Dr Greg Wadley

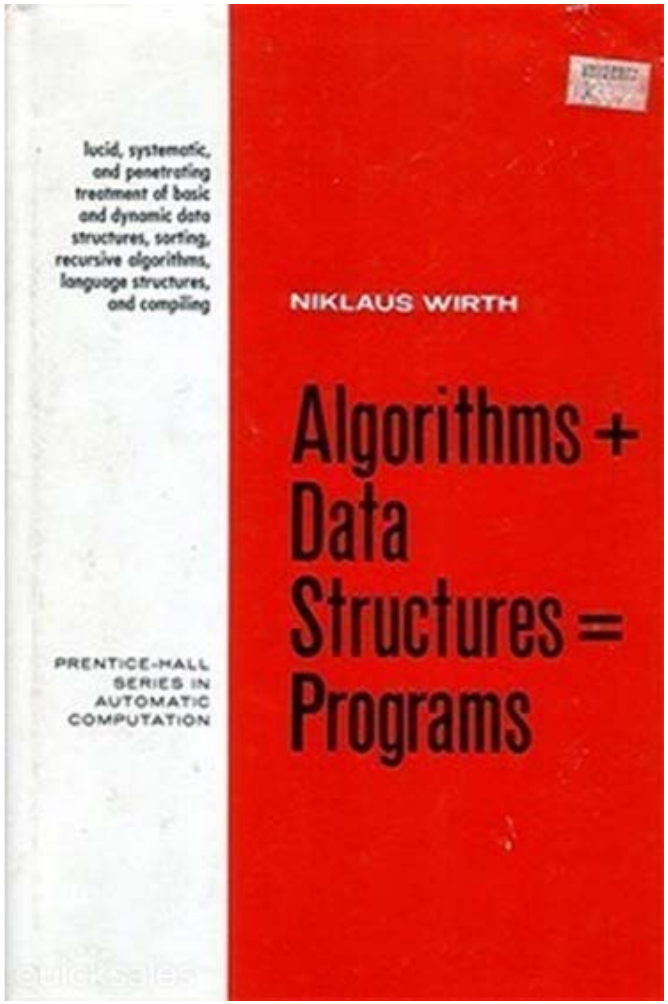

# INFO 90002
# Database Systems & Information Modelling

Week 01

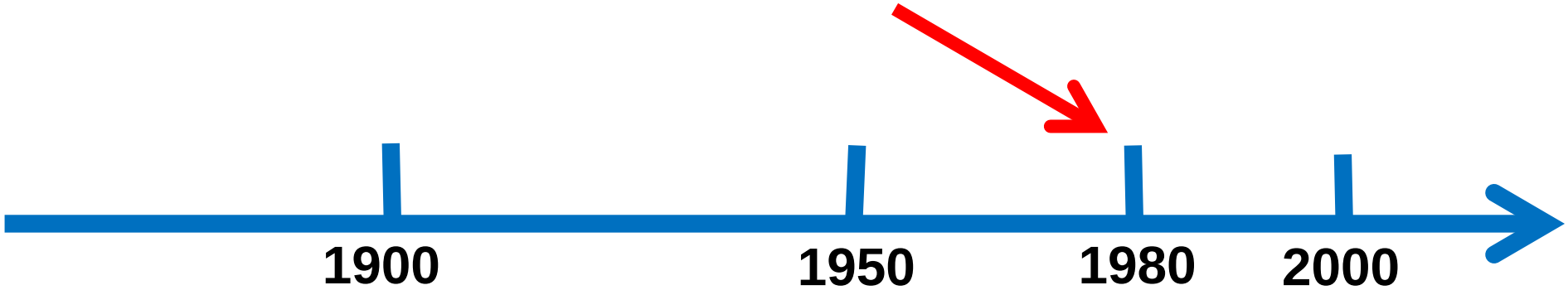
Introduction to Databases

# Context: software and data

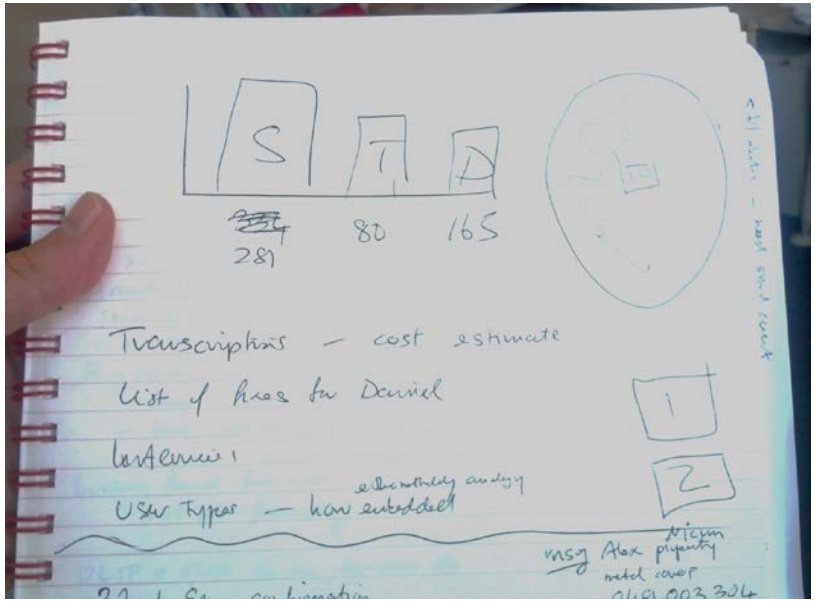
- Computer systems consist of *software* (algorithms) working to process *data*.
- You will learn about creating software and algorithms in other subjects.
- This subject is about *data*.



# Introduction to databases

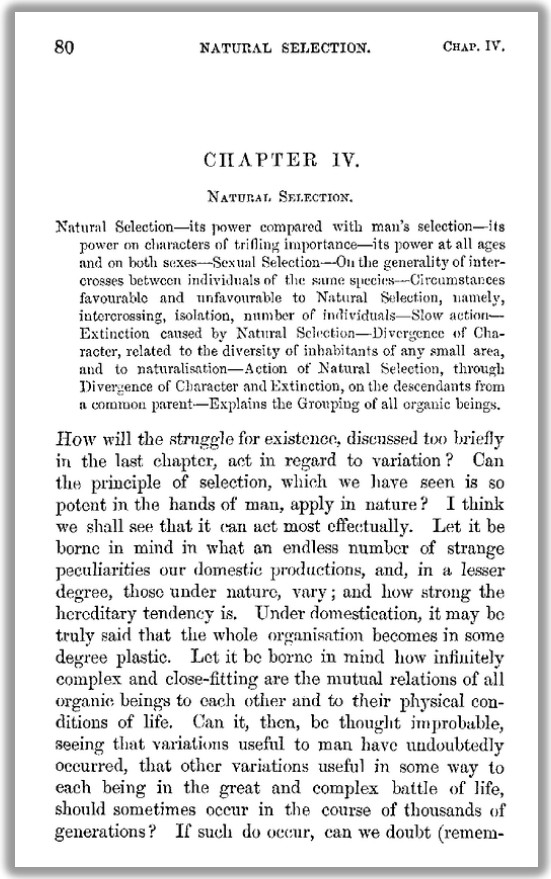



# Unstructured data




- handwritten notes
- printed books
- spreadsheets etc

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | My Holiday Budget -----> | | | | |
| 2 | | | | | |
| 3 | Flights | $10,000.00 | | | |
| 4 | Accommodation | $5,000.00 | <- can this be reduced? | | |
| 5 | Food | $1,000.00 | | | |
| 6 | | | | | |
| 7 | | $16,000.00 | TOTAL | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |



80 NATURAL SELECTION. CHAP. IV.

CHAPTER IV.

