DOI: 10.34641/ctestem.2022.473

Developing a Continuous, Rather Than Binary, Classification for Measuring STEM Jobs

Ted CARMICHAEL[1*], John STAMPER[2], John CARNEY[3]
[1]TutorGen, Ft. Thomas, KY, USA
[2]Carnegie Mellon University, Pittsburgh, PA, USA
[3]MARi, Alexandria, VA, USA

ted@tutorgen.com*, jstamper@cs.cmu.edu, john.carney@mari.com

ABSTRACT
This paper presents our review and synthesis of the literature on STEM classification, and our results for a novel approach towards understanding, categorizing, and tracking STEM attributes in the workplace. We found two deficiencies in the way STEM is traditionally discussed, which we attempt to address in this work. The first is that the key components of STEM tend to be discussed holistically in the literature, rather than discretely as Science, Technology, Education, and Mathematics. The second is that our ability to track changes in S.T.E.M. concentrations in the workplace, both geographically and temporally, is underdeveloped. Further, we have found that this second deficiency is due, in part, to how STEM occupations are categorized; i.e., "STEM" tends to be a binary designation, rather than measured on a continuum for each job, and each component of S.T.E.M. It is also due to the lack of a "gold standard" measurement of the quantity of S.T.E.M. for all occupations. Here, we present a novel approach for machine learning algorithms using a "bag of words" method. These algorithms are trained on a small selection of Standard Occupational Classification (SOC) occupations, using ratings for each component of S.T.E.M. as the exemplars on which to train (SOC 2019). Recognizing that such a classification scheme is new, and that one of the goals of this project is to solicit Subject Matter Expert (SME) feedback, the resultant model of S.T.E.M. measurements across these occupations is designed to easily incorporate multiple distinct models and alternative approaches.

KEYWORDS
STEM classification, machine learning, Standard Occupational Classification (SOC)

1. INTRODUCTION
This report is one part of a greater research effort started through the U.S. Department of Education (DOE), Office of Career, Technical, and Adult Education (OCTAE) to create a STEM Index to quantify the elements of S.T.E.M. (i.e., science, technology, engineering, and math) in the North American Industry Classification System (NAICS), the Standard Occupational Classification system (SOCs), and possibly to the level of individual jobs. This statistical model will be designed to be applied to government employment and industry datasets to provide STEM educators and administrators information related to the amount of Science, Technology, Engineering, and Math within a range of occupations included in both SOCs and NAICS. One of the most valuable applications of this STEM index will be the future ability to match state and local programs and course completions to quarterly regional Bureau of Labor and Statistics (BLS) economic activity, thus allowing greater alignment between course offerings and regional in-demand jobs. (Note: our paper follows the standard nomenclature of using "STEM" when speaking about programs and initiatives holistically, and "S.T.E.M." when referring to the individual component.)

One key finding from our literature review is that the STEM categories are not generally broken down into composite classifications. Further, all occupations tend to be categorized as either STEM or not, based on broad definitions, and in a binary manner, rather than by quantifying the amount of STEM knowledge needed for a job or determining a list of required STEM skills and competencies. With the current classification method, some occupations will be included when they should not be, while others are excluded even if they have a significant STEM component. "Broad groupings can only give broad estimates and are not useful for targeted workforce policy. [...] Problematic in the current discourse on the value and impact of STEM discipline-related skills is the use of the STEM acronym to encompass a wide variety of different concepts in instances where a more precise or appropriate term is needed" (Siekmann & Korbel, 2016).

Finally, the current classification scheme itself resists timely updates. To assign classifications across such a comprehensive list of occupations is quite labor intensive. And there is a significant need to be able to compare data, both "across agencies and organizations," as well as over time, in order to "maximize the comparability of data" (SOC, 2019). This hampers our ability to make systemic changes in the classification and quantification of the STEM content across occupations.

