Chapter 5

Brought, Sought, and Taught: Toward a System of Skill-Driven Applications

Amanda Welsh
https://orcid.org/0000-0002-8065-050X
Northeastern University, USA

Allison Ruda
Northeastern University, USA

ABSTRACT

Skillification is a powerful concept that can drive better outcomes for students, employers, and institutions of higher education (IHEs). Successful use, however, requires IHEs to adopt a systems thinking mindset more than developing a singular taxonomy or exquisite model. Creating a system of skill-driven applications assumes that universities have rich input language that can be translated to skills without extraordinary investment or effort and can do that translation many times over using different algorithms created by different providers as their application needs warrants. Two tests conducted at Northeastern University offer guidance on how to approach this new design: by affirming the feasibility of using syllabi as input for automated skill extraction and identifying data evaluation activity that drives better decisions about third-party partnerships and skill-driven application use.

INTRODUCTION

Continuously building connections between academic curricula and the skills employers need is an imperative for institutions of higher education (IHEs). An overwhelming percentage of workers consider continuous skills development as either important or essential to future career success (Rainie, 2018), and many believe high demand skills correlate to higher paying jobs (Clayton & Torpoe-Sabey, 2021). For those areas of IHEs that primarily serve working adults and historically underrepresented and underserved populations, this imperative is especially urgent. Providing learners with appropriate opportunities to develop and apply skills is not just a trend, it is fundamental to creating a more inclusive prosperity.

DOI: 10.4018/978-1-6684-3809-1.ch005

As IHEs strive to accomplish this mission, a good starting point is to explicitly associate learning content and activities with the skill(s) they address, a process we will follow Lightcast (2021) and refer to as "**skillification**." Once identified, the skills from a curriculum can be used as a connector to other things that have been similarly tagged (Lee, 2005; Sodhi & Son, 2010; Zhang & Zhang, 2012). In one such example, Western Governors University and Central New Mexico Community College defined skills taught in courses which were then were mapped to skills identified by the National Institute of Cybersecurity as meaningful for cybersecurity professionals. As students completed courses, the associated skills they had gained were stored in a Learning Credential Network blockchain created by IBM and used in career counseling as they explored their job potential (America Workforce Policy Advisory Board Digital Infrastructure Working Group, 2020).

What is most intriguing about applications like the one from IBM is that skills appear to be a unit of information that can be extracted from a number of experiences and can power a broad range of solutions. In addition to helping students find jobs relevant to their education, matching skills between jobs and courses can help IHEs keep curriculum current with market needs or guide course recommendations relevant to a student's job goals. Clear articulation of which skills are taught at which points in a course can be used to dissect courses into smaller units that can be stacked differently for different learner populations as context warrants. Identifying skills can facilitate a model for thinking about how to value real-world experience in lieu of classroom learning, which is useful in awarding prior learning credit. It also offers an easy way to connect the curriculum of one IHE to another to support credit transfer in a more streamlined and consistent manner.

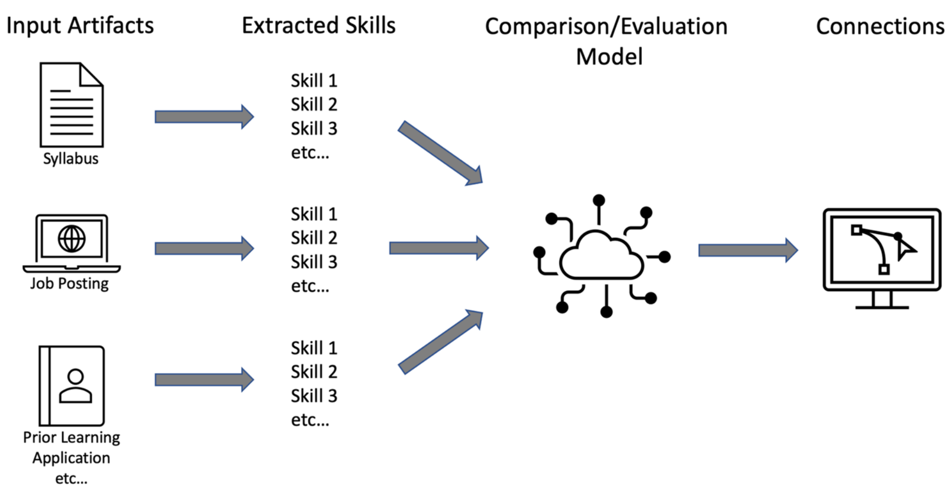
Despite the great potential, however, it is not yet clear that there is widespread use of skill identification for the sorts of applications we have just imagined. Defining and mapping skills in a curriculum can be daunting for an IHE. The level of intentionality that identifying the relationships between skills and coursework calls for is far greater and significantly more time consuming than typical curriculum development approaches (Joyner, 2016; Wang, 2015). Skill identification by faculty is often painstaking and, even worse, occasionally inconsistent (Britton, et al., 2008). Once mapping has occurred, documentation of that work generally lives in disconnected spreadsheets which can be cumbersome to access. Limited access makes it difficult for faculty and students to use skills information on a regular basis. It also makes it less likely that information will be updated regularly, a problem which can be especially damaging in disciplines where knowledge and needed skills are constantly evolving (D'Orio, 2019).

Solutions which seek to mitigate mapping and usage concerns through algorithmic identification of skills and easy access from a database constitute an improvement but are often bespoke projects driven by computer science researchers (Almaleh et al., 2019; Tavakoli et al., 2020). The models which define how lexical terms are elevated to skill status tend to be narrowly focused due to their exploratory nature and are built as discrete standalone solutions that will require ongoing investment from a university to maintain. Increasingly, universities can avoid expensive investment in limited, resource-hungry technology projects by leveraging a burgeoning ecosystem of third-party options. The explosion of online job boards has created rich datasets with skills information driven by actual employer demand. Companies, like Lightcast, have developed systems that parse this information into a skills taxonomy and have built tools to help users sift through connections between courses and jobs. Some organizations offering to store an individual's lifetime of learning, such as iDatify, standardize the inputs they receive into "smart resumes," effectively creating a skills taxonomy. Nonprofit consortia like Open Skills Network or the T3 Innovation Network promote a set of standardized "skill descriptors," itself a comprehensive taxonomy, for use by all network members. In addition, increasing reliance on human resource management software has driven creation of tools to help employers develop their own, proprietary skills taxonomies that inform hiring, development, and advancement decisions (Bersin, 2020).

The problem of relying on a technology solution created by one of these third parties is that each has a reasonable, but vested, interest in considering its skills list as best or most appropriate. The result is a Tower of Babel-like cacophony of similar but nonetheless distinct taxonomies of skills that still require universities to invest time and energy creating crosswalks between them or to make a difficult choice to work with only one solution (World Economic Forum, 2021). Either decision clearly limits the potential for work with a range of partners. Faced with the onerous choice of intense manual effort or resource-hungry bespoke solutions or proprietary taxonomies that are difficult to use in an extended ecosystem, it's not surprising that IHEs may struggle to embrace skillification in meaningful ways.

Responding to the gap between the promise and the execution of skill identification, the College of Professional Studies at Northeastern University (CPS) conducted several tests designed to deepen our understanding of what was needed to support a more strategic, *systems-thinking* approach. A full skill-driven system, shown in Figure 1, consists of artifacts that encode skills, a method to reduce artifacts to a list of skills, some application or model to compare skills from different sources, and an output with a description of relevant connections between artifacts.

Figure 1. Diagram of a skill-driven system

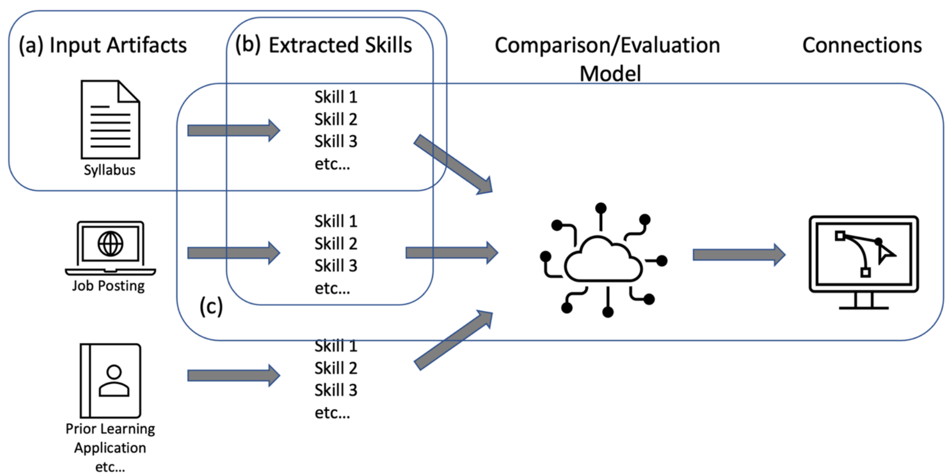


To truly capture its broad potential, such a skill-driven system requires that universities have rich input language that can be translated to skills without extraordinary investment or effort and that they will do this translation many times over with different algorithms created by different providers chosen for their appropriateness for each specific application need. This is a radically different approach from the push toward creating a singular, perfect list of skills that is adopted as currency across the entire education ecosystem. Instead, we imagine a system that is not all too dissimilar from how underlying credit information is translated into a credit score for consumers - dynamically and with some variation in execution by different score creators.

