


Business Information System

Unit – 2 Computer Files & Databases

By: Priya Patel

PICA – BCA

Files

- A **computer file** is a computer resource for recording data discretely in a computer storage device.
 - **Files** are the most important mechanism for storing data permanently on mass-storage devices. ***Permanently*** means that the data is not lost when the machine is switched off.
 - A file is an object on a computer that stores data, information, settings, or commands used with a computer program.
 - Files can be edited and transferred through the internet on that particular computer system.
- 

Various Types of Files

- **Text File:** A text file is a computer file that only contains text and has no special formatting such as bold text, italic text, images, etc. With Microsoft Windows computers text files are identified with the **.txt file extension**, as shown in the example picture.

Text files are also known as **flat files** or **ASCII** files.



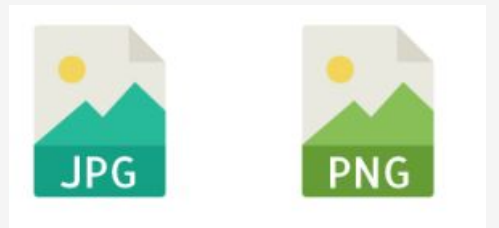
Various Types of Files

- **Binary File:** A binary file is a file stored in binary format. A binary file is **computer-readable but not human-readable**. All executable programs are stored in binary files, as are most numeric data files (Example: 0101).
- Extension of binary file is **.bin**



Various Types of Files

- **Image File:** Image files are composed of digital data (in pixel) in one of these formats that can be rasterized for use on a computer display or printer. A file that contains graphics data.
- Various Image File Types:
 - GIF (Graphic Interchange Format)
 - PNG (Portable Network Graphics)
 - JPEG (Joint Photographic Experts Group) also known as JPG
 - TIFF (Tagged Image File Format)
 - And more



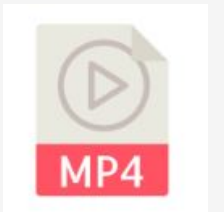
Various Types of Files

- **Audio File:** An audio file format is a file format for storing digital audio data on a computer system.
- Various Audio File Types:
 - MP3 (MPEG-1 Audio Layer 3)
 - WMA (Windows Media Audio)
 - WAV (Waveform Audio File Format)
 - AIFF (Audio Interchange File Format)
 - And more



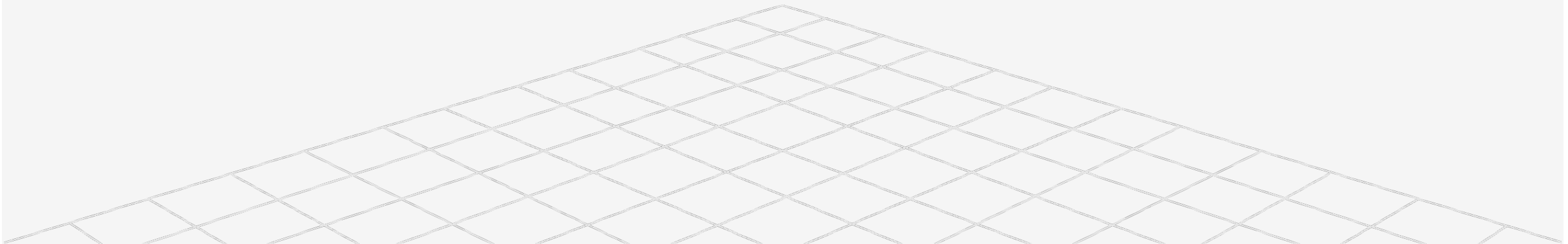
Various Types of Files

- **Video File:** A video file format is a type of file format for storing digital video data on a computer system.
- Various Audio File Types:
 - MP4 (MPEG Audio Layer 4)
 - 3GP (3rd Generation Partnership)
 - WMV (Windows Media Video)
 - FLV (Flash Video)
 - And more




Various Types of Files

- There are so many types of files available like HTML File, C/C++/Java File, Photoshop File, PDF File, MS Office File and many more.



Types of File Operations

- 1) Read Operation: Meant To Read the information which is Stored into the Files.
 - 2) Write Operation: For inserting some new Contents into a File.
 - 3) Rename or Change the Name of File.
 - 4) Copy the File from one Location to another.
 - 5) Sorting or Arrange the Contents of File.
 - 6) Move or Cut the File from One Place to Another.
 - 7) Delete a File
 - 8) Execute Means to Run Means File Display Output.
- 

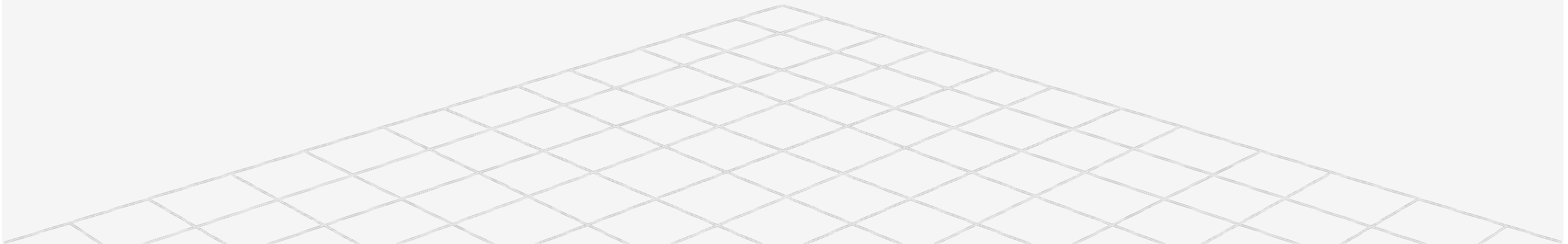
Directory

- A directory is a location for storing files (documents, spreadsheets, images, vidéo, audio, etc) on your computer. Directories are found in a hierarchical file system.
- In the picture to the right is an example of the tree command output that shows all the local and subdirectories (e.g., the "big" directory in the cdn directory). When looking at this overview, the C: drive is considered the current directory and root directory

```
C:.\nbanners\nbtips\ncdn\n|   big\n|   people\n|   tips\ncomp\n|   logos\ndownload\ngames\n|   sanitar\n|   t7g\nhardware\ninternet\nscreen\nsharewar\nupdates\nutility\nwin2000\nwin95\nwin98\nwindll\nwinme\nwinnt\nwinxp\ndrivers\nfeed\nficon\ngame\n|   images\n|   creatures\n|   items\n|   npc
```

Databases

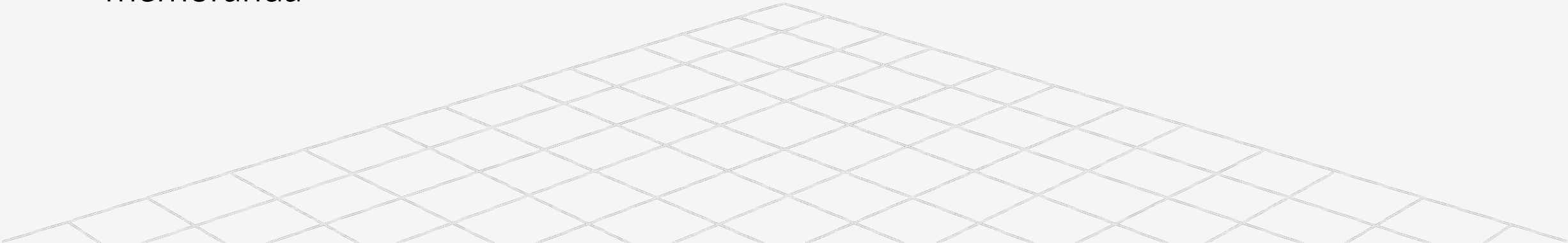
- A database is an organized collection of data, generally stored and accessed electronically from a computer system.
- A **relational database** is, simply, a database that stores related information across multiple tables and allows you to query information in more than one table at the same time.



Business Files

- A business file is a document (hard copy or digital) that records a business dealing.

Business documents include:

- employment contracts
 - accounting source documents
 - Minutes of meeting, and
 - memoranda
- 

Business Files

Different records kept by businesses

- **Accounting Records:**
 - Business transaction – income , expense and equity , tax records
 - Help in development of financial statement
- **Bank Statements:**
 - records of your checking, savings, investments, and credit cards
 - accounting records, bank statements help you track your
 - business's progress

Business Files

- **Legal Documents:**

- legal documentation are mandatory if operate under different business structures. (partnership agreement)
- Sole proprietors and LLCs also have legal documents. legal documents in business records as proof that you own the company

- **Permits and Licenses:**

- Business location and industry may require to have a permit
 - need documentation of licenses to show follow regulations.
- 

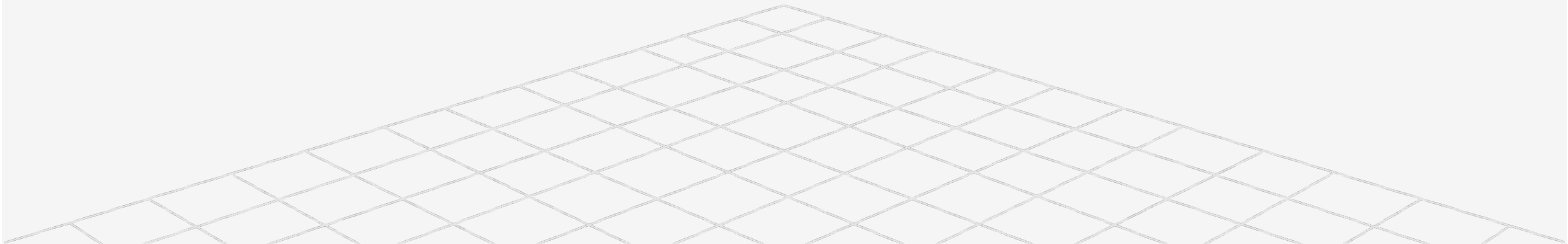
Business Files

- **Insurance Documents:**

- General business liability insurance protects your business from losses.

- **Client Files:**

- electronic files for clients and projects
- record of all business arrangements, agreements, contracts and clients they have dealt with




Business Files

- **Tax and Financial Files:**

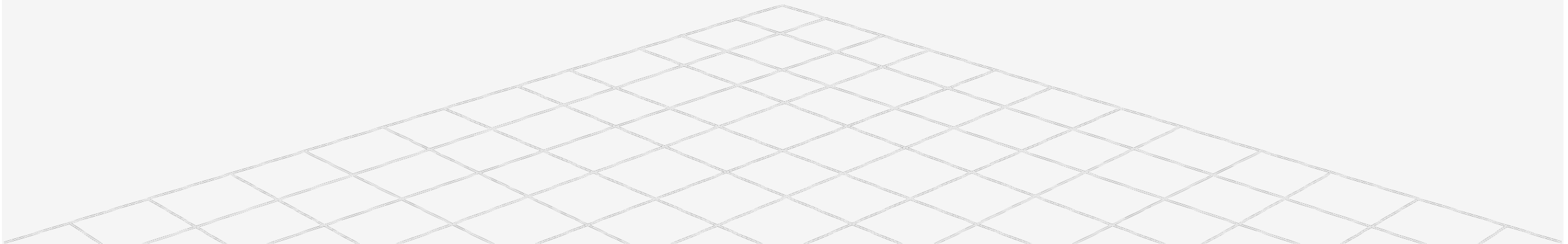
- files should include payroll, business expenses, credit card statements, bank statements, annual tax returns, quarterly tax returns, inventories, cash register tapes, travel logs and sales and income statements.

