ERP Staff versus AI Recruitment with Employment Real-Time Big Data

Zhaohao Sun (zhaohao.sun@gmail.com)
PNG University of Technology
Kenneth David Strang
W3-Research, Data Analytics, St. Thomas, Virgin Islands, 00802

Research Article

Keywords: Human Resource Management, Recruiting, Job Applications, Big Data, Artificial Intelligence (AI), Machine Learning, Quasi-Experiment, ANOVA, Enterprise Resources Planning (ERP)

Posted Date: April 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1515074/v1

ERP Staff versus AI Recruitment with Employment Real-Time Big Data

Kenneth David Strang
W3-Research, Data Analytics, St. Thomas, Virgin Islands, 00802, USA; and
RMIT University, Business Research, North Sydney, 2062, Australia
Phone: +1 518 636 0603 E-mail: professor@kennethstrang.com
ORCID: 0000-0002-4333-4399

Zhaohao Sun*
Department of Business Studies
PNG University of Technology, Private Mail Bag, Lae 411, Morobe, PNG
Phone: +675 473 4406 (O), Fax: +675 475 7667
zhaohao.sun@pnguot.ac.pg, zhaohao.sun@gmail.com
ORCID: 0000-0003-0780-3271
* The corresponding author for this paper.

Abstract: The purpose of this study was to compare the performance of recruiters and an artificial intelligence (AI) program for processing internet employment big data in the enterprise resources planning (ERP) function. Previous AI implementations were discriminatory. Thus, the research question was: Could ERP staff perform better than AI in selecting the best candidate from internet employment real-time big data. A quasi-experiment was created using primary data. Job criteria were developed using machine learning to identify key skills from existing staff in a case study company. The skills were transformed into hiring criteria and a job description. AI software and a random sample of ERP recruiters assessed the same internet-based real-time big data. The results were compared between the ERP staff and AI using ANOVA followed by post-hoc Tukey. Contrary to the research question, AI out-performed ERP staff. The proposed approach might facilitate the research and development of big data, data analytics, artificial intelligence, and human resource management.

Keywords: Human Resource Management, Recruiting, Job Applications, Big Data, Artificial Intelligence (AI), Machine Learning, Quasi-Experiment, ANOVA, Enterprise Resources Planning (ERP)

1 Introduction

The research problem investigated in the current study was whether enterprise resource planning (ERP) staff can perform better than artificial intelligence (AI), for recruiting new talent from the internet, when there was a high volume and high velocity of real-time application data. The term 'performing better' could mean faster and more objectively, with less discrimination. Recruiting is a strategic and critical function, among many others, performed within an organization's ERP function [1]. The recruiting function has become more digital and virtual due to new technology as well as the travel constraints imposed by the recent COVID-19 pandemic [2]. Therefore, AI has been used in some ERP functions, to process high volumes of big data [3, 4]. Most of the recent literature reported that AI was beneficial for ERP functions including recruiting new talent [5-7].

However, AI caused many ERP problems according to some researchers [8-11]. Amazon's 'secret hiring AI' used during 2014-2017 was found to bias job candidates if the applications contained a woman's salutation or name, and excluded certain colleges [9]. A BBC journalist researched and experimented with recruiting AI including Pymetrics, Textico, and HireVue, asserting women and people of colour were being overlooked by AI ─ including herself [12]. MIT researchers uploaded a fake job to My Interview and Curious Thing recruiting AI software then they applied as an experiment ─ both programs failed, for example, by scoring a candidate high for English proficiency when she spoke only in German [13]. Thus, AI created significant problems in the ERP function - discrimination is a serious legal problem in developed countries.

Despite some problematic implementations of AI in ERP for recruiting, there was a counter argument in the literature that AI could be beneficial [6]. In a hiring experiment conducted by Bertrand and Mullainathan [14] where they used fake resumes to apply for jobs in two highly-populated U.S. cities, they found that there was significant racial discrimination by humans even at companies who had claimed to be equal opportunity employers eligible for favourable tax and other government considerations. In a meta-analysis, Fraij and Laszlo [5] concluded that AI was beneficial for recruiting such as automating time-consuming processes like screening job applications, providing human resource management staff with more quality time to focus on strategic activities. Other researchers claimed that AI was valuable and much needed to automate some of the ERP labour-intensive recruiting processes, although careful human oversight was needed [15, 16].

The above rationale, grounded in the literature, led to the research question (RQ): Could ERP staff out-perform AI, without discriminating, to objectively select the best job candidate from internet-based real-time big data. A quasi-experiment was designed to answer the RQ, using integrated methods. Pragmatic approaches were used to overcome the complexity of comparing AI and human resource staff performance on recruiting with high volume high velocity internet-based real-time big data.

The rest of this paper is organized as follows: Section 2 provides literature review for this research and presents big data derived small data analysis. Section 3 provides research approach of this research. Section 4 explores the experiment, participants, and datasets for our research. Section 5 discusses the research results. Section 6 points out the limitations and future research directions. The last section ends this paper with some concluding remarks and recommendations for future research.

