hr | avatar
Illuminating Talent

Test Results and Interview Guide

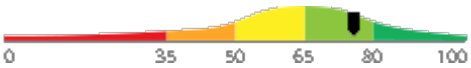
Candidate: **Elizabeth Wantsajob**
Assessment: Retrieval-Augmented Generation (RAG) Concepts
Completed: July 2, 2026
Prepared for: Sara Maple
Example Company

What's Included

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

Important Note: The Retrieval-Augmented Generation (RAG) Concepts assessment measures one or more important competencies, and collects audio or video responses to specific questions. Attribute types measured vary by test, but can include cognitive ability, skills, knowledge, personality characteristics, emotional intelligence, and past behavioral history. Various types of analysis may be conducted on the recorded responses depending on the test configuration. 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

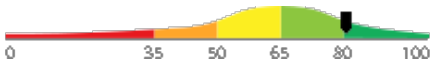
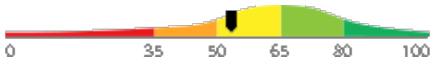
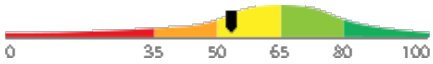
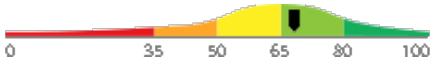
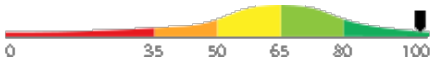
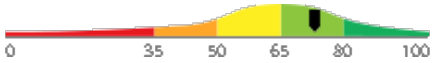
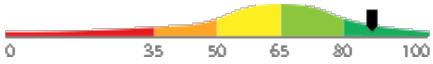
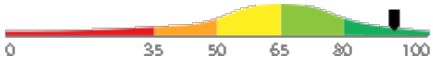
Candidate	Score	Interpretation
Elizabeth Wantsajob	76	

beth.wantsajob@gmail.com
Retrieval-Augmented Generation (RAG) Concepts
July 2, 2026

The candidate demonstrates a solid and broadly competent understanding of RAG concepts, design principles, and common implementation features, including vector databases, retrieval pipelines, and known limitations such as hallucinations and context window constraints. They are likely capable of contributing effectively to RAG-based application development at an entry-level to mid-level capacity with minimal supervision.

Key
- Candidate Score
- Higher Risk
- Lower Risk

Competency Summary

Competency	Score	Interpretation
Skills/Knowledge (relates to immediate readiness)		
Chunking Strategies	81	
RAG Architecture and Pipeline (Free Text Responses)	53	
Vector Embeddings and Similarity Search (Free Text Responses)	53	
Prompt Construction with Retrieved Context	68	
RAG Architecture and Pipeline	98	
RAG Limitations and Quality Improvement Techniques	73	
Vector Databases and Indexing	87	
Vector Embeddings and Similarity Search	92	

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	0 10 20 30 40 50 60 70 80 90 100
Global	76th	
North America	63rd	
United States	63rd	
Example Company	70th	

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): • (Generic Text for Sample Report) Strong performer in Drag and Drop Files tasks, indicating comfort with file management and basic computer interactions. • Demonstrates solid numerical accuracy in Recognizing and Confirming Numbers, a valuable asset in detail-oriented roles. • Moderate overall performance in Analytical Thinking and Attention to Detail, with adequate grammar skills but room for improvement. • Struggles with Reading and Analyzing Problems, which may limit effectiveness in roles requiring critical reading and complex problem-solving. • Lowest performance in Navigating Between Screens, suggesting difficulty with multi-screen software workflows that could impact productivity in computer-intensive roles. Narrative (AI): Elizabeth Wantsajob demonstrates a mixed profile of knowledge, skills, and abilities across the assessed competencies. 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. 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. 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. 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. Computed on: April 2, 2026, 11:09:49PM EDT

Detail

Candidate:	Elizabeth Wantsajob, beth.wantsajob@gmail.com
Assessment:	Retrieval-Augmented Generation (RAG) Concepts
Authorized:	July 2, 2026, by Sara Maple, Example Company, qamailsaram.mike@hravatar.com
Started:	July 2, 2026, 5:23:00PM EDT
Completed:	July 2, 2026, 5:23:00PM EDT
Overall Score:	76

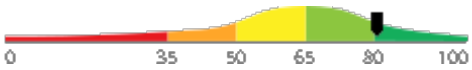
Knowledge and Skills Detail

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

Chunking Strategies

Score: 81



Description:

Understanding how source documents are split into smaller segments (chunks) before being embedded and indexed. This includes knowing common chunking approaches, how chunk size affects retrieval quality, and the trade-offs involved in choosing a chunking strategy.

