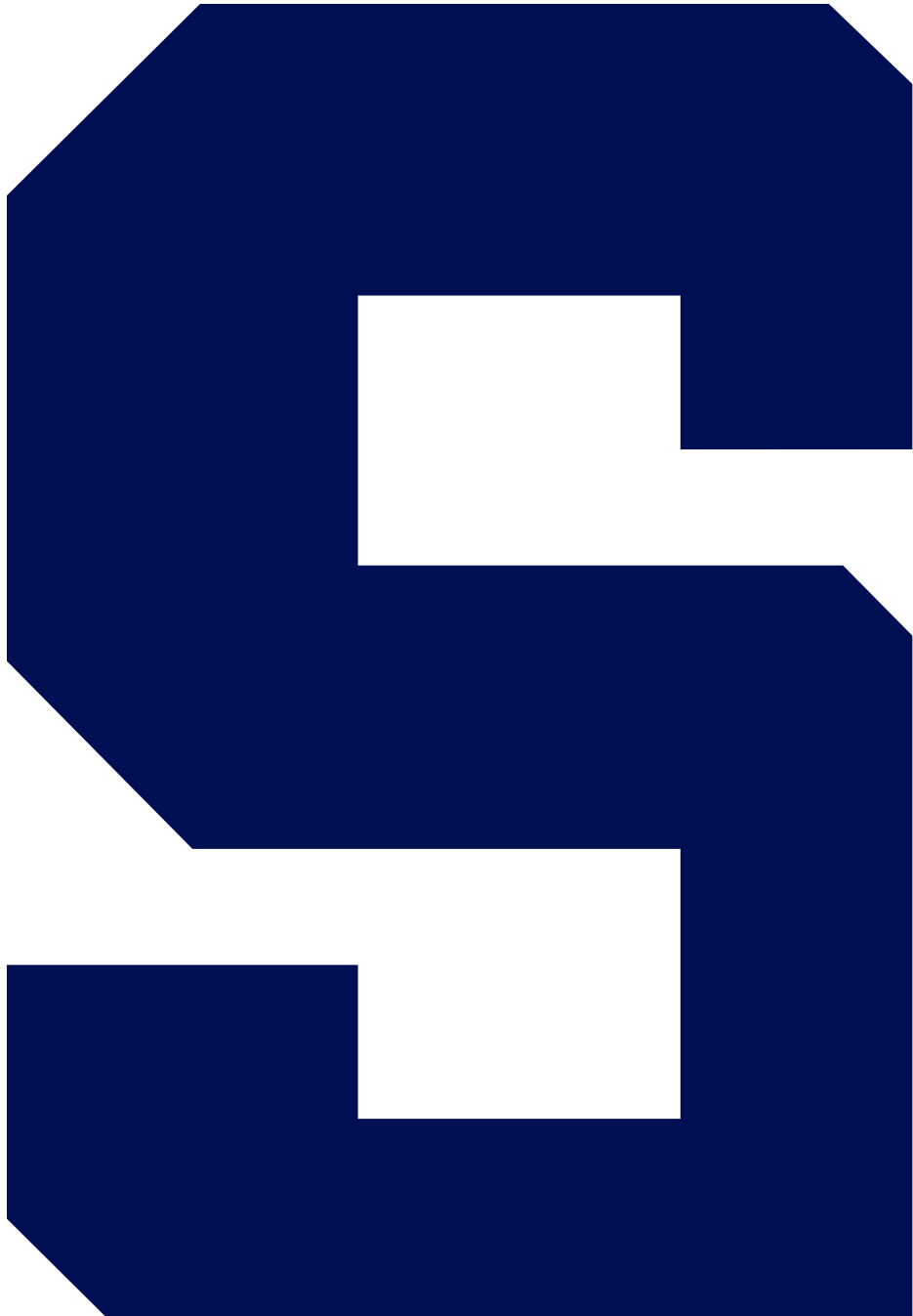
Syracuse University
School of Information Studies



Introduction to Databases

Agenda

- What are data and metadata?
- What is a database?
- What is a database management system (DBMS)?
- What kinds of DBMS are out there?
- How are databases perceived?
- How do we build database applications?

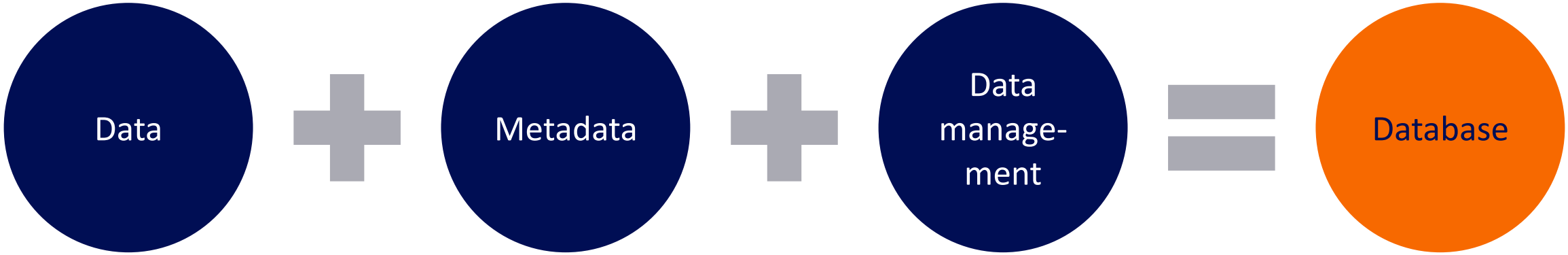
Introduction to Databases

The End

Syracuse University
School of Information Studies

The Database Defined and Data

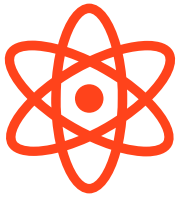
The Database Defined



Data

- Raw, unprocessed facts
- Three types
  1. Atomic: simple
  2. Composite: consists of more than one value
  3. Rule: identifies how data are related

