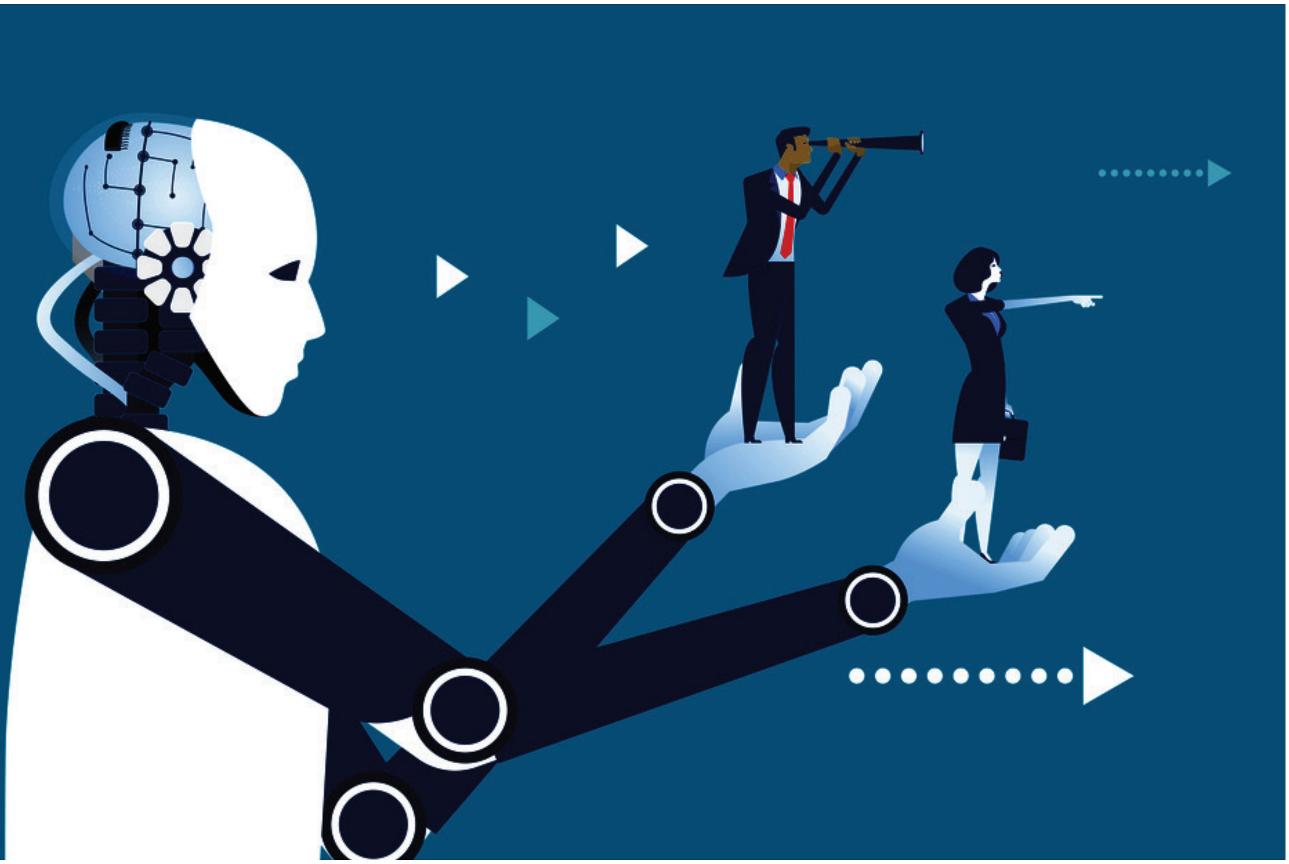
SIOP White Paper Series

SIOP | SOCIETY for INDUSTRIAL and ORGANIZATIONAL PSYCHOLOGY

SCIENCE FOR A SMARTER WORKPLACE



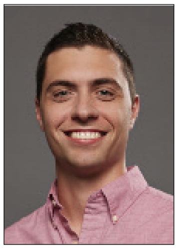
Artificial Intelligence in Talent Assessment and Selection

