

Test Results and Interview Guide


Candidate: **Elizabeth Wantsajob**
Assessment: AI-Enabled Systems Engineering
Completed: May 22, 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 AI-Enabled Systems Engineering 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 beth.wantsajob@gmail.com AI-Enabled Systems Engineering May 22, 2026 Above average	77	

Key





- ▼ Candidate Score
- Higher Risk
- Lower Risk

Competency Summary

Competency	Score	Interpretation
Skills/Knowledge (relates to immediate readiness)		
AI Agents & Multi-Stage Agentic Workflows	85	
AI System Guardrails, Security & Responsible Deployment	66	
API Integration, Data Engineering & System Connectivity	78	
Large Language Model (LLM) Integration & Prompt Engineering	85	
Retrieval-Augmented Generation (RAG) Systems	72	

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	77th												
North America	64th												
United States	64th												
Example Company	71st												

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	-	-	<p>Summary Points (AI):</p> <ul style="list-style-type: none"> (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. <p>Narrative (AI): Elizabeth Wantsajob demonstrates a mixed profile of knowledge, skills, and abilities across the assessed competencies.</p> <p>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.</p> <p>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.</p> <p>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.</p> <p>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.</p> <p>Computed on: April 2, 2026, 11:09:49PM EDT</p>

Detail

Candidate: **Elizabeth Wantsajob**, beth.wantsajob@gmail.com
Assessment: AI-Enabled Systems Engineering
Authorized: May 22, 2026, by Sara Maple, Example Company, qamailsaram.mike@hravatar.com
Started: May 22, 2026, 2:58:43PM EDT
Completed: May 22, 2026, 2:58:43PM EDT
Overall Score: 77

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

AI Agents & Multi-Stage Agentic Workflows

Score: 85



Description:

Covers the design and implementation of AI agents — systems in which an LLM autonomously reasons, plans, and takes actions by calling tools, APIs, or other agents across multiple steps to complete complex tasks. This includes understanding agent frameworks, tool use, memory management, and orchestration of multi-agent pipelines, which are increasingly central to advanced AI business applications.

Interpretation:

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

The candidate exhibits an advanced and comprehensive mastery of AI agent design and multi-stage agentic workflows, including autonomous reasoning and planning, tool and API use, memory management, and the orchestration of complex multi-agent pipelines. They are highly capable of leading the architecture and implementation of sophisticated agentic AI systems within enterprise and advanced business application contexts.

Describe how you would architect a multi-stage agentic workflow for a business process — for example, a pipeline that researches a topic, summarizes findings, and drafts a report — including how you would handle failures, loops, and quality control between stages.



1

The candidate describes a linear sequence of LLM calls without addressing agent orchestration, inter-stage data passing, error handling, retry logic, or mechanisms to prevent infinite loops or runaway agents.



2

The candidate describes breaking the workflow into discrete agent steps with defined inputs and outputs, and mentions basic error handling, but does not address state management, guardrails, human-in-the-loop checkpoints, or evaluation of intermediate outputs.



3



4



5

The candidate outlines a well-structured multi-agent architecture with clearly defined agent roles, shared or scoped memory/state, tool definitions, orchestrator logic, inter-stage validation, failure handling with fallbacks, guardrails to constrain agent behavior, and optionally human-in-the-loop review — demonstrating awareness of reliability and safety in agentic systems.

How would you describe what an AI agent is, and how does it differ from simply sending a single prompt to an LLM and receiving a response?



1

The candidate cannot clearly distinguish an agent from a standard LLM call, or describes an agent in a way that is indistinguishable from a basic chatbot with no tool use or multi-step reasoning.



2

The candidate correctly identifies that an agent can take multiple steps and use tools or external resources to complete a task, but cannot explain how the agent decides what actions to take or how results are fed back into the reasoning loop.



3



4



5

The candidate clearly explains that an agent uses an LLM as a reasoning engine that iteratively decides which tools or actions to invoke, observes the results, and continues reasoning until a goal is achieved — and can describe the think-act-observe loop (e.g., ReAct pattern) and give a concrete business example.

Detail

Interview Guide

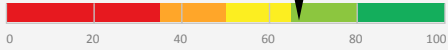
Results by Topic for AI Agents & Multi-Stage Agentic Workflows

Agent Tool Use & Orchestration: 0 of 2 Correct

Agent Memory & State Management: 2 of 4 Correct

Detail
Interview Guide
AI System Guardrails, Security & Responsible Deployment

Score: 66


Description:

Covers the practical measures required to deploy AI-enabled applications safely and responsibly in business environments, including input and output validation, prompt injection defense, content filtering, access controls, and monitoring for model misbehavior or drift. These practices are essential for any production AI system and are applied continuously across the application lifecycle.

