

Test Tema 105 #1

Actualizado el 13/04/2025

1. El objetivo principal de Google es:

- a) maximizar el Índice de Precisión
- b) maximizar el Índice de Respuesta
- c) minimizar el Silencio
- d) B y C

2. En los Sistemas de Recuperación de Información se define:

- a) El Índice de Retorno, que mide la cantidad de material devuelto, y el Índice de Precisión, que mide el ajuste del material devuelto.
- b) El Índice de Devoluciones, que mide la cantidad de material devuelto, y el Índice de Calidad, que mide el ajuste del material devuelto.
- c) El Índice de Datos, que mide la cantidad de material devuelto, y el Índice de Errores, que mide el desajuste del material devuelto.
- d) El Índice de Retorno, que mide la cantidad de material devuelto y el Índice de Calidad, que mide el ajuste del material devuelto.

3. En el contexto de bases de datos SQL Server o PostgreSQL, el concepto de "Log Shipping" es equivalente a:

- a) Copia automática asíncrona de base de datos.
- b) Intercambio de información para el mantenimiento del quorum de cluster.
- c) Agregación centralizada de logs de auditoría de base de datos.
- d) Copia síncrona de base de datos.

4. ¿Cuál es el ciclo de vida correcto de la información?

- a) Generación, Almacenamiento, Acceso, Disposición, Uso y Conservación o Destrucción
- b) Generación, Acceso, Almacenamiento, Uso, Disposición y Conservación o Destrucción
- c) Generación, Acceso, Disposición, Almacenamiento, Uso y Conservación o Destrucción
- d) Generación, Almacenamiento, Acceso, Uso, Disposición, y Conservación o Destrucción

5. En la recuperación de la información, los modelos clásicos basados en aproximaciones estadísticas son:

- a) Booleano
- b) Vectorial
- c) Probabilístico
- d) Todos los anteriores son modelos clásicos

6. En un Sistema de Gestión de Bases de Datos Documentales el diccionario de palabras vacías permite:

- a) Buscar palabras o combinaciones de las mismas sin conocer con exactitud los caracteres que las componen
- b) Enlazar consultas individuales para crear rutinas de interrogación complejas
- c) Definir un conjunto de palabras que no deseamos que formen parte de los índices
- d) Igualar caracteres para que el sistema de búsqueda los considere iguales

7. En un cluster activo-pasivo:

- a) Todos los nodos del cluster reciben la misma parte de carga.
- b) Disponemos de dos nodos con el mismo servicio a los que los clientes se pueden conectar indistintamente de forma simultánea.
- c) Los clientes se conectan al activo y los procesos de menor prioridad se ejecutan en el pasivo.
- d) La carga la asume uno de los nodos y en caso de caída se traspasa a otro nodo.

8. En el campo de la Recuperación de Información, ¿cuál de los siguientes es un ejemplo de modelo estructurado, que tratan de combinar tanto el contenido de los documentos como el conocimiento sobre su estructura?:

- a) Nodos próximos.
- b) Lógica borrosa.
- c) Vectorial generalizado.
- d) Redes de inferencia.

9. El tipo de indización por descriptores es el más utilizado en la indización asistida por ordenador, y puede realizarse mediante distintos métodos. ¿Cuál de los reseñados no sería un método de indización?

- a) El método estadístico
- b) El método por asignación de conceptos claves del documento
- c) El método sintáctico
- d) El método morfológico

10. En los Sistemas de Recuperación de Información la lógica borrosa permite:

- a) Indexar parcialmente un documento mediante dos o más términos.
- b) Clasificar los textos con arreglo a su relevancia estimada.
- c) Técnicas de muestreo de la cuantificación del grado de relevancia de las claves.
- d) Generar varias claves sin tener en cuenta el Tesauro.

11. La expresión 'Information Lifecycle Management' (o 'ILM') se corresponde con un concepto vinculado a procesos automatizados, que permiten la gestión eficiente del almacenamiento de la información de una organización. ¿Qué técnica se emplea habitualmente dentro de un esquema ILM para aumentar la capacidad y disponibilidad de la información?

- a) Técnicas de resizing.
- b) Utilización de punteros o enlaces a los elementos de información para evitar duplicidades.
- c) Transformación de los datos a formatos vectoriales.
- d) Purgado de información mediante procesos de borrado automático.

12. Cuál de los siguientes no es un método o técnica de recuperación de información:

- a) Lenguajes recuperativos.
- b) Retroalimentación.
- c) Modelos vectoriales y probabilísticos.
- d) Sistemas booleanos.

