raw text, and extracting only the outcome/objective sections of each syllabus. The tagged syllabi served as the training set for the natural language processing model. The model, leveraging the trained data, assigned skills to untagged syllabi creating a final dataset, whereby all syllabi have skills attached.

The final step utilized the fully tagged syllabi and associated skills to associate the complete list of courses to each graduate from the UTSA MDS program based on data from university institutional research. An additional part of step 3 captured salary information, acquired through a survey sent to graduates. The list of skills and the captured salary provided all the outcomes needed to address the research objectives.



**Figure 2**

*Research Design*



*Note.* The rounded squares represent a process, the document missing a corner is text output, the silo shape represents a database. The head with gears is the symbol for Amazon Comprehend. The arrows are all connected to show the dependency of steps.

The study's design is consistent with other research such as assessing course skills alignment with natural language processing (Gottipati et al., 2021; Gromov et al., 2019; Rockwell, 2019). A regression to evaluate statistical significance based on scored calculations when determining alignment is congruent with other studies (Kaira, 2011). The choice for using a survey for direct student data collection aligns with other curriculum alignment studies (Föll & Thiesse, 2017; Whitcomb et al., 2016).

### Population

There are two populations in this research study. The primary participants were the graduates from the Multidisciplinary Studies (MDS) program, the General Studies offering, at the University of Texas at San Antonio (UTSA) from the last three years. A total population of 467 MDS graduates to potentially take the survey (B. Cordeau,



personal communication, September 24, 2021). The second population for analysis was the syllabi for all undergraduate courses. The syllabi must be an undergraduate course, offered at least once a year since 2018. The course must not require special permissions from an instructor to enroll, i.e., independent studies. The assumption is that a course-level outcome will not change regardless of the professor or section. The syllabi will be downloaded manually from the <http://bluebook.utsa.edu> and stored locally.

### Sampling and Sampling Procedures

Sampling is the process to select a predetermined number of observations from a population to estimate characteristics for a larger population (Fink, 1995). The sampling strategy consisted of all graduates from the last 3 years of the MDS program solicited to take the survey. The probability of return of the survey drove the purpose for the large population. A survey is an essential step in the process of understanding actual results (Fink, 1995). Meeting minimum statistical power ensures that any probability achieved through statistical tests will detect an effect if one is present. The researcher recognizes that ideally, yet improbable that the receipt of all surveys would occur, assigning a minimum sample size to achieve a 95% confidence level is essential. This survey requires a minimum return rate of 212 surveys to ensure an acceptable margin of error of 5%. An incentive of two online gift cards of \$50 will be used to increase participation. Gift card recipients will be determined through a drawing overseen by the dissertation committee chair. A request to the institutional research department will provide the students' email addresses.



## Informed Consent and IRB

Written informed consent was provided with a solicitation email and prominently displayed on the LimeSurvey (LS) survey front page accessed through the embedded web link within the email. The informed consent provided a detailed description of the study and the assurance of anonymity. Any information collected through the survey was collected anonymously with no Personal Identifiable Information (PII) since LS does not capture nor return PII.

Permission was granted to conduct this study at the University by the department's chair to perform this study pending IRB study, as noted in Appendix A. This study required approval from the Institutional Review Board (IRB) approval of the University of Southern Mississippi which was granted and presented in Appendix B. Approval ensured that the population and the questions asked aligned with the outcomes associated with this study and the participants were protected.

## Instruments and Operationalization of Constructs

This study involved capturing data from the university website and data correlations made using data provided by the university's institutional research department. Validate the skills acquired by MDS students and the value of those skills as it relates to salary guided the survey's purpose. A survey questionnaire allowed for the most cost-effective and timely way to capture high-level responses (Fink, 1995).

The survey consisted of 16 questions, the complete survey available in Appendix C. The first questions established when the student graduated, when they acquired employment (whether pre-or post-graduation and how long after graduation) and the industry they joined. The second set of questions identify the participant's current salary



range, their perceived potential salary growth, and the organization's size. The last group of questions inquired about the skills the student feels were the most influential in getting their position.

#### *Demographic Questions*

As suggested by Hoyt and Allred (2008) and in support of RO1, the demographic questions assessed the age of survey participant, their racial and ethnic background, and whether they were a continuing education student or a traditional student before graduation. The survey requested the participants provide what semester and year they graduated from the program to understand any shifts in student outcomes.

#### *Employment Acquisition Questions*

The survey requests the participants' current employment status, as it will negate any further responses in the survey if they are unemployed. As was suggested in previous studies (Bird & Smith, 2005; Ntiri et al., 2004), understanding what industry a student works in was important, as technical fields offer a more significant competitive advantage to entry-level workforces. The survey used the North American Industry Classification System defined by the United States Census, accounting for 20 high-level industries (United States Census, 2021). The survey proceeds to request the participant to identify if their employment in this industry was before joining MDS or before graduation. The purpose of these questions was to reduce potential Type I errors, whereby spurious correlation is attributed to the MDS degree when it is not associated—reinforced further, with inquiries related to how long after graduation employment occurred.



### *Job Inquiry Questions*

The next set of questions explicitly relate to the job the participant acquired. The questions requested current job title and expected salary (in their opinion) for someone with > 8 years' experience in that title. Eight years derived from Lazear (2009), which ascertained that expertise in a specific job occurs after eight years. This question helped to confirm the ideal salary for that role, adding to the correlation if the entry-level position acquired is competitive.

The next set of question related to the future and the participant's future employment interest. The questions inquired about the participant's desired job title and the pay range for that selected job title. This set of questions, derived from previous studies, related to curriculum alignment, asking participants if the degree prepared them for future roles (Steele et al., 2020). The final set of questions asked if they were interested in this role before their enrollment in MDS or after acquiring a job. This question set attempted to understand if MDS was an influencer or the participant was influenced by outside experience, aligning to other curriculum-based studies (Goldring, 2017; Whitcomb et al., 2016; Yudiono et al., 2018).

### *Summary*

The questions used in the survey derived from several studies for the development. Table 1 presents these questions along with their corresponding sources. The survey responses design aids in potential future investigation. The length of the survey aligns with a best practice outlined in Fink (1995), survey questions and responses ought to begin with broad, general questions and progress to specific ones, especially in



response to sensitive questions. Utilizing questions created for previous studies and aligning it to best practices, this survey answered the needs for the research objectives.

**Table 1**

*Survey Questions and Sources*

Question	Source
1. To which gender do you most identify?	Hoyt & Allred (2008)
2. How would you specify your ethnicity?	Hoyt & Allred (2008)
3. When you enrolled in Multidisciplinary Studies (MDS), were you a traditional student (started right from HS and stayed all 4 years) or were you continuing education (did not start after high school or started and stopped in your degree attainment)?	Hoyt & Allred (2008)
4. What semester and year did you graduate from the program?	
5. Are you currently employed?	Bird & Smith (2005)
6. What industry do you work in?	Bird & Smith (2005)
7. Did you have the job in your current industry before enrolling in MDS program?	Bird & Smith (2005)
8. If not, did you have a job in your current industry before graduating from the MDS program?	Bird & Smith (2005)
9. If not, how long after graduation did you acquire a position in your current industry?	Bird & Smith (2005)
10. What is your current job title?	
11. What is your current annual gross salary?	
12. At your organization, what is the current pay range for a person with >8 years' experience with the same title as yours?	Lazear (2009)



**Table 2** (continued)

Question	Source
13. What is your desired job title within the next 8 years at your current organizations?	Lazear (2009)
14. At your organization, what is the current pay range for a person with >8 years' experience with this title?	Lazear (2009)
15. When did you determine you wanted to work in this industry?	Steele et al. (2020)
16. Do you want to remain in this industry for the next 5 years?	Steele et al. (2020)

### Data Collection

In most quantitative studies, data collection represented the bulk of the work to prepare for analysis. This research effort is no exception, and desired a triangulation of results, data collection required multiple phases. Several data collection methods used third-party data elements to establish baselines and perform complex algorithmic processing, followed by a survey to past graduates.

#### *Data Collection Phase I – Entry-Level Skills*

Initially, this research intended to leverage a content and quantification analysis of online job postings. Collection of the data and its investigation included a series of steps: acquisition of postings, filtration, coding, and ranking, using an API to Indeed.com. However, a vendor, Burning Glass (BG), had already performed this analysis and provided the data to subscribers. BG houses this data in their "Labor Market Analytics Warehouse." Using greater than forty thousand online sources, this service scrapes and codes job posting information into a series of different variables. These variables include



skills information, experience, pay, occupation title, industry, etc. (Eggenberger et al., 2018).

Using the definition that a competitive entry-level position begins at \$35,000 (NACE, 2021), the researcher collected skills along with the job titles and corresponding salaries. BG uses generic titles congruent with industry and not isolated by the employer. This data served as the framework to address RO2: determine the skills employers seek for competitive entry-level positions that do not require a specific college major.

#### *Data Collection Phase II – Course Outcomes*

This phase required significant effort to perform due to the unique course outcomes offered the university. This phase had two parts to implement. The first phase attempted to understand the course outcomes associated with every course and convert these into relevant skills. This data collection process required several steps to complete. The first step gathered all the course outcomes.

All course syllabi include student learning outcomes (SLOs). Upon completing the course, these outcomes describe what the student will demonstrate in terms of specific and measurable knowledge, values, and skills. Retrieval of these SLOs requires a web scrapper to download all syllabi in Adobe PDF format from the university's Bluebook system (<https://bluebook.utsa.edu>). The Bluebook system is a website designed to house all course syllabus in response to Texas House Bill 2504. This bill requires all public universities to make course data available to the public (Wilhelm & Vaaler, 2018). The university stores three years' worth of syllabi in this warehouse.

Extracting the necessary text for analysis by models required the use of python to read the PDFs and extract the text. The first step imported each PDF into python using



the PyPDF2 1.26.0 package. This package extracted text directly from the syllabi and split the document into pages and sections (Inc., n.d.). The next step used a series of code to identify the word outcomes/objectives and extract the next 15 lines of the text. This code looped through all the documents extracting the text. The text for each document becomes a unique record as a string.

The method of leveraging Amazon Comprehend required a training dataset. This training data set included 611 total syllabi records with a minimum of 10 syllabi for each label or skill derived, the minimum number necessary to train Amazon Comprehend to detect the skills in other syllabi using the multi-label classification routine (*Guidelines and Quotas - Amazon Comprehend*, n.d.).

Four independent taggers, people trained to label the syllabi outcomes with the top 25 skills, performed the labeling exercise. This method reduced bias and increased accuracy in the outcome (Bruce & Wiebe, 1999). While modern algorithms have improved to allow the researcher to tag their own data, based on the extensive data size, and the potential subjectivity, led to the selection of using taggers rather than the study's researcher. The users selected as taggers each understood industry skill association to language as they are members of the researcher's cohort.

The tagging occurred within a solution provided by LightTag. This solution allowed the uploading of all the syllabi course objectives into the system. Each tagger automatically received a subset to review, and these taggers supplied the relevant tags to the documents. The taggers received training, discussed in detail in Appendix D. The final output is a data file with the course, the provided labels, and the subsequent text, as shown in Figure 7.



**Figure 3**

*Mapped Data*

Course	Label	Text
ACC_2033	MATHEMATICS   EXCEL  WRITING   CRITICAL_THINKING	1.students will submit written responses to certain assignments and will interactively discuss the issues and concepts in class; 2. most of the assigned accounting exercises, problems and examinations will require quantitative and qualitative reasoning skills to successfully complete the assignments; 3. spreadsheet and/or Internet technology will be used from time to time to support accounting class assignments; 4. most class discussions will relate the subject material to the broader global market perspective, and some assignments and examination questions will relate specifically to the global accounting environment
AAS_4013	RESEARCH   ANALYTICAL_SKILLS   CRITICAL_THINKING	1.Understand the major events and issues surrounding the Black Lives Matter movement. 2. Define what the Black Lives Matter Movement is. 3. Gain an increased understanding of protest movements. 4. Contemplate and analyze contemporary social issues faced by black people in the United States. 5. Develop critical thinking and analytical writing skills. 6. Apply research skills and conduct research on Black Lives Matter.



**Figure 3** (continued)

Course	Label	Text
ENG_3303	WRITING   CRITICAL_THINKING   COMMUNICATION_SKILLS	1. Gaining factual knowledge (terminology, classifications, methods, trends) 2. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course 3. Developing skill in expressing oneself orally or in writing 4. Learning to analyze and critically evaluate ideas, arguments, and points of view
ENG_4905	WRITING   COMMUNICATION_SKILLS   BUILDING_EFFECTIVE_RELATIONSHIPS	Students in this course will have the opportunity to do a number of things: gain a near exhaustive knowledge of a major filmmaker's body of work and some of its historical, film-historical, cultural, and other contexts; demonstrate critical reading, viewing, thinking, and writing skills, through written and oral discussion, a presentation, and a paper; work collaboratively and across multiple media to present, summarize, and contextualize research and analysis; come into contact with numerous critical and theoretical approaches to filmic texts.

*Note.* The table shows an example of tags or labels mapped to a course number and the text of the course outcomes. This table will include over 1,000 syllabi once complete



### *Data Collection Phase III – Student Data*

General Studies allows students to take courses from any non-restricted course (courses that require a professor's approval or a major status), making it difficult to identify quickly what courses students took at the macro level. Therefore, each student may take different courses. To gather each students' course data required a request from the UTSA institutional research team. The ask was for course completion records for all students who graduated from MDS between 2018 and 2021. Precisely, the course ID for each course indexing as a unique value per student. The data request asked for their most recent email address in Data Collection Phase IV, separate from the course information. The institutional research team provided the data as a comma-separated value (.csv) file, and the researcher stored this file locally on a computer for analysis. The researcher was the only person that accessed this file and planned to delete the file at the end of the study.

As not all courses have been available since 2018, the analysis required removing all courses not available in the course syllabus pulled from Data Collection Phase II. The complete list of classes was applied as individual variables to each student record, denoting a 0 to indicate the student did not take the course and a 1 indicating they had taken it. The data provided the ability to create a skills profile when correlated with the associated course outcomes.

