4. Why DB Design?

Today's Lecture

1. Mishandling Keywords and Categories
2. Repeated Information
3. Designing for a Single Report
4. Excercises
5. Summary

Mishandling Keywords and Categories

- Fail to properly deal with keywords and categories.

Example: The Plant DB

A small portion of a database table recording information about plants. Along with the botanical and common names of each plant, the developer decides it would be convenient to keep information on the uses for each plant. This is to help prospective buyers decide whether a plant is appropriate for their requirements.

“I need to store all the info I have about each plant.”

“What uses does this plant have?”

plantID	genus	species	common_name
1	Dodonaea	viscosa	Akeake
2	Cedrus	atlantica	Atlas cedar
3	Alnus	glutinosa	Black alder
4	Eucalyptus	nichollii	Black peppermint gum
5	Juglans	nigra	Black walnut
6	Acacia	mearnsii	Black wattle

Table Plants

plant	use
1	soil stability
1	hedging
1	shelter
2	shelter
3	firewood
3	soil stability
3	shelter

Table Uses

Example: Research Interests

An employee of a university's liaison team often receives calls asking to speak to a specialist in a particular topic. The liaison team decides to set up a small spreadsheet to maintain data about each staff member's main research interests. Originally, the intention is to record just one main area for each staff member, but academics, being what they are, cannot be so constrained. The problem of an indeterminate number of interests is solved by adding a few extra columns in order to accommodate all the interests each staff member supplies.

Find who is interested in a particular topic.

Find a researcher who is able to "visualize data"?

personID	...	...	...	interest 1	interest 2
152				Computing education	
275				Computer visualisation	Simulation
282				Scientific visualization	Statistics
292				Visualisation of data	Computing education
890				Databases	Scientific visualisation

Repeated Information

- Storing unnecessarily the same piece of information several times.

Example: Insect Data

Team members of a long-term environmental project regularly visit farms and take samples to determine the numbers of particular insect species present. Each field on a farm has been given a unique code, and on each visit to a field a number of representative samples are taken. The counts of each species present in each sample are recorded.

Find information about fields with particular soil types or visits undertaken in fine weather conditions.

	A	B	C	D	E	F
1	farm	field	date	sample	springtail	fungus_beetle
268	1	ADhc	Aug-11	1	2	0
269	2	ADhc	Aug-11	2	2	0
270	1	ADhc	Aug-11	3	7	0
271	1	ADhc	Aug-11	4	3	2
272	1	ADhc	Aug-11	5	3	0
273	1	ADhc	Aug-11	6	3	9
274	1	ADhc	Aug-11	7	2	1
275	1	ADhc	Aug-11	8	6	1
276	1	ADhc	Aug-11	9	2	1
277	1	ADhc	Aug-11	10	5	3
278	1	ADhc	Aug-11	11	0	0
279	1	ADhe	Aug-11	1	0	6
280	1	ADhe	Aug-11	2	1	1
281	1	ADhe	Aug-11	3	5	2

field	farm	soil
Adhc	1	
Adhe	1	
Mvhe	2	
MVhc	2	

Table Fields

visitID	field	date	conditions
113	Adhc	Aug-06	Fine
114	Adhe	Aug-06	Fine
115	Adhc	Sep-06	Rain
116	Adhe	Sep-06	Overcast

Table Visits

visitID	sample	springtail	fungus_beetle
113	1	2	0
113	2	2	0
113	3	7	0
113	4	3	0
113	5	0	2
113	6	3	1

Table Counts

Designing for a Single Report

Design a table to match the requirements of a particular report.

Example: Academic Results

A university department needs to have its final-year results in a format appropriate for taking along to the examiners' meeting. The course was very rigidly prescribed with all students completing the same subjects, and a report was generated by hand prior to the system being computerized. This format allowed each student's performance to be easily compared across subjects, helping to determine honors' boundaries.

ID	Name	S001	S002	S103	S104	S202	S310	S331	GPA
982208	Jo Brown	A+	A	A	A+	A	B+	B+	8.6
986667	Helen Green	A	A	A+	A	A	B+	B+	8.5
987645	Peter Smith	A	B+	A-	A-	B+	A-	B	7.5

Exercise 4-1

A school is planning some outdoor activities for its students. The staff wants to create a database of how parents can help. The secretary sets up the database table to keep the information.

last_name	first_name	phone	contribution	contribution2
Smith	Jane	4623598	Food preparation	Driving
Green	Rob	8965431	Transport	
Henry	James	9576342	Camping Gear	Cooking
Wang	Li	9612345	Cooking	

- What problems can you foresee in making good use of this information?
- Suggest some better ways that this information could be stored.

