

2. Why DB Design?

Schedule*

0. Seminar Overview
1. Overview of the Relational Data Model
2. Why DB Design?
3. Development Process
4. Requirements
5. Conceptual Data Model
6. Generalization & Specialization
7. Relational Database Design
8. Normalization
9. Keys and Constraints

*subject to change without notification

What will be dealt

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

Mishandling Keywords and Categories

- 키워드와 카테고리를 정확히 다룰 수 있어야 함.

Example: The Plant DB

- 나무에 관한 정보를 기록한 데이터베이스 테이블.
- 테이블에 식물학 및 일반 나무 이름과 함께 용도에 대한 정보를 유지하고자 함.
- 이는 나무가 구매자의 요구 사항에 적합한 가를 알 수 있도록 하기 위함.

“나무에 관한 모든 정보를 저장합니다.”

“이 나무의 용도는?”

plantID ▾	genus ▾	species ▾	common_name ▾	use1 ▾	use2 ▾	use3 ▾
1	Dodonaea	viscosa	Akeake	shelter	hedging	soil stability
2	Cedrus	atlantica	Atlas cedar	shelter		
3	Alnus	glutinosa	Black alder	soil stability	shelter	firewood
4	Eucalyptus	nichollii	Black peppermint gum	shelter	coppicing	bird food
5	Juglans	nigra	Black walnut	timber		
6	Acacia	mearnsii	Black wattle	firewood	shelter	soil stability

PlantsUses

plantID ▾	genus ▾	species ▾	common_name ▾	use1	use2	use3
1	Dodonaea	viscosa	Akeake	1	2	3
2	Cedrus	atlantica	Atlas cedar	1		
3	Alnus	glutinosa	Black alder	3	4	1
4	Eucalyptus	nichollii	Black peppermint gum	1	5	6
5	Juglans	nigra	Black walnut	7		
6	Acacia	mearnsii	Black wattle	4	1	3

Plants

Use_Id	Use_Description
1	shelter
2	hedging
3	soil stability
4	firewood
5	coppicing
6	bird food
7	timber

Uses

Example: Research Interests

- 대학에서 특정 분야의 전문가를 대외협력팀 직원에게 요청.
- 해당 직원은 교수들의 주요 연구 관심사에 대한 데이터를 유지하기 원함.
- 단 하나의 관심 영역을 기록하려 하나, 그것이 무엇인지에 따라 제약이 따를 수 있음.
- 애매한 관심사 문제를 수용하기 위하여 몇 개의 추가 열을 추가.

특정 주제에 관심이 있는 교수?

“visualize data”를 연구하는 교수?

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

- 동일한 데이터를 여러 번 반복 저장.

Example: Insect Data

장기 환경 프로젝트에서 정기적으로 농장을 방문하고 샘플을 채집하여 특정 곤충 종의 수를 측정한다.

- 농장의 각 필드에는 고유 id를 부여.
- 필드를 방문 할 때마다 여러 샘플을 채집.
- 각 샘플에 있는 각 종의 수를 기록.

특정 토양 유형이나 좋은 기상 조건하에 방문한 필드에 대한 데이터를
검색

	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	

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

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

Counts

Designing for a Single Report

특정 보고서의 요구 사항을 충족하는 테이블을 디자인.

Example: Academic Results

대학에서 학기말 학생들의 최종 성적을 적합한 형식으로 작성합니다. 이 형식을 통해 각 학생의 성과를 여러 과목에서 쉽게 비교할 수 있었고, 장학금 등의 지급하는 데 도움을 줍니다.

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 2-1

학교에서 학생들의 야외 활동 수업을 계획하고 있습니다. 교직원들은 부모들의 자원봉사에 대한 데이터베이스를 만들고 싶어합니다.

이 데이터를 저장할 데이터베이스 테이블을 만들어 주세요.

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	

- 이 데이터를 잘 활용하면 어떤 이점이 있으며 무엇을 예견할 수 있을까요?
- 이 데이터를 위와 같이 관리하면 어떤 문제점이 있으며, 바람직하게 관리할 수 있는 방법을 제안하십시오.