### *Data Collection Phase IV – Survey*

Each of the former MDS graduates received a self-administered questionnaire. Initially, an already established survey was the preferred method. However, one could not be acquired. As discussed in the instrumentation section, this study leveraged a custom



survey using questions and formats utilized in previous studies. The participants received the survey electronically due to their distance; most students do not remain in the San Antonio area after their degree completion. This method allowed for higher response rates (Fink, 2003). The creation of the survey occurred leveraging a 3rd party utility, LimeSurvey (LS), a copy of which is in Appendix B. LS was selected based on the low cost of deployment (\$30 for a month) and the abundance of data retrieved within the system.

The researcher provided the questionnaire to the on-campus Institutional Review Board for review. LS automatically sent an email on December 18, 2021, containing a URL link to the online instrument and instructions to complete the survey to all the participants identified for the study. After three days, participants received a reminder email to increase participation potentially. A third and final attempt was made three days after the second attempt to procure any remainders potentially. After that period, the researcher extracted the data from the 3<sup>rd</sup> party supplier for analysis.

Participants that completed the survey had a chance to win one of two \$50 gift cards offered through a randomized drawing. The dissertation committee chair oversaw the drawing, and the winners received a digital link through their email.



### *Summary and Data Collection Timeline*

The data collected in these four phases provided all the necessary inputs to test the six research objectives. The work performed over three months, as shown in Table 2, concluded with a final push to receive greater participation in the survey. Each element of data requested served a purpose in the research. Skills data from BG for use in the skills dictionary in the custom classifier and as a feedback element in the survey, to the custom tagged features in the syllabi, to the questionnaire itself.

**Table 3**

### *Data Collection Timeline*

Data Collection Activity	Proposed Date of Collection
IRB Approval	Week 0
Collection of Syllabi	Week 1
Extract Syllabi Text	Week 2
Compile Skills Dictionary from BG	Week 3
Upload to LightTag for Tagging	Week 4
Complete and Compile Tags	Week 5
Receive Student Data from UTSA	Week 6
Send the first link to MDS graduates	Week 7
Follow-up for participation	Week 7
Final Follow-up for participation	Week 8
Extract Survey Data from LimeSurvey	Week 8



## Data Analysis Plan

There are two primary software tools used in this analysis. The first software, Amazon Comprehend, isolated out the skills associated with specific classes. Based on the complexity of NLP, using custom models provided by Amazon simplified the process. The other software, SPSS (latest version 28), performed several statistical analyses on the student data and the data returned from the survey.

### *RO1 – General Studies Student Characteristics*

The first three questions from the survey provide the details to answer RO1. A frequency table provided the results for the first three questions of the survey. This table identified how many people participated in the study and the difference between their ethnicity (ETHNICITY) coded as 0 = *Asian*, 1 = *African American*, 2 = *Hispanic or Latino*, 3 = *Native American*, 4 = *White*, their gender (GENDER) coded as 0 = *Female*, 1 = *Male*, 2 = *Gender Variant / Non-Conforming*, 3 = *Transgender Female*, 4 = *Transgender Male* and their status within the program (ENROLLMENT) as 0 = *Continuing Education* (students who were non-continuous or entered school later) and 1 = *Traditional Student* (students who went straight from HS and were continuous). These addressed whether the students align with the current literature (Hoyt & Allred, 2008) or are anomalous.

### *RO2 – Employer Skills*

Burning Glass reported the top 25 most requested skills for jobs in the US (Burning Glass Technologies, 2015). RO2 attempted to validate this supposition by leveraging BG's database to evaluate skills preferred by competitive entry-level positions, defined as having a salary greater than \$35K in the San Antonio area. There are over 17K



skills mapped in Burning Glass's database. Therefore, a simple frequency table based on the number of jobs between October – December 2021 requesting a specific skill supported this inquiry; the greater the number of positions, the greater the ranking for the skillset. This frequency table ascertained if the top 25 skills are most relevant in the San Antonio market.

### *RO3 and RO4 – Course Outcome and Skills Alignment*

Analysis of the syllabi data leveraged the custom classification function within Amazon Comprehend. The first step required training the model. As discussed in the data collection process, taggers added skills manually to a subset of the syllabi. The multi-label custom classification model used this tagged data set for training. The model ran this test data and identified *Trained* when successful. After which, the unlabeled data set ran and provided the output metrics and a final labeled data file.

To assist in estimating how well a custom classifier performs, Amazon provided several metrics to include a confusion matrix (*Custom Classifier Metrics - Amazon Comprehend*, n.d.). A confusion matrix is a table that keeps track of the classifications (both correct and incorrect). A confusion matrix provides a method to evaluate the performance of the classification model between the accurate data labels and those the model predicted (Sarkar et al., 2018). The columns in a confusion matrix represent the model's prediction (in counts), while the rows represent the actual labels (in counts) (Sarkar et al., 2018). When the count was No-No, reading from left to right, these were true negatives (TN). The number of items where the model predicted the label was not relevant to the document, nor was the actual documents labeled appropriate. When the item was No-Yes, these are false positives (FP), or where the model identified the label



as relevant, but the actual labels are not. When the item was Yes-No, these are false negatives (FN), where the model did not find the label relevant, but the actual labels suggest that it is. The final item, Yes-Yes, these were true positives (TP), where both the model predicted that the labels were relevant, and the actual labels suggested the same thing (Sarkar et al., 2018).

The accuracy metric identified how well the model predicted the correct label based on the test data. To calculate the accuracy of a model, divide the number of correct labels by the total number of labels provided (Sarkar et al., 2018). The following formula represents accuracy; Amazon Comprehend delivered it as a percentage at the macro-level (an average of all accuracy scores).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Precision is concerned with the relevant results of the model, calculated as the number of documents correctly classified by the model divided by the total number of classifications for a specific label (Sarkar et al., 2018), delivered by Amazon Comprehend at the macro-level (an average of all precision scores).

$$Precision = \frac{TP}{TP + FP}$$

Recall is an essential metric in classification problems as we are most concerned with identifying a higher percentage of positive labels regardless of total accuracy.

Another metric provided by Amazon Comprehend is recall, also known as the hit rate.

This metric identifies the percentage of relevant labels (Sarkar et al., 2018). Delivered at the macro level by Amazon comprehend (an average of all recall scores).

$$Recall = \frac{TP}{TP + FN}$$



The most common metric for classification models of this type is the F1 score, the harmonic mean of recall and precision, assisting in optimizing the classifier for balanced precision and recall (Sarkar et al., 2018), which was of most assistance in this research. Amazon Comprehend provides this metric at the macro-level (the average F1 scores).

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

A failed model receives an F1 Score of 0, whereas a score of 1 is a perfect model. Assuming all models are wrong, the goal was optimization, and anything over .5 will assume moderate accuracy in its prediction (Sarkar et al., 2018). If the model did not achieve this level of performance, the taggers would provide more labels in the data collection process and, when fed back into the model as training sets, would potentially increase the model's accuracy.

The solution delivered the final output in a JavaScript lines file provided, converted into a .csv (comma-separated value) file with all the labeled syllabi, like Figure 7, as indicated in data collection. This list was ranked to understand which courses most align with the list's skills, i.e., more skills equal higher rank. One additional step was to understand which course family is most harmonious, identified by the curriculum category, i.e., Accounting, Kinesiology, etc.

#### *RO5 – Students Enrollment*

RO5 sought to understand the outcomes from student enrollment. The first step leveraged the data achieved in RO4, matching it to the student enrollment data, and performing a frequency table to determine the number of students who took a course by ranking to determine if students took courses most aligned to the skills. This output



ascertains if students took classes associated with high skills achievement. The next step is understanding the skills present in the General Studies student population. The analysis leveraged data from RO3 and 4. Answering this objective required a frequency table using the labeled tags and the courses taken by students, corresponding skills.

The skills data created continuous variables. For instance, a student may take several courses that offer math skills. Therefore, math skills could have a value of 1 (if they took one class that offered math skills) up to the maximum possibility. Frequency tables provide the results of this analysis.

#### *RO6 – Competitive Position Relationship*

Addressing research objective six required the use of an ordinal logistic regression. This method predicts the outcome of an ordinal dependent variable given one or more independent variables (Harrell, 2015). This method determined if specific skills acquired in the courses taken during the MDS program assisted in gaining a competitive entry-level job. As stated previously, competitive is assessed through salary garnered after graduation (INCOME). The data matched the assumptions required to perform this analysis. The first assumption is an ordinal dependent variable (Harrell, 2015), salary, derived from the specific question asked in the survey, "What is your gross annual salary?" which provided ordinal responses. The independent variables, skills, are continuous, as described in RO4, which meets the second assumption for ordinal regression. The third assumption is that the variables have no multicollinearity (Harrell, 2015). Variance inflation factor (VIF) tested this assumption. At first, the ordinal regression ran with all the variables and assumed no multicollinearity as the skills are mutually exclusive (Harrell, 2015). However, if the variance inflation increased beyond a



value of 5, the assumption was that the variables have multicollinearity (Harrell, 2015). Variables with no influence were removed from the model iteratively and explained.

The fourth assumption is that the model has proportional odds, or each variable has the same effect at each split of the ordinal dependent variable (Harrell, 2015). The full-likelihood ratio test ran through SPSS determined if the variables met the proportional odds assumption. Using the Test of Parallel Lines available in SPSS, the test compared the model fit for all independent variables between a proportional odds model and a cumulative odds model without any constraints on the proportional odds (Harrell, 2015). If the Chi-square value was small and the p-value was greater than .05 or not statistically significant, the model met the assumption (Harrell, 2015). However, if the Chi-square value was large and the p-value was less than .05, statistically significant, the model violated the assumption (Harrell, 2015). If a violation occurred, examining the odds ratios for each variable occurred through separate binomial logistic regressions (Harrell, 2015). This analysis required dividing the ordinal variable (INCOME) into dichotomous dependent variables.

If the odds ratios are different for each value in each variable, then the variable did not meet the assumption and required removal from the final regression run (Harrell, 2015). If the regression met all assumptions, the ordinal regression ran using the skills variables as independents and the salary as the dependent.



To increase validity on this test, an additional odds ratio model will run using GENDER and ETHNICITY, each coded as a unique dummy variable to assess if this increases likelihood outcomes beyond skills when it comes to a competitive salary. If violations occur in the previous analysis, Mann-Whitney or Kruskal-Wallis tests are required to determine if there is influence from these variables.

#### *Additional Survey Questions*

As indicated in Table 3, not all survey questions directly align with the research objectives. These questions are there for several reasons. The primary reason for the additional questions is future research on this population. Lazear (2009) and Allred and Hoyt (2008) suggested that the skills earned initially may not be relevant immediately upon graduation. However, their career may expand to utilize these, addressed explicitly in their desired employment. This study does not initially address the specific industry students gained employment. However, this may become a relevant descriptor based on the results and its alignment to the skills association. The same justification applies to the question related to the student's choice to join this industry. The last question related to skills they find most valuable is an alignment question and used to qualitatively add color to the results of the skills association research.



**Table 3***Survey Map*

Research Question	Information Needed	Questions	Analysis
R01	Characteristics of General Studies Graduates	1 – 5	Frequency
R06	Salary Information	10 – 11	Ordinal Regression
	Additional Information about Graduates	6-9; 12-19	

*Summary and Research Objective/Analysis Alignment*

Leveraging multiple data analysis methods provides a more thorough understanding of course curriculum outcomes. While the Amazon Comprehend model can ascertain the skills associated with courses, the survey data provides the value proposition of these courses themselves. Performing the ordinal regression will identify if any skills are valuable for competitive salary or, instead, something else is the true indicator of competitive entry-level position acquisition. Table 4 provides the analysis alignment to the investigated research objectives.



**Table 4***Research Objective/Analysis Alignment*

Research Objectives	Data Collection Tool	Variables	Data Type	Data Analysis
RO1	Survey	Gender Ethnicity Enrollment	Nominal	Frequency Table
RO2	Burning Glass	Skills	Nominal	Frequency Table
RO3	Python; LightTag	Skills	Nominal	Natural Language Processing (LDA Custom Classification)
RO4	Python	Course Skills	Ordinal	Frequency Table
RO5	University Institutional Research Data	Skills Courses	Continuous	Frequency Table
RO6	Survey University Institutional Research Data	Current Salary Skills	Ordinal Continuous	Ordinal Regression

## Threats to Validity

There are several threats to the validity of this study. The first is the tagging of data for the custom classification model. While multiple taggers tagged the data independently, it still introduces human bias. This bias is controlled by having multiple taggers (Bruce & Wiebe, 1999). The second threat to validity is that the survey is self-reported. Participants may have an inherent desire to appear looking better. They can impact the study's internal validity (Brick & Tourangeau, 2017) focuses on salary performance, and thus participant responses may be influenced by social desirability



pressure. The survey attempts to mitigate this by informing participants of the anonymity at the beginning of the study.

Another threat to validity is survey participation. While an ideal confidence level of 90% requires the survey to achieve 176 surveys, the entire population of the last three years of graduates is being used instead of a random sample to ensure enough returned surveys. There are risks to this approach. Primarily one assumes that more motivated people will see the survey approach as a positive that may suggest a more competitive person in the market. The last threat to validity is the syllabi themselves. As humans write these with little guidance beyond a standardized template, the course may not teach the items listed. However, the use of several syllabi, over 1000, should help overcome this bias, especially within course families.

#### Ethical Concerns

The research required following The University of Southern Mississippi guidelines, and participants received informed consent. The guidelines require that the participants receive a description of the risks of participation, an explanation of the procedures, benefits of participation, and assurance they may withdraw from the study at any point. At the beginning of the survey, the survey reminds participants of this and can discontinue it. Participants will be over 18 and have graduated from the university, the primary qualification to participate in this survey. The destruction of all recorded materials occurs at the end of the study. The data collected in the survey is anonymous, with no attempt to identify a particular participant of the investigation beyond matching the courses completed to the survey respondent. Results of the study will be available to



the participants in the study. A personal, secure computer stores the data, and only the researcher has access.