Exercise 4-2

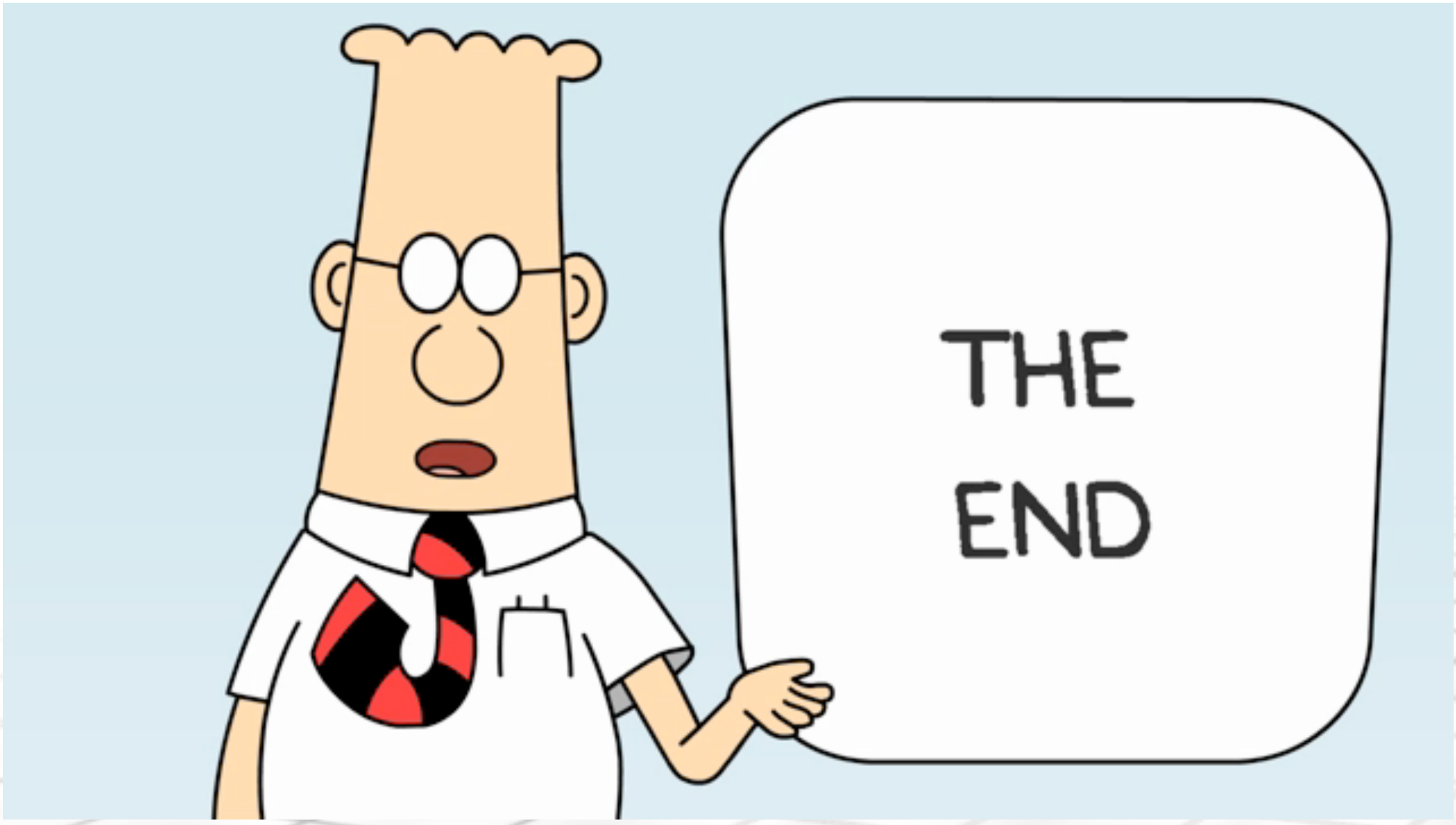
A small library keeps a roster of who will be at the desk each day. They have a database table

week_start	Mon	Tue	Wed	Thur	Fri
17/10/2011	Jane	Sue	George	Sue	Jane
24/10/2011	Jane	Sue	Linda	Sue	Lee
31/10/2011	Sue	Sue	Lee	George	George

- What problems can you foresee in making good use of this information?
- Suggest some better ways that this information could be stored.

Summary

- Show how to design a database may be influenced by a particular report or by a particular method of input.
- The hurried creation of a database or spreadsheet can lead to a design that cannot cope with even simple changes to the information you would like to retrieve.
- It is important to think carefully about the underlying data, and design the database to reflect the information being stored rather than what you might want to do with the data in the short term.



출처: metachannels.com

Thank you!