Database Quality on Modelling Stage

Relational Databases Basics

Ideas to consider

Subject matter analysis depth

Data non-redundancy (single storage point)

Data changes through time

Data validity

Data precision

Data formats unification

Ask experts

Read documentation

Communicate with the Customer
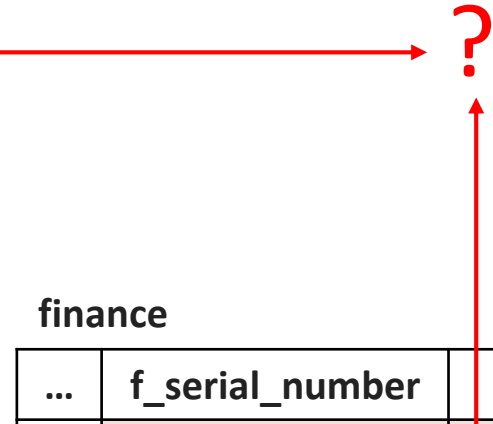
Mind project boundaries

Ideas to consider: data non-redundancy (single storage point)

...	c_serial_number	c_sum	...
...	**AC345347856DF**	**34 000 000**	...
...	DF345345652YH	12 000 000	...
...	AA345235235KL	32 511 012	...
...	GT456345342UT	41 356 343	...

?

finance

...	f_serial_number	f_sum	...
...	**AC345347856DF**	**29 627 532**	...
...	DF345345652YH	12 000 000	...
...	AA345235235KL	32 511 012	...
...	GT456345342UT	41 356 343	...

Ideas to consider: data changes through time

Yesterday			
...	**i_code**	**i_price**	**...**
...	1234	12 000	...
...	4567	11 500	...
...	9021	14 700	...
...	5378	11 199	...

Today			
...	**i_code**	**i_price**	**...**
...	1234	12 300	...
...	4567	10 500	...
...	9021	11 500	...
...	5378	10 099	...

And what if suddenly we need yesterday's data?

Ideas to consider: data validity

Shall we force phone number format?

Shall we force name format?

Shall we allow some value to be typed or force selection?

...

Ideas to consider: data precision

Mind decimal precision! (!!)

Know the exact precision you need

Know all the exact rules of your subject matter

Beware of storing datetime values with redundant precision

Avoid formats variety

Mind precision

Know all the exact rules of your subject matter

How to…

Information gathering, analysis and discussion

Thoughtful technical decisions

Control and self-control
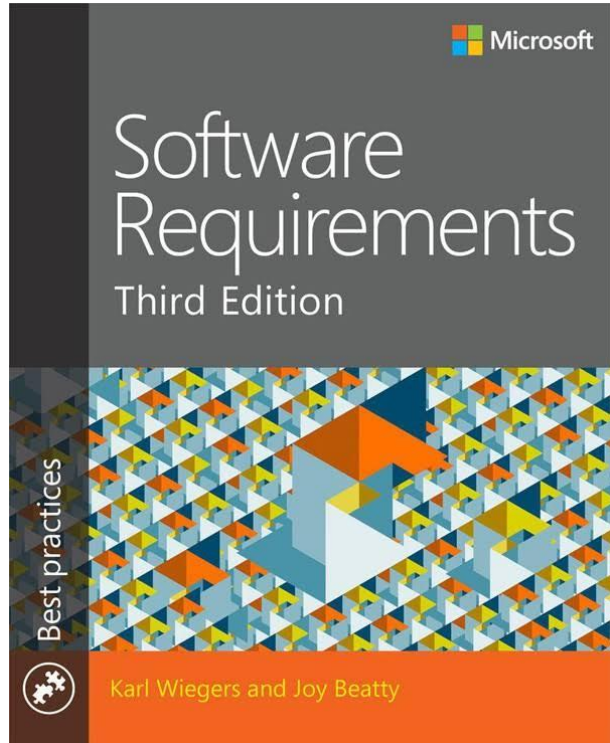
Experiments

"What if?" vs "OMG, it works!"

Detailed documentation

Outlook widening

Further education…

Database Quality on Modelling Stage

Relational Databases Basics