- **Employee Files:**

- Files for employees, including subcontractors and freelancers
 - files should include contracts, tax statements, payroll records, insurance and any other forms the employee signed during his time with the company.
- 

Business Files

- **Business Plan:**
 - Business plans are considered how-to guides for the owner to ensure her business stays on track and succeeds.
 - business plans and any budgets created from those plans.



Data Storage Media

A **storage medium** is any material on which data can be electronically placed, kept, and retrieved. Example: CD, DVD, Floppy Disk, Hard Disk, etc.

Types of Storage Device:

1. **Primary Storage:** A primary storage device is a medium that holds memory for short periods of time while a computer is running. Although it has a much lower access time and faster performance.

Example: RAM (Random Access Memory) & Cache Memory

- It is directly accessible to CPU
- CPU continuous reads, stores & execute instructions as requires

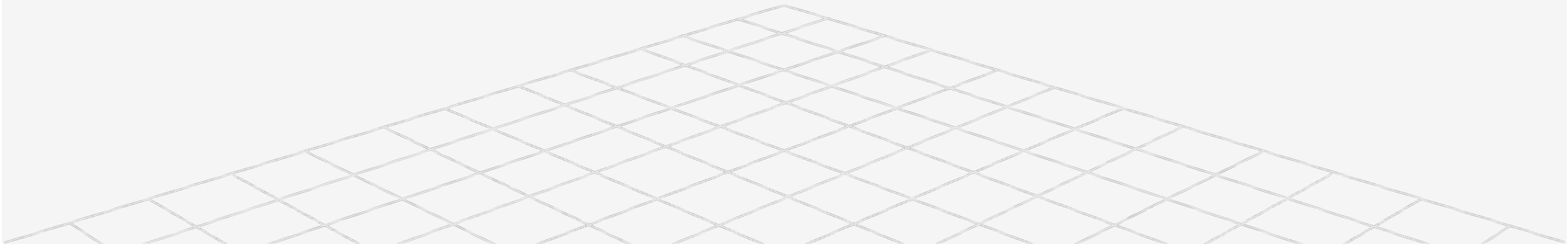
Data Storage Media

2. Secondary Storage: Alternatively referred to as **external memory** and **auxiliary storage**, a secondary storage device is a non-volatile device that holds data until it is deleted or overwritten.

- Secondary storage is cheaper than primary storage.

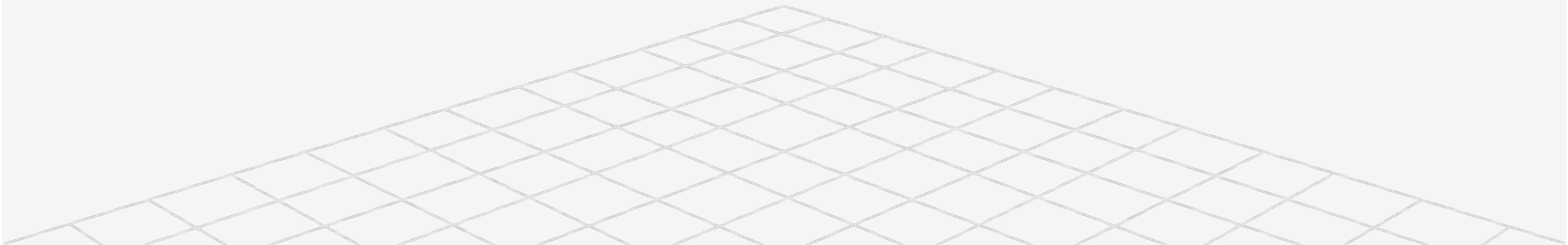
Example: Hard Disk, CD, DVD, etc

- Usually uses input/output channels to access secondary storage



Types of Storage Medium

1. Semiconductor Storage
2. Magnetic Storage
3. Optical Storage
4. Paper Data Storage



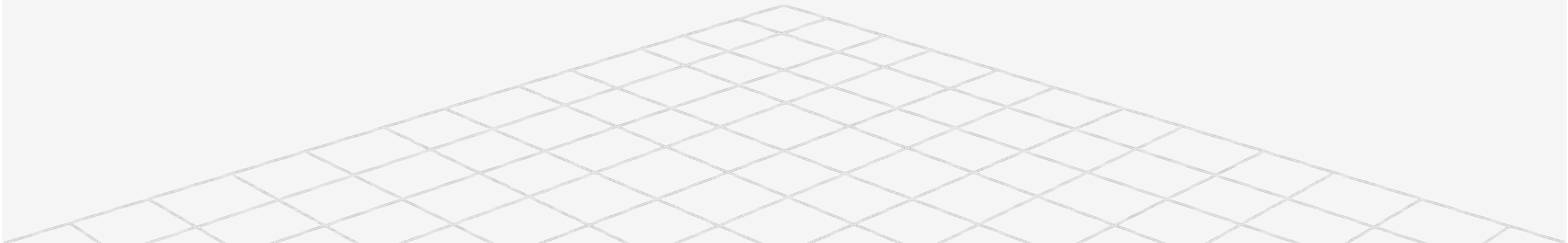
Types of Storage Medium

1. **Semiconductor Storage:** Semiconductor memory is a digital electronic data storage device, often used as computer memory, implemented with semiconductor electronic devices on an integrated circuit (IC).

Example:

RAM (Random Access Memory)

ROM (Read Only Memory)



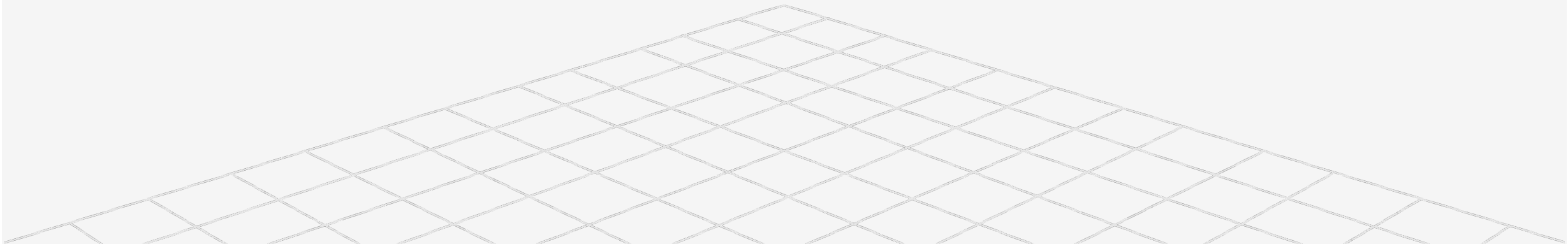
Types of Storage Medium

2. Magnetic Storage: Magnetic tape is a medium for magnetic recording, made of a thin, magnetizable coating on a long, narrow strip of plastic film.

Example:

Floppy Disk

Hard Disk



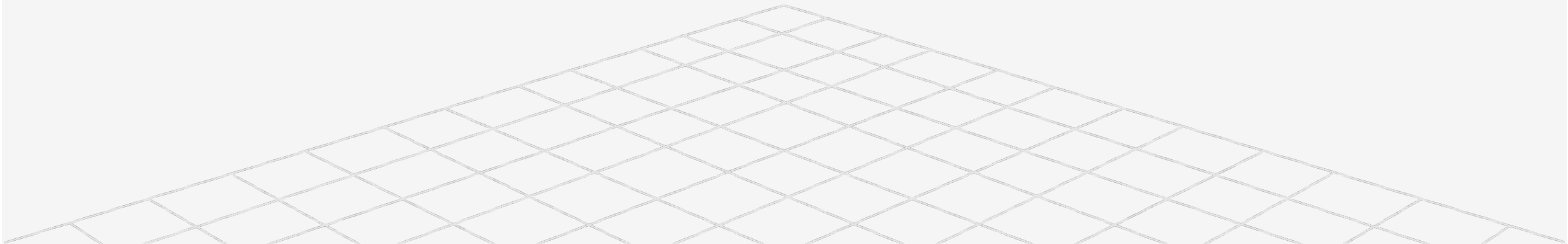
Types of Storage Medium

3. Optical Storage: Any storage method that uses a laser to store and retrieve data from optical media.

Example:

CD (Compact Disk)

DVD (Digital Video Disk)