### Summary

This chapter aimed to present the research method used to answer the research objectives. The presentation of the research design, the participants in the study, the data collection, and survey questions provided the specifics of how the researcher plans to conduct the research and the participants involved. The chapter concluded using Natural Language Processing (NLP) to understand the skills taught at a university and the quantitative survey leveraged to identify skills correlation related to competitive job acquisition. Chapter IV presents the results from this analysis.



## CHAPTER IV – RESULTS OF THE ANALYSIS

The purpose of this study was to understand industry talent demands related to outcomes in courses available to General Studies degree holders and their relation to competitive job acquisition. Two different types of data served as the basis for research, syllabi from The University of Texas at San Antonio and survey data from graduates of the MDS program at the same university. The pool of potential syllabi was 4,605. The undergraduate syllabi, reduced to 1,159 due to duplication, ran against the custom classification model. After receiving the data from Institutional Research, the survey population was 467 graduates of Multidisciplinary Studies for the past three years, all of whom received surveys. 70 (15%) of those surveyed began the survey, with 67 (14%) completing the survey in its entirety.

### Demographic Data

The response to RO1 requires descriptive statistics of demographic data to understand the relationship of those surveyed and its alignment to current studies of General Studies students. The first three questions of the survey derived the demographic data. A total of 467 MDS graduates received invitations to participate in the survey, 67 graduates completed. Tables 5, 6, 7, and 8 show the characteristics of these participants. Most respondents were male ( $n = 37$ , 55.2%); this corresponds similarly to the total demographics of the multidisciplinary studies program, 53% of whom are male. The participants were relatively mixed in ethnicity. Latino participants ( $n = 24$ , 35.8%) were smaller than the comparative 57% of MDS graduates who identify as Latino. The next largest population identified as White ( $n = 22$ , 32.8%), exceeding the 21% of MDS graduates who identify as White. 29.9% of African Americans ( $n = 20$ ) are a larger



population than the 9.5% of MDS graduates, and 1 Native American (1.5%), .18% of MDS graduates, represented the remainder of the returned survey population. Most of the graduates were traditional students ( $n = 35$ , 52.2%) vs. continuing education ( $n = 30$ , 44.8%). However, this percentage of continuing ed is greater than the 17% reflected in the general undergraduate population. The final demographic analysis required for this research is the salary bands outcome of the question, labeled SALARY. The salaries were varied, with most participants earning less than \$60K a year. The largest population earned \$50K - \$59K ( $n = 12$ , 17.9%).

**Table 5**

*Frequency Table for Variable Gender*

Gender	Count	%
Male	37	55.2
Female	30	44.8

**Table 6**

*Frequency Table for Ethnicity*

Ethnicity	Count	%
African American	20	29.9
Hispanic or Latino	24	35.8
Native American	1	1.5
White	22	32.8



**Table 7***Frequency Table for Enrollment*

Enrollment	Count	%
Traditional Student	35	52.2
Continuing Education	30	44.8

**Table 8***Salary Bands in Surveyed Population*

Salary	Count of Surveyed	%
\$20K - \$29K	10	14.9%
\$30K - \$39K	11	16.4%
\$40K - \$49K	7	10.4%
\$50K - \$59K	12	17.9%
\$60K - \$69K	9	13.4%
\$70K - \$79K	3	4.5%
\$80K - \$89K	2	3.0%
\$90K - \$99K	3	4.5%
>\$100K	10	14.9%

### Entry-Level Skills

In response to RO2, understanding the top demanded entry-level skills required the use of Burning Glass's database to pull the list of top skills for entry-level positions in San Antonio that have a salary of greater than \$35K. Burning Glass identified 45,366



posted roles between the end of October and the beginning of December 2021, with an estimated salary above \$35K. Table 9 lists the top 25 skills in order by the total count of jobs associated with each skill. The top skills requested by employers, albeit the order, did not change from the top 25 previously identified in the 2015 study by Burning Glass. Multi-tasking was the number one skill requested by entry-level employers representing 42,142 jobs, followed by communication skills ( $n = 40,182$ ) and customer service ( $n = 39,493$ ).

**Table 9**

*Skills Frequency Table*

Skill	Count of Jobs
Multi-Tasking	42,142
Communication Skills	40,182
Customer Service	39,493
Computational Skills and Typing	38,678
Microsoft Office + Word	37,721
Creativity	37,718
Teamwork	36,831
Organizational Skills	34,798
Meeting Deadlines	33,313
Planning	31,752
Project Management	30,087
Self-Starter	29,465
Critical Thinking	29,052



**Table 9** (continued)

Skill	Count of Jobs
Microsoft Excel	27,902
Supervisory Skills	27,172
Time Management	26,717
Building Effective Relationship	24,403
Analytical Skills	22,812
Problem Solving	20,210
Positive Disposition	19,668
Research	19,300
Listener	19,117
Mathematics	17,001
Presentation Skills	16,890
Detail-Oriented	16,807

#### Course Outcomes / Skills Alignment

A multi-label classification model was used to validate RO3, specifically the automated tagging of associated skills to courses based on a prediction. The four tagging associates labeled a total of 611 syllabi, using the skills identified in Table 9. Table 10 identifies the tags and their labeled counts. Analytical skill was the largest population of labels ( $n = 360$ ). Of the 25 tags, taggers identified only 17 skills with the minimum ten labels required for the model. Multi-tasking, Customer Service, Microsoft Office + Word, Meeting Deadlines, Self-Starter, Microsoft Excel, Supervisory Skills, Time Management, and Positive Disposition were absent from the tagging population used in the model. The



lack of these skills already displays misalignment between the top requested skills in the San Antonio job market and the skills indicated in syllabi as 4 of the top 5 skills requested.

**Table 10**

*Labels Frequency Table*

Tags	Count of Labels
Analytical Skills	360
Communication Skills	298
Critical Thinking	205
Mathematics	149
Problem Solving	116
Listener	113
Presentation Skills	104
Teamwork	93
Creativity	62
Computational Skills and Typing	59
Planning	56
Research	55
Organizational Skills	36
Building Effective Relationship	31
Project Management	23
Detail-Oriented	15



The predictive analysis required the upload of 611 tagged syllabi to Amazon Comprehend. Table 11 presents the confusion matrix returned by the model. The model split the 611 tagged syllabi into three groups, 488 for training, 61 for validation, and 62 for a test. The confusion matrix represents the scored results of the 62 test syllabi. As explained in Chapter III, a confusion matrix is a table that keeps track of the classifications (both correct and incorrect). A confusion matrix provides a method to evaluate the performance of the classification model between the accurate data labels and those the model predicted (Sarkar et al., 2018). The columns in a confusion matrix represent the model's prediction (in counts), while the rows represent the actual labels (in counts) (Sarkar et al., 2018). When the count is No-No, reading from left to right, these are true negatives (TN). The number of items where the model predicted the label was not relevant to the document, nor was the actual documents labeled appropriate. When the item is No-Yes, these are false positives (FP), or where the model identified the label as relevant, but the actual labels are not. When the item is Yes-No, these are false negatives (FN), where the model did not find the label relevant, but the actual labels suggest that it is. The final item, Yes-Yes, these are true positives (TP), where both the model predicted that the labels were relevant, and the actual labels suggest the same thing (Sarkar et al., 2018).



**Table 11***Confusion Matrix*

Tags	Actual	Predicted	
		No	Yes
Mathematics	No	0	0
	Yes	1	17
Analytical Skills	No	0	7
	Yes	3	33
Teamwork	No	0	0
	Yes	2	6
Communication Skills	No	0	7
	Yes	3	26
Presentation Skills	No	0	0
	Yes	3	6
Creativity	No	0	0
	Yes	1	4
Listener	No	0	2
	Yes	3	8
Problem Solving	No	0	2
	Yes	2	11
Critical Thinking	No	0	3
	Yes	8	14
Research	No	0	0
	Yes	1	4



**Table 11** (continued)

Tags	Actual	Predicted	
		No	Yes
Building Effective Relationships	No	0	0
	Yes	1	2
Computational Skills and Typing	No	0	0
	Yes	4	3
Project Management	No	0	0
	Yes	1	1
Planning	No	0	0
	Yes	4	3
Detail-Oriented	No	0	0
	Yes	1	1
Organizational Skills	No	0	0
	Yes	3	1

Most of the identified scored predictions were Analytical Skills ( $n = 33$ ) and Communication Skills ( $n = 26$ ), aligning to the high number of tagged syllabi provided to the model. Table 12 presents the precision, recall, and F1-Score values calculation for each classification algorithm.



**Table 12***Precision, Recall, F1-Score Values for the Multi-label Custom Classification Model*

Category	Precision	Recall	F1-Score
Mathematics	0.997	0.941	0.97
Analytical Skills	0.962	0.917	0.94
Communication Skills	0.935	0.897	0.92
Problem Solving	0.89	0.833	0.86
Creativity	0.907	0.8	0.85
Teamwork	0.94	0.75	0.83
Presentation Skills	0.932	0.727	0.82
Research	0.825	0.8	0.81
Listener	0.892	0.727	0.80
Critical Thinking	0.87	0.667	0.76
Writing	0.822	0.625	0.71
Building Effective Relationships	0.722	0.667	0.69
Computational Skills and Typing	0.703	0.429	0.53
Project Management	0.555	0.5	0.53
Planning	0.541	0.333	0.41
Organizational Skills	0.454	0.25	0.32
Detail-Oriented	0.508	0	0.00
All Labels	0.889	0.769	0.82

The F1-score, the harmonic average of precision and recall, identifies the overall accuracy of a classification model (Sarkar et al., 2018). The overall model achieved an



F1-Score of .82, which denotes a high level of accuracy in the model's ability to predict an accurate label for each untagged syllabi, as it exceeds the .5 required for moderate predictability. However, Computational Skills and Typing, Project Management, Planning, Organizational Skills, and Detail Oriented all scored right at or below .5. Five suggests that those automated tags are unreliable and removed from the final model.

Based on the model's performance, 1,159 untagged syllabi ran as batch predictions against the model. This method produces individual confidence scores for each label related to an individual syllabus, i.e., for MGT 4643, analytical skills received a confidence score of .99. A score above .5 is considered a strong prediction score, and therefore the label is reliable, which means that MGT 4643 most likely has a student outcome of analytical skills. Table 13 presents the count of syllabi matching each of the final 11 tagged skills run in the model that achieved a confidence score above .5.

**Table 13**

*Skills-ranking Based on Syllabi Count*

Skill	Count	% of Tagged Syllabi
Analytical Skills	929	80.2%
Communication Skills	668	57.6%
Critical Thinking	339	29.2%
Mathematics	198	17.1%
Problem Solving	142	12.3%
Research	84	7.2%
Teamwork	48	4.1%
Creativity	34	2.9%



**Table 13** (continued)

Skill	Count	% of Tagged Syllabi
Listener	31	2.7%
Presentation Skills	25	2.2%
Building Effective Relationships	5	0.4%

The model applied 2,503 tags, with Analytical Skills receiving 929 of these tags, 80.2% of the 1,159 syllabi, followed by Communication Skills ( $n = 668$ , 57.6% of syllabi) and Critical Thinking ( $n = 339$ , 29.2% of syllabi). This outcome suggests that most undergraduate classes provide analytical, communication, and critical thinking skills. The final rank-ordered list of classes is in Appendix E. One additional inquiry was to understand what programs offered the most skills. Table 14 depicts the top 10 programs providing the most skills.

**Table 14***Top Programs by Skills*

Subject	Count of Skills	% of Total Skills
Math	114	4.5%
History	109	4.3%
Music	107	4.2%
Communications	98	3.8%
Political Science	96	3.8%
English	95	3.7%
Biology	91	3.6%



**Table 14** (continued)

Subject	Count of Skills	% of Total Skills
Sociology	76	3.0%
Art	70	2.7%
Kinesiology	69	2.7%

One consideration is the number of syllabi for each subject run through the model. Music had the highest number of syllabi ran through the model, 117, but produced only 107 skills, resulting in an average skill per course of 1.58. Whereas biomechanical engineering, based on its specificity of a program, had only 20 syllabi but produced 62 skills resulting in an average of 3.1 skills per course, a significantly higher percentage of skills than music. Table 15 expresses these averages for courses with ten or more syllabi in the corpus.

**Table 15**

*Total Skills as Average of Syllabi in Corpus*

Subject	Total Skills	# Of Syllabi	Avg. # of Skills by Courses
Biomechanical Engineering	62	20	3.10
Finance	46	15	3.07
History	109	36	3.03
Biology	91	31	2.94
English	95	33	2.88
Engineering	42	15	2.80
Civil Engineering	42	15	2.80



**Table 15** (continued)

Subject	Total Skills	# Of Syllabi	Avg. # of Skills by Courses
Art	70	27	2.59
Chemical Engineering	31	12	2.58
Accounting	31	12	2.58
Math	114	46	2.48
Communication	98	40	2.45
Architecture	56	24	2.33
Ecology	32	14	2.29
Nutritional Science	34	15	2.27
Educational Psychology	29	13	2.23
Statistics	31	14	2.21
Political Science	96	45	2.13
Geography	38	18	2.11
Management	40	19	2.11
Interdisciplinary studies	21	10	2.10
Chemistry	46	22	2.09
Spanish	27	13	2.08
Health	37	18	2.06
Sociology	76	37	2.05
Physiology	51	25	2.04
Kinesiology	69	34	2.03
Health & Nutrition	30	15	2.00



**Table 15** (continued)

Subject	Total Skills	# Of Syllabi	Avg. # of Skills by Courses
Philosophy	22	11	2.00
Bicultural Studies	29	15	1.93
Marketing	21	11	1.91
Construction Science	34	18	1.89
Global Arena	49	26	1.88
Geology	64	34	1.88
Public Administration	37	20	1.85
Special Education	34	19	1.79
Psychology	54	33	1.64
Criminal Justice	40	25	1.60
Music	107	70	1.53
English as a Second Language	17	12	1.42
Counseling	21	15	1.40

This analysis suggests that apart from 3 programs (Art, English, History), the classes that provide the greatest number of skills per class (an average of 2.84 skills per course) were those heavy in science and math. The final ranked course results are of interest to RO5 and identify if students from MDS enrolled in high skill count classes.

