hr | avatar
Illuminating Talent

# Test Results and Interview Guide

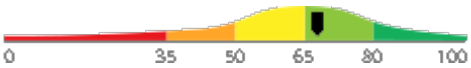
Candidate: **Elizabeth Wantsajob**
Assessment: Agentic AI Architecture and Concepts
Completed: July 2, 2026
Prepared for: Sara Maple
Example Company

## What's Included

- Overall Score
- Competency Summary Table
- Comparison Matrix
- Detailed Competency Results with Interview Guide

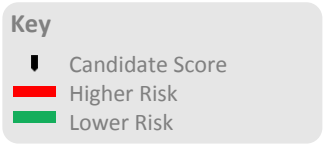
**Important Note:** The Agentic AI Architecture and Concepts assessment measures key factors related to high performance and tenure in this job. Attribute types measured vary by test, but can include cognitive ability, skills, knowledge, personality characteristics, emotional intelligence, and past behavioral history. This report includes a one page summary, followed by detailed results with an embedded interview guide. Note that these results should always be used as a part of a balanced candidate selection process that includes independent evaluation steps, such as interviews and reference checks.

## Overall

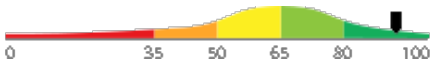
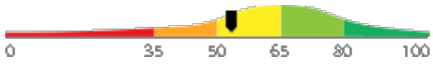
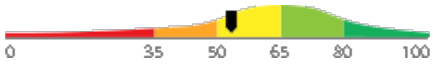
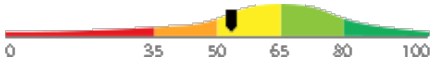
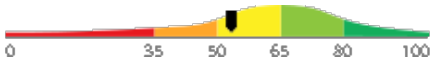
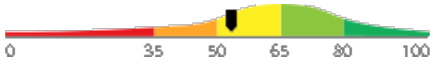
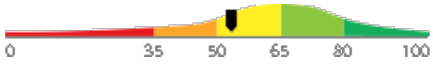
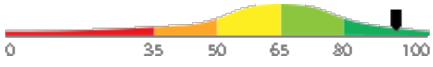
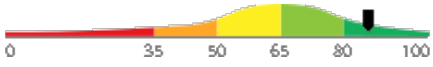
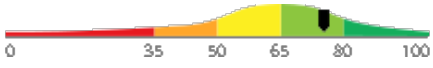
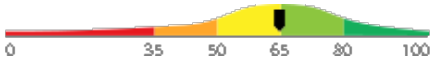
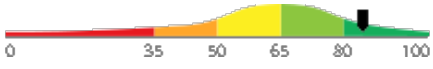
| Candidate | Score | Interpretation |
|---|---|---|
| Elizabeth Wantsajob<br>beth.wantsajob@gmail.com<br>Agentic AI Architecture and Concepts<br>July 2, 2026 | 68 |  |



The candidate exhibits a solid and broadly competent understanding of agentic AI design and operational concepts, including agent reasoning, tool use, prompt engineering, planning, and error handling. While some advanced or specialized areas may warrant further development, this individual demonstrates the knowledge base expected of a capable entry- to mid-level professional working with agentic AI systems.




## Competency Summary

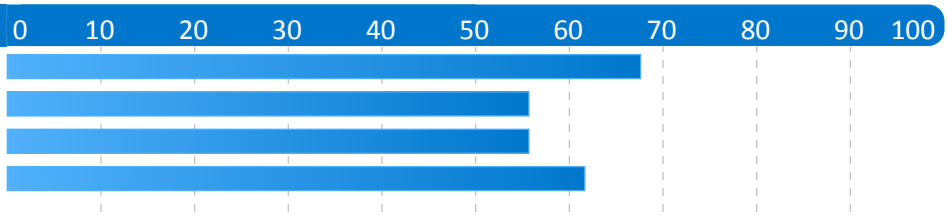
| Competency | Score | Interpretation |
|---|---|---|
| **Skills/Knowledge (relates to immediate readiness)** | | |
| Agent Reasoning and Task Execution | 92 |  |
| Agent Reasoning and Task Execution (Free Text Responses) | 53 |  |
| Error Handling, Guardrails, and Safe Operation (Free Text Responses) | 53 |  |
| Memory and Context Management (Free Text Responses) | 53 |  |
| Multi-Agent Orchestration and Workflow Patterns (Free Text Responses) | 53 |  |
| Prompt Design for Agent Behavior (Free Text Responses) | 53 |  |
| Tool Use and Function Calling (Free Text Responses) | 53 |  |
| Error Handling, Guardrails, and Safe Operation | 92 |  |
| Memory and Context Management | 86 |  |
| Multi-Agent Orchestration and Workflow Patterns | 75 |  |
| Prompt Design for Agent Behavior | 65 |  |
| Tool Use and Function Calling | 84 |  |