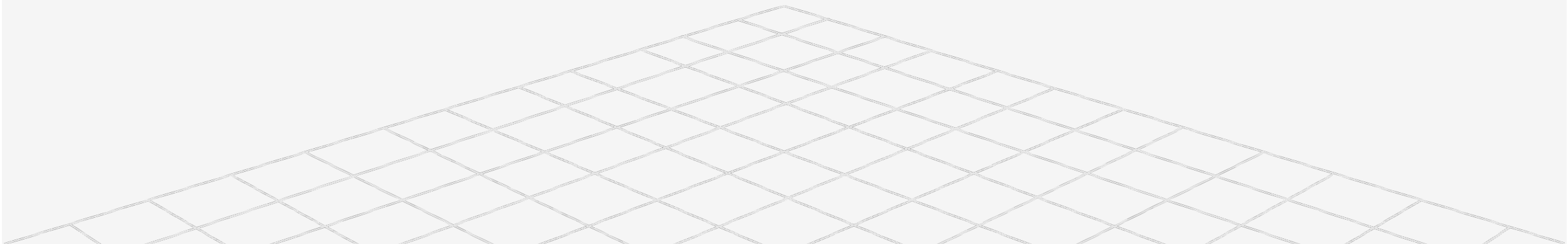


Types of Storage Medium

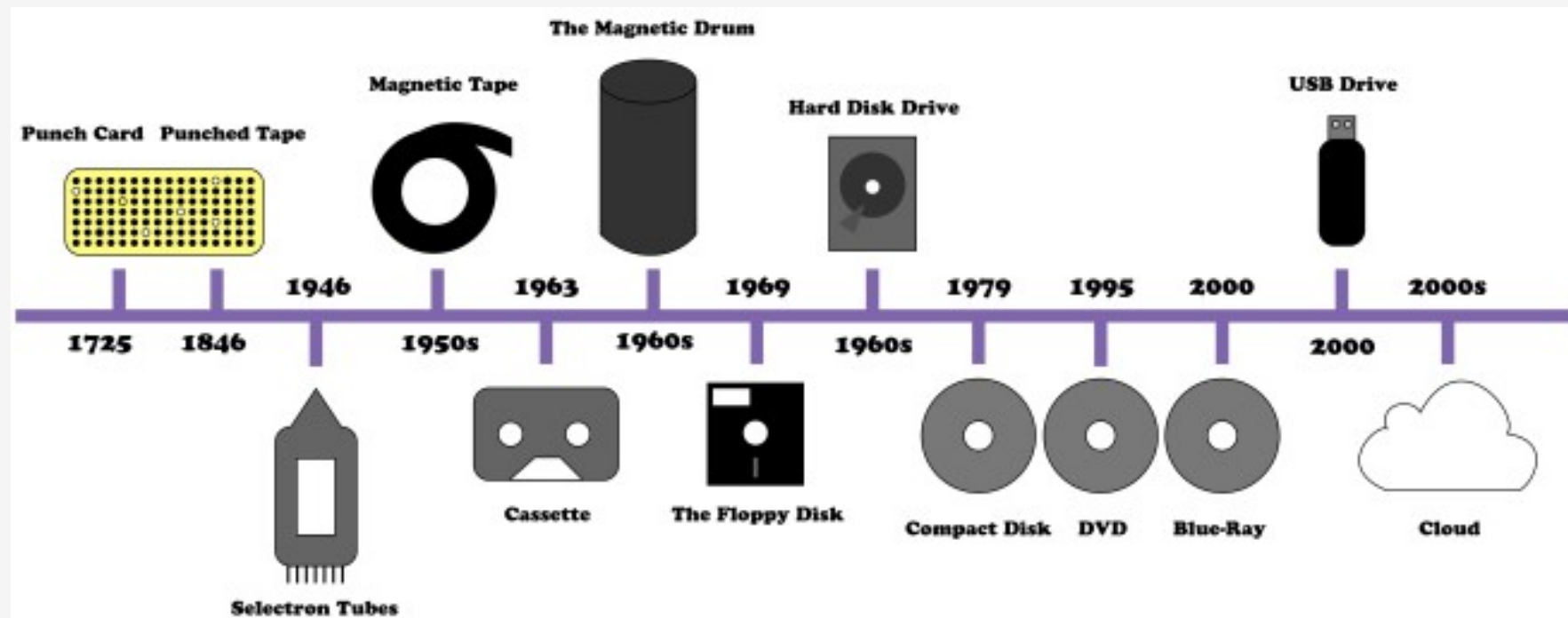
4. Paper Data Storage: Paper data storage refers to the use of paper as a data storage device. This includes writing, illustrating, and the use of data that can be interpreted by a machine or is the result of the functioning of a machine.

Example:

Punch Cards



Timeline of Storage



Data Base Management System (DBMS)

Database: A database is an organized collection of data, generally stored and accessed electronically from a computer system. It is a collection of related data which represents some aspect of the real world.

DBMS: A database management system (DBMS) is system software for creating and managing databases. The DBMS provides users and programmers with a systematic way to create, retrieve, update and manage data. A DBMS makes it possible for end users to create, read, update and delete data in a database.

- A very large, integrated collection of data.
- Models real-world enterprise.
 - Entities (e.g., students, courses)
 - Relationships (e.g., ABC is a employee of XYZ company)

Why DBMS?

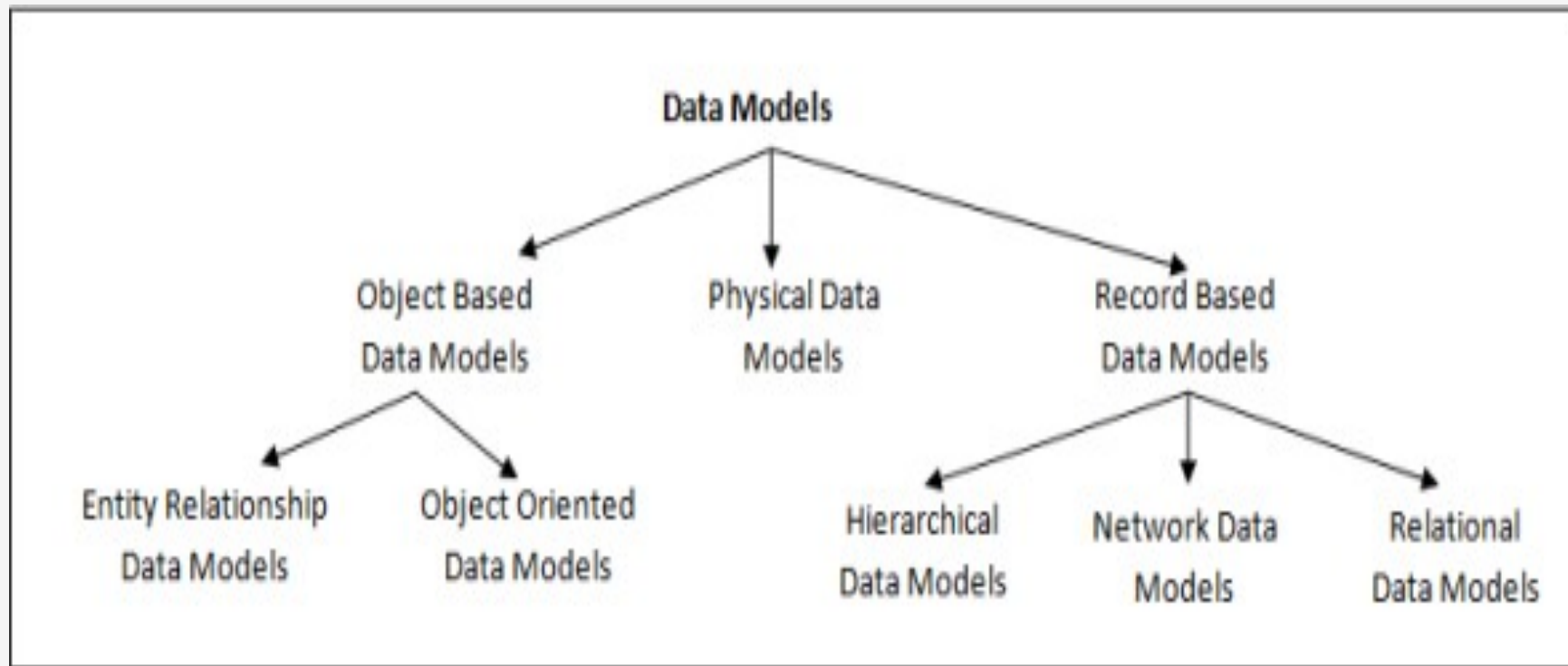
- Replication control
 - Concurrent access, recovery from crashes
 - Uniform data administration
 - Data integrity and security
 - Data independence and efficient access
 - Reduced application development time
- 

Data Models

- Databases are developed for a particular requirements.
- Planning the structure of databases is called **data models**.
 - It involves planning about tables, their columns, mapping between tables
- A data model helpsto put the real world requirement into a design
- Helps developer understand relationship between different objects.
- And correctany drawbacks in the design stage itself.

Data Models

Data Model categorization



Object Based Data Model

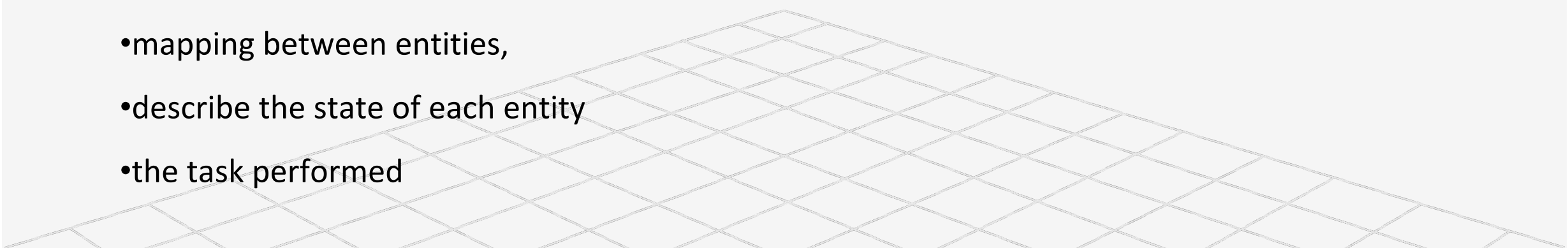
- From user requirement: Identify the entities involved
- List attributes of all entities with their attributes and relationships

Two types of object based data models

1. Entity Relationship Model

- basis of design of database
- defines mapping between entities

2. Object oriented data model

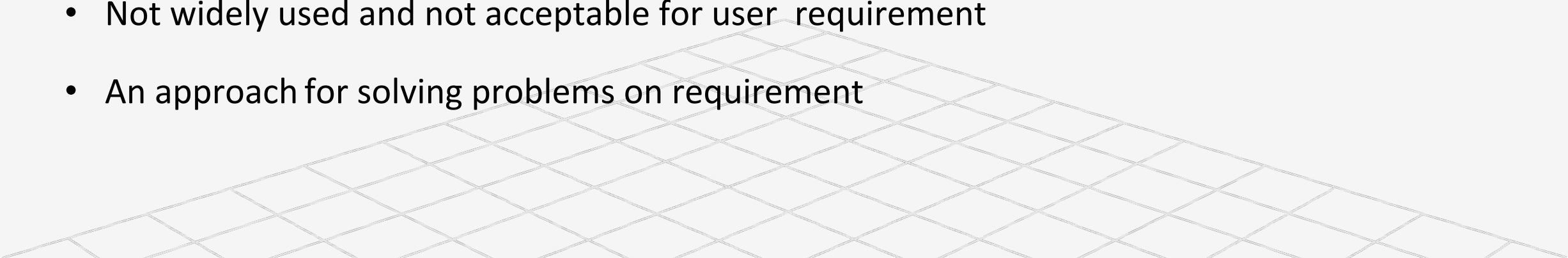
- mapping between entities,
 - describe the state of each entity
 - the task performed
- 

Object Based Data Model

Advantages

- Re-use the attributes and functionalities
- Flexible in case of any changes
- Each class binds attributes and its functionality – very close to real world entity

Disadvantages

- Not widely used and not acceptable for user requirement
 - An approach for solving problems on requirement
- 

Object Based Data Model

Entity Relationship Model:

- It is a pictorial representation of real world objects
- Involves various symbols and notations to draw diagrams

Advantages:

- Requirements are easy to understand
- ER diagrams can be easily converted to relational model

Disadvantages:

- Meant for high level designs(not coding level)

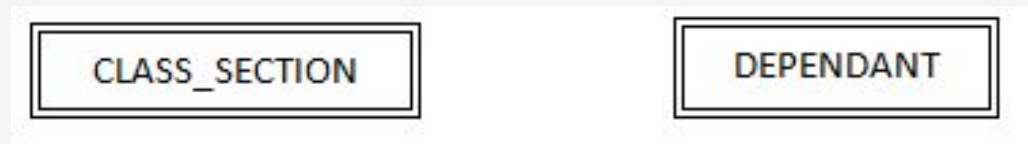
Object Based Data Model

Entity: It is a real world object. Name written inside rectangle.