Interpretation:

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

The candidate demonstrates an advanced and comprehensive mastery of chunking strategies in the context of Retrieval-Augmented Generation. They possess an expert-level understanding of document segmentation techniques, the nuanced effects of chunk size on retrieval performance, and the ability to critically assess and apply optimal chunking strategies for complex use cases.

Why is it important to think carefully about how you split documents into chunks, and what trade-offs do you consider when choosing a chunk size?

1	2	3	4	5
Does not understand why chunking matters or treats all chunk sizes as equivalent.		Recognizes that chunk size affects retrieval but gives only one-sided trade-off analysis.		Discusses context preservation, retrieval precision, and LLM context limits with clear trade-off reasoning.

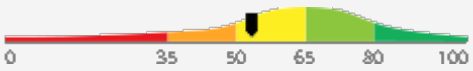
What is chunking in the context of RAG, and can you give an example of a simple chunking approach you might use?

1	2	3	4	5
Cannot define chunking or explain why documents need to be split.		Defines chunking and names a basic approach (e.g., fixed size) but cannot explain the rationale.		Defines chunking, explains the rationale, and describes at least one approach with its strengths.

Detail

RAG Architecture and Pipeline (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.

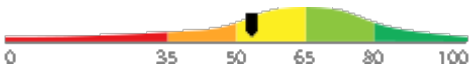
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

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

Vector Embeddings and Similarity Search (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.

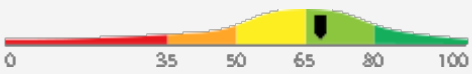
Can you walk me through the basic stages you would follow to take an AI-enabled product from an initial idea to a live deployment?

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 Construction with Retrieved Context

Score: 68



Description:

Understanding how retrieved document chunks are incorporated into the prompt that is sent to a large language model. This includes knowing how to structure prompts so the LLM uses the retrieved context to answer the query accurately, and being aware of context window limits.

Interpretation:

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

The candidate exhibits a solid understanding of how retrieved context is integrated into prompts and how to structure those prompts to guide a large language model toward accurate responses. They are generally knowledgeable about context window limitations and their implications for prompt design.

Interview Guide

How would you structure a prompt for a RAG system to make sure the language model uses the retrieved documents to answer the user's question?

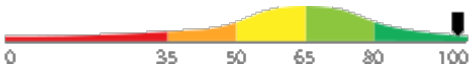
1	2	3	4	5
Cannot describe how to include retrieved context in a prompt or ignores the context window.		Describes including context in the prompt but lacks specifics on structure or context window awareness.		Describes a clear prompt structure with system instructions, retrieved context, and query; mentions token limits.

What is a context window, and why does it matter when you are passing retrieved documents to a language model in a RAG system?

1	2	3	4	5
Cannot define context window or explain why it is relevant to RAG.		Defines context window correctly but gives only a surface-level explanation of its impact.		Defines context window, explains token limits, and describes strategies like truncation or chunk selection.

RAG Architecture and Pipeline

Score: 98



Description:

Understanding how RAG combines a retrieval component with a large language model (LLM) to generate responses. This includes knowing the end-to-end flow: indexing documents, retrieving relevant content based on a query, and passing that content to an LLM as context to produce an answer.

Interpretation:

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

The candidate demonstrates a comprehensive and well-developed mastery of Retrieval-Augmented Generation concepts, including vector embeddings, vector databases, chunking strategies, similarity search, retrieval pipelines, prompt construction with retrieved context, re-ranking and filtering techniques, known system limitations, and the appropriate application of RAG versus alternative language model approaches. This individual is well-positioned to contribute effectively to the design and implementation of sophisticated RAG-based business applications with minimal oversight.

Interview Guide

Can you walk me through what happens, step by step, when a user submits a question to a RAG-based application?