## Comparison

Percentile scores indicate how the candidate compares to other test-takers within various groups. The candidate scored equal to or better than the fraction of test-takers indicated by the percentile.

| Test-Taker Group | Percentile |
|---|---|
| Global | 68th |
| North America | 56th |
| United States | 56th |
| Example Company | 62nd |



## Artificial Intelligence (AI) Generated Scores

This table includes one or more scores derived from a large language model AI query. AI-derived scores are non-deterministic. That is, they are not precisely repeatable. Therefore, these scores should always be treated as supplementary information and should never be used exclusively or compared to hard cutoff values.

| Estimated Value | Score | Confidence | Interpretation |
|---|---|---|---|
| Knowledge, Skills, and Abilities Summary | - | - | Summary Points (AI):<br>• (Generic Text for Sample Report) Strong performer in Drag and Drop Files tasks, indicating comfort with file management and basic computer interactions.<br>• Demonstrates solid numerical accuracy in Recognizing and Confirming Numbers, a valuable asset in detail-oriented roles.<br>• Moderate overall performance in Analytical Thinking and Attention to Detail, with adequate grammar skills but room for improvement.<br>• Struggles with Reading and Analyzing Problems, which may limit effectiveness in roles requiring critical reading and complex problem-solving.<br>• Lowest performance in Navigating Between Screens, suggesting difficulty with multi-screen software workflows that could impact productivity in computer-intensive roles.<br>Narrative (AI): Elizabeth Wantsajob demonstrates a mixed profile of knowledge, skills, and abilities across the assessed competencies.<br><br>Elizabeth shows a strong aptitude in Drag and Drop Files, performing well on this technical task and suggesting she is comfortable with this type of computer interaction. This is a notable strength that would translate well into roles requiring file management and basic computer navigation tasks.<br><br>In the area of Analytical Thinking and Attention to Detail, Elizabeth performs at a moderate level. She demonstrates solid ability in Recognizing and Confirming Numbers, which suggests she is careful and accurate when working with numerical data — a valuable skill in detail-oriented work environments. Her Grammar performance is adequate but leaves room for improvement, indicating she may occasionally make written communication errors. Her weakest area within this competency is Reading and Analyzing Problems, where she struggled to consistently interpret and work through written problem scenarios. This may impact her effectiveness in roles that require critical reading, written comprehension, or complex problem-solving.<br><br>Elizabeth's most significant area for development is Navigating Between Screens, where she scored considerably lower than the other competencies. This suggests she may have difficulty efficiently moving through software interfaces or multi-screen workflows, which could slow productivity in roles that rely heavily on navigating computer applications or data entry systems.<br><br>Overall, Elizabeth brings some useful technical strengths, particularly in file management and numerical accuracy, but would benefit from targeted development in software navigation and analytical problem-solving to be fully effective in roles that demand these skills.<br>Computed on: April 2, 2026, 11:09:49PM EDT |

## Detail

| | |
|---|---|
| Candidate: | Elizabeth Wantsajob, beth.wantsajob@gmail.com |
| Assessment: | Agentic AI Architecture and Concepts |
| Authorized: | July 2, 2026, by Sara Maple, Example Company, qamailsaram.mike@hravatar.com |
| Started: | July 2, 2026, 5:22:33PM EDT |
| Completed: | July 2, 2026, 5:22:33PM EDT |
| Overall Score: | 68 |

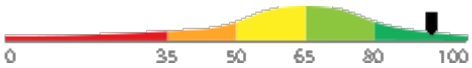
## Knowledge and Skills Detail

This section contains a list of job-related knowledge areas and skills that have been evaluated. Low scores in these areas often indicate that additional learning may be required before top performance can be achieved.

Detail | Interview Guide

### Agent Reasoning and Task Execution

Score: 92




*Description:*

Covers how an AI agent perceives inputs, decides what actions to take, and carries out steps to complete a goal. Includes understanding of reasoning loops (such as ReAct-style think-act cycles), how agents select tools or actions based on context, and how they determine when a task is complete.

