PRACTICE EXAM 4: AIGP SIMULATION (100 QUESTIONS)

1. A hospital deploys an AI diagnostic system that was trained on patient data from 2019-2023. In early 2025, the hospital begins treating patients with a newly emerging respiratory illness that presents differently from any condition in the training data. The system's overall accuracy metrics remain within acceptable thresholds because the new illness represents only 2% of cases. However, it misdiagnoses 85% of patients with the new condition. What type of issue is this?

A. Data drift, because the statistical distribution of incoming patient symptoms has shifted relative to the training data distribution

B. Model drift, because the AI model's internal parameters have degraded over time due to computational resource limitations

C. A monitoring design failure, because the system's performance metrics should have detected the 85% misdiagnosis rate immediately

D. Concept drift, because the emergence of a new condition has changed the relationship between symptom presentations and diagnostic outcomes in ways the model was never trained to recognize

2. An AI governance committee must decide between two approaches for a high-risk lending AI system. Approach 1: Deploy the system with human-in-the-loop oversight where a loan officer reviews every AI recommendation before it becomes a decision. Approach 2: Deploy with human-on-the-loop oversight where the system makes decisions autonomously but a compliance team reviews a statistical sample of decisions daily. For a high-risk lending system that directly affects individuals' access to credit, which approach is MOST appropriate and why?

A. Approach 1, because high-risk systems making consequential financial decisions affecting individuals' rights require human review before each decision, not statistical sampling after the fact

B. Approach 2, because statistical sampling provides more comprehensive oversight than individual case review and is the standard required by the EU AI Act

C. Either approach is equally valid because the EU AI Act does not specify which human oversight model must be used for any category of high-risk system

D. Neither approach is adequate because high-risk lending systems must use human-in-command oversight exclusively, with full manual override of every system parameter

3. An organization's AI system processes customer data for personalized product recommendations. The system also generates inferences about customers' likely income ranges, family status, and health conditions based on purchasing patterns. The privacy notice discloses only that data is used for "product recommendations and service improvement." A governance professional identifies a compliance gap. Between purpose limitation and transparency, which principle is MOST directly violated by the generation of income, family, and health inferences?

A. Purpose limitation is the primary violation because the inference generation goes beyond the disclosed purposes, but transparency is not violated because the system's existence is disclosed

B. Transparency is the primary violation because the privacy notice fails to inform individuals about the specific types of inferences being generated, but purpose limitation is satisfied because inference generation serves the disclosed purpose of service improvement

C. Both principles are violated — purpose limitation because generating sensitive inferences likely exceeds the stated purpose of product recommendations, and transparency because individuals are not informed that the system infers income, family status, and health conditions from their purchasing behavior

D. Neither principle is violated because purchasing pattern analysis is a standard component of product recommendation systems and customers reasonably expect this type of processing

4. A deployer of a high-risk AI system discovers through its own monitoring that the system's fairness metrics have degraded significantly. The deployer informs the provider, who acknowledges the issue but states that a fix will take four months. The deployer also notifies the relevant market surveillance authority. While waiting for the provider's fix, the deployer continues to operate the system unchanged. Is the deployer meeting its governance obligations?

A. Yes, because the deployer has fulfilled both its notification obligations — informing the provider and the market surveillance authority — and has no independent obligation to take further action

B. No, because the deployer has independent obligations to mitigate harm that are not contingent on the provider's timeline — the deployer should implement interim measures such as manual review, scope reduction, or system suspension

C. Yes, because the EU AI Act assigns responsibility for system fixes exclusively to providers, and deployers are prohibited from modifying the system's operation without the provider's authorization

D. No, but only because the deployer failed to also notify the European AI Office, which is the correct regulatory body for fairness-related complaints about high-risk AI systems

5. An organization is evaluating whether its AI-powered chatbot falls under the EU AI Act's limited-risk transparency provisions or whether it is classified as minimal risk with no specific obligations. The chatbot provides general product information on the company's website. It does not make decisions affecting individuals, does not process sensitive data, and does not generate deepfake content. Under the EU AI Act, what determines whether the transparency obligation applies?

A. Whether the chatbot processes personal data, because only chatbots that collect user information trigger the limited-risk transparency requirements

B. Whether the chatbot uses a large language model architecture, because only LLM-based systems are subject to transparency requirements under the Act

C. Whether the chatbot has been classified as high-risk under Annex III, because transparency obligations only apply to systems that have undergone a conformity assessment

D. Whether the chatbot interacts with natural persons, because the Act requires that AI systems designed to interact with people must ensure individuals are informed they are interacting with AI

6. A data science team proposes using transfer learning — taking a model pre-trained on a general dataset and fine-tuning it on the organization's specific data — to accelerate development of a high-risk AI system. The governance professional must evaluate this approach. What is the MOST important governance consideration specific to transfer learning?

A. Transfer learning is prohibited for high-risk AI systems under the EU AI Act because the provider cannot fully document a pre-trained model that was developed by another entity

B. Transfer learning eliminates the need for bias testing on the fine-tuned model because the pre-trained model has already been validated by its original developer

C. The pre-trained model may carry biases, limitations, and behaviors from its original training that persist into the fine-tuned model — governance must evaluate the base model's characteristics and test the fine-tuned model independently rather than assuming that fine-tuning resolves inherited issues

D. Transfer learning requires the organization to obtain a GPAI model license from the European AI Office before incorporating any pre-trained model into a high-risk system

7. An organization conducts bias testing on its AI hiring system before deployment. The testing reveals that the system produces equitable outcomes across gender and racial groups when evaluated on the test dataset. Six months after deployment, monitoring reveals significant disparities have developed. What is the MOST likely explanation for this divergence between pre-deployment testing and post-deployment monitoring results?

A. The test dataset was not sufficiently representative of the actual applicant population — demographic composition, qualification distributions, or application patterns in the real world differ from the test data, causing the model to behave differently in production than in testing

B. The AI vendor secretly updated the model after deployment without notifying the deployer, introducing new biases that were not present in the original version

C. The monitoring system is producing false positive bias alerts due to a statistical error in the fairness metric calculation that should be investigated and corrected

D. The model's source code was modified by a malicious insider who deliberately introduced discriminatory logic into the system after it passed the pre-deployment bias audit

8. Under the EU AI Act, a company develops an AI system that assists teachers in evaluating student essays by providing suggested grades and feedback. The system does not make final grading decisions — teachers review and may modify the AI's suggestions. The company is uncertain whether this system is classified as high-risk. Which factor is MOST determinative for the classification?

A. Whether the system uses natural language processing technology, because NLP-based educational tools are automatically classified as high-risk regardless of how they are used

B. Whether the system is used to evaluate learning outcomes or determine access to education, because AI systems used in these educational contexts are explicitly listed as high-risk in Annex III

C. Whether teachers modify the AI's suggestions more than 50% of the time, because systems whose outputs are frequently overridden are automatically classified as limited risk

D. Whether the system is used in public or private educational institutions, because the EU AI Act's educational provisions apply exclusively to publicly funded schools and universities

9. An AI governance professional is advising a company that operates in both the EU and the United States. The company deploys an AI system for employee performance monitoring that analyzes email sentiment, meeting participation, and project completion rates. In the EU, this system may fall under the EU AI Act's restrictions on emotion recognition in workplaces. In the United States, no equivalent federal restriction exists. The company wants a unified global approach. What governance strategy BEST balances compliance and operational efficiency?

A. Apply U.S. standards globally because they are less restrictive, and the company's headquarters is in the United States

B. Deploy different versions of the system in each jurisdiction — a restricted version in the EU and an unrestricted version in the United States

C. Apply the EU AI Act's more restrictive standards globally to ensure compliance everywhere, accepting that this may limit functionality for U.S. employees that would otherwise be permissible

D. Delay deployment in both jurisdictions until the United States enacts federal AI legislation equivalent to the EU AI Act to ensure a level regulatory playing field

10. An organization's AI system for fraud detection has been operating for two years. The governance team conducts a scheduled comprehensive assessment and discovers that the model's performance remains strong overall but that its false negative rate (missed fraud) has increased by 40% for a specific transaction type — peer-to-peer mobile payments — that has grown significantly since the model was trained. The overall false negative rate across all transaction types remains within the original threshold. What does this finding reveal about the adequacy of the monitoring framework?

A. The monitoring framework is adequate because the overall false negative rate remains within thresholds, and the increase for a single transaction type does not warrant governance action

B. The finding reveals a data quality issue that should be addressed by the data engineering team without involving the governance function

C. The finding reveals that the model needs complete retraining on all transaction types to restore its performance to original levels across the entire portfolio

D. The monitoring framework's reliance on aggregate metrics masked a significant subgroup performance degradation — the framework should be enhanced with transaction-type-disaggregated monitoring to detect category-specific deterioration that aggregate metrics cannot reveal

11. Under GDPR, an organization relies on "legitimate interests" as the lawful basis for processing personal data to train an AI fraud detection model. A customer objects to the processing of their data for AI training under Article 21. The organization has strong evidence that the fraud detection model provides significant security benefits. Can the organization continue processing the objecting customer's data?

A. The organization may continue processing only if it can demonstrate compelling legitimate grounds for the processing that override the customer's interests, rights, and freedoms — the burden shifts to the organization upon receiving the objection

B. The organization must immediately cease all processing of the customer's data, including removing their data from the training set and retraining the model, with no exceptions

C. The customer's objection is invalid because fraud detection is a legal obligation that automatically overrides any individual's right to object to data processing

D. The organization may continue processing for up to 12 months after receiving the objection to allow sufficient time for the model to be retrained without the customer's data

12. An AI vendor provides a model card stating that its facial recognition system has an overall accuracy of 99.2% and has been "tested for fairness." A deployer's governance team requests the vendor's complete fairness testing methodology and results. The vendor responds that the fairness testing details are "proprietary." What should the deployer do?

A. Accept the vendor's claim because the model card's statement that the system was tested for fairness provides sufficient assurance for regulatory compliance purposes

B. Inform the vendor that without transparent fairness testing methodology and disaggregated results, the deployer cannot verify the system meets its governance and regulatory requirements — and consider whether the governance gap is acceptable or whether an alternative vendor should be evaluated

C. Deploy the system and conduct the deployer's own fairness testing after deployment to verify the vendor's claims through real-world operational data

D. File a complaint with the national competent authority because AI vendors are legally required to disclose all proprietary testing methodologies to deployers

13. An organization has deployed three AI systems: (1) a customer service chatbot, (2) an AI-assisted medical diagnostic tool, and (3) an internal IT ticket routing system. The governance team has resources to conduct a comprehensive annual audit of only one system. Using risk-based prioritization, which system should be audited FIRST?

A. The customer service chatbot, because it directly interacts with the largest number of external individuals and any errors could damage the organization's reputation

B. The internal IT ticket routing system, because internal systems are often overlooked and may accumulate governance debt that creates hidden organizational risk

C. All three systems should receive equal audit attention regardless of risk level because consistent governance standards require uniform treatment of all AI systems

D. The medical diagnostic tool, because it poses the highest risk to individuals' health and safety, is likely classified as high-risk under the EU AI Act, and errors could result in patient harm

14. A governance committee is reviewing an AI system's model card before deployment approval. The model card lists the system's intended use as "credit scoring for consumer lending" and identifies the system's known limitation: "Accuracy decreases for applicants with thin credit files (fewer than 3 credit accounts)." The committee notes that the deployment context includes a significant population of young first-time borrowers who typically have thin credit files. What is the MOST appropriate governance response?

A. Approve deployment because the model card honestly discloses the limitation and deployers are responsible for managing known limitations in their operational context

B. Reject deployment and return the system to the vendor because any known limitation makes the system unsuitable for production use in a high-risk credit scoring context

C. Approve deployment with the condition that applications from first-time borrowers with thin credit files are routed to manual underwriting review, and that the limitation's impact on this population is monitored continuously

D. Approve deployment but only after the vendor eliminates the thin-credit-file limitation, as high-risk systems must perform equally across all applicant subpopulations without exception

15. An AI system trained to predict employee attrition uses features including commute distance, salary relative to market rate, and years of tenure. Analysis reveals the model assigns significant predictive weight to "number of sick days taken in the past 12 months." A governance professional raises a concern about this feature. What is the governance basis for this concern?

A. Sick day usage may correlate with disability status, chronic health conditions, pregnancy, or caregiving responsibilities — using it as a predictive feature for attrition may create disparate impact against employees in protected categories, and could also constitute processing of health-related data without an Article 9 exception

B. Sick day data is always inaccurate because employees may take sick days for reasons unrelated to illness, making it an unreliable predictor of attrition

C. The number of sick days is a protected characteristic under employment discrimination law and cannot be used as a feature in any employment-related AI system

D. Sick day usage is too strongly correlated with other features in the model, creating multicollinearity that degrades the model's statistical reliability

16. Under the NIST AI RMF, the Govern function includes establishing workforce diversity, equity, inclusion, and accessibility processes in the AI lifecycle. A colleague argues this is a human resources concern unrelated to AI risk management. What is the MOST accurate response?

A. The colleague is correct because workforce diversity is an HR function that is separate from technical AI risk management and should not be included in the AI governance framework

B. Workforce diversity in AI teams brings diverse perspectives that improve risk identification, reduce blind spots in bias detection, and produce more robust governance outcomes — making it a legitimate and important component of AI risk management, not merely an HR initiative

C. The NIST AI RMF mentions workforce diversity only as an aspirational goal and does not include it as an actionable component of the Govern function

D. Workforce diversity is important only for AI development teams and has no relevance for governance committees, audit functions, or oversight roles

17. A deployer of a high-risk AI system receives the provider's instructions for use, which specify that the system must be monitored weekly for performance drift and that human overseers must complete a

specific training program. The deployer implements daily monitoring (exceeding the provider's specification) but does not require the specified training for human overseers. Is the deployer compliant with its EU AI Act obligations?

A. Yes, because the deployer has exceeded the monitoring specification, which compensates for the training gap and demonstrates a strong overall commitment to responsible governance

B. Yes, because the EU AI Act requires deployers to follow the provider's instructions only for system operation, not for personnel training, which is an internal organizational matter

C. No, because the EU AI Act requires deployers to use high-risk AI systems in accordance with the instructions for use — exceeding one requirement does not compensate for failing to implement another, and the training specification is part of those instructions

D. No, but only because daily monitoring exceeds the provider's weekly specification and constitutes an unauthorized deviation from the instructions for use

18. An organization uses an AI system to generate synthetic training data for another AI system. The synthetic data is designed to augment an underrepresented demographic group in the original training set. Testing reveals that the synthetic data generation model has produced synthetic samples that closely resemble specific real individuals from the underrepresented group whose data was used to train the generator. What governance concern does this raise?

