GARP Risk and AI (RAI) Practice Exam (Sample)

Study Guide

Everything you need from our exam experts!

Table of Contents

Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- **Practice answering questions under realistic conditions,**
- **Improve accuracy and speed,**
- **Review explanations to strengthen weak areas, and**
- **Approach the exam with greater confidence.**

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

How to Use This Guide

This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here’s how to get the most out of your Examzify study guide:

1. Start with a Diagnostic Review

Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.

2. Study in Short, Focused Sessions

Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations.

3. Learn from the Explanations

After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.

4. Track Your Progress

Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.

5. Simulate the Real Exam

Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.

6. Repeat and Review

Don’t just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.

There’s no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!

Questions

1. **Which practice involves deleting data as soon as it is no longer needed and using encryption and anonymization to keep data safe?**

   A. Data Deletion and Security

   B. Data Encryption and Anonymization

   C. Data Retention Policy

   D. Data Backup and Recovery

2. **Using personal data only in accordance with the norms of the context in which it was originally given.**

   A. Data Minimization

   B. Right to be Forgotten

   C. Explainability

   D. Contextual Integrity

3. **Which term describes a model that fits training data very closely but generalizes poorly to new data?**

   A. Underfitted Models

   B. Overfitted models

   C. Plateaus

   D. Vanishing Gradients

4. **Which measure assesses the expected cumulative future rewards from a given state, ignoring the specific actions taken?**

   A. Action-Value Function

   B. Value Function

   C. Reward

   D. Policy

5. **Which term describes extreme observations that can bias regression results?**

   A. Ridge

   B. Residual Plots

   C. Outliers

   D. Target Variable

6. Which principle is central to data protection compliance?

A. Data Minimization

B. Consent Obligation

C. Data Retention

D. Data Sharing

7. Which type of hierarchical clustering starts with individual data points and groups them by similarity?

A. Agglomerative Clustering

B. Divisive Clustering

C. Hierarchical Clustering

D. Initialization

8. Which search method divides the search space in half to locate an answer more quickly?

A. Heuristic Search

B. Adversarial Search

C. Binary Search

D. Frame Problem

9. Shapley Values quantify what in a model?

A. The speed of model inference

B. The contribution of each feature to the prediction

C. The overall accuracy of the model

D. The number of features used

10. What is the focus of value-based reinforcement learning?

A. Value-based approach

B. Policy-based approach

C. Deep Reinforcement Learning

D. Monte Carlo method

Answers

1. A
2. D
3. B
4. B
5. C
6. A
7. A
8. C
9. B
10. A

Explanations

1. Which practice involves deleting data as soon as it is no longer needed and using encryption and anonymization to keep data safe?

A. Data Deletion and Security

B. Data Encryption and Anonymization

C. Data Retention Policy

D. Data Backup and Recovery

The main idea is combining minimizing what you keep with strong protections for what you do keep. Deleting data as soon as it's no longer needed reduces the amount of information that could be exposed in a breach. At the same time, using encryption and anonymization helps keep the remaining data safe by making it unreadable without keys and by removing identifiable details so people can't be tied to individuals. Because this option explicitly covers both deleting unused data and applying protective measures, it best matches the described practice. Other options either focus on protection alone without deletion, or on retention rules or backup and recovery rather than on removing data when it's no longer needed.

2. Using personal data only in accordance with the norms of the context in which it was originally given.

A. Data Minimization

B. Right to be Forgotten

C. Explainability

D. Contextual Integrity

Contextual Integrity centers on keeping information flows aligned with the expectations and norms of the situation in which the data was given. The statement captures that idea directly: personal data should be used only in ways that fit the original context—by the same actors, for the same purposes, and under the same conditions that were understood at the time of collection. This view of privacy emphasizes that what's appropriate depends on the social setting and purpose, not just on limiting data or ensuring explanations. Data minimization would focus on collecting only what's necessary, which is related but not the same idea. The right to be forgotten is about deleting data on request. Explainability deals with making automated decisions understandable. None of these describe restricting data use to fit the contextual norms as precisely as Contextual Integrity does.

3. Which term describes a model that fits training data very closely but generalizes poorly to new data?

A. Underfitted Models

B. Overfitted models

C. Plateaus

D. Vanishing Gradients

Overfitting is when a model learns the training data too well, capturing noise and peculiarities that don't repeat in new data. This leads to very low training error but poor performance on unseen data because the model has effectively memorized the training set rather than learned general patterns. It often happens when the model has high capacity relative to the amount of data, allowing it to fit idiosyncrasies instead of underlying trends. In contrast, underfitting means the model is too simple to capture the patterns, resulting in error on both training and test data. Plateaus relate to slow learning in certain regions of the optimization landscape, and vanishing gradients describe diminishing gradient signals in deep networks—neither specifically describe poor generalization. So the term for a model that fits training data very closely but generalizes poorly is overfitted models.

4. Which measure assesses the expected cumulative future rewards from a given state, ignoring the specific actions taken?

A. Action-Value Function

B. Value Function

C. Reward

D. Policy

The value function measures how good it is to be in a given state when you follow a specific policy, by looking at the expected total reward from that point onward. It captures the future, discounted rewards but does not pin down a particular action for that state. In other words, V(s) averages over all actions the policy might take in state s and the resulting future states, weighting by how the policy behaves, to give a single number for the state's desirability under that policy. Formally, you're looking at the expected sum of discounted rewards starting from state s and following the policy: $V(s) = E_\pi [\sum_{t=0}^{\infty} \gamma^t R_{t+1} | S0 = s]$. This focus on the state value under the policy distinguishes it from the action-value function Q(s,a), which would require choosing a specific action a in that state and then continuing. The immediate reward is just the snippet of reward at one step, whereas the value function concerns the entire future return from the state under the policy.