*Interpretation:*

Candidate should achieve superior job performance in this area with little or no training.

The candidate demonstrates a comprehensive and sophisticated understanding of agentic AI design features and operational concepts across all measured areas, including prompt engineering, multi-agent orchestration, guardrails, pipeline troubleshooting, and agent monitoring. This level of knowledge reflects strong readiness to independently design, implement, evaluate, and maintain agentic AI systems in a professional environment.

Describe a situation where an agent's reasoning loop could get stuck or produce unhelpful repeated actions. How would you detect and address that problem?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Cannot describe a failure scenario or offers only a vague answer with no concrete fix. | | Identifies a plausible loop failure but proposes only a basic fix such as restarting the agent. | | Describes specific loop failure, explains detection via step logging, and proposes targeted fixes like step limits or replanning triggers. |

---

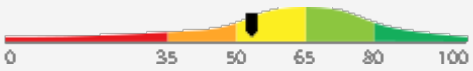
In your own words, can you walk me through what happens inside an AI agent between the moment it receives a user request and the moment it produces a final response or takes a final action?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Describes agent as simply generating text; no mention of reasoning steps or action selection. | | Mentions the agent thinks before acting but cannot clearly explain the loop or decision process. | | Clearly describes a perceive-reason-act cycle, referencing iterative steps and goal-checking. |

## Detail

### Agent Reasoning and Task Execution (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

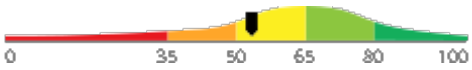
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Error Handling, Guardrails, and Safe Operation (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

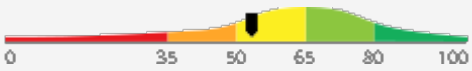
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Memory and Context Management (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

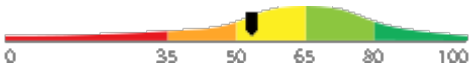
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Multi-Agent Orchestration and Workflow Patterns (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

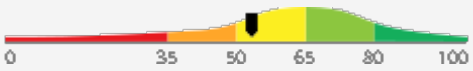
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Prompt Design for Agent Behavior (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

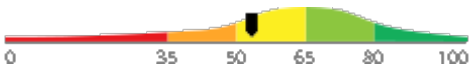
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Tool Use and Function Calling (Free Text Responses)

Score: 53




*Description:*

Covers the end-to-end process of planning, building, testing, and deploying AI-enabled applications for both internal staff and external customers. Includes managing iteration cycles, versioning, model monitoring, and coordinating cross-functional teams through each phase of the product lifecycle.

*Interpretation:*

The candidate exhibits average writing skills, which can hinder high performance in some jobs.

The candidate possesses a moderate understanding of AI product management, demonstrating basic familiarity with lifecycle management, strategic assessment, and process orchestration, though proficiency across these areas is inconsistent. With targeted coaching and hands-on experience, this individual has the potential to develop into a capable contributor in managing AI-enabled application initiatives.

| | |
|---|---|
| Overall AI Score: | 60.0 |
| High words per minute detected while composing one or more essays: | 27.3 words per minute. Possible copy/paste or use of AI tools. Average WPM while composing is about 15. |
| AI Confidence Level: | 80 |
| Argument Strength (AI): | 70.0 |
| Clarity and Coherence (AI): | 80.0 |
| Match with Ideal Response (AI): | 60.0 |
| Other Errors per 100 Words: | 0.0 |
| Spelling errors per 100 words: | 0.0 |

Please see below to view the essay submitted.

## Interview Guide

Describe a time you managed or contributed to an AI product through multiple lifecycle stages. What were the most significant challenges you encountered between phases, and how did you address them?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a generic or superficial example that lacks detail about AI-specific lifecycle challenges. Does not clearly articulate their personal role or the decisions they made between phases. | | Candidate shares a relevant example with reasonable detail, identifying at least one meaningful challenge such as stakeholder alignment or testing delays. However, the response may lack specificity about how AI-related factors (e.g., model performance, data readiness) influenced lifecycle decisions. | | Candidate provides a detailed, concrete example that demonstrates ownership across multiple lifecycle phases. Clearly describes AI-specific challenges such as model validation failures, shifting requirements, or deployment infrastructure issues, and articulates the specific actions they took to resolve them and keep the product on track. |