A. The concern is purely technical — the synthetic data generator is overfitting and needs to be retrained with stronger regularization to prevent memorization of individual training examples

B. The synthetic data is of higher quality than expected because it closely resembles real individuals, which will improve the downstream model's performance for the underrepresented group

C. The concern is limited to intellectual property because the synthetic samples may be considered derivative works of the original photographs used to train the generator

D. The synthetic data may constitute personal data if specific individuals can be identified from the generated samples — raising privacy, consent, and data protection concerns even though the data was generated artificially rather than collected directly from those individuals

19. An organization deploys an AI system for automated contract review. The system identifies clauses that deviate from the organization's standard terms and flags them for legal review. After two years of

operation, lawyers notice that the system has never flagged a specific type of problematic clause — an indemnification cap limitation — because this clause type was not represented in the system's training data. All contracts reviewed by the system during this period contained unflagged indemnification cap issues. What is the governance implication?

A. The absence of a clause type from training data means the system has a known blind spot that has allowed potentially harmful contract terms to go undetected for two years — requiring retroactive review of all contracts processed during the period and immediate remediation of the training data gap

B. The issue is a minor training data deficiency that should be addressed in the next scheduled model retraining cycle without requiring any review of previously processed contracts

C. The legal team is responsible for catching issues the AI system misses, so the governance failure lies with the lawyers who should have independently reviewed every contract for indemnification cap clauses

D. The system's training data cannot be expected to cover every possible clause type, and the organization accepted this risk when it deployed the system for contract review

20. An AI system processes employment applications and produces a ranking of candidates. An applicant who was ranked low and not selected discovers the AI system was involved in the process and requests an explanation. The organization provides: "Your application was evaluated by our AI-assisted screening system, which considers multiple factors." Under GDPR transparency requirements, is this explanation adequate?

A. Yes, because the organization has disclosed the existence of automated processing and has indicated that multiple factors were considered, which satisfies the transparency requirement

B. No, because GDPR requires "meaningful information about the logic involved" — the explanation must identify the specific factors that influenced the ranking, how they were weighted, and what the applicant could do differently, not merely acknowledge that AI was used and multiple factors were considered

C. Yes, because the organization is not required to explain AI-driven employment decisions to unsuccessful applicants under GDPR since the final hiring decision was made by a human

D. No, but only because the organization did not provide the explanation within the required 72-hour timeframe specified by GDPR Article 22 for automated employment decisions

21. An insurance company uses an AI system to process claims. The system was originally designed to assess property damage claims only. Over time, the claims team begins using it to assess personal injury claims as well — a use case the system was never validated for. The system has no training data for personal injury assessment. When personal injury claims are submitted, the system processes them using property damage assessment logic because that is all it knows. What type of governance failure is this?

A. Concept drift, because the relationship between claim inputs and appropriate outputs has changed as the system encounters a new claim category

B. A monitoring failure, because the monitoring system should have detected the shift in claim type distribution and alerted governance

C. A vendor failure, because the AI provider should have implemented technical controls preventing the system from processing claim types outside its validated scope

D. Secondary use without governance review — the system is being applied to a purpose it was never designed, trained, or validated for, and the absence of validation data for personal injury means every assessment is unreliable

22. A large language model provider publishes a training data summary as required by the EU AI Act's GPAI provisions. The summary states: "The model was trained on a diverse dataset of text from the internet, books, and code repositories." A copyright holder whose works were likely included in the training data reviews the summary. What SPECIFIC deficiency makes this summary inadequate under the Act's requirements?

A. The summary is adequate because it identifies the three main categories of training data sources, which is the level of detail the EU AI Act requires

B. The summary fails because it does not include the total number of tokens in the training dataset, which is the only quantitative metric the EU AI Act requires for GPAI training data summaries

C. The summary lacks sufficient detail to enable rights holders to determine whether their works were included — it does not identify specific data sources, curation methodologies, temporal coverage, or mechanisms for rights holders to exercise their copyright reservations

D. The summary is deficient only because it mentions "books" without specifying whether the books were in the public domain or under active copyright protection

23. An organization operates an AI system that assists judges in sentencing decisions. The system provides a recommended sentence range based on case characteristics, criminal history, and sentencing guidelines. A defense attorney challenges the system, arguing that the judge cannot verify the AI's reasoning because the system is opaque. The prosecution argues the system is merely advisory and the judge makes the final decision. How should the governance framework address this dispute?

A. The advisory nature of the system does not eliminate the need for interpretability — if the AI's recommendation materially influences the sentence, the defendant's right to understand and challenge the basis for the sentence requires that the system's reasoning be sufficiently interpretable to support meaningful due process

B. The prosecution's argument is definitive because any system that provides recommendations rather than making final decisions falls outside the scope of due process requirements

C. The defense attorney's challenge should be dismissed because AI sentencing tools are classified as minimal risk under the EU AI Act and are not subject to interpretability requirements

D. The dispute should be resolved by removing the AI system from the sentencing process entirely because AI should never be used in any capacity within criminal justice proceedings

24. An AI governance professional discovers that the organization's data lineage tracking system has a gap — transformations applied during the feature engineering phase are not recorded. The data science team argues this gap is acceptable because the final training dataset is documented and the intermediate steps are "just data cleaning." What governance risk does this lineage gap create?

A. The gap creates no meaningful governance risk because documenting the final training dataset is sufficient for regulatory compliance and audit purposes

B. The lineage gap prevents the organization from tracing the origin of potential biases introduced during feature engineering, investigating incidents that may have been caused by data transformations, or reproducing the training dataset if the process needs to be replicated or audited

C. The gap only creates risk if the feature engineering process involved personal data, and has no governance implications for non-personal data transformations

D. The gap is acceptable for the current system but should be closed before the next model retraining cycle to ensure future compliance

25. An AI system used for resume screening was trained on data labeled by three different annotation teams. Team A labeled resumes from the technology sector, Team B labeled resumes from the healthcare sector, and Team C labeled resumes from the financial sector. A fairness audit reveals that the model exhibits bias against older applicants in the financial sector only. Investigation shows that Team C's labeling instructions emphasized "innovation potential" as a screening criterion, which the annotators interpreted as favoring younger candidates. What type of bias is this?

A. Representation bias, because older financial sector applicants were underrepresented in Team C's annotation sample relative to their prevalence in the actual applicant pool

B. Historical bias, because the financial sector has historically favored younger employees and Team C's labels accurately reflect this industry pattern

C. Measurement bias, because Team C's annotation instruments produced systematically different accuracy rates for older versus younger financial sector applicants

D. Label bias, because the ground truth labels applied by Team C's annotators were influenced by a subjective interpretation of "innovation potential" that systematically disadvantaged older applicants

26. An organization is implementing the NIST AI RMF and has reached the Measure function. The team must select fairness metrics for a credit scoring AI system. One team member advocates for demographic parity (equal approval rates across groups). Another advocates for equalized odds (equal true positive and false positive rates across groups). A third advocates for predictive parity (equal precision across groups). The governance professional must advise on the selection. What is the MOST important principle guiding this decision?

A. The team should implement all three metrics simultaneously because the EU AI Act requires that all high-risk systems satisfy every known fairness definition

B. The team should select the single metric that is easiest to compute and monitor, because practical implementability is more important than theoretical completeness

C. Multiple fairness definitions are mathematically incompatible in most real-world settings — the governance professional should advise selecting metrics based on the specific harms the system could cause, the regulatory requirements, and the values of the deployment context, while documenting why the chosen metrics are appropriate

D. The team should defer the fairness metric selection to the NIST AI RMF Playbook, which specifies the exact metric that must be used for each industry category

27. A technology company's AI system generates personalized news feeds for millions of users. The company's data science team discovers that the system's engagement optimization algorithm has gradually shifted the content mix toward sensationalist and emotionally provocative stories because these generate more clicks. The data science team reports this finding to the governance committee. The product team argues that the system is working exactly as designed — maximizing engagement — and no governance intervention is needed. How should the governance committee respond?

A. Agree with the product team because the system is meeting its design objective and engagement maximization is a legitimate business goal that governance should not constrain

B. Commission a study of the long-term effects of sensationalist content on user well-being before making any decision about whether to modify the system's optimization objective

C. Immediately redesign the system to show only neutral, factual content to eliminate any possibility that the recommendation algorithm could promote emotionally provocative material

D. The product team's position conflates technical function with governance appropriateness — a system can function as designed while producing harmful outcomes that governance must address, and the committee should evaluate whether the optimization objective should be modified to incorporate content quality and user welfare alongside engagement

28. An organization's AI system for employee scheduling uses historical data to predict optimal shift assignments. A governance review reveals that the system consistently assigns less desirable shifts (nights, weekends, holidays) to employees who have previously accepted such shifts without complaint, while employees who have formally objected to undesirable shifts receive more favorable schedules. What governance concern does this pattern raise?

A. The system is creating a punitive feedback loop — employees who accommodate the organization's needs are systematically disadvantaged, while those who complain receive better treatment, potentially penalizing cooperative behavior and creating a disparate impact on employees who may feel less empowered to object due to job insecurity, immigration status, or other vulnerabilities

B. The system is functioning optimally because assigning undesirable shifts to employees who are willing to work them maximizes organizational efficiency and minimizes scheduling conflicts

C. The concern is limited to employee satisfaction and does not rise to the level of a governance issue because shift scheduling is a routine operational decision

D. The pattern indicates a software bug in the scheduling algorithm that can be resolved through a technical fix without governance involvement

29. Under the EU AI Act, an organization places a high-risk AI system on the EU market. Two years later, the European Commission updates Annex III, adding a new use case category that encompasses the organization's system — a category that did not exist when the system was originally deployed. What is the organization's regulatory obligation?

A. The organization has no new obligations because the system was compliant when it was placed on the market and cannot be retroactively subjected to requirements that did not exist at deployment

B. The organization must immediately withdraw the system from the market until it can satisfy the new requirements because non-compliance with updated Annex III categories constitutes operating a prohibited AI practice

C. The organization must evaluate whether its existing governance controls satisfy the new requirements and implement any additional measures needed to bring the system into compliance with the updated classification

D. The organization is only required to register the system in the EU database under the new category but does not need to modify its governance practices or compliance documentation

30. An organization's monitoring dashboard for a deployed AI system shows the following: performance metrics stable, fairness metrics stable, data drift increasing steadily over three months, concept drift not detected. The governance team is debating whether to take action. What is the MOST appropriate governance response?

A. No action is needed because performance and fairness metrics are stable and data drift alone does not warrant governance intervention as long as outcomes remain acceptable

B. Immediately retrain the model because any detectable data drift indicates the model is operating outside its validated conditions and cannot produce reliable outputs

C. Shut down the system immediately because increasing data drift is a leading indicator of imminent system failure that requires emergency containment measures

D. Investigate the cause and trajectory of the data drift, increase monitoring frequency, and prepare retraining plans — because data drift is a leading indicator that may precede performance and fairness degradation, even though these have not yet been detected

31. An AI system used for credit scoring produces an adverse decision for an applicant. The applicant exercises their GDPR Article 22 right to request human intervention. The bank assigns a human reviewer who has access to the AI system's output but not to the underlying factors or the applicant's full profile. The reviewer confirms the AI's decision within 30 seconds. Has the bank satisfied the Article 22 requirement?

A. Yes, because a human reviewed the decision and the speed of the review is irrelevant as long as a natural person was involved in the final determination

B. No, because meaningful human intervention requires the reviewer to have access to relevant information, the competence to evaluate the decision, and sufficient time and authority to reach an independent conclusion — a 30-second rubber stamp without access to underlying factors is not genuine intervention

C. Yes, because Article 22 only requires that a human be "assigned" to the review, not that the human conduct any specific level of analysis or have access to any particular information

D. No, but only because the review took less than the minimum 5-minute review period that GDPR Article 22 specifies for automated credit decisions

32. An organization is preparing for an external audit of its AI governance practices under ISO/IEC 42001. The audit team requests evidence of the organization's AI impact assessment methodology. The organization presents a DPIA that was conducted under GDPR Article 35. Is this sufficient to demonstrate compliance with ISO 42001's impact assessment requirements?

A. A GDPR DPIA addresses only data protection risks and does not cover the full scope of AI impacts — including safety, fairness, societal effects, and organizational risks — that ISO 42001's impact assessment requirements encompass, so it is necessary but not sufficient evidence

B. Yes, because the GDPR DPIA is a more rigorous assessment than anything ISO 42001 requires, and compliance with GDPR automatically satisfies the ISO standard's requirements

C. No, because ISO 42001 does not accept any assessment that was conducted under a different regulatory framework as evidence of compliance with its own requirements

D. Yes, but only if the DPIA was conducted by an external auditor rather than the organization's own data protection officer, which provides the independence ISO 42001 requires

33. An AI governance committee is evaluating a vendor's proposal for an AI system. The vendor claims the system is "explainable" because it uses a decision tree model rather than a neural network. The governance professional is skeptical of this claim. What is the basis for the skepticism?

A. Decision trees are never explainable because their branching structure is too complex for humans to interpret regardless of the tree's size or depth

B. The vendor's claim is correct — all decision tree models are inherently interpretable, and the governance professional's skepticism is unfounded

C. The governance professional should be skeptical because the vendor did not provide the decision tree's source code, and without code review, no claims about explainability can be verified

D. A decision tree's interpretability depends on its size and complexity — a small, shallow tree may be interpretable, but a deep ensemble of thousands of trees (such as a random forest or gradient boosting model) can be as opaque as a neural network, making the claim "decision tree equals explainable" potentially misleading

34. An organization is building an AI system and the development team proposes skipping the formal impact assessment because the system is classified as limited risk rather than high risk. The governance professional pushes back. On what basis can the governance professional justify requiring an impact assessment for a limited-risk system?