2 Literature Review

This section provides a background on big data, big data analytics, machine learning, data mining, artificial intelligence, data science and ERP human resources management as related to the current study.

2.1 Big Data and Analytical Approaches

Big data is considered a frontier for research and development in many disciplines since 2012 [2, 17]. Big data is a transforming science which has impacted engineering, technology, medicine, healthcare, finance, business and management, education, and ultimately society itself due to the emergence of big data analytics [18]. The big v's refer to the five distinguishing characteristics of big data: Big volume, big velocity, big variety, big veracity, and big value [19, 20]. Other researchers have added more features to help describe big data. Therefore, big data should have all the basic five characteristics to be considered big data, although some scholars argue it is enough to meet only one of the five characteristics, since each one is beyond analysis with most contemporary methods of statistical software [21, 22].

Big data analytics is a science with integrated technology for organizing big data, analysing and discovering knowledge, patterns, and intelligence from big data, as well as visualizing and reporting the discovered knowledge for assisting decision making [22]. The main components of big analytics include big data descriptive analytics, predictive analytics, and prescriptive analytics [22], which correspondingly address the three questions of big data; When and what occurred? What will occur? What is the best answer or choice under uncertainty? All these questions are often encountered in almost every science, technology, business, management, organization, and industry [3, 23].

Machine learning (ML) is one of the most popular techniques used for analysis of big data [3, 23]. ML is concerned about how computers can adapt to new circumstances, as well as to detect and extrapolate patterns [4]. The essence of ML is that it is an automatic process of pattern recognition by a learning machine [23]. Machine learning mainly aims to build systems that can perform at or exceed human level competence in handling many complex tasks or problems [4].

We are in the age of AI and big data. AI and big data have been applied to almost every sector and have been revolutionizing our work, lives and societies. AI is concerned with imitating, extending, augmenting, amplifying, automating intelligent behaviours of human beings [4]. AI attempts not only to understand how human think, understand, write, learn, act rationally and smartly, but also to build intelligent entities that can think, write, perceive, understand, predict, and manipulate a world.

The relations among deep learning, machine learning, and artificial intelligence are mathematically represented as follows: deep learning $\sqsubset$ machine learning $\subset$ artificial intelligence. In other words, deep learning is a subset of machine learning, and machine learning is a subset of artificial intelligence [4, 23].

Data mining is a process of discovering various models, summaries, and derived values, knowledge from a large database [24]. Data mining includes descriptive data mining and predictive data mining [24]. The former produces new nontrivial patterns and knowledge, while the latter produces models and roles of the systems. The primary tasks of descriptive data mining include clustering, summarization, dependency modelling. The primary tasks of predictive data mining include classification, regression, change and deviation detection [1, 4].

The relations among big data, data mining, and big data analytics are mathematically represented as follows: data mining $\sqsubset$ big data analytics $\subset$ big data $\subset$ data science. Data scientists aim to invent data intelligence-driven technologies and machines to represent, learn, simulate, reinforce, and transfer human-like intuition, imagination, curiosity, and creative thinking through human-data interaction and cooperation [17, 23].

2.2 Methods for Searching Big Data

Practitioners across all disciplines are living in the age of big data. However, sometimes, practitioners can only process small data analysis. This is a pragmatic contradiction. It may be a challenge to face the relativity between big data and small data. This subsection presents a dialectic unification between big data and small data by proposing a big data-derived small data approach based on [28]. As a process, a big data derived small data approach consists of: 1. Big data search; 2. Big data reduction; 3. Big data derived small data collection; and 4. Big data derived small data analysis. Big data search is the first step for the big data derived small data approach. Search all possible data, if not all existing data, is the important task for any research activity. It is also important for literature review, for imagination, and for association. In the big data search process, we can lead to literature review, imagination, association, and revision what we have planned and designed. This also implies that research as a search is the first step. In this step, we must identify where is the big data resource for our research?

In this research, we select Google Scholar as the big data resource, because it is the largest scholar publication base in the digital world although EBSCO, Semantic Scholar, and Researchgate are good open-access (free) platforms for locating research. Therefore, we use Google Scholar to conduct big data search. Big data reduction is the second step for the big data derived small data approach. Reducing big data is, in essence, a kind of selection. The proper selection of data is to address how can we search big data?

To address the second question, big data should be reduced in the big data search. That is, whenever searching, we must keep in mind that, big data reduction is critical for any big data search, otherwise, big data search would become the search of all the big data, for example using Google for the entire Web. The heuristic method to address the second question is that we first analyse the research title and initial proposal or abstract of the research and obtain the important keywords or clauses in order to narrow the search space. For our search, we should select 1. big data human resources, 2. Artificial intelligence human resources, 3. big data recruitment, 4. Artificial intelligence recruitment. Because artificial intelligence is usually abbreviated as AI, we should also select 5. AI recruitment. This search has revolutionized our tradition that we usually relied on a few articles published in a few top journals and so-call important principles and results published in a few classic books or textbooks. The latter is equal to all the data searched from the big data search.