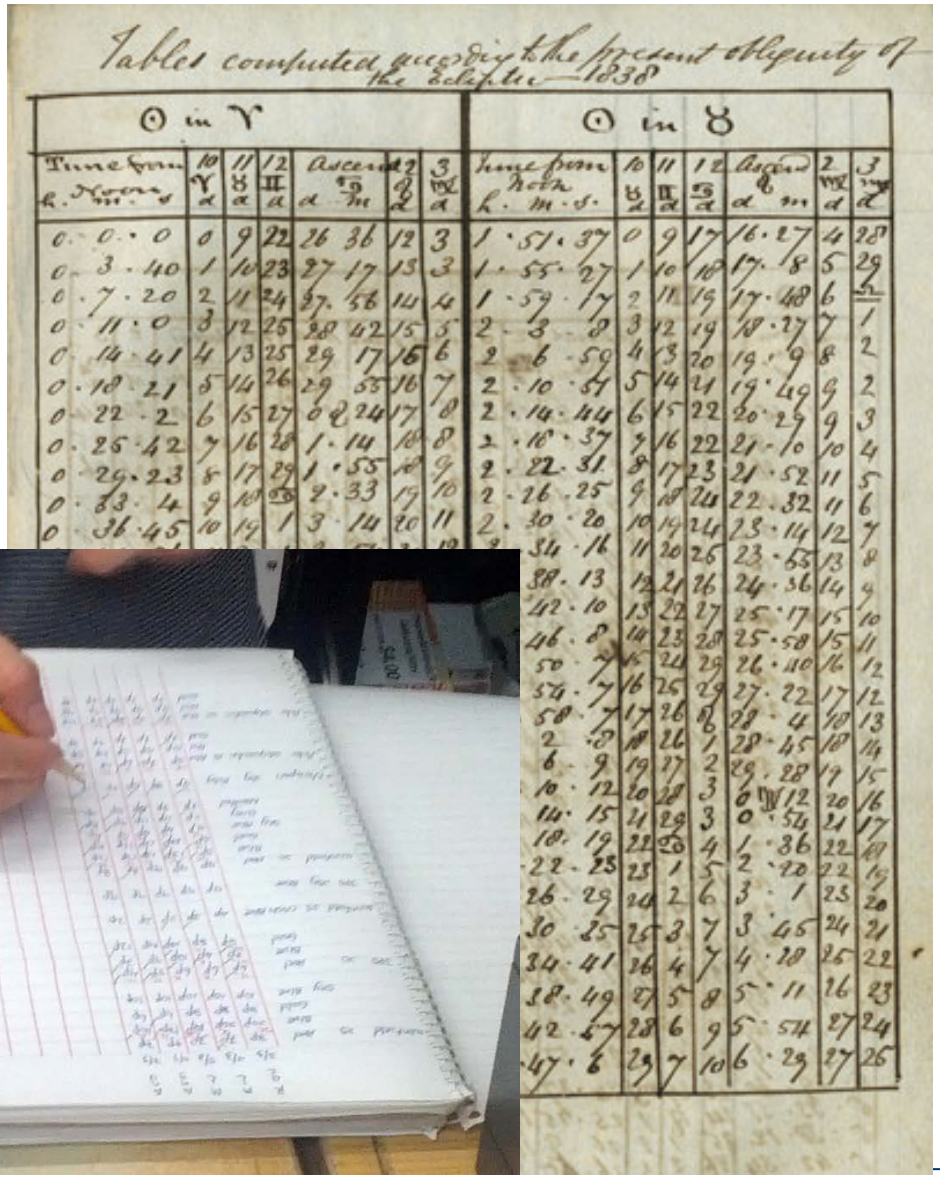
NATURAL SELECTION.

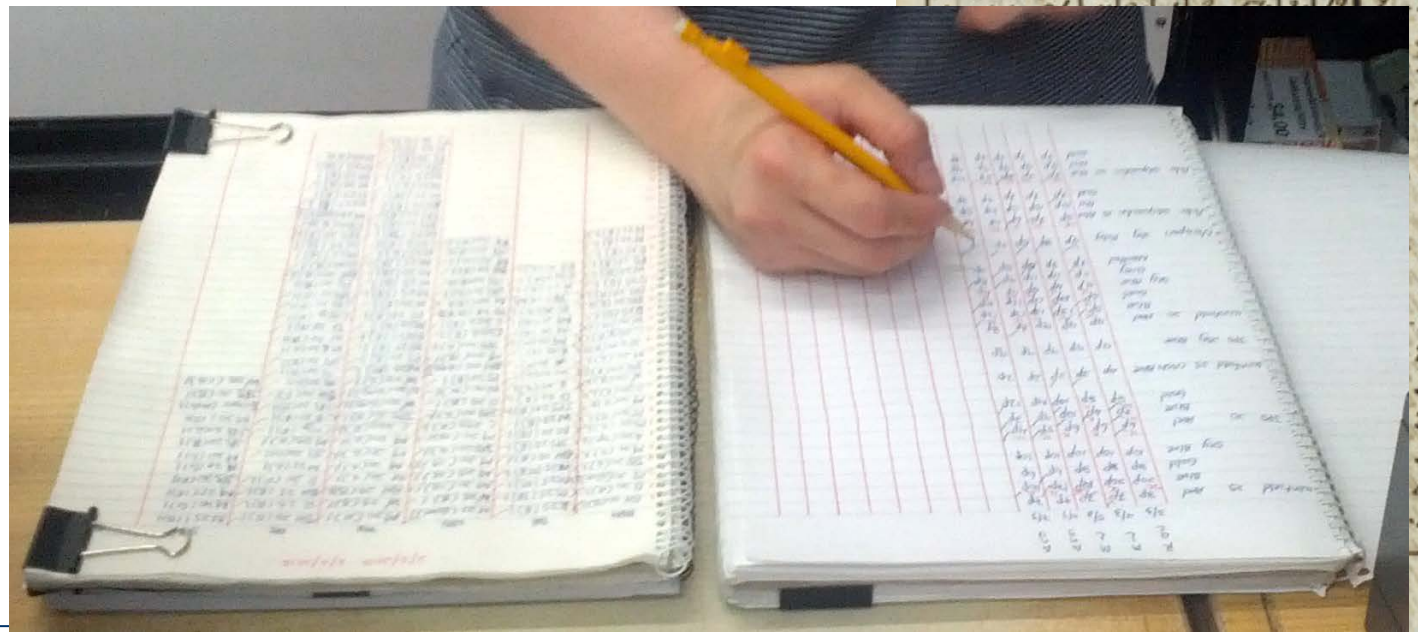
Natural Selection—its power compared with man's selection—its power on characters of trifling importance—its power at all ages and on both sexes—Sexual Selection—On the generality of intercrosses between individuals of the same species—Circumstances favourable and unfavourable to Natural Selection, namely, intercrossing, isolation, number of individuals—Slow action—Extinction caused by Natural Selection—Divergence of Character, related to the diversity of inhabitants of any small area, and to naturalisation—Action of Natural Selection, through Divergence of Character and Extinction, on the descendants from a common parent—Explains the Grouping of all organic beings.