Our inquiry focus, therefore, was not on whether a curriculum can be skillified into one ideal taxonomy or to validate one particular use case; rather, as indicated in Figure 2, we evaluated system components.

In particular, we asked (a) whether the College has a data input that can reasonably serve as the basis for automated skillification, (b) could we gain confidence that the quality and relevance of automatically generated skills was acceptable, particularly without requiring significant human involvement in adjusting the results; and, (c) what additional considerations on skill extraction and modeling are raised in different use cases that might guide how to engage with third-parties and how to select the best partner.

Figure 2. Elements of a skill-driven system examined by tests



This chapter will distill lessons learned from the CPS tests and offer actionable advice and practical suggestions for curriculum developers interested in skillifying the curriculum. For those new to the concept, it offers an exploration of skillification as an enabler for curriculum strategies including modular learning, microcredentialing, and relating workplace experience to curriculum. For those already beginning to explore what skillification might offer, these perspectives may provide insights and examples of steps institutions can take now to pave the way to accelerate more quickly and systematically toward solutions on the horizon.

RESEARCH DESCRIPTION

Our research consisted of two tests conducted in partnership with Lightcast, a leading third-party skillification company. The initial test, designed to answer question (a)[1], evaluated a variety of extant course artifacts, including course descriptions and course-level student learning outcomes found in syllabi, to understand if and how well each resulted in robust skill lists using Lightcast's automated skill extraction solution. Since syllabi are routinely created by faculty for courses independent of a skillification agenda, success in using them for skillification is an empirically less labor-intensive solution for sourcing skill tags for courses. Syllabus evaluation sought to explore a fundamental hypothesis that more input language would correspond to more unique terms and more unique terms would, in turn, translate to more skills identified. To accomplish this, we used a simple bag of words method to quantify the volume and variation of words found in syllabi and correlate that with the number of skills that were subsequently identified by the Lightcast algorithm as relevant to course content.

The second test tackled question (b)[2] and looked at the strength of the connection between skills found in job postings and the course skill lists to validate the quality of the automatically extracted skill information. This work required exploring a few specific points. Notably, did syllabi produce enough skills to achieve reasonable levels of matching to job skills? Were the skills relevant—did the automatically extracted skills cover the same sort of information that was present in job postings, or did syllabi emphasize things employers did not? And finally, was there any benefit from having faculty input on adjusting skill lists to make them more appropriate for use in skill-driven applications? This directly addressed whether there was still a need for resource intensive activity even when using an algorithmic approach. The second test concluded by vetting the automatically generated curricular skills quality in two specific use cases: informing curricular updates and recommending courses to learners based on their job aspirations. Exploration of specific applications was also expected to inform question (c)[3], when to engage with third parties and how to best do so.

Success in both tests would mean that we had identified a scalable, repeatable solution for skillifying our curriculum that could drive different application use cases. Armed with positive answers to our questions, we could further work backwards to identify what language metrics for syllabi corresponded to the desired number of actionable skills and therefore establish minimum benchmarks for syllabus language to guide faculty as new syllabi were written. In this way, we not only sought to validate the potential for using course syllabi as inputs to an algorithmic skillification system, but also to develop a perspective on how to maintain the impact of this input over time.

TEST ONE: EVALUATION OF SYLLABI LANGUAGE

Data for the Initial Test

For the initial phase of work, we created test data sets for three graduate degree programs in CPS, Project Management (PJM), Analytics (ALY) and Regulatory Affairs (RGA). Data consisted of course description language, course outcome language and a section from the syllabus that provided information on weekly topics from all courses required for each degree.[4] While these three syllabus sections are readily available in all CPS syllabi, which follow a standard template, the actual language content is specific to a course and not part of the boilerplate language that is repeated from syllabus to syllabus. Each set of raw language input was cleaned to exclude stop words (“a” or “the”, e.g.), words of three characters or fewer, and special characters. The cleaned language was deemed to have a higher likelihood of containing only words with interesting semantic content.

In addition to data from the syllabi for courses in the test degrees, we also compiled language from course descriptions and course outcomes found in the syllabi for courses in 27 additional graduate degree programs. These degrees cover a wide range of business, social science, and technical disciplines, and correspond to richly varied skills. The aim of this additional data set was to facilitate a slightly deeper dive into whether there was meaningful variation in language and skillification across disciplines.

Using an application programming interface (API) from Lightcast, we then provided the syllabus language as input to the Lightcast skillification algorithm and received back the corresponding skills. Lightcast mines job posting websites for language that they parse to create a dictionary of roughly

30,000 skills (Verougstraete, 2020). The exact nature of the skillification algorithm is unknown to us but was not a concern. An important aspect of creating a system in which we might engage multiple vendors is a recognition that we often will not have intimate knowledge of each skill extraction process. Knowledge of a commercial company's internal workings may reasonably constitute trade secrets that they are disinclined to share. Furthermore, like maintaining a tech platform or guiding faculty through a manual process, evaluating a vendor's code requires an investment of university resources, which we are seeking to minimize by using a partnership model. We will examine the boundaries of accepting the "black box" nature of third-party output as part of our analysis.

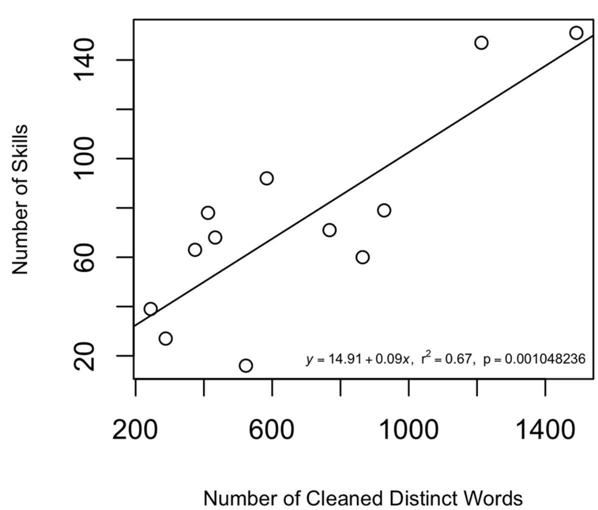
Volume and Variation of Language

Initial examination sought to understand and quantify the volume of input at our disposal. In the three test master's degree programs, the total language taken from all three sections of the syllabi for all courses in each program was equivalent to a 10-15 page paper. While there was some variation, the language for each course corresponded to roughly two paragraphs. An early potential hurdle, that syllabi simply did not contain all that much useful language, was easily cleared.

Additionally, there was a reasonable amount of variation in what words were used in different sections of the syllabus. Only about one fifth of the words in the data for each program was used in more than one section. Practically, this means that all the different sections of the syllabus contributed distinct terms to the final list of cleaned words, and it appears that to create the richest input data set possible, all syllabus language that can be included as input to a skillification algorithm should be.

The power of including as many terms as possible was validated in a comparison of the number of input words and the number of skills extracted from each programs' course descriptions, learning outcomes, weekly topics, and a combined dataset of all three (Figure 3). There is a general increase in extracted skills with a rise in the volume of input terms.