Big data derived small data collection is a special kind of sampling. Sampling can be applied to big data as a form of big data reduction. For example, Google Scholar should be a sampling, because Google Scholar cannot collect all the data of scholars on the Internet. There are two core parts for any sampling towards data analysis based on statistical inference. One is to collect what kind of data. The second is how to collect data. The former is related to what kind of data were important for the research. In other words, importance of data is related to data analysis. The latter has been discussed in terms of statistical sampling. For the importance of data, we argue not all data need be taken for any decision making and rule-seeking as well as statistical inference. Just as focusing on main problems with main solutions, one can also seek the important data for any decision making and statistical inference. Therefore, it is a big issue for research to identify which data set is important to meet the objectives of the research.

For this research, what kind of data is important for this research to examine big data, AI and recruiting? The possible answers are four types of data, that is, data on 1 "big data human resources, 2. Artificial intelligence human resources, 3. big data recruitment, 4. Artificial intelligence recruitment. 5. AI recruitment" from Google Scholar (https://scholar.google.com/). Using this method, this research should collect up to 100 × 5 data results, that is, if we collect the data on the first 100 scholars' publications in each area with highest citation, then we can know the state-of-the-art big data, AI, and recruiting and their relationships.

We have still reduced the results using the criteria of citation, if the search key words or clauses are in the publication title or abstract. When we use Google Scholar search for a key words x, there are a lot of searched results. This research collects about 100 × 5 data results. The 100 × 5 data items are a small data, but it is derived from the big data of Google Scholar (http://scholar.google.com). Therefore, it is a

big data derived data collection. This data collection is, in essence, a big data reduction for the proposed research.

Big data derived small data analysis is important both for big data approach and big data analytics as a discipline [28]. First of all, big data has basically been controlled by many global data giants such as Facebook, Google, Tencent, Baidu, and Alibaba rather than by an individual scholar. It is almost impossible for a scholar to use the big data of the mentioned giants to do research on big data driven intelligent recruitment or similar. It is too expensive or unaffordable for a scholar to collect data and analyse the collected data, because s/he has not a platform similar to that of the mentioned giant.

Secondly, most statistical inference based on sampling is reasoning based on incomplete knowledge or data [28]. Therefore, most statistical modelling or inference is a kind of big data derived small data analysis and reasoning [29].

Thirdly, from a data processing viewpoint, the largest data analyses could be performed in large data centers of a few global data monopolies running specialized software such as Hadoop over HDFS to harness thousands of cores to process data distributed throughout the cluster [29]. This means that individuals have to use big data derived small data analysis to analyse small but quality data.

Finally, any research in general and research publications in particular are, in essence, based on big data derived small data analysis, because an average research publication consists of 30 references, which has only up to 30 MB of data from a data volume viewpoint [17, 23]. In the big data world, the data with 30 MB is relatively small. In comparison, Amazon (AWS) and Google have processed 500 exabytes (EB) and 62 petabytes (PB) of big data in 2021 respectively [1, 2] where 1 EB = 1024 PB=1024 x 1024 TB.

This research will use the above proposed method to look at how AI and big data to drive HR management in general and recruiting in particular in next Sections.

2.3 AI and Big Data in ERP Talent Recruitment

In this subsection, AI and big data problems are highlighted from the literature, particularly within the ERP talent recruitment function. The focus in this section is on empirical analysis of big data using AI. In other words, the preliminary literature RQ driving this section was: what actual AI techniques have been applied in ERP talent recruiting, and what were the key findings? This section is original in the literature because there are no other publications examining this preliminary literature RQ. In this section the methodology is explained for conducting the literature review of AI in ERP using big data-derived small data analysis.

The first step in the literature review for this section was done using Google Scholar (https://scholar.google.com/). We searched "AI Human resource", "artificial intelligence human resource", "big data human resource", "artificial Intelligence recruitment", "big data recruitment", and "ERP HR recruiting AI". Then notational relationship between these search terms are as follows:

big data recruitment $\subseteq$ big data human resource (HR)

ERP recruitment big data $\subseteq$ big data human resources management (HRM)

artificial intelligence recruitment $\subseteq$ artificial Intelligence human resource (HR).

A Google Scholar search for "AI Human resource" (March 22, 2022) found 85 results. We first analyse the titles of the found results (based on the ranking of Google Scholar) in terms of "AI Human resource". A Google Scholar (https://scholar.google.com.au/) search for "artificial intelligence human resource" found 95 results (retrieved on March 27, 2022).

It should be noted that only the first 70 results are related to the search key words. The Google scholar's search results are not exact what we expected mathematically and algorithmically. For example, almost all the search engines on the web have not really realized the modus ponens, that is, $If\ A, A \rightarrow B, then\ B$. What Google or Google Scholar has provided is B', it satisfies $B \sqsubset B'$. We can call these as inclusion intelligence. Almost all the search engines use inclusion intelligence to provide the searched results to the users online. However, this is a weak intelligence from a mathematical viewpoint.