Another pattern that we have found, not only in the STEM literature, but also found in many STEM-focused education websites, program initiatives, and advocacy resources, is a general trend towards defining STEM more broadly, rather than with more precision. That is, there is a marked tendency to discuss STEM skills holistically; and to promote STEM education, training, and directed resources by listing many that are more correctly classified as "foundational" skills. Foundational not just for STEM occupations, but for a wide variety of non-STEM occupations as well. These are skills such as: creativity, organization, communication and teamwork, problem solving, and critical thinking. These skills remain far from unique to STEM occupations. Further, in some STEM resources, even more general traits and habits are discussed in this context, such as: persistence, flexible thinking, empathy, engagement, and metacognition (Costa, 2008; Asunda & Weitlauf, 2018).

From the Department of Education report: STEM Education Strategic Plan 2018: “Over the past 25 years, STEM education has been evolving from a convenient clustering of four overlapping disciplines toward a more cohesive knowledge base and skill set critical for the economy of the 21st century. The best STEM education provides an interdisciplinary approach to learning, where rigorous academic concepts are coupled with real-world applications and students use STEM in contexts that make connections between school, community, work, and the wider world. Leaders in STEM education continue to broaden and deepen its scope and further transcend the fields of study beyond just a combination of the four disciplines to include the arts and humanities. Modern STEM education imparts not only skills such as critical thinking, problem solving, higher order thinking, design, and inference, but also behavioral competencies such as perseverance, adaptability, cooperation, organization, and responsibility.” This, as evidence that STEM is more and more often discussed as a synthesized field rather than as individual components; and more than just S.T.E. and M. skills are referred to as part of the STEM curricula.

This literature review and project takes as its working hypothesis the possibility that the pendulum has swung too far. Perhaps what is needed is a new focus on the individual pieces that comprise STEM; or, further, the four components of S.T.E. and M. (The convention is, when discussing STEM as a unified subject, the acronym is written as a single word; conversely, when emphasis is on the individual components, the acronym is written as “S.T.E.M.”) This more granular focus may renew our understanding of all the skills, abilities, aptitudes, and competencies that go into performing STEM tasks and occupations. By more precisely measuring the competencies that comprise S.T.E.M., we might better track these skills across a continuum, and determine their changing proportions over time and across all occupations.

2. BACKGROUND

In recent years, educational and vocational professionals have sought to define STEM core competencies in a more holistic way. While emphasis on the sciences dates back at least to the decade that saw the formation of both NASA and the NSF, the acronym STEM, and the emphasis on these four fields - Science, Technology, Engineering, and Mathematics - emerged in the late 1990’s and early 2000’s (Chute, 2009). Coined by Dr. Judith Ramaley, who served as the assistant director of the Education and Human Resources Directorate at the NSF, “STEM” was defined as “an educational inquiry where learning was placed in context, where students solved real-world problems and created opportunities—the pursuit of innovation” (Daugherty, 2013).

What’s more, there is a persistent belief that “only some kids can really learn math and science to high levels” (Chute, 2009). Here is Dr. Nancy Bunt, program director of the Math & Science collaborative in the Allegheny Intermediate Unit in Pittsburgh, PA: “There is this very strong belief out there on the part of parents and the part of some educators and society as a whole: If I wasn't good at math, my kids don't have a chance of being good at math. It's a gene thing. They'd never say that about reading. There is an assumption that everyone needs to learn to read.” There is even data to support the idea that a strong emphasis on the four S.T.E.M. components, as well as their integration into the unified “STEM” category, with meaningful overlap and synergy, will not only increase the number of students who pursue the sciences, but also positively influence the number that pursue any bachelor’s or post-secondary degree (Chute, 2009).

However, definitions of STEM in terms of what fields are included have tremendous consequences for US (and international) policy, funding decisions, resource allocations, and an incredible variety of educational initiatives and workforce development programs. Which fields of study are included can impact anything, from which among the millions of undergraduate and graduate students are supported by the NSF, to who will be eligible to receive student visas, to which programs receive extra resources as fields of designated national interest.