A. The governance professional cannot justify this requirement because the EU AI Act only mandates impact assessments for high-risk systems, and imposing additional requirements on limited-risk systems violates the principle of proportionate governance

B. The governance professional can argue that external regulatory changes might reclassify the system as high-risk in the future, and conducting the assessment now avoids future compliance costs

C. The governance professional can cite the EU AI Act's Article 35, which requires impact assessments for all AI systems regardless of risk classification

D. The governance professional can justify the requirement based on the organization's internal governance policy, which may require impact assessments beyond what the EU AI Act mandates — proportionate governance means calibrating assessment depth to risk level, not eliminating assessment entirely for lower-risk systems

35. An AI system for hiring was tested for bias using a dataset representative of the U.S. workforce. The organization now plans to deploy it in Brazil. The governance team questions whether the U.S. bias testing is sufficient. The development team argues that the fairness metrics were satisfied and the system is therefore fair globally. What is the flaw in this reasoning?

A. The reasoning is correct because fairness metrics are mathematically universal and produce the same results regardless of the population on which they are measured

B. Fairness is context-dependent — the demographic composition, protected characteristics, historical discrimination patterns, cultural norms, and legal frameworks differ between the U.S. and Brazil, meaning that a system validated as fair for one population is not necessarily fair for another, and separate evaluation is required

C. The flaw is that the U.S. bias testing dataset was too small and should have included at least 1 million records to achieve statistical validity for global deployment

D. The reasoning is flawed only because Brazil has stricter bias testing requirements than the United States, requiring a minimum of 50 fairness metrics rather than the standard 10

36. An organization's AI governance committee is reviewing the results of a red teaming exercise on a customer-facing generative AI system. The red team identified 23 successful jailbreak techniques and 8 methods for extracting portions of the system prompt. The development team has addressed 15 of the 23 jailbreaks and 5 of the 8 extraction methods but argues that the remaining vulnerabilities are edge cases that require "sophisticated technical knowledge" to exploit. Should the committee approve deployment?

A. The committee should require remediation of all identified vulnerabilities before deployment because the sophistication of the current exploit does not predict the future accessibility of the technique, and deploying a system with known safety control bypasses creates unacceptable risk for a customer-facing application

B. The committee should approve deployment because addressing 65% of jailbreaks and 63% of extraction methods demonstrates a strong security posture that exceeds industry standards

C. The committee should approve deployment with the condition that a follow-up red teaming exercise is scheduled for six months post-deployment to reassess the remaining vulnerabilities

D. The committee should defer the decision to the AI vendor's security team because the vendor has the technical expertise to determine whether the unresolved vulnerabilities pose a material risk

37. An organization processes health insurance claims using an AI system. The system was trained on claims data that included diagnosis codes, treatment codes, and claim amounts. A governance review reveals that the system has learned to associate certain diagnosis codes with claim fraud — not because those diagnoses are actually associated with fraud, but because the training data reflects a historical pattern where claims from clinics serving low-income populations were disproportionately investigated and labeled as fraudulent. This is an example of which TWO types of bias working together?

A. Representation bias and measurement bias, because the training data underrepresents claims from wealthy populations and measures fraud differently across income groups

B. Data drift and concept drift, because the historical fraud patterns no longer reflect current fraud trends and the relationship between diagnoses and fraud has changed

C. Label bias and historical bias — the labels ("fraudulent") reflect biased investigation patterns (label bias), and those investigation patterns themselves reflect historical discrimination against low-income communities (historical bias)

D. Selection bias and measurement bias, because the training data excludes certain claim types and measures fraud probability differently for different diagnosis categories

38. A technology company releases a new version of its AI model that significantly outperforms the previous version on standard benchmarks. Several organizations that deploy the model through APIs automatically receive the updated version. One deployer discovers that while overall performance improved, the new version performs worse than the previous version for a specific use case that is critical to their business. What governance principle has been violated?

A. The vendor violated the transparency principle by not disclosing the specific performance changes in the updated model compared to the previous version

B. The deployer violated its own governance obligations by not implementing controls to evaluate vendor updates before they are applied in production

C. The vendor violated record-keeping requirements by not maintaining the previous model version alongside the new one in the EU database for high-risk AI systems

D. Both the vendor and deployer bear responsibility — the vendor should have provided advance notice and detailed change documentation, and the deployer should have maintained controls to evaluate updates before automatic application, preventing unauthorized changes to the production system's behavior

39. An organization uses an AI model to predict customer lifetime value (CLV) and allocates marketing resources accordingly. High-CLV customers receive premium service, exclusive offers, and dedicated account managers. Low-CLV customers receive standard service. Analysis reveals that CLV predictions strongly correlate with customer income and wealth indicators, meaning the system effectively provides better service to wealthier customers and worse service to less affluent ones. A governance professional raises concerns. The business team argues this is standard commercial practice. How should the governance committee evaluate this situation?

A. The business team is correct because differential service based on commercial value is a standard and legal business practice that does not raise governance concerns

B. The governance committee should immediately mandate that all customers receive identical service levels regardless of their predicted lifetime value to eliminate any service differentials

C. The governance concern is limited to data privacy because the CLV model processes financial data that may require an additional lawful basis under GDPR

D. The committee should evaluate whether the CLV-based service differentiation creates unfair outcomes — particularly if the income correlation means that service quality effectively tracks protected characteristics (race, national origin) through proxy effects, and whether the differential treatment constitutes unfair commercial practice under consumer protection principles

40. An AI system generates automated medical reports that summarize diagnostic imaging findings. A radiologist reviews and signs each report before it is sent to the referring physician. The AI system misidentifies a benign growth as potentially malignant, and the radiologist signs the report without independently reviewing the images. The patient undergoes an unnecessary surgical procedure. Who bears PRIMARY governance responsibility for this outcome?

A. The radiologist bears primary responsibility because the human oversight mechanism was in place but was not meaningfully exercised — the radiologist signed a report without independent review, rendering the human-in-the-loop safeguard ineffective through automation bias

B. The AI vendor bears primary responsibility because the system produced an incorrect classification that directly caused the downstream harm to the patient

C. The hospital bears primary responsibility because it deployed the AI system and failed to ensure that its radiologists had adequate training on proper use of AI-assisted diagnostic tools

D. The referring physician bears primary responsibility because they ordered the surgical procedure based on the report without seeking a second opinion from another radiologist

41. An organization is developing an AI governance maturity model to assess its progress. The model defines five maturity levels: (1) Ad Hoc, (2) Developing, (3) Defined, (4) Managed, and (5) Optimizing. The organization currently has documented policies, assigned roles, and completed training, but governance processes are not consistently followed across business units and there is no systematic measurement of governance effectiveness. At what maturity level is this organization MOST accurately classified?

A. Level 4 — Managed, because the organization has policies, roles, and training in place, which represents an advanced stage of governance maturity

B. Level 1 — Ad Hoc, because inconsistent process adherence across business units indicates that no real governance program exists beyond documentation

C. Level 5 — Optimizing, because the presence of documented policies and completed training demonstrates that the governance program has reached its highest maturity state

D. Level 3 — Defined, because governance structures are documented and established but not yet consistently implemented or measured across the organization, which is the hallmark of a defined but not yet managed governance program

42. A governance professional is reviewing an AI system's test results before deployment. The system was tested on a held-out test set and achieved strong results. However, the professional notices that the test set was created by random sampling from the same data distribution as the training set. The professional is concerned. For which type of AI system is this testing methodology MOST problematic from a governance perspective?

A. A medical imaging AI system, because patient populations in deployment may differ from the research population used for training, and random sampling within the same distribution does not test the model's performance on underrepresented populations or deployment-specific conditions

B. A spam email filter, because spam techniques evolve rapidly and the test set must include future spam patterns that the model has never encountered to be valid

C. A random sampling test set is the gold standard methodology for all AI system types and the governance professional's concern is unfounded regardless of the application domain

D. A weather prediction model, because atmospheric conditions change continuously and test data from the same distribution as training data cannot evaluate the model's ability to predict future weather patterns

43. An AI system for processing government benefit applications uses a risk scoring model to flag applications for additional verification. The system's design includes a feature that weights "consistency of information across data sources" heavily. Applicants who live in areas with less digitized government records — disproportionately rural and indigenous communities — naturally score lower on consistency simply because fewer data sources are available to cross-reference. What governance concept BEST describes this issue?

A. A deliberate design choice to exclude rural communities from government benefits in order to reduce processing costs and administrative burden

B. A measurement bias embedded in the system's design — the "consistency" feature systematically disadvantages applicants from communities with less digitized infrastructure, not because their information is less consistent but because fewer data sources exist to verify it

C. An acceptable limitation because the system correctly identifies that unverifiable applications represent higher risk regardless of the reason for the lack of verification data

D. A data quality issue that can be resolved by requiring all government agencies to digitize their records before the AI system is deployed in their jurisdiction

44. An AI vendor provides comprehensive technical documentation for its high-risk AI system, including model cards, test results, and a conformity declaration. A deployer reviews the documentation and identifies that the vendor tested the system using only English-language data, but the deployer plans to use it in a multilingual environment serving speakers of seven languages. What is the MOST appropriate governance response?

A. The deployer should conduct its own testing across all seven languages before deployment, because the vendor's English-only validation does not verify the system's performance, fairness, or safety characteristics for the other six languages in the deployer's operational context

B. The deployer should request that the vendor extend its testing to all seven languages before deployment, and should not deploy until the vendor provides multilingual validation results

C. The deployer may proceed with deployment because the vendor's documentation demonstrates compliance and the deployer has no independent testing obligation under the EU AI Act

D. The deployer should deploy the system for English-language use only and implement a separate non-AI process for the other six languages until the vendor provides multilingual testing

45. An organization's AI ethics board is evaluating a proposal to develop a "deepfake detection" AI system that analyzes video to determine whether it was generated or manipulated by AI. The system would be marketed to news organizations, social media platforms, and election integrity organizations. A board member raises a concern that the system itself could be used to develop better deepfakes by identifying what the detector looks for. What governance concept does this concern illustrate?

A. A standard software quality assurance concern about whether the detection system's accuracy will degrade over time as deepfake technology evolves

B. A marketing concern about whether news organizations will be willing to pay for a detection tool that may become obsolete as generation technology advances

C. An intellectual property concern about whether the detection methodology could be reverse-engineered by competitors offering similar deepfake detection products

D. The dual-use nature of AI technology — the detection system's methodology, if understood by adversaries, could be used to develop deepfakes specifically designed to evade detection, creating an arms race dynamic that governance must anticipate and address

46. A financial services company deploys an AI system that generates personalized investment portfolio recommendations. The system's training data is heavily weighted toward bull market conditions (2010-2021) with limited representation of bear markets and recessions. A governance professional identifies this as a risk. The data science team argues the model has been "stress tested" by simulating adverse conditions. What is the key difference between historical data representation and stress testing that the governance professional should emphasize?

A. There is no meaningful difference because stress testing simulates the same conditions that historical data would represent, making the two approaches equivalent for governance purposes

B. Stress testing is always superior to historical data because it can simulate conditions more extreme than any that have actually occurred in financial markets

C. Stress testing based on simulated adverse conditions may not capture the complex, non-linear interactions and cascading effects that characterize actual bear markets and recessions — real-world economic crises involve behavioral patterns, market dynamics, and systemic effects that are difficult to simulate accurately from bull market data alone

D. The difference is purely academic because investment AI systems are classified as minimal risk under the EU AI Act and are not subject to data governance requirements regarding training data composition

47. An organization has recently completed the deployment of a high-risk AI system. During the first week of operation, the monitoring system generates an alert indicating a minor deviation from expected performance metrics. The deviation is within the warning threshold but has not reached the critical threshold. The operations team dismisses the alert as "noise." Two weeks later, the deviation has worsened and crossed the critical threshold, resulting in unfair outcomes for approximately 500 individuals. What governance lesson does this scenario teach?

A. Warning threshold alerts should be investigated promptly — they exist specifically to provide early detection of developing problems while intervention is still possible and before critical thresholds are crossed, and dismissing them as noise defeats the purpose of a tiered alerting system

B. The monitoring system's warning threshold was set too sensitively and should be raised to reduce false positive alerts that create alert fatigue for the operations team

C. The operations team acted appropriately by waiting for the critical threshold to be reached, because warning alerts are informational notifications that do not require investigation or action

D. The monitoring system should be redesigned to eliminate warning thresholds entirely and only generate critical alerts that require immediate action to prevent alert fatigue

48. An AI governance committee is evaluating whether an organization's AI system for automated resume screening triggers the GDPR Article 22 protections for solely automated decision-making. The system produces a ranked list of candidates, and a recruiter reviews the top 20 candidates before scheduling interviews. Candidates ranked below position 20 are automatically rejected without human review. For the automatically rejected candidates, does Article 22 apply?

A. No, because the system only ranks candidates and does not make the final hiring decision, which is made by the human recruiter who selects from the top 20

B. Yes, because candidates ranked below position 20 are rejected through a solely automated process with no meaningful human involvement, and the rejection produces significant effects on their employment prospects

C. No, because GDPR Article 22 applies only to decisions with legal effects, and employment screening does not produce legal effects as defined by the regulation

D. Yes, but only if the screening system uses more than five input features, because simpler models with fewer features are exempt from Article 22's automated decision-making provisions

49. An organization is implementing an AI governance program from scratch. The governance professional must decide which governance activity to prioritize first. The organization currently has no AI inventory, no governance policies, no risk classification framework, and no monitoring capabilities. It operates 15 AI systems across various departments. What should be done FIRST?

A. Develop comprehensive governance policies that establish the organization's responsible AI principles, roles, and procedures before any other action is taken

B. Implement monitoring for all 15 AI systems simultaneously to establish a performance baseline before developing governance policies

C. Conduct an inventory of all AI systems — documenting what each system does, who it affects, what data it processes, and what risks it may pose — because governance cannot be prioritized or designed without first understanding what needs to be governed

D. Begin by developing a risk classification framework because risk-based prioritization is the foundation of efficient governance and should be established before examining any individual system

50. A healthcare AI system for patient triage assigns priority levels in an emergency department. The system was designed to prioritize based on clinical urgency alone. However, the training data was collected from a hospital that historically allocated more emergency department resources to privately insured patients, resulting in faster treatment and better-documented clinical presentations for this group. The AI system has learned to associate well-documented clinical presentations with higher urgency. What is the MOST accurate characterization of this issue?

A. An integration failure, because the AI system was not properly connected to the hospital's insurance verification system and cannot distinguish between insurance types

B. A deliberate design choice that appropriately prioritizes patients with more complete medical histories because more information enables more accurate clinical triage

C. A data quality issue that can be resolved by standardizing the format of all clinical presentation records in the emergency department

D. A historical bias in the training data — the resource allocation disparities based on insurance status produced systematically different documentation quality, which the AI system has learned to interpret as clinical urgency rather than as an artifact of inequitable historical resource allocation

51. An organization is evaluating the environmental impact of its AI operations. The governance team discovers that training their largest AI model consumed approximately the same amount of energy as 10 average U.S. households use in a year. The organization is developing a new, larger model that will require approximately 5 times more computational resources. Under responsible AI principles, how should the governance framework address this?