Comparing with the search for "AI Human resource", Google scholar's search for "artificial intelligence human resource" found more results and more scholarly results. A Google Scholar (https://scholar.google.com.au/) search for "big data human resource" find 85 results (retrieved on March 27, 2022). Then we process these 85 results and delve into each of them. The analysis demonstrates that only the first 60 results are related what we searched. Among them there are also many results that are not really related to big data human resource.

Zang and Ye discussed how to apply big data to recruit talents for the enterprise [6] They claimed that there are often prejudices in the recruiting. They asserted that big data can make the recruiting fairer. This is because big data from the social networking platform can be viewed by peers and supervisors so candidates will likely be honest with their public profiles and resumes. Companies incorporate the recruitment within online social networking services, and constantly gather resume information and background information such as, personal videos, personal pictures, living conditions, social relationships, and the self-reported competencies of the applicants. According to statistics, China has more than two-thirds of enterprises use online recruitment [25]. In fact, LinkedIn has been used as a part of recruiting talents. Google Scholar, Semantic-scholar, LinkedIn, and Researchgate can provide big data on research activities, publications, and the background information of any scholar or job candidate [18, 26]. The information stored on the internet by those companies could be considered internet-based real-time big data. That big data is constantly changing, some of it is duplicated, and it is spread around the world on different servers. That data can be utilized in the ERP function of a company or university to identify potential suitable candidates based on searches of mandatory hiring keywords. That big data may be evaluated and used to rank the candidates for a job opening. Interestingly, that same big data can also be evaluated to determine if the ERP talent recruitment process was fair, without race or skin color discrimination.

A Google Scholar (https://scholar.google.com.au/) search for "big data recruitment" find 31 results (retrieved on March 27, 2022). The algorithm in recruitment combining ML to assess big data sets in order to evaluate technical talent has been used in the recruitment practice [5, 10, 27]. Therefore, algorithmic recruitment is a kind of AI and big data driven recruitment. Evaluating big data to increase the efficiency and accuracy of recruitment and hiring decisions has drawn significant increasing attention [17]. Recently, a growing number of companies have used big data to build platforms and tools to help recruiters identify promising talents for jobs. Companies have also collected big data from a wide range of online social networking platforms such as Facebook, WeChat, and LinkedIn, where the potential candidate share their expertise from daily life to professional work [7]. Therefore, big data has made the personal skill, knowledge, and performance transparent to the world. In the end of the day, big data driven recruitment will became fair, fairer than the traditional one with obvious various discriminations, because openness and global transparency can mitigate the prejudice and discriminations often happened in the recruitment process.

A Google Scholar (https://scholar.google.com.au/) search for "artificial Intelligence recruitment" returned 163 results (retrieved on March 30, 2022). We analysed the results based on the rank and order of Google Scholar. The preliminary analysis showed that AI recruitment has become an effective component of some firm's entire promotion and human resource talent management processes [8]. For example, AI techniques have been used in interviewing and assessment of candidates [2, 8, 10]. More generally, AI has been used to automate the recruitment process in human resources including tracking of applications, performance reviews, on boarding, compensation and career management, unbiased screening candidates, scheduling and in decreasing of unconscious bias [2, 10].

ERP Human Resources Management (HRM or HR) is the management of people in a company or organization to achieve success in business performance. HRM is a strategic activity with the ERP function [1, 4, 10]. In the ERP function, HR departments are responsible for employee recruitment, overseeing employee-benefits design, training and development, performance appraisal, and reward management, such as managing pay and employee-benefits systems [16, 23]. Recruitment is a process of identifying, sourcing, screening, shortlisting, interviewing candidates and offering the contracts for

jobs (either permanent or temporary) provided by an organization. AI and big data have significant impacts on ERP in general and HRM recruitment in particular. For example, ERP was conceived to empower human resources management to transform data into strategic knowledge and to develop talent through recruiting as well as training [1].

3 Research Approach

From a viewpoint of research methodology, this research uses a multidisciplinary approach consisting of the above proposed big data-derived small data approach, a pragmatic-driven ideology, practical analytical methods, computer tools, a big data-driven quasi-experiment, and formal statistical techniques. This research used pragmatic-driven ideology, consisting of customized literature search and reduction techniques, and Microsoft Excel to collect and analyse the literature review.

The researchers held a pragmatic ideology, meaning that quantitative factual evidence was sought to prove deductive theories but practical approaches were designed to conduct an otherwise difficult to execute experiment. This is can be called pragmatic-driven research design. The pragmatic ideology was applied because the researchers found that it was impossible to conduct a controlled experiment to compare an AI recruiter with ERP staff when collecting internet-based job application real-time big data. Internet-based real-time big data cannot be stored in a single computer or a cloud database, at least not at the time of writing. Mixed sequential methods were applied, starting with machine learning at a case study site, followed by a quasi-experiment to collect primary data. Analysis of variance (ANOVA) and then Tukey post-doc techniques were applied to analyse the results in SPSS version 25. The authors developed two hypotheses based on the RQ and the literature review:

H1: Are the score means different between the AI recruiter and ERP staff?