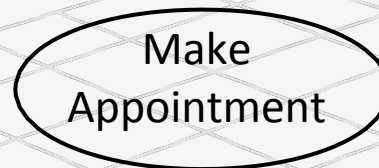
- **Strong entity - simple rectangle**



- **Weak entity – two rectangle**

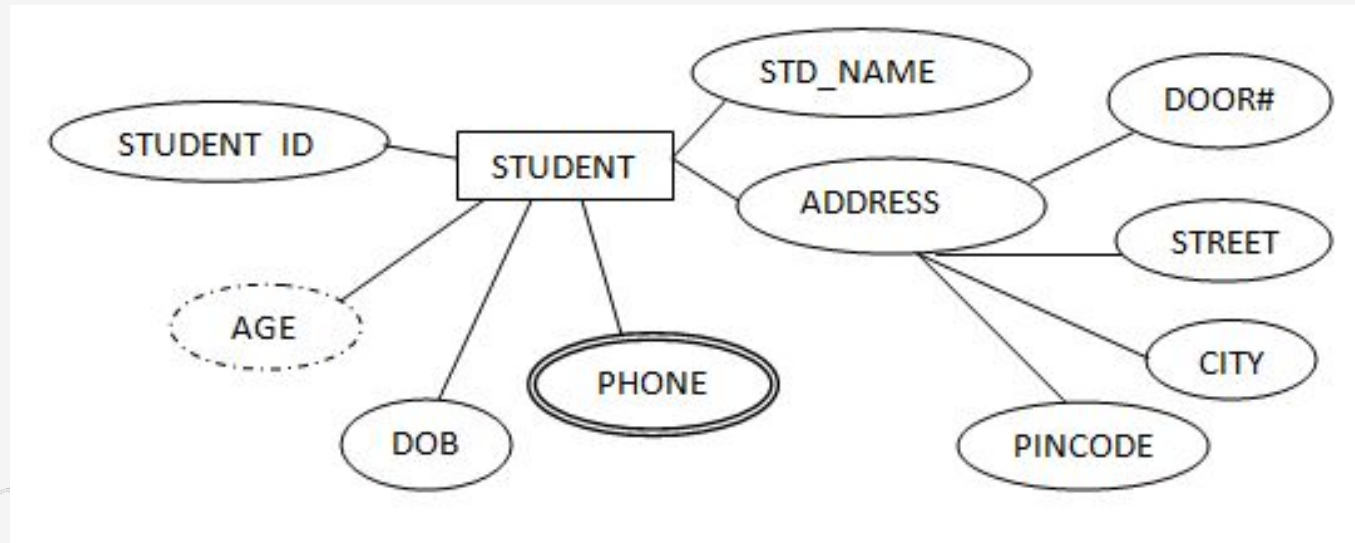


- **Attributes:** It is a summary of scenarios for a single task or goal. Represented by oval.



Object Based Data Model

- Multi-valued attribute – double oval
- Derived attribute – dashed oval
- Composite attribute – connected to parent attribute in tree structure

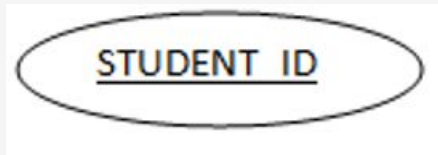


Object Based Data Model

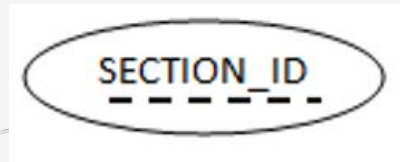
Primary key: It is a special relational database table column (or combination of columns) designated to uniquely identify all table records.

A primary key's main features are: It must contain a **unique value** for each row of data. It **cannot contain null values**.

- Underline to attribute in entity is primary key.

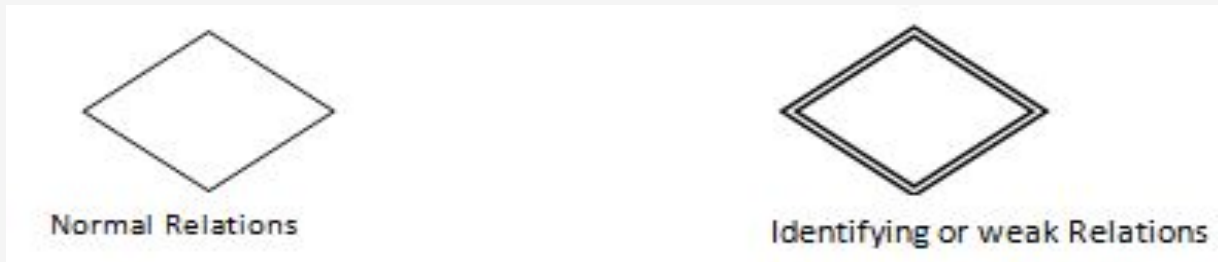


- For weak entity primary key is used as dashed line.



Object Based Data Model

Relation: Represented by diamond symbols.



Object Based Data Model

Cardinality of relationship:

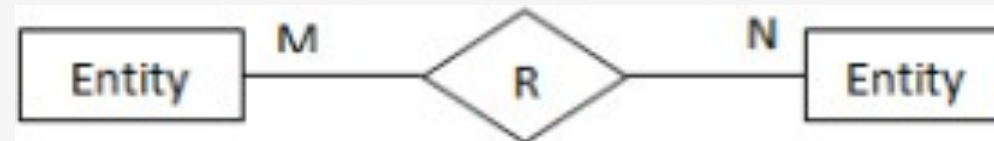
- One to one



- One to many



- Many to many

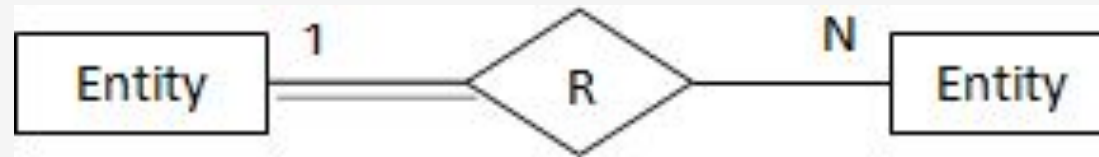


Object Based Data Model

Participation Constraint:

Total participation – double line

Partial participation – single line



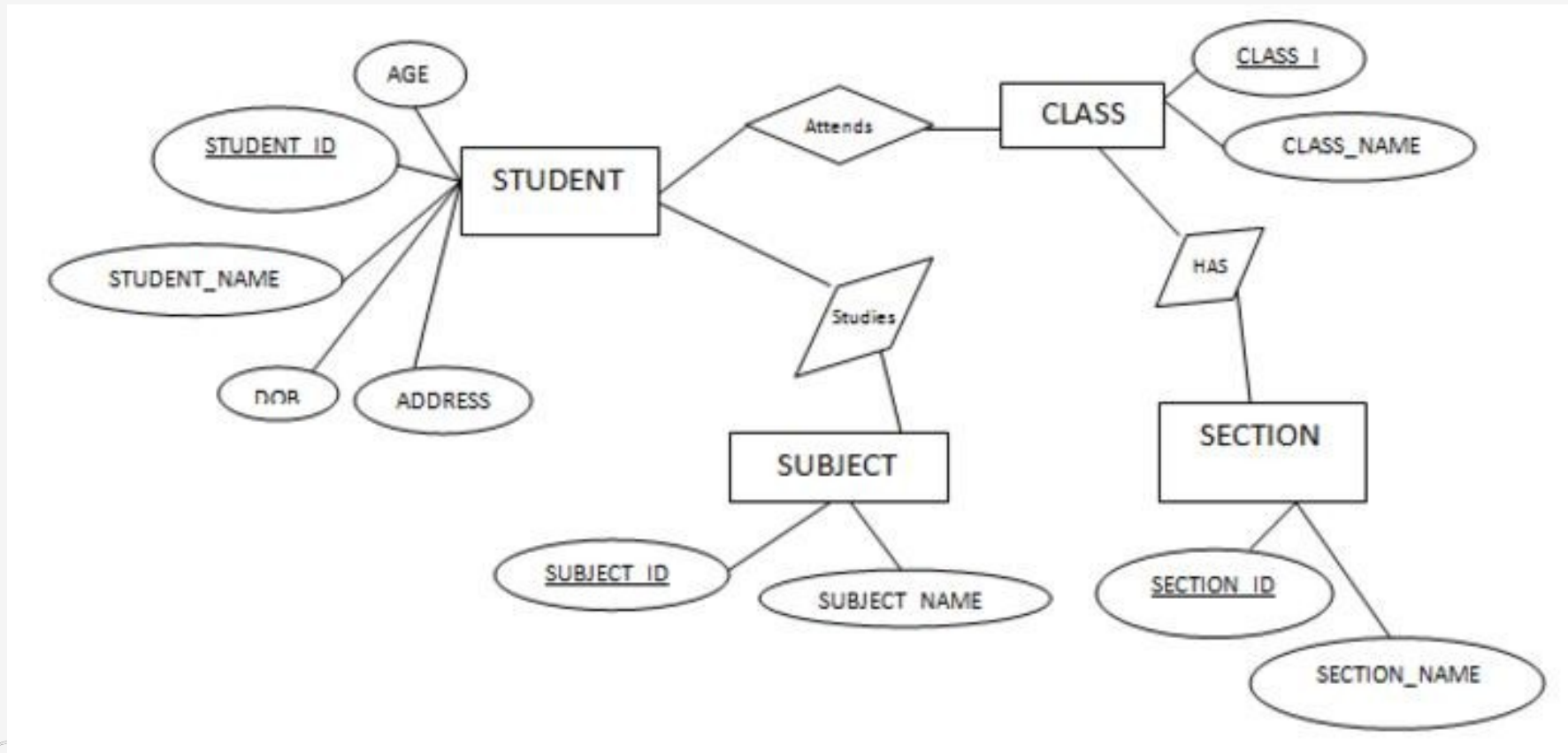
Object Based Data Model

Student database

STUDENT	CLASS	SECTION	SUBJECT
<u>STUDENT_ID</u>	<u>CLASS_ID</u>	<u>SECTION_ID</u>	<u>SUBJECT_ID</u>
STUDENT_NAME	CLASS_NAME	SECTION_NAME	SUBJECT_NAME
ADDRESS			
DOB			
AGE			
CLASS_ID			
SECTION_ID			