Neil Morelli

A White Paper prepared by the Visibility Committee of the Society for Industrial and Organizational Psychology.
440 E Poe Rd, Suite 101 Bowling Green, OH 43402

Table of Contents

Author



Neil Morelli
Berke

Neil A. Morelli, PhD, is VP of Product and Assessment Science at Berke. He specializes in talent acquisition, data science, and product development. Before joining Berke, Dr. Morelli was the head of Product and Selection Science for The Cole Group, an executive search firm serving high-growth technology companies. He has also worked for an early stage, web-based assessment company and as a talent management consultant for *Fortune* 500 companies including Georgia Pacific, Wal-Mart, Bridgestone Tires, and Texas Instruments. Dr. Morelli received his PhD in I-O Psychology from the University of Georgia and his MS in I-O Psychology from the University of Tennessee at Chattanooga. His research interests include technology-enabled recruitment and assessment, and he has authored several articles and book chapters in a variety of peer-reviewed publications.

Introduction

Why Should You Care About AI Used for Hiring?

The health and vibrancy of our national economy depends on the health and vibrancy of our organizations. How an organization is staffed is a key driver of its health. Yet only one in three executives rates his or her company as "very effective" at reducing unsuccessful hiring decisions (Bravery et al., 2019). Thus, industrial-organizational (I-O) psychologists help organizations thrive by choosing the right talent for the right job at the right time, using the most accurate, cost effective, and fair tools available (Ployhart, Schmitt, & Tippins, 2017).

Traditionally, I-O psychology identifies the best person for the job by making common selection tools like the resumé or interview more scientific, or by developing technical screening tools such as psychometric assessments. Today, artificial intelligence (AI) is being promoted as a tool that more quickly and easily screens job candidates. In other words, AI is now being used to supplement or replace traditional methods that measure the individual differences that predict future job performance.

As hiring experts, I-O psychologists have a responsibility to evaluate AI in this use case. Although AI makes several hopeful promises for more efficient, effective, and enjoyable talent selection, many topics relevant to traditional testing also deserve future research attention for AI tools: reliability, validity, fairness transparency, acceptance by job seekers, and legality. I-O psychologists have a document called the *Principles for the Validation and Use of Personnel Selection Procedures* (5th ed.; SIOP, 2018) that covers these key issues that remain highly relevant to AI tools.

All of that to say, AI is a megatrend that is already changing how organizations function, especially when it comes to hiring talent. We want to help you adapt with the changing technology in this critical area. To paraphrase Thomas Siebel, founder of an AI company and author of *Digital Transformation*, AI is an oncoming train and you're either on it or you're on the track (LeVine, 2019).

Therefore, the goals of this paper are to dispel some of the mystique that surrounds AI used for hiring, while also encouraging confident investment in AI tools that can help your organization succeed through more effective talent assessment and selection.

Background

What Is AI?

When the term "AI" is invoked it often comes with a lot of hype, jumbled meanings, and unrealistic expectations. So, it is helpful to start with a definition: AI is the collection of technologies and algorithms that imitate one or more of the following human abilities (Barney, 2019):

- natural language processing: understanding and communicating in a human language;
- knowledge representation: creating models based on what a system learns or perceives;
- reasoning: using data to answer questions or draw new conclusions;
- learning: adapting to new situations and extrapolating patterns;

- sensing: perceiving objects or people;
- object manipulation (in some cases): moving objects in a physical space.

This definition is important because we must remember that AI is not “one thing.” Tools rightly billed as “AI” normally have several machine or deep learning models behind the scenes working together to process inputs that are supervised or unsupervised (i.e., data that are labeled by human labelers, or not) to create outputs that are continuous or categorical (i.e., predicting a number or a category/group). In other words, much of what AI is actually “doing” can be broken down by these functions:

- classification: assigning things to a group based on their similarity to previously labeled groups;
- clustering: determining potential groups from unlabeled data;
- regression: predicting a number based on a known relationship;
- identifying patterns between variables: experimenting with potential relationships within data to discover patterns.

