Received: 4 December 2020 | Accepted: 15 February 2021
DOI: 10.1002/aaai.12040

SPECIAL TOPIC ARTICLE

Intelligent links: AI-supported connections between employers and colleges

Robby Robson[1] | **Elaine Kelsey**[1] | **Ashok Goel**[2] | **Sazzad M. Nasir**[1] | **Elliot Robson**[1] | **Myk Garn**[2] | **Matt Lisle**[3] | **Jeanne Kitchens**[4] | **Spencer Rugaber**[2] | **Fritz Ray**[1]

[1]Eduworks Corporation, Corvallis, Oregon, USA
[2]Georgia Institute of Technology, Atlanta, Georgia, USA
[3]University System of Georgia, Atlanta, Georgia, USA
[4]Credential Engine, Washington, DC, USA

Correspondence
Robby Robson, Eduworks Corporation, 400 SW 4th Street Suite 110, Corvallis, Oregon, USA.
Email: robby.robson@eduworks.com

Funding information
National Science Foundation Award, Grant/Award Number: 2033578

Abstract

When modernization and other changes demand workforce reskilling, employers often turn to local colleges for training programs. Doing so can be a frustrating experience. HR and talent professionals have difficulty identifying and communicating requirements, especially for new jobs and roles, while college continuing education (CE) and professional development offices have difficulty understanding and responding to company needs. This article describes an NSF Convergence Accelerator project called *SkillSync*™ in which multiple forms of AI are used to address this specific problem and provide national efforts (e.g., the US Chamber of Commerce Talent Pipeline Management initiative) with skills data and skills alignment services. Skillsync uses variations on the Siamese Multi-depth Transformer-based Hierarchical Encoder (SMITH) and other natural language understanding methods to map job descriptions and course information to skills taxonomies, uses machine-learned models to align skills needs with learning outcomes and training, and incorporates an intelligent coach based on Georgia Tech's Jill Watson "virtual teaching assistant" to answer questions about Skillsync's vocabulary, functionality, and process. This article describes these AI methods, how these methods are used in Skillsync, and the challenges involved.

INTRODUCTION

Industry 4.0 is creating demands for new skills. In this article, we describe how the *Skillsync* project, which is part of the US National Science Foundation (NSF) Convergence Accelerator *AI and Future of Work and the National Talent Ecosystem* track (https://www.nsf.gov/od/oia/convergence-accelerator/Award%20Listings/track-b.jsp), is using AI to meet this demand. As a web application, Skillsync helps companies identify and meet reskilling and upskilling needs in partnership with a local college professional and continuing education (CE) programs. As a platform, Skillsync processes job- and course-related data and uses language models to extract the *knowledge, skills, and abilities (KSAs)* that jobs require and that courses offer and to help colleges determine how well a combination of training offerings cover a set of skills requested by companies.

The extraction of KSAs and the analysis of training opportunities are operations that require detailed domain knowledge. They cannot be performed by a single human across large numbers of domains, and it is difficult for

FIGURE 1 The Skillsync workflow

HR professionals and CE program administrators to keep up with rapidly changing fields. The capabilities of modern AI enable these operations to be performed as services available to any application. In this article, we discuss these underlying AI services, how they are applied, and how we are integrating a virtual coach first developed as a "virtual teaching assistant" (Goel and Polepeddi 2016) as part of a participatory sociotechnical system design process (Baxter and Somerville 2011). We will also discuss how we are addressing issues of bias and fairness in our underlying algorithms. This is critical because Science, Technology, Engineering, and Math (STEM) and soft skills play an important role in regional economic health (Stewart, Yeom, and Stewart 2020), and we hope to narrow gender, racial, and ethnic gaps in representation in STEM and other jobs (Fry, Kennedy, and Funk 2021).

THE SKILLSYNC WORKFLOW

The overarching goal of *Skillsync* (https://www.eduworks.com/press-release/Skillsync-sept2020.html) is to create efficient, effective, and equitable reskilling opportunities for workers (Figure 1). It does this by: Helping companies identify the skills needed by a cohort of workers (e.g., floor workers who need to upskill to become supervisors); documenting these in the form of training requests and transmitting these to college CE offices. Helping colleges formulate training proposals that respond to these requests and facilitating communication between companies and colleges during (and after) this process.