Consider, for example, the STEM Educational Act of 2014. Passed in July of that year, this act has only two stated purposes. Aside from a few minor changes in wording of the existing statues, this law was written solely “To define STEM education to include computer science, and to support existing STEM education programs at the National Science Foundation.” And the second of these was not really a change; rather, it was just a continuation of previous policy: the bill states that “The Director of the National Science Foundation, through the Directorate for Education and Human Resources, shall continue [emphasis added] to award competitive, merit-reviewed grants to support” STEM learning environments, learning outcomes, engagement, and research in STEM education.

In other words, so crucial to funding decisions, programs, and resource allocation was explicitly adding “computer science” and computational thinking to the standard, government-wide definition of STEM that this act was passed by both houses of Congress and signed by the President, to make this inclusion clear and give it the force of law. Definitions matter.

As such, there is continual pressure, from stakeholders, curriculum managers, and workforce development programs, to include their individual fields and domains in standard definitions of “STEM,” in order to remain relevant (and funded). Conversely, there is pressure in the other direction to keep STEM a semi-exclusive and useful definition: becoming too broad would render it nearly meaningless. This reality has created tremendous ambiguity in the STEM label, such that it is often defined and redefined based on various needs, creating the situation where there is little agreement, across all stakeholders, in a precise definition.

The power of the STEM acronym comes, in part, from this ambiguity, as it can mean all things to all people. As a loosely defined and malleable concept, "STEM" gains wider acceptance and becomes the focus of more initiatives, more funding, and greater involvement of key stakeholders. However, this same ambiguity, that allows much of this national (and international) attention, also drives misunderstandings and confusion over what, exactly, STEM refers to. "Whether the acronym is understood and fashionable outside these education groups is not well known. What is known is that the acronym and associated term is not well-defined, even within groups that make heavy use of it" (Daugherty, 2013, citing Storksdieck, 2011).

But such ambiguity has consequences. The NSF's "STEM Education for the Future: A Visioning Report" from 2020 makes the point that access to STEM education varies "across zip codes and income levels," as well as among underrepresented groups (Education & Human Resources, 2020). It further articulates certain key priorities, in order to meet the challenges in STEM education and workforce development. One of these priorities is to level the playing field for STEM educational opportunities. And so, priority one is to increase opportunities for those being left behind.

Related to this, another listed priority emphasizes the need to motivate improvements across the board: to continue to strive for advances in our national capabilities; to promote increased invention and innovation; and to fill the demand for the high-tech, high-quality jobs of the rapidly approaching future. This is because, according to the National Science Board's Science and Engineering Indicators 2018, Americans' basic STEM skills have only modestly improved over the past two decades. And, they continue to lag behind many other countries. Further, according to the indicators, from 2006–2015, American 15-year-olds still tended to score below the international average in mathematics skills, and at or slightly above the international average in science skills. These are important areas to address.

However, having an insufficient definition of S.T.E. and M., and not going far enough in classifying STEM occupations vs. non-STEM occupations, means that, not only are we not accurately measuring these problems, but we lack the data to properly target interventions, and do not have the means to measure or judge the success or failure of those interventions.

What we need to do—the motivating premise behind this project—is to move away from the binary classification of STEM vs. non-STEM jobs, and instead focus on the level of STEM skills and abilities that are found within all jobs. This change in emphasis would improve measurement of STEM skills in our workforce in a way that is more granular, and thus provide a better understanding of which STEM skills and competencies are increasing in demand, so that we might better meet current and emerging workforce needs.

Our research in this project will look specifically at novel ways to do exactly this. The current definitions simply are not granular enough, and are not updated frequently enough, to allow more precisely targeted interventions. It is our hope that not only will we be able to measure and track a continuum of STEM-related skills needed for more precise categories of occupations, but also to quantify the changes in demand for these skills over time.

3. APPROACH

In order to develop a scalable, algorithmic method for quantifying each of the S.T.E.M. components across all occupations, two things are needed. First, we need a ranking of S.T.E.M. content for each occupation on which to train that algorithm. The ranking system should have both a theoretical and valid base. Second, we need a model to predict the rankings. Our approach to the model is that a single model is unlikely to provide a highly accurate result across all jobs, so we rely on a ensemble approach (Niculescu-Mizi et al., 2009) described in section3.2. and an initial first predictive model for the ensemble described in section 3.3.