13. En backups, ¿qué indica el RTO?

- a) Nos indica la cantidad máxima de información que podemos perder.
- b) Nos indica la cantidad máxima de información que podemos salvaguardar.
- c) Nos indica el tiempo máximo de recuperación.
- d) Nos indica el tiempo promedio en creación de backups.

14. En los Sistemas de Recuperación de Información, la precisión se define como la relación:

- a) Documentos relevantes recuperados frente a Documentos relevantes.
- b) Documentos relevantes recuperados frente a Documentos recuperados.
- c) Documentos no relevantes recuperados frente a Documentos recuperados.
- d) Documentos relevantes conocidos recuperados frente a Documentos relevantes conocidos.

15. En el contexto de la disponibilidad de la información es FALSO:

- a) Un RPO (Recovery Point Objective) alto minimiza el impacto de un desastre en la organización.
- b) RTO (Recovery Time Objective) representa cuánto tarda la restauración desde el incidente hasta que las operaciones normales estén disponibles para los usuarios.
- c) RPO (Recovery Point Objective) viene dado por el tiempo desde la última copia de la información.
- d) RPO (Recovery Point Objective) define la cantidad máxima permitida de datos perdidos, medidos en el tiempo, desde la ocurrencia de un fallo, hasta la fecha última de backup válido.

16. El tipo de indización por descriptores consistente en utilizar técnicas de análisis morfológico y semántico para captar la estructura del texto, se denomina:

- a) Método estadístico
- b) Método por asignación de los conceptos clave del documento
- c) Método sintáctico
- d) Ninguna de las anteriores respuestas es correcta

17. ¿Qué técnica de recuperación asigna importancia relativa a los términos de la consulta, grados de indexación a los términos de cada documento, y puede modificar el rigor de los operadores "o" e "y", mediante los valores de una variable?

- a) Norma-Gamma
- b) Lógica Borrosa
- c) Índices invertidos
- d) Recuperación automática vectorial

18. Los métodos de recuperación de información se basan fundamentalmente en:

- a) Manipulación de volúmenes de información muy pequeños
- b) Identificación de claves (palabras, fragmentos o frases) y su búsqueda en la base de información o ficheros auxiliares
- c) Tratamiento de información mediante sistemas hipertexto
- d) Grabación y recuperación de imágenes

19. Se define Índice de pertinencia (un valor entre 0 y 1) como:

- a) la relación entre el número de documentos relevantes recuperados y todos los documentos recuperados en una búsqueda
- b) la relación entre el número de documentos relevantes recuperados y todos los documentos relevantes de la base de datos en una búsqueda
- c) 1 menos el Ruido
- d) Son correctas A y C

20. Los métodos vectoriales y probabilísticos de recuperación de información se basan:

- a) Álgebra de Boole y frecuencia de términos respectivamente
- b) Sistemas de retroalimentación y lenguaje natural respectivamente
- c) Indexación y recuperación vectorial
- d) Frecuencia de aparición de términos y teorema de Bayes

21. La expresión "Information Lifecycle Management" (o "ILM") se corresponde con un concepto vinculado a la gestión eficiente en el almacenamiento de la información de una organización. ¿Qué técnica NO se emplea habitualmente dentro de un esquema ILM para aumentar la capacidad y disponibilidad de la información?

- a) Técnicas de compresión.
- b) Utilización de punteros o enlaces a los elementos de información para evitar duplicidades.
- c) Indexación de la información para optimizar la recuperación.
- d) Purgado de información mediante procesos de borrado automático.

22. En informática, el término de recuperación en texto libre se utiliza para describir:

- a) La recuperación en un base de datos que posee índices asociados
- b) La recuperación en un base de datos que almacena el texto íntegro, y donde cada palabra del texto es indexada
- c) La recuperación en un base de datos que almacena el texto íntegro y carece de índices
- d) La recuperación en un base de datos que almacena el texto íntegro, y donde cada palabra del texto es indexada, a excepción de las contenidas en el diccionario de palabras vacías

23. En relación con las estructuras de datos conocidas como listas, señale la afirmación correcta:

- a) La pila es una lista en la que sólo se pueden hacer inserts y updates en su final, llamado "top".
- b) La pila puede ser interpretada como una lista FIFO
- c) Se pueden usar pilas para trabajar con operaciones aritméticas en notación postfix
- d) Las pilas con las operaciones "push" y "pop" coincidentes se conocen como colas ADT

24. ¿Cuál de los siguientes no es un modelo clásico de recuperación de información?

- a) Probabilístico
- b) Estructurado
- c) Vectorial
- d) Booleano

25. Indique cuál de las siguientes opciones corresponde a un modelo o técnica de recuperación de la información:

- a) Modelo vectorial.
- b) Modelo conceptual.
- c) Modelo de Borh.
- d) Modelo empírico.

