

LICENSE STATEMENT

The AI AGENT - CASE STUDY – SHORTLISTS

is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/>

You are free to use the contents but

you must attribute the original authors;

Hugh Abbott <https://www.linkedin.com/in/hugh-abbott/>

Ebrahim Jahanshiri <https://www.linkedin.com/in/geoej/>

please contact through LinkedIn

AI AGENT - CASE STUDY - SHORTLISTS

EXECUTIVE SUMMARY

A large NHS Trust piloted an AI agent to generate applicant shortlists.

Addressing the time-consuming process of manually screening job applicants.

The pilot generated 28 shortlists and reviewed 1,333 applicants.

- the agent produced accurate & reliable shortlists
- the agent produced shortlists in hours rather than days
- the agent will save more than **1,750 hours per year**



AI AGENT - CASE STUDY - SHORTLISTS

THE CHALLENGE

- **High number of roles:** the trust expects to advertise 500 + roles every year.
- **High number of applicants:** on average, each role attracts 47 applicants. Many roles attract well over 100 applicants.
- **Pressures on senior medics:** The task of shortlisting takes senior medics away from working with patients.
- **Process delays:** delays often occur because senior people are not always available to create the shortlists.
- **Consistency and fairness:** Human shortlisters are often overwhelmed by the high number of applicants. Human shortlisters often scan read which leads to unfairness and inconsistency.



AI AGENT - CASE STUDY - SHORTLISTS

THE SOLUTION

Develop an AI Agent using the latest LLMs

1. **Input:** a single PDF including the job description and all the applicants.
2. **Pre-processing:** use an LLM to extract all the criteria used for judging.
3. **Processing:** use LLMs to score each candidate against the judging criteria.

Output: the agent outputs a [ranked list of all candidates](#)* with a clear [audit-trail of every decision](#)**.



* [See Appendix 1](#)

** [See Appendix 2](#)

AI AGENT - CASE STUDY - SHORTLISTS

RESULTS

1. **Accurate & Reliable Shortlists:** the hospital developed a framework to evaluate the agent-generated shortlists.
 - the pilot evaluated 28 shortlists
 - there were 1,333 applicants
 - the shortlists were judged to be **98% accurate** (using the hospital framework)
2. **Saving One FTE:** a conservative estimate is that the agent would save one FTE of work.
 - on average, there were 47 applicants per role
 - we estimated it takes 3.5 hours for a human to create one shortlist
 - with 500 roles per year that is 1,750 hours \geq one FTE
3. **Fair & Thorough with an [Audit-trail of Decisions](#)**:** the agent records an [audit-trail of decisions](#)**.
 - the agent was thorough it did not get tired and skip any applicants
 - the agent was fair; it scored every single candidate against all the criteria
 - the agent creates an [audit-trail](#)* of decisions, it records an explanation of every score it gives



** [See Appendix 2](#)

AI AGENT - CASE STUDY - SHORTLISTS

INNOVATIONS

1. **Referee Architecture:** the team developed an architecture whereby all candidates were scored by two different LLMs. And a third LLM was used as a referee if there was a disagreement between the initial two LLMs.

- the applicant was assessed by Gemini (Google)
- the applicant was then assessed by Llama (Meta)
- if there were any discrepancies in the scores then Claude (Anthropic) was used as a referee



2. **Enrichment Step:** the criteria for judging applicants were ‘enriched’ by adding extra information.

- enrichment added information about abbreviations associated with academic qualifications
- enrichment added information about jargon associated with technical roles
- enrichment added examples of how criteria might be evidenced

3. **Hyper Granular 1,000+:** over the course of the project, the team developed the hyper-granular approach.

- the team found that splitting the task into many smaller tasks (prompts) delivered best results
- the agent runs slower but the results are much more accurate
- the team now uses **1,000+ prompts** to create one shortlist

The referee architecture, the enrichment step, the hyper-granular approach all helped to produce shortlists that were more rigorous and more accurate than any prior automated solution.