Data Are Raw and Unprocessed



Facts

Atomic
$45
13244

Objects

Composite
314 Hinds Hall Syracuse, NY
https://www.syr.edu/ischool

Processes/rules

Connects objects
Customer pays invoice
Car consists of parts

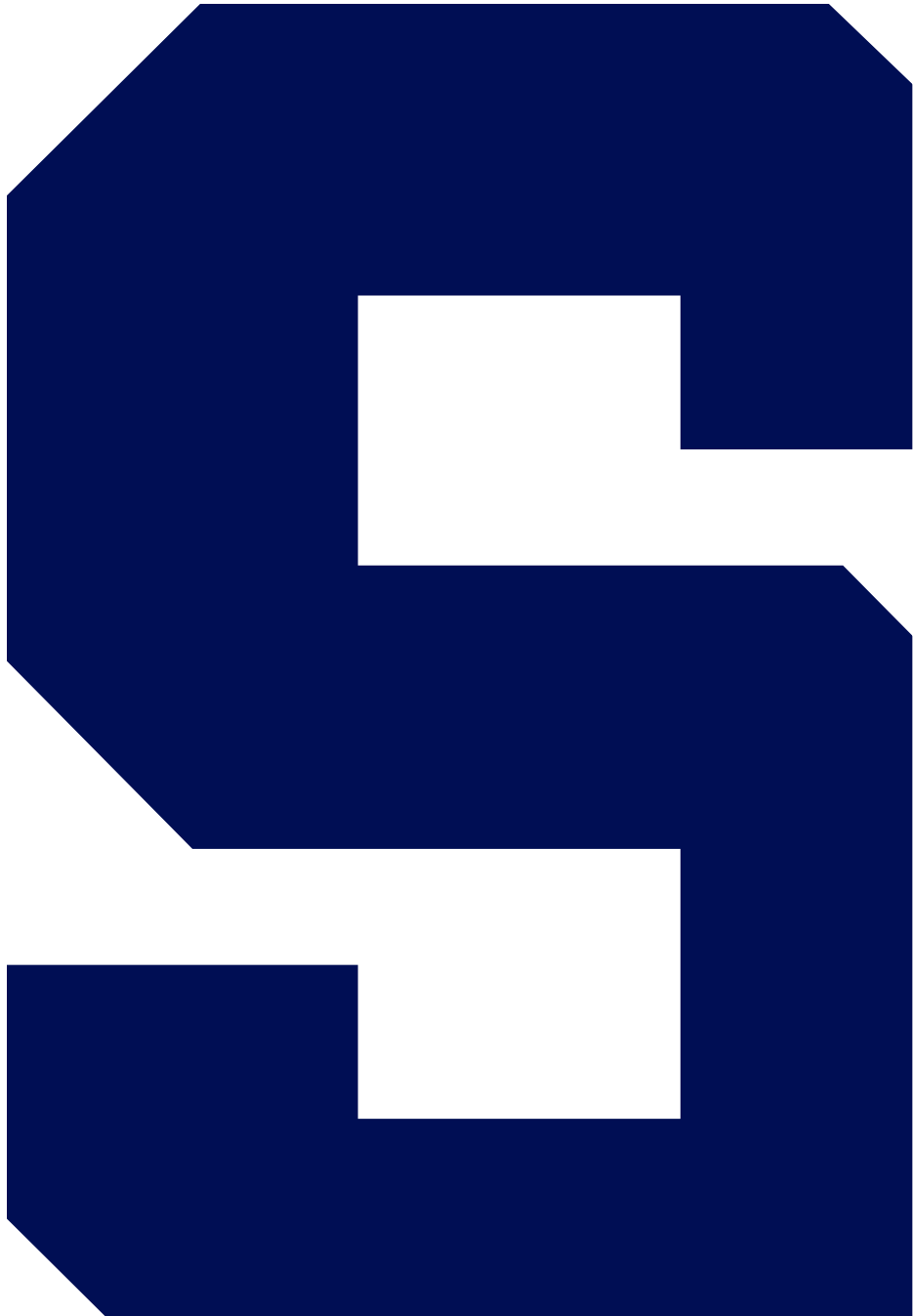
Syracuse University
School of Information Studies

The Database Defined and Data

The End

Syracuse University
School of Information Studies

Information Defined



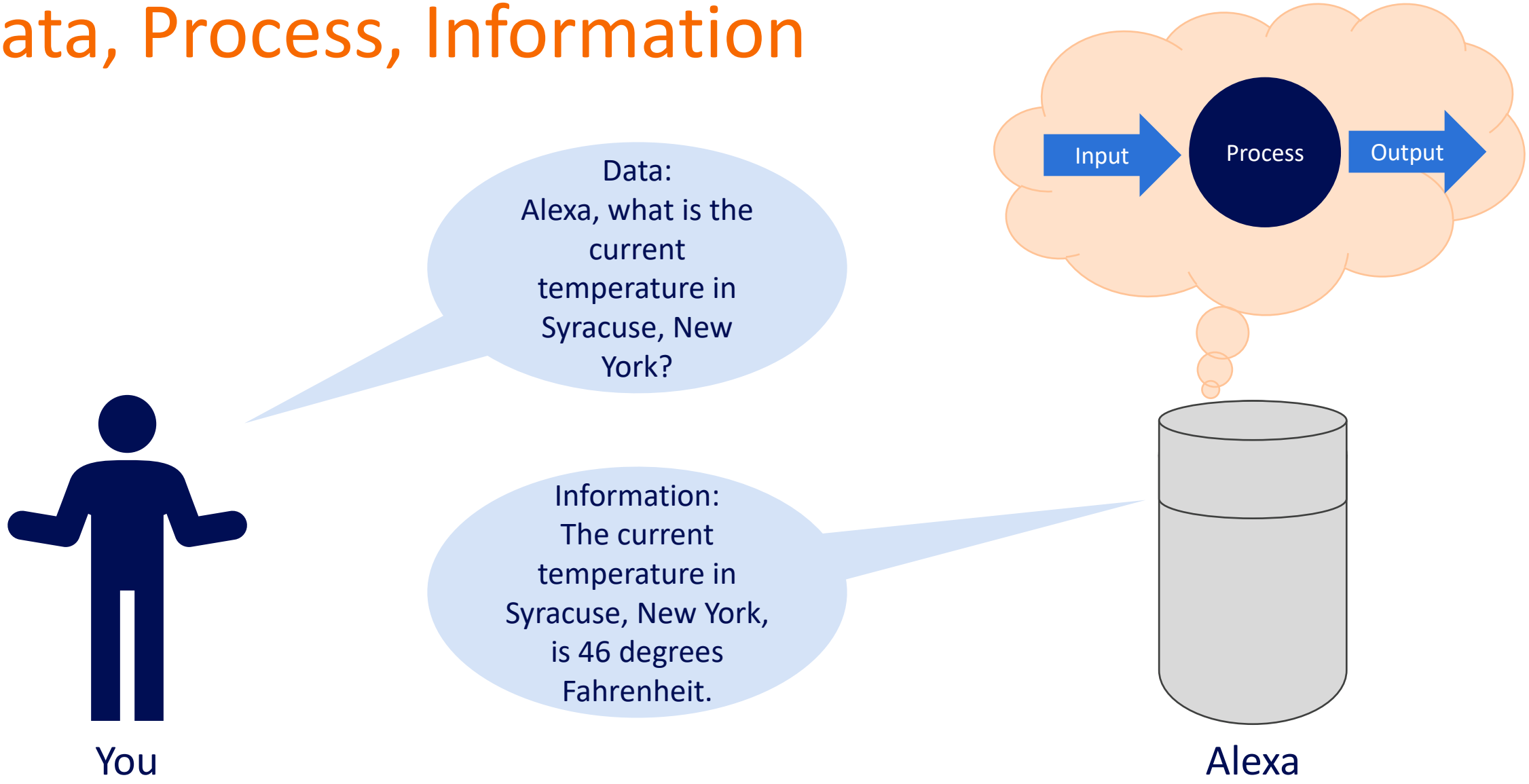
Information

- Information is contextualized data; any data within a context
- The output of a process
- Processed data
- Examples
  - 2 + 3 = 5
    - 2 and 3 are data, 5 is information within the addition context.
  - Because Bob's date of birth is 11/1/1990, on November 1, 2020, Bob will be 30 years old
    - 30 years old is information, the output of a process.