Exercise 2-2

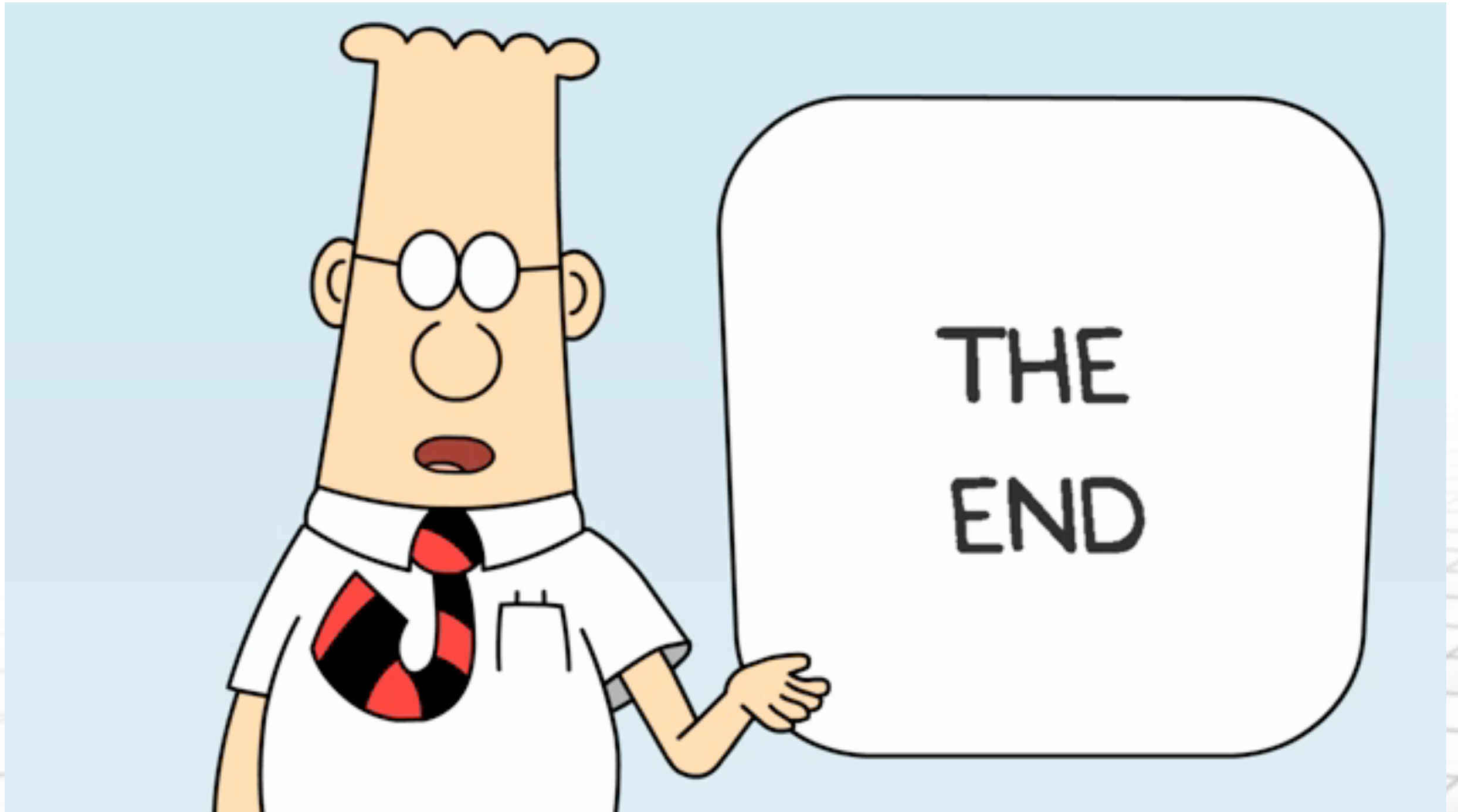
동네 작은 도서관에서 아래와 같은 테이블에 자료를 저장하여 근무하는 자원 봉사자들을 관리합니다.

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

- 이 데이터를 잘 활용하면 어떤 이점이 있으며 무엇을 예견할 수 있을까요?
- 이 데이터를 위와 같이 관리하면 어떤 문제점이 있으며, 바람직하게 관리할 수 있는 방법을 제안하십시오.

Summary

- 데이터베이스 설계는 특정 보고서 또는 특정 입력 방법의 영향을 받을 수 있다.
- 데이터베이스 또는 스프레드시트의 단순한 생성은 검색하려는 정보에 대한 단순한 변경조차에서도 어려움이 있을 수 있다.
- 기본 데이터에 대해 신중히 생각하고 닥친 문제를 해결하기 위하여 데이터를 저장하는 것보다 데이터의 특성을 반영할 수 있도록 데이터베이스를 설계하는 것이 중요하다.



출처: metachannels.com

Thank you!