

## IRS BE IT DLO MCQs

Q.1. Commercial institutions are always concerned with

- a. required resources to implement a system
- b. the approach that minimizes development costs increases the salability of their product
- c. the approach that increases the salability of their product
- d. All of the above

Q.2. The major objective of an Information Retrieval System is minimization of human resources required in the finding of needed information to accomplish a task.

- a. True
- b. False

Q.3. In information systems, precision and recall are the key metrics used in evaluations.

- a. True
- b. False

Q.4. Information is composed of

- a. Text (including numeric and date data)
- b. Images,
- c. Audio & Video
- d. All of the above

Q.5. In an Information Retrieval system, the \_\_\_\_\_ aspect has been the only data type that lent itself to full functional processing.

- a. Images
- b. Audio & Video
- c. Text
- d. None of the above

Q.6. The term \_\_\_\_\_ is used to represent the smallest complete unit that is processed and manipulated by the system.

- a. Text
- b. Item
- c. Image

## IRS BE IT DLO MCQs

### d. Audio & Video

Q.7. Which of the following is not classified as an “item”?

- a. Chapter
- b. Article
- c. Document
- d. Image

Q.8. Aspects of information retrieval overhead are \_\_\_\_\_

- a. search composition
- b. search execution
- c. reading non-relevant items
- d. All of the above

Q.9. A primary mechanism for researching sources of needed information and playing a major role in available Information Retrieval Systems are \_\_\_\_\_

- a. Database systems
- b. Structured databases
- c. Libraries
- d. None of the above

Q.10. Images across the Internet are searchable through the following website majorly

- a. Webseek
- b. BBC
- c. Disney
- d. None of the above

Q.11. Images across the Internet are searchable through the following website majorly

- a. Webseek
- b. DITTO.COM
- c. ALTAVISTA/IMAGES

## IRS BE IT DLO MCQs

d. All of the above

Q.12. News organizations such as the \_\_\_\_\_ are processing the audio news they have produced and are making historical audio news searchable via the audio transcribed versions of the news.

a. CNN

b. NBC

c. BBC

d. Fox News

Q.13. Major video organizations such as \_\_\_\_\_ are using video indexing to assist in finding specific images in their previously produced videos to use in future videos or incorporate in advertising.

a. Disney

b. Youtube

c. Vimeo

d. None of the above

Q.14. Information Retrieval exploitation of multi-media is still in its \_\_\_\_\_ with significant theoretical and practical knowledge missing.

a. Maturity

b. Infancy

c. Growth

d. None of the above

Q.15. Example of overhead in IRS can be \_\_\_\_\_

a. query generation

b. query execution

c. scanning results of query to select items to read

d. All of the above

Q.16. A financial advisor recommending a billion dollar purchase of another company needs to be sure that information on the target company has been located and reviewed in writing the recommendation. This is an example of \_\_\_\_\_

a. Sufficient information

b. Significant information

c. Reasonable information

## IRS BE IT DLO MCQs

d. None of the above

Q.17. A student only requires \_\_\_\_\_ references in a research paper to satisfy the expectations of the teacher, which never is all inclusive.

- a. Sufficient information
- b. Significant information
- c. Reasonable information
- d. None of the above

Q.18. A system that supports reasonable retrieval requires fewer features than one which requires comprehensive retrieval.

- a. True
- b. False

Q.19. A system that supports comprehensive retrieval requires fewer features than one which requires reasonable retrieval.

- a. True
- b. False

Q.20. \_\_\_\_\_ retrieval is a negative feature because it overloads the user with more information than is needed.

- a. Comprehensive
- b. Reasonable
- c. Both a & b
- d. None of the above

Q.21. In information retrieval the term “\_\_\_\_\_” item is used to represent an item containing the needed information.

- a. Irrelevant
- b. Relevant
- c. Both a & b
- d. None of the above

Q.22. Check whether the formula for Precision is True or False.

Precision =  $\text{Number\_Retrieved\_Relevant} / \text{Number\_Total\_Retrieved}$

## IRS BE IT DLO MCQs

- a. True
- b. False

Q.23. Check whether the formula for Recall is True or False.

Recall = Number\_Retrieved\_Relevant/ Number\_Possible\_Relevant

- a. True
- b. False

Q.24. Check whether the formula for Precision is True or False.

Precision = Number\_Retrieved\_Relevant/ Number\_Possible\_Relevant

- a. True
- b. False

Q.25. Check whether the formula for Recall is True or False.