How will the struggle for existence, discussed too briefly in the last chapter, act in regard to variation? Can the principle of selection, which we have seen is so potent in the hands of man, apply in nature? I think we shall see that it can act most effectually. Let it be borne in mind in what an endless number of strange peculiarities our domestic productions, and, in a lesser degree, those under nature, vary; and how strong the hereditary tendency is. Under domestication, it may be truly said that the whole organisation becomes in some degree plastic. Let it be borne in mind how infinitely complex and close-fitting are the mutual relations of all organic beings to each other and to their physical conditions of life. Can it, then, be thought improbable, seeing that variations useful to man have undoubtedly occurred, that other variations useful in some way to each being in the great and complex battle of life, should sometimes occur in the course of thousands of generations? If such do occur, can we doubt (remem-

# Tables

- Each row represents an instance of a set of entities
- Entities might be people, things, events, transactions…
- You can extract useful knowledge via simple, repetitive processes
- A natural fit for business and scientific data.



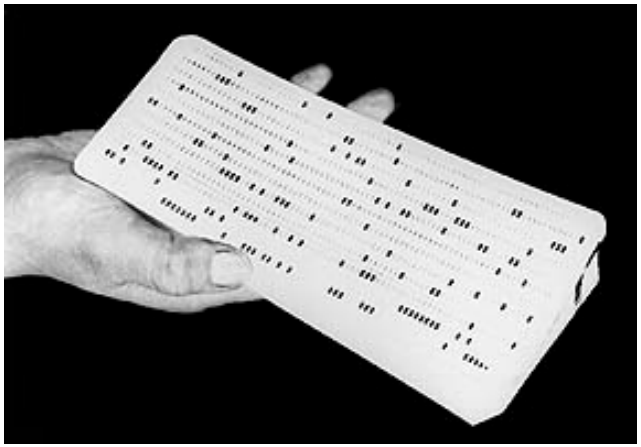



# Table of rows -> Deck of cards

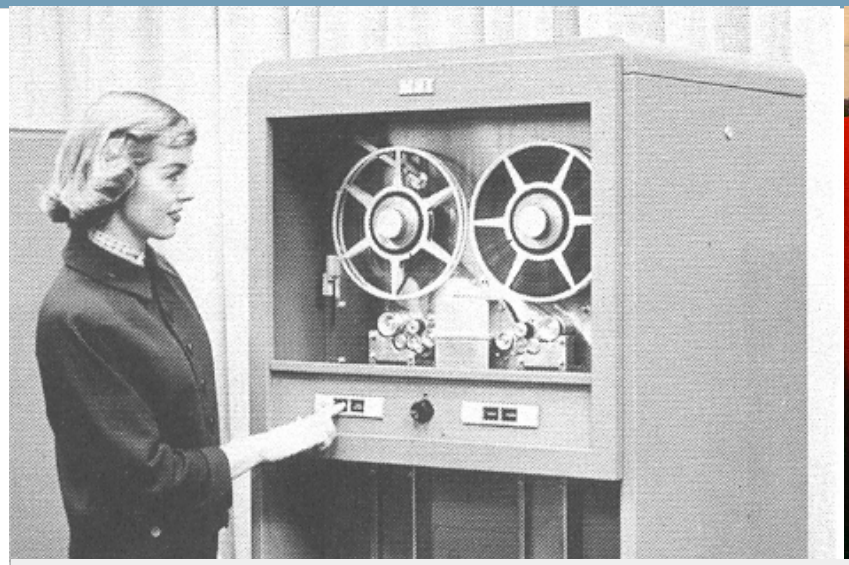
1890 US census

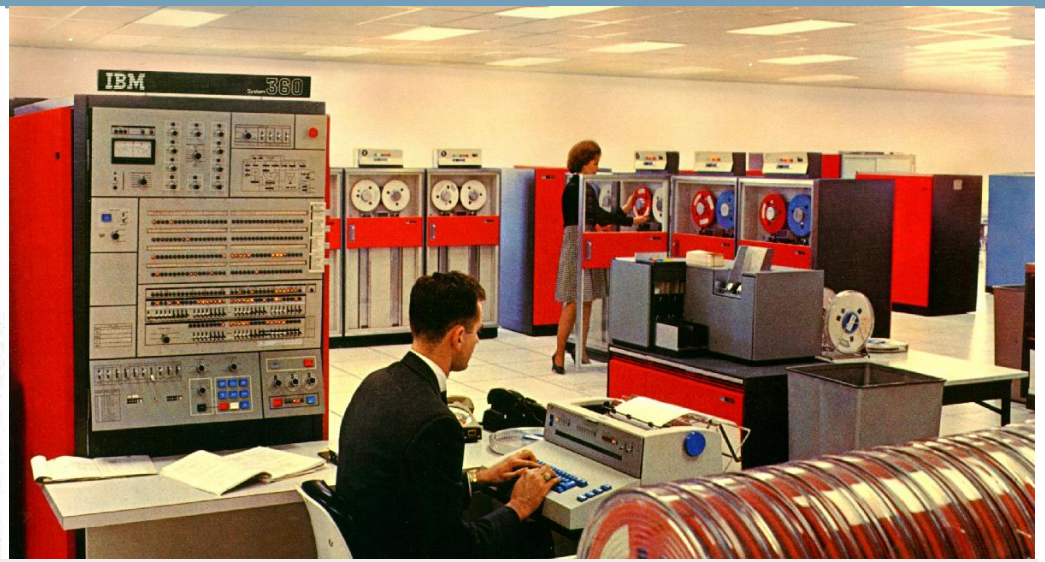
| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | CM | UM | Jp | Ch | Oc | In | 20 | 50 | 80 | Dv | Un | 3 | 4 | 3 | 4 | A | E | L | a | g |
| 5 | 6 | 7 | 8 | CL | UL | O | Mu | Qd | Mo | 25 | 55 | 85 | Wd | CY | 1 | 2 | 1 | 2 | B | F | M | b | h |
| 1 | 2 | 3 | 4 | CS | US | Mb | B | M | 0 | 30 | 60 | 0 | 2 | Mr | 0 | 15 | 0 | 15 | C | G | N | c | i |
| 5 | 6 | 7 | 8 | No | Hd | Wf | W | F | 5 | 35 | 65 | 1 | 3 | Sg | 5 | 10 | 5 | 10 | D | H | O | d | k |
| 1 | 2 | 3 | 4 | Fh | Ff | Fm | 7 | 1 | 10 | 40 | 70 | 90 | 4 | 0 | 1 | 3 | 0 | 2 | St | I | P | e | l |
| 5 | 6 | 7 | 8 | Hh | Hf | Hm | 8 | 2 | 15 | 45 | 75 | 95 | 100 | Un | 2 | 4 | 1 | 3 | 4 | K | Un | f | m |
| 1 | 2 | 3 | 4 | X | Un | Ft | 9 | 3 | i | c | X | R | L | E | A | 6 | 0 | US | Ir | Sc | US | Ir | Sc |
| 5 | 6 | 7 | 8 | Ot | En | Mt | 10 | 4 | k | d | Y | S | M | F | B | 10 | 1 | Gr | En | Wa | Gr | En | Wa |
| 1 | 2 | 3 | 4 | W | R | OK | 11 | 5 | l | e | Z | T | N | G | C | 15 | 2 | Sw | FC | EC | Sw | FC | EC |
| 5 | 6 | 7 | 8 | 7 | 4 | 1 | 12 | 6 | m | f | NG | U | O | H | D | Un | 3 | Nw | Bo | Hu | Nw | Bo | Hu |
| 1 | 2 | 3 | 4 | 8 | 5 | 2 | Oc | 0 | n | g | a | V | P | I | Al | Na | 4 | Dk | Fr | It | Dk | Fr | It |
| 5 | 6 | 7 | 8 | 9 | 6 | 3 | 0 | p | o | h | b | W | Q | K | Un | Pa | 5 | Ru | Ot | Un | Ru | Ot | Un |