3.1. *Rating System*

When considering how much Math, or how much Engineering, is required for a particular job, there are really two different senses for how to quantify this requirement: level of expertise needed to perform the job; and level of intensity, or how much time is spent on that activity. We chose to use "level of expertise" only in quantifying the amount of S.T.E. or M. required. So, for example, a retail job where one is expected to add and subtract figures all day would still have a low level of Math required, perhaps a 1 or a 2 on the 9-point rating scale.

Table 1. STEM Level Classification* Ratings.

Rating	Education Equivalent
0	None
1	Middle School
2	High School
3	Certification
4	Assoc. Degree (2 yr)
5	Bachelor's Degree (4 yr)
6	Master's Degree
7	Doctoral Degree
8	Everything Above

*These designate minimum levels (or equivalent) of knowledge acquisition via degrees and/or years of experience.

Conversely, a job that requires Calculus and Differential Equations knowledge would be rated high, even if the employee was not expected to draw on those skills very often. In this way our system quantifies the level of skill needed to perform the job, so that qualifications and how they change over time are captured.

The STEM level classification was done on a 9-point scale, with the level of rating roughly equivalent to years of education as a proxy for knowledge requirement to perform the job. Table 1 is included to illustrate these rating levels.

3.2. *Ensemble Approach*

In order to achieve robust results, we have implemented a workflow that includes an ensemble methodology (Niculescu-Mizi et al., 2009, Yu et al., 2010). This methodology recognizes that any single model might not be highly accurate across all possible predictions, and by combining models using weighted averages a better result can be achieved. Our ensemble scoring methodology can be seen in figure 1, and is available in source code made available in our public repository https://github.com/jcstamper/CTE-STEM.

$$\mathrm{IndexScore}_{s,t,e,m}(\mathrm{SOC}) = M_1\omega_1 + M_2\omega_2 + \ldots + M_x\omega_x$$

For Models [M] and Weights [ω]

$$\text{where } \sum \omega = 1$$

Figure 1. Ensemble workflow for weighting models.

3.3. *Model Prime*

We proposed and implemented an initial model for our workflow that we named M_{Prime}. This model uses NLP methods in a bag of words approach from a data source that contains job descriptions and compares word embeddings against a weighted vector for each of the formal topics of Science, Technology, Engineering, and Math. The vector distance is then calculated and normalized to our ranking scale.

For our first pass, we derived the vectors for the formal topics from a list of topics curated by STEM experts, and our initial data source for the job descriptions came from O*NET.

4. RESULTS

We ran our initial model on 82 jobs exported from the O*NET repository. Note that because we do not have a true gold standard for our topic vectors, our weightings were not trained in any way. Having more data from additional models or from experts could help us better train the models in the future. The results, however, were promising. Although experts were able to find potential disagreements in the results, we compared Cohen's Kappa statistic on 25 jobs classified by two experts and our system. The results between the two experts was .55 and between the two experts and our system were .52 and .44 respectively. While all these values suggest a weak level of agreement, it also shows that our values were not far off.

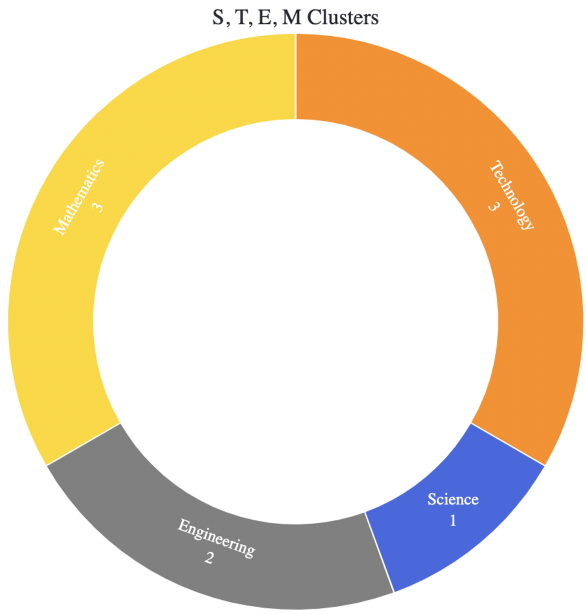

Figure 2. Index for SOC 27-1022, Fashion Designers, as a pie chart visualization.