THE SKILLSYNC APPROACH

Skillsync describes job and skill requests in terms of KSAs. When determining how well a set of training opportunities matches a skills request, we apply language models and scoring algorithms that consider the description, prioritization, and relations among KSAs. The KSAs are organized into *skill frameworks* that come from official sources such as the U.S. Department of Labor and industry associations, from company job descriptions, and from national job postings provided by the National Labor Exchange (https://usnlx.com/).

There are several reasons why SkillSync focuses on KSAs rather than using language modeling techniques to directly match job descriptions to training descriptions. First, our analysis has shown that it is important to be explicit about the skills that workers need. Matching job descriptions directly to courses does not allow colleges to identify the individual skills they need to teach, does not help companies shape their learning and development programs, and does not help employees pinpoint the skills they need. Second, we believe we can get better results by matching skills than by matching text alone. From a modeling perspective, the universe of skills creates an intermediate feature set into which job descriptions, courses, and job requirements can be embedded by both humans and machines. Third, we believe that skill frameworks provide a means to reduce bias. Formulating requirements in terms of skills gives opportunities to people who lack the associated formal education but have nonetheless acquired the skills. In addition, it allows vocabularies from different social contexts to be mapped to a common set of skills, thereby making it easier to address biases inherent in the language used to describe job requirements and course outcomes. Finally, our recent trials confirmed that structured and searchable lists of KSAs are preferred and more impactful for end users.

FIVE SKILLSYNC AI CHALLENGES

Many challenges must be addressed to support the Skillsync workflow. Currently, it is almost entirely manual and not fully integrated into the systems that companies and colleges use for related functions, and it is carried out by HR professionals and CE administrators who have limited time and many competing responsibilities. This leads to business processes, data acquisition, workflow management, and system integration issues, as one expects when designing a sociotechnical system. The SkillSync project continues to address these through a participatory design process, which includes the participatory design of AI. This article focuses on five challenges that stem from the KSA-based approach we take, the use of AI to implement this approach, and our application of sociotechnical system design to an AI-enabled system. Specifically:

1. How do we extract and prioritize KSAs from job postings and other unstructured sources?
2. When deriving KSAs from job postings, how do we avoid including company-identifiable information (CII)?
3. How do we determine how well a set of courses or modules address a skills request and how do we display the

results in a way that is accurate enough to be useful but simple enough to be understood?

4. How do we avoid bias and unfairness, especially in our AI models and algorithms?
5. How do we use AI to create an "intelligent" user experience that meets requirements for efficiency, transparency, and accuracy?

Challenge 1: Language modeling and KSA extraction

Transfer learning in the form of transformer-based, pretrained language models has become ubiquitous in state-of-the-art solutions for a wide range of common natural language processing tasks. In this paradigm, a transformer-based language model is pretrained, using large unsupervised text datasets, to create a contextual representation of the language(s) and/or domain(s) of interest for a wide range of downstream applications. In many cases, a large, general model, trained on broad corpora, is further refined through additional rounds of training on domain-specific texts to increase performance in that domain, while still retaining a relatively high level of performance on domain-general tasks. The resulting language models can then be trained on labeled datasets to perform a wide range of specific tasks. Additional transfer learning is possible; transductive transfer learning can be used to extend the domain coverage of a model trained to perform a specific task without requiring additional labeled task data in the new domains. Likewise, inductive transfer learning can be used to train a model, either simultaneously or sequentially, to perform a wide variety of tasks in many domains (Ruder et al. 2019).