https://www.census.gov/history/www/innovations/technology/the_hollerith_tabulator.html



# Tape and Disc





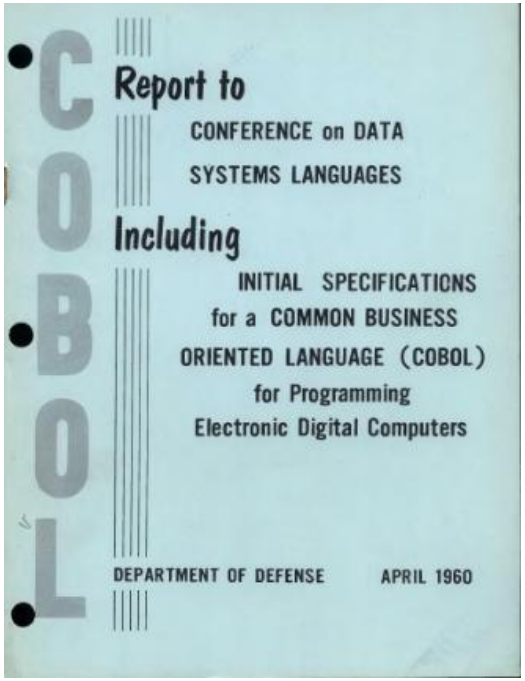

| EMP_RECORD... | EMP_ID | EMP_REGION | EMP_DEPT | EMP_HIRE_D... | E... | E... | EMP_... | EMP_SALARY | EMP_NAME | EMP_SKIL |
|---|---|---|---|---|---|---|---|---|---|---|
| 68 | 3715 | 4 | 153 | 09061987 | 9 | 6 | 1987 | 14000000 | IRENE HIRSH | 041085 |
| 62 | 39412 | 1 | 650 | 03119590 | 3 | 11 | 9590 | 167000000 | ANN FAHEY | 031099 |
| 56 | 1939 | 2 | 265 | 09281988 | 9 | 28 | 1988 | 21300000 | EMILY WILM... | 021077 |
| 50 | 3502 | 2 | 165 | 07041985 | 7 | 4 | 1985 | 19500000 | CATHEZINE ... | 011015 |
| 44 | 4435 | 2 | 117 | 05141989 | 5 | 14 | 1989 | 17000000 | AGNES KING | 00 |
| 68 | 1673 | 3 | 138 | 07021985 | 7 | 2 | 1985 | 16800000 | MARTIN XU | 041033 |
| 62 | 4181 | 3 | 161 | 02031988 | 2 | 3 | 1988 | 15900000 | JOHN DURN | 030045 |
| 56 | 1443 | 1 | 265 | 12028900 | 12 | 2 | 8900 | 6000000 | PAT DUNN | 021055 |
| 50 | 3607 | 3 | 127 | 08072000 | 8 | 7 | 2000 | 18300000 | ANDREA HIN... | 011014 |
| 44 | 1775 | 3 | 288 | 02051989 | 2 | 5 | 1989 | 2700000 | PETER JONES | 00 |
| 68 | 1209 | 2 | 165 | 05121986 | 5 | 12 | 1986 | 17300000 | DIDRA WILK... | 041065 |

http://www.computerhistory.org/timeline/memory-storage/

# File Processing

http://groups.engin.umd.umich.edu/CIS/course.des/cis400/cobol/cobol.html

```
DATA DIVISION.
FILE SECTION.
FD StudentFile.
01 StudentRec.
   88  EndOfStudentFile  VALUE HIGH-VALUES.
   02  StudentId         PIC 9(7).
   02  StudentName.
       03 Surname        PIC X(8).
       03 Initials       PIC XX.
   02  DateOfBirth.
       03 YOBirth        PIC 9(4).
       03 MOBirth        PIC 9(2).
       03 DOBirth        PIC 9(2).
   02  CourseCode        PIC X(4).
   02  Gender            PIC X.
```

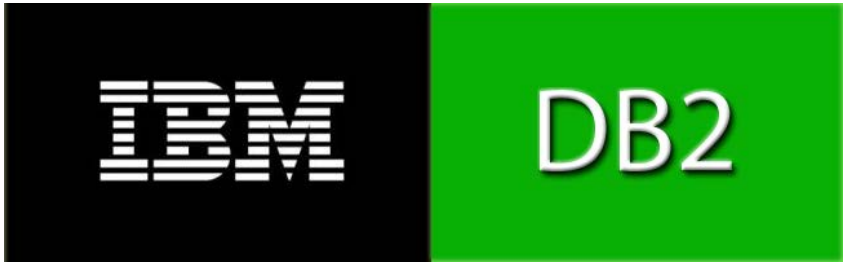



Problems with flat-files:

- data access routines must be programmed in detail
- each program must include full detail of data structure
- multiple users cannot simultaneously access data
- multiple copies of data - not centrally managed