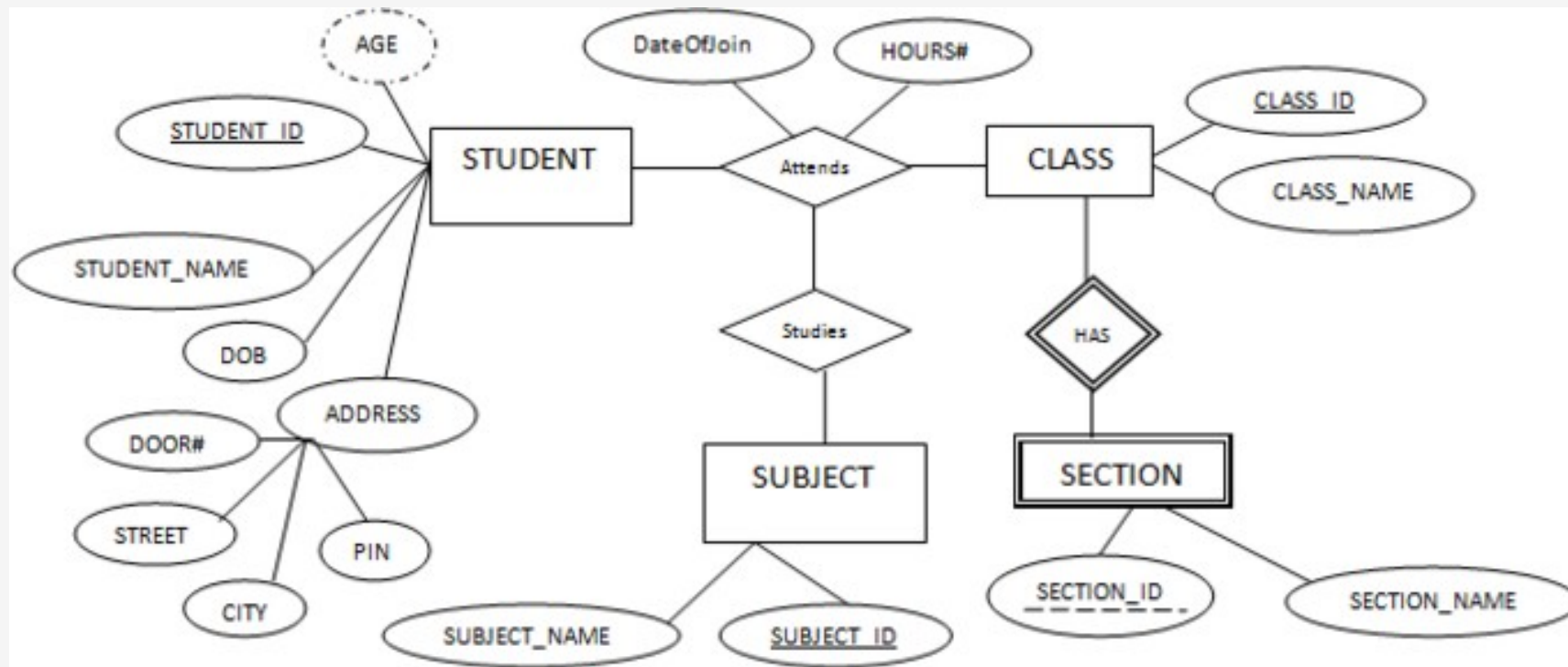
Object Based Data Model

ER Diagram for Student Data Base



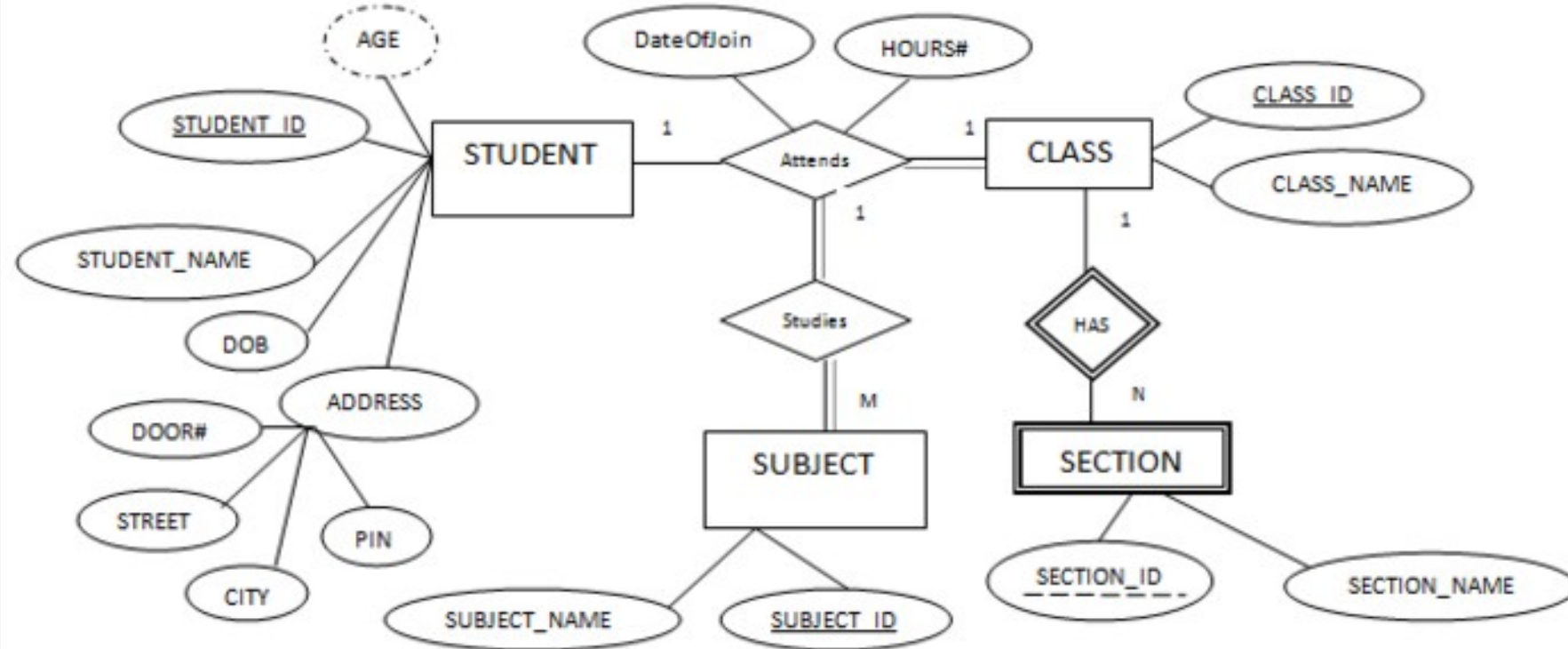
Object Based Data Model

ER Diagram for Student Data Base



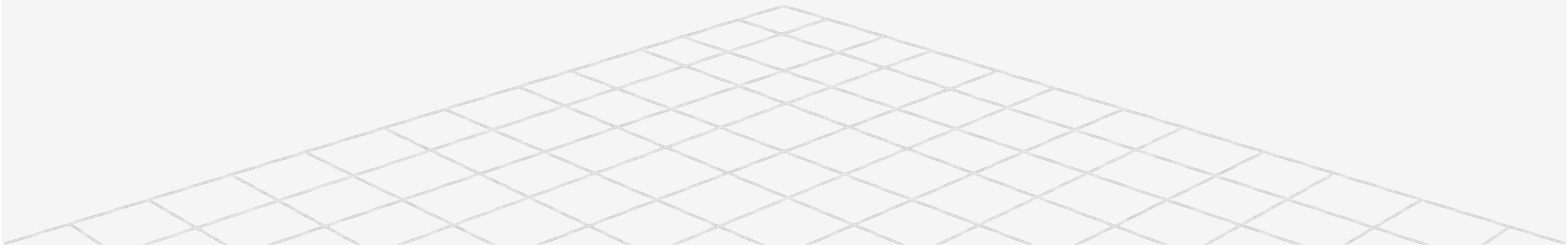
Object Based Data Model

ER Diagram for Student Data Base



Physical Data Model

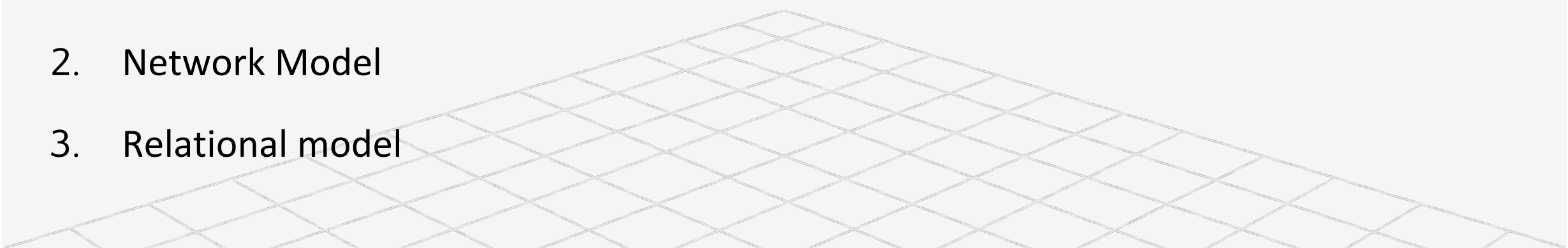
- it describes how data are stored in computer memory,
- how they are scattered and/or ordered in the memory,
- and how to retrieve from memory



Record Based Model

- Based on application and user levels of data (logical structure of objects in database)
- Define actual relationship between data in the entities.

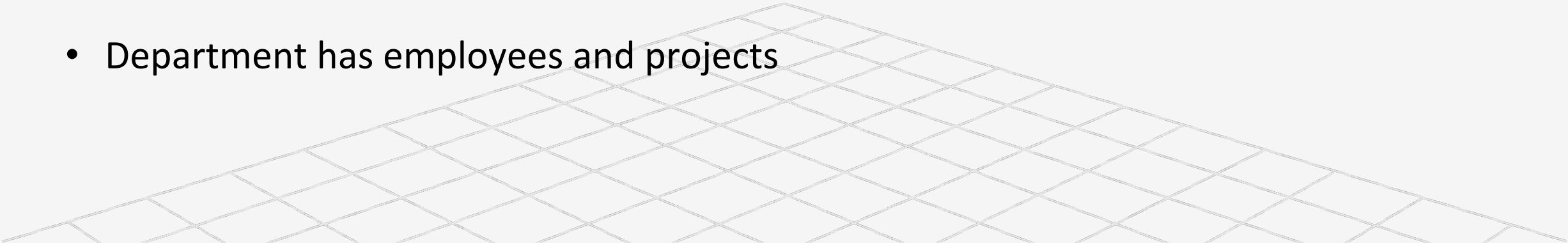
Three types of record based models

1. Hierarchical Model
 2. Network Model
 3. Relational model
- 

Record Based Model

1. **Hierarchical Model:** Observe each entity in terms of parent-child relationship.

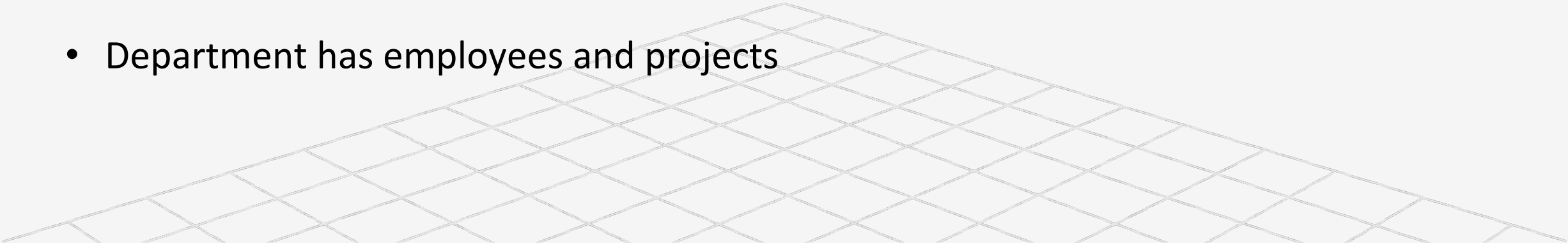
Database for a company – entities involved are:

- Company, its department, its supplier, its employees, different projects of company
 - Model them in terms of parent-child
 - Company parent for all other entities
 - Department has employees and projects
- 

Record Based Model

1. **Hierarchical Model:** Observe each entity in terms of parent-child relationship.