To address the challenges of KSA extraction, anonymization and removal of CII, and alignment of training requests with course content and providers, we applied this general paradigm to each problem. In the case of CII extraction, we began the transfer learning and task-specific training process with an "off-the-shelf"pretrained BERT language model (Devlin et al. 2018). However, for KSA extraction and alignment of training requests with course content and providers, we pretrained a BERT-based general language model from scratch. To perform automated extraction of KSA from unstructured text, we assembled a test set of a range of job description documents from the US manufacturing sector. The majority were drawn from job descriptions published online in early 2021. Additional documents were provided by the Business-Higher Education Forum (https://www.bhef.com/). A separate training dataset was drawn from a database of job descriptions from 2017 and 2018, provided by the National Labor Exchange. Documents in each dataset were labeled to identify KSAs, using a standard protocol, by a diverse set of college students recruited and trained specifically for this task. The overall task was structured as a span prediction task. Preprocessing involved document preprocessing using Eduworks' proprietary Ndoc technology (an AI-powered document ingestion and processing system) to localize the core text of each document, as well as chunking for syntactically logical candidates, using both a constituency and dependency parser and with overlapping permits. Candidate spans tagged as KSAs by human labelers were further classified as either composite or stand-alone KSAs. We then used the multiclass classifier from the Simple-Transformers library (an NLP-oriented extension of HuggingFace's Transformers library) to correctly classify candidate KSAs as either multiword or single word KSAs (Wolf et al. 2019; Rajapakse 2020). Finally, we used Eduworks' proprietary Selector Tool (a deep learning powered tool to detect, generate, and transform mixed syntactic–semantic patterns) to perform postprocessing of the candidates identified by the transformer model, including checks for syntactic and semantic coherence and standardization of the final syntactic form.

Challenge 2: CII

CII extraction from job descriptions is an NER (named entity recognition) task with caveats. A named entity may refer to different entities in different contexts, giving rise to a name ambiguity problem. For example, the name Apple is CII requiring anonymization and removal from the text if it appears as the hiring company in a job advertisement, whereas a mention of Apple OS in a list of skills is not a CII and therefore does not need to be flagged for removal. Name variation in CII extraction is also an issue that requires coreference resolution and entity linking: Southwest Airlines and Southwest may refer to the same company and, hence Southwest may be CII, as is the text "we are the largest domestic airlines in the US" and "our Rapid Rewards program," which both pinpoint that the hiring company is Southwest Airlines. Deep learning models have proven successful in addressing some of these challenges.

For CII extraction, we fine-tuned a pretrained BERT model using a subset of the National Labor Exchange database of job descriptions ranging from 2017 to 2018. Training documents were labeled for CII using the standard BIO tagging format, casting the problem of CII extraction as a standard multilabel classification task. College students were recruited and trained specifically for this labeling task, and the documents were preprocessed using Eduworks' proprietary Ndoc technology and other standard techniques. A rule-based model using the metadata

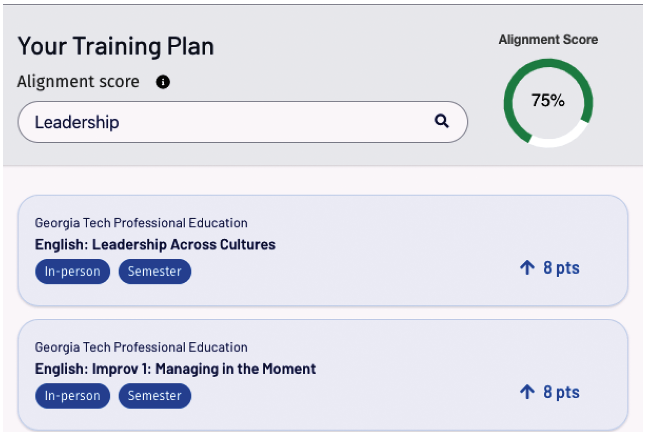

FIGURE 2 Skillsync alignment score

of job descriptions was stacked on the top of the BERT model for further CII extraction. Such a rule-based model can potentially serve as a baseline against which other CII-extraction models can be assessed. Recent research suggests that BERT-based pretrained language models can provide a unified framework for addressing issues such as name ambiguity, coreference resolution, and entitylinking in an NER task (Broscheit 2019; Du, Qi, and Sun 2019; Joshi et al. 2019). In the future, we plan to extend our models to incorporate some of these features, making CII-extraction services robust and broadly applicable across a wide range of NLP tasks.

Challenge 3: Alignment scores—indicating the match between courses and requests

During workshops and interviews conducted in 2019, we discovered that we need a way to measure the distance, or fit, between the KSAs required for a job and those obtainable from an educational or training program. This led to the notion of an *alignment score* that we hope users will trust and be able to apply, much the way that users trust and apply credit ratings, sports statistics, product ratings, and similar quantities. In SkillSync, after a company communicates a skills request to a college and a college responds, both parties must determine how well the proposed training opportunities meet the company's needs. This is done with the alignment score, which is presented on a scale of 0–100 (Figure 2).