	Supervised	*Unsupervised*
Categorical	Classification	Clustering
Continuous	Regression	Identifying patterns

Figure adapted from Soni (2018)

AI might seem like brand-new technology, but much of the math and computer science behind AI and machine learning (AI’s equally famous subcategory) have been around for several decades. What has changed recently is that computing costs are down, data volume and velocity are up, and the number of computations possible on a single chip have never been higher.

Despite recent progress and attention, a human-like AI (also called a “general” or “strong” AI) is still science fiction. However, “narrow” or “weak” AIs (those trained to complete specific tasks: recognize images, play games, and make recommendations) have recently made eye-popping headlines in the *Wall Street Journal, New York Times*, and *Washington Post*. These use-case driven AIs have made notable accomplishments in fields such as medicine, transportation, manufacturing, law, and finance. Yet, according to Garry Mathieson, co-chair of the AI industry group at Littler law firm, “HR is actually late to the game” (Brin, 2019). HR might be a new use case for AI, but in a 2019 survey of 7,300 professionals representing nine industries, 40% said their organizations are already using AI to screen or assess job candidates during recruitment (Bravery et al., 2019).

What Are the Promises of AI for Assessing and Selecting Talent?

AI makes bold promises to overcome many problems inherent to finding, engaging, and assessing job seekers. To summarize, when applied to hiring, AI claims to:

1. Automate repetitive and cumbersome recruiting tasks. An early success story has been how AI automates manual recruiter tasks such as sourcing candidates and screening resumés. This is where most of the AI-based hiring market focuses today. Several vendors have sprung up to offer this service, touting AI-based technologies that source potential job candidates by screening and assembling social media and other online data. Although this technology blurs the line between identifying and assessing job candidates, initial scientific research shows that AI systems not only find candidates but also reliably assess individual differences, such as personality, to match candidates to jobs (Akhtar et al., 2019; Morelli & Illingworth, 2019).
2. Make it possible to work with data that are massive, decentralized, "noisy," and otherwise unstructured. AI can more easily and efficiently sift through massive datasets that are often housed in systems spread across an organization (EY, 2018). AI is better at both processing large amounts of data and helping reduce the time it takes humans to clean and prepare data for additional analyses. Because 60% of data science is cleaning and organizing data (Press, 2016), AI systems that can clean "noisy" and unstructured data can increase HR analyst's effectiveness and productivity.

3. Model more variables (features) than current self-report assessment methods. Traditional assessment approaches often "leave data on the table." In other words, models based on traditional self-report assessment data typically include about a dozen variables. Failing to collect and include other valuable information from candidates might decrease the model's predictive power and candidate experience. Conversely, deep learning is an AI approach that can handle zettabytes of data and thousands of features in a model, potentially increasing a model's predictive power while reducing candidate time and effort. Although deep learning's potential is most often applied to image and speech recognition, assessments that use video and speech data can now incorporate hundreds or thousands of features into its algorithms with the help of deep learning, in an attempt to enhance traditional selection methods.
4. Help I-O psychologists create and validate traditional assessments. If an "AI-first" solution isn't appealing, Barney (2019) describes how the use of AI can make creating or validating traditional assessments easier and cheaper. For instance, I-O psychologists can use AI to conduct traditional job analyses efficiently and thoroughly. Job analyses are the bedrock data-collection activities that define what tasks are performed on a job and what key characteristics (KSAOs) are required to perform those tasks. AI can also supplement how I-O psychologists perform targeted meta-analyses (quantitative summaries of studies) that help streamline how assessments get validated. AI can even perform analyses on its own to determine an assessment's psychometrics—that is, AI can compute the statistics that suggest an assessment is reliable and accurate.

Ask yourself: what things would be better if they were done 24/7? What would be better if it were done at scale? What would benefit from greater consistency? What would be possible if we leveraged broader expertise to see beyond our current limits? These are good candidates for AI.

–Deborah Bubb
VP and Chief Leadership, Learning & Inclusion Officer
IBM (Guenole & Feinzig, 2018)

Implications for Practice

What Is a Specific, Viable Example of AI Applied to Talent Assessment and Selection?