#### Student Enrollment

The previous results identified that numerous courses offer students the ability to acquire entry-level skills; however, whether these students enroll in them is the premise



behind RO5. Student records assisted in identifying if students enrolled in high-skill courses. MDS students took classes from several different areas by the nature of the program. However, most students took classes with less than three skills, aligning with the previous results that indicated the median skills offered by courses was 2.2. Tables 16 through 26 outline the count of skills acquired of each type by the student, assumed by the number of classes the student took that featured that skill.

**Table 16**

*Skill Frequency for Analytical Skills*

Number of Analytical Skills	Count of Students	% of Students
3	2	0.07%
4	1	0.1%
5	6	0.4%
6	4	0.3%
7	13	1.1%
8	16	1.6%
9	15	1.7%
10	24	2.9%
11	22	3.0%
12	22	3.2%
13	27	4.3%
14	20	3.4%
15	31	5.7%
16	30	5.9%



**Table 16** (continued)

Number of Analytical Skills	Count of Students	% of Students
17	31	6.5%
18	15	3.3%
19	10	2.3%
20	14	3.4%
21	20	5.2%
22	13	3.5%
23	17	4.8%
24	21	6.2%
25	25	7.7%
26	19	6.1%
27	11	3.6%
28	12	4.1%
29	9	3.2%
30	6	2.2%
31	4	1.5%
32	1	0.4%
33	3	1.2%
35	3	1.3%



**Table 17***Skills Frequency for Communication Skills*

Number of Communication Skills	Count of Students	% of Students
2	1	0.21%
3	4	0.9%
4	8	1.7%
5	15	3.2%
6	12	2.6%
7	12	2.6%
8	24	5.1%
9	32	6.9%
10	27	5.8%
11	30	6.4%
12	32	6.9%
13	25	5.4%
14	28	6.0%
15	28	6.0%
16	21	4.5%
17	26	5.6%
18	17	3.6%
19	21	4.5%
20	24	5.1%
21	20	4.3%
22	13	2.8%



**Table 17** (continued)

Number of Communication Skills	Count of Students	% of Students
23	11	2.4%
24	15	3.2%
25	4	0.9%
26	8	1.7%
27	3	0.6%
28	2	0.4%
29	1	0.2%
30	1	0.2%
31	1	0.2%
35	1	0.2%



**Table 18**  
*Skills Frequency for Critical Thinking Skills*

Number of Critical Thinking Skills	Count of Students	% of Students
1	7	1.50%
2	29	6.2%
3	35	7.5%
4	43	9.2%
5	51	10.9%
6	58	12.4%
7	47	10.1%
8	38	8.1%
9	41	8.8%
10	26	5.6%
11	23	4.9%
12	25	5.4%
13	17	3.6%
14	10	2.1%
15	10	2.1%
16	4	0.9%
17	2	0.4%
20	1	0.2%



**Table 19***Skills Frequency for Mathematics*

Number of Mathematics	Count of Students	% of Students
0	134	28.69%
1	114	24.4%
2	69	14.8%
3	44	9.4%
4	26	5.6%
5	21	4.5%
6	14	3.0%
7	15	3.2%
8	14	3.0%
9	7	1.5%
10	1	0.2%
11	2	0.4%
12	3	0.6%
13	3	0.6%



**Table 20***Skills Frequency for Problem Solving*

Number of Problem Solving	Count of Students	% of Students
0	167	35.76%
1	111	23.8%
2	89	19.1%
3	33	7.1%
4	26	5.6%
5	18	3.9%
6	8	1.7%
7	3	0.6%
8	5	1.1%
9	3	0.6%
10	1	0.2%
11	2	0.4%
13	1	0.2%



**Table 21***Skills Frequency for Research*

Number of Research	Count of Students	% of Students
0	203	43.47%
1	141	30.19%
2	82	17.6%
3	33	7.1%
4	6	1.3%
5	1	0.2%
7	1	0.2%

**Table 22***Skills Frequency for Teamwork*

Number of Teamwork	Count of Students	% of Students
0	194	41.54%
1	145	31.05%
2	78	16.70%
3	35	7.5%
4	10	2.1%
5	4	0.9%
6	1	0.2%



**Table 23***Skills Frequency for Presentation Skills*

Number of Presentation Skills	Count of Students	% of Students
0	104	22.27%
1	122	26.12%
2	95	20.34%
3	66	14.1%
4	38	8.1%
5	25	5.4%
6	9	1.9%
7	6	1.28%
8	1	0.21%
9	1	0.21%



**Table 24***Skills Frequency for Listener*

Number of Listener	Count of Students	% of Students
0	117	25.05%
1	99	21.20%
2	102	21.84%
3	55	11.8%
4	38	8.1%
5	31	6.6%
6	11	2.4%
7	11	2.36%
8	3	0.64%

**Table 25***Skills Frequency for Creativity*

Number of Creativity	Count of Students	% of Students
0	225	48.18%
1	140	29.98%
2	68	14.56%
3	20	4.3%
4	10	2.1%
5	4	0.9%



**Table 26***Skills Frequency for Building Effective Relationships*

Number of Building Effective Relationships	Count of Students	% of Students
0	369	79.01%
1	97	20.77%
2	1	0.21%

As previously identified, the analytical skill was the leader in potential skills acquisition. MDS students took between 3 to 35 classes where analytical skill was possible. The most significant number of students took between 13 – 15 classes with analytical skills. The same large variability was present in communication skills whereby MDS students took between 2 to 35 classes where communication skills were possible. The most significant number of students took between 8 and 11 classes with communication skills as an outcome. In critical thinking, the maximum class count drops to 20, and the most significant number of students took 5 to 6 classes. The skills ranking is most beneficial for the last part of the analysis to determine if skills acquisition facilitates more significant competitive job acquisition.

#### Competitive Position Correlation

The previous research objectives served the primary purpose of preparing the data for use in an ordinal regression. The identified salary bands provided from the survey serve as the ordinal dependent variable with the skills identified in RO5, isolated for those graduates who returned the survey, serving as the series of continuous independent variables. These aspects meet the assumptions necessary to perform an ordinal



regression. However, the independent variables required a variance factor test to ensure no multicollinearity. Table 27 presents the initial results.

**Table 27**

*Results of Initial Variable Influence Factor Analysis*

Model	Collinearity Statistics	
	Tolerance	VIF
# of Analytical Skills Acquired	0.105	9.483
# of Communication Skills Acquired	0.117	8.566
# of Critical Thinking Skills Acquired	0.324	3.085
# of Mathematics Skills Acquired	0.334	2.997
# of Problem Solving Skills Acquired	0.322	3.107
# of Writing Skills Acquired	0.470	2.130
# of Research Skills Acquired	0.858	1.165
# of Teamwork Skills Acquired	0.584	1.712
# of Presentation Skills Acquired	0.283	3.535
# of Listener Skills Acquired	0.264	3.795
# of Creativity Skills Acquired	0.608	1.644
# of Building Effective Relationships Acquired	0.662	1.511

*Note.* Variable Influence Factor stated as VIF.

As stated by Harrell (2015), any VIF greater than 5 represents critical levels of multicollinearity, leading to potential issues in understanding which variables contribute to the explanation of the dependent variable in the final ordinal logistic regression. The close relationship in the summary statistics between analytical skills and communication



skills suggests communication skills removal. Rerunning the test produced results presented in table 28.

**Table 28**

*Results of Second Variable Influence Factor Analysis*

Model	Collinearity Statistics	
	Tolerance	VIF
# of Analytical Skills Acquired	0.322	3.104
# of Critical Thinking Skills Acquired	0.324	3.082
# of Mathematics Skills Acquired	0.360	2.778
# of Problem Solving Skills Acquired	0.322	3.104
# of Writing Skills Acquired	0.483	2.070
# of Research Skills Acquired	0.876	1.141
# of Teamwork Skills Acquired	0.598	1.672
# of Presentation Skills Acquired	0.290	3.451
# of Listener Skills Acquired	0.264	3.792
# of Creativity Skills Acquired	0.659	1.518
# of Building Effective Relationships Acquired	0.671	1.491

*Note.* Variable Influence Factor stated as VIF.

The removal of communication skills reduced the VIF values of analytical skills below 5, reducing multicollinearity across the independents, meeting the third assumption to run an ordinal regression. However, the assumption of proportional odds failed, as assessed by a full likelihood ratio test comparing the fit of the proportional odds location model with varying location parameters,  $X^2(77) = 95.098, p = .079$ . The large Chi-



Square suggests a large difference between the two models. Failing this assumption requires understanding the odds ratio for each variable as calculated through separate binomial logistic regressions, requiring the split of the dependent variable (SALARY) into seven separate variables labeled Income 1 - 7. Table 29 presents these results.

**Table 29**

*Odds Ratio Results for Income*

Independent Variable	Exp(B) (Odds Ratio, OR)						
	Income 1	Income 2	Income 3	Income 4	Income 5	Income 6	Income 7
Analytical Skills	0.788	0.963	0.945	0.917	0.938	0.982	1.023
Critical Thinking Skills	3.079	1.256	1.151	1.056	1.132	1.136	1.069
Mathematics Skills	2.433	1.018	0.809	0.664	0.717	0.766	0.699
Problem Solving Skills	0.772	1.449	1.922	4.703	3.025	2.102	2.175
Writing Skills	1.317	0.455	0.724	1.954	1.602	1.446	1.234
Research Skills	0.19	1.085	0.739	0.57	0.589	0.713	0.649
Teamwork Skills	1.502	3.674	2.22	4.147	1.511	0.866	1.108
Presentation Skills	0.656	0.798	0.787	0.607	1.134	1.088	1.477
Listener Skills	0.639	0.967	0.986	0.908	0.779	0.854	0.761



**Table 29** (continued)

Independent Variable	Exp(B) (Odds Ratio, OR)						
	Income 1	Income 2	Income 3	Income 4	Income 5	Income 6	Income 7
Creativity Skills	0.503	1.155	1.115	0.756	1.095	3.202	2.311
Building Effective Relationships	0	0.876	2.06	1.865	0.857	0.81	0.427

The separate binomial regression results negate the assumptions required to perform an ordinal regression. Therefore, the most suitable option is to run a multinomial logistic regression. However, the challenge with this approach is the loss of the ordered value. The likely outcome can only express the likelihood of being in a specific salary band, not the ability to understand if more skills suggest a higher salary. This method will not address the research objective. Therefore, no result is possible from this research to determine if the greater number of skills equates to a greater level of compensation.

The last step was to understand if gender or ethnicity had any relationship to the MDS graduates' job acquisition. The failure to garner results led to the abandonment of an ordinal regression for the use of non-parametric tests. The purpose of this test is to determine if there is a statistically significant difference between groups (Fried, 2018). A Mann-Whitney U test determined differences in annual gross Salary between males and females. Although the distributions of Salary were dissimilar when identified through the visual analysis, annual gross Salary for males (mean rank = 36.53) and females (mean rank = 30.88) were not statistically significantly different,  $U = 648$   $z = 1.191$ ,  $p = .234$ .



This result assumes that gender does not play a significant role in earnings for the surveyed population.

A Kruskal Wallis H test determined if there were differences in annual gross Salary between the ethnicity of MDS graduates, African American ( $n = 20$ ), Hispanic or Latino ( $n = 24$ ), Native American ( $n = 1$ ), and White ( $n = 22$ ). Distributions of salary were relatively similar for all groups, as assessed by visual inspection of a boxplot. Apart from Native Americans ( $n = 8$ ), an outlier in the dataset, salaries only slightly decreased from Latinos and African Americans ( $n = 3$ ) to White (3.0), and the differences are not statistically significant,  $\chi^2(3) = 5.343$ ,  $p = .148$ .

### Summary

This research aimed to identify outcomes for the six research objectives. Research Objective 1 described the demographic characteristics of the graduates from MDS who took the survey compared against the demographic population of UTSA. The respondent population was equally diverse across gender (males 55.2%, females 44.8%), ethnicity (African Americans 29.9%, Latinos 35.8%, White 32.8%), and enrollment (traditional student 52.2%, continuing education 44.8%); differing dramatically across salary bands.

Understanding the entry-level job market and associated skills was the premise behind Research Objective 2. Accessing a third-party database, Burning Glass, to identify entry-level skills count by postings, it was determined that 45,366 entry-level jobs were making \$35K or more open in the San Antonio market between late October to early December. Of those entry-level positions, the top 25 skills from the 17,000+ available were extracted and determined that the same top 25 skills identified for entry-level jobs nationwide in a 2015 study by Burning Glass remained the same in San Antonio in 2021.



The specific order of the skills changed, with multi-tasking representing the greatest number of job postings ( $n = 42,142$ ) followed by communication skills ( $n = 40,182$ ) and customer service ( $n = 39,493$ ).

The purpose of Research Objective 3 was to determine the skills associated with courses offered by The University of Texas at San Antonio. This effort required the most complex analysis of this research study. Utilizing Natural Language Processing, the results suggest that most courses allowed students to acquire analytical and critical thinking and communication skills. Based on the model results, only 12 of the 25 skills identified in RO2 appeared confidently against the syllabi. Apart from Art, History, and English, the top-performing programs that offered the greatest number of skills per course were science and math-related. Of the entire population, the median number of skills a student can acquire in a single course is 2.2.

Research Objectives 4 and 5 investigated if high skill providing courses were available to MDS students and enrolled. The results suggest that students from MDS enroll in high skill acquisition classes related to the top 3 skills (Analytical and Critical Thinking, Communication), but at varying rates. However, there was a steep decline in enrollment for the additional skills, such as math or problem-solving, indicated by the high percentage of students with 0 or 1 skill acquired. This result is potentially a symptom of the number of classes exhibiting this skill in the syllabi population and indicates that enrollment is significantly lower.