26. ¿Qué técnica de recuperación se basa en la observación de las frecuencias de aparición de términos?

- a) Lógica borrosa
- b) Modelos probabilísticos
- c) Modelos vectoriales
- d) Lenguaje natural

27. ¿Cuál es la principal línea de trabajo en el campo de la recuperación de la información?

- a) Aproximaciones semánticas
- b) Estadísticas
- c) Las respuestas a) y b) son verdaderas
- d) Las respuestas a) y b) son falsas

28. Los sistemas de recuperación de información tienen como objeto fundamental:

- a) Recuperación de bases de datos CODASYL y relacionales
- b) Análisis documental
- c) Restaurar ficheros y bases de datos pseudo relacionales
- d) Consultas y modificaciones de bases de datos a través del lenguaje SQL

29. ¿Qué es un 'antidiccionario'?

- a) Un conjunto de palabras de caracterización o palabras clave
- b) Un conjunto de palabras y signos algebraico-matemático que puede encontrarse en los textos
- c) Un conjunto de palabras vacías de significado (artículos, preposiciones...)
- d) Un diccionario en orden alfabético inverso

30. Changed Block Tracking o Resilient Change Tracking:

- a) permite acelerar los tiempos de backups en entornos virtualizados.
- b) permite reducir los tiempos de recuperación de backups en entornos virtualizados.
- c) proporciona aislamiento en el backup.
- d) ninguna de las anteriores.

31. ¿Qué técnica de recuperación se basan en la aplicación del teorema de Bayes?

- a) Lógica borrosa
- b) Modelos estadísticos
- c) Modelos vectoriales
- d) Lenguaje natural

32. La principal diferencia entre las bases de datos documentales respecto a las normales como pueden ser bases de datos relacionales es:

- a) Pueden crearse índices por cualquier campo que se desee
- b) Los índices se almacenan dentro de la base de datos
- c) Pueden crearse índices de más de una palabra
- d) Indexación de partes del contenido de un campo

33. ¿Cuál de las siguientes es un tipo de consulta utilizada para recuperar información en la web?

- a) Consultas de información (informational queries)
- b) Consultas transaccionales (transactional queries)
- c) Consultas de navegación (navigational queries)
- d) Todas las anteriores son consultas válidas

34. En sistemas de recuperación de la información, al cociente entre el número de referencias correctas obtenidas y el número total de documentos recuperados, se denomina:

- a) Índice de retorno
- b) Índice de exactitud
- c) Índice de precisión
- d) Índice de localización

35. Indica la opción en que todos sean CMS (Content Management System)

- a) Drupal, WordPress, Hadoop.
- b) Drupal, Liferay, SonarQube.
- c) Selenium, Cypress, Maven.
- d) Drupal, Joomla, Liferay.

36. ¿Cuál de los siguientes es un algoritmo de ordenación?

- a) Algoritmo dicotómico (Binary)
- b) Algoritmo por montículos (Heapsort)
- c) Algoritmo de búsqueda atrás (Backtracking)
- d) Algoritmo devorador (Greedy)

37. En los Sistemas de Recuperación de Información la técnica de Sistemas Expertos consiste en:

- a) Expresar las consultas como solicitudes de textos relevantes con relación a un tópico particular, produciéndose a continuación la búsqueda en la base de conocimientos mediante activación de reglas, que producen finalmente una relación clasificada de texto
- b) Expresar las consultas como solicitudes de textos relevantes con relación a varios tópicos, produciéndose a continuación la búsqueda en la base de conocimientos mediante activación de reglas, que producen finalmente una relación clasificada de textos
- c) Expresar las consultas como solicitudes de textos relevantes con relación a un tópico particular, produciéndose a continuación la búsqueda en la base de conocimientos mediante consulta de índices invertidos, que producen finalmente una relación clasificada
- d) Expresar las consultas como solicitudes de textos relevantes con relación a un tópico particular, produciéndose a continuación la búsqueda en la base de conocimientos mediante activación de reglas, que producen finalmente una relación de textos presentes

38. En sistemas de recuperación de la información, al cociente entre el número de referencias obtenidas y el número total de referencias existentes, se denomina:

- a) Índice de retorno
- b) Índice de exactitud
- c) Índice de precisión
- d) Índice de localización

39. El tipo de indización por descriptores consistente en realizar una selección de los conceptos más significativos mediante el análisis de frecuencias de los términos existentes, se denomina:

- a) Método estadístico
- b) Método por asignación de los conceptos clave del documento
- c) Método sintáctico
- d) Ninguna de las respuestas anteriores es correcta

40. Para evaluar diferentes sistemas de recuperación de información, ¿qué parámetros suelen analizarse?

- a) Índice de retorno e índice de precisión
- b) Índice de estimación cuantificable
- c) Índice de ajuste de material devuelto
- d) Número de referencias correctas

41. Los ficheros inversos se utilizan preferentemente en:

- a) Bases de datos relacionales.
- b) Bases de datos jerárquicas.
- c) Sistemas de recuperación de información.
- d) Bases de datos reticulares.

