Lessons in AI deployment in a regulated landscape.

Alex Vaughan, WAAC 2021
How Do We Find The Right Person For The Right Job?

The traditional approach:

Cloud Native Infrastructure Engineer / Architect at

Post Date: June 27, 2020  Full Time

Required Technical and Professional Expertise

Minimum 12 years' experience in Kubernetes administration and management

Hands-on experience on setting up Kubernetes platform, deploying microservices and other web applications, and managing secure secrets along with container orchestration using Kubernetes

*Problem: Kubernetes is 6 years old...*
How Do We Find The Right Person For The Right Job?

Data Quality & Scale
Proprietary and unique datasets.

Machine Learning
Algorithmic advances & computational power

Demonstrated Business Value
Proven track record of high lifetime ROI.
How Do We Find The Right Person For The Right Job?

**pymetrics: Behavior, Reasoning, and Communication**

<table>
<thead>
<tr>
<th><strong>Behavioral Aptitudes</strong></th>
<th>12 games that measure cognitive and emotional attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort, Emotion, Risk, Fairness, Focus, Attention</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Numerical Reasoning</strong></th>
<th>4 games that specifically measure numerical and logical reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative Reasoning, Numerical Agility</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Communication</strong></th>
<th>Communications assessment that replaces the early phone screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content, Context</td>
<td>Question: What people say, Exclamation: How they say it</td>
</tr>
</tbody>
</table>
pymetrics centers on a set of core assessments rooted in cognitive science
Our algorithmic platform optimizes for performance and fairness.

Preprocessing
- Outlier, Inattention & Cheating Removal
- Scaling & Imputation

Training and bias testing
- High-performing Employees
- Baseline Population
- Debiasing Dataset

Debiased & predictive model
Business value arises from three key features.

- **PREDICTIVE**
  - Our assessments predict key factors driving performance: attention, focus, communication and numerical skills.
  - Our assessments are continuously tested for predictive value, before and after deployment.

- **FAIR**
  - Fairness first approach for all assessments.
  - For new video assessments we have a special obligation to avoid fairness arising from facial recognition / audio fingerprinting.

- **EXPLAINABLE**
  - Reports for recruiters explaining each model as well as explaining each decision.
  - Reports for candidates explaining what makes them unique and how to improve
Demonstrated ROI for pymetrics clients.

Proven and quantifiably better outcomes for clients, candidates, and the world

Selection of real results from pymetrics customer engagements

<table>
<thead>
<tr>
<th>efficiency</th>
<th>diversity</th>
<th>effectiveness</th>
<th>experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% Reduction in recruiter screens</td>
<td>100% Increase in gender, ethnic &amp; SES diversity</td>
<td>30% Reduction in attrition</td>
<td>95%+ Candidate satisfaction</td>
</tr>
<tr>
<td>100% Increase in hire yield</td>
<td>4x Increase in female engineering hires</td>
<td>20% Increase in performance</td>
<td>31% Increase in completed apps</td>
</tr>
<tr>
<td>50% Decrease in time to hire</td>
<td>55% Increase in university diversity</td>
<td>4.4x More likely to receive offers</td>
<td>92% Of recruiters felt added value</td>
</tr>
<tr>
<td>40% Recruiter cost savings</td>
<td>200%+ Increase in offers to minorities</td>
<td>3.8x Improvement in performance metric</td>
<td>33% Increase in sales</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65% Of candidates found alternative careers</td>
</tr>
</tbody>
</table>
Three short lessons from the pymetrics journey.
1. Fairness is a performance criterion.

EEOC regulations require that all assessments used in hiring be fair within an 80% margin.
1. **Fairness is a performance criterion.**

**pymetrics Audited AI passes the EEOC 80% rule.**

**EEOC standard Impact Ratio:** 0.8

- **Gender**
  - Average > .95

- **Race/Ethnicity**
  - Average > .92

*Results in preparation for peer-reviewed submission.*
Research team from Northeastern University published results from an independent third-party audit

**Correctness:** Does pymetrics’ source code implement the four-fifths rule via the minimum bias ratio metric, using the process described in their documentation?

**✓**

**No Direct Discrimination:** Do models trained using the pymetrics source code directly incorporate demographic features, and/or do the models take demographic features as direct input?

**✓**

**Debiasing Circumvention:** Is there a way to manipulate the input data to the pymetrics source code in such a way that the fairness checks are circumvented?

**✓**

2. **Transparency is a collaborative process.**

Deployment of AI in regulated spaces can accelerate the ethics life-cycle.

- Regulatory compliance is the first step, but not the last.
- There is a clear need for public-facing audits alongside traditional regulatory compliance.
- **Key learning:** trust is your most valuable asset, and transparency is the most powerful tool to build trust.
3. The life cycle of AI deployment at a societal scale is complex.

- **Data Quality & Scale**: Proprietary and unique datasets.
- **Machine Learning**: Algorithmic advances & computational power.
- **Demonstrated Business Value**: Proven track record of high lifetime ROI.
3. The life cycle of AI deployment at a societal scale is complex.

Only one piece of the AI life-cycle has been (or can be) meaningfully commoditized.
3. The life cycle of AI deployment at a societal scale is complex.

- **Data Quality & Scale**: Proprietary and unique datasets.
- **Machine Learning**: Algorithmic advances & computational power.
- **Demonstrated Business Value**: Proven track record of high lifetime ROI.

The most important phase is just beginning.
Thanks!

Alex Vaughan, WAAC 2021