Computing this score involves assembling a range of alignment metrics into a human-interpretable recommendation. At the core of this task is the basic challenge of determining whether or not a given course of instruction supports a given KSA. In addressing this issue, we rely on a transformer-based model of alignment we developed for matching learning objectives to content in intelligent tutoring systems (Bell et al. 2020) and trained using the semantic alignment library of SimpleTransformers. Using a BERT model pretrained via standard techniques, and by preprocessing course content using Facebook AI Similarity Search (FAISS), an indexing method for rapid analysis and clustering of dense vector representations (Johnson, Douze, and Jégou 2019), we are able to quickly identify a set of candidate courses that potentially align well with each KSA. To select the "best" match for each KSA, we used transfer learning to further train the transformer model to perform a refined version of the matching task. This involves the creation of a labeled data set of KSA/course pairs, where human raters are used to determine the quality of the match on a numeric scale. Human raters are asked to produce an overall match score, as well as match scores in individual dimensions (domain, skill, knowledge, level, depth/coverage, etc.). We currently train the model using the overall match score only, but we are conducting simultaneous data labeling in each match dimension individually to support the eventual expansion of the model to support weighted matching by end users across multiple dimensions.

Given that each training request is composed of multiple, weighted KSAs, we explored different methods of matching overall training requests (weighted lists of KSAs) to both individual courses and training providers. The simplest approach involved a heuristic for finding the subset of courses that produces the maximum overall weighted score when summing the match quality between each KSA and the course in the subset. This can be further restricted to the subset of courses offered by each training provider. We are now experimenting with a more holistic approach that views the entire training request and each subset of courses as a single, long-form document. By pretraining a language model using preprocessing techniques to map longer phrases to symbolic representations of higher-level concepts (concept modeling), we are able to reduce long-form inputs to short-form symbolic representations of concepts, and perform matching at the concept level, using existing pretraining and transfer learning techniques. More recently, we have experimented with applying the general approach described in Yang et al. (2020) to perform long-to-long matching between training requests and course portfolios. While the latter approach appears comparable to concept modeling for alignment scoring, our early results suggest that it may be particularly useful in the future in evaluating the potential improvement in overall match from the addition of hypothetical new course offerings.

Challenge 4: Addressing bias

Large-scale language models have been criticized for capturing and sometimes amplifying undesirable societal biases. These biases have been particularly noted in relation to occupation and job-skill-related terminology (Lu et al. 2020). For SkillSync AI services, we identified five potential points where undesirable biases might be introduced or mitigated. The first point we considered is the large text corpora used to pretrain the underlying language models used in downstream transfer learning for all services. To the extent that undesirable biases exist in the texts used for unsupervised learning, they may be reflected downstream in service outputs. Labeled datasets used to train models to perform specific tasks can also include similar undesirable biases, and while the labeled test sets used to measure model performance do not directly introduce bias into the models, if they contain undesirable biases, they may boost the performance ratings of candidate models that mirror those biases. Skillsync employs several techniques, including counterfactual data augmentation (CDA) and REG, to mitigate undesirable biases in underlying datasets. The former replaces gendered language and linguistic references to race and ethnicity with either a neutral equivalent or a multidirectional expansion; the latter is a bias regularization method developed by Bordia and Bowman (2019) that debiases embedding during language model training by minimizing the projection of neutral words on the relevant axes. We also added a data preprocessing step that performs occupation and job-specific debiasing, replacing gendered occupational and job task language with neutral equivalents.

The last two points of bias introduction and mitigation that we identified were the criteria used to rank the performance of different candidate models during testing, as well as conscious or unconscious bias in the pool of human raters used to label the data used to train for downstream tasks. To address the former, we trained data scientists to maximize model performance while minimizing bias when conducting hyperparameter tuning and in the design of data preprocessing pipelines. To address the latter, we attempted to diversify the pool of human raters who label training data. We were initially concerned that there might be tradeoffs involved in both maximizing performance while minimizing bias in model training. However, in practice, we found that minimizing undesirable bias often served to prevent overfitting.