42. Entre los estándares que aplican a la conservación de la información en formato electrónico encontramos ISO 19005-1:2005 PDF/A-1, que identifica:

- a) Las recomendaciones para la conservación a largo plazo de información electrónica basada en documentos.
- b) Los requisitos para la conservación a largo plazo de documentos electrónicos.
- c) El formato de archivo de documentos electrónicos para conservación a largo plazo.
- d) Las recomendaciones para la veracidad y fiabilidad de información almacenada electrónicamente.

43. ¿Qué modelo de recuperación de información se basa en el Teorema de Bayes?

- a) Vectorial.
- b) Probabilístico.
- c) Indexación por Semántica Latente.
- d) Basado en la teoría de conjuntos.

44. El método de recuperación de la información basado en ficheros auxiliares de términos (claves o índices de palabras) o ficheros de apuntadores (referencias de los documentos base) es propio de:

- a) Modelo booleano
- b) Indexación y recuperación automática vectorial
- c) Índices invertidos
- d) Retroalimentación

45. El tipo de indización por descriptores consistente en seleccionar conceptos de listas normalizadas de descriptores, se denomina:

- a) Método estadístico
- b) Método por asignación de los conceptos clave del documento
- c) Método sintáctico
- d) Ninguna de las anteriores respuestas es correcta

46. En qué supuesto será el mirroring o la duplicación de datos la estrategia de recuperación más aconsejable en caso de desastre:

- a) Cuando el SDO (Objetivo de Entrega de Servicio) es inferior al SDO habitual.
- b) Cuando la pérdida de datos es aceptable.
- c) Cuando el RPO (Objetivo de Punto de Recuperación) está en minutos.
- d) Cuando el negocio puede aceptar indisponibilidad de servicios de TI.

47. El elemento dentro de un Sistema de Recuperación de la Información que facilita eliminar información irrelevante desde el punto de vista del ordenamiento y la búsqueda y que se realiza durante la fase de ingreso del documento en el sistema, se denomina:

- a) Taxonomía.
- b) Tesauro.
- c) Descriptor o Signatura.
- d) Lista de palabras vacías.

48. Respecto de los sistemas de recuperación de información indicar la respuesta CORRECTA:

- a) Cuando el usuario trata de localizar información útil, está llevando a cabo un ojeo (browsing).
- b) Cuando el usuario esté interesado en un tema inherentemente amplio o que no tiene claramente definido, la interacción con el sistema suele realizarse mediante acciones de búsqueda (retrieval) en lugar de mediante técnicas de ojeo (browsing).
- c) Los sistemas clásicos de Recuperación de Información están orientados en su mayoría a las acciones de búsqueda, mientras que los sistemas basados en hipertextos se ajustan habitualmente a permitir un ojeo rápido.
- d) Los sistemas modernos y los buscadores Web solo utilizan acciones de ojeo (browsing) para proporcionar capacidades de recuperación de información mejoradas.

49. Las Normas Gamma son unas técnicas de recuperación de información:

- a) también conocidas como retroalimentación
- b) donde no es posible modificar el rigor con que se interpretan los operadores "o" é "y"
- c) que asignan valores relativos de importancia a los términos de la consulta y grados de indexación a los términos de cada documento
- d) menos eficaces que las basadas en sistemas booleanos tradicionales

50. Misión principal de un Sistema de Recuperación de Información:

- a) Asistir la búsqueda de textos en bases de datos relacionales para localizar los documentos o referencias interesantes o relevantes ignorando todas las demás.
- b) Asistir la búsqueda de textos en bases de datos documentales o textuales para localizar los documentos o referencias interesantes o relevantes ignorando todas las demás.
- c) Asistir la búsqueda de textos en bases de datos documentales o textuales para localizar todos los documentos o referencias que aparezcan.
- d) Asistir la búsqueda de imágenes en bases de datos relacionales para localizar todos los documentos o referencias que aparezcan.