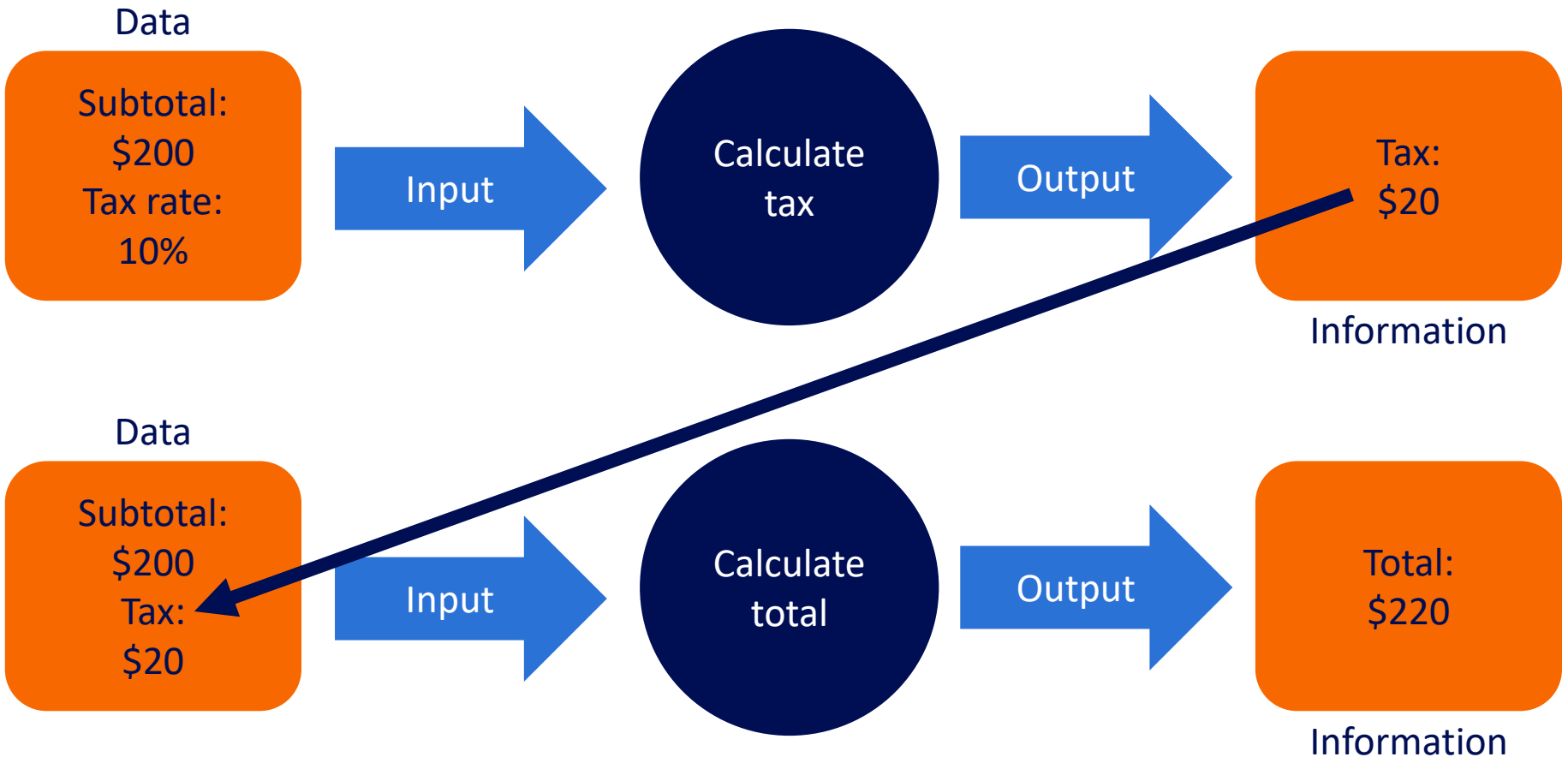
Information Is Data in Context

- Example, a function call in code:
  `getStudentGrade('mafudge','IST256','Fall 2020')`
- `Input data: 'mafudge','IST256','Fall 2020'`
- `Process: getStudentGrade`
- `Output information: A-`

Data, Process, Information



Data Can Be Information; Information Can Be Data



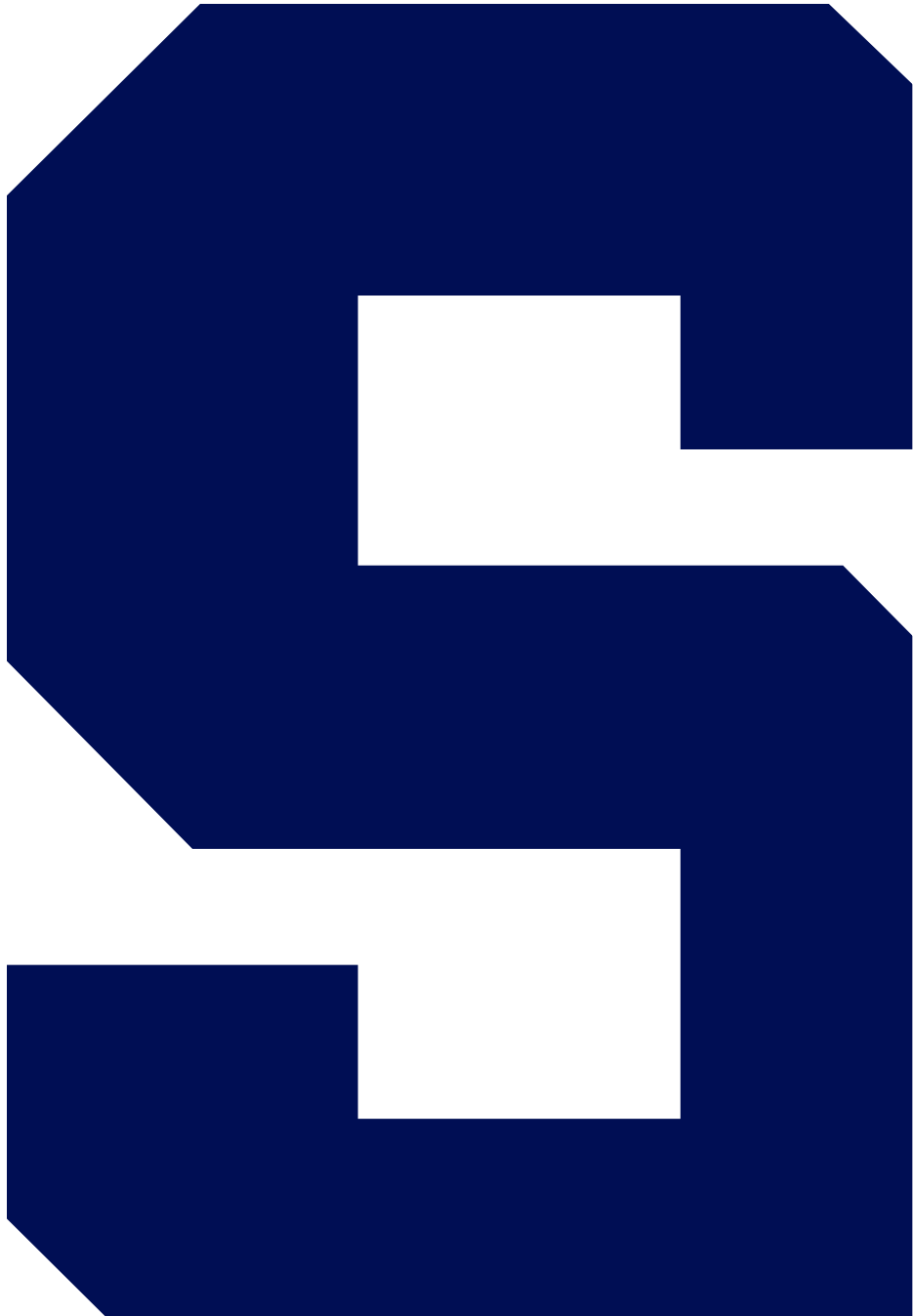
Syracuse University
School of Information Studies

Information Defined

The End

Syracuse University
School of Information Studies

Metadata Defined



Metadata

- Metadata is data about data
- It is used to describe and add additional meaning to data
- Examples
  - 3.75 is just data. But GPA is 3.75; the “GPA” labels the data. GPA is metadata.
  - Dec-2021 is just data. With date metadata, we know this is a month and year.
  - 3155551234 is just data. With structural metadata, (315) 555-1234 is a phone number.