H2: Could ERP staff outperform the AI, by objectively selecting the best candidate from real-time big data using the same hiring criteria?

4 Experiment

A case study company was selected. It is a government consulting firm in the IT sector based in the USA. The first author was the principal investigator (PI). Approval was given by the PI's employer. The PI conducted the data collection and quasi-experiment. Informed consent was obtained from the case study company, on condition that everything except the data results would remain confidential. The case study company nominated one of their best performing departments, the project management office (PMO). Project managers working in the PMO of the case study company were asked to update their resumes. Their resumes were parsed into fields and then assessed by nearest neighbor analysis in machine learning to identify the most important skill clusters. The most important clusters were used to build a job description with 21 mandatory criteria. An AI recruiting program, CVViZ, was selected as an AI recruiter. An 'IT project manager' job was created using the highest-ranked skills as criteria. A best resume was not used when prompted by the AI recruiter to avoid learning discriminatory behaviors. An internet search was activated to load relevant resumes and cover letters from multiple recruiting databases and boards. The AI recruiter was used to screen candidates, and the scores were exported for the top candidate.

Next a random panel of recruiters were selected from LinkedIn social media, by searching for 'recruiter + IT' and inviting participants. The first 10 who worked in the ERP function as recruiters for the IT industry were selected, and the expression of interest poll was closed. Participants were given access to the job description and the 21 mandatory hiring criteria. All recruiters and the AI recruiter had access to the same candidate real-time big data based on the internet. No local database was used. The recruiters and the AI recruiter were constrained by 8 hours, to simulate 1 work day, to complete the task. Recruiters were tasked to select candidates from the internet using the mandatory hiring criteria and to score enough to determine the best candidate. The scores of their top candidate were exported into SPSS.

The PI checked the accuracy of the exported data and compared the scores against the best candidate resumes to ensure the results were reasonable. For an example of one criterion, the resume or cover letter

needed to state risk management experience was applied at a company or project. No outliers or unusual evaluations were found. ANOVA was applied to compare the scores of the 21 criteria between the 10 recruiters and the AI recruiter to determine if they were significantly different. A post-hoc Tukey analysis was performed to indicate which was significantly different. The means were sorted to illustrate which score was highest.

5 Results and Discussion

The demographic characteristics of the 10 recruiters indicated all had at least 5 years of experience (M=7.2, SD=1.6), slightly more than half were female (57%), the average age was 37 (SD=4.7) and all were employed in the IT staffing industry with a focus on IT project management. Most of the recruiters had a bachelor level education (52%), one had a Master, and the rest had Associate level degrees. All the best-candidates selected by the 10 recruiters and the AI recruiter were different. The AI recruiter and 10 recruiters were allowed 8 hours. The AI recruiter completed the analysis in 2 hours while all recruiters used most of their 8 hours (as judged by when they submitted the candidate selection in the system).

The first goal was to answer the RQ by testing the first hypothesis (H1): Are the score means different between the AI recruiter and ERP staff? The result was yes, H1 was accepted. Table 1 contains the key estimates from the ANOVA technique, where evaluation score was the dependent variable, while the number of the 10 recruiters with the AI recruiter = 11 were the fixed factors. The result was significant based on F (220, 10) = 17.2, $p < .001$ with an effect size of 44% ($\eta^2 = 0.439$, $\omega^2 = 0.412$).

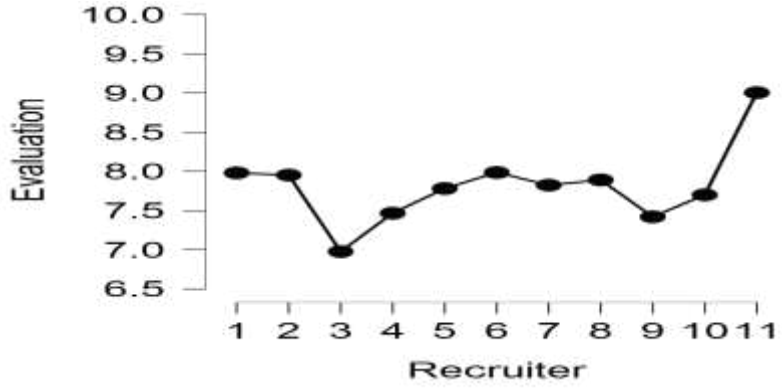
Table 1. ANOVA estimates of comparison between recruiters and the AI recruiter

Cases	Sum of Squares	DF	Mean Square	F	p	η^2	η^2_p	ω^2
Recruiter	52.185	0	5.218	17.201	< .001	0.439	0.439	0.412
Residuals	66.745	20	0.303					

Next the second hypothesis (H2) was examined, that the ERP staff could outperform the AI, by objectively selecting the best candidate from real-time big data using the same hiring criteria. As discussed, prior to any statistical analysis, the PI had reviewed all 11 best-candidate selections, in an objective manner, to ensure they roughly matched the scores given on each hiring criterion from the job description. To evaluate H2, a Tukey post-hoc analysis was performed to detect if the AI recruiter's score was significantly lower than any of the 10 human recruiters. Contrary to our second hypothesis (H2), the result was that the AI outperformed all 10 recruiters with a higher score ($p < .001$), so H2 was rejected. Nonetheless, despite H2 being rejected, it was surprising that the AI outperformed all 10 ERP staff. This will be analysed in more detail below.