51. En el ámbito de las copias de seguridad, es CIERTO:

- a) Para recuperar una copia de seguridad incremental solo habrá que comprobar su existencia en dos copias de respaldo: la última copia incremental realizada y la última copia completa realizada.
- b) Una copia de seguridad incremental no almacena todos los ficheros, sino solo los ficheros modificados desde la anterior copia.
- c) Una copia de seguridad incremental no sólo copia los datos que se hayan modificado desde la última copia, si no todos los que se hayan modificado desde la última copia completa realizada.
- d) Una copia de seguridad diferencial solo copia los datos que han variado desde la última copia de respaldo realizada, ya fuera incremental, diferencial o completa.

52. Indicar la respuesta falsa sobre modelos de recuperación de información:

- a) En general, el modelo vectorial se comporta peor que el probabilístico cuando se usa contra colecciones genéricas de documentos
- b) Las redes neuronales son muy útiles en un sistema de recuperación de información por su potencia reconociendo patrones
- c) Las redes de inferencia y las redes de confianza son modelos basados en redes bayesianas
- d) Los modelos estructurados permiten combinar la especificación de consultas con la información de la estructura del documento

53. En los sistemas de recuperación de información full text ¿qué operaciones son válidas para reducir el conjunto de palabras clave?

- a) Eliminación de palabras vacías (stop words)
- b) Identificación de la raíz gramatical (stemming)
- c) Compresión de los conjuntos de palabras obtenidos
- d) Todas las anteriores son operaciones válidas

54. El stemming es:

- a) Una técnica de indexación sintáctica
- b) Una técnica de indexación morfológica
- c) Una técnica de indexación semántica
- d) Una técnica de eliminación de palabras vacías

55. En los Sistemas de Recuperación de Información el método que consiste en que el usuario indica si entre los textos recuperados hay algunos relevantes o no, y el sistema modifica así el vector de búsqueda se denomina:

- a) Retroalimentación.
- b) Modelo booleano.
- c) lenguaje natural.
- d) Lógica borrosa.

56. En el campo de la Recuperación de la Información, el mecanismo de indexación y representación de documentos que los identifica con un conjunto de términos que se suponen independientes entre sí, y que por tanto permiten definir un espacio ortonormal, se corresponde con:

- a) Los modelos booleanos.
- b) Los modelos bayesianos.
- c) Los modelos estructurados.
- d) Los modelos vectoriales.

57.Cuál de los siguientes no es un método o técnica de recuperación de información:

- a) Sistemas de hojéo-ojéo.
- b) Sistemas de fila incierta.
- c) Sistemas de normal.
- d) Sistemas expertos.

58. El método booleano de recuperación de información presenta como problema fundamental:

- a) Es de sintaxis compleja y la consulta debe reintentarse normalmente para obtener una cantidad razonable de documentos
- b) Deben utilizarse múltiples ficheros auxiliares
- c) Utilizan operadores lógicos y modelos probabilísticos simultáneamente
- d) Está basado en sistemas expertos

59. Los sistemas de recuperación de información consistente en rápidas ojeadas de selección buscando referencias en textos sin análisis con profundidad de los contenidos, auxiliándose de señaladores implícitos y explícitos:

- a) Se denominan hojéo-ojéo.
- b) Son relativamente poco complejos.
- c) Todas las anteriores.
- d) Ninguna de las anteriores.

60. El tesauro es un diccionario dinámico de descriptores que pasa las relaciones entre sus términos:

- a) En operadores de jerarquía
- b) En operaciones asociativos
- c) En operadores semánticos y preferenciales
- d) Son correctas todas las respuestas anteriores

61. Una de las características especiales de las bases de datos documentales, con respecto a otros tipos de bases de datos es que:

- a) Están integradas por campos
- b) Son adecuadas para una información muy estructurada
- c) Están basadas en sistemas de archivo y ficheros inversos
- d) Se almacenan en tablas

62. En un Sistema de Gestión Documental la recuperación de la información se produce mediante:

- a) Palabras clave.
- b) Texto completo.
- c) Índices.
- d) Todas las anteriores.

63. Entre los estándares que aplican a la conservación de la información en formato electrónico encontramos ISO 19005-1:2005 PDF/A-1, que identifica:

- a) Las recomendaciones para la conservación a largo plazo de información electrónica basada en documentos.
- b) Los requisitos para la conservación a largo plazo de documentos electrónicos.
- c) El formato de archivo de documentos electrónicos para conservación a largo plazo.
- d) Las recomendaciones para la veracidad y fiabilidad de información almacenada electrónicamente.

64. En el análisis documental al nivel en el que se extrae un "abstract" o resumen analítico del texto íntegro, se denomina:

- a) Nivel de asiento
- b) Nivel de descriptor
- c) Nivel de resumen
- d) Nivel documental