Metadata: Data About Data

Different types

- Label: visual descriptor; amount of money I have → $5
- Definition: provides a definition; money is a quantity of legal tender → $5
- Type: valid data type; numeric → $5
- Constraints: business rules or acceptable values; must be 0 or more → $5
- Length: storage format; 12 bytes → $5
- Location: where the data can exist; my wallet → $5
- Ownership: who or what has access; my wallet → $5

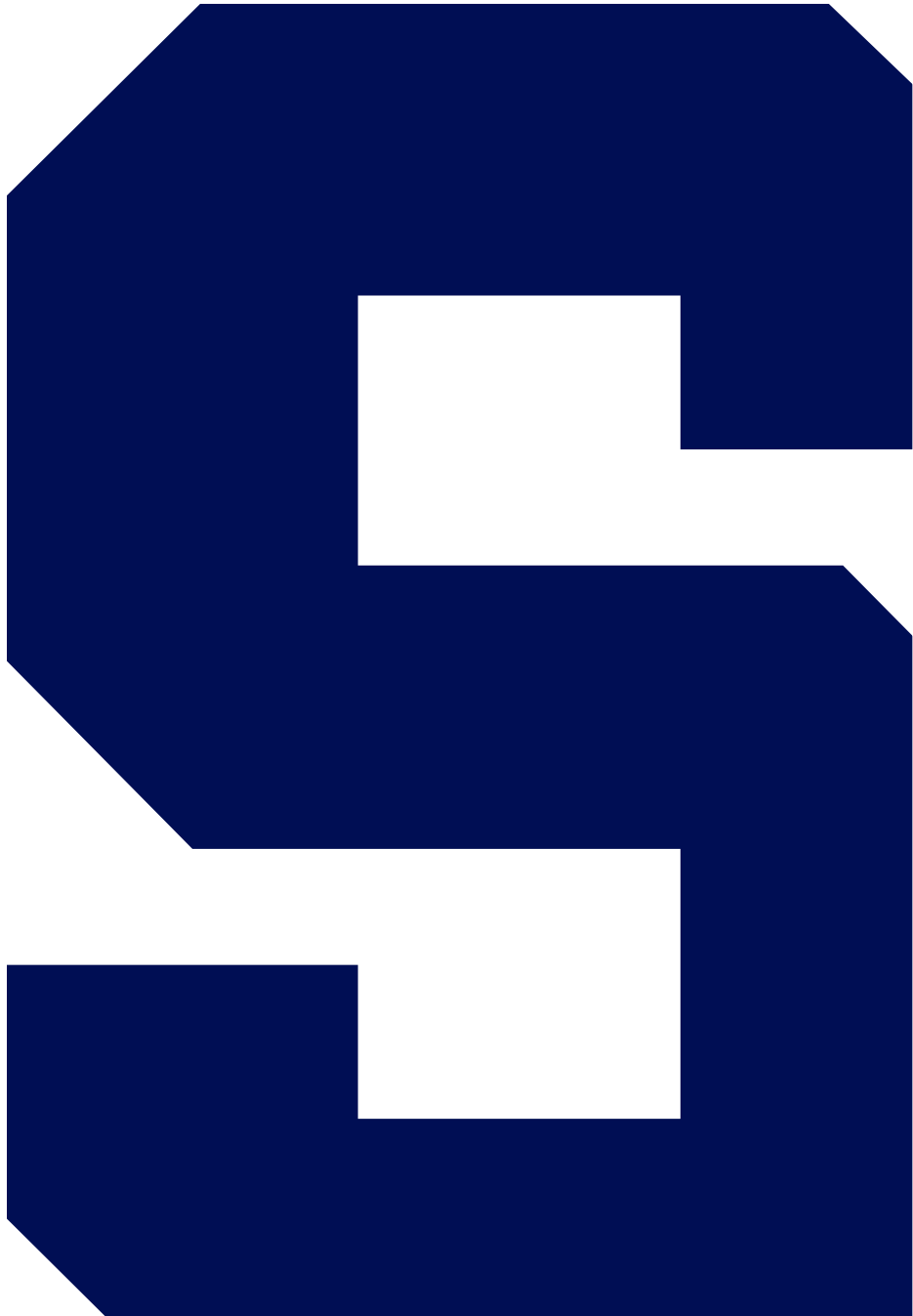
Syracuse University
School of Information Studies

Metadata Defined

The End

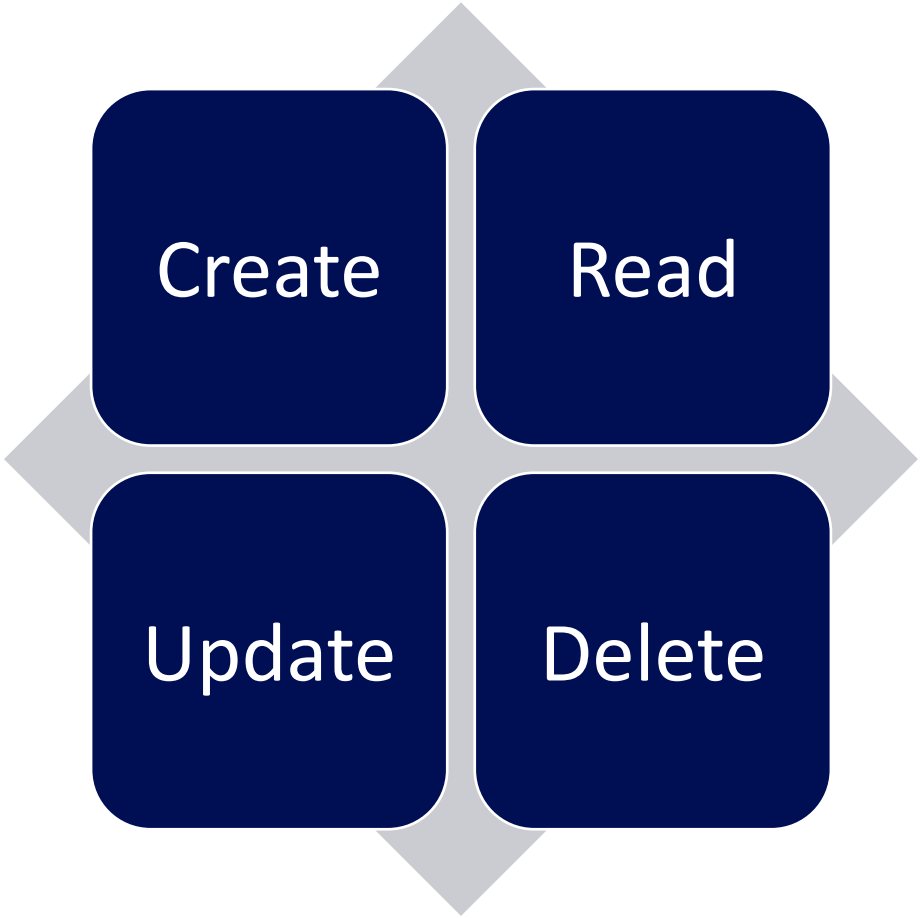
Syracuse University
School of Information Studies