5. Which term describes extreme observations that can bias regression results?

A. Ridge

B. Residual Plots

C. Outliers

D. Target Variable

Outliers are extreme observations that can bias regression results. When a data point lies far from the overall pattern, it can pull the regression line toward itself, changing the estimated slope and intercept and, in turn, altering predictions and model metrics for the rest of the data. This is why detecting and handling outliers is important in regression analysis: they can distort the fit and lead to misleading conclusions about relationships between variables. Ridge is a regularization technique that shrinks coefficients to avoid overfitting, not specifically defined by extreme observations. Residual plots are diagnostic tools used to assess how well the model fits the data by examining residual patterns, not to describe extreme observations themselves. The target variable is simply what you're trying to predict, not a description of extreme data points.

6. Which principle is central to data protection compliance?

A. Data Minimization

B. Consent Obligation

C. Data Retention

D. Data Sharing

The central principle being tested is data minimization—the idea that organizations should only collect, use, and retain the minimum amount of personal data needed to achieve a stated purpose. This matters for compliance because privacy laws require data to be adequate, relevant, and not excessive for its intended purpose. By limiting data collection and processing to what is strictly necessary, you reduce the risk of unnecessary exposure, errors, and misuse, and you make it easier to grant data subjects access, deletion, or correction rights. In practice, data minimization guides the entire data lifecycle: define the purpose clearly at the outset, collect only what is needed to achieve that purpose, and retain data only as long as it's necessary. While consent is important for lawful processing, and retention and sharing policies are also crucial, minimization is the overarching guardrail that constrains quantity and scope of data across collection, use, and storage. For example, if you're running a signup form, gathering only the information truly needed for account creation and communication keeps data well within the minimum necessary boundary, reducing risk and simplifying compliance.

7. Which type of hierarchical clustering starts with individual data points and groups them by similarity?

A. Agglomerative Clustering

B. Divisive Clustering

C. Hierarchical Clustering

D. Initialization

Agglomerative clustering is the bottom-up form of hierarchical clustering that starts with each data point as its own cluster and then merges the most similar pairs of clusters step by step. At each step, the two clusters that are closest according to a chosen linkage criterion (such as single-link, complete-link, average-link, or Ward) are combined, until all points fall into a single cluster or until a desired number of clusters remains. This approach creates a dendrogram that shows how clusters are nested within larger clusters, capturing the hierarchical structure of the data. Divisive clustering is the opposite, beginning with all points in one cluster and repeatedly splitting them. The general term hierarchical clustering covers both approaches, while initialization is not a clustering method.

8. Which search method divides the search space in half to locate an answer more quickly?

A. Heuristic Search

B. Adversarial Search

C. Binary Search

D. Frame Problem

Binary search halves the search space at each step to find an item quickly. This method works on a sorted collection: compare the target with the middle element, then discard the half that cannot contain the target and continue with the remaining half. Repeating this process reduces the search area dramatically, giving a time complexity of about O(log n), which is why it's much faster than checking items one by one in large datasets. Other approaches don't inherently cut the space in half. Heuristic search uses knowledge about the domain to guide exploration but doesn't guarantee halving the search space. Adversarial search looks at game trees to anticipate opponent moves and uses strategies like minimax, not a binary halving procedure. The frame problem concerns representing and reasoning about actions in an environment, not about efficient search.

9. Shapley Values quantify what in a model?

A. The speed of model inference

B. The contribution of each feature to the prediction

C. The overall accuracy of the model

D. The number of features used

Shapley values attribute a model's prediction to its input features. They measure how much each feature contributes to pushing the prediction away from a baseline (often the average prediction) by averaging the feature's marginal impact across all possible subsets of other features. This creates a fair breakdown where the contributions from all features add up to the difference between the actual prediction and the baseline. It also preserves useful properties: a feature that never changes the output gets a zero value; features with identical effects receive equal shares; and the total of all feature contributions matches the prediction. Shapley values can explain a single prediction (local) or be aggregated for a sense of global behavior. They are not about how fast the model runs, the overall accuracy, or how many features are used.

10. What is the focus of value-based reinforcement learning?

A. Value-based approach

B. Policy-based approach

C. Deep Reinforcement Learning

D. Monte Carlo method

Value-based reinforcement learning centers on learning value functions that estimate how good it is to be in a state or to take a specific action in a state. The key idea is to predict future rewards, and then let the agent act by choosing actions that maximize those estimated values. In practice, you learn a value function like V(s) or Q(s, a), and the policy is derived from it—usually by selecting the action with the highest estimated value (greedy or with some exploration, like epsilon-greedy). This makes the behavior emerge from the value estimates rather than being directly programmed into a separate policy. Classic examples include methods that maintain a Q-function and use it to pick actions with the largest Q-value, such as Q-learning and, in deep RL, Deep Q-Networks. By focusing on accurate value estimation, these approaches aim to capture the future reward landscape of the environment so optimal decisions follow from it. Policy-based approaches, in contrast, directly optimize a parameterized policy without necessarily grounding decisions in a value function, which is why they aren't the defining focus here. Monte Carlo methods can be used within value- or policy-based settings to estimate returns from episodes, but the essential focus of value-based methods remains on learning and using value functions to drive action choice.

Next Steps

Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.

As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.

If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.

Or visit your dedicated course page for more study tools and resources:

https://garprai.examzify.com

We wish you the very best on your exam journey. You've got this!