Challenge 5: Virtual coach for using Skillsync

The adoption of intelligent tools such as Skillsync in companies and colleges depends on several economic, social, cultural, and technological factors. From the perspective of human-centered computing, technological considerations include not only usefulness, efficiency, and

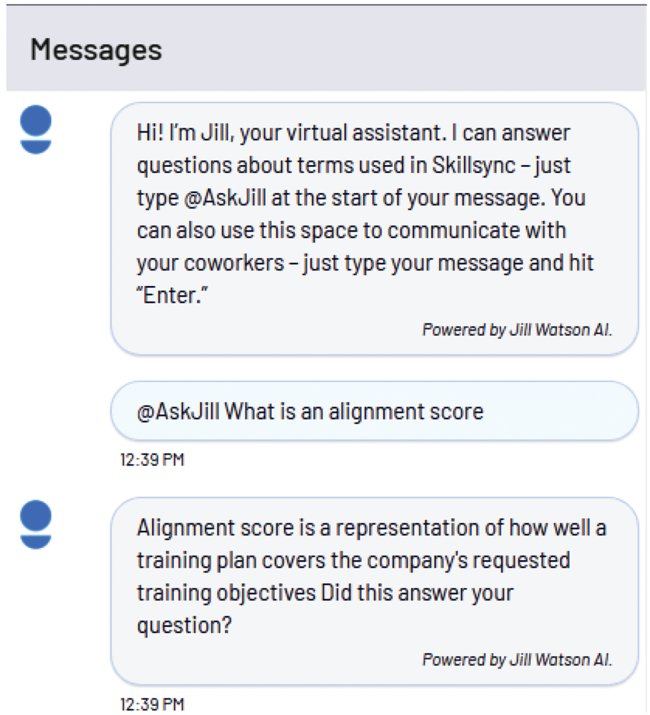

FIGURE 3 A Skillsync user has a question answered by AskJill

accuracy, but also usability, learnability, and transparency of the tool. If a user is unable to easily learn to use the tool to efficiently produce a result that she feels comfortable with and confident about, she may not want to use the tool again or recommend it to other prospective users. Thus, another challenge in developing Skillsync is to make its vocabulary, functions, and processing transparent so that the user can both easily learn to use the tool efficiently and effectively, and the user can be confident about the result she produces with the use of the tool.

To facilitate Skillsync's adoption in practice, we are developing a virtual coach for using Skillsync. Let us consider, as an example, the term *KSA* used in Skillsync (as well as in this article): a user (or a reader) may not necessarily know what a KSA is or what Skillsync means by a *skill*. If the user does not understand Skillsync's meaning of the word *skill*—and dozens of other terms the tool uses—she may have a difficult time using the tool effectively and feel less confident in the results it produces. The virtual coach, called *AskJill*, is designed to explain what Skillsync means by the various terms it uses as well as how it works, with the goal of helping the user build a "theory of mind" (Baron-Cohen 1999) of Skillsync (Figures 3 and 4). This is important for complex intelligent tools like Skillsync. AskJill helps build a shared mental model of Skillsync's vocabulary across users from companies and colleges. This is important for domains where there is no agreed-upon vocabulary, which is very much the case for competency and skills management.

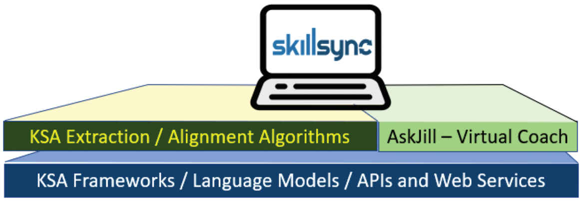

FIGURE 4 Skillsync architecture

The virtual coach in Skillsync builds on Jill Watson, a virtual teaching assistant developed at Georgia Tech's Design and Intelligence Laboratory. Jill Watson automatically answers students' questions in online discussion forums of college-level classes (Goel and Polepeddi 2016, 2018; Goel 2020). For example, a student may ask "Will we have office hours in this class?" Jill Watson may answer "Most of our teaching assistants will hold weekly office hours. A schedule for the office hours will be made available early in the semester." The original Jill Watson was developed using IBM's Watson platform (Ferrucci et al. 2010). We found that by answering some questions automatically, Jill Watson saved teachers time and helped serve students who had anywhere, anytime access.