Figure 3. Relationship of number of words to number of skills for each language source



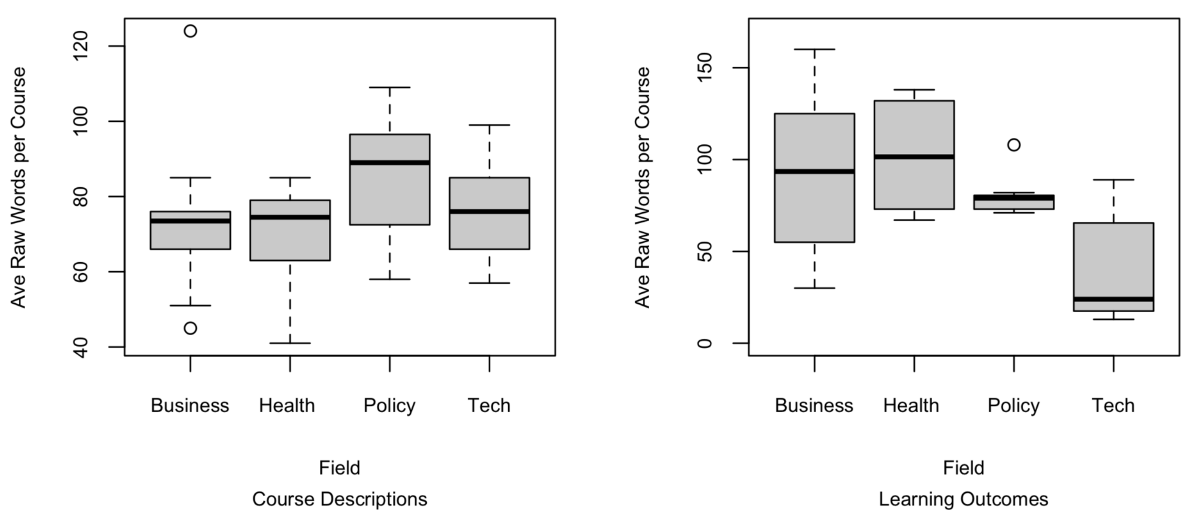
Results from the first test, thus far, confirm that syllabi appear to contain language that can be used for skillification. All sections contribute unique information and should be used if it is practical to do so. In anticipation of building best practices to guide faculty writing new syllabi, we also find support for the foundational premise that more language corresponds to more distinct terms which, in turn, loosely corresponds to more skills extracted.

Variation by Discipline

Given an initial affirmation of the potential of syllabus language, the next step was to determine if the three test programs were reasonably representative of the range of disciplines offered in the College. Some disciplines rely more on specialized vocabulary and a preponderance of field specific technical terms might alter the fundamental nature of the volumetric observations. Comparison of course description and course learning outcome language from the 27 CPS grad programs in our second data set revealed more consistency in the word count of course descriptions than for the program course learning outcomes (course description standard deviation = 17.0 words; course learning outcome standard deviation = 37.0 words).[5] This certainly makes sense since the logistics of publishing course descriptions in a catalogue forces a prescriptive length for this content. There are no such limitations placed upon language which lives only in the syllabus, and it is reasonable to expect more variation from course to course.

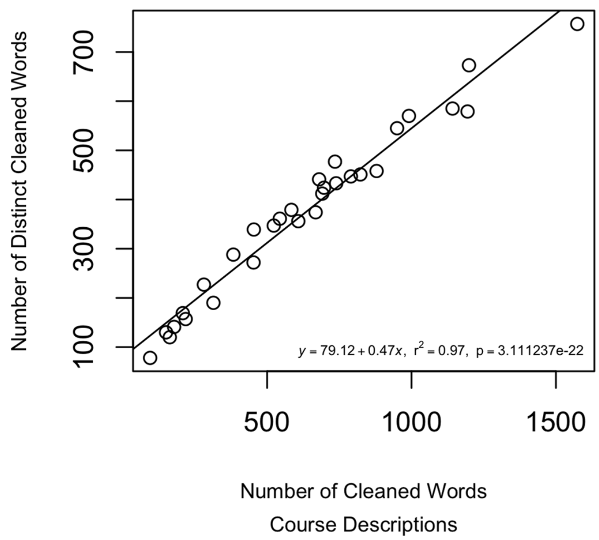
Notably, however, the variation in the number of words used in syllabi was not sensitive to specific disciplines. As shown in Figure 4, courses that can be generally grouped as applying to law and policy are described by above average word count in course descriptions but below the averages for other disciplines in course learning outcome language. Tech related courses average slightly higher word counts than other fields in course descriptions but noticeably less in course learning outcomes. The key here is that there is variation, but not variation that can be explained by the nature of the content being described.

Figure 4. Variation in average raw words for all grad programs grouped by general area



Additionally, there is lack of systemic variation in lexical diversity across disciplines. Comparison of word count for different disciplines shows a definitively clear, strong linear relationship (Figure 5). For every two words in the course description and course outcomes language in any field, the number of distinct words in the cleaned dataset (i.e., where repeated terms were only counted once) will roughly increase by 1.

Figure 5. Relationship between word counts found in course descriptions for all grad programs



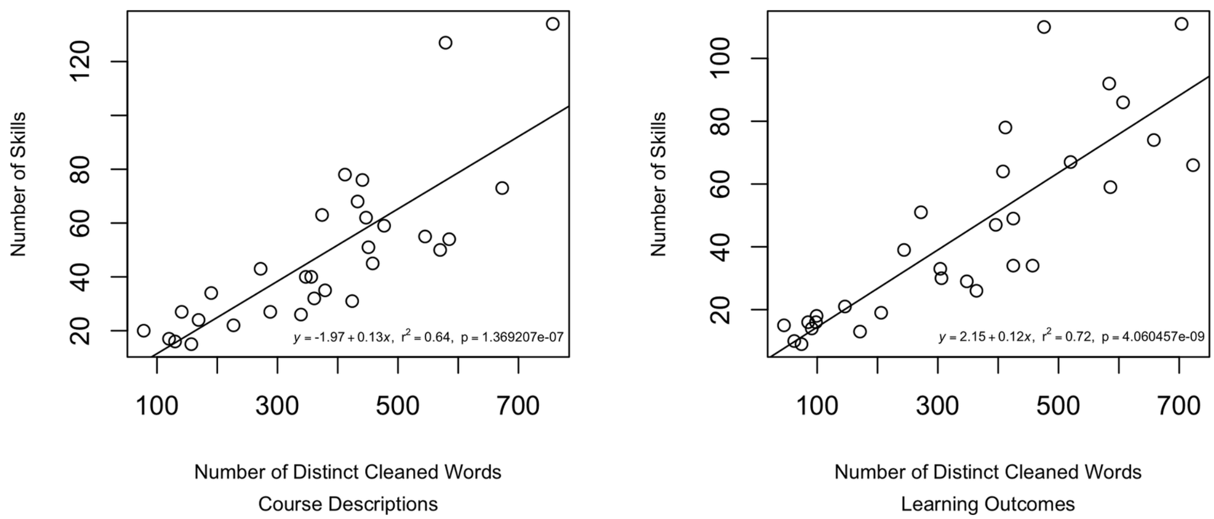
Given both observations, it appears reasonable to imagine generalized guidelines for language volume requirements in syllabi without any discipline specific variation.

Robustness of Skill Extraction

The final, and arguably most important, metric is examination of the number of skills extracted from a given language input. Skill lists were successfully created from all syllabi in all fields, which affirms that there is indeed a signal for skillification broadly in syllabus language. What's more, as shown in Figure 6, the number of skills derived positively correlated to the volume of input language—the more distinct cleaned words in the input data, the more skills extracted.

That said, the correlation between input language volume and skills extracted is not quite as strong as the one between cleaned and distinct words in Figure 5. Whereas cleaned to distinct word counts all fall on or very close to the regression line that best expresses the relationship, the data points of the relationship between input language volume and count of skills are more scattered. Some sit well above or below the regression line, indicating variation among programs that is worth understanding better. Since we have already determined that the input language did not appear to vary in meaningfully identifiable ways, it seemed appropriate to take a step back and consider if the variation might be a function of the skills taxonomy itself.

Figure 6. Relationship between word counts and skills counts for all grad programs



When working with a stable and trusted input source to extract program skills from, the quality of a skills taxonomy is most easily described as the match rate to the input. However, because we are asking an *a priori* question—are course artifacts such as course descriptions, course learning outcomes and weekly class topics good input—we also need to think about the degree to which the taxonomy contents play a role in identifying program skills. The richest course language mapped to a highly limited skill dictionary will still yield a limited result. We need to be confident that the skills taxonomy is appropriately exhaustive in its compilation of skills across the types of programs and job opportunities that should be relevant.

Typically, the quality of an exhaustive measure of something is validated by comparing it to an estimate of the size of the total population—in our current case, a count of the number of the skills that are found in all the jobs in the world. Because no attempt at such quantification has ever been conducted that we are aware of, we are reduced to proxy measures to gauge the sufficiency of any third-party skills list.[6] To be clear, our goal is to be able to create any number of program skills lists by mapping our content to a range of skill taxonomies. It is reasonable to expect that each taxonomy will have its own strengths and weaknesses so the focus here is not to applaud one source over another but to define an evaluation process that any IHE might undertake to assure proper fit with whatever list is used for the task at hand.