A. The governance framework should incorporate environmental impact assessment into AI development decisions — evaluating whether the expected benefits of the larger model justify the increased energy consumption, considering whether more efficient model architectures or training techniques could reduce the footprint, and documenting the environmental analysis as part of the impact assessment

B. Environmental impact falls outside the scope of AI governance and should be addressed by the organization's sustainability team without involving the governance function

C. The organization should cancel the larger model development entirely because AI's environmental footprint is unjustifiable regardless of the model's expected benefits or societal value

D. Environmental impact is only a governance concern when the AI system directly processes environmental data and is not relevant to the energy consumed during the development process

52. An organization uses an AI system to classify customer support tickets by priority. The system has been operating for three years. A new employee in the governance function discovers that the system has no model card, no documentation of its training data, no record of any testing, and no monitoring. The system was deployed before the organization established its AI governance program. The system performs adequately and no incidents have been reported. What is the MOST appropriate governance response?

A. Leave the system in place without documentation because it predates the governance program and retroactive governance is impractical for legacy systems

B. Bring the system into governance compliance by documenting its current state, conducting a risk assessment, performing retroactive testing where possible, establishing monitoring, and creating the missing governance artifacts — with priority determined by the system's risk classification

C. Immediately shut down the system because operating any AI system without documentation, testing records, or monitoring constitutes an unacceptable governance violation

D. Schedule the system for decommissioning at the end of its current lifecycle and apply governance requirements only to its replacement

53. An AI system for detecting potential money laundering analyzes transaction patterns across thousands of bank accounts. The system identifies a pattern it has never encountered before — a series of transactions that do not match any known laundering typology in its training data but that the system flags as anomalous based on statistical deviation from normal patterns. A human investigator reviews the flag and cannot determine whether the transactions represent legitimate business activity or a novel laundering technique. What governance challenge does this scenario highlight?

A. The system should be shut down because generating alerts for patterns that investigators cannot interpret demonstrates that the AI system is not sufficiently accurate for its intended purpose

B. The scenario highlights that investigators need additional training on money laundering techniques to improve their ability to evaluate AI-generated anomaly alerts effectively

C. The scenario demonstrates a monitoring failure because the system should not generate alerts unless they match a known laundering typology in its classification database

D. The scenario highlights the challenge of governing AI systems that detect novel patterns — the system may be discovering genuinely new threats, but the organization needs processes for investigating anomalies that fall outside both the AI's training and the human investigator's experience, including escalation procedures and specialized review capability

54. An organization is deploying an AI system that will be used by both trained professionals (in-house analysts) and untrained end users (customers accessing a self-service portal). The same AI model powers both interfaces. From a governance perspective, what is the MOST important design consideration for managing this dual-audience deployment?

A. The AI model should be split into two separate models — a more powerful version for professionals and a simplified version for end users — to prevent untrained users from accessing advanced analytical capabilities

B. No design differentiation is needed because the same AI model producing the same outputs for both audiences ensures consistency and eliminates the risk of differential treatment

C. The system should present AI outputs differently to each audience — professionals receive detailed analytical information with appropriate context for expert interpretation, while end users receive simplified explanations with stronger guardrails, clearer limitations disclosures, and more prominent pathways to human assistance

D. End users should be prohibited from accessing any AI-powered features because untrained individuals cannot be expected to interpret AI outputs correctly regardless of how the interface is designed

55. A large language model deployed as a customer service chatbot occasionally generates responses that include specific medical advice, such as recommending specific medications for described symptoms. The system was not designed for medical use and has no medical validation. The organization's terms of service state that the chatbot "does not provide medical advice." Despite the disclaimer, some customers follow the chatbot's medical recommendations. What is the PRIMARY governance concern?

A. The primary concern is that the system is generating outputs outside its intended purpose that could cause physical harm to individuals who act on them — a terms-of-service disclaimer does not eliminate the organization's responsibility to prevent foreseeable harm from AI-generated content, and technical controls should prevent the system from generating medical advice

B. The primary concern is that the terms of service should be reformatted to display the disclaimer more prominently so that customers cannot reasonably claim they did not see it

C. The primary concern is that the chatbot's medical responses may be of variable quality, and the organization should validate the medical accuracy of these responses rather than trying to prevent them

D. The primary concern is limited to reputational risk because the terms-of-service disclaimer provides complete legal protection against any liability for harm caused by the chatbot's medical recommendations

56. An organization conducts an AI impact assessment for a new system and identifies a significant risk: the system may produce discriminatory outcomes for elderly users due to training data underrepresentation. The impact assessment recommends three mitigation measures. The governance committee approves deployment with all three mitigations in place. Six months later, an internal audit discovers that only one of the three recommended mitigations was actually implemented. The other two were deferred due to "resource constraints." What governance principle has been violated?

A. The transparency principle, because the governance committee was not informed that two of the three mitigations were deferred due to resource constraints

B. The accountability principle — the deployment was approved conditionally on all three mitigations being in place, and proceeding without implementing the approved conditions violates the governance

committee's authorization and creates an accountability gap between approved governance controls and actual operational practice

C. The fairness principle exclusively, because the incomplete mitigation may result in discriminatory outcomes for elderly users

D. No principle has been violated because resource constraints are a legitimate business reason for deferring mitigation measures and the governance committee would have approved the deferral if consulted

57. An organization receives notice from a data protection authority that it is under investigation for potential GDPR violations related to its AI processing activities. The authority requests the organization's technical documentation, impact assessments, and monitoring records for a specific AI system. The organization's legal team discovers that monitoring records for the past six months are incomplete due to a logging infrastructure failure. How does this gap affect the organization's regulatory position?

A. The gap has no regulatory significance because monitoring records are supplementary documents that are not subject to GDPR's record-keeping requirements

B. The gap only matters if the data protection authority specifically requests the missing monitoring records, and the organization is not obligated to disclose the gap proactively

C. The organization should generate synthetic monitoring records to fill the gap, as this is standard practice when infrastructure failures create documentation interruptions

D. The gap significantly weakens the organization's regulatory position because it demonstrates a period during which the organization cannot prove it was monitoring the AI system's compliance, fairness, and performance — undermining the accountability principle and potentially constituting a record-keeping violation

58. A technology company operates an AI-powered hiring platform used by thousands of employers. The company discovers a bias in its algorithm that has affected hiring decisions across multiple client organizations for the past nine months. The company must decide who to notify and in what order. Under the EU AI Act and general governance principles, what is the CORRECT notification sequence?

A. Notify the company's shareholders first, then the media, then affected employers, and finally the regulatory authority

B. Notify the company's legal team only and keep the discovery confidential to avoid reputational damage while the bias is remediated

C. Notify only the affected employers and allow them to decide independently whether to notify their applicants or regulatory authorities

D. Notify the relevant market surveillance authority about the serious incident, notify all affected deployer organizations so they can take their own protective measures, and coordinate on notification of affected applicants — with the regulatory notification and deployer notification occurring promptly upon establishing the causal link

59. An AI governance professional is evaluating an organization's approach to managing AI systems from multiple vendors. The organization currently conducts vendor assessments at the time of procurement but does not reassess vendors after the initial contract is signed. The professional recommends implementing ongoing vendor monitoring. What is the STRONGEST justification for this recommendation?

A. Vendor governance practices, security posture, financial stability, model performance, and regulatory compliance can all change after the initial procurement — ongoing monitoring ensures the deployer detects vendor-side changes that may affect the AI system's governance profile before they cause harm in the deployment context

B. Ongoing vendor monitoring is required by the EU AI Act for all deployers of vendor-sourced AI systems regardless of risk classification

C. The recommendation is primarily about maintaining the organization's negotiating leverage with vendors rather than addressing any specific governance risk

D. Ongoing vendor monitoring is only necessary if the vendor is located in a different jurisdiction from the deployer because cross-border vendor relationships are inherently riskier

60. A national examination board uses an AI essay grading system for high-stakes university entrance exams. The system was validated on essays from the current year's applicant pool. A governance professional notes that the validation data was drawn from the same population as the deployment data and questions whether this provides adequate assurance. Specifically, the professional is concerned that the validation does not test the system's ability to fairly grade essays from students with non-traditional educational backgrounds, students with learning disabilities who may have unconventional writing styles, and students whose first language is not the exam language. What type of validation gap is the professional identifying?

A. The professional is identifying a gap in technical infrastructure testing, because the grading system may not process exam submissions quickly enough for the volume of applicants

B. The professional is identifying a gap in subpopulation validation — the system may perform differently for specific groups whose characteristics may be underrepresented in the validation data even though they are part of the overall applicant population, and their educational futures depend on the system grading them fairly

C. The professional is identifying a gap in security testing, because students with technical expertise might find ways to game the AI grading system by mimicking patterns that earn high scores

D. The professional's concern is unfounded because validating on the actual applicant population is the most rigorous testing methodology available and addresses all fairness concerns by definition

61. An organization deploys an AI chatbot for customer service. The system includes a feature that detects when a customer appears to be in emotional distress and routes them to a human agent. During a system update, this detection feature is accidentally disabled. For the next three weeks, customers in emotional distress — including those expressing suicidal thoughts — receive automated responses instead of human escalation. What governance control should have prevented this?

A. The organization should have had a policy prohibiting all AI system updates to prevent the risk of accidentally disabling safety features during the modification process

B. The organization should have employed more experienced software developers who would not make such errors during the update implementation

C. The organization should have used a more expensive AI platform with built-in safety features that cannot be modified under any circumstances by the deploying organization

D. A change management process with pre-deployment testing of safety-critical features — verifying that the emotional distress detection and escalation pathway functions correctly after every system change — should have caught the disabled feature before it affected customers

62. An AI system for predictive maintenance in a manufacturing plant monitors equipment vibration, temperature, and pressure sensors. The system recommends maintenance actions: "no action needed," "schedule routine maintenance," or "immediate shutdown required." A maintenance supervisor observes that the system has recommended "schedule routine maintenance" for a piece of equipment that the supervisor, based on 20 years of experience, believes is about to fail catastrophically. The system's confidence in its "routine maintenance" recommendation is 94%. What should the supervisor do?

A. Follow the AI system's recommendation because its 94% confidence level indicates high reliability, and overriding the system without data-driven justification undermines the purpose of deploying AI for predictive maintenance

B. Escalate the disagreement to the plant manager for a final decision because the supervisor does not have the authority to override AI-generated maintenance recommendations

C. Override the AI recommendation and initiate immediate inspection, documenting the decision and clinical reasoning — human oversight means experienced professionals retain authority to exercise judgment that considers factors the AI may not have observed

D. Submit a formal request to the AI vendor to investigate the discrepancy between the supervisor's assessment and the AI's recommendation before taking any maintenance action

63. An organization receives two data subject access requests related to its AI systems. Request A is from a customer asking what personal data the organization holds about them in its AI system's training data. Request B is from a customer asking for an explanation of why the AI system denied their loan application. Which request is MOST challenging to fulfill from a governance perspective, and why?

A. Request B is more challenging because the organization must provide "meaningful information about the logic involved" in the denial decision, which requires the AI system to support some form of explainability — a capability that may not have been built into the system's architecture

B. Request A is more challenging because the organization must provide an exact copy of every data point used to train the model, which requires reconstructing the complete training dataset

C. Both requests are equally straightforward because GDPR provides clear, step-by-step procedures for responding to access requests and explanation requests

D. Neither request is challenging because AI systems are exempt from data subject access rights under GDPR Article 15 when the data has been used for machine learning purposes

64. An organization is developing a governance framework for its use of third-party AI APIs. The organization sends customer queries to an external AI API and receives responses that are delivered to customers. The API provider's terms state that it "may update the model at any time without notice." From a governance perspective, what is the MOST critical risk this creates?

A. Unnotified model updates mean the organization cannot evaluate whether updates affect performance, fairness, compliance, or safety in its deployment context — the system's behavior can change at any time without the deployer's knowledge or governance review, potentially degrading the system in ways that violate the organization's regulatory obligations

B. The risk is limited to potential service disruptions during model updates and does not affect the governance profile of the AI system

C. The risk is acceptable because API providers have the technical expertise to ensure their updates improve system quality without creating governance issues for deployers

D. The risk is mitigated by the organization's monitoring system, which will detect any changes in the API's behavior after they occur and alert the governance team

65. An AI system for automated document classification in a government agency misclassifies a document containing classified national security information as "public," resulting in its publication on the agency's website. The document is available for 4 hours before the error is detected and the document is removed. What governance controls should have been in place to prevent this incident?

A. The agency should have prohibited the use of AI for document classification because government agencies handle sensitive information that requires exclusively human classification judgment

B. The agency should have required the AI vendor to achieve 100% classification accuracy before deployment, as any error rate is unacceptable for government document handling

C. The agency should have hired additional IT security staff to manually review the AI system's source code daily for classification errors in the algorithm's decision logic

D. Human review should have been required for documents classified as "public" before publication — a governance control that prevents misclassified sensitive documents from being exposed by ensuring a human verifies the AI's classification before public release, particularly for any action that cannot be reversed

66. An organization's AI system produces a recommendation that a human decision-maker follows. The recommendation turns out to be harmful. In the post-incident investigation, the decision-maker states: "The AI told me to do it." Under the principle of accountability, how should governance assign responsibility?

A. Responsibility lies entirely with the AI system's developer because the developer created the tool that produced the harmful recommendation

B. Responsibility lies entirely with the decision-maker because they chose to follow the recommendation and had the authority to exercise independent judgment

C. The governance framework should assign shared responsibility — the decision-maker is responsible for exercising independent judgment rather than blindly following AI recommendations, but the organization is also responsible for providing adequate training, explainability, and override authority to prevent over-reliance on AI outputs

D. No one bears responsibility because the interaction between human and AI decision-making creates an accountability gap that current governance frameworks cannot resolve

67. An AI system is trained to detect safety violations in factory environments using security camera footage. The system monitors whether workers wear required protective equipment. A worker's union representative raises a concern that the system could be used to discipline individual workers for safety violations rather than to improve overall safety conditions. The union argues this represents a different purpose than the system's stated safety monitoring objective. Is the union's concern legitimate from a governance perspective?