# Relational Databases

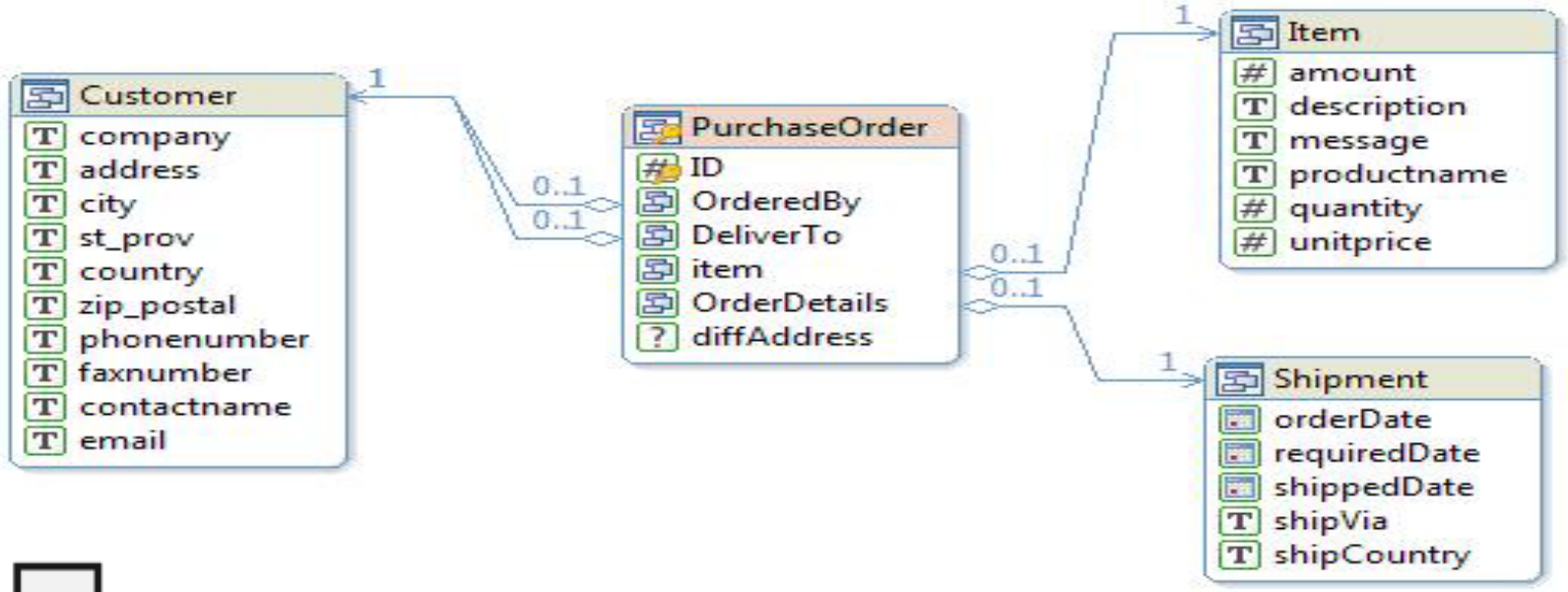
- the first relational databases from Oracle and IBM appear around 1980

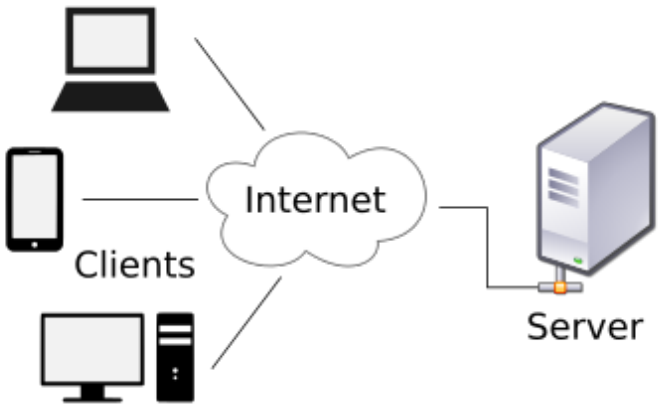




- others appear later


# Relational Databases

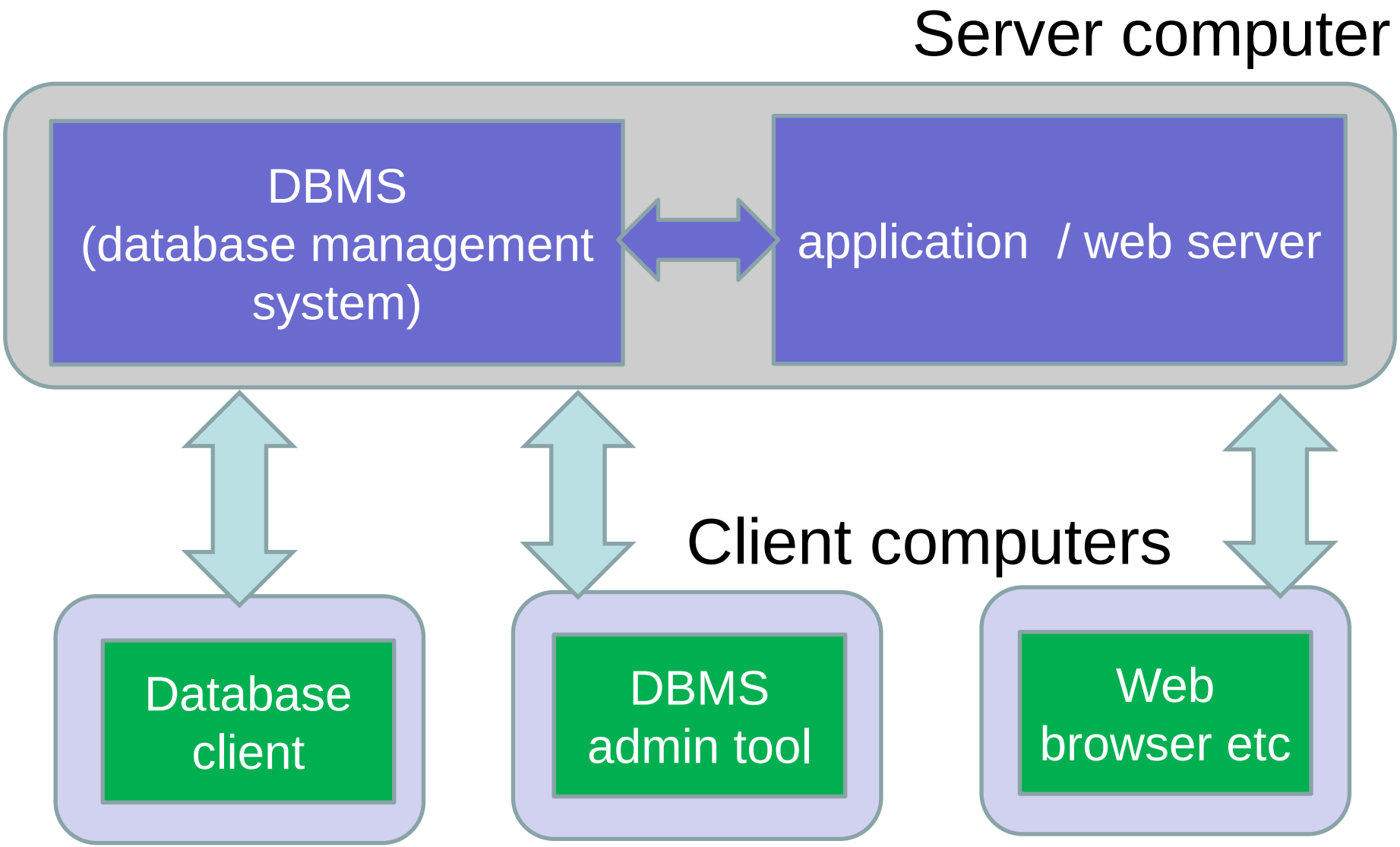
- entity-relationship diagrams, client-server architecture, SQL language