The last part of the analysis, Research Objective 6, attempted to understand if skills were a factor in competitive job acquisition, earning an entry-level job >\$35K. The results were inconclusive due to violations in several assumptions to leverage the ordinal



variable SALARY. As for whether gender or ethnicity were factors, the analysis determined that there was not a significant difference between the groups regarding earnings.

Chapter IV provided the results of this analysis. Chapter V's final chapter provides the relationship between the analysis results and previous literature. It offers a discussion of the findings, a set of conclusions and potential implications, and future recommendations for research.



## CHAPTER V – DISCUSSIONS AND CONCLUSIONS

Recent research suggests that having a college education does not guarantee success in the economic market, especially those pursuing non-business, non-STEM degrees (Garner et al., 2019; Lobo & Burke-Smalley, 2018; Nicholas, 2018; Wu & Lewis, 2019). This study addresses whether students who graduate from one such specific program, General Studies, meet the talent demands of industry based on the skills acquired at university. Specifically, did the university provide a curriculum that addressed the skills required for entry-level jobs, and did General Studies students enroll.

### Findings, Conclusions, and Recommendations

This study implies that in the case of entry-level jobs, leveraging a natural language processing model, courses at the university did not explicitly provide most skills requested in all classes offered, not just those enrolled in by students. The study also implies that skills from universities may not be a deciding factor on the ability to acquire a competitive salary. However, this study does provide a university with a method to understand the skills being provided and help them craft coursework that meets entry-level employers' demands. This chapter provides a summation of the findings from the study, associated conclusions, and future recommendations.

#### *Finding 1*

*Universities do not explicitly meet the talent demands of employers.*

Utilizing a natural language processing model and the provided outcomes/objectives from over 1,000 undergraduate syllabi from The University of Texas at San Antonio, only 12 of 25 skills requested by employers appeared in the syllabi with confidence. Analytical skills, communication skills, and critical thinking appeared



prominently in the outcomes of the syllabi. Conversely, they were not the most vital skills requested by entry-level employers. Of the programs that offered the most significant number of skills, meeting the entry-level employers' needs, per course, courses heavy in science and math outnumbered other programs, however not significantly.

### *Conclusion 1*

The analysis of these results supports current research that universities are not providing the skills to meet entry-level demands (Burke et al., 2020; Lobo & Burke-Smalley, 2018). However, in contradiction to several studies stating that analysis, critical thinking, or communication skills, skills requested by business leaders, are missing from the academy (Jackson, 2020; Weise, 2019), most of what graduates receive from university courses are these skills as shown in their skills ranking of (communication skills (2), critical thinking (13), analytical skills (18)). This finding is the most damaging to the skills bundling approach, as students cannot earn various skills and instead have them isolated into a few focused areas.

This study concludes that students acquire skills but that syllabi do not explicitly outline skills requested by entry-level employers making it difficult to translate for graduating students. Also, science and math classes outnumber skills per course, these courses did not differ significantly from the median number of skills per course in other programs. This result supports the research that even STEM degrees have challenges translating their education to skills required by industry (Ntiri et al., 2004; Weise, 2019; Weise et al., 2018; Yu, 2019).



### *Recommendation 1*

Translating skills acquired at universities has been shown in the research to be a challenge primarily because college courses do not explicitly express the skills acquired by the student after the program. For those that do, those skills are not necessarily the most in-demand by entry-level employers. Therefore, professors, the creators of the courses and syllabi, ensure that their courses offer those skills to students. Continuously analyzing the current job market to identify skills requested by employers may help professors create a more aligned curriculum to in-demand skills.

With the model developed for this study, university leaders, program chairs, even professors can now scan the objectives of their courses and determine what expressed skills. Apart from explicitly writing a skill, such as multi-tasking, these academic practitioners can now identify if the courses align with the entry-level market's skills. This activity may help the university be competitive and attractive to employers and help students craft their resumes to express the skills acquired from the university more accurately.

An additional recommendation may include a certificate or technical programs to increase skills acquisition to match those employers' demands. Research has indicated that universities and colleges can mitigate the ill-defined nature of all degrees with certificate programs or competency education supplements (Craven & DuHamel, 2003; Gerstein & Friedman, 2016). Encouraging all students to enroll in these programs, not just those in General Studies, may help students align better with the entry-level market.



## *Finding 2*

*General Studies students differ demographically from the undergraduate student population.*

Of those General Studies graduates surveyed, the research determined that continuing education students and minorities appeared higher as a percentage than the total population of the undergraduate population. However, there was no significant difference within the survey population between groups (continuing education vs. traditional, White vs. African American vs. Latino, Male vs. Female). The percentages received in the survey matched those of the MDS program.

## *Conclusion 2*

The results of the study align with previous research findings. Previous studies suggested that General Studies programs have evolved to a degree focused on completion, especially for continuing education and minorities (Hoyt & Allred, 2008). In the survey results, graduates represented minorities and continuing education more significantly than other undergraduate programs at UTSA. Their representation in the survey matched the MDS program, suggesting that the MDS program matches previous research findings.

General Studies as a degree-focused program for minorities or continuing education may not be harmful. The high variation in salary, especially with a median dollar amount > \$50K, suggests that the degree may not be the primary influencer in compensation acquisition. When reviewing the demographics of those surveyed, ethnicity and gender did not play a role in salary acquisition. This result, while contrary to previous researcher drawbacks on the General Studies and Liberal Arts degrees (Garner



et al., 2019; Moran, 2010; Weise, 2019; Weise et al., 2018), supports literature on the continued value of degree completion (Gandara et al., 2017; Gándara & Rutherford, 2018).

### *Recommendation 2*

General Studies degree programs as a completion-focused degree may supplement work experience for continuing education students. Therefore, if the student's goal is degree completion, General Studies programs fill this need. Often, continuing education students have jobs beyond entry-level status, and they are potentially supplementing their skillset with employer-requested skills such as analytical, communication, and critical thinking. However, for traditional students, it might suggest that the format for General Studies may not be as conducive. Therefore, for traditional students, General Studies degrees may require additional skill supplements to meet employers' demands or different experiences.

### *Finding 3*

*Acquired skills was an inconclusive predictor of a competitive salary.*

The research performed was inconclusive regarding skills acquisition and competitive pay. This finding may be due to the limited number of participants. Still, it might also suggest that skills, especially those acquired by General Studies students, did not directly correlate to salaries earned. Running the binomial logistic regression exposed difficulty in skills to serve as a predictive variable of salary grouping.

### *Conclusion 3*

The reason that research on General Studies and traditional liberal arts degrees regarding curriculum is limited may be due to the difficulty in understanding achieved



skills or competencies. Lacking a core curriculum reduces uniform outcomes and creates difficulty understanding generalized applicability for a degree, leading many employers to focus on major specificity (Leighton & Speer, 2020). While there was no significant distinction between salary earned and skills acquired, this was due to the lack of normalization in the participants' skill count and variability on salary. Therefore, acquired skills from university play a minimal role in competitive job acquisition.

These results may align much closer to Lazear's (2009) premise that skills bundling is more strength than any other outcome. As the students only acquire skills from a small pool of available skills, they do not increase their chances of differentiation. Therefore, they require other skills approaches to be competitive. As seen in finding 1, competitive entry-level jobs do not specifically call out skills one can readily acquire from university; therefore, a student must develop those skills elsewhere. Unseen in the data, this reality may account for the salary differentiation.

### *Recommendation 3*

General Studies degrees offer students the ability to acquire skills. However, the lack of a core curriculum ensures employers' difficulty understanding the outcomes learned from students. Creating more structured results for the program could assist traditional students in translating their skills to the market. The varied number of skills acquired by each student suggested that there is no defined direction provided for them. This lack of focus, while not indicative of the salary results, may hinder employers from understanding the value of a General Studies degree. As a means for graduate



completion, the degree makes sense; however, the value was inconclusive as a tool for skills adoption.

A recommendation is to develop two different General Studies degrees, one that is for continuing education and one for traditional students. Traditional students may leverage the model's outcomes to focus on classes that offer greater skills achievement, while continuing education can focus on degree attainment. However, the model is worth investigating as it could deliver more positive student outcomes.

### Limitations

A limitation in a study design is the systematic bias that a researcher could not or did not control and could affect the result (Brutus et al., 2013). This study has several limitations. The first is generalizability, referring to applying the study's results to the general population (Mullinix et al., 2015). There was a Type I or  $\alpha$ -error due to the small sample size. After a second attempt, the sample only achieved 67 participants, which out of a sample of 467 students achieved a 90% confidence level with a margin of error of 9%. The sample size reduces the generalizability for the population sampled, MDS students. There are generalizability issues for other General Studies programs outside The University of Texas at San Antonio. This survey and analysis only applied to graduates from this specific program. The responses to the survey present an additional limitation. As these are self-reported results, the researcher cannot verify the outcomes are accurate.

The model, built to run against any objective/outcome for any syllabi, still presented several limitations, tags used, taggers bias, and individually developed syllabi outcomes. Using the BG data, the researcher developed the tags. This single source of tag creation creates a bias as potentially relevant skills will be lost in the tagging process.



Taggers provided the trained data to the model, the researcher did provide training, and the tags were dual coded. However, the use of human intervention creates an opportunity for bias.

Additionally, the created tags may vary if the researcher developed a corpus of syllabi from a different university. Another limitation related to syllabi was that professors wrote the syllabi with limited governance by leaders or other practitioners, such as the department chair. Therefore, even though a skill may have been relevant to the course offering, the model would not detect these skills if the professor failed to include them.

### Discussion

Education has been a central principle in human capital theory (Becker, 1962). Yet, more recent research has suggested that the type of education, specifically college major, may matter (Leighton & Speer, 2020; Pennington & Stanford, 2019). Acknowledging, that employers are requesting specific skills development as opposed to general acumen (Weise, 2019; Weise et al., 2018; Woodside, 2018). Colleges have expedited their technical training and competency development programs to address these skills challenges (Craven & DuHamel, 2003; Friedman & Friedman, 2018; Gerstein & Friedman, 2016; Goldring, 2017). At the same time, they are attempting to maintain graduation rates by developing General Studies programs absent of formal rigor (Bell et al., 2018; Green et al., 2007; Moran, 2010; Rios, 2019). Yet, General Studies degrees are not necessarily a negative reality for students.

The purpose of this study was an attempt to identify if General Studies degrees could be a vehicle for skills development that leads students to competitive job



acquisition, comparable to other degree programs. Previous research has indicated that education is one method for skills development. It is the types and number of skills a person acquires that establish their capital level. Under this assumption, this study investigated if a local university offered several skills available to General Studies students and if students developed more skills earned more money upon graduation. The study suggests that universities do not provide various skills, at least those demanded by entry-level employers. Most courses focus on analytical skills, communication, and critical thinking; while requested by employers, they were not the most in-demand.

General Studies students varied in the number of skills they developed, but overall, no number or type of skills earned had a bearing on the salary they earned upon graduation. However, the degree itself may be the value proposition as most of those surveyed were making more than \$50K supporting previous research that degrees still matter in attaining higher wages. Colleges and universities may not be a center for learning many skills, especially not those associated with current employment, and that may not be their purpose. Even though researchers and business leaders criticize the academy for failure to meet the market needs, universities are strongholds of the skills business leaders lament they need the most.

The reality for a graduate from any program is to make their degree relevant as all classes offer the same fundamental skills regardless of the program. The results of this study might suggest that earning a college degree doesn't become significantly valuable until long past entry-level employment when analytical, critical thinking, and communication skills are more relevant. Which helps validate why General Studies offers a purpose for continuing education.



The General Studies degree began as an academically inclined degree, for students who wanted to create their own path to understanding and inquiry. This model required the use of a faculty advisor; however, students received a holistic education. This model is still in use in private universities, but with the advent of public-school pressure and a requirement to increase admissions, the General Studies model has evolved into the catch-all degree as confirmed by this research. This is not bad for continuing education students who may require the degree to facilitate a requirement for a new job role, but for traditional students the lack of curriculum and limited direct investment from faculty, the General Studies degree is not the ideal choice. The outcome of the General Studies degree for traditional students is no longer holistic education, but simply a degree. The lack of math skills for most students and high variability of the count of skills amongst graduates confirms this. There is value in General Studies education, but either it needs to become two separate tracks, a continuing education – degree completion track and a curriculum defined program that generates truly holistic education for students or it needs to be reserved for continuing education solely.

#### Recommendations for Future Research

This study provides an exploratory foundation for future research. The most immediate of these would be the improvement of the natural language processing model. As this study used only syllabi from a single university, the opportunity to expand to other schools is a reasonable first step. Expanding to other universities offers several beneficial outcomes. The first is the potential to add more tags to the model and increase the tagged syllabi samples, potentially improving the model's performance, especially to previously unaccounted skills. The second is the ability to increase the survey pool.



Expansion of the survey pool would increase potential participation and help identify if skills lend to higher salaries. The third benefit is helping to confirm if the skills requested by the San Antonio market for entry-level positions are the same in other areas, assisting in understanding if college may be more beneficial in other regions.

This study investigated General Studies degrees, but based on the outcomes, it appears that the skills are similar for courses for most programs. There is potential to expand to all degree programs and their graduates. Doing so would deepen the research on skills development and address questions related to college major specificity and if it is a skills-related commentary or associated with signaling choices by employers.

The last recommendation is to investigate graduates from General Studies programs specifically. Instead of focusing on the degree program itself, future research could identify what skills these graduates already have, where they acquire them, and if college was merely supplemental or an establishing body. This research would help General Studies programs focus on better development of their advising structure and benefitting their students' needs.

### Summary

Skills attainment from a General Studies program was the premise of this current study. Using a combination of natural language processing on a population of undergraduate syllabi and statistical methods against survey results from the past three years of MDS graduates served as the framework for the investigation. Results showed that varied skills development is not an outcome students can expect from college courses, at least in terms of the skills requested by entry-level employers. However, in terms of salary, the study results suggest it is not the skills acquired in college that



indicate a higher wage. The study effectively understood how natural language processing helps recognize outcomes of courses and may assist college leadership in improving the overall curriculum.