Natural Language Processing (NLP), or AI-based interpretation and application of human language, is a large-scale attempt to identify and select talent effectively from text-based data. NLP is much more than keyword matching; it can identify themes and relationships within and across text passages. One AI company has already used NLP to process millions of resumés and CVs to cluster people based on their similarity to a job profile defined by a referent individual or job posting. Matches can be searched by adding and subtracting profiles, greatly increasing the search's flexibility beyond keyword matching to a profile "analogy"—or the characteristic and experience combinations recruiters and leaders often use to describe an ideal candidate (May, 2016).

Scientific studies have shown how NLP and machine learning can create structured data from unstructured resumes (this is often a very cumbersome and error-prone data-entry task for recruiters). Structured resumé data help to rank candidates according to a desired job posting (Sadiq, Ayub, Narsayya, Ayyas, & Tahir, 2016). A growing body of research is also showing that personality can be assessed by mining social media and online profiles such as LinkedIn, and some studies have suggested that social media-assessed personality traits are related to work performance (Akhtar Winsborough, Lovric, & Chamorro-Premuzic, 2019). These researchers argue that this is an exciting prospect for talent assessment because online data reflect offline behavior, represent a more complete picture of someone's personality than data from self-report measures, and can systematically be tied to job performance. Certainly, privacy issues are not to be ignored here, but the promise of NLP lies in analyzing enormous amounts of data in a standardized way that removes the subjectivity inherent to casually vetting social media profiles, interviews, and personal CVs.

NLP has also been used to score open-ended assessment responses (Campion, Campion, Campion, & Reider, 2016). The scientific and vendor-based research shows that text data (either taken from written essays or transcribed from video or audio) can be mined and understood for its content, increasing both an assessment's efficiency and the number of features that can be added to a model. Combined with easier scoring, NLP of open-ended responses make it possible to create more engaging and realistic assessment simulations. These simulations can be dynamic to candidate responses, allowing scenarios to look and feel like the job while providing a personalized experience to the candidate.

Finally, NLP is behind conversational chatbots that help newly hired employees get up to speed faster. For example, each year Unilever hires 30,000 people out of nearly 2 million applications. In addition to assessing aptitudes and matching candidates to jobs using machine learning, Unilever onboards new hires with a NLP-powered chatbot that answers simple questions and retrieves information through dialogue (Marr, 2018). Although still in its early phases, over 80% of the regions who have implemented the chatbot said they would continue to use it in their onboarding process.

What Are the Open Questions and Obstacles Related to AI-Based Hiring Tools?

As with any new tool or technology, organizations need to understand the logistical and ethical implications before blindly adopting AI for hiring. Fortunately, many computer scientists and I-O psychologists have already developed important questions for organizational decision-makers to consider.

For example, many I-O psychologists evaluating AI hiring tools, especially tools that incorporate deep learning models, have asked a conceptual question: Are the predictors or variables inside AI algorithms job related? In other words, do we really understand the variables that are being combined in a deep learning model? These questions are often mentioned as the "black box" problem, or the inability to interpret a multilayered, data-driven AI model. Models that cannot be interpreted cannot be as easily defended in court or explained to business stakeholders and job candidates; two very important issues related to assessment and selection. However, all AI solutions shouldn't be dismissed as "black boxes." In some cases, interpretability might not be the chief goal and dismissing or discounting AI as being a "black box" ignores recent efforts to increase a model's interpretability (Landers, 2019). For instance, a company whose stated mission is to "make AI explainable" just raised $30 million from venture capitalists to grow the business—a tangible sign of work in this area.

Ethical questions often asked are: Are AI hiring tools biased? If so, are AI hiring tools legally defensible? Research has demonstrated that models trained on text, such as those used in NLP-based tools, contain the semantic or historical biases present in the text itself (Caliskan, Bryson, & Narayanan, 2017). Although there are no court cases from the EEOC (yet), many in the field have warned about the inherent biases that can creep in from the data that are used to train models. For example, Amazon suffered some negative PR from an AI resumé screening tool that was biased against women (Dastin, 2018). Bias will be an important, ongoing issue for developers and practitioners to consider, but some researchers are already charting a course to correct or mitigate inherent bias in AI models (Veale & Binns, 2017).