Interpretation:

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

The candidate demonstrates a solid and well-rounded knowledge of AI system guardrails, security, and responsible deployment practices, including input and output validation, prompt injection defense, content filtering, and monitoring for model misbehavior or drift. They are likely capable of applying these measures effectively in most production AI environments with only occasional need for guidance on more advanced or nuanced scenarios. This individual is well-positioned to contribute meaningfully to the safe and responsible deployment of AI-enabled systems.

What guardrails and safety mechanisms would you put in place when deploying an LLM-powered application for business use, and how would you monitor the system over time to catch issues before they impact users or the business?



1

The candidate mentions only surface-level measures such as 'add a content filter' without describing input validation, output validation, access controls, logging, alerting, or a plan for ongoing monitoring and response to incidents.



2

The candidate describes a reasonable set of guardrails including input/output filtering, rate limiting, and logging, and mentions monitoring for inappropriate outputs, but does not address adversarial testing, model drift, audit trails, incident response, or role-based access controls.



3



4



5

The candidate outlines a comprehensive safety and monitoring strategy covering input validation and sanitization, output filtering and schema enforcement, prompt injection defenses, role-based access controls, audit logging, real-time alerting on anomalous outputs, periodic red-teaming or adversarial testing, monitoring for model performance drift, and a defined incident response process — demonstrating mature, production-oriented thinking.

What is prompt injection, and why is it a concern when building a business application that uses an LLM — can you give an example of how it might happen?



1

The candidate cannot define prompt injection or confuses it with SQL injection without drawing any meaningful analogy, and cannot provide a relevant example in the context of an LLM-powered application.



2

The candidate correctly defines prompt injection as a user attempting to override or manipulate the system prompt through crafted input, and gives a basic example, but cannot describe mitigation strategies.



3



4



5

The candidate clearly defines prompt injection, provides a concrete and realistic example (e.g., a user embedding instructions in their input to make the model ignore its system prompt or reveal confidential instructions), and describes at least one mitigation strategy such as input sanitization,

Detail	Interview Guide
	strict output validation, privilege separation, or sandboxing tool access.
Results by Topic for AI System Guardrails, Security & Responsible Deployment	
Input/Output Validation & Content Filtering:	1 of 2 Correct
AI Monitoring, Access Controls & Responsible Deployment:	1 of 2 Correct

Detail

Interview Guide

API Integration, Data Engineering & System Connectivity

Score: 78



Description:

Addresses the practical skills needed to connect AI components to external data sources, services, and business systems through APIs, data pipelines, and integration patterns. This includes ingesting and preprocessing data for AI consumption, managing data flows between services, and ensuring that AI applications can reliably read from and write to the systems they are embedded in.

Interpretation:

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

The candidate demonstrates a solid and proficient understanding of connecting AI components to external data sources, services, and business systems. They are capable of independently designing and managing data pipelines, applying appropriate integration patterns, and ensuring reliable read and write operations across AI-embedded systems.

How would you design a data ingestion pipeline that keeps the knowledge base of a RAG system up to date as source documents change, and what challenges would you anticipate and address?



1

The candidate describes only a one-time batch load of documents without addressing incremental updates, change detection, re-embedding of modified content, deletion of stale records, or pipeline scheduling and monitoring.



2

The candidate identifies the need for incremental updates and describes a basic approach such as polling for changes or using timestamps, but does not fully address document versioning, deduplication, embedding consistency, or failure recovery.



3



4



5

The candidate describes a well-designed pipeline covering change detection (e.g., webhooks, polling, CDC), incremental ingestion with deduplication and versioning, re-chunking and re-embedding of changed documents, deletion or invalidation of stale vectors, pipeline scheduling and monitoring, and handling of failures — demonstrating practical data engineering competency in the context of AI systems.

When building an AI application that needs to pull data from an external source — like a company database or a third-party API — what are the basic steps you would take to make that connection work reliably?



1

The candidate provides a very vague answer (e.g., 'just call the API') without mentioning authentication, error handling, data formatting, or how the retrieved data would be prepared for use by the AI component.



2

The candidate describes the basic mechanics of making an API call and handling a response, and mentions authentication, but does not address data validation, transformation for AI consumption, rate limiting, retries, or monitoring.