A more recent version of Jill Watson acts as a virtual coach called AskJill in the context of a virtual laboratory for inquiry-based modeling and learning called VERA (An et al. 2020). Although VERA contains a user's guide as well as a glossary, most users do not have the time or the inclination to read them. AskJill in VERA allows users to learn about using VERA by asking questions (Goel 2020). For example, a user may ask "What is a food web?" AskJill may answer "The elaborate, interconnected feeding relationships in an ecosystem." In addition to IBM's Watson platform, AskJill uses an ensemble of classifiers combined with knowledge-based preprocessing of questions and postprocessing of answers. We found that by explaining the terms used by VERA and providing instructions in response to queries, AskJill makes VERA more usable and learnable. In addition, AskJill serves users anywhere at any time, regardless of whether human help is available.

The virtual coach for Skillsync, also called AskJill, is patterned after AskJill in VERA. To develop AskJill for SkillSync, we engaged in extensive participatory design with stakeholders from companies and colleges. We elicited design requirements for AskJill, developed a glossary of Skillsync terms, and developed a typology of questions users may ask of AskJill while using Skillsync, which builds on the existing question typology used by AskJill in VERA. We then trained AskJill to answer questions about Skillsync terms, used Amazon Web Services to enable two-way communication between Skillsync and AskJill, developed a UI for interacting with AskJill within Sklllsync, and conducted extensive testing. In the near future, we plan to expand AskJill's capabilities to answer questions about Skillsync's functionality and processing. Thus, AskJill in Skillsync is an experiment in using human-centered AI for enhancing technology adoption through question–answering.

SKILLSYNC IN CONTEXT

The challenges and solutions discussed above are specific to Skillsync as an application that connects companies to colleges, but Skillsync is just one part of a larger picture that includes multiple national efforts to address shortages and inequities in America's talent pipeline. In the long run, Skillsync is designed to provide services to these efforts. For example, Skillsync will provide CII-free skills frameworks extracted from NLx job postings for use in an NLx research portal and KSA frameworks extracted from courses to the US Chamber of Commerce Foundation's T3 Innovation Network Open Competency Framework Collaborative (https://www.ocf-collab.org/) and the Credential Engine, a nonprofit organization that provides a centralized registry with current information about degrees, licenses, badges, and other credentials together with a common description language that enables credential comparability and a platform that supports customized applications for searching and retrieving credential information.

In the context of the NSF's Convergence Accelerator Track B, which includes two other projects ("LEARNER" and "NeuroAI@Work") with articles in this volume, Skillsync is envisioned as an integration platform that manages frameworks of KSAs relevant to the other focus areas (first responders and neurodiverse workers) and that connects companies and CE programs to the innovative training being developed by these projects. By creating common, machine-actionable sets of skills, we can analyze the skills that people possessed prior to training and those that were demonstrated or acquired through training, leading to models that can improve the Track B training programs and inform companies of which skills they may wish to prioritize.

In addition, the work we are doing on AskJill was inspired by (and is generalizable to) other applications and poses the interesting questions of finding the right balance between a general dialog agent such as Siri and a fully contextual help-screen agent and of determining the effectiveness of question–answering agents for creating shared mental models of complex applications.

Finally, we are investigating the role of AI in achieving the goal of connecting companies to colleges to meet reskilling needs. In current practice, significant domain knowledge is required to interpret job descriptions and to understand the precise topics and skills that a CE course covers. We use granular KSAs to create a common language for expressing and comparing job requirements and training outcomes. This requires sophisticated language models and trained algorithms, but it is possible that the dominant benefit of Skillsync is simply facilitating connections between companies and colleges and has little to do with its use of AI. Preliminary results from focus groups run from October 2020 to March 2021 indicate otherwise, but we consider this to still be an open research question that applies to many other systems that use AI to re-engineer and digitize existing processes and workflows characterized by interpersonal communication and dependence on tacit knowledge.