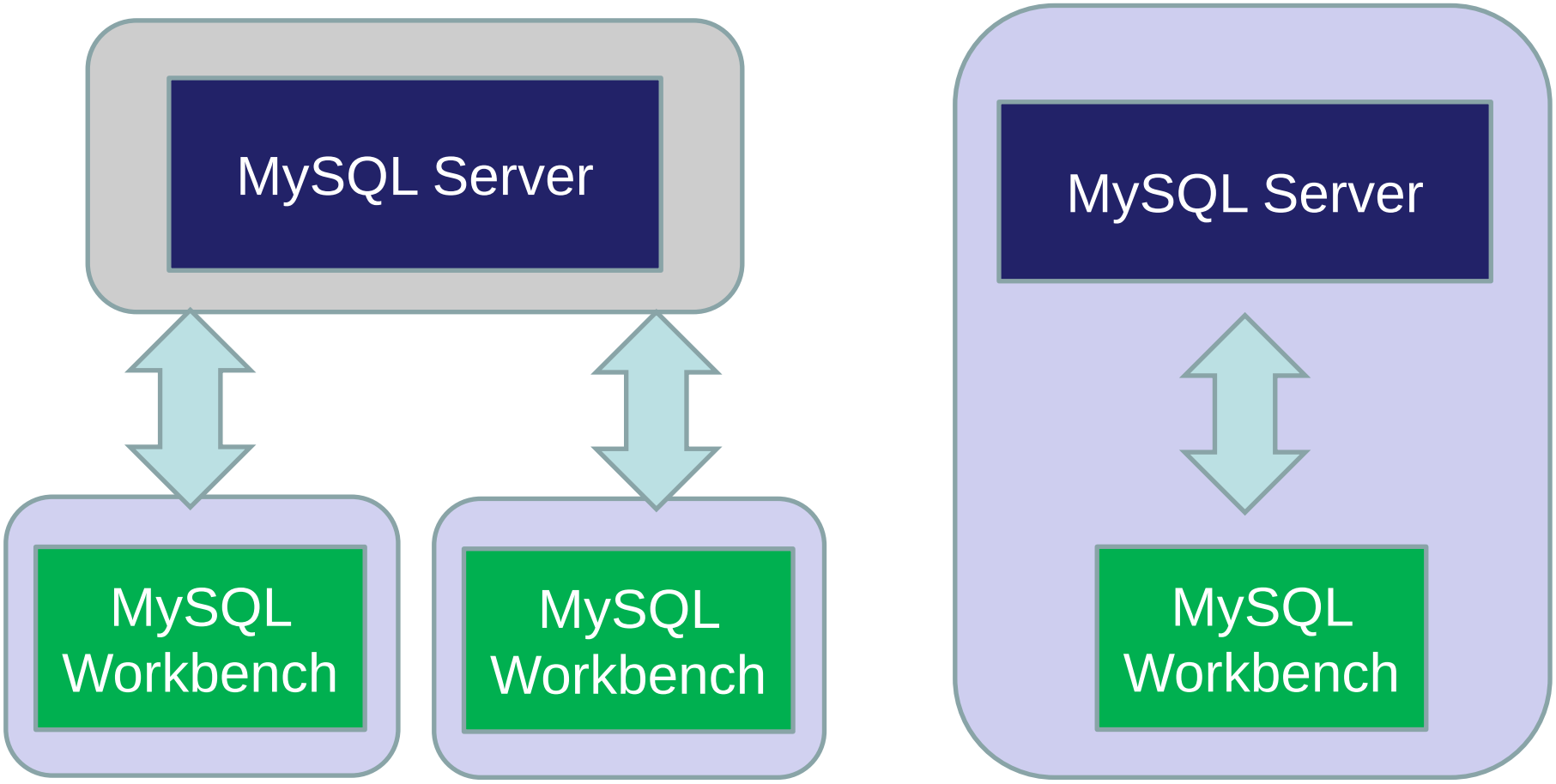




```
SELECT address
FROM Customer
WHERE Country = 'USA';
```