Finally, if the goal of using AI tools is to hire the right candidate at the right time, how do candidates react to algorithmic hiring? A recent Pew Research study suggests candidates are worried (Smith & Anderson, 2017). In this study, over 4,000 respondents reacted to a hypothetical scenario where "computer programs may be able to provide a systematic review of each [job] applicant without the need for human involvement." More than half (67%) said that this scenario worries them, and three out of four said that they would not apply to a job that used an algorithm to make a hiring decision. Attitudes and perceptions can change quickly, but these numbers reveal people's uneasiness with humans being completely removed from the decisions that affect their livelihoods. How to overcome the fear and uncertainty of AI hiring tools is a difficult and open question for practitioners and developers to answer.

Next Steps

What Should a Savvy Consumer Do Before Using an AI Hiring Tool?

1. Get I-O psychologists involved. It's true that new developments in AI-based hiring often come from computer scientists and software engineers, but you don't need to solely rely on engineers to solve talent

selection challenges. Hire your next data scientist with I-O psychology training and expertise in hiring. Reach out to an I-O psychology consultant to help vet a potential AI vendor, or choose to work with companies that employ I-O psychologists. Having an I-O psychology perspective at the table can help you ask the right questions to make sure your chosen AI hiring tool is relevant, effective, ethical, and legal. This advice echoes calls from IT leaders for humans to be "in the loop" (i.e., have cyclical human input during a model's development and tuning) so that negative or biased AI-based decisions can be avoided (Persson & Kavathatzopoulos, 2017; Rahwan, 2017).

2. Pair AI-based tools with human decision makers. AI has a lot to offer in terms of efficiency and accuracy, but we can't expect people to be completely removed from a hiring decision. Rather than remove managers or other stakeholders from an AI-infused hiring process, consider how you might involve them. This "human plus machine" approach is supported by research that's shown how combining algorithms and human judgment increases the accuracy of predicting future job performance beyond human judgment alone (Kuncel, Klieger, Connelly, & Ones, 2013).
3. Apply a healthy amount of skepticism to marketing materials and ask for specifics. Even if you don't understand all of the answers, asking for specifics about "how the sausage is made" helps detect vague or contradictory answers from vendors, and it allows you to compare across vendors. Thinking critically about technology that is often marketed as a magic wand will help you ultimately work with vendors that can deliver on realistic promises and add value.

AI applied to assessing and selecting talent offers some exciting promises for making hiring decisions less costly and more accurate for organizations while also being less burdensome and (potentially) fairer for job seekers. However, HR is "late to the AI game." In these early days it's important to understand that there is plenty of hype surrounding AI-based tools. After being equipped with the right questions and knowledge that removes some of AI's mystique, you can be a more informed AI user and a better decision maker for your organization.

Questions to ask a vendor when evaluating an AI hiring tool:

1. **How do you model the task-based and team-based requirements of the job?**
2. **How do you define the human skills and traits that are relevant for selecting job applicants and predict performance?**
3. **How have you validated the tool in a way that complies with legal and professional standards?**
4. **What specific AI methods are employed in your product (i.e., regression or classification)?**
5. **How are data gathered and prepared when developing the model(s)?**
6. **What empirical evidence can you share that supports the reliability, validity, and fairness of your AI tools (e.g., test–retest reliability, validity coefficients, and group-mean differences)?**
7. **Have you compared your results with traditional alternatives or other AI tools? How confident can I be that your results will apply in my organizational setting?**
8. **At what point are humans involved in the final deliverable or outcome?**

References

Akhtar, R., Winsborough, D., Lovric, D., & Chamorro-Premuzic, T. (2019). Identifying and managing talent in the age of artificial intelligence. In. F. L. Oswald, T. S. Behrend, & L. L. Foster (Eds.), *Workforce readiness and the future of work* (pp. 169-185). New York, NY: Routledge.