To achieve this, calculating the ratio of skills to cleaned distinct terms in input language, which we call "input performance," can be useful. Looking at the "input performance" of syllabus language across all degrees, we find programs in Table 1 for which language from both the course description and learning outcomes sections of syllabi yield fewer skills than might be expected given the volume of the input. Interestingly, these programs cluster in the law and policy area. In contrast, a non-trivial number of technology programs have above average "input performance" scores for both sources, yielding more skills than would be expected given their input language volume.

Table 1. Programs by input performance relative to average input performance across all grad programs

Below average score on all syllabi sections	Above average score on one syllabus section; below average for the other	Above average score on both syllabi sections
Policy: Food Regulatory Affairs Policy: Security and Intelligence Policy: Criminal Justice Policy: Homeland Security Policy: Law and Policy Policy: Global Studies Business: HR Management Business: Public Relations Business: Leadership Business: Communication Business: Nonprofit Management Health: Nutrition	Health: Human Services Health: Healthcare Management Health: Physical Therapy Health: Clinical Trial Business: Finance Business: Accounting Business: Construction Management Tech: Technical Writing Tech: Remote Sensing Policy: Regulatory Affairs	Health: Respiratory Therapy Tech: Geographic Information Systems Tech: Digital Media Tech: Enterprise AI Tech: Analytics Tech: Information Technology Business: Commerce and Economic Development Business: Project Management

Since it is a bit of a stretch to imagine that different faculty drafting individual course syllabi across a set of different but related programs are all comparably poor at using rich, explanatory language, a more likely explanation for the clear clustering of performance by content area is lack of representation in the skill taxonomy itself. It is important to call out that a lower number of skills associated with a given discipline may be appropriate —there may legitimately be fewer discrete skills needed for someone in public service than in high tech. However, even if this is the case, the practical implications of skew in the taxonomy should be considered. As will be discussed shortly, there is some evidence that having fewer skills leads to lower matching levels when matching courses to other skillified artifacts, such as job postings. A sensible response is not to require rethinking the taxonomy—we want to stipulate that this is impractical since a systems approach demands that it be provided by the third-party vendor. Rather, given the success of "more equals more" in the initial evaluation of syllabus language, we propose simply increasing input to capture as many skills as may be available. Until further research determines that lower skill counts are acceptable for matching applications in certain disciplines, faculty teaching in domains with lower skill representation in a taxonomy might reasonably be encouraged to include more language in their syllabi than colleagues in fields with higher representation. It also seems reasonable, in cases where the input performance of certain programs is sufficiently concerning, to explore choosing a different third-party vendor.

We conclude the first test with confidence that the answer to our first question, whether the College has a data input that can reasonably serve as the basis for automated skillification, is yes. Course descriptions, course learning outcomes and weekly topics contained in syllabi offer a rich source of input language for skill extraction. Since syllabi containing these kinds of elements are routinely created by faculty already, universities may find that they have already achieved scale in creating an appropriate input for an automated skillification solution with little additional effort required.

In addition to gaining confidence about a key building block for skill-driven applications, we have also gained some initial understanding of how to make overall system design decisions. Given the correlation between language volume and the number of skills extracted, there is value in defining a minimum amount of language that syllabi contain as a best practice to guide faculty in future syllabus creation. In the case of CPS, we determined that the volume of language in each syllabus section should be above a minimum defined by evaluating the average across all courses in the College. With this requirement, only 3% of input language was incorrectly identified as acceptable when it did not generate the number of skills that we ultimately determined we wanted. Happily, any minimum language requirement does not have to be sensitive to discipline variation outside of demands suggested by skew in the skills taxonomy itself, which can be easily identified by calculating the "input performance" ratio across programs. Using a measure like this, educators can examine input content for patterns to consider as they make decisions about specific adjustments to any basic language requirements they establish.

TEST TWO: EVALUATION OF MATCHING BETWEEN SYLLABI-BASED SKILLS AND JOB-BASED SKILLS

Data for the Second Test

For the second test of the inquiry, we examined one program, Project Management (PJM), to see how well skills from PJM courses matched to skills culled from jobs posted online. We received a file from Lightcast of roughly 12,000 random jobs that included the job description and title along with a list of skills that Lightcast derived from the job description field.

We reviewed job descriptions to identify "true" jobs relevant for the PJM degree holders. Jobs that required a standard industry credential (a Project Management Professional certification offered by the Project Management Institute) or used the term "project manager" in the job description were flagged. Additionally, jobs that used one of 87 keywords deemed indicative of project management responsibilities in the job description were flagged. The flagged jobs were then reviewed manually for appropriate fit, resulting in identification of 363 jobs that were appropriate for PJM degree holders.

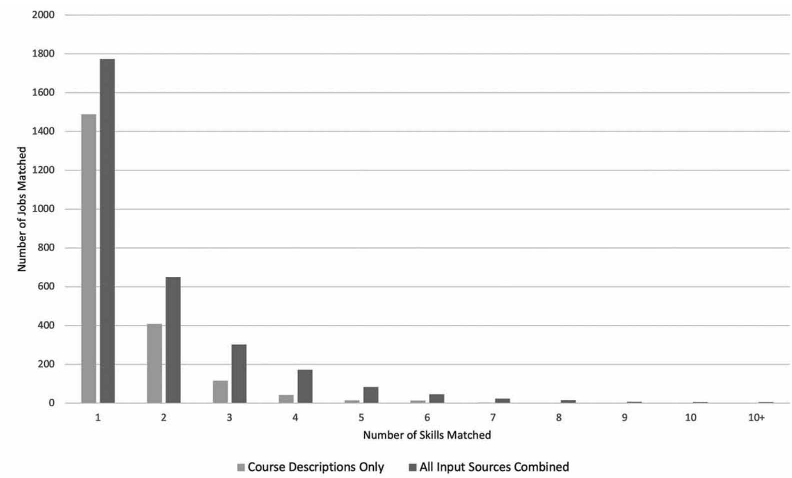
Match Rates

Prior work in skill-driven applications has typically focused on the viability of a given matching solution with less attention paid to the nature of the elements being matched. Since we are most interested in evaluating whether we have an acceptable way to create an appropriate list of curricular skills, we focused on how well our skills exactly matched skills from other items of interest. We can certainly imagine more sophisticated matching models that yield better predictions about reasonable connections between artifacts than what we consider here. There is ample literature that offers insight into a range of relevant improvements (Gugnani & Misra, 2020; Kaur et al., 2020). What is obscured by the more advanced models, however, is an understanding of the fundamental level of quality needed in the data input for an extensible system to achieve results.

Application of a deterministic matching routine returned a preponderance of cases, roughly three-quarters, where no matches between PJM course skills and jobs occurred. This was a good result since a very small subset of jobs were, in fact, relevant to PJM degree holders. When matching did occur, it was typically at a low volume: one to three skills matched in most jobs. The upper bound was 20 matched skills.

Variation in the skills match rate for different syllabi sections affirmed the fundamental assumption that identifying more skills in the curriculum would drive more matches to job content. As shown in Figure 7, the skills derived from using the combined language of all PJM syllabus sections matched more jobs than the skills from the course description language alone, a list about one-third the length of the combined list.

Figure 7. Number of jobs by number of skills matched



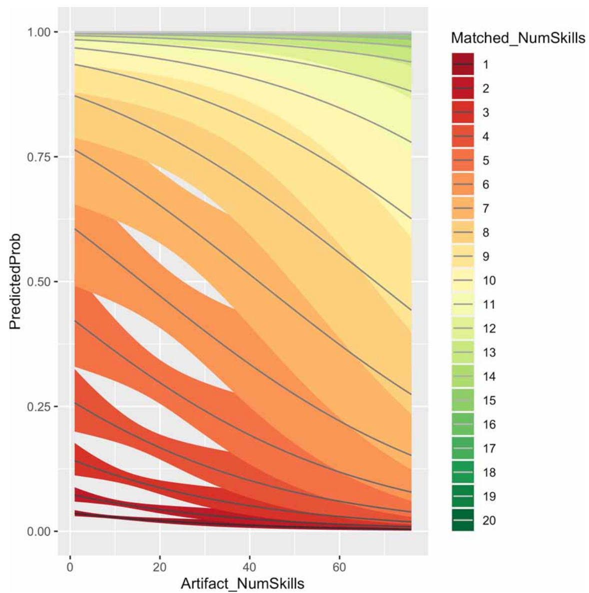
While it is useful to be able to quantify the amount of matching given different skill lists, perhaps the more interesting question is "what amount is enough?" Using coding that identified the true positives in the jobs data (i.e., the jobs the PJM degree did prepare candidates for), a logit model was created to quantify the probability that a job the complete dataset was a true PJM job as a function of the number of matches between PJM curriculum skills and the employer skills. The model results indicate that for each additional skill that matched, the odds of that job being a true project management job increases by roughly a factor of two. The impact of any additional matching, at least in this example, is reasonably large, and further reinforces the assumption that there is value in building out longer skill lists as is feasible.