Data Management

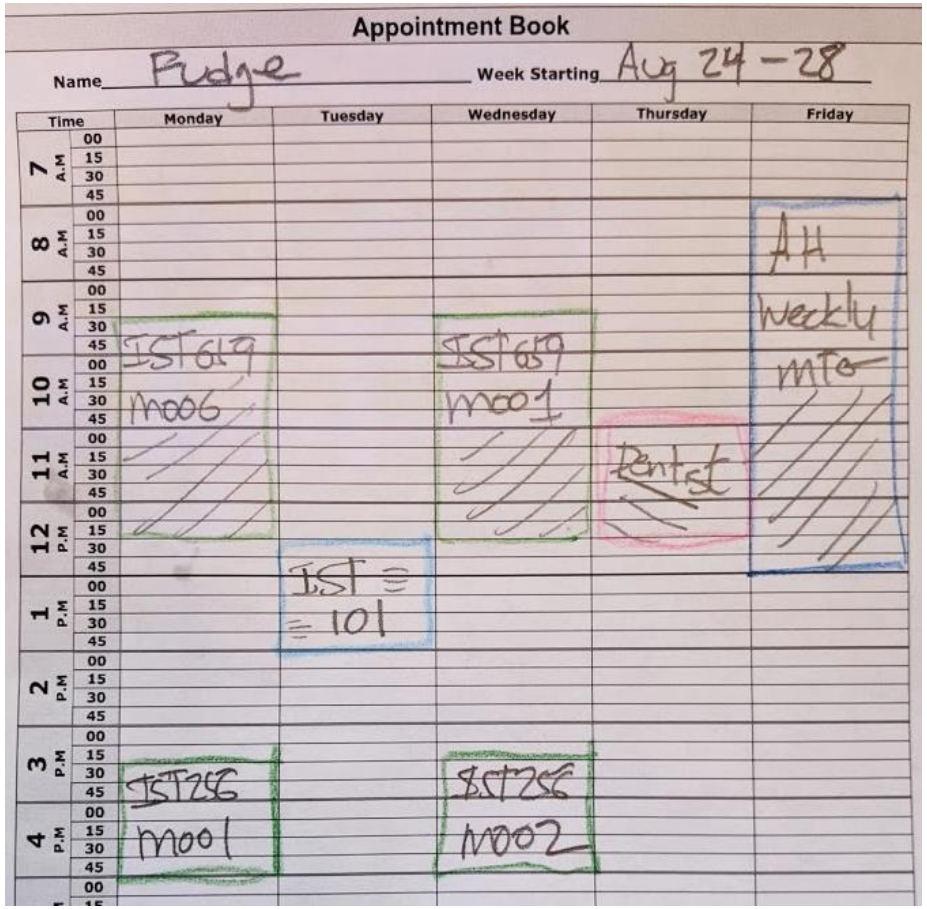


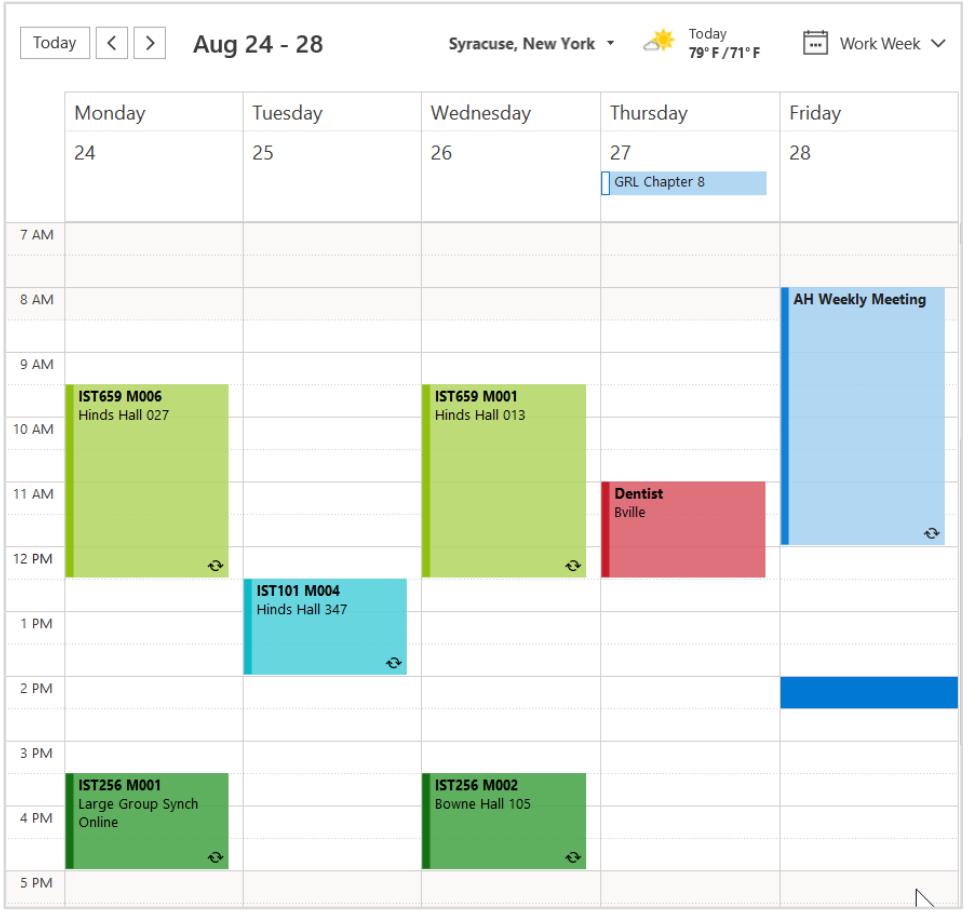
Data Management

- The process of storing, retrieving, and maintaining data
- Metadata helps by defining a storage format, access rules, and maintaining data integrity
- Four data management operations…