Recall = Number\_Retrieved\_Relevant/ Number\_Total\_Retrieved

- a. True
- b. False

Q.26. \_\_\_\_\_ is a very useful concept, but due to the denominator, is non-calculable in operational systems.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.27. \_\_\_\_\_ measures one aspect of information retrieval overhead for a user associated with a particular search.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.28. \_\_\_\_\_ gauges how well a system processing a particular query is able to retrieve the relevant items that the user is interested in seeing.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

## IRS BE IT DLO MCQs

Q.29. \_\_\_\_\_ starts off at 100 per cent and maintains that value as long as relevant items are retrieved.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.30. \_\_\_\_\_ starts off close to zero and increases as long as relevant items are retrieved until all possible relevant items have been retrieved.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.31. \_\_\_\_\_ is directly affected by retrieval of non-relevant items and drops to a number close to zero.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.32. \_\_\_\_\_ is not effected by retrieval of non-relevant items and thus remains at 100 percent once achieved.

- a. Precision
- b. Recall
- c. Both a and b
- d. None of the above

Q.33. Examples of Information Retrieval system are \_\_\_\_\_

- a. RetrievalWare
- b. TOPIC
- c. AltaVista
- d. All of the above

Q.34. Examples of Information Retrieval system are \_\_\_\_\_

## IRS BE IT DLO MCQs

- a. Infoseek
- b. INQUERY
- c. Both a & b
- d. None of the above

Q.35. For \_\_\_\_\_, however, the retrieved objects might be inaccurate and small errors are likely to go unnoticed.

- a. information retrieval system
- b. data retrieval system
- c. Both a & b
- d. None of the above

Q.36. A \_\_\_\_\_ (such as a relational database) deals with data that has a well-defined structure and semantics.

- a. information retrieval system
- b. data retrieval system
- c. Both a & b
- d. None of the above

Q.37. \_\_\_\_\_, while providing a solution to the user of a database system, does not solve the problem of retrieving information about a subject or topic.

- a. information retrieval system
- b. data retrieval system
- c. Both a & b
- d. None of the above

Q.38. The primary goal of an \_\_\_\_\_ is to retrieve all the documents which are relevant to a user query while retrieving as few non-relevant documents as possible.

- a. information retrieval system
- b. data retrieval system
- c. Both a & b
- d. None of the above

Q.39. Both retrieval and browsing are, in the language of the World Wide Web, \_\_\_\_\_ actions.

- a. Pushing
- b. Pulling
- c. Both a & b

## IRS BE IT DLO MCQs

d. None of the above

Q.40. Retrieval done in an automatic and permanent fashion using software agents, use\_\_\_\_\_ technique to the information towards the user.

a. Push

b. Pull

c. Both a & b

d. None of the above

Q.41. "Information useful to a user could be extracted periodically from a news service". This is an example of \_\_\_\_\_

a. Push

b. Pull

c. Both a & b

d. None of the above

Q.42. Defining the text database includes the following except, \_\_\_\_\_

a. the documents to be used

b. the operations to be performed on the text

c. the text model

d. search method

Q.43. The text operations transform the original documents and generate a logical view of them.

a. True

b. False

Q.44. An \_\_\_\_\_ is a critical data structure because it allows fast searching over large volumes of data.

a. List

b. Index

c. Set

d. None of the above

Q.45. \_\_\_\_\_ process considers query as a document.

a. Matching

## IRS BE IT DLO MCQs

- b. Mapping
- c. Both a and b
- d. None of the above

Q.46. Document surrogates are the \_\_\_\_\_ of full document.

- a. Limited representation
- b. Predefined structure
- c. Classification
- d. All of the above

Q.47. The users generate the document and query in \_\_\_\_\_

- a. Ectosystem
- b. Endosystem
- c. Both a and b
- d. None of the above

Q.48. A query (A AND (NOT B) OR (B AND (NOT A))), specifies exclusive retrieval, that is, or either A where B is not present or B where A is not present. Which of the following hubs has administration and diagnosis capabilities?

- a. A is not present where B is present in a document
- b. A is present where B is not present in a document
- c. Both a and b
- d. Both A and B are present in the document

Q.49. The matching process is complicated because\_\_

- a. Both document and query are in different forms
- b. Query is shorter as compared to document
- c. The document is not organized
- d. None of the above

Q.50. SGL stands for \_\_\_\_\_

- a. System generalized markup language
- b. Standard generalized markup language
- c. Static generalized markup language

## IRS BE IT DLO MCQs

d. None of the above

Q.51. The gross structure of a document affects the way the document is \_\_\_\_\_ and \_\_\_\_\_.

- a. Matched and mapped
- b. Stored and processed
- c. Stored and accessed
- d. Organized and processed

Q.52. Static data model provides faster compression and decompression because \_\_\_\_\_

- a. Less document size
- b. Less computation
- c. Document does not change
- d. None of the above

Q.53. The effect of data compression depends on \_\_\_\_\_ and \_\_\_\_\_.