One challenge of looking only at a count of matched skills is that, as in the discussion of skill extraction relative to taxonomy contents, matching between syllabi and job skills also refers to the intersection of two stimuli—only one of which we control. IHEs are unlikely to ever have a material impact on how employers draft the descriptions of jobs they post. Therefore, we refined our analysis to account for variation that we should understand even if we cannot affect it. The logit model was adjusted to consider the number of skills in each job description that were being matched against, the opportunity for matching, in addition to the actual number of matches. With this refinement, the projected probability of successfully identifying appropriate jobs with varying levels of information could be created (Figure 8; bands indicate the full range of possible values at a 95% confidence interval).

The projections show that to be above a 50% probability of predicting the correct TRUE/FALSE status for a PJM job (i.e., better than guessing), we should look for a minimum of seven curricular PJM skills to match in jobs defined by 40 or more skills. For jobs that are described by fewer skills, the same number of matched skills offers closer to a 75% probability of predicting the right classification. Since CPS programs corresponded to an average of 45 skills per program, our curricular skill information was comfortably more than the minimum matches we might require.

The matching test provides an initial answer to the second question about the quality of our algorithmically generated skills lists. From the basic match rates, we see that there were enough and the right kinds of skills surfacing algorithmically from syllabi that match rates between job and course skills had some level of predictive power. It also affirmed, not surprisingly, that there is a positive relationship between the number of skills matched between two stimuli and the likelihood they have a valid relationship and, consistent with the first test, that volume was important. The more skills extracted from a course artifact, the more matches to jobs.

Figure 8. Probability of successfully identifying appropriate jobs with varying levels of information



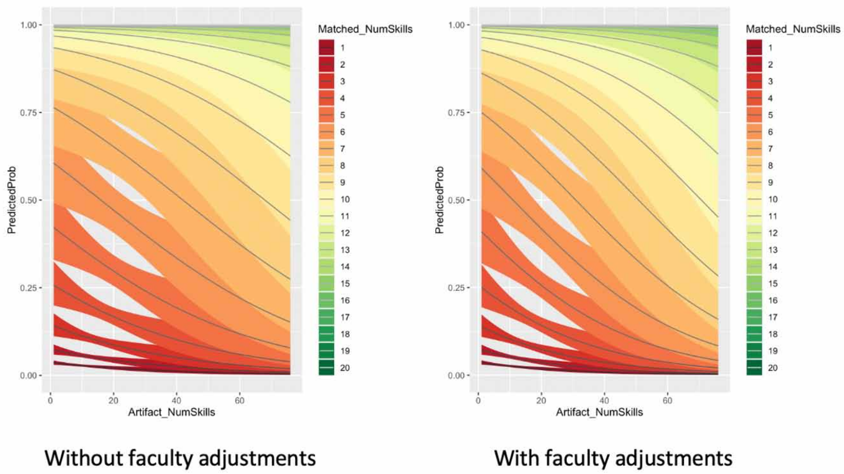
Impact of Faculty Review on Makeup of Skills Lists

As one last point in evaluation of the quality of algorithmically derived skill lists, we turned our attention to how adjustments made by faculty may or may not improve things. We were interested both in how the number of skills for a program might change following faculty review as well as if the types of skills they introduced (or eliminated) resulted in skill lists that were qualitatively different.

We provided skills lists for each course in the PJM program to faculty and invited them to add, move or eliminate skills as they saw fit. From a quantitative perspective, faculty review of the PJM skill output had little impact. Project Management faculty added 11 new skills, removed 5 skills, and adjusted skill assignment to address or eliminate repetition. While this did change the relationship between courses somewhat and arguably offered more precision on how learning accrues through the degree journey, it did not shift any conception of the skills taught in the program.[7] Overall, faculty changed fewer than 8% of the total number of skills.

Given the very limited changes introduced by faculty, it was not surprising that job matching also was not markedly impacted. Matching the job description skills against the faculty-cleaned PJM lists yielded predictive power that was essentially similar to, actually very modestly worse than, matching the lists of algorithmically derived skills (Figure 9). At least for use cases where progression through the degree is not a factor, we determined that the burden of soliciting faculty input did not change the result enough to make the investment warranted.

Figure 9. Comparison of probability of successfully identifying appropriate jobs using skill lists with and without faculty adjustments



In contrast to the outcome in the test with Project Management faculty, the impact of review in a similar test run with faculty in the Organizational Leadership (LDR) program did uncover an interesting finding. In their review, Leadership faculty added 19 new skills or about 11% of the total number of program skills. While this was slightly more than what Project Management faculty did, it still had little to no quantitative impact. What was interesting was that skills introduced by the Leadership faculty in their review did not correspond to skills in Lightcast's dictionary.

In a few cases, the lack of correspondence could be chalked up to variation in tokenization. Faculty chose slightly different language than Lightcast to capture the same concept. While there may be some instinct to solve this problem by coaching faculty to use specific desired vocabulary, this could be counterproductive. Setting aside the pushback such a prescriptive approach would likely engender among experienced faculty, standardization on term usage inside the IHE will still not account for any variation across vendors. From the same content, one vendor may extract "cost management" and a second "budgeting control." Standardizing on one term will still only work some of the time. A better solution is to realize that skill token variation will occur only when we invite faculty to imagine the skill itself. It should completely disappear when we take normal descriptive text—used by faculty in syllabi and employers in job descriptions—and derive skill lists by applying the same extraction process/algorithm to all input. If the skillification system codes a given skill as "cost management," for example, it should reduce the appropriate text only ever to "cost management" and never introduce a different term for the same concept.

In a handful of other instances, the lack of skill correspondence was more semantic in nature. Faculty introduced terms focused on personal development milestones such as "growth mindset" and "critical reflection." Once again, we might imagine that guidance to faculty on language choice could minimize gaps in skill identification. However, it is not clear that the lack of skillification in this case is even a problem. Review of job post language reveals that employers do not reference anything resembling "growth mindset" to a significant degree. Consequently, a taxonomy derived from job postings will not likely include any version of this skill. The fact that faculty articulated a skill that did not exist in the

Lightcast taxonomy will play little role when using that taxonomy to identify appropriate jobs for learners who complete a given course.

Despite not being a common staple of terms used in job postings, the concepts identified by the Leadership faculty are not without merit. It is useful to communicate development of a "growth mindset" as a course objective and the value of possessing one is hard to argue. Indeed, there can be interesting use cases where this *would* be a meaningful skill to identify—in a solution offering modularized learning matched to student-defined rather than employer-defined goals, for example. In this case, thoughtful vendor engagement is probably a better route to solve the taxonomy gap and avoid the need for tapping into precious faculty time. We might reasonably expect that a third-party skill list developed with a purpose more aligned to the use case purpose would contain the skills that our faculty felt were missing.[8]

The current exploration of matching drives confidence in the quality of skills derived from parsing syllabi, without requiring laborious additional review by human subject matter experts/faculty. To the contrary, there is some evidence that matching artifacts subjected to different skillification treatment leads to slightly worse outcomes than matching in a system where both artifacts are treated comparably. In answer to the second question driving the formulation of our systems approach, we conclude that well-written syllabi, on their own, can effectively deliver skills of appropriate quality using LIGHTCAST's skill extraction solution.

As with the first test, this investigation also highlighted important additional considerations about system design. We begin to see the practical need to be sensitive to the nature of the desired use of an application. The absence of personal development goals in the taxonomy flagged by the Leadership faculty was not an issue for a solution which matched course skills to jobs, given how employers write job descriptions, but it could be limiting in other imagined uses. A heightened awareness of the use case considerations can help IHEs identify relevant criteria for vendor review—for example, by surfacing questions about how they construct their skills lists and how that may lead to important gaps in the skills identified or matched. The need for use case sensitivity as a driver in vendor selection becomes all the more evident as we unpack our two sample applications.