Figure 1 is a line plot illustrating the candidate score by the recruiters and the AI recruiter where the AI recruiter is on the right (recruiter 11 in the exported data). The y-axis score was converted from a 100% basis to a score out of 10 for display purposes.

Fig. 1 Job candidate evaluation scores of 10 recruiters and AI recruiter (11) using real-time big data



The descriptive estimates of the job candidate evaluation scores exported form the 10 human recruiters and the AI recruiter are listed in Table 2. These are sorted in descending order of evaluation score. The AI recruiter had the highest candidate score (M=90, SD=7.95) which was significantly higher than any of the recruiters. Respondents 6, 1 and 3 had similar candidate scores, while recruiter 3 had the lowest (M=69.8, SD=3.05).

Table 2. Descriptive statistics of evaluations

Recruiter	Mean	SD
AI	90.05	7.95
6	79.86	2.97
1	79.81	4.93
2	79.52	5.16
8	78.91	4.88
7	78.24	4.45
5	77.81	6.89
10	77	5.1
4	74.67	6.41
9	74.2	6.67
3	69.76	3.05

Clearly the AI selected the best candidate since the score was significantly higher. However, we must keep in mind that all recruiters including the AI recruiter selected different best-candidates. Despite that the PI ensured that the evaluation scores were representative of the application data in each case, we do not know conclusively if the AI recruiter picked the best candidate from all internet application big data. That task would be daunting to prove. Nevertheless, we can conclude from this practical quasi-experiment that the AI recruiter outperformed the 10 recruiters, which answered the RQ and the hypothesis.

6 Limitations and Future Research Directions

Before anyone jumps to conclusions that this finding is generalizable everywhere, the limitations must be emphasized. This was a quasi-experiment, with very little control. Only the tool was controlled (the AI software holding the job hiring criteria and selected applications). The sample size was small with 10 recruiters randomly selected on a social media platform using a poll feature which was closed in a day (to purposively limit too many respondents which would have caused the study to be unmanageable). There was minimal purposeful sampling because the respondents self-selected. Since the applications

were sourced form internet big data, we cannot know if during the 8-hour time period over 1 day that some applications were deleted by candidates or maybe new ones would have been added. Thus, the data was not a point in time which the authors argue is impossible with big data (otherwise it is not really big data). Even so, this is a practice based on the principle of big data driven small data analysis [28].

Another limitation that we do not know what role pure chance or systematic error played, due to the quasi-experimental design without a control. For example, the time order of candidate applying might play a role since the AI recruiter was able to select better applications before those expired or were deleted from the internet. This was done without statistical replacement, so we do not know how well the AI recruiter would perform as compared to each of the 10 recruiters if everyone evaluated the exact same resume and cover letter. As noted earlier, it was not possible in order to satisfy the RQ, but it is certainly recommended in future research designs.

In summary, the current quasi-experiment was very limited. It was exploratory but with interesting and provocative results. Other researchers would need to replicate this and use alternative methods, possibly multiple case studies with embedded controlled experiments, to more fully compare the AI recruiter with experienced recruiters. Additionally, other recruiting AI software should be tested in a scientific controlled manner such as in the current study.

7 Conclusion

In current study, a pragmatic quasi-experiment was used to answer the RQ. This was because the unit of analysis involved real-time big data, actual job applications sourced from the internet. It was not practical to download these, and using a sample would have defeated the goal to utilize actual big data. It was also not practical to control the experiment by selecting a resume or creating a fake one as done by other researchers [13-16]. Creating one resume and cover letter for experimental control would have cancelled the realistic effect of analyzing big data. The purpose was to force the AI recruiter and recruiters to analyze actual relevant real-time big data.

Unfortunately, we had to reject our hypothesis that the ERP staff could outperform AI for objectively selecting the best candidate from internet real-time big data. Nonetheless, we learned a valuable lesson which can be shared with the research community of practice, for the purposes of stimulating more studies. In the current quasi-experiment, the AI recruiter outperformed all 10 recruiters on the time, by a factor of more than 4 (2 hours for the AI recruiters versus 8 hours for the ERP staff). The AI recruiter also outperformed each of the 10 recruiters by selecting a better candidate with a statistically significant higher score (90 versus the next best human recruiter evaluation of 79.9). Based on this the RQ and hypotheses were answered but the deductive hypotheses were rejected. This was honestly rigorous and valuable science where the methods were explained in detail and the results were truthfully reported, not predetermined.