We created an interface and web portal to inspect our results, which are available here in our web application https://share.streamlit.io/jcstamper/cte-stem/main/AppFinal.py. An example of selected jobs with several visualizations can be seen in Figures 2 and 3.

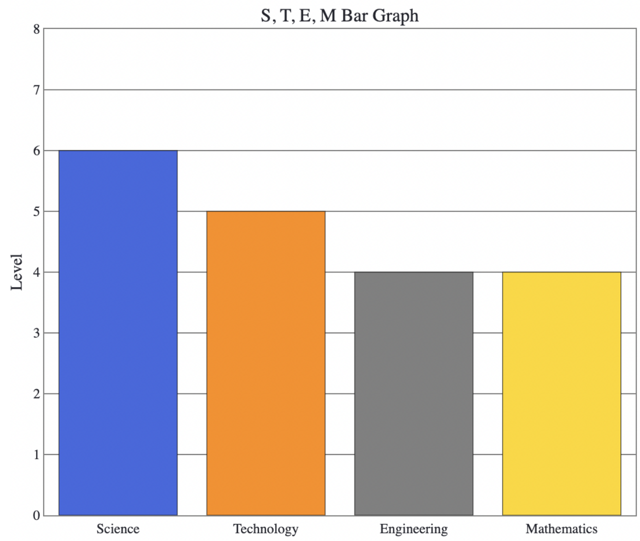

Figure 3. Index for SOC 11-101, Chief Sustainability Officers, as a bar chart visualization .

5. DISCUSSION

There are two primary stakeholders groups for robust S.T.E. and M. classifications: employers with non-STEM jobs, and educational institutions preparing students for the workforce. Currently, middle and high school STEM activities, curriculum, and specialized STEM academies focus on STEM jobs and occupations. STEM funding and a significant amount of school resources are directed towards STEM. Non-STEM jobs and occupations, with no S.T.E.M. educational requirements, potentially have key gaps in the curriculum that results in proficiency gaps for entry-level employees.

With a framework established to facilitate a gold standard for an S.T.E.M. rating system, we believe that industry-specific subject matter experts will be able to quickly fine tune the algorithms to provide road maps for job candidates and educational institutions.

Future applications of the S.T.E.M. index could potentially allow K-12 school districts to automatically match their non-STEM career and technical education offerings to the latest Bureau of Labors Statistics economic data for their region to gain insights into how aligned their S.T.E.M. curriculums are to the economic needs of their region.

6. CONCLUSIONS

The U.S. Department of Education (DOE), Office of Career, Technical, and Adult Education (OCTAE), and other stakeholders, have identified a need to create a STEM Index, quantifying the elements of S.T.E.M. (i.e., science, technology, engineering, and math) in the workforce. By more precisely tracking required S.T.E.M. skills as they change over time, and as differences appear across geographic regions, we will be better positioned to respond to education and training needs.

This is preliminary work; a first step in generating a new STEM Index. Our goal is to create a starting point for SMEs and stakeholders to contribute in meaningful ways, and as such our approach for integrating alternative methods and models is a key outcome of this project. We hope people will inspect and scrutinize the algorithms, ratings, approaches, and outcomes we have developed, which are here: github.com/jcstamper/CTE-STEM. And we also want to see other innovative approaches, and gather additional stakeholder requirements and use-case examples. In particular it is important to find novel ways to utilize the vast amounts of data currently being collected, as well as identify new data sources that need to be developed.