# Client-Server … in industry

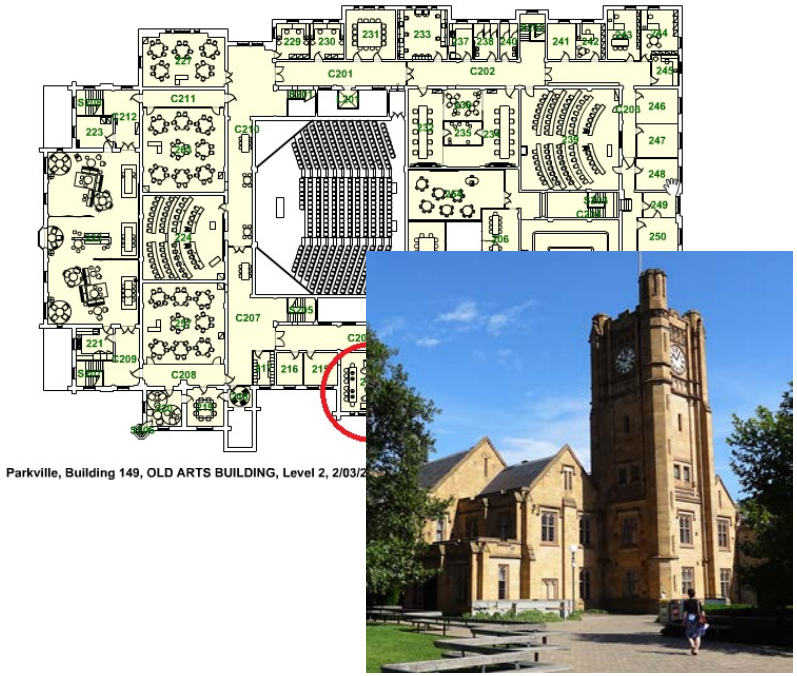


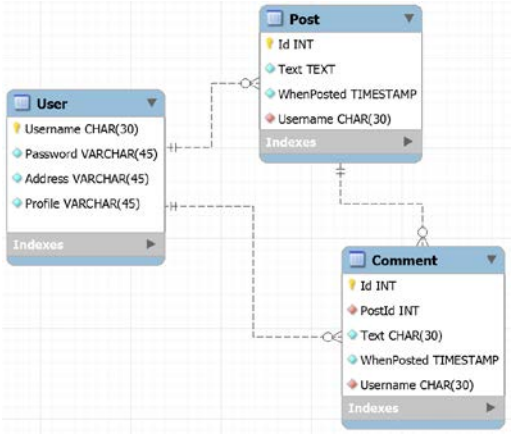
# … in teaching and development



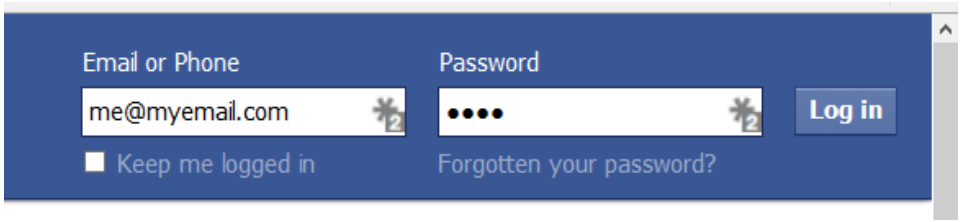
# Let's design a database …

"What is unique about ICT ... In no other discipline is there such an emphasis on developing artefacts (e.g., computer and information systems) which are so abstract and complex and where modelling tools and methods are essential. The systems that ICT professionals deal with cannot be seen or handled in the same simple and direct manner as products of other applied disciplines …" *from the Australian Computer Society CBOK*



# Let's design a database …




## User

| Username | Password | Address | Profile |
|---|---|---|---|
| Anne | pass1234 | 1 Anne st | hi I am Anne |
| Bill | petsname | 2 Bill st | this is Bill's profile |
| Christine | mystreet | 3 Christine st | hi everyone this is Christine!!! |
| David | childsname | 4 David st | David's profile data |

## Post

| Id | Text | WhenPosted | Username |
|---|---|---|---|
| 1 | Here's what I had for lunch | 2015-01-24 20:09:02 | Bill |
| 2 | What's everyone doing tonight? | 2015-01-24 20:09:02 | Anne |
| 3 | check out this great CAT VIDEOZ!!! | 2015-01-24 20:09:02 | David |
| 4 | now look what I had for dinner | 2015-01-24 20:09:02 | Bill |
| 5 | I love Game of Thrones | 2015-01-24 20:09:02 | Anne |
| 6 | I am posting a lot today… | 2015-01-24 20:09:02 | Anne |

## Comment

| Id | PostId | Text | WhenPosted | Username |
|---|---|---|---|---|
| 1 | 3 | ha ha great video!! | 2015-01-24 14:01:33 | Christine |
| 2 | 2 | going to the movies | 2015-01-24 14:01:33 | Bill |
| 3 | 2 | oh that sound nice! | 2015-01-24 14:01:33 | Anne |

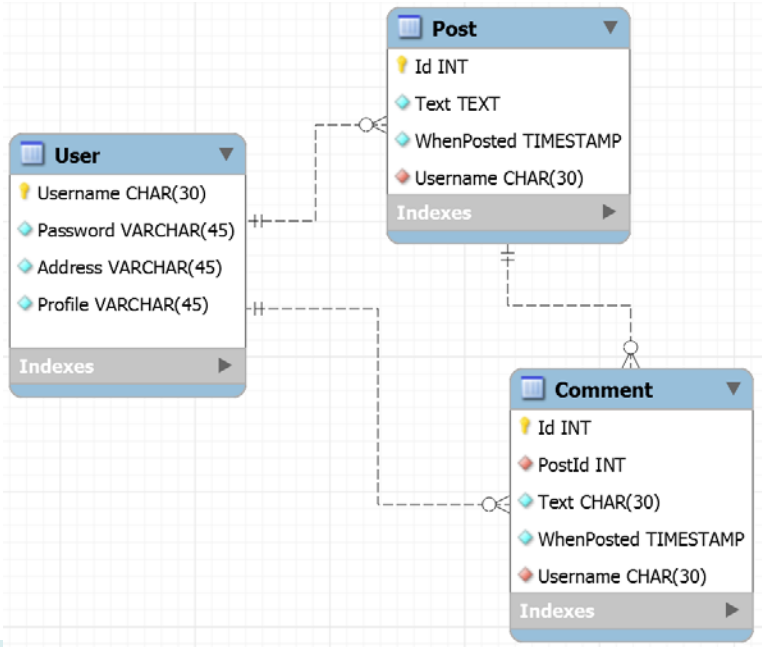
# Let's design a database …

| Username | Password | Address | Profile |
|---|---|---|---|
| Anne | pass1234 | 1 Anne st | hi I am Anne |
| Bill | petsname | 2 Bill st | this is Bill's profile |
| Christine | mystreet | 3 Christine st | hi everyone this is Christine!!! |
| David | childsname | 4 David st | David's profile data |

| Id | Text | WhenPosted | Username |
|---|---|---|---|
| 1 | Here's what I had for lunch | 2015-01-24 20:09:02 | Bill |
| 2 | What's everyone doing tonight? | 2015-01-24 20:09:02 | Anne |
| 3 | check out this great CAT VIDEOZ!!! | 2015-01-24 20:09:02 | David |
| 4 | now look what I had for dinner | 2015-01-24 20:09:02 | Bill |
| 5 | I love Game of Thrones | 2015-01-24 20:09:02 | Anne |
| 6 | I am posting a lot today… | 2015-01-24 20:09:02 | Anne |

| Id | PostId | Text | WhenPosted | Username |
|---|---|---|---|---|
| 1 | 3 | ha ha great video!! | 2015-01-24 14:01:33 | Christine |
| 2 | 2 | going to the movies | 2015-01-24 14:01:33 | Bill |
| 3 | 2 | oh that sound nice! | 2015-01-24 14:01:33 | Anne |

Entity-Relationship diagram:

