

FOUNDATIONS of DATA CURATION

Allen H. Renear, Cheryl A Thompson, Katrina S Fenlon, Myrna Morales



School of Information Sciences

• University of Illinois at Urbana-Champaign



DATA MODELS: RELATIONS

⑤

ABSTRACTION, INDIRECTION AND DATA INDEPENDENCE

Abstraction, indirection, and data independence

It is worth saying these things one more time

Abstraction

Indirection

Data independence

A fundamental principle in the data curation creed

Abstraction and Indirection

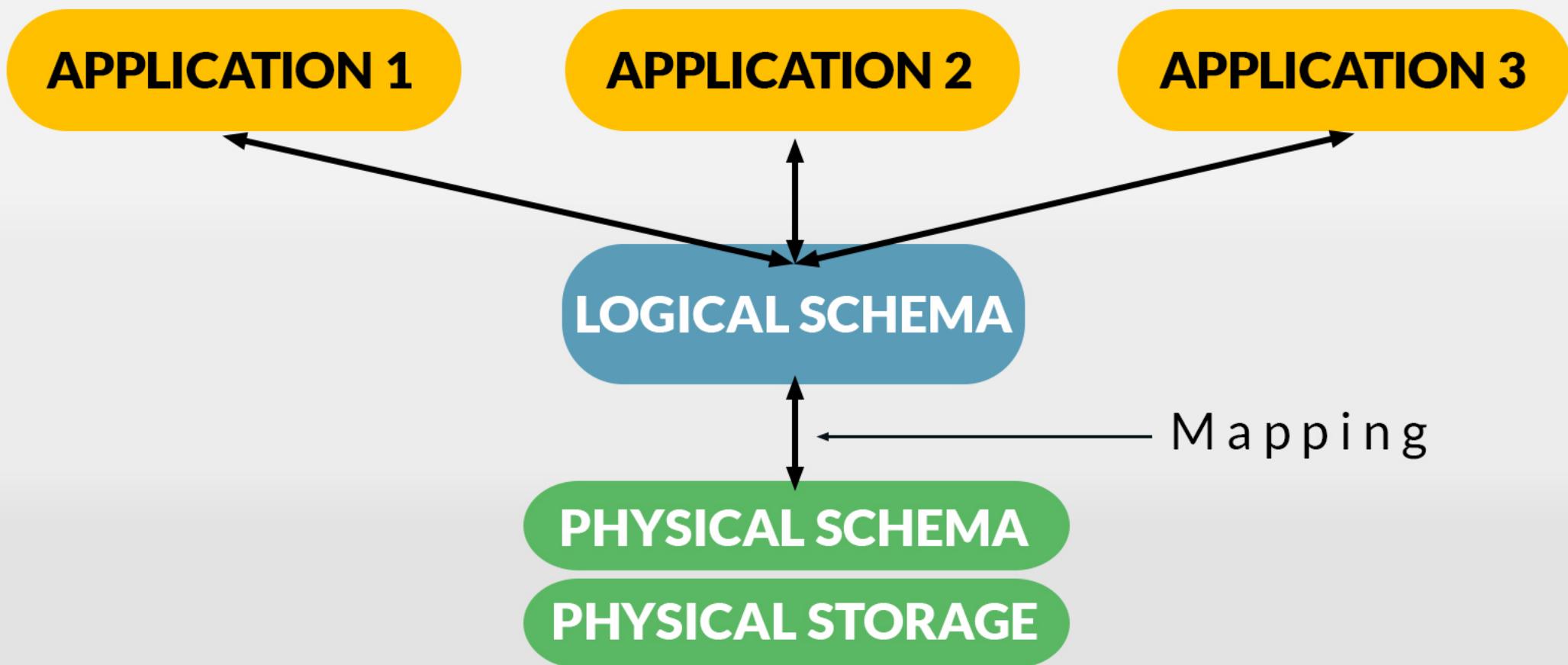
The success of the relational model is based primarily on two related principles:

Abstraction and Indirection

The relational representation *abstracts* away from the specific and transient details of storage, presenting only the intrinsic features of the data itself.

Subsequent interaction with the stored data is then *indirect* via a mapping from the relational schemas to the stored data

Mapping



Adjusting the mapping manages data independence

Physical data independence

The **physical schema** (storage method) can change without affecting interactions with the data.

The problem with direct interactions:

if the storage method switches from a fixed field approach to a delimited field approach
then the access programs (and other tools) will return the wrong results.

The solution:

The physical schema <> logical schema mapping is adjusted to accommodate the changes

Adjusting the mapping manages data independence

Logical data independence

The **logical schema** can change without affecting interactions with the data.

For instance, if a new attribute is accommodated by adding a delimited field to the right side of a record, then any program or other tool that has been identifying fields by counting delimiters right to left will probably return the wrong result.

The solution:

Again, the physical schema <> logical schema mapping is adjusted to accommodate the changes

...with feeling

Abstraction

implemented with indirection

supports data independence

(among other things)

This is a fundamental principle in the data curation creed.
We will see it again and again in this course.

FOUNDATIONS OF DATA CURATION (IS531)

Allen H. Renear, Cheryl A Thompson, Katrina S Fenlon, Myrna Morales
School of Information Sciences
University of Illinois at Urbana-Champaign

**Includes material adapted from work by Carole Palmer, Melissa Cragin,
David Dubin, Karen Wickett, Bertram Ludaescher, Ruth Duerr and Simone Sacchi.**

Comments and corrections to: renear@illinois.edu.