Student Name: _____   Date: 3 Mar 2016
Student ID: _____   Time: 11:15AM to 11:30AM

Total number of questions: 8
Each question has a single answer!

1. Which one of these data management techniques **CANNOT** be used to reduce data access latency?

   ☐ *a*) Data partitioning

   ☐ *b*) Data replication

   ☐ *c*) Data caching

   ☒ *d*) All of them can be used to reduce access latency

2. What is a possible approach to reconcile semantically heterogeneous models in an information system?

   ☐ *a*) Interpretation   ☒ *c*) Standardization

   ☐ *b*) Validation   ☐ *d*) Documentation

3. Which of the following communication models **CANNOT** be emulated at the software level?

   ☐ *a*) Unicast

   ☐ *b*) Multicast

   ☐ *c*) Broadcast

   ☒ *d*) All of the above can be emulated at the software level

4. A distributed database consists of two nodes, one fully replicating the other. There is one single application running that updates data on one of the two nodes. Which of the following statement is **true**?

   ☒ *a*) The update may compromise consistency

   ☐ *b*) The update may compromise persistency

   ☐ *c*) The update may compromise isolation

   ☐ *d*) All of the above

5. Assume we insert a tuple into a distributed relational database. Which of the following is generally **true**:

   ☐ *a*) The update will be less expensive when replicating the relation to all nodes.

   ☐ *b*) The update will be less expensive when using an index than without index.

   ☒ *c*) The update will be less expensive when using a row store instead of a column store.

   ☐ *d*) The update will be less expensive when using a push communication than with a pull communication.

6. A startup has developed a new method for image classification. It tests its performance by using crowd-sourcing, i.e. many human users (called crowd-workers) check whether the proposed classes make sense. Performing the crowd-sourcing task is:

   ☐ *a*) A retrieval task, because information is retrieved from the crowd-workers.

   ☐ *b*) A data mining task, because classification is a data mining method.

   ☐ *c*) A conceptual modeling task, because the image classification method generates a model.

   ☒ *d*) An evaluation task, because the classification model is compared to human judgment.

7. A system administrator installs a new version of a relational DB. He informs the SQL developers that they have to <mark>rewrite certain queries,</mark> as some features are no longer available. This **violates**:

- □ *a)* Logical data independence
- ⊠ *b)* Physical data independence
- □ *c)* Autonomy
- □ *d)* Isolation

> physical data indepdnence requires that i can use the same query method regardless of updates
> logical data independence requires that the user sees the same data / can access the same data even w updates

8. TV Teletext (currently used in close-captioning in TV channels) is an information dissemination method that uses a push model. Apart from that, it is:

- □ *a)* multi-cast and conditional
- ⊠ *b)* broadcast and conditional
- □ *c)* multi-cast and ad-hoc
- □ *d)* broadcast and ad-hoc