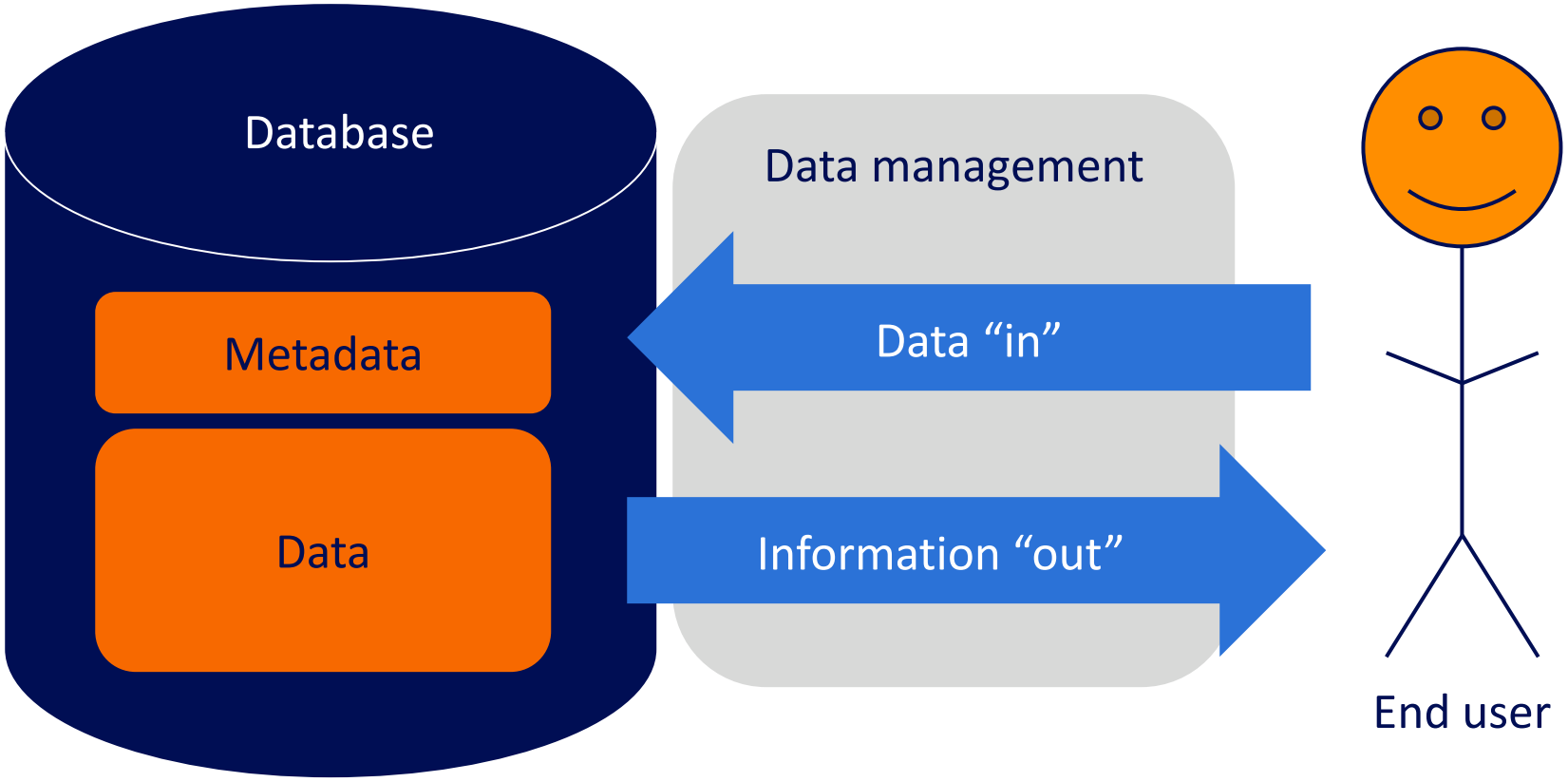


Computers Are Built for Data Management





Let’s Revisit the Database Definition



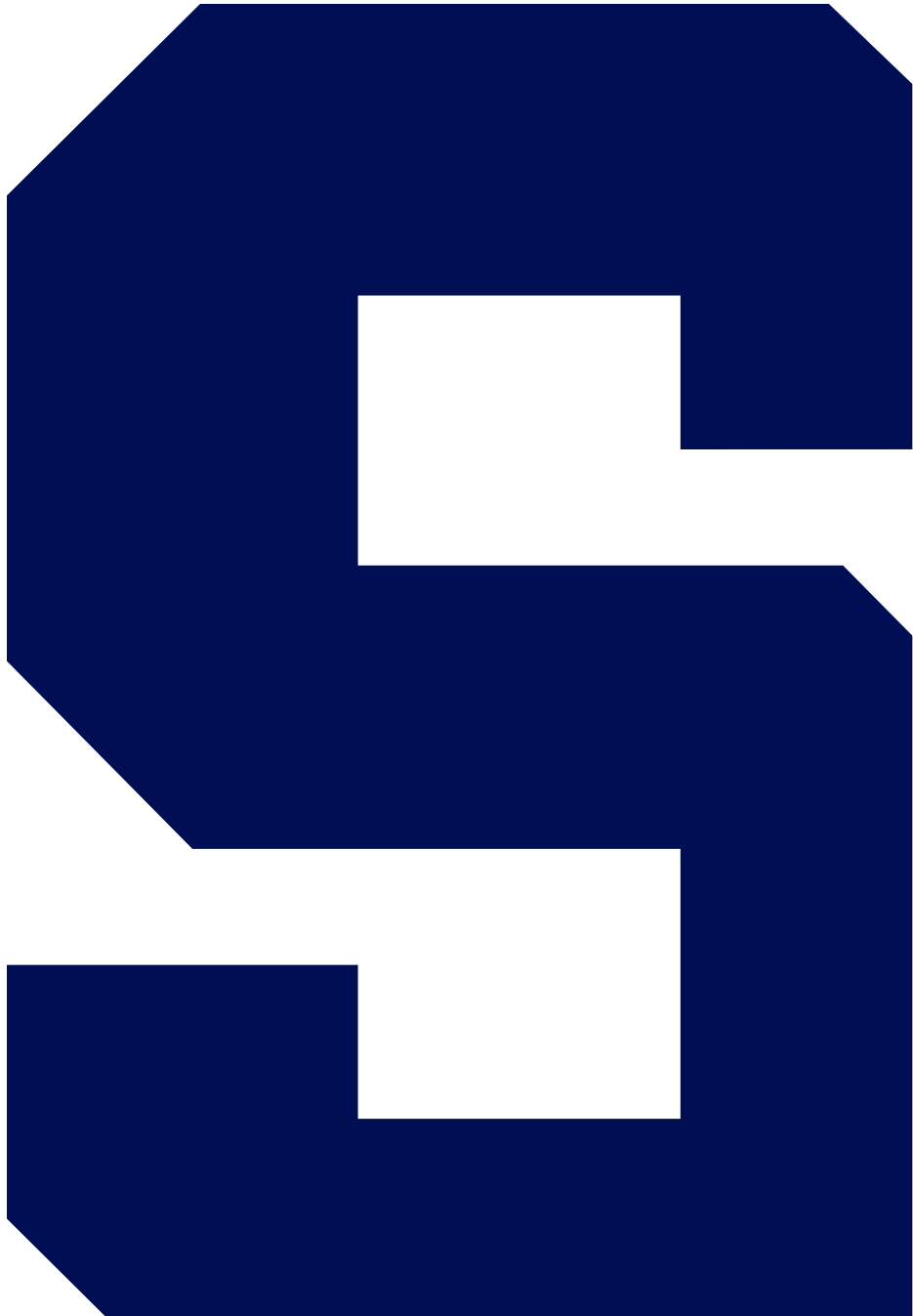
Syracuse University
School of Information Studies

Data Management

The End

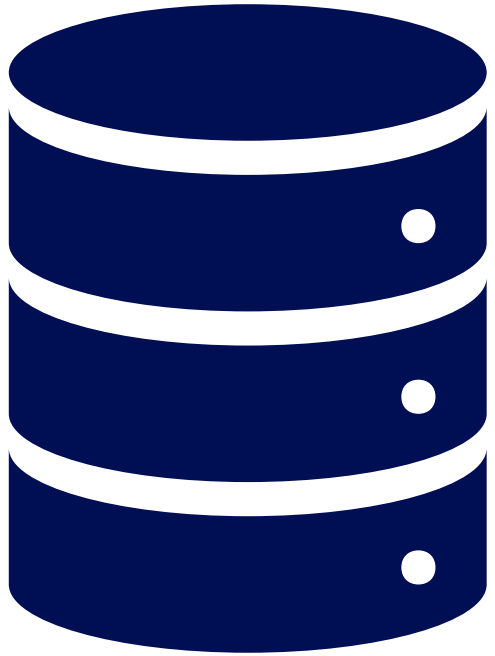
Syracuse University
School of Information Studies

Database Management System (DBMS)



Database Management System

Software suited to the task of database management



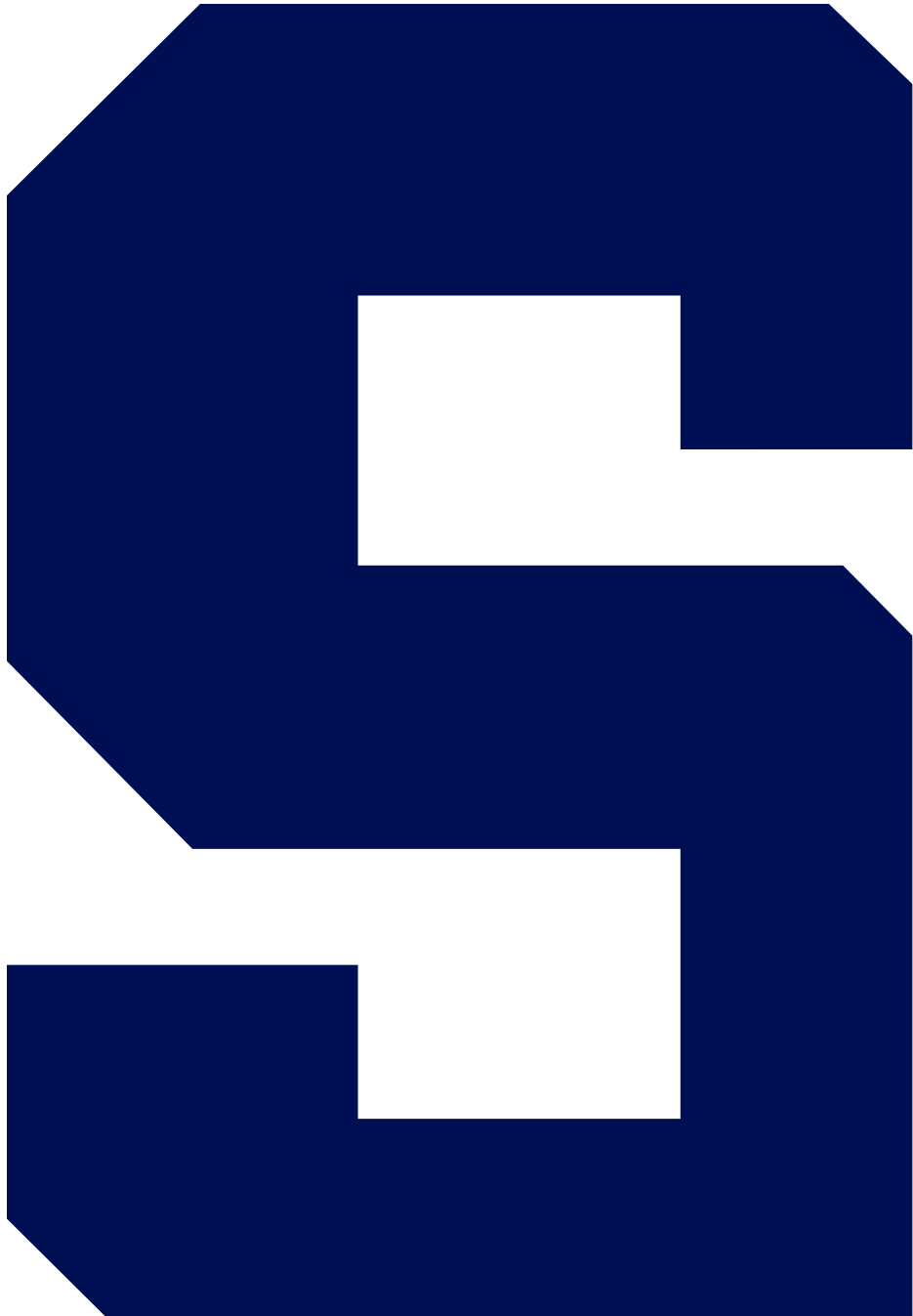
Common Features of the DBMS

- Metadata implementation
- Data management
- Query language
- Data independence
- Concurrency control
- Data security
- Backup and recovery management