APPLICATION IN TWO USE CASES

Guiding Faculty in Adjusting Curriculum

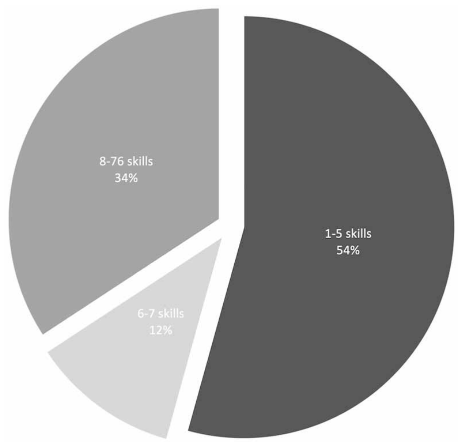
With the fundamental matching activity sorted, we could now turn to question (c), understanding how the matches between jobs and syllabi skills might lead to applications that drive curricular adjustments and course recommendations and what guidance this offers for working with third parties.

There are two actions that might be taken based on a gap analysis between course and job skills: removing skills taught in courses that do not match to skills sought by employers and adding skills asked for by employers but not taught in courses.

Adding to the Curriculum

We have gained confidence in the quality of the syllabus skills. However, as we saw in the discussion of the taxonomy content, the richness of the skills data we want to match against is also important. We need to reconsider the jobs data to achieve reasonable confidence in course-to-job skill matching. Earlier modeling offered guidance on the amount of matching that is desirable to predict job classification in our very simple system, roughly six to seven skills. Unfortunately, as shown in Figure 10, a little more than half of job descriptions in our data set were so short that they corresponded to five or fewer skills. They simply did not contain enough information to support even a marginally reliable classification prediction.

Figure 10. Distribution of number of jobs by how many skills were extracted from them



This creates a reasonable suspicion that a number of relevant cases may not be identified in our classification system even though they should be. Lowering the amount of matching required to identify a relationship will allow for more cases to be identified, but it will also reduce the probability of correct classification and introduce a larger number of false positives. As the illustration below suggests, this can lead to false conclusions and incorrect decisions.

The first column in Figure 11 identifies the skills sought in at least 15 project management jobs but not taught in any of the courses in the PJM degree. Looking at the skills list in the first column immediately prompts the observation that not all entries found in the job posts and flagged by Lightcast are what we might consider skills. Merriam Webster offers a useful definition of a skill as “a learned power of doing something competently; a developed aptitude or ability” (Merriam-Webster Inc, 2022). Following this definition, it is not clear that something like “supply chain” should be included.

As with the discussion that coaching faculty to find perfect skill descriptors may not be a necessary or desirable focus of energy, we might conclude something similar here. Sometimes we use lists of skills to be meaningful. Publishing a list of skills with a course, for example, communicates learning outcomes to students (though we might argue that actual prose descriptions found in syllabi are better for this). Unlike this example which relies on skill descriptors to communicate content, skill-driven applications simply use skills to connect things together. It is not necessary to communicate the contents of two artifacts to be able to conclude that they share similar attributes. It would be nice if a skillification output did have some recognizable bearing on skills as a guiding organizing principle, but for many use cases strictly adhering to the Webster definition is not a *sine qua non* requirement. Provided artifacts being compared are subject to the same skillification treatment, flexibility in what is considered a skill by a given system should not really matter.

Accepting that the skills in column one are essentially adequate if not literally correct, we then turn to consider how we use skill-based matching to identify reasonable changes to the PJM curriculum.

Figure 11. Comparison of job skills that did not match to any skills taught in PJM courses

Skills From True Positive Set	#of Jobs	Skills From Jobs With Low Match Threshold	#of Jobs	Skills From Jobs With Higher Match Threshold	#of Jobs
subcontracting	46	accounting	83	accounting	59
computer_science	30	***merchandising***	40	jira	22
accounting	27	customer_satisfaction	38	strategic_business_unit	21
supply_chain	23	***warehousing***	33	systems_development_life_cycle	20
automation	22	strategic_business_unit	31	computer_science	19
systems_development_life_cycle	22	***packaging_and_labeling***	29	subcontracting	19
customer_satisfaction	22	***financial_analysis***	28	***warehousing***	18
information_systems	19	***financial_statements***	28	***financial_analysis***	18
workflows	19	subcontracting	28	***customer_experience***	17
jira	18	automation	27	automation	17
financial_services	18	***customer_experience***	27	***financial_statements***	17
process_improvements	17	***customer_relationship_management***	27	***mortgage_loans***	17
strategic_business_unit	17	financial_services	27	***packaging_and_labeling***	16
estimators	16	jira	27	customer_satisfaction	16
cyber_security	15	supply_chain	26	***customer_relationship_management***	16
		discounts_and_allowances	25	***merchandising***	15
		…		***risk_mitigation***	15
		nursing	15	supply_chain	15
		occupational_health_and_safety_administration	15		
		strategic_management	15		
		truckload_shipping	15		

Because IHEs will not always have the luxury of being able to manually review jobs data, and instead will need to rely only on models to classify which jobs are relevant, we created two further groups of jobs in addition to the set of jobs we identified as related to project management. One included the jobs from our data that met a skills match threshold low enough to connect PJM coursework to jobs even when the job descriptions were very short. The other included only jobs that met a higher match threshold. The higher threshold connected far fewer jobs to PJM courses (meaning that cases we might legitimately be interested in were not identified) but also resulted in fewer wrong connections. Wrong connections could happen, for example, when a job required some skills that overlapped with project management skills but also required other, more important skills that a project management graduate would not possess. The second and third columns in Figure 11 show the job skills in each of the two additional datasets that did not match to any skill in any PJM courses. Note that the list of unmatched skills at the lower threshold was significantly longer, more than three times the true positive set list. Only a portion of that list is included in the table.

Skills not found in the true positive set but found in jobs positively classified from our matching models at each threshold are shown in bolded italic. With this side-by-side comparison, the potential danger of false positives—predicting a meaningful relationship when one does not exist—becomes apparent. Almost three quarters of the skills in the middle column were not captured in 15 or more jobs in our true positives. Faculty relying on information in the second column might incorrectly be guided to think about adding content related nursing, truckload shipping, and employee safety skills to the PJM degree.

Happily, the output given the slightly higher match threshold has fewer false positives and is more comparable to that of the true positives. From the third column, faculty could conclude that a focus on finance, supply chain, and tech skills should be interesting to develop further. This is roughly the same conclusion to be drawn from looking at the true positive data. However, there is still error we should be sensitive to—some PJM-related jobs were not identified simply because the posts did not contain enough information to generate the required number of matches. Because our understanding of the count of appropriate cases is compromised, our understanding of the amount of demand for a skill in the marketplace is also compromised. Consider, for example, that the true positive data in the first column suggests that demand for accounting and computer science skills, requested in 27 and 30 jobs in our sample respectively, is roughly equal. In contrast, the number of jobs tallied for the third column suggests that computer science is called for in considerably fewer PJM related jobs (19) than is accounting. Program faculty relying on information only from a model might mistakenly prioritize adding more accounting skills to the program over computer science skills.

Removing Skills from a Curriculum

On the other side of the equation, the curricular to job skills matching model can also isolate skills that are taught in courses but enjoyed no matches at all to the project management jobs. A sample of unmatched skills is listed in Table 2.

Table 2. Examples of skills taught in courses but not mentioned by employers in job postings

activity_sequencing	income_tax	project_scoping
activity-based_costing	innovation	quantification
agile_leadership	integration	rate_of_return
agile_management	international_business	requirements_traceability
baselining	kickoff_meetings	resource_leveling
critical_path_method	persona_user_experience	team_building
cultural_diversity	precedence_diagram_method	team_motivation
customer_analysis	prince2	technical_data_management

Upon examination, it appears that many of the unmatched skills represent underlying competencies of more general project management capabilities. Given the richness of the syllabus language compared to the relative lack of job description language, we might reasonably conclude that job descriptions operate at a higher level of generalization than our curricular data. When an employer wants someone with "project management" skills, that employer is implicitly, but not explicitly, requesting skills in "team building," "activity sequencing," and "resource leveling," and if job descriptions included the same level of detail as the syllabus language, we would likely see many of these orphaned skills matching. Here again, our understanding of the quality of the input data drives our understanding of limits on how we should interpret skill matching. We concluded that the true power of a skill-driven solution that informs curricular adjustment lies in considering what skills are present in the jobs data and not in the coursework. The lack of a match to a job skill from a course is not as meaningful.

The positive outcome is that CPS ultimately arrived at a strategy, even given our very simplistic matching model, to gather useful information about general areas that we should consider accentuating in the PJM degree. However, the real takeaway is that we did so with a deliberate understanding of the quality of the input data and how that shifts expectation of what we can learn from our application. In this case, the low number of skills in job postings required us to prioritize precision over recall which means that we can identify skills to consider adding but need to look to other sources of information to understand the degree to which such skills are in demand. Similarly, the general nature of terms chosen by employers in job posts limited our ability to gain useful insight into whether skills taught in courses but not sought by employers were, in fact, not really desired.