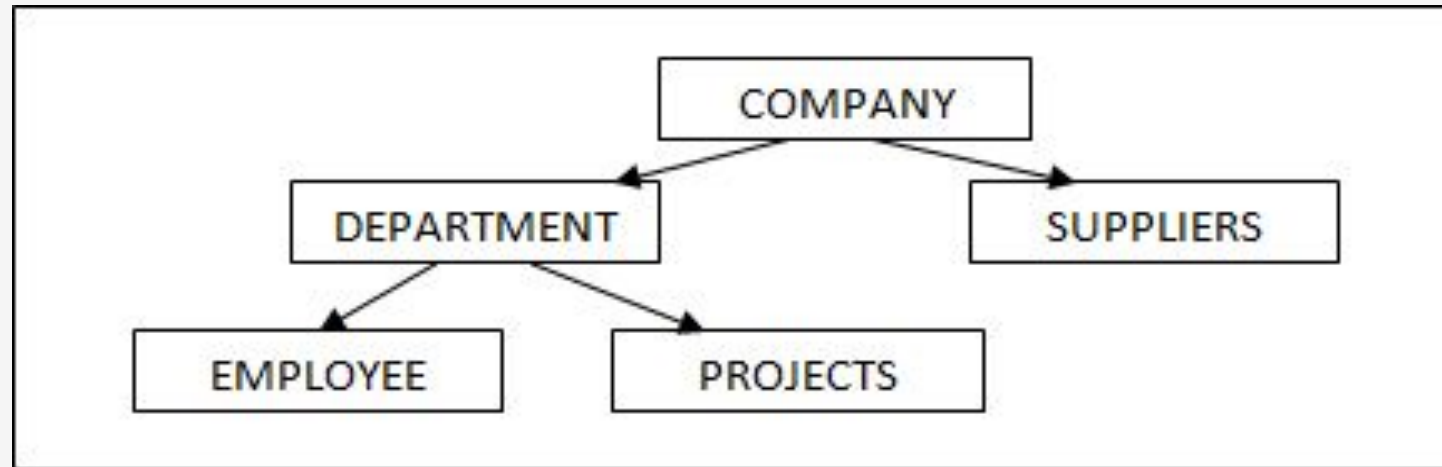
Database for a company – entities involved are:

- Company, its department, its supplier, its employees, different projects of company
 - Model them in terms of parent-child
 - Company parent for all other entities
 - Department has employees and projects
- 

Record Based Model

Hierarchical Model:

Database for company - entities involved are

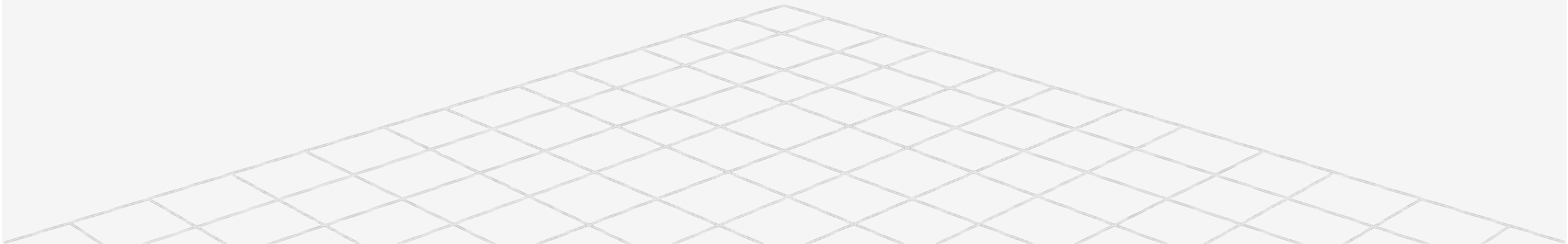


Record Based Model

Network Data Model: It helps to address M:N relationship. This model will not have single parent concept. Any child in the tree can have multiple parents here.

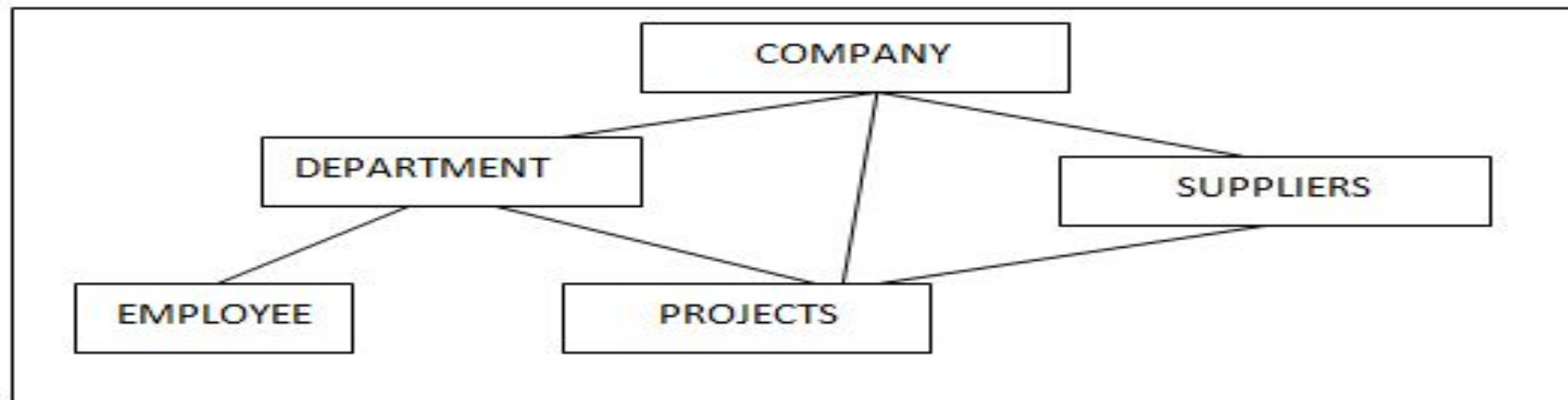
Database for a company:

- Company has many projects and
- Departments own those projects
- Sometimes suppliers give input for projects



Record Based Model

Network Data Model



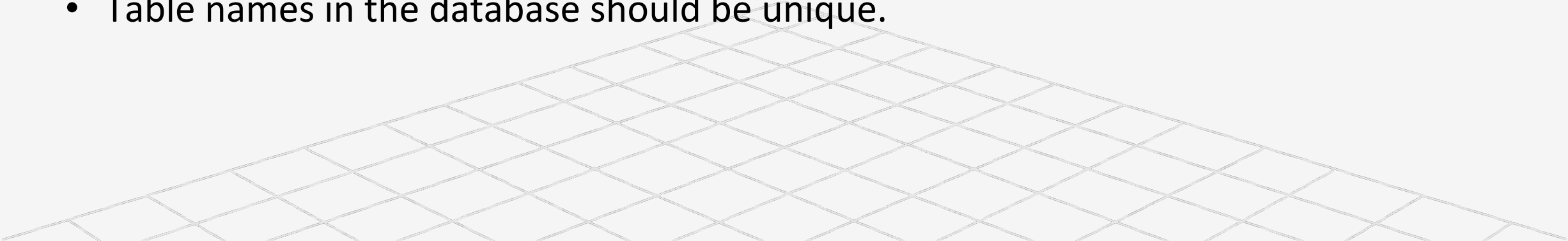
Record Based Model

Relational Data Model:

- isolates physical structure from the logical structure.
- based on the mathematical concepts of set theory
- tables as a 2-D table with rows and columns
- How data can be represented in rows and columns
- the way to establish the relation between other tables.

Record Based Model

Relational Data Model (Five Rules):

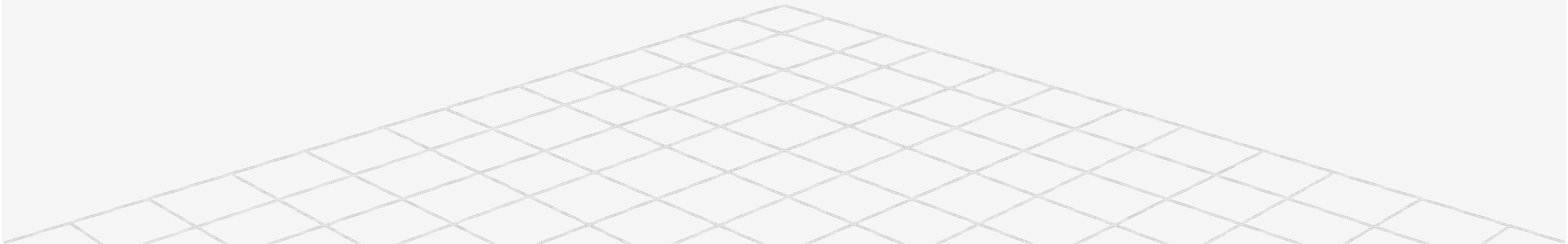
- All information is logically represented as tables
 - Each record in the table is unique.
 - Each column/attribute will have single value in a row
 - All attributes should be from same domain.
 - Table names in the database should be unique.
- 

Record Based Model

Relational Data Model:

Database for a company – employee and department

- Employee_id and Dept_id primary key
- Department details
- Employee details



Record Based Model

Relational Data Model:

EMPLOYEE				DEPARTMENT	
EMP_ID	EMP_NAME	ADDRESS	DEPT_ID	DEPT_ID	DEPT_NAME
100	Joseph	Clinton Town	10	10	Accounting
101	Rose	Fraser Town	20	20	Quality
102	Mathew	Lakeside Village	10	30	Design
103	Stewart	Troy	30		
104	William	Holland	30		

Record Based Model

Difference between Hierarchical, Network & Relational Data Model

Hierarchical Data Model	Network Data Models	Relational Data Models
Supports One-Many Relationship	Supports both one to many and Many to Many relationship	Supports both one to many and Many to Many relationship
Because of single parent-child relationship, difficult to navigate through the child	It establishes the relationship between most of the objects, hence easy to access compared to hierarchical model	It provides SQL, which makes the access to the data simpler and quicker.
Flexibility among the different object is restricted to the child.	Because of the mapping among the sub level tables, flexibility is more	Primary and foreign key constraint makes the flexibility much simpler than other models.
Based on the physical storage details	Based on the physical storage details	Based on the logical data view

Normalization

- A process of organizing the data in database to avoid data redundancy, insertion anomaly, update anomaly & deletion anomaly.
- A process of organizing data into tables in such a way that the results of using the database are always unambiguous and as intended.
- Types of Normalization:
 1. First Normal Form (1NF)
 2. Second Normal Form (2NF)
 3. Third Normal Form (3NF)
 4. Boyce-Codd Normal Form (BCNF)

1NF

First normal form enforces these criteria:

- Eliminate repeating groups in individual tables.
- Create a separate table for each set of related data.
- Identify each set of related data with a primary key.

