**BÀI TẬP**

**CHƯƠNG 1 – CÁC KHÁI NIỆM VỀ CƠ SỞ DỮ LIỆU**

**NGUYỄN NGỌC TRINH – 12**

**--o0o--**

**I. REVIEW QUESTIONS**

**1. Định nghĩa cho mỗi khái niệm dưới đây:**

a. data

Data là các dữ kiện thô được lưu trữ trên thiết bị lưu trữ

b. field

Field là trường thông tin, dùng để mô tả đặc điểm tính chất của đối tượng

c. record

Record là bản ghi, gồm nhiều trường thông tin của cùng một đối tượng

d. file

File là tập tin, gồm nhiều bản ghi

***Trả lời***

2. Dư thừa dữ liệu là gì? Đặc tính gì của hệ thống tệp có thể dẫn đến việc đó?

- Dư thừa dữ liệu: cùng 1 dữ liệu được lưu trữ ở nhiều nơi khác nhau

- Đặc tính của hệ thống tệp có thể dẫn tới việc dư thừa dữ liệu: không có sự kết nối, tích hợp liên quan giữa các tệp

3. Độc lập dữ liệu là gì? Tại sao lại thiếu trong hệ thống tệp?

- Độc lập dữ liệu: khi dữ liệu thay đổi thì chương trình không thay đổi

- Hệ thống tệp thiếu tính độc lập dữ liệu vì truy nhập vào một tệp phụ thuộc vào cấu trúc của từng tệp, tất cả chương trình phải sửa đổi cho phù hợp với tệp mới

4. Hệ quản trị cơ sở dữ liệu là gì? Nó có những chức năng gì?

Hệ quản trị cơ sở dữ liệu là một phần mềm được thiết kế để quản lý các cơ sở dữ liệu

Chức năng: là trung gian để người dùng kết nối với cơ sở dữ liệu, giúp tích hợp, chia sẻ, bảo mật dữ liệu

5. Độc lập cấu trúc là gì? TẠI SAO nó quan trọng?

Độc lập cấu trúc: sự thay đổi cấu trúc của tệp không làm ảnh hưởng đến truy cập

Nó quan trọng vì: nếu không độc lập cấu trúc, mỗi khi thay đổi lại phải viết lại chương trình

6. Giải thích sự khác nhau của dữ liệu, thông tin và cơ sở dữ liệu.

Dữ liệu là các dữ liệu thô chưa được xử lý, được lưu trữ trên các thiết bị lưu trữ

Thông tin: dữ liệu đã được xử lý

Cơ sở dữ liệu: là tập hợp các dữ liệu có liên quan với nhau về mặt logic nào đó

7. Vai trò của Hệ quản trị cơ sở dữ liệu? Ưu nhược điểm của nó?

Vai trò của hệ quản trị cơ sở dữ liệu: là trung gian giữa người dùng và cơ sở dữ liệu

Ưu điểm: giúp bảo mật, chia sẻ, dự phòng, khôi phục dữ liệu

Nhược điểm: phải có chuyên gia quản lí hệ quản trị, tốn chi phí

8. Liệt kê và mô tả sự khác nhau giữa các loại cơ sở dữ liệu khác nhau.

- Chia theo người dùng:

Chia theo vị trí:

Chia theo chức năng:

Chia theo cấu trúc:

9. Các thành phần chính của hệ thống cơ sở dữ liệu là gì?

Người dùng: người sử dụng, người quản trị, thiết kế, lập trình, người dùng cuối

Phần cứng: các thiết bị

Phần mềm:

Dữ liệu

10. Siêu dữ liệu là gì?

Siêu dữ liệu là dữ liệu của dữ liệu

11. Giải thích tại sao thiết kế cơ sở dữ liệu lại quan trọng?

Để tránh dư thừa dữ liệu, dữ liệu đáp ứng yêu cầu, cung cấp thông tin giá trị hơn

12. Các chi phí khi triển khai hệ thống cơ sở dữ liệu là gì?

13. Sử dụng ví dụ để so sánh sự khác nhau của dữ liệu có cấu trúc và dữ liệu không có cấu trúc. Loại nào phổ biến hơn trong môi trường kinh doanh?

14. Chỉ ra một số chức năng co bản mà bảng tính (trong excel) không thể thực hiện được?

15. Các vấn đề phổ biến của tập các bảng tính được tạo ra bởi người dùng được chia sẻ với hệ thống tệp tin khác?

16. Giải thích tầm quan trọng của việc mất quyền truy cập trực tiếp vào dữ liệu kinh doanh mà người dùng đã lưu so với việc lưu trữ dữ liệu trên cơ sở dữ liệu (kho dữ liệu) trên máy tính

17. Giải thích tại sao chi phí sở hữu có thể thấp hơn với cơ sở dữ liệu đám mây so với cơ sở dữ liệu truyền thống của công ty.

**II. PROBLEMS**



1. How many records does the file contain? How many fields are there per record?

7 records

5 fields

2. What problem would you encounter if you wanted to produce a listing by city? How would you solve this problem by altering the file structure?

Rất khó đưa vì thành phố lẫn trong địa chỉ, phải xử lý xâu

Xử lý bằng cách thêm cột City

3. If you wanted to produce a listing of the file contents by last name, area code, city, state, or zip code, how would you alter the file structure?

Tách cột last name, fisrt name, area code, city, state, zip code riêng ra

4. What data redundancies do you detect? How could those redundancies lead to anomalies?

1 người lặp nhiều địa chỉ

Giải pháp: Tách thành 2 bảng khác nhau

5**. Identify and discuss the serious data redundancy problems exhibited by the file structure shown in Figure P1.5.**



Sự dư thừa dữ liệu

6. Looking at the EMP\_NAME and EMP\_PHONE contents in Figure P1.5, what change(s) would you recommend?

Tách thành 2 bảng

7. Identify the various data sources in the file you examined in Problem 5.

8. Given your answer to Problem 7, what new files should you create to help eliminate the data redundancies found in the file shown in Figure P1.5?

9. Identify and discuss the serious data redundancy problems exhibited by the file structure shown in Figure P1.9. (The file is meant to be used as a teacher class assignment schedule. One of the many problems with data redundancy is the likely occurrence of data inconsistencies two different initials have been entered for the teacher named Maria Cordoza.)



**10. Given the file structure shown in Figure P1.9, what problem(s) might you encounter if building KOM were deleted?**

**11. Using your school’s student information system, print your class schedule. The schedule probably would contain the student identification number, student name, class code, class name, class credit hours, class instructor name, the class meeting days and times, and the class room number. Use Figure P1.11 as a template to complete the following actions.**



a) Create a spreadsheet using the template shown in Figure P1.11 and enter your current class schedule.

b) Enter the class schedule of two of your classmates into the same spreadsheet.

c) Discuss the redundancies and anomalies caused by this design.