The outcomes of this study partially align with previous research. All students, regardless of degree, may struggle to communicate the skills earned in college (Ntiri et al., 2004; Yu, 2019), and universities are not meeting the needs of entry-level employers (Burke et al., 2020; Lobo & Burke-Smalley, 2018). However, skills development may not be the strongest indicator of success, at least concerning college degrees, as completion may signal to employers' value.

As the number of graduates continues to increase from General Studies programs, research must continue to investigate the value of this degree. In the case of continuing education, the program's purpose may just be graduation. However, for traditional students who want to remain competitive, additional skills may be required not offered at a university. College and universities are a significant investment, and the return on that investment ought to be warranted. Providing students with realistic outcomes is critical to ensuring universities remain relevant and students feel valued.



## APPENDIX A – APPROVAL EMAIL

---

Sunday, September 12, 2021 at 14:57:19 Eastern Daylight Time

**Subject:** Study  
**Date:** Tuesday, September 7, 2021 at 11:13:04 PM Eastern Daylight Time  
**From:** Karen Daas  
**To:** Gregory Sansone

Hi Gregg,

Pending IRB approval, I approve your research with Multidisciplinary Studies students.

Best,  
Karen

Karen L. Daas, Ph.D.  
Program Director for Multidisciplinary Studies and Academic Inquiry and Scholarship  
Associate Professor of Communication  
The University of Texas at San Antonio



## APPENDIX B – IRB Approval Letter

### Office of Research Integrity

118 COLLEGE DRIVE #5116 • HATTIESBURG, MS | 601.266.6756 | WWW.USM.EDU/ORI



#### NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

The project below has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident submission on InfoEd IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.

PROTOCOL NUMBER: 21-164  
PROJECT TITLE: COMPARING INDUSTRY DEMANDS FOR TALENT TO GENERAL STUDIES CURRICULUM FOR COMPETITIVE JOB ACQUISITION  
SCHOOL/PROGRAM School of Interdisciplinary Studies & Professional Development  
RESEARCHERS: PI: Gregory Sansone  
Investigators: Sansone, Gregory-Brown, Hamett-  
IRB COMMITTEE ACTION: Approved  
CATEGORY: Expedited Category  
PERIOD OF APPROVAL: 19-Nov-2021 to 18-Nov-2022

Donald Sacco, Ph.D.  
Institutional Review Board Chairperson\*)



## APPENDIX C – SURVEY

### **Post-graduation Multi-disciplinary Studies Survey**

Hello, previous MDS student. You are invited to participate in our survey to align your undergraduate schooling to your occupational attainment. In this survey, approximately 500 people will be asked to complete a survey that asks questions about your job attainment. It will take approximately 10 minutes to complete the questionnaire. Your participation in this study is entirely voluntary. There are no foreseeable risks associated with this project. However, if you feel uncomfortable answering any questions, you can withdraw from the survey at any point. Your survey responses will be strictly confidential, and data from this research will be reported only in the aggregate. Your information will be coded and will remain confidential. If you have questions at any time about the survey or the procedures, you may contact Gregg Sansone at 210-913-7768 or by email at [Gregory.Sansone@UTSA.edu](mailto:Gregory.Sansone@UTSA.edu)

There are 16 questions in this survey.

#### *Demographic Questions*

##### **1. To Which Gender Do You Most Identify?**

*Choose one of the following answers*

*Please choose **only one** of the following:*

- Male
- Female
- Transgender Female
- Transgender Male



- Gender Variant/Non-Conforming
- Not Listed

**2. How would you specify your ethnicity?**

*Choose one of the following answers*

*Please choose **only one** of the following:*

- White
- African American
- Hispanic or Latino
- Native American
- Asian
- Other

**3. When you enrolled in Multidisciplinary Studies, were you a traditional student (started right from HS and stayed all 4 years) or were you continuing education (did not start after high school or started and stopped in your degree attainment)?**

*Choose one of the following answers*

*Please choose **only one** of the following:*

- Traditional Student
- Continuing Education

**4. What semester and year did you graduate from the program?**

*Choose one of the following answers*

*Please choose **only one** of the following:*

- Spring 2019
- Summer 2019
- Fall 2019
- Spring 2020
- Summer 2020
- Fall 2020
- Spring 2021

*Employment Acquisition*

*These questions ask you questions relative to your current employment.*



**5. Are you currently employed?**

*Please choose **only one** of the following:*

- Yes
- No

**6. What industry do you work in?**

*Please choose **only one** of the following:*

- Agriculture, Forestry, Fishing, and Hunting
- Mining, Quarrying, and Oil & Gas Extraction
- Utilities
- Construction
- Manufacturing
- Wholesale Trade
- Retail Trade
- Transportation and Warehousing
- Information
- Finance and Insurance
- Real Estate and Rental & Leasing
- Professional, Scientific, and Technical Services
- Management of Companies and Enterprises
- Administrative and support and Waste Management and Remediation Services
- Educational Services
- Health Care and Social Assistance
- Arts, Entertainment, and Recreation
- Accommodation and Food Services
- Other Services (except Public Administration)
- Public Administration

**7. Did you have the job in your current industry before enrolling in the MDS program?**

*Please choose **only one** of the following:*

- Yes
- No



**8. If not, did you have a job in your current industry before graduating from the MDS program?**

*Please choose **only one** of the following:*

- Yes
- No

**9. If not, how long after graduation did you acquire a position in your current industry?**

*Please choose **only one** of the following:*

- < 3 Months
- < 6 Months
- < 12 Months
- >12 Months

*Job Inquiry Questions*

*These questions help to understand your current role at your employer and align it to industry*

**10. What is your current job title?**

*Please write your answer here:*

**11. What is your current annual gross salary?**

*Please choose **only one** of the following:*

- \$20K - \$29K
- \$30K – \$39K
- \$40K - \$49K
- \$50K - \$59K
- \$60K - \$69K
- \$70K - \$79K
- \$80K - \$89K
- \$90K - \$99K
- >\$100K
- Other:



**12. At your organization, what is the current pay range for a person with > 8 years' experience with the same title as yours?**

*Please choose **only one** of the following:*

- \$20K - \$29K
- \$30K – \$39K
- \$40K - \$49K
- \$50K - \$59K
- \$60K - \$69K
- \$70K - \$79K
- \$80K - \$89K
- \$90K - \$99K
- >\$100K
- Other:

**13. What is your desired job title within the next 8 years at your current organization?**

*Please write your answer here:*

**14. At your organization, what is the current pay range for a person with > 8 years' experience with this title?**

*Please choose **only one** of the following:*

- \$20K - \$29K
- \$30K – \$39K
- \$40K - \$49K
- \$50K - \$49K
- \$50K - \$59K
- \$60K - \$69K
- \$70K - \$79K
- \$80K - \$89K
- \$90K - \$99K
- >\$100K
- Other:

**15. When did you determine you wanted to work in this industry?**

*Please choose **only one** of the following:*

- Prior to majoring in MDS
- Prior to graduating from MDS
- After Graduation
- It was where you could obtain a job



- Other

**16. Do you want to remain in this industry for the next 5 years?**

*Please choose **only one** of the following:*

- Yes
- No

Receiving this message indicates that you have completed our Questionnaire and that we owe you a debt of gratitude

I greatly appreciate the time you have spent in assisting in my analysis and commit to both the confidentiality expressed at the beginning and leveraging the information gained to help assist better outcomes for MDS graduates. I will share the results of my research in both a published dissertation as well as an outcome report.

If you have any questions related to this survey or the research I am performing, please do not hesitate to contact me at 210-913-7768 or by email at Gregory.Sansone@UTSA.edu. Thank you again.

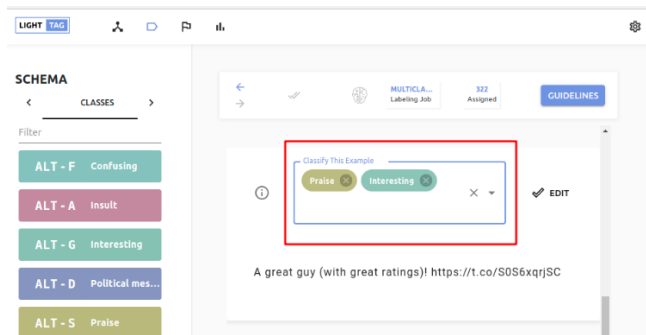


## APPENDIX D – LIGHTTAG TRAINING

Each of the taggers receive login credentials to LightTag (<https://www.lighttag.io>) and allocated a series of syllabi course objectives to tag. Each document appears one at a time. On the left of the screen are the individual tags along with their corresponding hot keys. As shown in Figure 8.

**Figure 4**

*Tagging Screen Example*



*Note.* LightTag. (2021). LightTag.

The tagger places the tags they think best align to the course objectives noted after which they hit next and an additional set of course objectives for a different course appear and the process continues until completion. The taggers receive a definition of each skill along with an example of a course objective from a syllabus that matches that definition.



## REFERENCES

- Abel, J. R., & Deitz, R. (2018). Underemployment in the early careers of college graduates following the great recession. In C. R. Hulten & V. A. Ramey (Eds.), *Education, Skills, and Technical Change* (pp. 149–182). University of Chicago Press. <https://doi.org/doi:10.7208/9780226567945-005>
- Abuselidze, G., & Beridze, L. (2019). Financing models of vocational education and its impact on the economy: Problems and perspectives. *SHS Web of Conferences*, 66, 01001. <https://doi.org/10.1051/shsconf/20196601001>
- Adeola, O. S. (2017). Perceived skill underutilization among big four accounting consultants in the United States [Capella University]. In *ProQuest Dissertations and Theses*. <https://www.proquest.com/dissertations-theses/perceived-skill-underutilization-among-big-four/docview/1894077323/se-2>
- Aso, M., Takamichi, S., Takamune, N., & Saruwatari, H. (2020). Acoustic model-based subword tokenization and prosodic-context extraction without language knowledge for text-to-speech synthesis. *Speech Communication*, 125, 53–60. <https://doi.org/10.1016/j.specom.2020.09.003>
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5, Part 2), 9–49. <https://doi.org/10.1086/258724>
- Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. National Bureau of Economic Research; distributed by Columbia University Press.
- Bell, E., Fryar, A. H., & Hillman, N. (2018). When intuition misfires: A meta-analysis of research on performance-based funding in higher education. In E. Hazelkorn, H.



- Coates, & A. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 108–124). Edward Elgar Publishing.  
<https://doi.org/10.4337/9781785369759.00017>
- Bird, R. B., & Smith, E. A. (2005). Signaling theory, strategic interaction, and symbolic capital. *Current Anthropology*, 46(2), 221–248. <https://doi.org/10.1086/427115>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bravenboer, D., & Lester, S. (2016). Towards an integrated approach to the recognition of professional competence and academic learning. *Education + Training*, 58(4), 409–421. <https://doi.org/10.1108/ET-10-2015-0091>
- Brick, J. M., & Tourangeau, R. (2017). Responsive survey designs for reducing nonresponse bias. *Journal of Official Statistics*, 33(3), 735–752.  
<https://doi.org/10.1515/JOS-2017-0034>
- Bruce, R. F., & Wiebe, J. M. (1999). Recognizing subjectivity: A case study in manual tagging. *Natural Language Engineering*, 5(2), 187–205.  
<https://doi.org/10.1017/S1351324999002181>
- Brutus, S., Aguinis, H., & Wassmer, U. (2013). Self-reported limitations and future directions in scholarly reports: analysis and recommendations. *Journal of Management*, 39(1), 48–75. <https://doi.org/10.1177/0149206312455245>
- Burke, C., Scurry, T., & Blenkinsopp, J. (2020). Navigating the graduate labour market: The impact of social class on student understandings of graduate careers and the graduate labour market. *Studies in Higher Education*, 45(8), 1711–1722.  
<https://doi.org/10.1080/03075079.2019.1702014>



- Burning Glass Technologies. (2015). *The human factor: The hard time employers have finding soft skills*. <http://burning-glass.com/research/baseline-skills>
- Busby, M. (2012). The perceived degree satisfaction and job preparedness of on-campus and distance campus graduates from the bachelor of science in interdisciplinary studies degree program at mississippi state university. *Theses and Dissertations*, 4583. <https://scholarjunction.msstate.edu/td/4583>
- Cadwell Bazata, D. (2020). Inquiry as practice in the implementation of a bachelor of general studies degree in the college of undergraduate studies at the university of central florida. *Electronic Theses and Dissertations*, 16. <https://stars.library.ucf.edu/etd2020/16>
- Cellini, S. R., Turner, N., Brooks, L., Chaudhary, L., Conger, D., Earle, J., Evans, B., Goodman, S., Joyce, T., & Mccubbin, J. (2019). Gainfully employed? Assessing the employment and earnings of for-profit college students using administrative data. *Journal of Human Resources*, 54(2), 342–370. <https://doi.org/10.3368/jhr.54.2.1016.8302R1>
- Chan, C. K. Y., Fong, E. T. Y., Luk, L. Y. Y., & Ho, R. (2017). A review of literature on challenges in the development and implementation of generic competencies in higher education curriculum. *International Journal of Educational Development*, 57, 1–10. <https://doi.org/10.1016/j.ijedudev.2017.08.010>
- Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology*, 37(1), 51–89. <https://doi.org/10.1002/aris.1440370103>
- Cohen, A. M., & Kisker, C. B. (2010). *The shaping of American higher education: emergence and growth of the contemporary system*. Jossey-Bass.



