First public global database of fossil fuels launches

On Latshaw drilling rig #43 in the Permian Basin in Odessa, Texas, on Oct. 13, 2021. (David Goldman / AP)

Vivian Tang

Fossil Fuel Database Released Today

Came across this article about a database tracking the world's fossil fuels becoming public today - in the past, this information was private and had to be paid for 😮

Hopefully this release of data will open up new discussions surrounding climate change!

Last Week × Q4Me

-3

Book vs. Slides

W01: CH01

W02: CH04 (2nd Ed.), CH02 (1st Ed.)

Lab Assignments

?

Last Week ×Q4U

Data Modeling (Data Odyssey)?

Data Modeling in Memory: Pros & Cons?

Data Modeling in File: Pros & Cons?

Object and Record are terms of what levels of data modeling?

Physical Level × File -1

Pros	Cons
Space	*Time* *Hard Disk Drive (HDD), Sequential Access, Electromechanical* *Solid State Drive (SSD), Random Access, Expensive*
DURABLE Persistency | Long-term Retention	*Not Easy DML* DELETE | UPDATE
Portable (self-explanatory)	✗ *Not Portable*
	✗ *No ACID Properties*

Last Week | DBMS | