

A guide for CEO's ,
General Counsels
and HR leaders on
smart, safe selection
of Ai Hiring Solutions



Ethical Ai. In 8 critical questions



Contents

PREFACE	3
A Recruiters 'Use Case' on Tackling Bias	5
1. What training data do you use?	6
2. Can you explain why 'person y' was recommended by the Ai and not 'person z'?	9
3. What assumptions and scientific methods are behind the product? Are they validated?	12
4. What are the bias tests that you use and how often do you test for bias?	13
5. How can you remove bias from an algorithm?	15
6. On which minority groups have you tested your products?	16
7. What kind of success have you had in terms of creating hiring equity?	16
8. What is the composition of the team building this technology?	17
LEARNING BITE	18



“There will always be people who are ahead of the curve, and people who are behind the curve. But knowledge moves the curve.”

Bill James, Baseball Statistician

Bill is the statistician who inspired the book and film Moneyball.

He invented predictive analytics for baseball, using statistical data to determine why teams win and lose, which he coined ‘sabermetrics’.

This kind of statistical analytics is now moving into the workplace as HR teams around the world deploy predictive talent models that can more effectively—and more rapidly—identify, recruit, develop, and retain the right people.

Leveraging objective data to augment decisions *like who to hire and who to promote* is critical if you are looking to minimise unconscious preferences and biases, which can surface even when those responsible have the best of intentions.

The greatest algorithm on earth is the one inside of our skull, but it is heavily biased. Human decision making is the ultimate black box. Only with data, the right data alongside human judgment can we get any change happening. And clearly, what your employees and candidates are now looking for, is change.

We hope that the debate over the value of diverse teams is now over. There is plenty of evidence that diverse teams lead to better decisions and therefore, business outcomes for any organisation.



This means that CHROs today are being charged with interrupting the bias in their people decisions and expected to manage bias as closely as the CFO manages the financials.

But the use of Ai tools in hiring and promotion requires careful consideration to ensure the technology does not inadvertently introduce bias or amplify any existing biases.

To assist HR decision-makers to navigate these decisions confidently, we invite you to consider these 8 critical questions when selecting your Ai technology.

You will find not only the key questions to ask when testing the tools but why these are critical questions to ask and how to differentiate between the answers you are given.

Less room for bias means increased fairness for applicants, and more diverse hiring for you.



Say NO to bias



Sarah is a recruiter who cares about candidate experience, giving every candidate a fair opportunity and more equitable hiring outcomes.

What Sarah wants ...



I want to interview all my candidates by asking them all the same questions.



I want candidates to feel comfortable and safe so no video and no time limits.



I only want to use what candidates write about themselves in answer to their interview questions. No CVs or data scrapped from the web.



I want candidates to be scored on the things that matter using a rubric that is consistent and fair.



I want to make sure that no biases creep into the process along the way. Our aim: Bias-free scoring and ranking.



I want transparency on how each candidate scores on different metrics like traits and communication so I can explain the score to anyone that asks.



I want useful benchmark data to know that I am hiring the best for the pool and the best in market



Feedback is a gift. I want to give all my candidates individualised insights and coaching tips.

1. What training data do you use?

Another way to ask this is: what data do you use to assess someone's fit for a role?

First up- why is this an important question to ask ...

Machine-learning algorithms use statistics to find and apply patterns in data. Data can be anything that can be measured or recorded, e.g. numbers, words, images, clicks etc. If it can be digitally stored, it can be fed into a machine-learning algorithm.

The process is quite basic: find the pattern, apply the pattern.

This is why the data you use to build a predictive model, called training data, is so critical to understand.

In HR, the kinds of data that could be used to build predictive models for hiring and promotion are:

- CV data and cover letters
- Games built to measure someone's memory capacity and processing speed
- Behavioural data, e.g. how you engage in an assessment,
- Video Ai can capture how you act in an interview—your gestures, pose, lean, as well as your tone and cadence.
- Your text or voice responses to structured interview questions
- Public data sources such as your social media profile , your tweets, and other social media activity



If you consider the range of data that can be used in training data, not all data sources are equal, and on its surface, you can certainly see how some carry risk of amplifying existing bases and the risk of alienating your candidates.

Consider the training data through these lenses:

Is the data visible or opaque to the candidate?

Using data that is invisible to the candidate may impact your employer brand. And relying on behavioural data such as how quickly a candidate completes the assessment, social data or any data that is invisible to the candidate might expose you to not only brand risk but also a legal risk. Will your candidates trust an assessment that uses data that is invisible to them, scraped about them or which can't be readily explained?

Increasingly companies are measuring the business cost from poor hiring processes that contribute to customer churn. 65% of candidates with a positive experience would be a customer again *even if they were not hired* and 81% will share their positive experience with family, friends and peers¹.

Visibility of the data used to generate recommendations is also linked to explainability which is a common attribute now demanded by both governments and organisations in the responsible use of Ai.

Video Ai tools have been legally challenged on the basis that they fail to comply with baseline standards for AI decision-making, such as the OECD AI Principles and the Universal Guidelines for AI ([click here for source](#)).

Reference: Talent Board: Benchmark Quoted by Niels Rozemeijer, Customer Success Manager, Starred.com



Or that they perpetuate societal biases and could end up penalising non-native speakers, visibly nervous interviewees or anyone else who doesn't fit the model for look and speech.

If you are keen to attract and retain applicants through your recruitment pipeline, you may also care about how explainable and trustworthy your assessment is. When the candidate can see the data that is used about them and knows that only the data they consent to give is being used, they may be more likely to apply and complete the process. Think about how your own trust in a recruitment process could be affected by different assessment types.

Is the data 1st party data or 3rd party data ?

1st party data is data such as the interview responses written by a candidate to answer an interview question. It is given openly, consensually and knowingly. There is full awareness about what this data is going to be used for and it's typically data that is gathered for that reason only.

3rd party data is data that is drawn from or acquired through public sources about a candidate such as their twitter profile. It could be your social media profile. It is data that is not created for the specific use case of interviewing for a job, but which is scraped and extracted and applied for a different purpose,

It is self-evident that an Ai tool that combines visible data and 1st party data is likely to be both more accurate in application for recruitment and have outcomes more likely to be trusted by the candidate and the recruiter.



Trust matters to your candidates and to your culture ...

At PredictiveHire, we are committed to building ethical and engaging assessments. This is why we have taken the path of a text chat with no time pressure. We allow candidates to take their own time, reflect and submit answers in text format.

We strictly do not use any information other than the candidate responses to the interview questions (i.e. fairness through unawareness – algorithm knows nothing about sensitive attributes).

For example, no explicit use of race, age, name, location etc, candidate behavioural data such as how long they take to complete, how fast they type, how many corrections they make, information scraped from the internet etc. While these signals may carry information, we do not use any such data.

Want to try the text-chat interview? Answer the questions and get your insights.

[Try It Here >](#)

No amount of surveys will change your culture just as no amount of diversity training will give you people who are less biased.



2. Can you explain why ‘person y’ was recommended by the Ai and not ‘person z’?

Another way to ask this is – Can you explain how your algorithm works? and does your solution use deep learning models?

This is an interesting question especially given that we humans typically obfuscate our reasons for rejecting a candidate behind the catch-all explanation of “Susie was not a cultural fit”.

For some reason, we humans have a higher-order need and expectation to unpack how an algorithm arrived at a recommendation. Perhaps because there is not much to say to a phone call that tells you were rejected for cultural fit.

This is probably the most important aspect to consider, especially if you are the change leader in this area. It is fair to expect that if an algorithm affects someone’s life, you need to see how that algorithm works.

Transparency and explainability are fundamental ingredients of trust, and there is plenty of research to show that high trust relationships create the most productive relationships and cultures.

This is also one substantial benefit of using AI at the top of the funnel to screen candidates. Subject to what kind of Ai you use, it enables you to explain why a candidate was screened in or out.

This means recruitment decisions become consistent and fairer with AI screening tools.



But if AI solutions are not clear why some inputs (called “features” in machine learning jargon) are used and how they contribute to the outcome, explainability becomes impossible.

For example, when deep learning models are used, you are sacrificing explainability for accuracy. Because no one can explain how a particular data feature contributed to the recommendation. This can further erode candidate trust and impact your brand.

The most important thing is that you know what data is being used and then ultimately, it's your choice as to whether you feel comfortable to explain the algorithm's recommendations to both your people and the candidate.



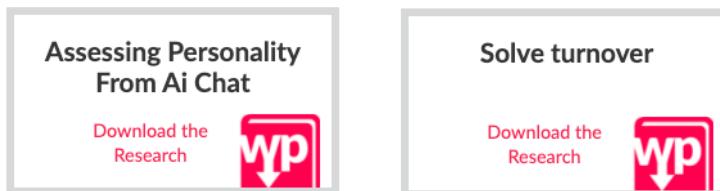
3. What assumptions and scientific methods are behind the product? Are they validated?

Assessment should be underpinned by validated scientific methods and like all science, the proof is in the research that underpins that methodology.

This raises another question for anyone looking to rely on AI tools for human decision making - where is the published and peer-reviewed research that ensures you can have confidence that a) it works and b) its fair.

This is an important question given the novelty of AI methods and the pace at which they advance.

At PredictiveHire, we have published our research to ensure that anyone can investigate for themselves the science that underpins our AI solution.



We continuously analyse the data used to train models for latent patterns that reveal insights for our customers as well as inform us in improving the outcomes.

HERE'S A FORMULA FOR THE 'RIGHT' CULTURE:

(Inclusiveness + Transparency) – Bias = Trust



4. What are the bias tests that you use and how often do you test for bias?

It's probably self-evident why this is an important question to ask. You can't have much confidence in the algorithm being fair for your candidates if no one is testing that regularly.

Many assessments report on studies they have conducted on testing for bias. While this is useful, it does not guarantee that the assessment may not demonstrate biases in new candidate cohorts it's applied on.

The notion of "data drift" discussed in machine learning highlights how changing patterns in data can cause models to behave differently than expected, especially when the new data is significantly different from the training data.

Therefore on-going monitoring of models is critical in identifying and mitigating risks of bias.

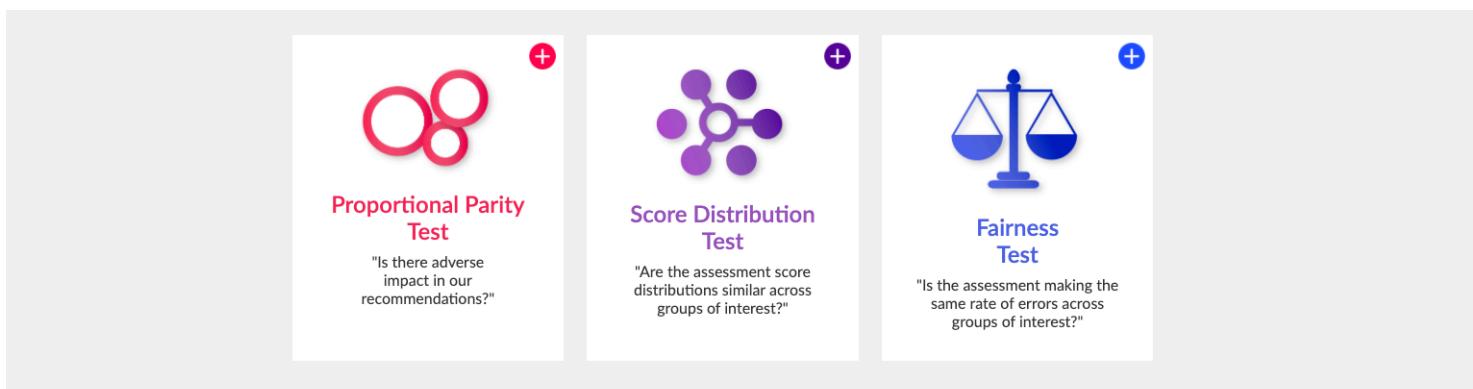
Potential biases in data can be tested for and measured.

These include all assumed biases such as between gender and race groups that can be added to a suite of tests. These tests can be extended to include other groups of interest where those group attributes are available like English As Second Language (EASL) users.



On bias testing, look out for at least these 3 tests and ask to see the tech manual and an example bias testing report.

- **Proportional Parity Test.** This is the standard EEOC measure for adverse impact on selection and recommendations.
- **Score Distribution Test.** This measures whether the assessment score distributions are similar across groups of interest
- **Fairness Test.** This measures whether the assessment is making the same rate of errors across groups of interest



At PredictiveHire, we conduct all the above tests. We conduct statistical tests to check for significant differences between groups of feature values, model outcomes and recommendations. Tests such as t-tests, effect sizes, ANOVA, 4/5th, Chi Squared etc. are used for this. We consider this standard practice.

We go beyond the above standard proportional and distribution tests on fairness and adhere to stricter fairness considerations, especially at the model training stage on the error rates. These include following guidelines set by IBM's AI Fairness 360 Open Source Toolkit.²

We continuously analyse the data used to train models for latent patterns that reveal insights for our customers as well as inform us in improving the outcomes.

Reference: <https://aif360.mybluemix.net/>) and the [Aequitas project at the Centre for Data Science and Public Policy at the University of Chicago](#)

5. How can you remove bias from an algorithm?

We all know that despite best intentions, we cannot be trained out of our biases. Especially the unconscious biases.

This is another reason why using data-driven methods to screen candidates is fairer than using humans.

Biases can occur in many different forms. Algorithms and AI learn according to the profile of the data we feed it. If the data it learns from is taken from a CV, it's only going to amplify our existing biases. Only clean data, like the answers to specific job-related questions, can give us a true bias-free outcome.

If any biases are discovered, the vendor should be able to investigate and highlight the cause of the bias (e.g. a feature or definition of fitness) and take corrective measure to mitigate it.



6. On which minority groups have you tested your products?

If you care about inclusivity, then you want every candidate to have an equal and fair opportunity at participating in the recruitment process.

This means taking account of minority groups such as those with autism, dyslexia and English as a second language (EASL), as well as the obvious need to ensure the approach is inclusive for different ethnic groups, ages and genders.

At PredictiveHire, we test the algorithms for bias on gender and race. Tests can be conducted for mostly any group in which the customer is interested. For example, we run tests on “English As a Second Language” (EASL) vs. native speakers.

7. What kind of success have you had in terms of creating hiring equity?

If one motivation for you introducing Ai tools to your recruitment process is to deliver more diverse hiring outcomes, it's natural you should expect the provider to have demonstrated this kind of impact in its customers.

If you don't measure it, you probably won't improve it. At PredictiveHire, we provide you with tools to measure equality. Multiple dimensions are measured through the pipeline from those who applied, were recommended and then who was ultimately hired.



8. What is the composition of the team building this technology?

Thankfully, HR decision-makers are much more aware of how human bias can creep into technology design. Think of how the dominance of one trait in the human designers and builders have created an inadvertent unfair outcome.

In 2012, YouTube noticed something odd.



About 10% of the videos being uploaded were upside down.

When designers investigated the problem, they found something unexpected: Left-handed people picked up their phones differently, rotating them 180 degrees, which lead to upside-down videos being uploaded,

The issue here was a lack of diversity in the design process. The engineers and designers who created the YouTube app were all right-handed, and none had considered that some people might pick up their phones differently.

In our team at PredictiveHire, from the top down, we look for diversity in its broadest definition.

Gender, race, age, education, immigrant vs native-born, personality traits, work experience. It all adds up to ensure that we minimise our collective blind spots and create a candidate and user experience that works for the greatest number of people and minimises bias.



It turns out that CV data is neither predictive nor fair for use in making hiring decisions.



Amazon did an experiment. They analysed 10 years of CV data to build a predictive model to help filter through hundreds of thousands of applications to work at the company.

Because the sample group was mostly male, the CVs were naturally biased towards male 'traits' (if there is such a thing). The model built off this training data naturally ended up mirroring that sample group which meant it preferred male to female CVs.

CV's can hold latent patterns of group identity.



PredictiveHire did its own experiment to understand how predictive a CV is of both whether someone is hired and whether they are a good performer in the role.

We analysed ~13,000 CVs received over a five-year period, all for similar roles for a large sales-led organisation. From this data set, 2,660 were hired and around 9,600 rejected.

We wanted to test how meaningful the CV is as a data source for hiring decisions. What did we find? Negligible correlation between the CV of a person who got hired and a person who was the best performer.

Humans are heavily prone to unconscious bias.

In fact, we rely upon it to survive. While these biases help us not go insane, unfortunately, it has led us to the point today where they are having a very significant effect in the workforce.

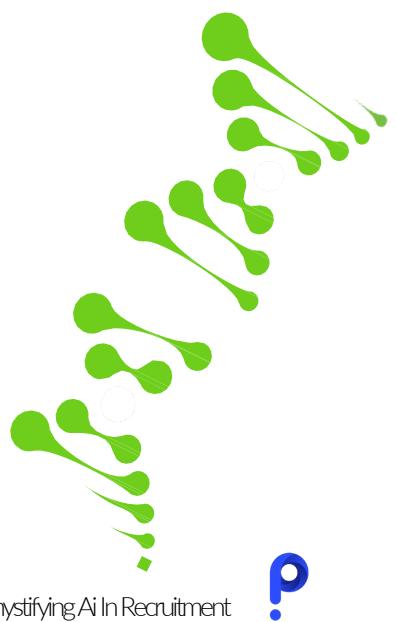
There are many serious forms of bias.

One study showed simply by changing the name of an applicant from a woman to a man's, with every other detail kept the same, the 'male' applicant was more likely to progress to an interview. That's with the exact same CV.

When humans do screening, they are prone to making snap judgements based on superficialities, ignoring the very many factors that can help predict whether a candidate will perform.

This is where data platforms have an advantage, by doing 'blind screening' and making the process both faster and fairer.

PredictiveHire's FirstInterview is a true blind assessment. No demographic details are collected from candidates nor used to influence their ranking. Only the candidates answer to relevant interview questions are analysed by our scientifically validated algorithm to assess their fit for the role.



Did you know: Social media data should not be used for any kind of HR decisions if you care about fairness and minimising bias.

While many studies have shown how the language used in social media platforms is predictive of personality traits, none of the studies have explored the topic of model generalisability. It's important to explore whether models trained using data from one source generalises well to other sources independently proven to be predictive of personality traits.

PredictiveHire did our own research on this.



Looking at the predictive validity of a Twitter-based text to personality traits classifier when applied on data from a different source. We analysed 11M tweets containing over 200M words from 5,920 users who self-reported personality. What we found was not that surprising. The profile of the Twitter user is not representative of the general population.

Using predictive models built with public training data sets that, like Twitter, are not generalisable to the whole population risks creating more bias.

Just think about it—how you speak with friends and family is going to be different to how you respond to questions in an interview—which is most certainly likely to be different to what you tweet.

Language models are highly sensitive to the source of the training data, which is a form of sample bias. This raises serious questions on the generalisability and as a result applicability of social-media-based text to personality models when used with data from other sources.