3



4



5

The candidate describes a robust integration approach including authentication and credential management, request/response handling, data validation and transformation (e.g., cleaning, formatting for embedding or prompt injection), error handling with retries and fallbacks, rate limit awareness, and logging — demonstrating readiness to build production-grade

Detail

Interview Guide

integrations.

Results by Topic for API Integration, Data Engineering & System Connectivity

API Integration & Service Connectivity: 3 of 4 Correct

Data Ingestion, Preprocessing & Pipeline Management: 3 of 4 Correct

Management:

Detail
Interview Guide
Large Language Model (LLM) Integration & Prompt Engineering

Score: 85


Description:

Covers the practical knowledge required to integrate large language models into business applications, including selecting appropriate models, constructing effective prompts, managing context windows, and controlling model outputs. This is a foundational skill for building AI-enabled applications and is applied in nearly every stage of development.

Interpretation:

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

The candidate exhibits a comprehensive and sophisticated command of the technical concepts required to build, deploy, and maintain secure, scalable, and highly functional AI-enabled business applications. Knowledge spans the full breadth of the domain, including advanced topics such as AI agent design, multi-stage workflows, LLM integration, Retrieval-Augmented Generation, and system guardrails, reflecting the level of expertise expected of a highly proficient practitioner in AI-enabled systems engineering.

Walk me through how you would design and iteratively refine a prompt for an LLM that needs to extract structured data (e.g., key fields from a customer support ticket) reliably and consistently in a production application.



1

The candidate describes a basic prompt without addressing consistency, structure, or iteration. They do not mention output formatting, validation, or how to handle edge cases.



2

The candidate describes using clear instructions and possibly few-shot examples to guide the model, and mentions testing the prompt against varied inputs, but does not address schema enforcement, fallback handling, or systematic evaluation.



3



4



5

The candidate outlines a structured approach including defining the desired output schema (e.g., JSON), using system and user message roles, incorporating few-shot examples, testing against diverse and adversarial inputs, using output parsers or validation layers, and iterating based on failure analysis — demonstrating production-readiness thinking.

Can you describe what a prompt is in the context of working with a large language model, and give an example of how you might write a prompt to get a useful response from an LLM in a business application?



1

The candidate provides a vague or incorrect description of a prompt, or gives an example that is generic and shows little understanding of how prompt structure affects model output.



2

The candidate correctly describes a prompt and provides a reasonable example, but does not demonstrate awareness of techniques such as role assignment, few-shot examples, or output formatting instructions.



3



4



5

The candidate clearly explains what a prompt is, provides a well-constructed example relevant to a business use case, and mentions at least one technique (e.g., system messages, few-shot prompting, output constraints) that improves response quality and reliability.

Results by Topic for Large Language Model (LLM) Integration & Prompt Engineering

Prompt Construction & Optimization: 3 of 4 Correct

LLM Integration & Output Management: 1 of 4 Correct

Detail
Interview Guide
Retrieval-Augmented Generation (RAG) Systems

Score: 72


Description:

Encompasses the design, implementation, and maintenance of systems that enhance LLM responses by retrieving relevant external information from knowledge bases, vector databases, or document stores at inference time. RAG is one of the most widely used architectural patterns in production AI business applications, enabling models to access up-to-date and domain-specific knowledge.

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 Retrieval-Augmented Generation systems, including the integration of vector databases, knowledge bases, and document stores within production AI applications. They are well-positioned to independently contribute to the design, implementation, and maintenance of RAG-based architectures with minimal oversight.

Describe the end-to-end architecture of a RAG pipeline you would build for a business application, and explain the key decisions you would make at each stage — from document ingestion to final response generation.



1

The candidate describes only a high-level flow (store documents, search, ask the LLM) without addressing chunking strategy, embedding model selection, vector database choice, retrieval ranking, prompt construction with context, or response quality evaluation.



2

The candidate describes most of the major stages including document chunking, embedding, vector storage, similarity search, and context injection into the prompt, but does not address trade-offs such as chunk size, hybrid search, re-ranking, or handling retrieval failures.



3



4



5

The candidate provides a thorough pipeline description covering document preprocessing and chunking strategies, embedding model selection, vector database indexing, hybrid or semantic search with optional re-ranking, dynamic prompt construction with retrieved context, handling of context window limits, and evaluation of retrieval and generation quality — demonstrating practical production experience.

In simple terms, can you explain why a business might use a Retrieval-Augmented Generation (RAG) approach instead of just sending a question directly to an LLM?