ACKNOWLEDGMENTS

Georgia Tech's Jill Watson team is grateful to IBM for providing access to its Watson platform for academic research. We thank Harshvardhan Sikka, Vrinda Nandan, and Jeonghyun Lee for their ongoing contributions to the development of AskJill in Skillsync. The Eduworks team is grateful to the many partners participating in SkillSync and to Jennifer Thornton and Stephanie Blochinger from the Business Higher Education Forum for their contributions and to the National Association of State Workforce Agencies (NASWA) for joining the SkillSync project and providing NLx data. Partially supported by the National Science Foundation Award # 2033578.

CONFLICT OF INTEREST

Authors declare no conflict of interests.

REFERENCES

An, S., R. Bates, J. Hammock, S. Rugaber, E. Weigel, and A. Goel 2020. "Scientific Modeling Using Large Scale Knowledge." In *Proceedings of the International Conference on Artificial Intelligence in Education*, 20–4. Cham: Springer.

Baron-Cohen, S. 1999. "Evolution of a Theory of Mind?" In *The Descent of Mind: Psychological Perspectives on Hominid Evolution*, edited by M. C. Corballis and S. E. G. Lea. New York: Oxford University Press.

Baxter, G., and I. Sommerville. 2011. "Socio-Technical Systems: From Design Methods to Systems Engineering." *Interacting with Computers* 23 (1): 4–17.

Bell, B., K. Brawner, D. Brown, and E. Kelsey. 2020. "Matching Content to Competencies with Machine-Learning: A Service-Oriented Content Alignment Tool for Authoring in GIFT and Beyond." In *Proceedings of the 8th Annual Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym8)*, 101. US Army Combat Capabilities Development Command Soldier Center, May.

Bordia, S. & S. R. Bowman 2019. "Identifying and Reducing Gender Bias in Word-Level Language Models." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop, 7–15, June, Minneapolis.

Broscheit, S. 2019. "Investigating Entity Knowledge in BERT with Simple Neural End-To-End Entity Linking." In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, 677–85, November.

Devlin, J., M. W. Chang, K. Lee & K. Toutanova 2018. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1, 4171–4186, June, Minneapolis.

Du, J., F. Qi & M. Sun 2019. "Using BERT for Word Sense Disambiguation." arXiv preprint arXiv:1909.08358.

Ferrucci, D. A., E. W. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, J. M. Prager, N. Schlaefer, and C. A. Welty. 2010. "Building Watson: An Overview of the DeepQA Project." *AI Magazine* 31 (3): 59–79.

Fry, R., B. Kennedy, and C. Funk. 2021. "STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity." Pew Research Center. April 1, 2021. https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/ (accessed April 29, 2021).

Goel, A. 2020. "AI-Powered Learning: Making Education Accessible, Affordable, and Achievable." CoRR abs/2006.01908; Team emPrize Round Report to XPrize AI Competition.

Goel, A., and L. Polepeddi. 2016. "Jill Watson, A Virtual Teaching Assistant for Online Education." Technical report, Georgia Tech. available from the Georgia Tech Library SMARTech @ http://hdl.handle.net/1853/59104

Goel, A., and L. Polepeddi. 2018. "Jill Watson, A Virtual Teaching Assistant for Online Education." In *Education at Scale: Engineering Online Teaching and Learning*, edited by C. Dede, J. Richards, and B. Saxberg. NY: Routledge.

Johnson, J., M. Douze, and H. Jégou. 2019. "Billion-Scale Similarity Search with GPUs." *IEEE Transactions on Big Data* 7 (3): 535–47.

Joshi, M., O. Levy, L. Zettlemoyer, and D. Weld. 2019. "BERT for Coreference Resolution: Baselines and Analysis." In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*.

Lu, K., P. Mardziel, F. Wu, P. Amancharla, and A. Datta. 2020. "Gender bias in neural natural language processing." In *Logic, Language, and Security*, 189–202. Cham: Springer.

Rajapakse, T. 2020. "Simple Transformers." https://simpletransformers.ai/.

Ruder, S., M. E. Peters, S. Swayamdipta, and T. Wolf. 2019. "Transfer Learning in Natural Language Processing." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials*, 15–8, June.

Stewart, F., M. Yeom, and A. Stewart. 2020. "STEM and Soft Occupational Competencies: Analyzing the Value of Strategic Regional Human Capital." *Economic Development Quarterly* 34 (4): 356–71.