1	2	3	4	5
Describes only one or two steps vaguely; confuses retrieval and generation roles.		Correctly describes the main steps (index, retrieve, prompt, generate) with minor gaps.		Clearly explains all steps in order, including how context is passed to the LLM and why.

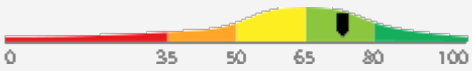
In a RAG system, what is the difference between the retrieval component and the generation component, and why do we need both?

1	2	3	4	5
Cannot clearly distinguish the two components or explain why both are needed.		Distinguishes the two components but gives a limited explanation of their interaction.		Clearly explains both roles and articulates how retrieval grounds the LLM's response in real data.

Detail

RAG Limitations and Quality Improvement Techniques

Score: 73



Description:

Understanding the common failure modes of RAG systems, such as hallucinations, poor retrieval, and irrelevant context, as well as practical techniques to improve output quality, including re-ranking retrieved results and filtering before passing context to the LLM.

Interpretation:

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

The candidate exhibits a solid understanding of RAG system failure modes and is knowledgeable about practical techniques to enhance output quality, such as re-ranking and context filtering. Minor gaps in knowledge may exist, but overall they demonstrate competent and reliable comprehension of this subject area.

Interview Guide

What are some common reasons a RAG system might give a poor or incorrect answer, and what steps can you take to improve retrieval quality?

1	2	3	4	5
Cannot identify common failure modes or improvement techniques beyond generic suggestions.		Identifies one or two failure modes and mentions re-ranking or filtering but lacks depth.		Identifies multiple failure modes and explains re-ranking, filtering, and prompt tuning with clear rationale.

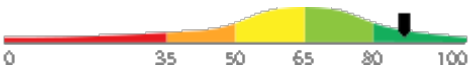
What is a hallucination in the context of a language model, and how can using RAG help reduce the risk of hallucinations?

1	2	3	4	5
Cannot define hallucination or explain how RAG relates to reducing it.		Defines hallucination correctly and notes RAG provides grounding context but lacks elaboration.		Defines hallucination, explains how retrieved context grounds the LLM, and notes RAG's remaining limitations.

Detail

Vector Databases and Indexing

Score: 87



Description:

Understanding the role of vector databases in storing, organizing, and efficiently searching large collections of embeddings. This includes knowing how documents are indexed before retrieval and why specialized vector storage is used instead of traditional databases.

Interpretation:

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

The candidate demonstrates an advanced and comprehensive mastery of vector databases and indexing within the context of Retrieval-Augmented Generation. They possess a thorough understanding of how embeddings are stored, organized, and efficiently searched, as well as a deep appreciation for the architectural advantages of specialized vector storage over traditional database solutions.

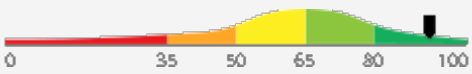
Interview Guide

What role does a vector database play in a RAG system, and how does it differ from a traditional relational database?

1	2	3	4	5
Cannot explain the role of a vector database or describes it as identical to a relational database.		Correctly identifies that vector databases store embeddings and support similarity search with some gaps.		Clearly explains vector storage, approximate nearest neighbor search, and contrast with relational DBs.

If you were setting up a RAG system, what would you store in the vector database and how would it get there?

1	2	3	4	5
Cannot describe what is stored or how embeddings are created and loaded into the database.		Mentions storing embeddings of document chunks but is vague about the indexing process.		Describes chunking, embedding generation, and loading vectors with metadata into the database clearly.

Detail	Interview Guide
Vector Embeddings and Similarity Search Score: 92  *Description:* Understanding how text is converted into numerical vectors (embeddings) that capture meaning, and how those vectors are used to find documents that are semantically similar to a user's query. This is the core mechanism that allows RAG systems to match queries to relevant content. *Interpretation:* Candidate should achieve superior job performance in this area with little or no training. The candidate demonstrates a comprehensive and sophisticated mastery of vector embeddings and similarity search, reflecting a deep understanding of how text is encoded into meaningful numerical representations and how those representations drive accurate semantic retrieval. They are well-equipped to apply these concepts effectively within retrieval-augmented generation systems.	How does a RAG system use vector embeddings to find documents that are relevant to a user's query? 1 Cannot explain the connection between embeddings and search; gives a keyword-based explanation. 2 3 Explains that embeddings represent meaning and that similarity is measured, but lacks detail. 4 5 Accurately describes embedding generation, vector space similarity (e.g., cosine similarity), and retrieval. What is a vector embedding, and why is it useful in a RAG system compared to searching documents by exact keywords? 1 Cannot define an embedding or articulate any advantage over keyword search. 2 3 Defines embeddings at a basic level and mentions semantic meaning but struggles with comparison. 4 5 Clearly defines embeddings, explains semantic search, and gives a concrete example of the advantage.