- a. Level of compression and data model
- b. Data model and compression technique
- c. Type of document and data model
- d. Data model and type of document

Q.54. The implicit structure of a document is used for \_\_\_\_\_

- a. Matching
- b. Encoding
- c. Information retrieval
- d. Error checking

Q.55. An \_\_\_\_\_ is a set of processes and procedures that transform data into information and knowledge.

- a. Information system
- b. Knowledge system
- c. Database system
- d. Computer system

## IRS BE IT DLO MCQs

Q.56. The \_\_\_\_\_ is defined as a set of activities performed across the organization creating as an output of value to the customer.

- a. Development process
- b. Business process
- c. Quality process
- d. Customer focus

Q.57. \_\_\_\_\_ ensures that systems are developed on time, within budget and with acceptable quality.

- a. Systems designer
- b. Project manager
- c. Systems owner
- d. External system user

Q.58. Which one of the following is not a business driver for an information system?

- a. Business process redesign
- b. Knowledge asset management
- c. The proliferation of networks and the Internet.
- d. Security and privacy

Q.59. Which of the following is not a technology driver for an information system?

- a. Enterprise applications
- b. Objective technologies
- c. Knowledge assessment and management
- d. Networks and Internet

Q.60. Information is \_\_\_\_\_

- a. Data
- b. Processed data
- c. Manipulated input
- d. Computer output

Q.61. Data by itself is not useful unless:

## IRS BE IT DLO MCQs

- a. It is massive
- b. It is processed to obtain information
- c. It is collected from diverse sources
- d. It is properly stated

Q.62. For taking decisions, data must be:

- a. Very accurate
- b. Massive
- c. Processed correctly
- d. Collected from diverse sources

Q.63. The organized set of documents is called \_\_\_\_\_

- a. Vector space
- b. Document space
- c. Both a and b
- d. None of the above

Q.64. Although the matching process is easy, it has not proven to be successful because \_\_\_\_\_

- a. User-centered view is not considered
- b. Use of uncontrolled vocabulary
- c. Huge document collection
- d. All of the above

Q.65. The process of reducing a large document collection into a reasonably small-sized set of potentially retrievable documents is called as \_\_\_\_\_

- a. Information storage
- b. Information processing
- c. Information filtering
- d. Information retrieval

Q.66. The \_\_\_\_\_ modifies the calculation on which relevance decisions are made.

- a. The weighting of terms
- b. Probability of terms

## IRS BE IT DLO MCQs

- c. Both a and b
- d. None of the above

Q.67. The \_\_\_\_\_ works on the principle that for the given document and query the probability can be calculated such that the document is relevant to the query.

- a. Boolean matching
- b. Fuzzy matching
- c. Probabilistic matching
- d. Proximity matching

Q.68. The drawback of the probabilistic retrieval model is \_\_\_\_\_

- a. It involves many calculations and assumptions
- b. Not suitable for large databases
- c. Probabilistic model is very rigid
- d. None of the above

Q.69. \_\_\_\_\_ and \_\_\_\_\_ are qualitative and quantitative fuzzy descriptors.

- a. Small and fairly important
- b. Tall and highly relevant
- c. Colorful and partially relevant
- d. All of the above

Q.70. The effects of \_\_\_\_\_ are that the relative proportion of the documents which are appropriate in response to a given query from a large collection which is more general in nature is very small.

- a. Conditional probability
- b. Order of proximity
- c. Weighting terms
- d. Scaling

Q.71. \_\_\_\_\_ is the major problem of databases in data fusion.

- a. Different types of databases
- b. Matching process
- c. Standardization
- d. Types of search engines

## IRS BE IT DLO MCQs

Q.72. The fundamental operation in an information retrieval system is \_\_\_\_\_ in both the document and the query.

- a. Information processing
- b. Matching the terms
- c. Assigning probabilities
- d. Assigning weights to terms

Q.73. Which of the following is the disadvantage of the uncontrolled vocabulary?