---

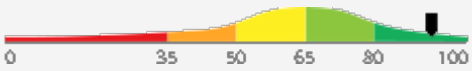
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Candidate provides a vague or incomplete description of the lifecycle, omitting key phases such as testing, validation, or deployment. May conflate AI product development with general software development without acknowledging AI-specific considerations like model training or data pipelines. | | Candidate identifies the major phases (discovery, development, testing, deployment) and acknowledges some AI-specific considerations, but struggles to articulate how the phases connect or how cross-functional teams are coordinated throughout. | | Candidate clearly outlines a structured lifecycle including discovery, requirements, development, model validation, testing, deployment, and monitoring. Demonstrates awareness of AI-specific challenges such as data quality, model drift, and iterative retraining, and explains how they would coordinate stakeholders across phases. |

## Detail

### Error Handling, Guardrails, and Safe Operation

Score: 92




*Description:*

Covers how agents detect and recover from failures, including tool errors, unexpected outputs, and reasoning mistakes. Also includes applying guardrails such as input/output filters, action restrictions, and fallback behaviors to keep agents operating safely and within intended boundaries.

*Interpretation:*

Candidate should achieve superior job performance in this area with little or no training.

The candidate exhibits an advanced and comprehensive mastery of error handling, guardrails, and safe operation within agentic AI architectures. They demonstrate a thorough ability to design and apply robust failure detection and recovery mechanisms, enforce input/output filters and action restrictions, and implement fallback behaviors to ensure agents consistently operate safely and within intended boundaries.

## Interview Guide

Describe how you would design an agent pipeline so that it handles a tool call failure gracefully rather than crashing or returning an unhelpful error to the user.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Suggests only catching exceptions without describing retry logic, fallback responses, or user communication. | | Mentions retry logic and a fallback message but does not address logging, escalation, or varying failure types. | | Describes retry with backoff, fallback action or response, user-facing messaging, error logging, and escalation for persistent failures. |

---

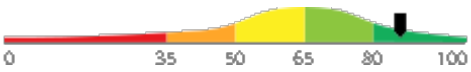
What would you do if an AI agent you deployed started producing responses that were off-topic, harmful, or clearly wrong? Walk me through how you would respond.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Only suggests turning off the agent with no mention of diagnosis, guardrails, or structured response. | | Mentions reviewing logs and updating the prompt but does not describe guardrails or systematic error handling. | | Describes immediate mitigation, log review, root cause analysis, adding output filters or guardrails, and a testing plan before redeployment. |

## Detail

### Memory and Context Management

Score: 86




*Description:*

Addresses how agents store, retrieve, and use information across steps within a single session and across multiple interactions. Covers short-term context (in-context memory), external memory stores (vector databases, key-value stores), and strategies for managing context window limits.

*Interpretation:*

Candidate should achieve superior job performance in this area with little or no training.

The candidate exhibits a comprehensive and sophisticated understanding of memory and context management within agentic AI systems, demonstrating strong command of short-term context handling, external memory store architectures, and advanced strategies for managing context window limits. This level of performance reflects expert-level knowledge well-suited for complex agentic AI design and implementation tasks.

## Interview Guide

You are building an agent that needs to remember facts about a user across separate sessions that happen days apart. What memory approach would you use and why?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Suggests storing everything in the system prompt without addressing scalability or cross-session persistence. | | Mentions an external database but does not explain retrieval strategy or how relevant facts are selected at runtime. | | Describes a persistent external store, explains semantic or key-based retrieval, and discusses how retrieved facts are injected into the prompt at session start. |

---

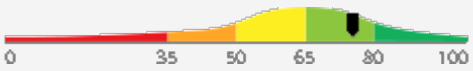
If a user has a long, multi-turn conversation with an AI agent, what challenges might arise with the agent remembering earlier parts of the conversation, and what is one way to address that?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Only mentions that the agent might forget things with no explanation of why or how to address it. | | Correctly identifies context window limits as the cause but offers only a general solution like summarization. | | Explains context window constraints, describes a concrete strategy such as summarization or external memory retrieval, and notes trade-offs. |

## Detail

### Multi-Agent Orchestration and Workflow Patterns

Score: 75




*Description:*

Covers how multiple specialized agents can be coordinated to complete complex tasks, including orchestrator-agent patterns, passing context between agents, and dividing work across agents with different roles. Includes understanding when to use a single agent versus a multi-agent setup.