Example 1:

Course	Content
Programming	Java, c++
Web	HTML, PHP, ASP

1NF

We re-arrange the relation ***table*** as below, to convert it to First Normal Form.

Course	Content
Programming	Java
Programming	C++
Web	HTML
Web	PHP
Web	ASP

1NF

Example 2:

Patron ID	Borrowed books
C45	B33, B44, B55
C12	B56

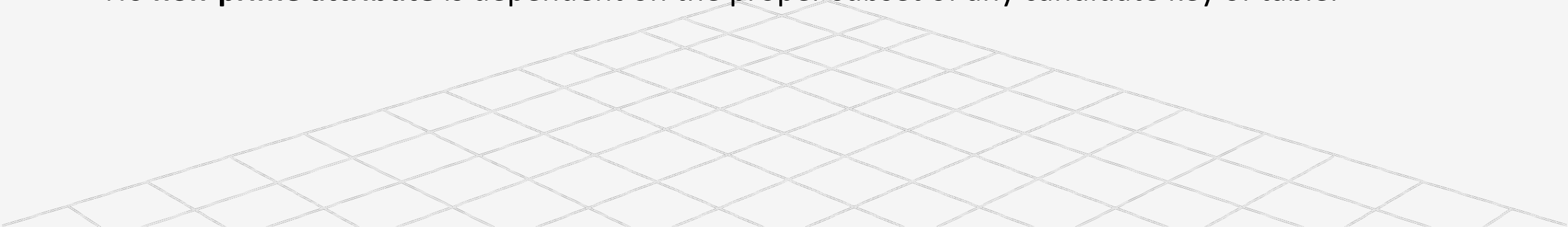
1NF

Patron ID	Borrowed book
C45	B33
C45	B44
C45	B33
C12	B56

2NF

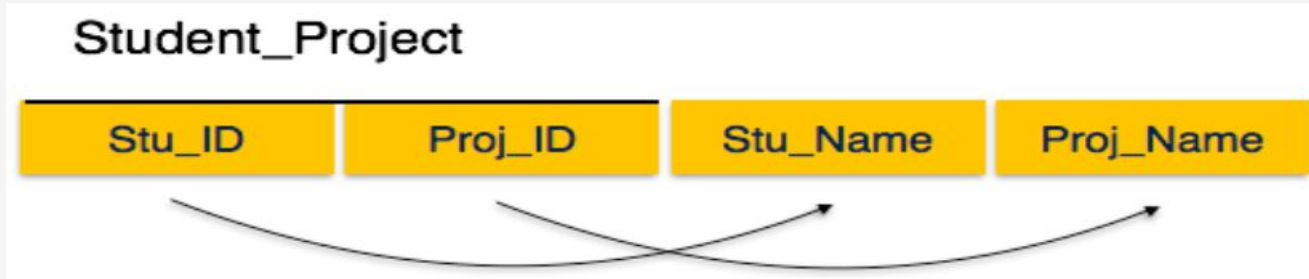
- **Prime attribute:** An attribute, which is a part of the prime-key, is known as a prime attribute.
- **Non-prime attribute:** An attribute, which is not a part of the prime-key, is said to be a non-prime attribute.

A table is said to be in 2NF if both the following conditions hold:

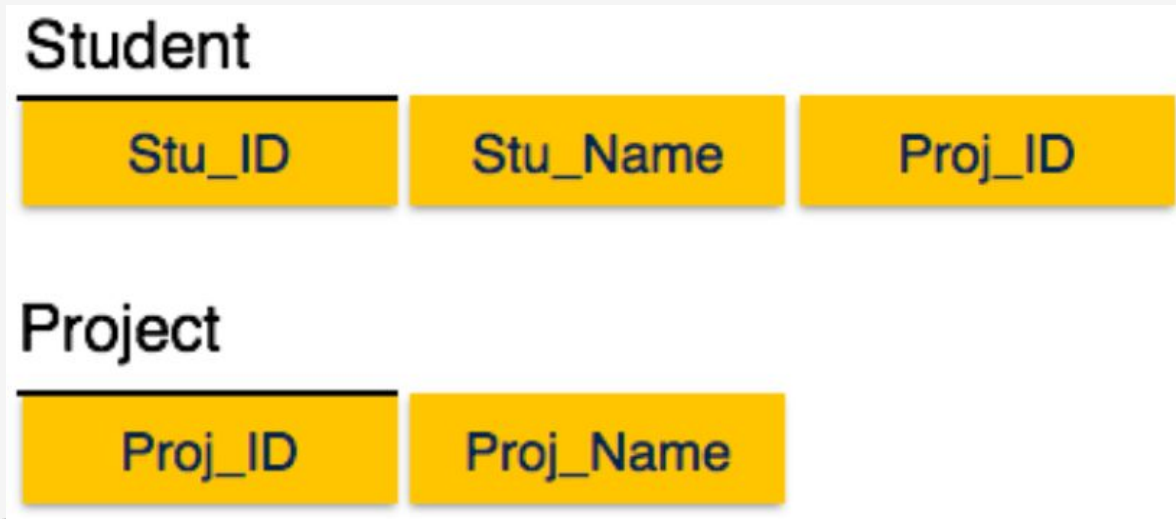
- Table is in 1NF (First normal form)
 - No **non-prime attribute** is dependent on the proper subset of any candidate key of table.
- 

2NF

Example 1:



2NF



2NF

Example 2:

BookNo	Patron	PhoneNo
B3	Shyam	555-1234
B2	Shyam	555-1234
B2	Ram	555-4321

2NF

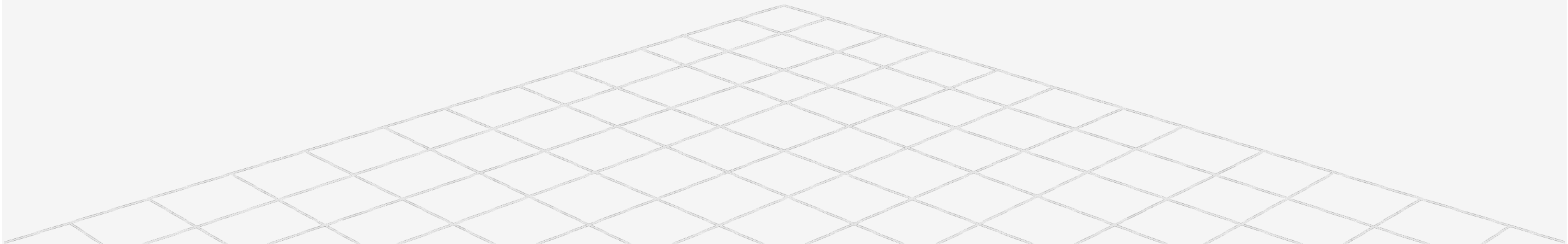
BookNo	Patron
B3	Shyam
B2	Shyam
B2	Ram

Patron	PhoneNo
Shyam	555-1234
Ram	555-4321

3NF

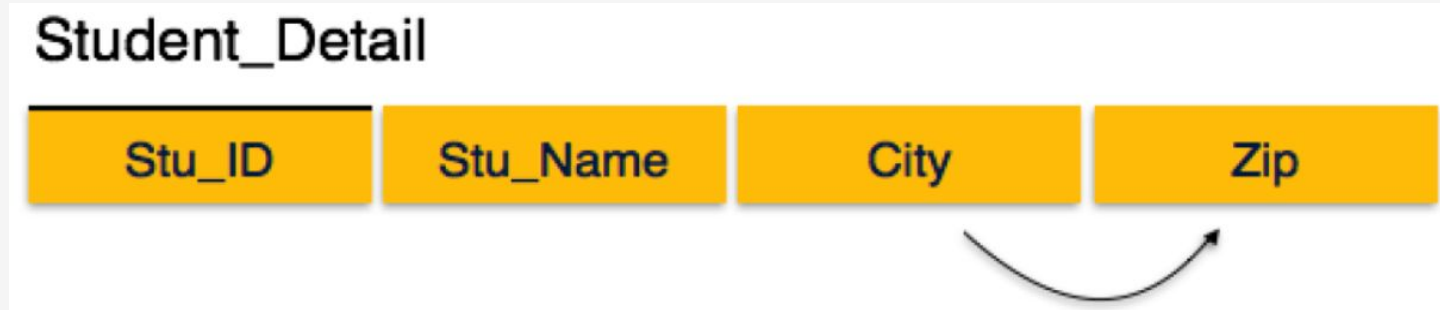
A table design is said to be in 3NF if both the following conditions hold:

- Table must be in 2NF
- No non-prime attribute is transitively dependent on prime key attribute
- For any non-trivial functional dependency, $X \rightarrow A$, then either –
 - X is a superkey or,
 - A is prime attribute



3NF

Example 1:



We find that in the above Student_detail relation, Stu_ID is the key and only prime key attribute. We find that City can be identified by Stu_ID as well as Zip itself. Neither Zip is a superkey nor is City a prime attribute. Additionally, $Stu_ID \rightarrow Zip \rightarrow City$, so there exists transitive dependency.

3NF

To bring this relation into third normal form, we break the relation into two relations as follows:

Student_Detail

Stu_ID

Stu_Name

Zip

ZipCodes

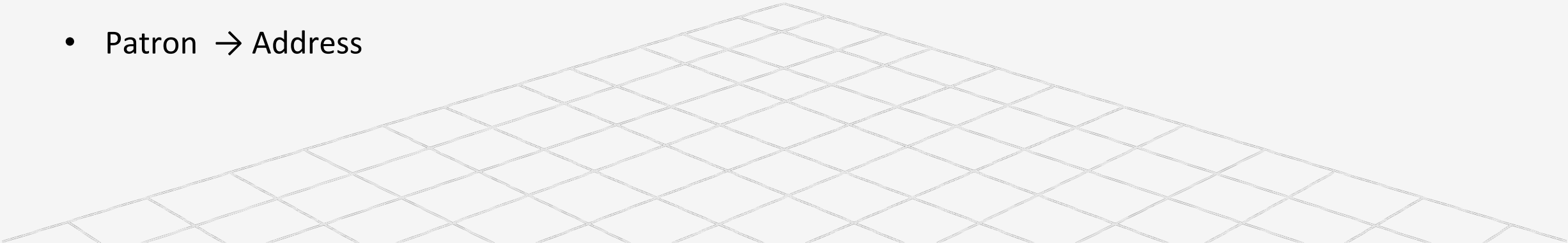
Zip

City

3NF

Example 2:

BookNo	Patron	Address	Due
B1	J. Fisher	101 Main Street	3/2/15
B2	L. Perez	202 Market Street	2/28/15

- Candidate key is BookNo
 - Patron → Address
- 

3NF

Put address in separate Patron table

BookNo	Patron	Due
B1	J. Fisher	3/2/15
B2	L. Perez	2/28/15

Patron	Address
J. Fisher	101 Main Street
L. Perez	202 Market Street

BCNF (Boyce-Codd Normal Form)

- It is an advance version of 3NF that's why it is also referred as 3.5NF. BCNF is stricter than 3NF.
- A table complies with BCNF if it is in 3NF and for every functional dependency $X \rightarrow Y$, X should be the super key of the table.