Free Text Responses

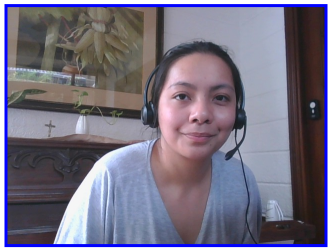
During the assessment, the candidate was asked to answer one or more questions using text, audio, video, or an uploaded text file. Their responses are included below for review.

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. **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. **Misspelled Words:** guardrails (1), hallucinization (1)

Identity Confirmation Photos

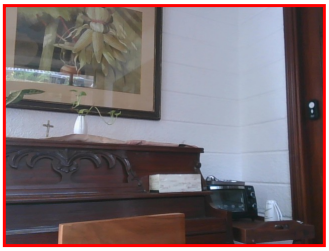
The following photos of the candidate and any identification were uploaded during the assessment session.

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%)

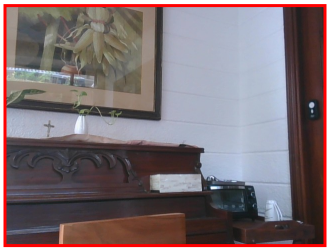


Pre/Post-Test Photo

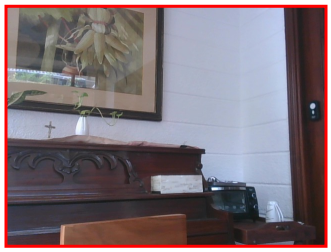
ID Photo



In-Test Error Detected (No Face Detected)

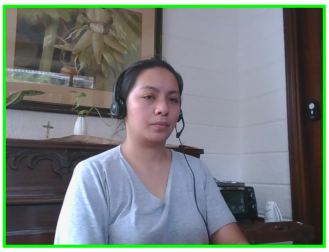


In-Test Error Detected (No Face Detected)

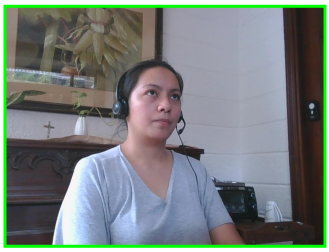


In-Test Error Detected (No Face Detected)

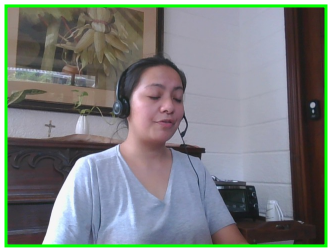
In-Test Photo



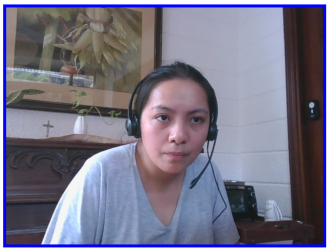
In-Test Photo



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: 20892-1, Key: 0-0, Rpt: 104, Prd: 9712, Created: 2026-07-02 17:23 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 (%)
Chunking Strategies	81.0116	Numeric Score	81.0116	12.5000
Prompt Construction with Retrieved Context	68.6547	Numeric Score	68.6547	12.5000
RAG Architecture and Pipeline	98.6634	Numeric Score	98.6634	12.5000
RAG Architecture and Pipeline (Free Text Responses)	53.8624	Numeric Score	53.8624	12.5000
RAG Limitations and Quality Improvement Techniques	73.5364	Numeric Score	73.5364	12.5000
Vector Databases and Indexing	87.1290	Numeric Score	87.1290	12.5000
Vector Embeddings and Similarity Search	92.4185	Numeric Score	92.4185	12.5000
Vector Embeddings and Similarity Search (Free Text Responses)	53.8624	Numeric Score	53.8624	12.5000
Weighted Average:				76.1423
Final Overall Score:				76

Notes

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