*Interpretation:*

Candidate should achieve above average job performance in this area with little or no training.

The candidate demonstrates a solid and competent understanding of multi-agent orchestration, including orchestrator-agent patterns, context passing between agents, and the appropriate use of multi-agent versus single-agent designs. Minor gaps may exist in more nuanced or advanced aspects of agentic workflow architecture.

## Interview Guide

You are designing a multi-agent system where one orchestrator agent delegates subtasks to specialist agents. What information does the orchestrator need to pass to each specialist, and how do you ensure results are combined correctly?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Describes only sending the original user request to each agent with no mention of context, formatting, or result aggregation. | | Mentions passing relevant context but does not address result formatting, conflict resolution, or orchestrator logic. | | Describes scoped context passing, structured output formats, orchestrator logic for aggregating and validating results, and handling specialist failures. |

---

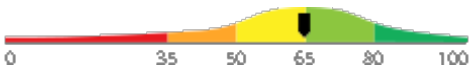
Can you explain the difference between using a single AI agent for a task versus splitting the work across multiple agents? When might you choose one approach over the other?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Cannot distinguish the two approaches or gives only a superficial answer based on task size. | | Correctly identifies that multi-agent setups handle complexity but cannot explain coordination or trade-offs clearly. | | Explains specialization, parallelism, and coordination overhead, and gives a concrete scenario for each approach with clear reasoning. |

## Detail

### Prompt Design for Agent Behavior

Score: 65




*Description:*

Focuses on writing and structuring prompts that guide an agent toward completing goals reliably. Includes system prompt construction, defining agent roles and constraints, formatting instructions for tool use, and iterating on prompts to fix unwanted agent behaviors.

*Interpretation:*

Candidate should achieve above average job performance in this area with little or no training.

The candidate exhibits a solid and proficient understanding of prompt design for directing agent behavior toward reliable goal completion. They are capable of constructing well-structured system prompts, clearly defining agent roles and constraints, formatting tool-use instructions effectively, and iterating on prompts to address and correct most unwanted agent behaviors with minimal supervision.

## Interview Guide

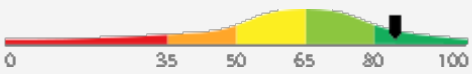
You notice an agent is frequently going off-topic or ignoring specific instructions in its system prompt. What steps would you take to diagnose and fix the prompt?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Suggests rewriting the whole prompt from scratch without a diagnostic process. | | Reviews the prompt for clarity but does not describe systematic testing or isolation of problem instructions. | | Describes isolating problematic instructions, testing variations, using examples, and validating fixes against consistent test cases. |

---

If you were setting up an AI agent to handle customer support tasks, what would you include in the agent's system prompt to help it behave correctly and stay on topic?

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Suggests only a brief greeting or role label with no mention of constraints or behavioral guidance. | | Mentions role definition and staying on topic but lacks specifics on constraints or output format. | | Describes role, scope boundaries, tone, escalation rules, and output format with clear rationale for each. |

| Detail | Interview Guide |
|---|---|
| **Tool Use and Function Calling**<br>Score: 84<br>  <br><br>*Description:*<br>Covers how agents interact with external tools, APIs, and services to retrieve information or take actions beyond text generation. Includes defining tool schemas, understanding how agents decide which tool to call, passing parameters correctly, and handling tool responses and errors.<br>*Interpretation:*<br>Candidate should achieve superior job performance in this area with little or no training.<br><br>The candidate demonstrates a comprehensive and sophisticated mastery of tool use and function calling in Agentic AI systems, including the full lifecycle of agent-tool interaction — from schema definition and tool selection logic to precise parameter passing and robust error handling. This level of performance reflects expert-level knowledge and the ability to design and evaluate complex agentic tool integrations. | An agent is supposed to call a weather API tool but sometimes calls it with missing or incorrect parameters. How would you identify the root cause and fix it?<br><br>1 ☆ Suggests only checking the API documentation without examining the agent's tool definition or prompt.<br>2 ☆<br>3 ☆ Reviews the tool schema and prompt but does not describe how to test or validate the fix systematically.<br>4 ☆<br>5 ☆ Audits tool schema definition, inspects agent logs for parameter values, updates prompt or schema, and validates with targeted test cases.<br><br>---<br><br>Can you explain what a tool or function call is in the context of an AI agent, and give an example of when an agent might need to use one?<br><br>1 ☆ Confuses tool calling with prompt writing or cannot give a concrete example of an external tool.<br>2 ☆<br>3 ☆ Correctly identifies that tools extend agent capability but gives a vague or generic example.<br>4 ☆<br>5 ☆ Clearly explains tool schemas, agent decision to call a tool, and gives a specific, realistic example with input/output. |