7. ACKNOWLEDGEMENTS

This document was produced and funded at least in part with Federal funds from the U.S. Department of Education under U.S. Government Contract Number: 919900-21-C-0011. The content of this report does not necessarily reflect the views or policies of the U.S. Department of Education nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Albert Palacios served as the contracting officer’s technical representative.

8. REFERENCES

Anderson, D. M., Baird, M. D., & Bozick, R. (2018). Who Gets Counted as Part of America’s STEM Workforce? RAND Corporation8.

Asunda, P. A., & Weitlauf, J. (2018). STEM habits of mind: enhancing a PBL design challenge-integrated STEM instruction approach. Technology and Engineering Teacher, 78(3), 34-38.

Chute, E. (2009). STEM education is branching out: Focus shifts from making science, math accessible to more than just brightest. Pittsburgh Post-Gazette, 10.

Costa, A. L., & Kallick, B. (2008). Habits of mind in the curriculum. Learning and leading with habits of mind, 16, 42-58.

Daugherty, M. K. (2013). The Prospect of an "A" in STEM Education. Journal of STEM Education: Innovations and Research, 14(2).

Donsbach, J., Tsacoumis, S., Sager, C., & Updegraff, J. (2003). O* NET analyst occupational abilities ratings: Procedures. DFR-03-22). Alexandria, VA: Human Resources Research Organization..

Education & Human Resources (2020). STEM education for the future: A visioning report. National Science Foundation.

Fleisher, M. S., & Tsacoumis, S. (2012). O* NET analyst occupational abilities ratings: Procedures update. Alexandria, VA: Human Resources Research Organization.

Fleisher, M. S., & Tsacoumis, S. (2012). ‘O* NET analyst occupational skills ratings: Procedures update. Raleigh, NC: National Center for O* NET Development.

Johnson, C. C., Peters-Burton, E. E., & Moore, T. J. (Eds.). (2021). STEM Road Map 2.0: A Framework for Integrated STEM Education in the Innovation Age. Routledge.

National Center for O*NET Development. O*NET OnLine. Retrieved July 26, 2021, from https://www.onetonline.org/8.

Niculescu-Mizil, A., Perlich, c., Swirsz, G., Sindhwani, V., Liu, Y., Melville, P., Wang, D., Xiao, J., Hu, J., Moninder, S. (2009). Winning the KDD Cup Orange Challenge with Ensemble Selection - Journal of Machine Learning Research, W&CP 7, pp. 23-34.

Oleson, A., Hora, M., & Benbow, R. J. (2014). What is a STEM job? How different interpretations of the acronym result in disparate labor market projections. Viewpoint paper, 1.

Rothwell, J. (2013). The hidden STEM economy. Metropolitan Policy Program at Brookings.

Siekmann, G., & Korbel, P. (2016). Defining “STEM” skills: review and synthesis of the literature. Support document, 2.

SOC: Standard Occupational Classification Manual (2019). Attachment A: Options for Defining STEM occupations under the 2018 SOC system; SOC Policy Committee recommendation to the Office of Management and Budget.

STEM Education Act of 2014. (2014, July 15). Text - H.R.5031 - 113th Congress (2013-2014): https://www.congress.gov/bill/113th-congress/house-bill/5031/text.

Storksdieck, M. (2011). STEM or STEAM. The Art of Science Learning. Retrieved October 14, 2011, from http://scienceblogs.com/art_of_science_learning/2011/04/stem_or_steam.php.

Costa, A. L., & Kallick, B. (2008). Habits of mind in the curriculum. Learning and leading with habits of mind, 16, 42-58.

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35.

Tsacoumis, S., & Willison, S. (2010). O* NET® Analyst Occupational Skill Ratings: Procedures. Alexandria, VA: Human Resources Research Organization.

Yu, H¬F., Lo, H¬Y., Hsieh, H¬P., Lou, J¬K., Mckenzie, T.G., Chou, J¬W., et al., (2010). Feature Engineering and Classifier Ensemble. Proc. of the KDD Cup 2010 Workshop, 1-16.