AI AGENT - CASE STUDY - SHORTLISTS

RISKS & MITIGATIONS

1. **Data Security:** the pilot was conducted using servers adhering to NHS data standards.
 - NHS data is never used to train LLMs
 - NHS data is not retained by the LLMs
 - data was handled in compliance with GDPR & SOC2 principles
2. **Bias:** the trust wanted to assess the agent for bias & fairness.
 - no bias was found in the pilot
 - applicants were assessed against published criteria
 - the criteria themselves were judged not to be a source of bias
3. **Transparency:** there is a clear audit-trail for every decision made by the agent.
 - every applicant is scored against all the criteria in the job description
 - every score given by the agent is accompanied by an [explanation](#)**
 - there is an [audit-trail](#)** clearly showing the justification for every ranking decision



** [See Appendix 2](#)

AI AGENT - CASE STUDY - SHORTLISTS

CALL TO ACTION

This pilot shows AI shortlisting can deliver **faster, fairer, and cheaper** recruitment in the NHS.


The approach is scalable to other Trusts and adaptable across roles. Senior leaders seeking to modernise healthcare recruitment can commission similar projects to unlock time savings and reduce costs across their organisations.

- **Next Step:** Explore how AI shortlisting could be applied to your recruitment to achieve measurable efficiency gains.



APPENDIX 1: RANKED LIST OF CANDIDATES

Screenshot shows a ranked list of candidates, taken from the summary page of the agent’s report. Each candidate is given a total score and an indication whether all essential criteria have been met. Surprisingly, JD Vance is the top scoring applicant.

| | | | | |
|---|--|--|-------|----------------------|
| <div>Shorlisting Service</div> | | | | |
| Job Title | | Assistant Practitioner - Nutrition and Speech & Language Therapy | | |
| Date | | 2024-12-12 | | |
| Total Applicants | | 5 | | |
| APPLICANT | | # | SCORE | ESSENTIAL |
| JD Vance | | 1 | 15 | all met |
| Meghan Markle | | 2 | 14 | all met |
| Harry Kane | | 3 | 12 | all met |
| Paul McCartney | | 4 | 9 | all met |
| Taylor Swift | | 5 | 3 | missing at least one |

APPENDIX 2: AUDIT-TRAIL OF DECISIONS

Below is a table showing details of criterion no. 5 which is asking for experience in Speech & Language therapy.


The candidate is scored 3 out of a maximum of 3.

The explanation, the audit-trail, is shown in yellow.

| | |
|-------------|--|
| # | 5 |
| Title | Experience of Speech and Language Therapy or Dietetics |
| Score | 3 |
| Max Score | 3 |
| Essential | Met |
| Notes | Previous experience as a Dietetic Assistant described. |
| Explanation | The applicant explicitly mentions experience as a "Dietetic Assistant" at Farooq Hospital. This directly addresses the criterion of therapy assistant experience within Dietetics. |

APPENDIX 3: CANDIDATE DETAILS SCREEN

Screenshot shows the details screen for applicant JD Vance.
Each criterion is given a score with short notes and a longer explanation.

| <div>Shorlisting Service</div> | | | | | | |
|---|--|-------|-----------|-----------|----------------------------------|---------------------------------------|
| Applicant | JD Vance | | | | | |
| Position | 1 | | | | | |
| Job Title | Assistant Practitioner - Nutrition and Speech & Language Therapy | | | | | |
| Date | 2024-12-12 | | | | | |
| Score | 15 | | | | | |
| Score + Extra | 21 | | | | | |
| ID | CRITERION | SCORE | MAX SCORE | ESSENTIAL | NOTES | EXPLANATION |
| 1 | Good communication skills - both written and verbal | 3 | 3 | Met | Applicant demonstrates | The applicant's application shows |
| 2 | Knowledge of medical conditions found in a frail | 3 | 3 | Met | Applicant discusses experience | The application highlights experience |
| 3 | Awareness of the roles of other professionals with | 3 | 3 | Met | Applicant details collaboration | The applicant explicitly describes |
| 4 | NVQ Level 3 (National Vocational Qualification) / | 3 | 3 | Met | Possesses BSc Hons in Nutrition | The applicant has a BSc Hons |
| 5 | Experience of Speech and Language Therapy | 3 | 3 | Met | Previous experience as a | The applicant's application explains |
| 6 | Experience working in hospital or healthcare setting | 2 | 3 | N/A | Previous experience as Dietitian | The applicant has experience |
| 7 | Desire to develop inter-professional skills | 2 | 3 | N/A | Applicant expresses desire | The applicant mentions collaboration |
| 8 | Ability to work flexibly across multiple disciplines | 2 | 3 | N/A | Applicant has experience | The applicant's experience as |