We concluded that, just as IHEs would do well to ask questions about how a skills taxonomy is constructed, they can and should ask vendors to explain how their solution is designed to address identifiable aspects of the data inputs, such as data paucity and lack of detail. IHEs would also do well to be clear on the goal of their use case to evaluate their associated tolerance of risk from errors in data interpretation. Developers of an application that lets students filter job opportunities by skills acquired in their degree, for example, might err on the side of providing as many options to students as possible. To do this in our simple model, they would reasonably relax the correspondence criteria so much that any information returned will include false positive hits as well. The student is not necessarily harmed by considering "stretch" jobs and can apply their own intuition about what jobs in the returned list make the most sense for their individual situation. While this kind of tradeoff seems perfectly reasonable in supporting students in a job search, it can lead to negative consequences when considering curricular change. Here, the time and cost of creating new curricula means that decisions to do so should be considered more judiciously. An IHE might determine in this latter use, as we did, that it is more important to favor accuracy over exhaustiveness in finding all the relevant cases.

Providing Course Recommendations

For the second application of algorithmically created skillification data, we wanted to understand if we could meaningfully make course recommendations to someone who was interested in applying for a given job someday. Here, we have the job signal—it is what the student identifies—and only need to call out courses that correspond to the interest defined. This is a fundamentally different use case from curricular adjustment. It is not a big data problem with its reliance on classification probabilities and a need to be sensitive to the type of errors that result. Rather the question in this use case is one of finding differentiated signals. Are course skills sufficiently different from one another to be able to drive a recommendation that is something more specific than "any course in the degree?" For this, we took the 363 jobs that were identified as relevant to PJM degree holders and matched job skills to course skills once again. This time, as a skill matched, the course was noted. In this way, we were able to show a distribution of how many courses matched to skills in each job. The results were modestly encouraging.

There was one skill ("project management") which appeared on the list for almost every course and that anchored the target job to the correct program. At the same time, there were also a fair number of skills that were taught in only one class in the degree. This meant that, after excluding the "project management" skill, a course could generally be recommended based upon the match of a single skill. In this construct, slightly more than a third of the jobs a student might select from our true positive set could be linked to anywhere from one to four course recommendations. While we can imagine ways to improve this result such as clustering skills to achieve more differentiation among courses, the fact that some level of success was possible using skills lists derived algorithmically from syllabi without painstaking manual articulation of a skill by faculty was very positive.

One downside to our solution was that in many instances where more than one course was recommended, the learner was presented with both an introductory and advanced treatment of the same topic.

This highlights a fundamental weakness in the simple matching model: outcomes were created based on the binary presence or absence of a match, with no mechanism to include concepts like mastery. This suggests additional, and intriguing, refinements for skill-driven applications to consider.

For now, the exploration of offering course recommendations, while once again affirming the potential of syllabi as a data input source, is also instructive in helping IHEs develop vocabulary for different types of use cases. We might distill the basic nature of any skill-driven application into one of three types: finding any relevant matches, as in surfacing potential jobs to a graduate searching for work; all accurate matches as in gap analysis that provides guidance on continuous curricular improvement; or the best match, as in the case of recommending a course (or skills-based module) to a learner with a declared goal. While finding matches in large data sets requires awareness of skills volume, identifying a best match requires understanding how differentiated skills in contrasting artifacts are. Sufficient differentiation appeared to happen organically in the data we worked with in CPS.

ADDITIONAL APPLICATIONS AND FURTHER RESEARCH

The investigation by CPS offers positive indications for the viability of a systems approach to maximize skill-driven applications. As we look forward, there are two areas in which additional investigation can be useful in shoring up this initial conviction: testing the current conclusions with other skillification providers and data input sources as well as extending the notion of skills as a unit of information.

We acknowledge that our findings really indicate that what we discuss as possible is possible with Lightcast. Repeating the evaluations we've described across multiple vendors would drive further nuance in understanding how to engage with third parties and build additional confidence in relying on a systems mindset where universities can reasonably expect to work with more than one partner. We suspect that some vendors will be better than others, but we certainly uncovered at least one example where understanding vendor capabilities may be less about "good" vs "bad" and more about which provider is appropriate for a given use case.

Similarly, there is value in extending evaluation to additional data inputs. We believe it is a strong finding that syllabi are useful as they stand. However, this should be further vetted with coursework that is less professionally focused, such as an undergraduate liberal arts curriculum. It is also entirely possible that simply asking faculty to write more when syllabi need to produce more skills holds true because faculty are subject matter experts accustomed to thinking about their work in terms of learning outcomes, a very close relative of skills. We might find that input artifacts from other authors are qualitatively different and more specific guidance on language choice is warranted. It is not clear, for example, if asking employers to post longer job descriptions would address the paucity of extracted skills that we found. This may not be that pressing a question since we suspect we will not ever have an opportunity to meaningfully impact how employers write job adverts at scale, but it is interesting when we consider creation of data inputs that universities can control, such as applications for prior learning credit from prospective students. Early investigation of the language in learner requests for course credit, justified by skills they bring from their work experience, suggests that the "more is more" finding loosely holds. We note, however, that these learners do appear less precise in their use of language than faculty and may benefit from guidance beyond achieving a minimum word count.

Finally, there is great potential in extending the power of skill-driven applications through models that transcend a simple binary presence/absence evaluation of skills. Such a refinement would allow for better understanding of mastery that might translate to more nuanced job matching by level of experience. It may helpfully distinguish introductory from high level courses.[9] Looking at skill clusters or repetition of skill exposure across artifacts may also offer interesting proxies for learning assessment. A student who had the *opportunity* to learn something from a class is more likely to actually have learned it following a defined pattern of exposure, for example.

CONCLUSIONS AND RECOMMENDATIONS

Skillification is a powerful concept that easily piques the imagination for how it might be used to uncover connections between courses, jobs and learner experience which drive better outcomes for students, employers, and IHEs. Real success in utilizing skill-driven applications, however, lies not in developing a singular taxonomy or exquisite model. Rather, it requires IHEs to adopt a systems thinking mindset and work through creation of a solution that has scale and is flexible across a range of potential use cases. The tests conducted at the College of Professional Studies at Northeastern University offer guidance on how to begin to approach the requisite need.

Of greatest importance is the evidence that IHEs likely do not need to invest in additional manual effort to skillify the curricula. While our tests affirmed several assumptions that may seem self-evident, they also offer assurance of the fundamental validity of the proposed approach. Faculty are experts in their fields and, it appears, will naturally use language that encodes the skills they teach as they explain courses to their students. Without any specific coaching, CPS faculty had written syllabi using language of both sufficient volume and variety to generate lists of associated skills that were enough and the right kind to match to jobs, the artifacts we were interested in. The tests offer a promising sign, therefore, that a university can imagine foregoing investment in maintaining a single set of skills associated with courses and instead create them algorithmically as needed with syllabi as input and using the right taxonomy for the purpose at hand. This is a very different model from what has been traditionally followed.

The tests also provide insight into simple and straightforward guidance to faculty to assure that syllabi are optimized for this new approach. The impact of involving faculty in explicit skill identification was modest, potentially even counterproductive. If skills are to function as an effective lingua franca, it appears useful to have the same skill extraction treatment applied to all stimuli input in a given use case. Validation of basic assumptions that more language will correspond to more skills means that, rather than encouraging faculty to encode specific skills or write in a certain way in a syllabus, they simply need encouragement to say more when the existing syllabus is not as potent as desired for skillification. Specific guidance might be that the word count in each syllabus section should be greater than a benchmark defined as the average number of words currently used in syllabi across all courses in the college.

While getting more language from faculty will almost certainly correspond to more skills extracted, the tests did uncover potential variation of skill volume across disciplines. There could be valid reasons for the variation, but it may also carry practical implications that we should be sensitive to. The notion of evaluating "input performance" (the ratio of skills to the input word count) of syllabi can be helpful to identify any skew. Any course syllabus language input which fails to score above an average measure of "input performance" may want to be examined more closely and refinements considered – either in syllabus construction or vendor selection.