A. Yes, the union's concern is legitimate — using safety monitoring data to discipline individual workers represents a secondary use (enforcement against individuals) beyond the primary stated purpose (monitoring overall safety conditions), which requires separate governance review, potentially a separate lawful basis, and different impact assessment considerations

B. No, because safety compliance and individual accountability for safety violations are the same purpose and do not require separate governance consideration

C. The union's concern is a labor relations matter that has no connection to AI governance and should be addressed through collective bargaining rather than the governance framework

D. The concern is only legitimate if the system uses facial recognition to identify individual workers, and is not a governance issue if the system monitors equipment usage without identifying individuals

68. An AI governance professional is comparing two deployed AI systems to determine which poses greater governance risk. System A has high accuracy (99%) but low explainability — its decisions cannot be meaningfully explained to affected individuals. System B has moderate accuracy (92%) but high explainability — its decisions can be clearly explained with specific factor attribution. Both

systems are used for the same high-risk purpose. Which system poses GREATER governance risk, and why?

A. System B poses greater risk because its lower accuracy means more individuals will receive incorrect decisions, and accuracy is always the most important factor in governance risk assessment

B. System A and System B pose equal governance risk because their advantages and disadvantages offset each other perfectly in all governance evaluation frameworks

C. System B poses greater risk because explainability creates additional liability by providing specific reasons that affected individuals can challenge in legal proceedings

D. System A poses greater risk because its inability to explain decisions undermines affected individuals' rights to understand and challenge adverse outcomes, creates accountability gaps when errors occur, and may violate transparency and explainability requirements under the EU AI Act and GDPR — even though its accuracy is higher

69. An organization's AI system is designed to make real-time trading decisions in financial markets. The system operates at speeds far exceeding human capability — executing thousands of trades per second. The organization's governance framework requires human oversight of the system. How should human oversight be meaningfully implemented for a system that operates faster than any human can review individual decisions?

A. Human oversight is impossible for high-speed trading systems, so these systems should be exempt from human oversight requirements under the EU AI Act

B. Human oversight for real-time systems should focus on monitoring at an appropriate time scale — observing aggregate patterns, reviewing statistical summaries, setting and adjusting operational parameters and risk limits, and maintaining the authority to halt the system

C. The organization should require a human to approve every individual trade before execution, even if this slows the system to a speed that eliminates its competitive advantage

D. Human oversight should be implemented only through daily end-of-trading reviews where a compliance officer examines the system's complete transaction log from the previous trading session

70. An organization is planning to deactivate an AI system that has been providing automated benefit eligibility determinations for a government social services program. The system processes

approximately 5,000 applications per week. The organization plans to transition to manual processing. The governance team estimates that manual processing can handle approximately 1,000 applications per week. What governance concern does this transition capacity gap create?

A. The capacity gap is irrelevant because manual processing is always preferable to AI processing and any delay is justified by the transition to human decision-making

B. The capacity gap is a budget concern that should be addressed by the finance department rather than the governance function

C. The governance concern is limited to ensuring that the IT infrastructure is properly decommissioned and that server costs are eliminated from the operating budget

D. The capacity gap means that 4,000 applicants per week will experience significant delays in receiving eligibility determinations during the transition — the governance team must plan for this impact on affected individuals, potentially through phased deactivation, temporary additional staffing, or interim processing arrangements

71. An AI vendor presents benchmark results showing their model outperforms competitors on three standard industry benchmarks. A governance professional evaluating the vendor for a high-risk deployment is skeptical of relying on benchmarks alone. What is the governance basis for this skepticism?

A. Standard benchmarks may not reflect the specific conditions, populations, and requirements of the deployer's operational context — a model that excels on benchmarks may underperform for the deployer's specific use case, population, and data characteristics, making deployment-specific evaluation essential

B. Benchmarks are completely unreliable indicators of AI system quality and should never be considered in any vendor evaluation or governance assessment process

C. The skepticism is only justified if the vendor used different model versions for benchmarking than the version offered for commercial deployment

D. Governance professionals should rely exclusively on the vendor's self-reported performance metrics because vendors have the deepest knowledge of their own systems

72. An AI system generates automated summaries of patient medical records for physician review. The system occasionally omits clinically significant information — a medication allergy, a prior adverse reaction, or a chronic condition — from the summary while including less relevant details. What design principle should governance require to mitigate this risk?

A. The system should be designed to include every piece of information from the medical record in the summary to eliminate any possibility of omission

B. The system should implement conservative error handling for clinically significant categories — ensuring that critical safety information such as allergies, adverse reactions, and life-threatening conditions is always included in summaries even at the cost of producing longer, less concise outputs, with clear flagging of items requiring physician attention

C. The system should display a disclaimer noting that the summary may be incomplete and that physicians should always review the full medical record before making clinical decisions

D. The system should be replaced with a non-AI template-based extraction system that pulls fixed fields from the medical record without any machine learning interpretation

73. An organization uses an AI content generation system to produce marketing materials. An employee prompts the system to generate a product comparison that includes competitor products. The AI system generates content containing false claims about a competitor's product — stating that the competitor's product was recalled for safety defects when no such recall occurred. The marketing team publishes the content without verification. Under consumer protection and defamation law, what governance failure does this scenario expose?

A. The governance failure is limited to the AI vendor, who is responsible for ensuring that its language model does not generate false statements about any company or product

B. The governance failure lies only with the employee who generated the content and should be addressed through disciplinary action without organizational governance implications

C. The organization has no liability because AI-generated content that contains hallucinated false claims is protected as AI-generated speech under First Amendment principles

D. The governance failure is the absence of a verification process for AI-generated content before publication — the organization is responsible for the accuracy of content it publishes regardless of whether it was generated by a human or AI, and deploying generative AI for marketing without output verification creates foreseeable risk of publishing false and potentially defamatory claims

74. An organization is assessing the risks of deploying an agentic AI system that autonomously researches, drafts, and sends emails to business contacts on behalf of sales representatives. The system would access the CRM, compose personalized outreach messages, and send them without human review. Compared to a standard AI email drafting assistant where a human reviews and sends each message, what UNIQUE risk does the agentic deployment create?

A. The agentic system uses more computational resources than the standard assistant, creating higher infrastructure costs that the governance framework must evaluate

B. The agentic system requires a more user-friendly interface than the standard assistant because sales representatives must be able to monitor its autonomous actions effectively

C. The risk profiles are identical because both systems produce the same type of output (emails) and the method of production does not affect the governance assessment

D. The agentic system can send communications containing errors, inappropriate content, or inaccurate claims without human review, creating compounding risk — each autonomous action may build on previous actions' errors, and the absence of human intervention between steps eliminates the opportunity to catch mistakes before they reach external recipients

75. An organization is developing an AI system for a government client. The government contract specifies that the AI system must be "transparent and accountable." The development team interprets this as requiring comprehensive logging and documentation. The governance professional argues this interpretation is incomplete. What ADDITIONAL elements of transparency and accountability should the governance professional emphasize?

A. Transparency and accountability in government AI extend beyond logging and documentation to include public disclosure of the system's existence and purpose, meaningful explainability of decisions affecting individuals, mechanisms for affected individuals to challenge adverse outcomes, independent audit access, and clear assignment of responsibility for the system's governance throughout its lifecycle

B. Logging and documentation fully satisfy the transparency and accountability requirements because they create a complete record of the system's operations that can be reviewed if questions arise

C. The additional elements are limited to publishing the system's source code on a public repository so citizens can independently verify the system's decision-making logic

D. The governance professional's argument is incorrect because government contracts define the compliance requirements and the development team's interpretation of the contract language is authoritative

76. An AI system for predicting equipment failures in a power plant recommends immediate shutdown of a turbine. Shutting down the turbine will cause a regional power outage affecting 50,000 homes. The plant operator is unsure whether the AI's recommendation is correct. Ignoring the recommendation risks catastrophic equipment failure that could endanger workers. What does this scenario illustrate about the governance of AI in safety-critical environments?

A. AI systems should never be deployed in safety-critical environments because the consequences of both following and not following AI recommendations are unacceptable

B. The scenario illustrates the challenge of human oversight in high-stakes, time-pressured situations — governance frameworks must prepare operators for these decisions through clear protocols, escalation procedures, training on system limitations, and access to additional diagnostic information that enables informed judgment under pressure

C. The plant operator should always follow the AI recommendation in safety-critical situations because the AI system has been validated and its recommendation should not be questioned

D. The scenario illustrates that AI systems in safety-critical environments must achieve 100% accuracy before deployment because any uncertainty in recommendations creates unacceptable governance risk

77. An organization provides its AI system's model card to a potential customer. The model card includes disaggregated performance metrics showing that the system achieves 97% accuracy for male subjects and 89% accuracy for female subjects. The potential customer asks the organization to remove the disaggregated metrics from the model card because the customer believes the gender performance gap will make the product harder to sell. What should the organization do?

A. Remove the disaggregated metrics as requested because the customer's commercial concerns are legitimate and the organization should accommodate customer preferences

B. Replace the disaggregated metrics with a single aggregate accuracy figure that averages performance across genders, providing a more marketable representation of system quality

C. Offer to modify the model to achieve equal performance across genders before providing an updated model card without the performance gap

D. Refuse to remove the disaggregated metrics because model card transparency about performance disparities across demographic groups is a governance obligation that enables deployers to make informed decisions and is required for high-risk systems under the EU AI Act — hiding known performance gaps would be a transparency violation

78. An organization has deployed 20 AI systems across its operations. The governance team maintains individual governance records for each system but has not conducted any analysis of patterns, risks, or governance gaps across the portfolio as a whole. A governance professional recommends implementing portfolio-level governance analysis. What specific governance risk does portfolio-level analysis address that system-level governance does not?

A. Portfolio-level analysis is unnecessary because comprehensive system-level governance for each individual AI system provides complete governance coverage for the organization

B. Portfolio-level analysis is only relevant for organizations with more than 100 AI systems and is impractical for smaller portfolios

C. Portfolio-level analysis reveals systemic risks, shared vulnerabilities, and governance patterns that are invisible at the individual system level — such as concentration risk from common vendors, shared training data dependencies, correlated failure modes across systems, and cumulative impact on specific populations who may be affected by multiple AI systems simultaneously

D. Portfolio-level analysis is only useful for generating executive dashboard visualizations and does not provide actionable governance insights

79. An organization deploys an AI system that generates automated meeting summaries for corporate teams. The system records, transcribes, and summarizes meeting content. An employee raises a concern that the system captured and summarized a private conversation between two colleagues that occurred before the formal meeting began. The conversation included sensitive personal health information. What governance principles does this incident implicate?

A. The incident implicates only the principle of data security because the sensitive personal information should have been encrypted during processing

B. The incident implicates privacy (processing personal health data without clear consent and potentially without a lawful basis), data minimization (capturing and processing pre-meeting conversations that were not intended to be recorded exceeds the stated meeting-summary purpose), and transparency (participants may not have been adequately informed that the system records all audio including pre-meeting conversation)

C. The incident implicates only the principle of accuracy because the system should have correctly identified that the conversation occurred before the meeting and excluded it from the summary

D. The incident has no governance implications because employees in corporate environments have no reasonable expectation of privacy in meeting rooms where AI recording systems are deployed

80. An organization must decide whether to deploy an AI system that has completed all testing and governance review, or to delay deployment by three months to address a moderate-risk fairness concern identified in the impact assessment. Deploying now would provide significant business value and competitive advantage. Delaying would allow the fairness concern to be fully mitigated. What governance principle should guide this decision?

A. The decision should be based on a proportionality analysis that weighs the severity and scope of the fairness concern against the business value at stake, the affected population's vulnerability, whether interim mitigations can reduce the risk during the three-month period, and whether the organization can accept the identified residual risk with documented justification — there is no universal rule that moderate risks automatically block or automatically permit deployment

B. Business value and competitive advantage should always take priority over moderate governance concerns because governance must not impede commercial operations

C. Any identified fairness concern, regardless of severity, must be fully resolved before deployment because the responsible AI principle of fairness has absolute priority over all business considerations

D. The decision should be deferred to the AI vendor because the vendor has the technical expertise to determine whether the fairness concern warrants deployment delay

81. An organization conducts its first comprehensive AI governance audit and discovers that 40% of its AI systems were deployed without impact assessments, 60% lack adequate monitoring, and 25% have no documented model cards. The audit report is presented to the board of directors. A board member asks what single metric would best indicate whether the governance program is improving over time. What metric is MOST useful for tracking governance program maturity?

A. The total number of AI governance policies published by the organization, because more policies indicate a more mature governance program

B. The annual budget allocated to AI governance activities, because higher spending directly correlates with better governance outcomes

C. The percentage of deployed AI systems that have completed all required governance activities (impact assessment, documentation, testing, monitoring) — tracked over time, this metric shows whether the governance program is successfully closing the compliance gaps identified in the audit

D. The number of AI governance staff employed by the organization, because larger governance teams provide more comprehensive coverage of the organization's AI portfolio

82. An AI system deployed for insurance claim fraud detection flags a claim as potentially fraudulent. The claim involves a house fire that destroyed a family's home. Investigation reveals the claim is legitimate — the family lost everything in the fire. However, the AI system flagged the claim because the family had recently increased their insurance coverage, a pattern associated with arson fraud in the training data. The system's flag delayed the family's claim payment by three months during investigation. What governance improvement would MOST directly prevent this type of harm to legitimate claimants?

A. Removing the "recent coverage increase" feature from the model because it produces too many false positives

B. Increasing the system's confidence threshold for flagging claims so that fewer legitimate claims are incorrectly flagged for investigation

C. Implementing a disclosure to all policyholders that claims may be delayed by AI fraud detection systems during the processing period

D. Implementing a process that ensures AI fraud flags are investigated promptly with clear SLAs for investigation duration, that flagged claimants receive interim support while under investigation, and that the human investigation process can quickly clear legitimate claims rather than subjecting them to the same extended investigation timeline as genuinely suspicious ones

83. An AI governance professional is developing training materials for the organization's data science team. The training emphasizes the importance of documenting model limitations in model cards. A senior data scientist argues that disclosing limitations in model cards creates competitive disadvantage because competitors will highlight those limitations to win customers. What is the MOST effective response to this argument?