## Free Text Responses

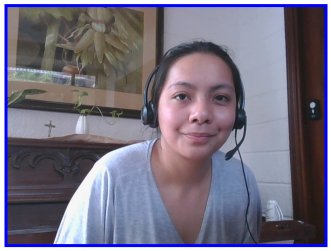
During the assessment, the candidate was asked to answer one or more questions using text, audio, video, or an uploaded text file. Their responses are included below for review.

| Question or Task | Response |
|---|---|
| After an AI product is deployed, what is model monitoring and why is it a necessary part of the product lifecycle? | Model monitoring is a technique for ensuring that the model does not wander or become overtrained after an extended period of repeated queries that have the same or similar prompts. This is very important for preventing hallucinization. It's also a key aspect of any guardrails strategy.<br><br>**Comments (AI):** The answer is clear and coherent but lacks depth in explaining the importance of model monitoring. The phrase 'hallucinization' is not commonly used in this context and may confuse readers. The answer could be improved by providing more specific examples of model performance metrics and how they are tracked. The argument strength is moderate as it does not fully explain why model monitoring is necessary in the product lifecycle.<br><br>**Misspelled Words:** guardrails (1), hallucinization (1) |

## Identity Confirmation Photos

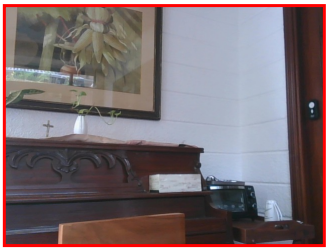
The following photos of the candidate and any identification were uploaded during the assessment session.

| Photo Analysis Results | |
|---|---|
| - Risk: | Medium risk of cheating based on image inconsistencies |
| - Percent match among processed faces | 100% |
| - Total images processed | 17 |
| - Total images with valid faces | 14 (82%) |
| - Total pairs of faces compared | 13 |
| - Pairs in which faces matched | 13 (100%) |

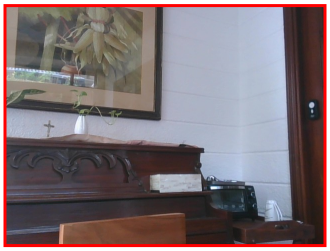


*Pre/Post-Test Photo*

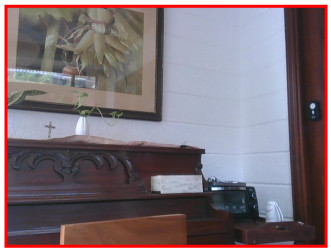
*ID Photo*



*In-Test Error Detected (No Face Detected)*

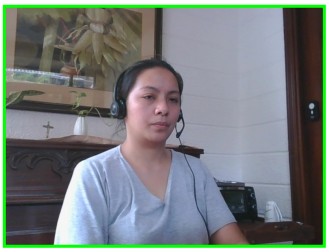


*In-Test Error Detected (No Face Detected)*

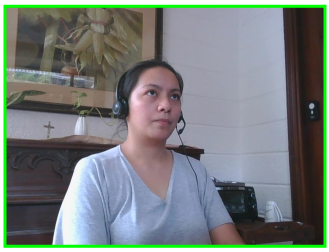


*In-Test Error Detected (No Face Detected)*

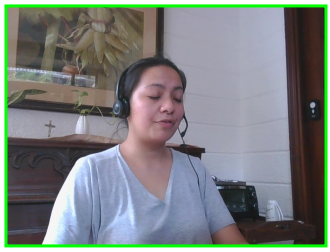
*In-Test Photo*



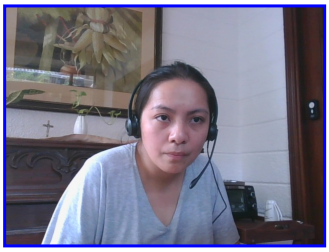
*In-Test Photo*



*In-Test Photo*



*In-Test Photo*



*Pre/Post-Test Photo*

# Resume or CV

## Summary Updated on

Motivated career professional with extensive experience in office administration and management. Proven track record of improving efficiency, reducing costs, and enhancing office operations through strategic initiatives and technology implementation.