With a strong and dynamic solution for curricular data input in place, institutions can turn their attention to how to work with partners, internal and external, on their desired range of uses. What emerged from the tests was a need for IHEs to develop a clear understanding of each skill-driven use case to define how to choose the right partner(s) for it. Our research suggests that developing understanding follows a few steps:

1) Consider the skills taxonomy development. This is a key connector between artifacts in any application and warrants its own distinct investigation with vendors. IHEs should understand how any skills taxonomy is derived and updated. What sources are used? Are there known limitations, such as covering jobs in the US but not in Europe? And fundamentally, is the taxonomy creation method aligned in purpose to the use case? Personal development milestones, for example, will only be included in a taxonomy if they are described in the source material used to develop the list.
2) Evaluate the artifact data inputs. IHEs can work with faculty to assure quality syllabus creation; they can also guide students on the best way to present evidence of prior learning. Investment in defining and driving quality standards for input data that IHEs control is useful. At the same time, IHEs are unlikely to convince employers to draft job descriptions differently or drive syllabi best practices at other educational institutions. In those circumstances, IHEs can focus on developing understanding of the implications of quality considerations. Given the limitations of matching due to the brief, high level nature of job description language, for example, it was consistently true that identifying the full number of appropriate jobs in our test data using a classification model could not be accomplished without introducing an overwhelming number of false positives. Recognizing this and favoring accuracy over exhaustiveness to drive meaningful curricular decisions eliminated all possibility of looking at distribution of skill match rates for guidance about skill importance or priority.
3) Define the type of use case. There is value in simply articulating the desired goal of finding any, all or the best matched outcome. This can lead to a better understanding of tolerated risk from errors in data interpretation as well as uncovering additional data demands. The "best-matched" case, for example, requires a heightened focus on differentiation among artifacts and potentially a need for additional data elements, such as skill mastery, to make differentiation more clear.

In response to previous challenges slowing large-scale adoption of skill-based applications, the CPS tests suggest that IHEs have reason to be confident about using well-written syllabi as a foundational input to skill extraction algorithms. This offers IHEs tremendous freedom to create any number of nuanced skill ontologies with capable companies for a range of well-thought-out applications. Add in a keen eye for assuring that data and its interpretation in each use case are clearly aligned with the objective, and IHEs should find themselves quite well positioned to further their mission through understanding and leveraging connections between a student's professional experience, employer needs, and coursework using a lingua franca of skills.

REFERENCES

Almaleh, A., Aslam, M. A., Saeedi, K., & Aljohani, N. R. (2019). Align my curriculum: A framework to bridge the gap between acquired university curriculum and required market skills. *Sustainability*, *11*(9), 2607. doi:10.3390u11092607

American Workforce Policy Advisory Board Digital Infrastructure Working Group. (2020). *Learning and employment records: Progress and the path forward.* https://www.commerce.gov/sites/default/files/2020-09/LERwhitepaper09222020.pdf

Bersin, J. (2020, January 17). *The war of the skills clouds*. https://joshbersin.com/2020/01/the-war-of-the-skills-clouds-skillscloud/

Britton, M., Letassy, N., Medina, M. S., & Er, N. (2008). A curriculum review and mapping process supported by an electronic database system. *American Journal of Pharmaceutical Education*, *72*(5), 99. doi:10.5688/aj720599 PMID:19214253

Clayton, D., & Torpey-Saboe, N. (2021, October 27). *Student outcomes beyond completion: National findings from the 2021 Strada alumni survey* (Public Viewpoint). Strada Education Network. https://cci.stradaeducation.org/pv-release-oct-27-2021/

D'Orio, W. (2019, June 24). *With skills mapping, colleges create a "universal language" to explain value*. Higher Ed Dive. https://www.highereddive.com/news/with-skills-mapping-colleges-create-a-universal-language-to-explain-valu/557422/

Daubney, K. (2020). Employability is not inimical to good learning: Rearticulating school academic curricula to surface their employability value. *Journal of Work-Applied Management*, *13*(1), 103–116. doi:10.1108/JWAM-08-2020-0041

Lightcast. (2021). *Skillify*. https://kb.Lightcastdata.com/glossary/skillify/

Gugnani, A., & Misra, H. (2020). Implicit skills extraction using document embedding and its use in job recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(08), 13286–13293. doi:10.1609/aaai.v34i08.7038

Joyner Melito, H. S. (2016). Curriculum mapping: A before-and-after look at faculty perceptions of their courses and the mapping process. *Journal of Food Science Education*, *15*(2), 63–69. doi:10.1111/1541-4329.12085

Kaur, S., Bhardwaj, V., & Singh, R. K. (2020). A novel investigation and extracting of critical factors affecting employability. *Journal of Xi'an University of Architecture & Technology*, *12*(5), 3018–3026.

Lee, C. K. (2005). Analysis of skills requirements for systems analysts in Fortune 500 organizations. *Journal of Computer Information Systems*, *15*, 544–554.

Merriam-Webster, Inc. (2022). Skill. In *Merriam-Webster Dictionary*. https://www.merriam-webster.com/dictionary/skill

Rainie, L. (2018, June 18). *Skill requirements for future jobs—10 facts* [PowerPoint slides]. International Telecommunications Union's Capacity Building Symposium, Santo Domingo, Dominican Republic. https://www.pewresearch.org/internet/2018/06/19/skill-requirements-for-future-jobs-10-facts/

Sodhi, M. S., & Son, B. G. (2010). Content analysis of OR jobs advertisements to infer required skills. *The Journal of the Operational Research Society*, *61*(9), 1315–1327. doi:10.1057/jors.2009.80

Tavakoli, M., Mol, S. T., & Kismihók, G. (2020). Labour market information driven, personalized, OER recommendation system for lifelong learners. *Proceedings of the 12th International Conference on Computer Supported Education*, *2*, 96-104. 10.5220/0009420300960104

Verougstraete, R. (2020, September 16). *Lightcast joins network to accelerate skills-based education and hiring*. https://www.economicmodeling.com/2020/09/16/open-skills-network-partnership/

Wang, C.-L. (2015). Mapping or tracing? Rethinking curriculum mapping in higher education. *Studies in Higher Education*, *40*(9), 1550–1559. doi:10.1080/03075079.2014.899343

World Economic Forum. (2021). *Building a common language for skills at work—a global taxonomy.* https://www.reskillingrevolution2030.org/reskillingrevolution/wp-content/uploads/2021/01/Skills-Taxonomy_Final-1.pdf

Zhang, S., Li, H., & Zhang, S. (2012). Job opportunity finding by text classification. *Procedia Engineering*, *29*, 1528–1532. doi:10.1016/j.proeng.2012.01.167

KEY TERMS AND DEFINITIONS

Classification Modeling: Any of various statistical and machine learning techniques used to assign a test item to a certain class.

Curriculum Mapping: The process of defining skills taught in a curriculum.

Data Paucity: An issue in data sets where some variables may be lacking detail or content.

Regression Analysis: A statistical technique that compares the relationships between variables.

Skill Taxonomy: An organized structured list of skills representing a universe of possible skills.

Skillification: The process of reducing text found in things like job postings, resumes or course syllabi to a list of representative skills.

Systems Thinking: An approach to problem solving that considers the totality of the solution as opposed to a focus on one discrete piece or outcome.

ENDNOTES

1 (a) whether the College has a data input that can reasonably serve as the basis for automated skillification,

2 (b) could we gain confidence that the quality and relevance of automatically generated skills was acceptable, particularly without requiring significant human involvement in adjusting the results

3 (c) what additional considerations on skill extraction and use in modeling are raised in different use cases that might guide how to engage with third-parties and how to select the best partner.

4 For this review, we focused only on each degree's required courses and department electives. We excluded information from possible electives provided by other programs.

5 It is interesting to note that the average length of course descriptions for Undergrad courses in CPS Professional Programs is 59 (with a similar standard deviation of 16.4) and for course learning outcomes is 76 (with a standard deviation of 44.4). There is a similarity to the patterns which drives confidence.

6 As Lightcast did with us, the author of a taxonomy may be willing to provide statistics on distribution of terms, which is also a useful guide to potential bias. However, this makes a generous assumption that skill category assignment by the vendor corresponds to how the university would group skills and still does not address the fundamental question of the suitability of representation.

7 To assess if faculty review impacted the skill to course mapping, we considered two views to show the relationship between courses: a simple correspondence analysis and a dendrogram of hierarchical clustering by terms. That work is not discussed in detail here but dendrogram plotting of courses clustered using the Lightcast skill list show a few outliers and a more general clustering of the remaining courses. The relationships created with the data reviewed by faculty shows more nesting of courses.

8 We do note that the "soft skills" called out by Leadership faculty may constitute a special skill category and look for more investigation into this specifically, such as found in Daubney (2020).

9 Workday (www.workday.com) and LinkedIn (www.linkedin.com) are very active in pushing these types of analysis forward.