Database Management System (DBMS)

The End

Syracuse University
School of Information Studies



DBMS Implementation Models

DBMS Implementation Data Models

- An implementation data model
  - Governs how the data are structured and stored within the database itself
  - Determines the design philosophy, capabilities, and limitations
- No one model is better than the other
- Each has its use cases and common applications

Popular Implementation Models

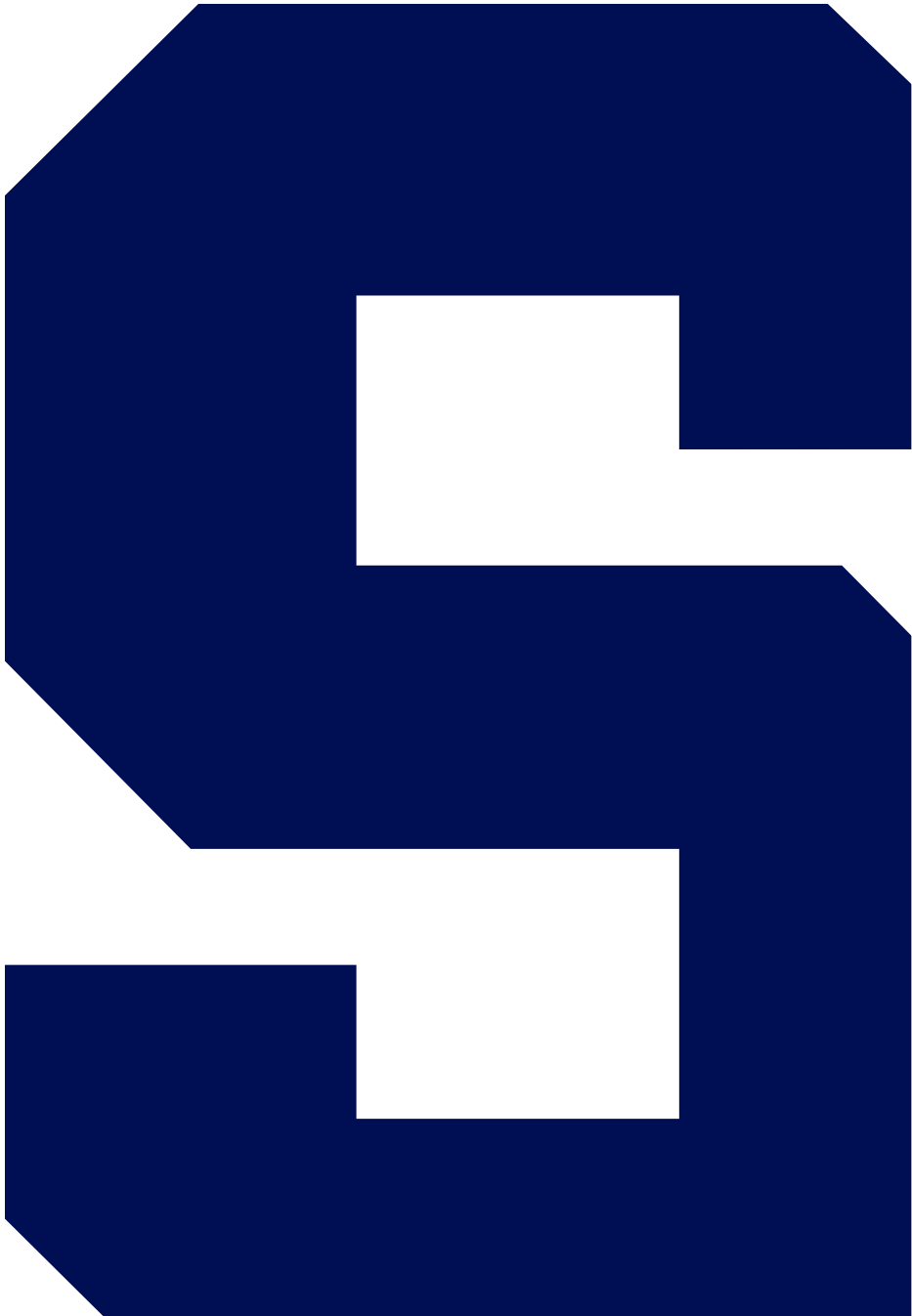
NoSQL

Name	Description	Use cases
Relational	Data are stored in structured tables of rows with metadata defining the columns; metadata defines how data in tables connect to one another	Business applications, multiuse
Key value	Data are stored under a key; information can be retrieved by key; little to no metadata	Caching, session management, real-time data
Document	Structured metadata is stored with data in a document; like documents are stored in collections	Content management, master data, search engines
Graph	Data are structured into nodes, edges, and labels; permits for complex relationships among data	Hierarchical data, networked data, social networks
Column-oriented	Tabular data structure with data in columns and metadata in the row, a computationally efficient structure for data analytics	Internet of things data, data analytics, data warehousing
Time series	Tabular data structure in time-order; data are immutable and support high-volume writes	Internet of things, time-oriented analysis, and forecasting