Barney, M. (2019). The reciprocal roles of artificial intelligence and industrial-organizational psychology. In R. Landers (Ed.), *The Cambridge handbook of technology and employee behavior* (pp. 38-58). Cambridge, UK: Cambridge University Press.

Brin, D. W. (2019). Employers embrace artificial intelligence for HR. Retrieved from https://www.shrm.org/resourcesandtools/hr-topics/global-hr/pages/employers-embrace-artificial-intelligence-for-hr.aspx

Bravery, K., Brodnikova, V., Cernigoi, A., De Maio, A., Gutowski, D., Hudson, T., Ladd, M., & Silva, J. (2019). Global talent trends 2019 [White paper]. Retrieved from https://www.mercer.com/our-thinking/career/global-talent-hr-trends.html

Caliskan, A., Bryson, J. J. and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183-186.

Campion, M. C., Campion, M. A., Campion E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology, 101*(7), 958-975.

Dastin, J. (2018, October 9). Amazon scraps secret AI recruiting tool that showed bias against women. Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

EY. (2018). The new age: Artificial intelligence for human resource opportunities and functions [White paper]. Retrieved from https://www.ey.com/Publication/vwLUAssets/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions/$FILE/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions.pdf

Guenole, N., & Ferinzig, S. (2018). The business case for AI in HR [White paper]. Retrieved from https://www.ibm.com/downloads/cas/AGKXJX6M

Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology, 98*(6), 1060-1072.

Landers, R. (2019). The existential threats to I-O psychology highlighted by rapid technological change. In R. Landers (Ed.), *The Cambridge handbook of technology and employee behavior* (pp. 3-21). Cambridge, UK: Cambridge University Press.

LeVine, S. (2019, July 22). The coming corporate wipeout. Retrieved from https://www.axios.com/newsletters/axios-future-d3111622-0671-4e93-b185-a0d3793ca86e.html

Marr, B. (2018, December 14). The amazing ways how Unilever uses artificial intelligence to recruit and train thousands of employees. Retrieved from https://www.forbes.com/sites/bernardmarr/2018/12/14/the-amazing-ways-how-unilever-uses-artificial-intelligence-to-recruit-train-thousands-of-employees/

May. R. (2016, May 12). Introducing CV2Vec: A neural model for candidate similarity. Retrieved from https://medium.com/talla-inc/introducing-cv2vec-a-neural-model-for-candidate-similarity-e215b1b12472

Morelli, N. A., & Illingworth, A. J. (2019). The next wave of internet-based recruitment. In R. Landers (Ed.), *The Cambridge handbook of technology and employee behavior* (pp. 59-99). Cambridge, UK: Cambridge University Press.

Persson, A., & Kavathatzopoulos, I. (2017). How to make decisions with algorithms: Ethical decision-making using algorithms within predictive analytics. *Orbit Journal, 1*(2), https://doi.org/10.29297/orbit.v1i2.44

Ployhart, R. E., Schmitt, N., & Tippins, N. T. (2017). Solving the supreme problem: 100 years of selection and recruitment at the Journal of Applied Psychology. *Journal of Applied Psychology, 102*(3), 291–304.

Press, G. (2016, March 23). Cleaning big data: Most time-consuming, least enjoyable data science task, survey says. Retrieved from https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/

Rahwan, I. (2017). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology, 20*(1), 5-14.

Sadiq, S. Z. A. M., Ayub J. A., Narsayya G. R., Ayyas M. A., & Tahir, K. T. M. (2016). Intelligent hiring with resume parser and ranking using natural language processing and machine learning. *International Journal of Innovative Research in Computer and Communication Engineering, 4*(4), 7437-7444.

Smith, A., & Anderson, M. (2017, October 4). Automation in everyday life. Retrieved from www.pewinternet.org/2017/10/04/americans-attitudes-toward-hiring-algorithms/

Society for Industrial and Organizational Psychology. (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). Washington, DC: Author.

Soni, D. (2018, March 22). Supervised vs. unsupervised learning: Understanding the differences between the two main types of machine learning methods. Retrieved from https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society,* 1-17.