- Complete College America. (2021). *Increasing college completion - Our work*.  
<http://completecollege.org/our-work/>
- Craven, R. F., & DuHamel, M. B. (2003). Certificate programs in continuing professional education. *Journal of Continuing Education in Nursing*, 34(1), 14–18.  
<https://doi.org/10.3928/0022-0124-20030101-04>
- Custom Classifier Metrics - Amazon Comprehend*. (n.d.). Retrieved August 17, 2021, from <https://docs.aws.amazon.com/comprehend/latest/dg/cer-doc-class.html>
- Cutler White, C. (2019). Higher education governance and the attainment agenda: Arrangements with benefits for community colleges? *Community College Review*, 47(3), 219–241. <https://doi.org/10.1177/0091552119852158>
- Davies, M., & Devlin, M. (2010). Chapter 1 Interdisciplinary higher education. In *International Perspectives on Higher Education Research* (Vol. 5, pp. 3–28). Emerald Group Publishing Limited. [https://doi.org/10.1108/S1479-3628\(2010\)00000005004](https://doi.org/10.1108/S1479-3628(2010)00000005004)
- Day, E. A., Arthur, W., & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology*, 86(5), 1022–1033.  
<https://doi.org/10.1037/0021-9010.86.5.1022>
- Deming, D., & Noray, K. (2019). STEM careers and the changing skill requirements of work. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3451346>
- Dorn, C. (2017). For the common good: A new history of higher education in America. In *For the Common Good: A New History of Higher Education in America*.  
<https://doi.org/10.1093/jahist/jay160>



- Dougherty, K. J., & Natow, R. S. (2015). *The politics of performance funding for higher education: Origins, discontinuations, and transformations*. Johns Hopkins University Press.
- Eggenberger, C., Rinawi, M., & Backes-Gellner, U. (2018). Occupational specificity: A new measurement based on training curricula and its effect on labor market outcomes. *Labour Economics*, 51, 97–107.  
<https://doi.org/10.1016/j.labeco.2017.11.010>
- Erickson, M. B., & Winburne, J. N. (1972). *General studies: A trend in higher education in the seventies*. Michigan State.
- Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics*. Sage.
- Fink, Arlene. (1995). *The survey handbook*. Sage.
- Fishman, R., Nguyen, S., Acosta, A., & Clark, A. (2019). *Varying degrees 2019* (Issue September). <https://www.newamerica.org/education-policy/reports/varying-degrees-2019/>
- Föll, P., & Thiesse, F. (2017). Aligning is curriculum with industry skill expectations: A text mining approach. *Proceedings of the 25th European Conference on Information Systems, ECIS 2017, 2017*, 2949–2959.
- Forbes, J. B. (1987). Early intraorganizational mobility: Patterns and influences. *Academy of Management Journal*, 30(1), 110–125. <https://doi.org/10.5465/255898>
- Foss, A., Jennings, M., Francis, D., Knox, E., & Tomlinson, T. (2017). *Hiring demand index (U.S. Patent Application No. 20190102724)*. U.S. Patent and Trademark Office. <https://appft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetacgi%2FPTO%2Fsearch->



bool.html&r=2&f=G&l=50&col=AND&d=PG01&s1=62566364&OS=62566364&  
RS=62566364

- Friedman, H., & Friedman, L. (2018). Does the growing number of academic departments improve the quality of education? *Psychosociological Issues in Human Resource Management*, 6(1), 96. <https://doi.org/10.22381/pihrm6120184>
- Gandara, D., Rippner, J. A., & Ness, E. C. (2017). Exploring the ‘how’ in policy diffusion: National intermediary organizations’ roles in facilitating the spread of performance-based funding policies in the states. *Journal of Higher Education*, 88(5), 701–725. <https://doi.org/10.1080/00221546.2016.1272089>
- Gándara, D., & Rutherford, A. (2018). Mitigating unintended impacts? The effects of premiums for underserved populations in performance-funding policies for higher education. *Research in Higher Education*, 59(6), 681–703. <https://doi.org/10.1007/s11162-017-9483-x>
- Garner, B. R., Gove, M., Ayala, C., & Mady, A. (2019). Exploring the gap between employers’ needs and undergraduate business curricula: A survey of alumni regarding core business curricula. *Industry and Higher Education*, 33(6), 439–447. <https://doi.org/10.1177/0950422219876498>
- Gerstein, M., & Friedman, H. (2016). Rethinking higher education: Focusing on skills and competencies. *Psychosociological Issues in Human Resource Management*, 4(2), 104. <https://doi.org/10.22381/pihrm4220165>
- Gibbons, R., & Waldman, M. (2004). Task-specific human capital. *American Economic Review*, 94(2), 203–207. <https://doi.org/10.1257/0002828041301579>



- Gill, P. M. (2018). The power of perseverance: A qualitative study of factors that impact adult student persistence to graduation (Publication AAI10289607) [Doctoral Dissertation, Indiana University of Pennsylvania]. In *Proquest Dissertations and Theses Global*. <https://www.proquest.com/docview/1948879924>
- Goldring, D. (2017). Pathways for 21st century learners: Integrating industry-based certifications into the marketing curriculum. *Journal of Higher Education Theory and Practice*, 17(1), 33–38.
- Goode, R. B. (1959). Adding to the stock of physical and human capital. *The American Economic Review*, 49(2), 147–155. <http://www.jstor.org/stable/1816110>
- Gottipati, S., Shim, K. J., & Sahoo, S. (2021). Glassdoor job description analytics – Analyzing data science professional roles and skills. *2021 IEEE Global Engineering Education Conference (EDUCON)*, 1329–1336. <https://doi.org/10.1109/EDUCON46332.2021.9453931>
- Green, G., Ballard, G. H., & Kern, D. M. (2007). Return on investment: Assessing a nontraditional, interdisciplinary degree and career impact. *Journal of Continuing Higher Education*, 55(1), 16–26. <https://doi.org/10.1080/07377366.2007.10400105>
- Gromov, A., Maslennikov, A., Dawson, N., Musial, K., & Kitto, K. (2019). Curriculum profile: modelling the gaps between curriculum and the job market. *Edm2020*, 610–614.
- Gruppen, L. D., Burkhardt, J. C., Fitzgerald, J. T., Funnell, M., Haftel, H. M., Lypson, M. L., Mullan, P. B., Santen, S. A., Sheets, K. J., Stalburg, C. M., & Vasquez, J. A. (2016). Competency-based education: programme design and challenges to



- implementation. *Medical Education*, 50(5), 532–539.  
<https://doi.org/10.1111/MEDU.12977>
- Guidelines and Quotas - Amazon Comprehend*. (n.d.). Retrieved August 16, 2021, from  
<https://docs.aws.amazon.com/comprehend/latest/dg/guidelines-and-limits.html>
- Guthrie, K. L., & Callahan, K. (2016). Liberal arts: Leadership education in the 21st century. *New Directions for Higher Education*, 2016(174), 21–33.  
<https://doi.org/10.1002/he.20186>
- Halibas, A. S., Mehtab, S., Al-Attili, A., Alo, B., Cordova, R., & Cruz, M. E. L. T. (2020). A thematic analysis of the quality audit reports in developing a framework for assessing the achievement of the graduate attributes. *International Journal of Educational Management*, 34(5), 917–935. <https://doi.org/10.1108/IJEM-07-2019-0251>
- Handali, J. P., Schneider, J., Dennehy, D., Hoffmeister, B., Conboy, K., & Becker, J. (2020). Industry demand for analytics: A longitudinal study. *Proceedings of the 28th European Conference on Information Systems*, 1–16.  
[https://aisel.aisnet.org/ecis2020\\_rp/11](https://aisel.aisnet.org/ecis2020_rp/11)
- Harrell, F. E. J. (2015). *Regression modeling strategies with applications to linear models, logistic and ordinal regression, and survival analysis* (1st ed.). Springer International Publishing. <https://doi.org/10.1007/978-3-319-19425-7>
- Harris, C. (2020). The earning curve: Variability and overlap in labor-market outcomes by education level. *Manhattan Institute for Policy Research*.  
<https://files.eric.ed.gov/fulltext/ED604364.pdf>



- Heisig, J. P. (2018). Measuring the signaling value of educational degrees: secondary education systems and the internal homogeneity of educational groups. *Large-Scale Assessments in Education*, 6(1), 1–35. <https://doi.org/10.1186/s40536-018-0062-1>
- Herriot, P., & Pemberton, C. (1996). Contracting careers. *Human Relations*, 49(6), 757–790. <https://doi.org/10.1177/001872679604900603>
- Hill, C. B., Kurzweil, M., Davidson, E. P., & Schwartz, E. (2019). *Enrolling more veterans at high-graduation-rate colleges and universities*. ITHAKA S+R. <https://vtechworks.lib.vt.edu/handle/10919/95165>
- Hillman, N. (2016). Why performance-based college funding doesn't work. *The Century Foundation*, 11. <https://tcf.org/content/report/why-performance-based-college-funding-doesnt-work/>
- Hora, M. (2018). Beyond the skills gap: How the vocationalist framing of higher education undermines student, employer, and societal interests. *Liberal Education*, 104(2). <https://www-proquest-com.libweb.lib.utsa.edu/docview/2075723673?accountid=7122&pq-origsite=primo>
- Hora, M. (2020). Hiring as cultural gatekeeping into occupational communities: implications for higher education and student employability. *Higher Education*, 79(2), 307–324. <https://doi.org/10.1007/s10734-019-00411-6>
- Hoyt, J. E., & Allred, E. (2008). Educational and employment outcomes of a degree completion program. *Journal of Continuing Higher Education*, 56(2), 26–33. <https://doi.org/10.1080/07377366.2008.10400150>



- Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., & Mary, J. C. (2020). The condition of education 2020. *Institute of Education Science*, 5(1), 682–688.  
<https://nces.ed.gov/pubs2017/2017144.pdf>
- Iellatchitch, A., Mayrhofer, W., & Meyer, M. (2003). Career fields: A small step towards a grand career theory? *International Journal of Human Resource Management*, 14(5), 728–750. <https://doi.org/10.1080/0958519032000080776>
- Inc., P. (n.d.). *PyPDF2 · PyPI*. Retrieved September 18, 2021, from <https://pypi.org/project/PyPDF2/>
- Jackson, D. (2018). Developing graduate career readiness in Australia: Shifting from extra-curricular internships to work-integrated learning. *International Journal of Work-Integrated Learning*, 19(1), 23–35.
- Jackson, D. (2020). The changing nature of graduate roles and the value of the degree. *Journal of Higher Education Policy and Management*.  
<https://doi.org/10.1080/1360080X.2020.1777634>
- Jurafsky, D., & Martin, J. H. (2022). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). <https://web.stanford.edu/~jurafsky/slp3/>
- Kaira, L. (2011). Using item mapping to evaluate alignment between curriculum and assessment. *Open Access Dissertations*, 318.  
[https://scholarworks.umass.edu/open\\_access\\_dissertations/318](https://scholarworks.umass.edu/open_access_dissertations/318)
- Kambourov, G., & Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, 50(1), 63–115. <https://doi.org/10.1111/j.1468-2354.2008.00524.x>



- Karasek, R., & Bryant, P. (2012). Signaling theory: Past, present, and future. *Academy of Strategic Management Journal*, 11(1), 91–100.
- Kelderman, E. (2019). *The rise of performance-based funding: How colleges are adapting in the new age of accountability*. The Chronicle of Higher Education.
- Kendricks, K. D., Arment, A. A., Nedunuri, K. v., & Lowell, C. A. (2019). Aligning best practices in student success and career preparedness: An exploratory study to establish pathways to STEM careers for undergraduate minority students. *Journal of Research in Technical Careers*, 3(1), 27. <https://doi.org/10.9741/2578-2118.1034>
- Kitto, K., Sarathy, N., Gromov, A., Liu, M., Musial, K., & Shum, S. B. (2020). Towards skills-based curriculum analytics: Can we automate the recognition of prior learning? *ACM International Conference Proceeding Series*, 171–180. <https://doi.org/10.1145/3375462.3375526>
- Kivunja, C. (2018). Distinguishing between theory, theoretical framework, and conceptual framework: A systematic review of lessons from the field. *International Journal of Higher Education*, 7(6), 44–53. <https://doi.org/10.5430/ijhe.v7n6p44>
- Klees, S. J. (2016). Human capital and rates of return: Brilliant ideas or ideological dead ends? *Comparative Education Review*, 60(4), 644–672. <https://doi.org/10.1086/688063>
- Kulshrestha, R. (2019). *A beginner's guide to latent dirichlet allocation (LDA)*. Towards Data Science. <https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2>



- Kunz, J. S., & Staub, K. E. (2020). Early subjective completion beliefs and the demand for post-secondary education. *Journal of Economic Behavior and Organization*, 177, 34–55. <https://doi.org/10.1016/j.jebo.2020.05.015>
- Laderman, S., & Weeden, S. (2020). State higher education finance: Fy 2019. In *State Higher Education Finance*.
- Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, 117(5), 914–940. <https://doi.org/10.1086/648671>
- Leighton, M., & Speer, J. D. (2020). Labor market returns to college major specificity. *European Economic Review*, 128, 103489. <https://doi.org/10.1016/j.euroecorev.2020.103489>
- Li, A. Y. (2020). Performance funding policy impacts on STEM degree attainment. *Educational Policy*, 34(2), 312–349. <https://doi.org/10.1177/0895904818755455>
- Light, A., & Schreiner, S. (2019). College major, college coursework, and post-college wages. *Economics of Education Review*, 73, 101935. <https://doi.org/10.1016/j.econedurev.2019.101935>
- Liu, M., Han, J., Zhang, H., & Song, Y. (2018). Domain adaptation for disease phrase matching with adversarial networks. *Proceedings of the BioNLP 2018 Workshop*, 137–141. <https://doi.org/10.18653/v1/W18-2315>
- Lobo, B. J., & Burke-Smalley, L. A. (2018). An empirical investigation of the financial value of a college degree. *Education Economics*, 26(1), 78–92. <https://doi.org/10.1080/09645292.2017.1332167>
- Lock, E., & Kelly, K. (2020). Ignorance is risk: An exploratory investigation of students' perceptions of their education–employment pathways. *Journal of Teaching and*