A. The data scientist is correct, and the organization should document limitations only in internal records while providing external model cards that emphasize strengths without disclosing weaknesses

B. Transparent disclosure of limitations in model cards is both a governance obligation and a competitive advantage — it builds trust with deployers, demonstrates governance maturity, enables deployers to implement appropriate safeguards, and reduces liability risk by ensuring that known limitations were disclosed rather than concealed

C. The organization should redact specific quantitative details about limitations while providing general qualitative descriptions that indicate limitations exist without giving competitors exploitable information

D. The concern can be resolved by classifying model cards as confidential business information and sharing them only under nondisclosure agreements that prevent deployers from sharing the content

84. An organization's AI system for medical appointment scheduling automatically distributes patients across available time slots. Analysis reveals that the system consistently schedules patients with complex medical conditions into shorter appointment slots because complex conditions are statistically associated with more frequent appointments in the training data. The algorithm learned that complex patients "come back often" and therefore assigned shorter slots, reasoning that they would return for follow-up. This has resulted in inadequate consultation time for the patients who need it most. What type of governance issue is this?

A. A misalignment between the AI system's learned optimization objective and the governance objective of equitable patient care — the system optimized for appointment volume distribution based on statistical patterns rather than for adequate clinical consultation time, producing an outcome that systematically disadvantages the most vulnerable patients

B. An acceptable scheduling optimization because the system correctly identified that complex patients have more frequent appointments, which is a legitimate factor in scheduling efficiency

C. A data quality issue that can be resolved by ensuring that appointment duration data in the training set accurately reflects the time actually spent with each patient during their visits

D. A monitoring failure because the scheduling system's performance metrics should have detected the shorter appointment slots assigned to complex patients before the pattern became established

85. An AI governance committee is evaluating a proposal to deploy an AI system that uses natural language processing to analyze employee communications for signs of workplace harassment. The system would scan emails, chat messages, and internal forum posts. Proponents argue it will create a safer workplace. Opponents argue it creates a surveillance infrastructure. What governance analysis is MOST critical before making this decision?

A. The committee should focus exclusively on the system's technical accuracy — if it can reliably detect harassment with a false positive rate below 1%, deployment should be approved regardless of other considerations

B. The committee should focus exclusively on legal compliance — if the system complies with all applicable privacy and employment laws, deployment should be approved regardless of ethical or cultural concerns

C. The committee must evaluate the proportionality of the surveillance mechanism against the safety benefit — considering whether less invasive approaches to addressing workplace harassment exist, whether the surveillance will chill legitimate communication, whether the power asymmetry between employer and employee undermines consent, and whether the system's deployment addresses root causes of harassment or merely monitors for symptoms

D. The committee should approve deployment but limit it to monitoring communications of employees who have been previously disciplined for harassment-related behavior

86. An organization discovers that its AI system for automated content moderation on a children's educational platform has been trained on data that includes adult content that was mislabeled as "age-appropriate" due to annotation errors. The system has been operational for four months. What is the MOST comprehensive governance response?

A. Retrain the model with corrected labels and redeploy it within the next scheduled maintenance window, which is in six weeks

B. Issue a software update that adjusts the content classification thresholds to be more conservative, reducing the likelihood that inappropriate content reaches children

C. Add a disclaimer to the platform informing parents that the content moderation system may occasionally permit age-inappropriate material to appear

D. Immediately implement emergency content restrictions while simultaneously conducting a retrospective review to identify any inappropriate content that may have been shown to child users during the four-month period, notifying parents and platform operators, retraining the model with corrected labels, and implementing enhanced data annotation quality controls to prevent recurrence

87. An AI system for evaluating mortgage applications consistently produces longer processing times for applications from minority neighborhoods — not because the system discriminates in its final decisions, but because it requests additional documentation from these applicants more frequently. The additional documentation requests are triggered by data completeness thresholds that are harder to meet for applicants in areas with less digitized property and financial records. The final approval rates are equal across demographic groups. A governance professional argues that this still constitutes a fairness concern. Why?

A. Disparate processing burden — even when final outcomes are equal — can constitute a fairness concern because applicants from minority neighborhoods bear a disproportionate documentation burden,

experience longer processing times, and may abandon applications due to the additional requirements, creating a chilling effect that discourages applications from affected communities even when approvals would ultimately be granted

B. The governance professional is incorrect because equal final approval rates demonstrate that the system treats all applicants fairly, and processing time differences do not constitute a fairness concern under any governance framework

C. The concern is valid only under U.S. fair lending law and has no relevance under the EU AI Act or GDPR

D. The concern is limited to a customer experience issue that should be addressed by the customer service team rather than the governance function

88. An organization operates an AI system that processes employee timesheets and detects anomalies that may indicate time theft. The system flags an employee who consistently submits timesheets showing exactly 8.0 hours per day with no variation. A human reviewer investigates and discovers that the employee actually works exactly 8 hours per day as scheduled and submits accurate timesheets. The AI system flagged the timesheets because the lack of variation was itself statistically anomalous in the training data, where most employees show minor daily variations. What does this scenario illustrate about AI anomaly detection?

A. The scenario illustrates a monitoring system failure because the human reviewer should not have been required to investigate a clearly accurate timesheet submission

B. Anomaly detection systems flag statistical outliers, not necessarily actual problems — accurate behavior that happens to be statistically unusual can trigger false alerts, requiring human judgment to distinguish between genuine anomalies warranting investigation and legitimate behavioral patterns that are simply uncommon

C. The scenario illustrates that the AI system's anomaly detection threshold is set correctly because it successfully identified an unusual pattern that warranted human investigation

D. The scenario illustrates that AI anomaly detection systems should be programmed with explicit rules defining what constitutes time theft rather than relying on statistical anomaly detection

89. An AI vendor sends an organization a notification that the vendor's training data has been found to contain personal data that was collected without appropriate consent — affecting data subjects in the EU. The organization uses the vendor's model in a customer-facing application. What is the deployer's MOST immediate governance concern?

A. The deployer's reputational risk if the media reports on the vendor's data collection practices and associates the deployer's brand with the data protection violation

B. Whether the deployer needs to renegotiate the vendor contract to include stronger data provenance warranties in light of the vendor's demonstrated compliance failure

C. Whether the deployer can continue marketing the AI-powered product feature to customers without disclosing the vendor's data compliance issue

D. Whether the deployer's own processing activities — which are built on a model trained with unlawfully collected data — may themselves constitute a GDPR violation, potentially triggering obligations to assess the impact on data subjects, notify the supervisory authority, and evaluate whether continued use of the model is legally defensible

90. An AI governance professional is reviewing the organization's complete AI governance program — policies, roles, training, monitoring, documentation, incident response, and vendor management. The professional identifies that all governance activities are functioning well at the individual system level but that the organization has no process for learning from governance experiences across systems. Audit findings from one system are not shared with teams managing other systems. Incident lessons from one deployment do not inform governance practices for other deployments. What governance capability is missing?

A. The organization is missing a regulatory scanning function that monitors changes in AI regulations across jurisdictions and distributes updates to all governance teams

B. The organization is missing a vendor management function that centralizes all vendor relationships and provides portfolio-level visibility into third-party AI risk

C. An organizational learning capability — a systematic process for capturing governance insights, sharing lessons learned across teams and systems, and continuously improving governance practices based on accumulated experience across the entire AI portfolio

D. The organization is missing a dedicated AI security function that coordinates adversarial testing and threat modeling across all deployed systems

91. An organization is considering whether to deploy a vendor AI system that has not been certified to ISO/IEC 42001. The vendor argues that ISO certification is unnecessary because the system complies with all applicable EU AI Act requirements. Is the vendor's argument sufficient from the deployer's governance perspective?

A. The vendor's EU AI Act compliance claim should be verified through the deployer's own evaluation, but the absence of ISO 42001 certification does not in itself prevent deployment — ISO certification is valuable as external assurance but is voluntary, and compliance with applicable legal requirements can be demonstrated through other evidence

B. The absence of ISO 42001 certification is an absolute disqualification because all high-risk AI systems must be certified to ISO 42001 before they can be deployed in the European Union

C. The vendor's argument is sufficient because EU AI Act compliance is the only governance requirement that matters, and ISO certification provides no additional governance value

D. The deployer should require ISO 42001 certification as a contractual condition because it is the only way to verify that the vendor has a functioning AI management system

92. An AI system deployed for automated insurance underwriting uses a model that considers applicants' occupations when calculating premiums. The system assigns higher premiums to occupations with higher injury rates. A governance review discovers that certain occupations with high injury rates — such as construction, agriculture, and domestic work — are disproportionately held by immigrant workers and racial minorities. The occupation-based pricing, while actuarially justified, produces premiums that track demographic composition. What governance analysis is REQUIRED?

A. No governance analysis is required because actuarially justified pricing based on legitimate risk factors is explicitly permitted under insurance regulations regardless of demographic correlation

B. The governance analysis must evaluate whether the actuarial justification for occupation-based pricing is sufficient to justify its disparate impact — considering whether the risk differentials are supported by current data, whether alternative rating factors with less disparate impact could achieve equivalent risk differentiation, and whether the pricing approach complies with applicable nondiscrimination regulations

C. The only governance requirement is to verify that the AI system does not use race or national origin as direct input features, which is sufficient to demonstrate nondiscrimination compliance

D. The governance analysis should focus exclusively on whether the insurance premiums are profitable for the organization rather than evaluating their distributional impact across demographic groups

93. An AI governance professional has been asked to evaluate an organization's readiness to deploy its first AI system. The organization has completed the following: documented AI governance policies, established a governance committee, classified the proposed system as high-risk, conducted an impact assessment, and developed model cards. However, the organization has NOT yet: tested the system for

bias, established monitoring capabilities, defined incident response procedures, or trained human overseers. What should the governance professional conclude?

A. The organization is fully ready for deployment because the documentation and policy elements are the most important components of AI governance

B. The organization is not ready because while foundational governance structures are in place, critical operational elements are missing — bias testing, monitoring, incident response, and human oversight training are prerequisites for deployment, not post-deployment enhancements

C. The organization is ready for deployment because the remaining items can be completed during the first month of operation without significant governance risk

D. The organization should deploy the system in a limited pilot to generate the operational data needed for bias testing and monitoring configuration, then expand to full deployment

94. An AI system for predicting equipment maintenance needs in a fleet of commercial aircraft has been operating for 18 months. The system was trained on maintenance data from a fleet of 200 aircraft. The airline has now added 50 new aircraft of a different model to its fleet and wants to extend the AI system's coverage to these new aircraft without retraining. The operations team argues the system "understands maintenance patterns" generically. What is the governance concern?

A. The concern is limited to ensuring the new aircraft's sensor data is compatible with the AI system's input format, which is a technical integration matter rather than a governance concern

B. The new aircraft model likely has different mechanical characteristics, failure modes, maintenance profiles, and sensor configurations than the original fleet — applying a model trained exclusively on one aircraft type to a different type creates a validation gap for safety-critical predictions

C. The concern is only relevant if the new aircraft model is manufactured by a different company than the original fleet aircraft

D. The concern can be addressed by increasing the monitoring frequency for the new aircraft from weekly to daily during the first three months of operation

95. An organization operates an AI-powered fraud detection system and an AI-powered customer segmentation system. Both systems independently process the same customer data but for different purposes — security and marketing, respectively. A governance professional discovers that the

marketing team has been using the fraud detection system's risk scores as a feature in the customer segmentation model, reasoning that "low fraud risk" customers are more valuable marketing targets. What governance principle has been violated?

A. Purpose limitation — fraud detection risk scores were generated for security purposes, and using them as marketing segmentation features constitutes secondary use that likely exceeds the original processing purpose without proper governance review or potentially a new lawful basis

B. Data minimization, because the marketing team should not have access to the fraud detection system's outputs under any circumstances

C. The principle of accountability, because the fraud detection team was not informed that their system's outputs were being used by the marketing team

D. No governance principle has been violated because both systems process the same customer data and the organization has a lawful basis for processing that data

96. An organization's AI governance program has been operational for three years. The governance team has successfully brought all AI systems into compliance, established monitoring, and conducted regular audits. However, a recent employee survey reveals that 70% of data scientists view the governance program as "bureaucratic overhead that slows innovation without adding value." What does this perception suggest about the governance program, and how should the governance team respond?

A. The perception is expected and acceptable because governance is inherently burdensome and resistance from technical teams is normal and should be ignored

B. The perception suggests the governance program may have failed to communicate its value proposition, may impose disproportionate requirements on lower-risk systems, or may not be sufficiently integrated into development workflows — the governance team should engage data scientists to identify specific friction points, streamline processes where possible, demonstrate governance value through concrete examples, and ensure governance is proportionate to risk

C. The perception indicates the governance program should be discontinued because it is impeding the organization's ability to innovate and compete effectively in the AI market

D. The perception can be resolved by making all governance activities mandatory through automated enforcement systems that prevent data scientists from deploying code without governance approval

97. An AI system for automated medical triage in an emergency department assigns urgency scores to patients. A patient arrives with symptoms that the system has never encountered in combination before

— the symptom combination falls entirely outside the system's training distribution. The system assigns a moderate urgency score with 52% confidence. The triage nurse notices the low confidence score. Under the principle of human oversight, what is the MOST appropriate response?

A. Follow the system's moderate urgency assignment because 52% confidence still represents the system's best assessment based on available information

B. Override the system and assign the lowest urgency score because the system's low confidence indicates the patient's condition is unlikely to be serious

C. Shut down the AI triage system entirely because encountering out-of-distribution inputs demonstrates the system is unreliable and should not be used for any patient assessment

D. The triage nurse should exercise independent clinical judgment to assess the patient directly — the system's low confidence on an out-of-distribution input signals that the AI's assessment is unreliable for this case and should not be relied upon, with the override documented in accordance with the system's governance procedures

98. An AI system that assists law enforcement with surveillance camera analysis is deployed in a city center. The system can identify and track individuals across multiple camera feeds. A civil liberties organization files a complaint arguing that the system violates privacy rights even though it is deployed in public spaces where individuals have reduced expectations of privacy. From a governance perspective, how should the balance between public safety and privacy be evaluated?

A. Privacy concerns are automatically overridden in public spaces because individuals voluntarily choose to appear in public and implicitly consent to any form of surveillance

B. The law enforcement benefits of the surveillance system automatically justify its deployment because public safety is the highest governance priority in all circumstances

C. The governance evaluation must consider that AI-powered identification and tracking across multiple cameras creates a qualitatively different surveillance capability than traditional CCTV — the ability to persistently track individuals across an entire city raises privacy concerns that go beyond the reduced expectation of privacy in public spaces, and must be evaluated against proportionality, necessity, and the specific restrictions in the EU AI Act on biometric identification in public spaces

D. The complaint should be dismissed because surveillance camera analysis is classified as minimal risk under the EU AI Act and is not subject to governance restrictions

99. An AI system used for automated loan underwriting has been successfully operating for five years with strong performance and no incidents. The governance committee proposes reducing the system's governance oversight — extending audit intervals from annual to biennial, reducing monitoring from continuous to monthly, and eliminating the requirement for human review of AI-assisted decisions for applications above €100,000. What is the MOST appropriate governance response to this proposal?