- a. Lack of consistency
- b. Multiple index terms
- c. No flexibility
- d. Vocabulary list

Q.74. The word proximity depends on

- a. A number of intermediate words
- b. Total words in the same document
- c. Vocabulary list
- d. Indexing language

Q.75. Stemming a large document is \_\_\_\_\_

- a. Very easy
- b. Time consuming
- c. Not possible
- d. Changes the original document

Q.76. Stemming strips off \_\_\_\_\_

- a. Suffixes
- b. Suffixes and prefixes
- c. Prefixes
- d. Sentences

Q.77. The advantage of automatic indexing is \_\_\_\_\_

- a. Very fast

## IRS BE IT DLO MCQs

- b. Large document can be indexed easily
- c. Flexibility to the user to choose index terms
- d. To locate document easily

Q.78. The matrices used to represent the document-term relationship cannot be used directly because

- a. Difficult to process
- b. Matrices are sparse
- c. Size of the matrices
- d. Complication of the matrices

79. Which of the following is not a technology driver for an information system?

- a. Enterprise applications
- b. Objective Technologies
- c. Knowledge assessment and management
- d. Networks and Internet

80. Information is \_\_\_\_\_

- a. Data
- b. Processed data
- c. Manipulated input
- d. Computer output

81. Data by itself is not useful unless:

- a. It is massive
- b. It is processed to obtain information
- c. It is collected from diverse sources
- d. It is properly stated

82. For taking decisions, data must be:

- a. Very accurate
- b. Massive
- c. Processed correctly
- d. Collected from diverse sources

83. \_\_\_\_\_ is a set of processes and procedures that transform data into information and knowledge.

- a. Information system
- b. Knowledge system
- c. Database system
- d. Computer system

84. Which of the following statement is incorrect? [2 Marks]

## IRS BE IT DLO MCQs

- a. The major objective of an Information Retrieval System is minimization of human resources required in the finding of needed information to accomplish a task.
- b. In information systems, precision and recall are the key metrics used in evaluations.
- c. A system that supports reasonable retrieval requires fewer features than one which requires comprehensive retrieval.
- d. A system that supports comprehensive retrieval requires fewer features than one which requires reasonable retrieval.

85. In information retrieval the term “ \_\_\_\_\_ ” item is used to represent an item containing the needed information. [2 Marks]

- a. Irrelevant
- b. Relevant
- c. Sufficient
- d. Insufficient

86. Which of the following formula for Precision is correct? [2 Marks]

- a.  $\text{Precision} = \frac{\text{Number\_Retrieved\_Relevant}}{\text{Number\_Total\_Retrieved}}$
- b.  $\text{Precision} = \frac{\text{Number\_Retrieved\_Relevant}}{\text{Number\_Possible\_Relevant}}$
- c.  $\text{Precision} = \frac{\text{Number\_Received\_Relevant}}{\text{Number\_Possible\_Relevant}}$
- d.  $\text{Precision} = \frac{\text{Number\_Received\_Relevant}}{\text{Number\_Total\_Retrieved}}$

87. Which of the following formula for Recall is correct? [2 Marks]

- a.  $\text{Recall} = \frac{\text{Number\_Retrieved\_Relevant}}{\text{Number\_Total\_Retrieved}}$
- b.  $\text{Recall} = \frac{\text{Number\_Retrieved\_Relevant}}{\text{Number\_Possible\_Relevant}}$
- c.  $\text{Recall} = \frac{\text{Number\_Received\_Relevant}}{\text{Number\_Possible\_Relevant}}$
- d.  $\text{Recall} = \frac{\text{Number\_Received\_Relevant}}{\text{Number\_Total\_Retrieved}}$

88. A query (A AND (NOT B) OR (B AND (NOT A))), specifies exclusive retrieval, that is, or either A where B is not present or B where A is not present. Which of the following hubs has administration and diagnosis capabilities? [2 Marks]

- a. A is not present where B is present in a document
- b. A is not present where B is not present in a document
- c. A is present where B is not present in a document
- d. Both A and B are present in the document

89. \_\_\_\_\_ and \_\_\_\_\_ are qualitative and quantitative fuzzy descriptors. [1 Mark]

- a. Tiny and fairly important
- b. Large and highly relevant
- c. Colorful and fully relevant
- d. Colorful and partially relevant

## IRS BE IT DLO MCQs

90. The drawback of the probabilistic retrieval model is \_\_\_\_\_ [1 Mark]

- a. It involves many calculations and assumptions
- b. Not suitable for large databases
- c. Probabilistic model is very rigid
- d. Probabilistic model is very simple

91. Static data model provides faster compression and decompression because \_\_\_\_\_ [1 Mark]

- a. Less document size
- b. Less computation
- c. Document does not change
- d. Document changes frequently

92. The matching process is complicated because \_\_\_\_

- a. Both document and query are in different forms
- b. Query is shorter as compared to document
- c. The document is not organized
- d. Both document and query are in the same form

93. Retrieval done in an automatic and permanent fashion using software agents, use \_\_\_\_\_ technique to the information towards the user.

- a. Push
- b. Pull
- c. Push as well as Pull
- d. Neither Push nor Pull