DBMS Implementation Models

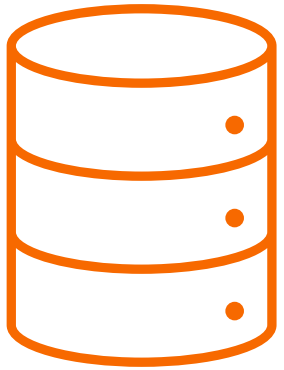
The End

Database Application Development



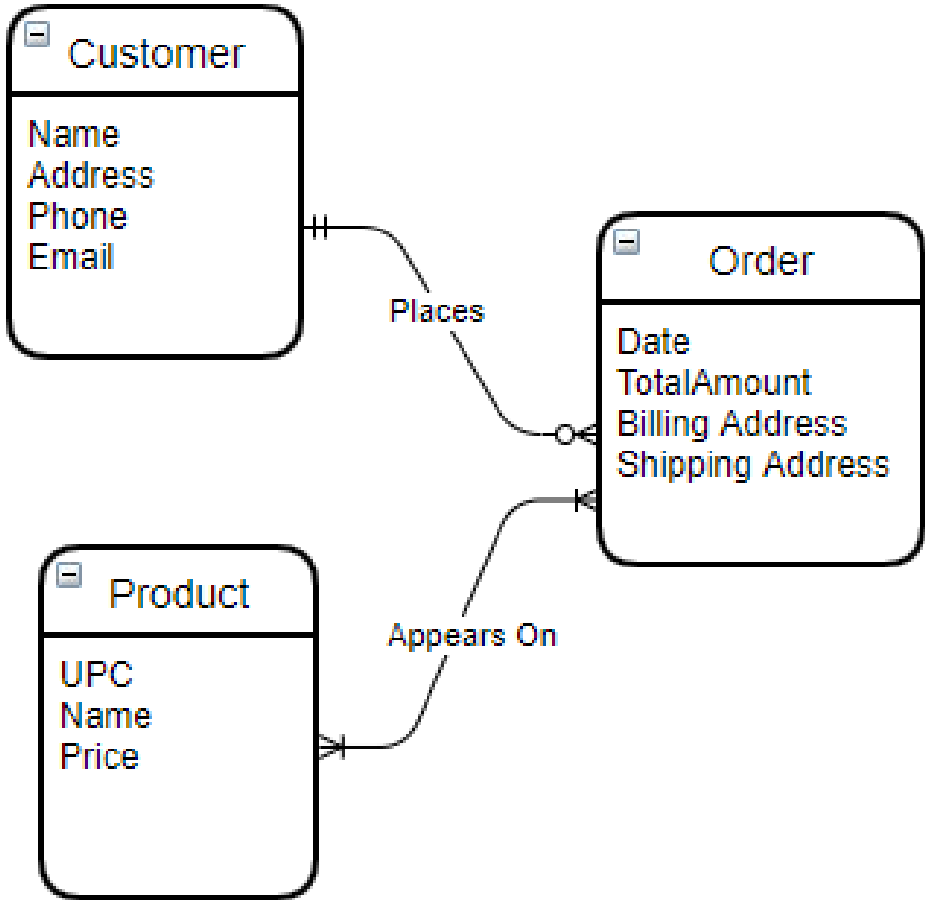
Database Application Development

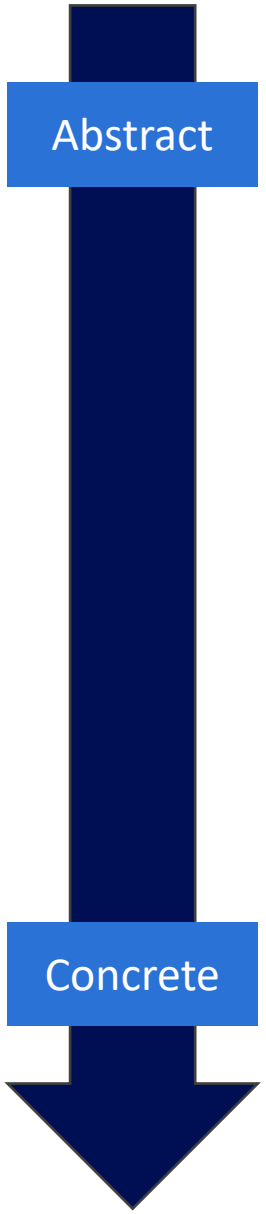
- All applications have data!
- Using a DBMS makes developing a database-driven application easier
- Since the DBMS trivializes data and metadata management
- But it starts with a good design
- In this course, we will learn design principles for creating database applications as well as database development



Data Models

- Various degrees of abstraction of the same database
- Important when within the context of database development





Conceptual data model

- An abstract representation of the data requirements
- No implementation of the database itself; no implementation model selected

Logical data model

- Mapping of the conceptual model to an implementation model
- No implementation of the database; model selected

Internal data model

- The internalized implementation of the database application
- A DBMS and implementation model selected

External data model

- User's view of the database application
- DBMS and implementation model selected, internal model implemented

Physical data model

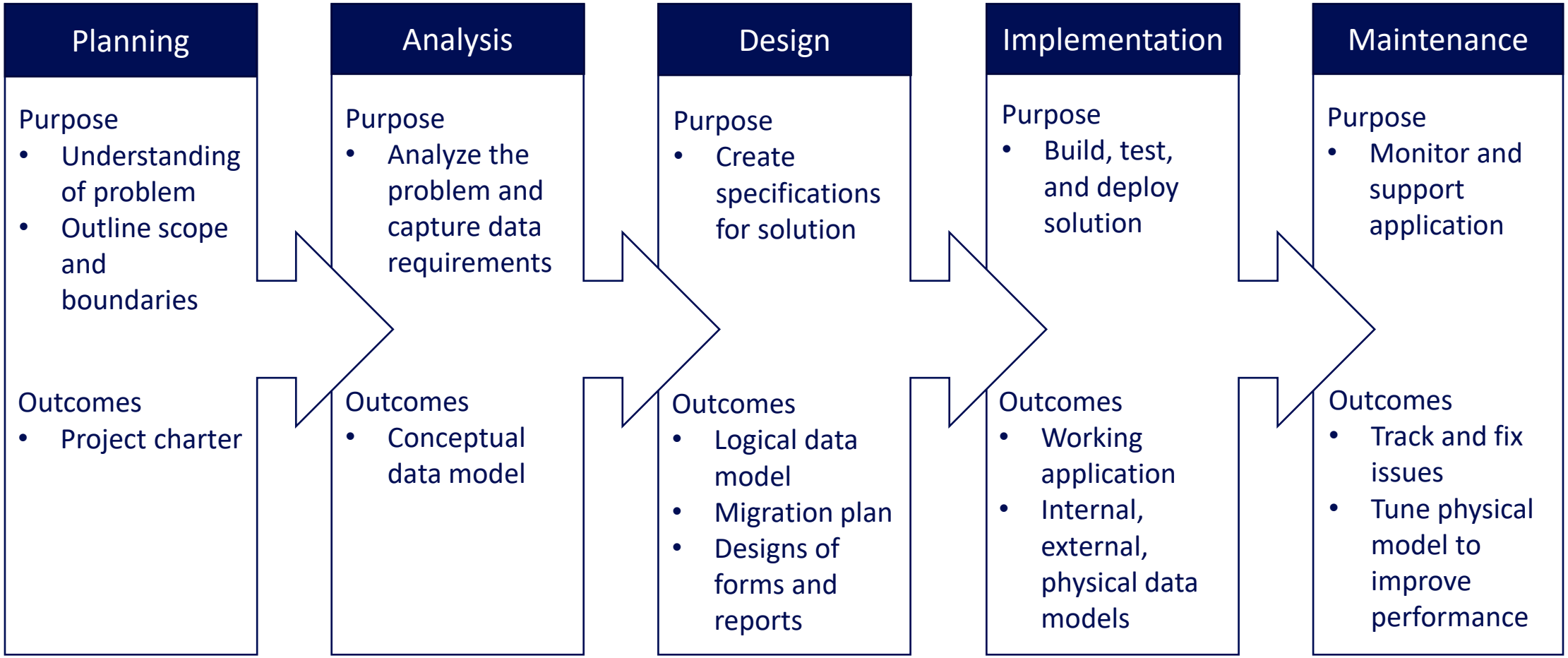
- How the internal/external model is stored by the DBMS and operating system
- All aspects of the database are an implementation

Data Models, an Example

Conceptual	Logical	Internal	External	Physical

Customer
Name
Email
Phone
	Customers
	ID
	First Name
	Last Name
	Email
	Phone
	Created On
	Last Update
		`create table customers (`
		`  id int identity`
		`  primary key,`
		`  first_name varchar(50),`
		`  last_name varchar(50),`
		`  email varchar(100),`
		`  phone char(10),`
		`  created_on datetime,`
		`  last_update datetime`
		`)`
			`create view v_customers`
			`as`
			`  select`
			`  first_name + ' ‘`
			`  last_name as name,`
			`  email,`
			`  phone`
			`  from customers`
				`create unique index`
				`  ix_customers_email`
				`  on customers (email)`
				
				`create index on`
				`  ix_customers_name`
				`  on customers (`
				`  last_name, first_name`
				`  ) include (`
				`  email, phone`
				`  )`

Database Development Life Cycle



Database Application Development

The End

Syracuse University
School of Information Studies

Summary

Summary

- Database is data plus metadata plus data management.
- Not all databases are computerized.
- A DBMS is software that helps to computerize databases.
- There are different implementation models of DBMS.
- Databases can be perceived at various levels, conceptually, logically, and as an implementation.
- These perceptions have their place in the database development life cycle.

Syracuse University
School of Information Studies

Summary

The End