A. Evaluate each proposed reduction against the system's current risk profile and regulatory requirements — audit interval extension may be acceptable if supported by the track record, but reducing monitoring frequency and eliminating human review for high-value decisions may increase risk exposure beyond acceptable levels for a high-risk financial AI system that directly affects individuals' access to credit

B. Approve all proposed reductions because five years of incident-free operation demonstrates that the system is reliable enough to justify reduced oversight

C. Reject all proposed reductions because governance requirements for high-risk systems are fixed and cannot be modified regardless of operational track record

D. Defer the decision to the AI vendor because the vendor's recommendations should determine the appropriate level of ongoing governance for the system

100. An AI governance professional is asked to summarize, in a single statement, the principle that connects every chapter of the AIGP Body of Knowledge — from AI fundamentals through legal frameworks to applied development and deployment governance. Which statement MOST accurately captures this unifying principle?

A. AI governance requires the most expensive technology solutions available to ensure that automated systems are properly supervised and controlled

B. AI governance is fundamentally about ensuring that AI systems serve human values and interests throughout their entire lifecycle — by understanding what AI systems are and how they work, by knowing the laws and standards that constrain them, by governing their development with rigor and accountability, and by maintaining vigilant oversight of their deployment, use, and eventual retirement so that the benefits of AI are realized while its risks are managed responsibly

C. AI governance is primarily about ensuring that organizations comply with the EU AI Act's requirements to avoid regulatory penalties and enforcement actions

D. AI governance is primarily a technical discipline that requires deep expertise in machine learning algorithms and neural network architectures to implement effectively

Practice Exam 4: Answer Key and Explanations

1. D — Concept drift occurs when the relationship between inputs and outputs changes — here, a new disease creates symptom-outcome patterns the model never learned. The model cannot recognize a condition absent from its training data, regardless of overall accuracy. Aggregate metrics masking subgroup failures is a separate monitoring concern, but the root cause is a changed diagnostic landscape.

2. A — High-risk lending decisions directly affecting individuals' access to credit warrant human-in-the-loop oversight where each decision is reviewed before it becomes final. Statistical sampling (human-on-the-loop) may detect patterns after harm has already occurred but does not prevent harm to individual applicants whose specific decisions were never reviewed.

3. C — Both principles are violated simultaneously. Purpose limitation is violated because generating sensitive inferences about income, family status, and health conditions goes beyond "product recommendations and service improvement." Transparency is violated because individuals are not informed that these specific types of inferences are being generated from their purchasing behavior.

4. B — Deployers bear independent governance obligations under the EU AI Act that are not contingent on the provider's remediation timeline. When fairness metrics degrade below acceptable thresholds, the deployer must take its own protective measures — manual review, scope restriction, or suspension — to mitigate ongoing harm while awaiting the provider's fix.

5. D — The EU AI Act's limited-risk transparency provision applies to AI systems designed to interact with natural persons — the trigger is the interaction with people, not the processing of personal data, the model architecture, or the risk classification. A chatbot that communicates with customers must inform them they are interacting with AI.

6. C — Transfer learning carries inherited governance risk. The pre-trained base model may contain biases, learned behaviors, and limitations from its original training that persist through fine-tuning. Governance must evaluate the base model's characteristics independently and test the fine-tuned model comprehensively rather than assuming fine-tuning resolves all inherited issues.

7. A — The most likely explanation is that the test dataset did not adequately represent the actual applicant population. Demographic composition, qualification distributions, and application patterns in

real-world operations often differ from curated test sets, causing models that appear fair in testing to produce disparate outcomes when processing actual applicant data.

8. B — The EU AI Act's Annex III classifies AI systems used in education to evaluate learning outcomes or determine access to educational institutions as high-risk. Whether the system is advisory or determinative, and whether it is used in public or private settings, does not change this classification — the use case category determines the risk tier.

9. C — Applying the EU's more restrictive standards globally ensures compliance in all jurisdictions while providing a single, consistent governance framework. This approach accepts that U.S. employees may receive less functionality than local law would permit, but eliminates the complexity and governance risk of maintaining different system versions across regions.

10. D — The monitoring framework's reliance on aggregate false negative rates masked a 40% increase for peer-to-peer mobile payments — a rapidly growing transaction category. This demonstrates that aggregate metrics can conceal significant subgroup deterioration. The framework needs transaction-type-disaggregated monitoring to detect category-specific degradation.

11. A — Under GDPR Article 21, when a data subject objects to processing based on legitimate interests, the organization may continue only if it demonstrates compelling legitimate grounds that override the individual's interests. The burden of proof shifts to the organization upon receiving the objection — the objection is not automatically upheld, but neither can it be automatically dismissed.

12. B — Without transparent fairness testing methodology and disaggregated results, the deployer cannot independently verify that the system meets governance and regulatory requirements for its specific deployment. A claim of "tested for fairness" without supporting evidence is insufficient for a high-risk biometric system, and the deployer must evaluate whether this governance gap is acceptable.

13. D — Risk-based prioritization directs limited audit resources to the system with the highest potential for harm. The medical diagnostic tool directly affects patient health and safety, is likely classified as high-risk under the EU AI Act, and carries the most severe consequences if governance failures exist. It should be audited first.

14. C — The governance committee should approve deployment with targeted safeguards that address the known limitation in the deployment context. Routing thin-credit-file applications to manual review ensures that the affected population (first-time borrowers) receives appropriate human evaluation while the AI system handles applications where it performs reliably.

15. A — Sick day usage correlates with disability, chronic illness, pregnancy, and caregiving — all protected categories. Using it as a predictive feature creates disparate impact risk against employees in these categories. Additionally, sick day data may constitute health-related information under GDPR Article 9, requiring a special category exception beyond the standard lawful basis.

16. B — Workforce diversity in AI teams is an AI risk management concern, not merely an HR initiative. Diverse teams identify risks that homogeneous teams miss, detect biases that mirror their own experiences, and produce more robust governance outcomes. NIST explicitly includes workforce diversity in the Govern function because it directly improves risk identification quality.

17. C — The EU AI Act requires deployers to use high-risk AI systems in accordance with the provider's instructions for use. The instructions form a comprehensive set of requirements — exceeding one element does not compensate for failing another. The training specification is part of the instructions, and non-compliance with it is a violation regardless of enhanced monitoring.

18. D — If synthetic data samples closely resemble specific real individuals from the training set, those samples may constitute personal data under GDPR — enabling identification of real people even though the data was artificially generated. This creates privacy, consent, and data protection obligations that must be addressed regardless of the data's synthetic origin.

19. A — A blind spot in the training data — the absence of indemnification cap clauses — means the system cannot detect what it was never trained to recognize. Two years of undetected contract risk requires retroactive review of all processed contracts to identify and address any unflagged problematic terms, plus immediate remediation of the training data gap.

20. B — GDPR requires "meaningful information about the logic involved" in automated processing that affects individuals. A generic acknowledgment that AI was used and "multiple factors" were considered provides no actionable information. The applicant must receive specific information about which factors influenced the ranking and how they were weighted.

21. D — The system is being applied to personal injury claims — a purpose it was never designed, trained, or validated for. Processing personal injury claims using property damage assessment logic produces unreliable outputs for every personal injury claim processed. This is secondary use without governance review, creating systematic harm to claimants.

22. C — The summary lacks sufficient detail for rights holders to determine whether their works were used. The EU AI Act requires "sufficiently detailed" training data summaries that enable rights holders to exercise their copyright reservations. Generic category descriptions ("internet text, books, code") do not enable this determination.

23. A — Even though the system is advisory, if its recommendations materially influence sentences, due process requires that the basis for those recommendations be sufficiently interpretable. A defendant's right to understand and challenge the evidence against them extends to AI-generated risk assessments that influence the severity of their sentence.

24. B — Feature engineering transformations can introduce or amplify biases in ways that are invisible without lineage tracking. Without documenting these intermediate steps, the organization cannot trace bias origins, investigate incidents caused by data transformations, or reproduce the training dataset. The lineage gap undermines multiple governance functions.

25. D — Label bias occurs when human annotators' subjective interpretations contaminate the ground truth labels. Team C interpreted "innovation potential" as favoring younger candidates, embedding age bias into the labels for financial sector resumes. The model learned this biased labeling pattern as the definition of a successful candidate.

26. C — Multiple fairness definitions are mathematically incompatible in most real-world settings — satisfying one metric often means violating another. The governance decision should be context-driven: which harms are most salient, which metrics detect those harms, what the regulatory requirements specify, and what values the deployment context prioritizes. There is no universally correct metric.

27. B — A system can function exactly as designed while producing harmful outcomes that governance must address. The product team's argument conflates technical performance (meeting the engagement objective) with governance appropriateness (whether the engagement objective itself produces

acceptable outcomes). The committee should evaluate whether the optimization objective needs modification.

28. A — The system creates a punitive feedback loop where cooperative employees are systematically assigned worse shifts, while employees who complain receive better treatment. This may create disparate impact on vulnerable employees — those with less job security, immigration dependencies, or economic pressure — who feel unable to object and are therefore trapped in increasingly undesirable schedules.

29. C — When the European Commission updates Annex III to include a new use case category, existing systems that fall within the new category must be brought into compliance. The organization must evaluate whether its current governance controls satisfy the new requirements and implement additional measures where gaps exist.

30. D — Increasing data drift is a leading indicator that performance and fairness degradation may follow. The appropriate response is proactive investigation (understanding what is causing the drift and where it is heading), increased monitoring frequency (watching for early signs of performance impact), and preparation for retraining (so the organization can act quickly if degradation materializes).

31. B — Meaningful human intervention under GDPR Article 22 requires that the reviewer has access to relevant information, the competence to evaluate the decision independently, and sufficient time and authority to reach a genuine conclusion. A 30-second review without access to underlying factors is a rubber stamp that does not constitute meaningful intervention.

32. A — A GDPR DPIA addresses data protection risks but does not cover the full scope of AI impacts required by ISO 42001 — including safety risks, fairness concerns, societal effects, environmental impact, and organizational risks. The DPIA is a necessary component of a comprehensive impact assessment but is not sufficient to satisfy the broader ISO standard.

33. D — "Decision tree" encompasses a wide range of complexity levels. A small, shallow tree with few branches may be genuinely interpretable. However, ensemble methods like random forests or gradient boosting combine hundreds or thousands of trees, creating a model that is functionally as opaque as a neural network. The vendor's blanket claim requires scrutiny of the actual model complexity.

34. C — Organizations can establish internal governance policies that exceed regulatory minimums. Proportionate governance means calibrating assessment depth to risk level — not eliminating assessment for lower-risk systems. A streamlined impact assessment proportionate to the system's limited risk satisfies both governance discipline and the proportionality principle.

35. B — Fairness is context-dependent. Different populations have different demographic compositions, different protected characteristics receive different legal treatment, different historical discrimination patterns create different proxy effects, and different cultural norms shape different expectations. Validation for one population does not transfer to another without independent evaluation.

36. A — All identified vulnerabilities should be remediated before customer-facing deployment. The sophistication of current exploits is not a reliable predictor of future accessibility — prompt injection techniques evolve rapidly and become more accessible over time. Deploying with known safety control bypasses creates unacceptable risk for a system that interacts directly with customers.

37. C — Two bias types compound: label bias (fraud labels reflect biased investigation patterns that disproportionately targeted claims from low-income community clinics) built on historical bias (the investigation patterns themselves reflected historical discrimination against low-income communities). The model has learned to associate certain diagnosis codes with fraud because of how those communities were historically treated, not because of actual fraud prevalence.

38. D — Both parties bear responsibility. The vendor should have provided advance notice and detailed documentation of performance changes across use cases before pushing the update. The deployer should have maintained controls to evaluate updates before automatic application. The shared failure enabled an untested change to reach production and affect operations.

39. B — The governance committee should evaluate whether the CLV-based service differentiation creates unfair outcomes. If income correlation means that service quality effectively tracks protected characteristics through proxy effects, and if less affluent customers systematically receive inferior service, the practice may constitute unfair treatment under consumer protection principles.

40. A — The radiologist bears primary responsibility because the human oversight mechanism was in place but was not meaningfully exercised. Signing a report without independently reviewing the images

is the definition of automation bias — the oversight safeguard existed but was rendered ineffective by uncritical acceptance of the AI output.

41. D — Level 3 (Defined) describes an organization where governance structures exist and are documented but are not yet consistently implemented across the organization or systematically measured. The presence of policies, roles, and training with inconsistent adherence and no effectiveness measurement is the hallmark of this maturity level.

42. C — For a medical imaging AI, the test set must evaluate performance on populations that the model will encounter in deployment — including underrepresented groups, different clinical settings, and edge cases. Random sampling from the same distribution as training data only confirms the model works on data similar to what it already saw, not on the diverse populations it will serve.

43. B — Measurement bias is embedded in the system's design through the "consistency" feature. The feature measures data availability across sources rather than actual information consistency. Communities with less digitized infrastructure score lower not because their information is inconsistent but because fewer digital sources exist to cross-reference — a measurement artifact of infrastructure inequity, not a risk signal.

44. A — The deployer should conduct its own multilingual testing before deployment. The vendor's English-only validation provides no assurance about the system's performance, fairness, or safety for the other six languages. The deployer's independent evaluation is essential because the deployer bears governance obligations for its specific deployment context.

45. D — This illustrates the dual-use nature of AI technology. A deepfake detection system's methodology, if understood by adversaries, could be used to develop deepfakes specifically designed to evade that detector — creating an arms race where detection and generation capabilities escalate in tandem. Governance must anticipate this dynamic.

46. C — Stress testing based on simulated adverse conditions may not capture the complex, non-linear, and cascading behavioral patterns that characterize actual market crises. Real bear markets involve panic selling, correlated defaults, liquidity crises, and systemic contagion effects that are extremely difficult to simulate accurately from bull market data alone.