We might partially explain our results as falling into Turing Test trap, a well-known phenomenon introduced by Alan Turing in 1950 [30]. In his experiment, computers outperformed human intelligence due to faster comparison algorithms where words, numbers or other data were compared in tables. The point was, even though computer-powered AI may be faster than humans in certain comparative tasks, this does not necessarily mean the AI software is better. Humans create AIs and humans are the ultimate decision makers. This is still a very controversial topic in computer science, data science, cognitive science, and philosophy. This is the reason why we are studying AI with humans in this quasi-experiment.

The current study addressed other issues from the literature review. Some researchers proved the human recruiters have biases and they discriminate [8, 10, 27]. For example, human recruiters may consciously or subconsciously not select candidates who are a certain gender or race, who did not attend an 'ivy college' of the USA nor 211 985 universities of China, who are too old, who are international, and even applicants who are not located in the city zip code of the posted job [14, 16]. The current study addressed those controversies to some extent, by listing hiring criteria identified by a relevant case study company, vetted by their HRM leaders, despite those skills were based on the resumes of their own high performing staff. In the current study, the hiring criteria were worded to be discrimination free, gender

and race based criterion were not used, no specific ivy school or accredited university in a certain country or state was mandated (or even mentioned). This forced the human recruiters (and the AI recruiter) to look for those hiring criteria and score against those, not a natural language learning algorithm to uncover personality clues from what or how candidates wrote their applications.

Furthermore, a few researchers found that recruiting AI software could also discriminate, not prejudice in the same way as humans, but ML can make errors based on the 'training' procedure [9, 15, 16]. Remember, there was the Amazon 'secret hiring AI' used during 2014-2017, which rejected applications containing a woman's salutation or name, and excluded certain colleges [9]. ML works by identifying patterns and looking for those. Recruiting AI software which generally includes ML usually prompts HRM staff to upload example best-in-class resumes, such as those already in the company. In the current study, this drawback was overcome because no best-in-class resume was uploaded into the hiring AI. Instead, which was a lot more work, hiring criteria were developed using experts at a relevant case study company, and those were input into the AI for selection criteria. Nonetheless, there was a BBC journalist who experimented with recruiting AI whereby she concluded women and people of color were being overlooked ─ including herself [12].

In the current study, AI discrimination was prevented because the PI made sure that the hiring criteria were not discriminatory in a way, including not citing a specific location or race/gender attributions. Finally, recall the MIT experiment to evaluate two hiring AI recruiters using a fake job, whereby both programs failed since among other errors a candidate was scored high for English proficiency when she spoke only in German [13]. The current study avoided this problem because live interviews were not included in the quasi-experiment, since the purpose was to use the AI recruiter to select candidates from internet application big data. To carry this theme further, the authors would suggest the next step if this were an actual hiring event would be to interview all 11 candidates. Alternatively, the authors suggest the AI recruiter could be used to identify 5-10 best-candidates, and then experienced human recruiters would take over and interview those, with humans making the complex final decision of the best-candidate to hire.

1. Disclosure of potential conflicts of interest

A: No authors reported any conflicts of interest.

2. Research involving Human Participants and/or Animals, and 3. Informed consent

A: The research project was approved by Strang's employer and all participants gave ethical consent.

4. Funding

A: This project is partially funded to Prof Zhaohao Sun by a national project titled "Big Data Driven Socio-economic Development in PNG", supported by PNG Science and Research Secretariat (Grant No: 1-3962 PNGSTS, 2021).

5. Author contributions

A: Sun served as the corresponding author. Strang initiated, designed the study, and served as principal investigator (PI). Strang and Sun conducted an iterative literature review, update, and refine the manuscript. Strang conducted and interpreted the experiment and quantitative analysis. Sun explored big data derived small data analysis. Strang wrote the first manuscript draft. Sun and Strang collaborated equally in writing and refining the paper.

References

[1] K. G. Laudon and K. C. Laudon, *Management Information Systems: Managing the Digital Firm (16th Edition)*. Harlow, England: Pearson, 2020.

[2] S. Bhardwaj, "Data analytics in small and medium enterprises (SME): A systematic review and future research directions," *Information Resources Management Journal*, vol. 35, pp. 1-18. http://doi.org/10.4018/IRMJ.291691, 2022.

[3] R. Geetha and S. R. Bhanu, "Recruitment through Artificial Intelligence: A Conceptual Study," *International Journal of Mechanical Engineering and Technology*, vol. 9, pp. 63-70, 2018.

[4] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (4th Edition)*. Upper Saddle River: Prentice Hall, 2020.

[5] J. Fraij and V. Laszlo, "A literature review: Artificial intelligence impact on the recruitment process," *International Journal of Engineering and Management Sciences*, vol. 6, pp. 108-121. doi:10.21791/IJEMS.2021.1.10, 2021.