Example 1: In previous student example,

Stu_ID is the super-key in the relation Student_Detail and Zip is the super-key in the relation ZipCodes. So,

$\text{Stu_ID} \rightarrow \text{Stu_Name, Zip}$

and

$\text{Zip} \rightarrow \text{City}$

Which confirms that both the relations are in BCNF.

BCNF (Boyce-Codd Normal Form)

Example 2:

Student	Course	Teacher
Aman	DBMS	AYUSH
Aditya	DBMS	RAJ
Abhinav	E-COMM	RAHUL
Aman	E-COMM	RAHUL
abhinav	DBMS	RAJ

}

y

-> Teacher

not superkey but determines

BCNF (Boyce-Codd Normal Form)

Student	Course
Aman	DBMS
Aditya	DBMS
Abhinav	E-COMM
Aman	E-COMM
Abhinav	DBMS

Course	Teacher
DBMS	AYUSH
DBMS	RAJ
E-COMM	RAHUL

Types of Databases

Modern Databases : Modern DBMS permits storage and retrieval of data types such as

- business graphs,
- Spreadsheets,
- Art clippings(clip arts),
- Picture, sound, video, etc.

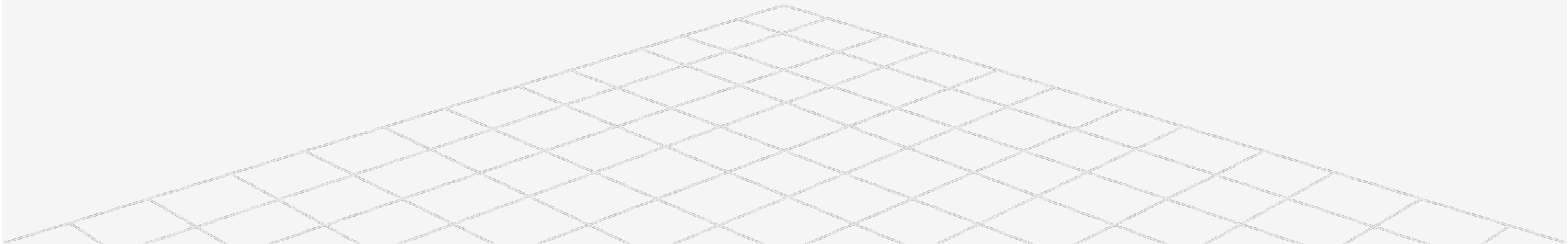
New data types are termed as 'objects'.

data planning relates to the way the data will be managed.



Types of Databases

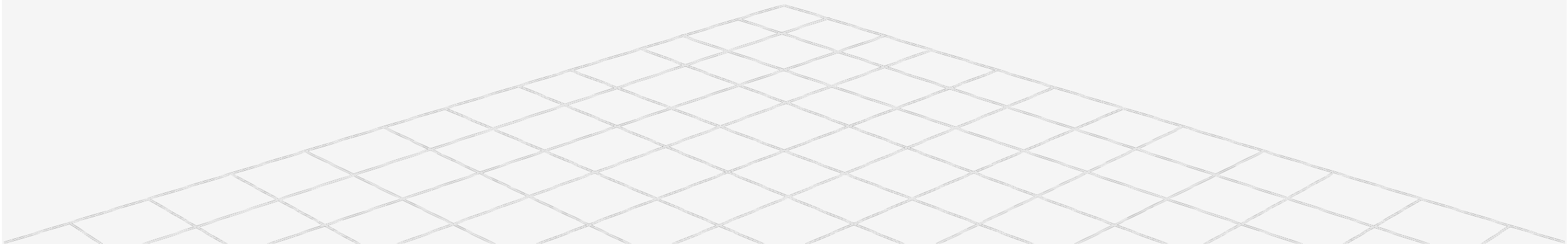
- Operational databases
- End user databases
- Centralized databases
- Distributed databases
- Personal databases
- Commercial databases



Types of Databases

Operational Database:

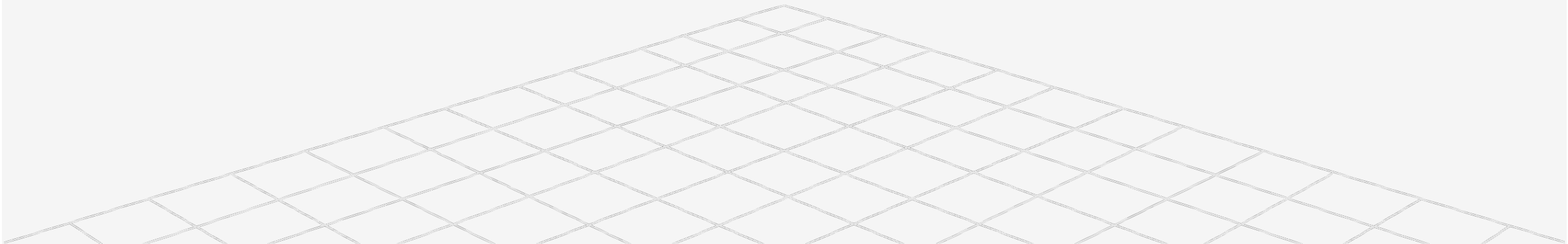
- These databases store data relating to the operations of the enterprise.
- Databases are organized on functional lines such as
 - marketing,
 - production,
 - Employees, etc.



Types of Databases

End-User Database:

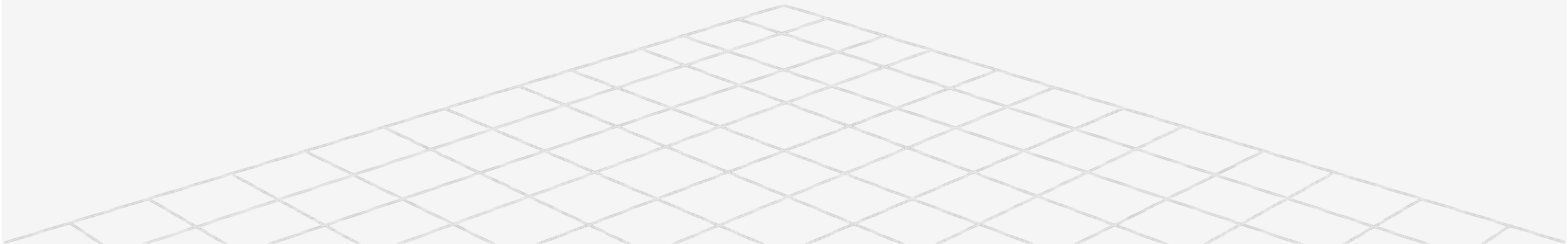
- databases are shared by users
- contain information meant for use by the end-users
 - like managers at different levels
 - managers may not be concerned about the individual transactions as found in operational databases.(summary)



Types of Databases

Centralized Database:

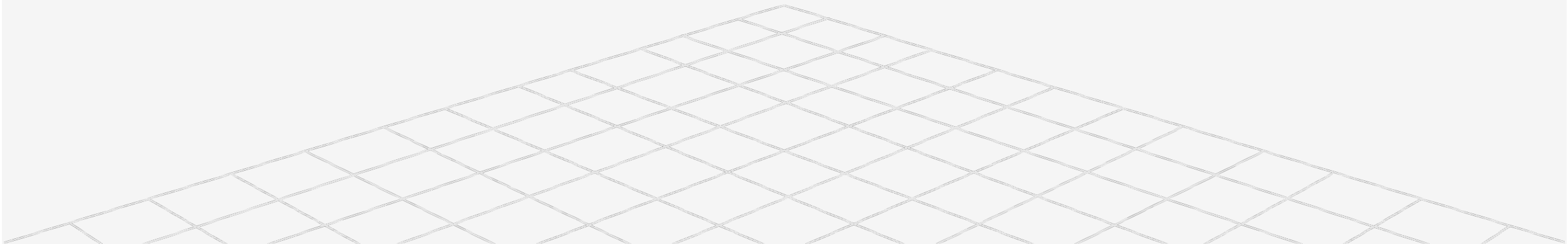
- databases store the entire information and application programs at a central computing facility
- Users at different location access database
- Application programs perform data validation and verification and assign unique ID to each transaction at central facility.
- Example :MTNL has centralized database for registration of new numbers.



Types of Databases

Distributed Database:

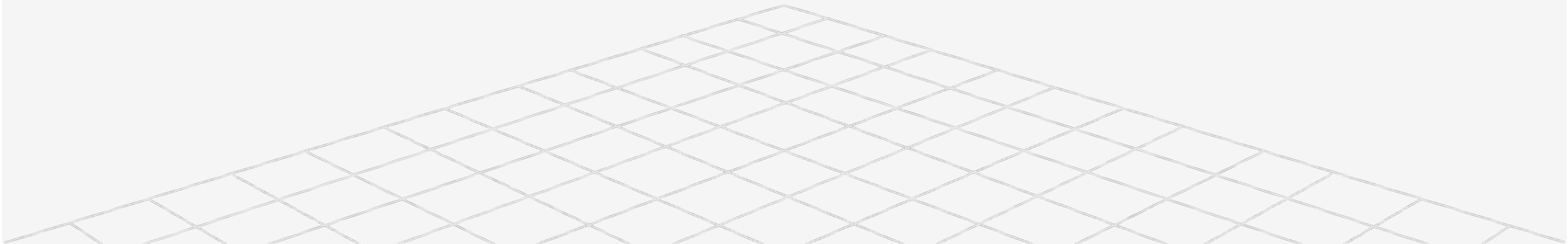
- The data is distributed at various sites
- sites are linked via communication links, the entire collection of data constitutes the logical database of organization.
- ensure that the detailed local information remains stored on the local facility
- The client server technology is most popular for managing distributed databases.



Types of Databases

Personal Database:

- Maintained on personal computer
- Contain information meant for limited number of users (in same department)
- Databases subject specific and user designed
- Simple and less powerful DBMS package used.



Types of Databases

Commercial Database:

- Access is provided to users as a commercial venture is called a commercial or external database.
- databases are subjected specific and access to these databases is sold as a paid service to its user.
- The access to commercial databases may be given through communication links.
- Databases contain information that external users would require but by themselves would not be able to afford maintaining such huge databases.
- databases may offer statistics regarding commodity, foreign exchange and stock markets, companies and their performance, importers and their buying patterns, decided case laws, etc

Thank You...