- Learning for Graduate Employability*, 11(1), 22–36.  
<https://doi.org/10.21153/jtlge2020vol11no1art894>
- Lyon, A. (1992). Interdisciplinarity: Giving up territory. *College English*, 54(6), 681.  
<https://doi.org/10.2307/377774>
- Malhotra, S., & Godayal, D. (2018). *An introduction to part-of-speech tagging and the Hidden Markov Model*. Freecodecamp. <https://www.freecodecamp.org/news/an-introduction-to-part-of-speech-tagging-and-the-hidden-markov-model-953d45338f24/>
- Manning, C. D. (2011). Part-of-speech tagging from 97% to 100%: Is it time for some linguistics? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6608 LNCS(PART 1), 171–189. [https://doi.org/10.1007/978-3-642-19400-9\\_14](https://doi.org/10.1007/978-3-642-19400-9_14)
- McConnell, F. J. (2017). *Tuition dependency in American public higher education (Publication No. 9949333254102959)* [Doctoral dissertation, University of Georgia]. <http://athenaeum.libs.uga.edu/handle/10724/37851>
- McCray, L. (2011). *An essential academic program: A case study of the general studies program at Louisiana State University in Shreveport (Publication 1651840172)* [[Doctoral dissertation, Louisiana State University]].  
<https://www.proquest.com/dissertations-theses/essential-academic-program-case-study-general/docview/1651840172/se-2?accountid=7122>
- McElrath, K., & Martin, M. (2021). *Bachelor's degree attainment in the United States: 2005 to 2019*. <https://www.census.gov/library/publications/2021/acs/acsbr-009.html>



- McMahon, W. W. (2018). The total return to higher education: Is there underinvestment for economic growth and development? *Quarterly Review of Economics and Finance*, 70, 90–111. <https://doi.org/10.1016/j.qref.2018.05.005>
- Miller, M. A., & McCartan, A.-M. (1990). At the crossings: making the case for new interdisciplinary programs. *Change: The Magazine of Higher Learning*, 22(3), 28–35. <https://doi.org/10.1080/00091383.1990.9937631>
- Mills, M. P. (2021). *The cloud revolution: How the convergence of new technologies will unleash the next economic boom and a roaring 2020s*. Encounter Books.
- Mishra, A. (2019). Amazon comprehend. In *Machine Learning in the AWS Cloud* (pp. 257–274). Wiley. <https://doi.org/10.1002/9781119556749.ch13>
- Moran, J. (2010). *Interdisciplinarity* (2nd ed.). Routledge.  
<https://doi.org/10.4324/9780203866184>
- Muafi, U. (2019). Employer-employee perceptual differences in job competency: A study of generic skills, knowledge required, and personal qualities for accounting-related entry-level job positions. *International Journal of Academic Research in Accounting*, 9(4), 73–82. <https://doi.org/10.6007/IJARAFMS/v9-i4/6660>
- Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments\*. *Journal of Experimental Political Science*, 2(2), 109–138. <https://doi.org/10.1017/XPS.2015.19>
- Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics*, 13(4), 653–677. <https://doi.org/10.1086/298388>
- Neunerdt, M., Trevisan, B., Reyer, M., & Mathar, R. (2013). Part-of-speech tagging for social media texts. *Lecture Notes in Computer Science (Including Subseries Lecture*



- Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), 8105 LNAI, 139–150. [https://doi.org/10.1007/978-3-642-40722-2\\_15](https://doi.org/10.1007/978-3-642-40722-2_15)
- Nguyen-Son, H. Q., Miyao, Y., & Echizen, I. (2015). Paraphrase detection based on identical phrase and similar word matching. *29th Pacific Asia Conference on Language, Information and Computation, PACLIC 2015*, 504–512. <https://aclanthology.org/Y15-1058.pdf>
- Nicholas, J. M. (2018). Marketable selves: Making sense of employability as a liberal arts undergraduate. *Journal of Vocational Behavior*, 109(August 2017), 1–13. <https://doi.org/10.1016/j.jvb.2018.09.001>
- Ntiri, D. W., Schindler, R. A., & Henry, S. (2004). Enhancing adult learning through interdisciplinary studies. *New Directions for Adult and Continuing Education*, 2004(103), 41–50. <https://doi.org/10.1002/ace.147>
- Obama, B. (2009). *Address to Joint Session of Congress*. Obama White House Archives. <https://obamawhitehouse.archives.gov/the-press-office/remarks-president-barack-obama-address-joint-session-congress>
- OECD. (2020). Education at a glance: OECD indicators. In *OECD Publishing*. <https://doi.org/10.5860/choice.41-5419>
- Patterson, K., Watters, C., & Shepherd, M. (2008). Document retrieval using proximity-based phrase searching. *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2008.129>
- Pennington, A., & Stanford, J. (2019). *The future of work for Australian graduates: The changing landscape of university employment transitions in Australia*, Centre for



- Future Work at the Australia Institute, prepared for Graduate Careers Australia.*  
<https://www.tai.org.au>
- Pham, B. (2020). *Parts of speech tagging: Rule-based.*  
[https://digitalcommons.harrisburgu.edu/cisc\\_student-coursework/2](https://digitalcommons.harrisburgu.edu/cisc_student-coursework/2)
- Pologeorgis, N. A. (2019). *Employability, the labor force, and the economy.*  
 Investopedia. <https://www.investopedia.com/articles/economics/12/employability-labor-force-economy.asp>
- Rai, A., & Borah, S. (2021). Study of various methods for tokenization. In *Lecture Notes in Networks and Systems* (Vol. 137, pp. 193–200). Springer Science and Business Media. [https://doi.org/10.1007/978-981-15-6198-6\\_18](https://doi.org/10.1007/978-981-15-6198-6_18)
- Ralston, S. J. (2021). Higher education’s microcredentialing craze: A postdigital-deweyan critique. *Postdigital Science and Education*, 3(1), 83–101.  
<https://doi.org/10.1007/s42438-020-00121-8>
- Rios, A. (2019). Examining the impacts of intrusive advising on the retention and academic success of first-year, at-risk, community college students. *Education Doctoral*. [https://fisherpub.sjfc.edu/education\\_etd/397](https://fisherpub.sjfc.edu/education_etd/397)
- Rockwell, D. (2019). Curtus: An NLP Tool to Map Job Skills to Academic Courses. *Theses and Dissertations*.  
[https://csuepress.columbusstate.edu/theses\\_dissertations/356](https://csuepress.columbusstate.edu/theses_dissertations/356)
- Sarkar, D., Bali, R., & Sharma, T. (2018). Building, tuning, and deploying models. *Practical Machine Learning with Python*, 255–304. [https://doi.org/10.1007/978-1-4842-3207-1\\_5](https://doi.org/10.1007/978-1-4842-3207-1_5)



- Saulnier, B. M. (2017). Using high impact practices to meet employer “soft skill” expectations in computer information systems education. *Issues In Information Systems, 18*(1), 109–117. [https://doi.org/10.48009/1\\_iis\\_2017\\_109-117](https://doi.org/10.48009/1_iis_2017_109-117)
- Schultz, T. W. (1963). *The economic value of education*. Columbia University Press.
- Schweri, J., & Hartog, J. (2017). Do wage expectations predict college enrollment? Evidence from healthcare. *Journal of Economic Behavior and Organization, 141*, 135–150. <https://doi.org/10.1016/j.jebo.2017.06.010>
- Sharma, A. K. (2020). *Understanding latent dirichlet allocation (LDA)*. Great Learning. <https://www.mygreatlearning.com/blog/understanding-latent-dirichlet-allocation/>
- Silos, P., & Smith, E. (2015). Human capital portfolios. *Review of Economic Dynamics, 18*(3), 635–652. <https://doi.org/10.1016/j.red.2014.09.001>
- Simon, J. (2017). *A quick look at natural language processing with Amazon Comprehend*. Medium. <https://julsimon.medium.com/a-quick-look-at-natural-language-processing-with-amazon-comprehend-238b8d9ec11d>
- Sketch Engine. (2018). *POS tags and part-of-speech tagging*. <https://www.sketchengine.eu/blog/pos-tags/>
- Snyder, M., Strategists, H., & Boelscher, S. (2018). *Driving better outcomes: Fiscal year 2018 state status & typology Report*. [http://hcmstrategists.com/wp-content/uploads/2018/03/HCM\\_DBO\\_Document\\_v3.pdf](http://hcmstrategists.com/wp-content/uploads/2018/03/HCM_DBO_Document_v3.pdf)
- Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87*(3), 355–374. <https://doi.org/10.2307/1882010>
- Steele, K. J., VanRyn, V. S., Stanescu, C. I., Rogers, J., & Wehrwein, E. A. (2020). Start with the end in mind: using student career aspirations and employment data to



- inform curriculum design for physiology undergraduate degree programs. In *Advances in physiology education* (Vol. 44, Issue 4, pp. 697–701). NLM (Medline).  
<https://doi.org/10.1152/advan.00167.2020>
- Sublett, C., & Tovar, J. (2021). Community college career and technical education and labor market projections: A national study of alignment. *Community College Review*, 49(2), 177–201. <https://doi.org/10.1177/0091552120982008>
- Sweetland, S. R. (1996). Human capital theory: Foundations of a field of inquiry. *Review of Educational Research*, 66(3), 341–359.  
<https://doi.org/10.3102/00346543066003341>
- Swifttype. (2021). *Search concept: Phrase matching*. <https://swifttype.com/search-concepts/phrase-matching>
- Thouin, M. F., Hefley, W. E., & Raghunathan, S. (2018). Student attitudes toward information systems graduate program design and delivery. *Journal of Information Systems Education*, 29(1), 25–36.
- Traver, A. E. (2016). How do we integrate students’ vocational goals into introduction to sociology curricula, and what are the effects of doing so? *Teaching Sociology*, 44(4), 287–295. <https://doi.org/10.1177/0092055X16643573>
- Umbricht, M. R., Fernandez, F., & Ortagus, J. C. (2017). An examination of the (un)intended consequences of performance funding in higher education. *Educational Policy*, 31(5), 643–673. <https://doi.org/10.1177/0895904815614398>
- United States Census. (2021). *NAICS codes & understanding industry classification systems*. <https://www.census.gov/programs-surveys/economic-census/guidance/understanding-naics.html>



- UTSA. (2021). *Customized multidisciplinary studies degree program (B.A.) / multidisciplinary studies program / UTSA / University of Texas at San Antonio*.  
<https://www.utsa.edu/uc/mdst/customized.html>
- Verma, A., Yurov, K. M., Lane, P. L., & Yurova, Y. v. (2019). An investigation of skill requirements for business and data analytics positions: A content analysis of job advertisements. *Journal of Education for Business*, 94(4), 243–250.  
<https://doi.org/10.1080/08832323.2018.1520685>
- Weisbrod, B. A. (1966). Investing in human capital. *The Journal of Human Resources*, 1(1), 5. <https://doi.org/10.2307/145011>
- Weise, M. (2019). Leveraging a new Rosetta Stone: Deciphering human + technical skills to navigate the future of work. *The Journal of Competency-Based Education*, 4(2), e01186. <https://doi.org/10.1002/cbe2.1186>
- Weise, M., Hanson, A., Sentz, R., & Saleh, Y. (2018). *Robot-ready: Human+ skills for the future of work*. <https://www.economicmodeling.com/robot-ready-reports/>
- Whitcomb, C. A., Khan, R., & White, C. (2016). Curriculum alignment use case for competency frameworks at the Naval Postgraduate School. *INCOSE International Symposium*, 26(1), 105–114. <https://doi.org/10.1002/j.2334-5837.2016.00148.x>
- Wierschem, D., & Mediavilla, F. A. M. (2018). Entry level technology positions: No degree required. *Journal of Information Systems Education*, 29(4), 253–268.  
<https://aisel.aisnet.org/jise/vol29/iss4/5>
- Woodside, J. M. (2018). Real-world rigour: An integrative learning approach for industry and higher education. <https://doi.org/10.1080/00137581.2018.1520685>



*Org.Libweb.Lib.Utsa.Edu/10.1177/0950422218784535*, 32(5), 285–289.

<https://doi.org/10.1177/0950422218784535>

Woolridge, R. W., & Parks, R. (2016). What's in and what's out: Defining an industry-aligned IS curriculum using job advertisements. *Journal of Higher Education Theory and Practice*, 16(2), 105–119. [http://www.na-businesspress.com/JHETP/WoolridgeRW\\_Web16\\_2\\_.pdf](http://www.na-businesspress.com/JHETP/WoolridgeRW_Web16_2_.pdf)

Wright, J. (2016). *The proliferation of general studies degrees*.

<https://www.economicmodeling.com/2016/03/22/proliferation-general-studies-degrees/>

Wu, L., & Lewis, M. W. (2019). How well does undergraduate education prepare college students for the employment outlook? A secondary data analysis of baccalaureate and beyond longitudinal study (B&B). *Higher Education Studies*, 9(4), 181. <https://doi.org/10.5539/hes.v9n4p181>

Xiao, Y., He, L., Ming, H., & Soong, F. K. (2020). Improving prosody with linguistic and bert derived features in multi-speaker based mandarin Chinese neural TTS. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2020-May*, 6704–6708.

<https://doi.org/10.1109/ICASSP40776.2020.9054337>

Yin, W., & Schütze, H. (2017). Task-specific attentive pooling of phrase alignments contributes to sentence matching. *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017 - Proceedings of Conference, 1*, 699–709. <https://arxiv.org/abs/1701.02149v1>



- Yu, Y. (2019). Liberal arts education: The essential foundation for vocational education. *Advances in Social Science, Education, and Humanities Research*, 336, 29–31.  
<https://doi.org/10.2991/icsshe-19.2019.8>
- Yudiono, H., Soesanto, & Haryono. (2018). An industrial competency-based curriculum alignment model. *World Transactions on Engineering and Technology Education*, 16(1), 18–22. <http://www.wiete.com.au/journals/WTE&TE/Pages/Vol.16, No.1%282018%2903-Yudiono-H.pdf>
- Zhou, Q., Zhang, Z., & Wu, H. (2019). NLP at IEST 2018: BiLSTM-Attention and LSTM-Attention via Soft Voting in Emotion Classification. *Proceedings Of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 189–194. <https://doi.org/10.18653/v1/w18-6226>