[6] S. Zang and M. Ye, "Human Resource Management in the Era of Big Data," *Journal of Human Resource and Sustainability Studies*, vol. 3, pp. 41-45, 2015.

[7] N. Nawaz, "How far have we come with the study of artificial intelligence for recruitment process," *International Journal of Scientific & Technology Research*, vol. 8, pp. 23-41, 2019.

[8] H. Rodney, K. Valáškova, and P. Durana, "The Artificial Intelligence Recruitment Process: How Technological Advancements Have Reshaped Job Application and Selection Practices," *Psychosociological Issues in Human Resource Management*, vol. 7, pp. 42-47, 2019.

[9] P. Henman, "Improving public services using artificial intelligence: possibilities, pitfalls, governance," *Asia Pacific Journal of Public Administration*, vol. 42, pp. 209-221. doi:10.1080/23276665.2020.1816188, 2020.

[10] Z. Su, G. Togay, and A.-M. Cote, "Artificial intelligence: a destructive and yet creative force in the skilled labour market," *Human Resource Development International*, vol. 24, pp. 341-352. doi:10.1080/13678868.2020.1818513, 2021.

[11] G. Comandè, "Regulating algorithms' regulation? First ethico-legal principles, problems, and opportunities of algorithms," in *Transparent Data Mining for Big and Small Data. Studies in Big Data*, vol. 32, T. Cerquitelli, D. Quercia, and F. Pasquale, Eds. NY: Springer, 2017, pp. https://doi.org/10.1007/978-3-319-54024-5_8.

[12] A. Murad, "The computers rejecting your job application," in *BBC News [Business Section]*, World ed. London, 2021, pp. 1-9.

[13] P. R. Daugherty, H. J. Wilson, and R. Chowdhury, "Using artificial intelligence to promote diversity," *MIT Sloan Management Review*, vol. 60, pp. 59-63, 2019.

[14] M. Bertrand and S. Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *National Bureau of Economic Research Working Paper Series*, vol. No. 9873, http://www.nber.org/papers/w9873.pdf, 2003.

[15] H. Ming-Hui and R. T. Rust, "Artificial Intelligence in service," *Journal of Service Research*, vol. 21, pp. 155-172, 2018.

[16] A. K. Upadhyay and K. Khandelwal, "Applying artificial intelligence: Implications for recruitment," *Strategic HR Review*, vol. 17, pp. 255-258, 2018.

[17] L. Cao, "Data science: challenges and directions," *CACM*, vol. 60, pp. 59-68, 2017.

[18] H. Ekbia, M. Mattioli, I. Kouper, G. Arave, A. Ghazinejad, T. Bowman, V. R. Suri, A. Tsou, S. Weingart, and C. R. Sugimoto, "Big data, bigger dilemmas: A critical review.," *Journal of the Association for Information Science & Technology*, vol. 66, pp. 1523-1545, 2015.

[19] U. Jovanovi, A. Stimec , and D. Vladusi, "Big-data analytics: A critical review and some future directions," *International Journal of Business Intelligence and Data Mining*, vol. 10, pp. 337-355, 2015.

[20] J. Fan, F. Han, and H. Liu, "Challenges of Big Data Analysis," *National Science Review Journal*, vol. 1, pp. 293-314, 2014.

[21] N. Terry, "Navigating the Incoherence of Big Data Reform Proposals.," *Journal of Law, Medicine & Ethics*, vol. 43, pp. 44-47, 2015.

[22] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, pp. 137-144, 2015.

[23] R. Sharda, D. Delen, and E. Turba, *Business Intelligence and Analytics: Systems for Decision Support (10th Edition)*. Boston, MA: Pearson, 2018.

[24] P. Zikopoulos, C. Eaton, D. DeRoos, T. Deutsch, and G. Lapis, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. NY: McGraw-Hill Osborne Media, 2011.

[25] R. Y. Zhong, G. Q. Huang, S. Lan, Q. Y. Dai, X. Chen, and T. Zhang, "A big data approach for logistics trajectory discovery from RFID-enabled production data.," *International Journal of Production Economics*, vol. 165, pp. 260-272, 2015.

[26] M. A. Rothstein, "Ethical Issues in Big Data Health Research: Currents in Contemporary Bioethics.," *Journal of Law, Medicine & Ethics*, vol. 43, pp. 425-429, 2015.

[27] C. Barabas, "Engineering the American dream: a study of bias and perceptions of merit in the high-tech labor market," vol. 2022. Boston, MA: MIT, 2015.

[28] Z. Sun, and Y. Huo, "The spectrum of big data analytics," Journal of Computer Information Systems, 61(2), 154-162. doi:10.1080/08874417.2019.1571456, 2021.

[29] National Research Council, Frontiers in Massive Data Analysis. Washington DC: The National Research Press, 2013.

[30] A. M. Turing, "Computing Machinery and Intelligence," Mind, vol.59, No. 236, pp. 433-460, 1950.