1

The candidate cannot explain the concept or gives an incorrect explanation, such as confusing RAG with model fine-tuning or being unable to articulate why retrieved context improves responses.



2

The candidate correctly identifies that RAG allows the model to reference external documents or data not in its training set, improving accuracy for domain-specific questions, but cannot describe the retrieval mechanism or its components.



3



4



5

The candidate clearly explains that RAG grounds the LLM in relevant, current, or proprietary information by retrieving documents at query time, reducing hallucination and enabling use of private data — and can name the core components: a retriever (e.g., vector search), a knowledge store, and the LLM that synthesizes the answer.

Detail

Interview Guide

Results by Topic for Retrieval-Augmented Generation (RAG) Systems

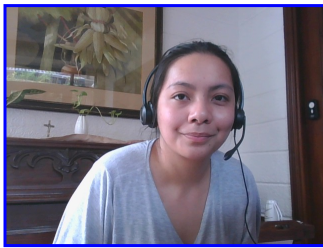
Retrieval Pipeline Design and Optimization:	1 of 4 Correct
RAG System Evaluation and Quality Assessment:	0 of 2 Correct

Identity Confirmation Photos

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

Photo Analysis Results

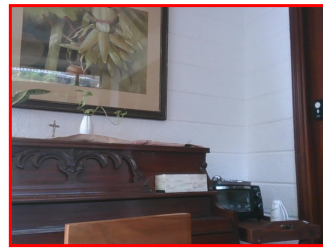
- Risk:	Medium risk of cheating based on image inconsistencies
- Percent match among processed faces	100%
- Total images processed	17
- Total images with valid faces	14 (82%)
- Total pairs of faces compared	13
- Pairs in which faces matched	13 (100%)



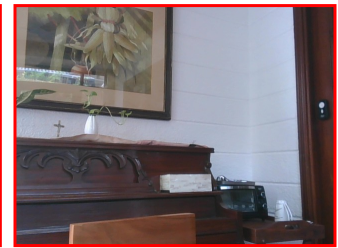
Pre/Post-Test Photo



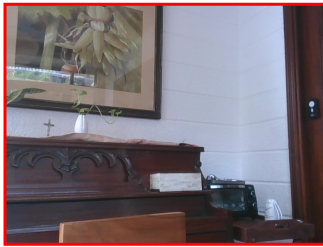
ID Photo



In-Test Error Detected (No Face Detected)



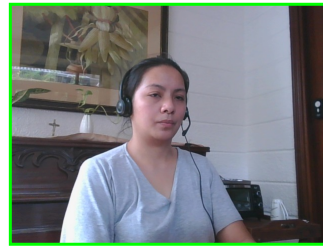
In-Test Error Detected (No Face Detected)



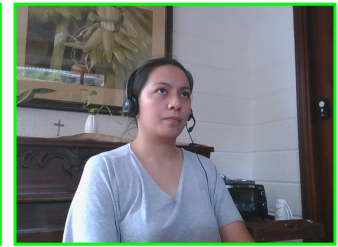
In-Test Error Detected (No Face Detected)



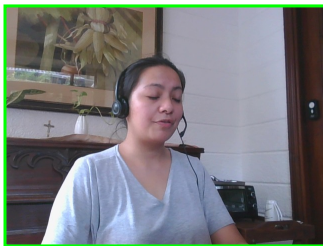
In-Test Photo



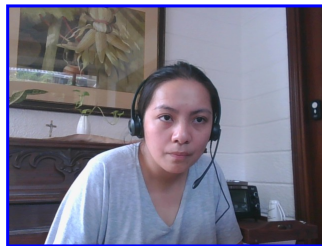
In-Test Photo



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: 20409-1, Key: 0-0, Rpt: 68, Prd: 9496, Created: 2026-05-22 14:58 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 (%)
AI Agents & Multi-Stage Agentic Workflows	85.4506	Numeric Score	85.4506	20.0000
AI System Guardrails, Security & Responsible Deployment	66.9719	Numeric Score	66.9719	20.0000
API Integration, Data Engineering & System Connectivity	78.7344	Numeric Score	78.7344	20.0000
Large Language Model (LLM) Integration & Prompt Engineering	85.9922	Numeric Score	85.9922	20.0000
Retrieval-Augmented Generation (RAG) Systems	72.5786	Numeric Score	72.5786	20.0000
Weighted Average:				77.9455
Final Overall Score:				77

Notes

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