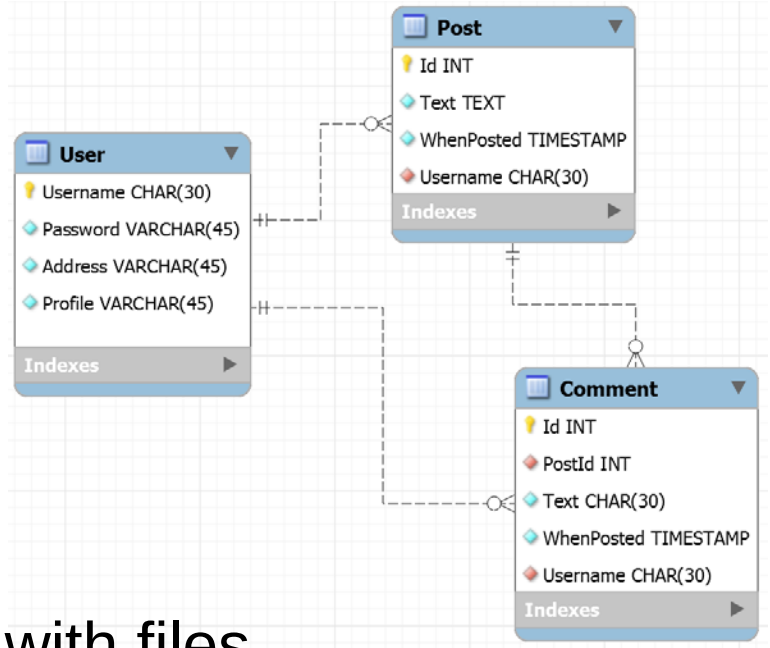



Relational notation:

User(<u>username</u>, password, address, profile)
Post(<u>id</u>, text, whenposted, *username*)
Comment(<u>id</u>, *postid*, text, whenposted)

# Maintaining structure of database

What commands do we need to manipulate the *structure* of our tables?




Working with tables is like working with files – there are 4 things you can do:
CREATE a table
DROP (i.e. delete) a table
ALTER a table (e.g. add a column)
RENAME a table

# Maintaining contents of database

What commands do we need to manipulate the *contents* of our tables?

**User**

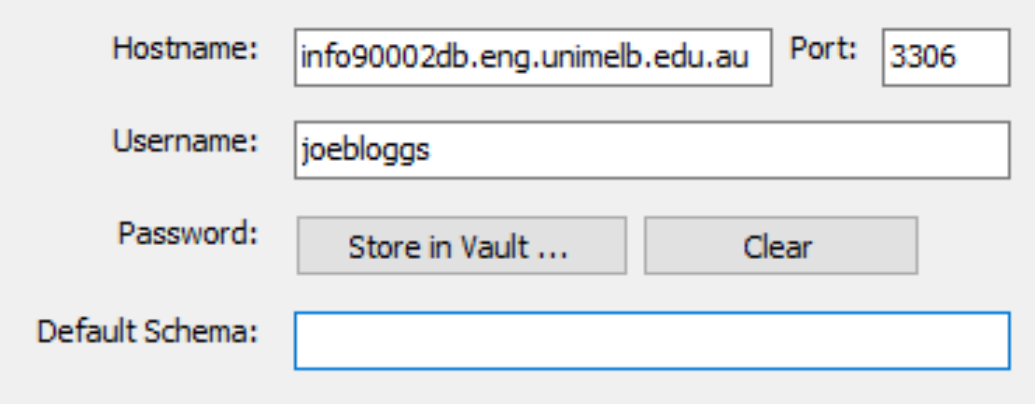
| Username | Password | Address | Profile |
|---|---|---|---|
| Anne | pass1234 | 1 Anne st | hi I am Anne |
| Bill | petsname | 2 Bill st | this is Bill's profile |
| Christine | mystreet | 3 Christine st | hi everyone this is Christine!!! |
| David | childsname | 4 David st | David's profile data |

You simply need to be able to:
SELECT, or read, data from the table
INSERT new rows into the table
DELETE existing rows from the table
UPDATE existing rows in the table

# Maintaining users' access to database

What commands do we need to control *users' access* to our tables?

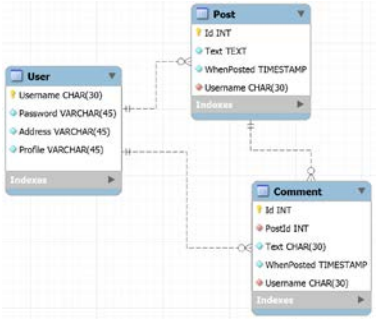
(users = I.T. staff here)




There are 4 things you can do:
CREATE a user
DROP a user
GRANT a user access rights to a table
REVOKE those rights

# Database lifecycle

- Design the database
  - data modelling, E-R diagrams




- Implement the database
  - data definition language DDL

•Create
•Drop
•Alter
•Rename

- Data access / programming
  - data manipulation language DML

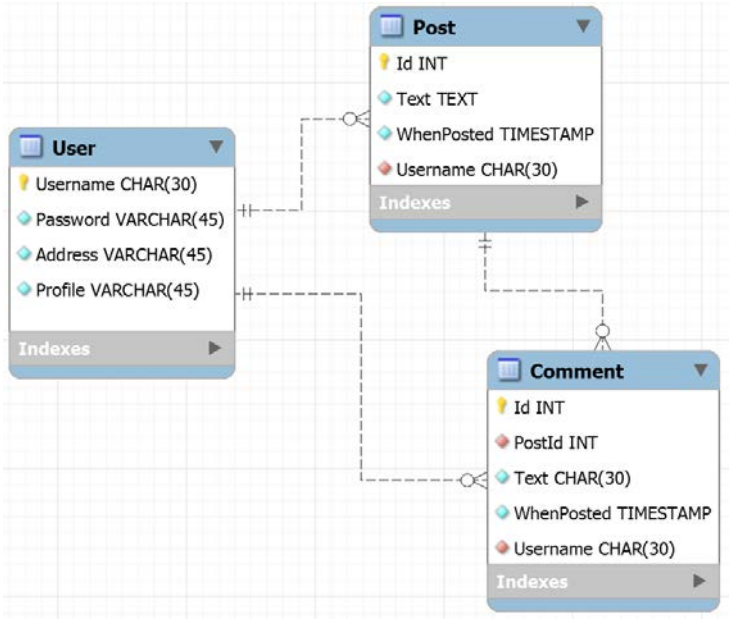
•Select
•Insert
•Update
•Delete

- Database administration
  - data control language DCL

•Grant
•Revoke

# Data Modelling

- What are the entities that need to be tracked?
- What attributes will be recorded about each entity?
- What are the relationships between entities?
- What are the cardinalities of relationships?




You will be given requirements such as:

"We have many users.
Users may enter posts into the system.
Each post is entered by exactly one user.
Users can comment on posts.
A post can be commented on many times."

# Noun-verb analysis

- An investment bank has a number of branches. Within each branch a number of departments operate and are structured in a hierarchical manner. The bank employs around 3000 staff who are assigned to work in the various departments across the branches. There are essentially three types of special employees where extra details required by the system. There are dealers who carry out investments who have limits imposed upon them for how much they can spend. There are IT compliance managers who's Basel2 role is required to be stored and there are HR managers that need have their assessment number recorded (along with other details not specified here).
- We need a database to record staff details including which department and branch they are assigned...

# Homework: Noun-verb analysis

- available on LMS
- do in your own time
- we'll discuss next week