

Semi-structured data ...

- A. is always schema-less
- B. always embeds schema information into the data
- C. must always be hierarchically structured
- D. can never be indexed

Answer B

Answer A: there exist frameworks for providing schemas for semi-structured data, such as XML Schema, or RDF schema we will introduce.

Answer C: hierarchical structure is very common for semi-structured data, but not necessary, as for example for email formats.

Answer D: Indexing of semi-structured data is possible, but cannot rely on a fixed data structure. Therefore the approaches are different from the ones used typical in databases with structure schemas.

Why is XML a document model and not just a general data model?

- A. It supports application-specific markup
- B. It supports domain-specific schemas
- C. It has a serialized representation
- D. It uses HTML tags

Answer C: the notion of “document” is related to the fact that the data is represented as a string. Statements A and B are also correct for XML, but do not justify the notion of document. More accurately, we are referring to text documents here.

An ontology ...

- A. helps to separate layout issues from the structural representation of data**
- B. defines a common syntactic framework to represent standardized domain models**
- C. can be used as a mediation framework for integrating semantically heterogeneous databases**

Answer C

Answer A refers to the role of structured document models, such as XML. An ontology is a priori also independent of its specific syntactic representation. It can be used to integrate data from semantically heterogeneous resources.

A basic statement in RDF would be expressed in the relational data model by a table ...

- A. with one attribute
- B. with two attributes
- C. with three attributes
- D. cannot be expressed in the relational data model

Answer B

The canonical representation of an RDF statement in a relational table would be by a table that represents the predicate of the statement and two attributes that are used to represent the subject and object. For example, the statement *X isauthor Y* would be represented as tuple *isauthor(X, Y)*. One might argue that also the predicate might be represented as attribute. This would be resulting in a generic representation with single table used to represent all RDF statements. This is not the standard approach of data modeling used in relational data models.

The type statement in RDF would be expressed in the relational data model by a table ...

- A. with one attribute
- B. with two attributes
- C. with three attributes
- D. cannot be expressed in the relational data model

Answer A

The type relation expresses membership to a specific category of elements in a domain. The canonical way to represent this in the relational model is to use a table that lists the elements of the category. For example, if we have elements of type country, and statements of the form X is-a country, we would represent this by tuples of the form country(X). One might argue that one could list the type also as an attribute in a binary table. However, then this corresponds rather to a specific property than a type information. For example, if we have elements of type female and male, we would have a table gender(X, F/M), but we would consider the gender not as type information.

Which is true?

- A. Reification is used to produce a more compact representation of complex RDF statements
- B. Reification requires to make a statement the subject of another statement
- C. Reified statements are always used to make a statement about another statement

Answer C

Answer A is obviously the opposite what happens. As for Answer B, a statement about which we make another statement can also be the object of the statement. For example, author X produced the statement S.

Which of the following are part of the RDF schema language?

- A. The « type » statement for RDF resources?
- B. The « domain » statement for RDF properties?
- C. The « subject » statement for RDF statements?

Answer B

You have noticed that the namespace for the type and subject statements is RDF. This makes sense since these two types of statements involve instances that are not part of an RDF schema, whereas the domain statement involves exclusively schema instances.

Which of the following is NOT an (instance-level) ontology?

- A. Wordnet
- B. WikiData
- C. Schema.org
- D. Google Knowledge Graph

Answer C

Schema.org provides a schema without any instances of facts, whereas the three other resources provide sets of facts.