### Objective

I am seeking a role where I can use my many skills and my exceptional judgment and empathy for customers to make a difference to a growing company.

### Education

- Associate of Applied Science in Office Administration, Portland Community College, 2020

### Experience

- General Office Clerk, Paramount Office Management, 09/2023 – Present
- Administrative Assistant, Global Enterprises Inc., 04/2021 – 08/2023
- Administrative Assistant, Innovative Business Solutions Ltd., 07/2019 – 03/2021

### Other Qualifications

- Microsoft Office Specialist (MOS) Certification
- Certified Administrative Professional (CAP)
- International Association of Administrative Professionals (IAAP) Certification

## Report Preparation Notes

- Hiring decisions should never be based on a single source of information. The most effective use of this assessment report is as a part of a multi-faceted program of candidate evaluation that includes resume review, interviews, and reference checks.
- Overall vs Percentiles Scores: The overall score reflects the success in the test, based on the mean (average) and standard deviation of the test scores. The percentile score reflects the percentage of test-takers who scored equal or below this overall score. We recommend you use the Overall Score as your primary evaluation criteria. However, percentile scores can often be useful in comparing specific candidates against one another and with a group, such as for test takers in a certain organization or within a certain account.
- Note that comparison information is calculated based on completed instances of this assessment at that time the assessment is scored. As additional instances are completed, the comparative data may change. You can always update a report to the current values by clicking on 'Recalculate Percentiles' within the online results viewing pages at www.hravatar.com.
- Most competency scores are norm-based, which means that they can be interpreted in terms of their distance from the average or mean score. For all scales, a score equal to the mean receives a score of 65 and scores above and below this value are set so that a score change of 15 equals one standard deviation.
- For linear competencies, higher is better across the entire scale. For these scales a score between 65 and 80 (light green) represents 0 to 1 standard deviation above the mean and a score above 80 (dark green) represents more than one standard deviation above the mean. Similarly, a score of 50 - 65 (yellow) represents 0 to 1 standard deviation below the mean, while a score of 35 - 50 (orange) equates to 1 to 2 standard deviations below the mean, and a score below 35 represents more than 2 standard deviations below the mean.
- Sim ID: 20890-1, Key: 0-0, Rpt: 68, Prd: 9710, Created: 2026-07-02 17:22 EDT
- UA: Mozilla/5.0 (Windows NT 6.3; Trident/7.0; Touch; rv:11.0) like Gecko

## Score Calculation Detail

The following table provides a summary of how the overall score was calculated from each of the individual competency scores. First, all competency scores are calculated on a scale of 0-100. Note that some competencies use their color category rather than their actual numeric score in the overall calculation. For these, a standard score associated with the assigned color category is used in the overall score calculation rather than the actual numeric score. This is reflected in the ''Score Value Used'' column. Next, a weighted average of scores is computed using individual competency weights, typically set using job analysis data provided by the US Government Occupational Information Network (O*Net).

| Competency | Score | How applied to overall | Score Value Used | Weight (%) |
|---|---|---|---|---|
| Agent Reasoning and Task Execution | 92.9587 | Numeric Score | 92.9587 | 8.3333 |
| Agent Reasoning and Task Execution (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Error Handling, Guardrails, and Safe Operation | 92.7588 | Numeric Score | 92.7588 | 8.3333 |
| Error Handling, Guardrails, and Safe Operation (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Memory and Context Management | 86.2867 | Numeric Score | 86.2867 | 8.3333 |
| Memory and Context Management (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Multi-Agent Orchestration and Workflow Patterns | 75.7320 | Numeric Score | 75.7320 | 8.3333 |
| Multi-Agent Orchestration and Workflow Patterns (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Prompt Design for Agent Behavior | 65.1830 | Numeric Score | 65.1830 | 8.3333 |
| Prompt Design for Agent Behavior (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Tool Use and Function Calling | 84.8402 | Numeric Score | 84.8402 | 8.3333 |
| Tool Use and Function Calling (Free Text Responses) | 53.8624 | Numeric Score | 53.8624 | 8.3333 |
| Weighted Average: | | | | 68.4112 |
| Final Overall Score: | | | | 68 |

## Notes

(This area is intentionally blank - it's reserved as space for your notes.)