If you are interested in reading this research paper by our data science team, please contact our head data scientist [Buddhi Jayatilleke, here.](#)

What other questions have you used to validate the fairness and integrity of the Ai tools you have selected to augment your hiring and promotion processes?

We'd love to know!
Share them here or let's chat!

Thank you for taking the time to share what you know a little more about me. I really appreciate this tool to find out a little more about a candidate applying. Better than a standard interview. This has been a really thoughtful and useful interview that has helped me reflect upon my value and goals. This is a fab way to save employee and employer time! Easy to use, I enjoyed the opportunity to express myself and learn a little more about my self worth and personal values. I found this conversation very interesting and thought provoking brilliant! Wow this was so easy! This is a very innovative way of conducting initial interviews. I have not experienced anything like this before. Unlike anything i've completed in the past! The questions were really felt like you want to make sure you pick the right person for your company. I enjoyed the questionnaire interview as it was comfortable to answer each question without a time restraint. It was better than a video interview. I think this is a great way to get to know one's personal attributes, work experience and career history ... This was an interesting process for a job application. Questions were broad and fast, i liked it. The interview process is truly wonderful. I can really put my best foot forward. This process was interesting and exciting way to conduct an interview. There were a few questions that made me really think about what kind of person I am and reflect on my views. It was interesting and thoughtful. The first thing the company is concerned with knowing what i think i really liked this format. It almost like i was chatting with friends! Thought that this was a very innovative and clear way for an interview online. Embrace change! I love technology! This is the most interactive and cool interview i've ever had! I really like how you want to go deep into our feelings to find out what we are thinking! I really enjoyed answering the questions. It makes life easy really! This is the most fun way to present a pre-interview process and i thoroughly enjoyed it. It was easy to answer and see the questions. Great system to get to know someone. It really made me think about who i am as a person. My favorite part of doing an interview like this are the questions. I came across something like this in the interview process before, was not as good. I have some awesome questions in here. I enjoyed the challenge and taking part in this interview. It was an interesting and thought provoking way of conducting initial interviews. Brilliant! Wow this was so easy! I have not experienced anything like this before. Unlike anything i've completed in the past! The interview process is truly wonderful. I enjoyed the questionnaire interview as it was comfortable to answer each question without a time restraint. It was better than a video interview.

100% bang on like me :) thank you for sharing. I appreciate your email !!! 100% agree with your predictions do sound like me. I'm impressed. Excellent and a very innovative way to do the initial interview. I definitely agree with this feedback. It feels as this explains my passion for the things I love. I really appreciate you for this. Thank you so much for the detailed feedback. This is probably the first time I have received such a great detailed feedback. Thank you once again for this. I think this is a very accurate representation of me. I opened my eyes to some things I can improve and do differently. The feedback is pretty accurate. Very impressed with this. I'm impressed that this level of accuracy can be drawn from such an easy-to-complete survey. I'm surprised with the accuracy of the results! Impressed, very accurate. Described me 100% correctly. I feel described like me. I'm impressed. I creeped out as well hahaha. It's refreshing to receive some sort of feedback. The email was positive. The advice was practical, relevant and useful. Thank you. Pretty accurate and detailed description of my professional skills and personality. Thank you! Really surprised with the results, i can relate to all of it. Results are pretty close to my personality, very accurate. Sounds like me! Accurate. Sounds like me in a lot of ways. Not everything i would disagree with, and some i agree with. Thank you for sending me this information through. I do feel that the tip and insights were very accurate. Thank you for the feedback, it was interesting to read and sounded just like me. Thank you for this feedback! I would say a very accurate! Thank you for your valuable feedback.

PredictiveHire Pty Ltd

W: predictivehire.com E: info@predictivehire.com
in <https://au.linkedin.com/company/predictivehire>
tw https://twitter.com/predictive_hire