47. A — Warning thresholds exist to provide early detection while intervention is still feasible. Dismissing warning alerts as noise defeats the tiered alerting system's purpose. The deviation that was within warning range two weeks earlier crossed into critical territory and harmed 500 individuals — harm that prompt investigation of the initial warning could have prevented.

48. B — Candidates ranked below position 20 are rejected through a solely automated process with no human involvement. The rejection produces significant effects on their employment prospects. Both conditions for GDPR Article 22 are met for these candidates, triggering their rights to human intervention, to express their point of view, and to contest the decision.

49. C — Governance cannot be designed without first understanding what needs to be governed. An inventory of all AI systems — documenting purpose, affected populations, data processed, and risks — provides the foundation for risk classification, prioritization, policy development, and monitoring design. Every subsequent governance activity depends on this baseline understanding.

50. D — Historical resource allocation based on insurance status produced systematically better clinical documentation for privately insured patients. The AI system interpreted documentation quality as clinical urgency rather than recognizing it as an artifact of inequitable historical resource allocation. The bias is embedded in the training data through historical institutional practices.

51. A — Responsible AI principles include consideration of environmental impact. The governance framework should evaluate whether the larger model's expected benefits justify the 5x increase in energy consumption, whether more efficient approaches could reduce the footprint, and should document the environmental analysis as part of the impact assessment.

52. B — Legacy AI systems that predate the governance program must be brought into compliance through retroactive governance activities — documenting current state, conducting risk assessment, performing testing, and establishing monitoring. The priority and depth of these activities should be determined by the system's risk classification.

53. D — AI systems that detect genuinely novel patterns create a unique governance challenge — the anomaly may represent a real threat or a false positive, and the organization's existing expertise may not be sufficient to determine which. Governance must provide processes for investigating novel anomalies, including escalation pathways and access to specialized expertise.

54. C — Dual-audience deployment requires differentiated presentation. Professionals need detailed analytical information with expert context. End users need simplified explanations, stronger guardrails, clearer limitation disclosures, and accessible pathways to human assistance. The same AI output presented identically to both audiences fails to account for the different needs and capabilities of each group.

55. A — The system is generating medical advice outside its intended purpose, creating foreseeable physical harm risk. A terms-of-service disclaimer does not eliminate the organization's responsibility to implement technical controls that prevent the system from generating content in domains where it has no validation and where incorrect outputs could cause serious harm.

56. B — The deployment was approved conditionally on all three mitigations being in place. Proceeding without implementing two of the three approved conditions violates the governance committee's authorization. The gap between approved governance controls and actual operational practice creates an accountability failure that may expose the identified vulnerable population to the unmitigated risk.

57. C — Incomplete monitoring records during a regulatory investigation significantly weaken the organization's position. The gap demonstrates a period during which the organization cannot prove it was monitoring compliance, fairness, and performance — undermining the accountability principle and potentially constituting both a record-keeping violation and evidence of inadequate oversight.

58. D — The EU AI Act requires providers to report serious incidents to market surveillance authorities and to inform affected deployers so they can take their own protective measures. Both notifications should occur promptly after establishing the causal link. Coordination on notification of affected applicants follows, as deployers are best positioned to identify and communicate with affected individuals.

59. A — Vendors' governance practices, security posture, model performance, financial stability, and regulatory compliance can all change after initial procurement. Ongoing monitoring ensures the deployer detects vendor-side changes that may affect the AI system's governance profile before they cause harm in the deployment context — procurement-time assessment alone provides a point-in-time snapshot that depreciates over time.

60. B — The validation gap concerns specific subpopulations — students with non-traditional backgrounds, learning disabilities, or non-native language use — whose characteristics may be underrepresented in the validation data even though they are part of the overall applicant pool. For high-stakes examinations affecting educational futures, subpopulation validation is essential.

61. D — A change management process requiring pre-deployment testing of safety-critical features after every system modification would have verified that the emotional distress detection feature still functioned before the update reached production. Testing safety-critical pathways after changes is a fundamental governance control for any system with safety-sensitive features.

62. C — Human oversight means experienced professionals retain the authority to exercise independent judgment. The supervisor's 20 years of experience may recognize signs — sounds, vibrations, visual indicators — that the sensor-based AI system cannot detect. The supervisor should override the AI, initiate immediate inspection, and document the clinical reasoning for the override.

63. B — Request A (access to training data) is challenging but follows standard data inventory processes. Request B (explanation of a denial decision) requires the AI system to produce "meaningful information about the logic involved" — an explainability capability that must be architected into the system. If the system was not designed to support explanations, this request may be technically difficult or impossible to fulfill.

64. A — Unnotified model updates mean the system's behavior can change without the deployer's knowledge or governance review. The deployer cannot evaluate whether updates affect performance, fairness, or compliance in its context, potentially violating regulatory obligations. This creates a governance blind spot where the system in production may differ from the system that was approved.

65. D — Human review of documents classified as "public" before actual publication creates a safety net for misclassification errors. For irreversible actions — such as publishing information publicly — governance controls must include verification steps that catch errors before they produce consequences that cannot be undone.

66. C — Accountability requires shared responsibility. The decision-maker is responsible for exercising independent judgment rather than blindly following AI outputs. The organization is responsible for

creating conditions that support good judgment — adequate training, explainable outputs, clear override authority, and a culture that doesn't penalize questioning AI recommendations.

67. A — Using safety monitoring data to discipline individual workers is a secondary purpose beyond the stated safety monitoring objective. This represents a different purpose requiring separate governance review, potentially a separate lawful basis under GDPR, and a different impact assessment — because the consequences for individuals (disciplinary action) differ fundamentally from the original purpose (aggregate safety improvement).

68. D — System A's inability to explain decisions undermines transparency, accountability, and affected individuals' rights. When errors occur in an unexplainable system, the organization cannot identify what went wrong, affected individuals cannot challenge adverse outcomes meaningfully, and regulators cannot verify compliance. High accuracy does not compensate for the governance deficiencies that opacity creates.

69. B — Human oversight for systems operating faster than human decision-making must occur at an appropriate time scale. This means monitoring aggregate patterns, reviewing statistical summaries, setting operational parameters and risk limits, and maintaining the authority and capability to halt the system. Oversight adapts to the system's operational tempo rather than attempting to match it.

70. D — The 4,000-applicant weekly capacity gap means individuals depending on timely benefit determinations will experience significant delays. Governance must plan for this impact through phased deactivation, temporary staffing increases, interim processing arrangements, or other measures that prevent the transition from creating harm for the vulnerable population the system serves.

71. A — Standard benchmarks evaluate performance under controlled conditions that may not reflect the deployer's specific operational context. A model excelling on benchmarks may underperform for the deployer's particular population, data characteristics, and use case requirements. Deployment-specific evaluation using representative operational data is essential for governance assurance.

72. B — For medical AI, conservative error handling for safety-critical categories ensures that information with direct patient safety implications — allergies, adverse reactions, life-threatening conditions — is always included in summaries. The system should prioritize completeness for safety-critical information even at the cost of longer outputs, with clear flagging for physician attention.

73. D — The organization is responsible for the accuracy of content it publishes regardless of whether a human or AI generated it. Deploying generative AI for marketing without a verification process for factual claims creates foreseeable risk of publishing false and potentially defamatory content. Output verification is a governance requirement for any AI-generated content published under the organization's name.

74. C — The agentic system's autonomous operation eliminates human review between action steps. Each email sent without review can contain errors, inappropriate content, or inaccurate claims that compound across the workflow. Unlike the standard assistant where a human reviews each message, the agentic system's errors reach external recipients without any opportunity for correction.

75. A — Government AI transparency and accountability extend beyond logging to include public disclosure, meaningful explainability for affected individuals, challenge mechanisms, independent audit access, and clear responsibility assignment. Documentation creates a record; full transparency and accountability require active disclosure, interpretability, and recourse mechanisms.

76. B — Governance frameworks for safety-critical environments must prepare operators for high-stakes, time-pressured decisions through clear protocols, escalation procedures, training on system limitations, and access to additional diagnostic information. The scenario illustrates why human oversight requires not just authority but also the preparation and support to exercise it under pressure.

77. D — Model card transparency about performance disparities is a governance obligation enabling deployers to make informed decisions. Removing disaggregated metrics would conceal a known performance gap, violating transparency requirements and potentially exposing deployers to liability for deploying a system whose limitations they were prevented from understanding.

78. C — Portfolio-level analysis reveals systemic risks invisible at the system level: concentration risk from shared vendors, common training data dependencies, correlated failure modes, and cumulative impact on populations affected by multiple AI systems. These cross-system risks cannot be detected by governing each system in isolation.

79. B — Three governance principles converge: privacy (processing personal health data without clear consent), data minimization (capturing pre-meeting conversations exceeds the meeting-summary purpose), and transparency (participants may not have known the system records audio before the

formal meeting begins). The incident demonstrates how AI recording systems can exceed their intended scope.

80. A — The decision requires proportionality analysis weighing the fairness concern's severity, the affected population's vulnerability, the availability of interim mitigations, and the business consequences of delay. Governance does not impose absolute rules that moderate risks always block or always permit deployment — it requires contextual judgment with documented rationale.

81. C — The percentage of AI systems with complete governance coverage — tracked over time — directly measures whether the program is closing the compliance gaps identified in the audit. This metric connects governance activity to outcomes and shows whether the organization is actually governing its AI portfolio more comprehensively over time.

82. D — The most effective prevention addresses the investigation process itself: prompt investigation with clear SLAs, interim support for flagged claimants, and rapid-clearance procedures for legitimate claims. This ensures that even when the AI correctly flags a suspicious pattern, the human follow-up process does not unnecessarily harm legitimate claimants through extended delays.

83. B — Transparent limitation disclosure builds deployer trust, demonstrates governance maturity, enables deployers to implement appropriate safeguards, and reduces liability by ensuring known limitations were disclosed. Concealing limitations creates greater competitive risk through liability exposure, regulatory violations, and destroyed trust when limitations are eventually discovered.

84. A — The system's optimization objective (appointment distribution) is misaligned with the governance objective (equitable patient care). The model learned a statistical pattern (complex patients return frequently) and applied it to an optimization that produces the opposite of appropriate care — shorter appointments for patients who need longer ones.

85. C — The committee must evaluate proportionality: whether the surveillance mechanism is proportionate to the safety benefit, whether less invasive alternatives exist, whether surveillance will chill legitimate communication, whether the employer-employee power dynamic undermines consent, and whether the system addresses root causes of harassment or merely monitors symptoms.

86. D — A comprehensive response addresses immediate safety (emergency content restrictions), retrospective harm assessment (identifying inappropriate content shown to children), root cause remediation (retraining with corrected labels), and future prevention (enhanced annotation quality controls). Each element addresses a different dimension of the governance failure.

87. A — Even when final approval rates are equal, disparate processing burden constitutes a fairness concern. Disproportionate documentation requirements, longer processing times, and the chilling effect of additional hurdles create systematic disadvantage for applicants from affected communities — some may abandon applications entirely, never appearing in the "equal approval rate" statistics.

88. B — Anomaly detection flags statistical outliers, not necessarily actual problems. The employee's accurate-but-unusual behavior triggered a false alert because consistency was itself uncommon in the training data. This illustrates why human judgment is essential in anomaly detection governance — to distinguish genuine concerns from legitimate behavioral patterns that happen to be statistically rare.

89. D — The deployer's most immediate concern is whether its own processing activities — built on a model trained with unlawfully collected data — may constitute a GDPR violation. The deployer must assess the impact on its own compliance posture, evaluate whether continued use is legally defensible, and determine whether supervisory authority notification is required.

90. C — An organizational learning capability systematically captures governance insights, shares lessons across teams, and improves practices based on accumulated experience. Without this capability, the organization repeats mistakes, misses improvement opportunities, and fails to leverage its growing governance experience across its AI portfolio.

91. A — ISO 42001 certification is valuable as external assurance but is voluntary. The vendor's EU AI Act compliance claim should be verified through the deployer's own evaluation — reviewing documentation, testing in the deployment context, and assessing governance practices. Compliance can be demonstrated through evidence other than ISO certification.

92. B — Actuarial justification does not automatically resolve governance concerns about disparate impact. The analysis must evaluate whether risk differentials are supported by current data, whether alternative rating factors with less disparate impact could achieve equivalent risk differentiation, and whether the approach complies with applicable nondiscrimination regulations.

93. C — While foundational governance structures are in place (policies, committee, risk classification, impact assessment, model cards), critical operational elements are missing. Bias testing, monitoring, incident response, and human oversight training are prerequisites for deployment because they provide the safeguards that protect individuals once the system begins making decisions.

94. D — Different aircraft models have different mechanical characteristics, failure modes, maintenance profiles, and sensor configurations. Applying a model trained on one aircraft type to a different type without validation creates a gap in safety-critical predictions. The operations team's assumption that maintenance knowledge is generic does not account for these material differences.

95. A — Purpose limitation requires that data processed for one purpose not be repurposed for an incompatible purpose without proper governance. Fraud detection risk scores were generated for security purposes. Using them as marketing segmentation features constitutes secondary use that exceeds the original processing purpose and requires separate governance review.

96. B — The 70% negative perception signals potential governance program issues — disproportionate requirements for low-risk systems, insufficient integration into workflows, or failure to communicate governance value. The governance team should engage data scientists to identify friction points, streamline where possible, demonstrate value, and ensure proportionality.

97. D — A 52% confidence score on an out-of-distribution input signals that the AI's assessment is unreliable for this specific case. The triage nurse should exercise independent clinical judgment, assess the patient directly using clinical expertise, and document the override. Human oversight is most critical precisely when the AI system signals low confidence.

98. C — AI-powered identification and tracking across multiple cameras creates qualitatively different surveillance capabilities than traditional CCTV. Persistent tracking of individuals across an entire city raises privacy concerns that exceed the reduced privacy expectation in public spaces. The evaluation must consider proportionality, necessity, and EU AI Act restrictions on biometric identification.

99. A — Each proposed reduction should be evaluated individually against the current risk profile. Extended audit intervals may be supported by the track record for a stable system. However, reducing monitoring frequency and eliminating human review for high-value lending decisions increases risk exposure for a high-risk system that directly affects individuals' financial rights.

100. B — The unifying principle across all four AIGP domains is ensuring AI systems serve human values throughout their lifecycle — understanding what they are, knowing the laws that constrain them, governing development rigorously, and maintaining vigilant oversight through deployment, use, and retirement. Every governance activity ultimately serves the goal of beneficial AI with managed risk.