Wolf, T., L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi & A. M. Rush … 2020. "HuggingFace's Transformers: State-of-the-Art Natural Language Processing." In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, October, 38–45.

Yang, L., M. Zhang, C. Li, M. Bendersky & M. Najork. 2020. "Beyond 512 Tokens: Siamese Multi-Depth Transformer-Based Hierarchical Encoder for Document Matching. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, October, 1725–1734. https://doi.org/10.1145/3340531.3411908

AUTHOR BIOGRAPHIES

Robby Robson is a Researcher, Entrepreneur, and Standards Professional known for creative and disruptive innovation in industry and academia. He is the Co-founder and CEO of Eduworks Corporation, is on the IEEE Standards Association Standards Board, and is currently developing the next generation of competency management and talent analytics solutions.

Ashok Goel is a Professor of Computer Science and Human-Centered Computing in the School of Interactive Computing at Georgia Institute of Technology and the Chief Scientist with Georgia Tech's Center for 21st Century Universities. He is also the Executive Director of the NSF's National AI Institute for Adult Learning and Online Education. Ashok is a Fellow of AAAI and the Cognitive Science Society and an Editor Emeritus of AAAI's AI Magazine.

Myk Garn is an Academic Innovator and Strategist with over 30 years of experience developing, deploying, and leading college and state system instructional, operational, organizational, policy, and strategic models. His experience includes strategic planning, affordability, emerging trends, disruptive innovation, quality standards, governance, accessibility, and accreditation.

Elaine Kelsey is the Director of Research Engineering at Eduworks Corporation, focusing on applications of natural language processing and machine learning in topic and concept detection and automated text generation. She has multiple bachelor's and master's degrees in computer science, linguistics, and molecular biology and is currently working towards a PhD in Computational Linguistics.

Sazzad M. Nasir is a Machine Learning Researcher at Eduworks Corporation. He has a PhD in theoretical physics from Cambridge University, UK. His career straddled multiple scientific disciplines with publications in leading journals. He was a faculty member at Northwestern University and has a research affiliation with Haskins Laboratories at Yale University. He has expertise in machine learning, predictive analytics, and computational modeling and developed deep learning models in areas of natural language understanding.

Elliot Robson has over 15 years of experience leading and designing new technology in training and education. He is the General Manager of Eduworks Corporation and is Co-PI on the NSF Competency Catalyst and Competency and Skills System (CaSS) projects. He has worked as a Leader in education technology transformation since 2005 with organizations including the NYC Department of Education, UNESCO, and Amplify Learning. He has numerous publications and presentations with a special interest in AI methods for competency computation and application.

Matt Lisle is the Director of Digital Learning Technologies at the Center for 21st Century Universities. He brings a combination of professional and academic experience in the fields of instructional design, web design and development, and content strategy to his work. His main professional focus is the creation and development of technology-enhanced learning experiences.

Jeanne Kitchens is the Chief Technology Services Officer at Credential Engine. She has led numerous workforce development programs for the State of Illinois and national programs, including the U.S. Chamber of Commerce Foundation's Talent Pipeline Management.

Spencer Rugaber is a Faculty Member in the College of Computing at the Georgia Institute of Technology. His research interests are in the area of Software Engineering, specifically reverse engineering and program comprehension, software evolution and maintenance, and software design. Dr. Rugaber has served as a Program Director for the Software Engineering and Languages Program at the U. S. National Science Foundation and as a Vice-Chairman of the IEEE Technical Committee on Reverse Engineering.

Fritz Ray is the Chief Technology Officer of Eduworks Corporation with 20 years of experience developing software for use in academic, government, commercial, and open-source environments, including applications for aviation, financial, and patent areas. He leads software development at Eduworks, contributes to IEEE and other standards efforts around learning and is the Lead Contributor of the open-source CaSS Project.

How to cite this article: Robson, R., Kelsey, E., Goel, A., Nasir, S. M., Robson, E., Garn, M., Lisle, M., Kitchens, J., Rugaber, S. and Ray, F. 2022. "Intelligent links: AI-supported connections between employers and colleges." *AI Magazine* 43: 75–82. https://doi.org/10.1002/aaai.12040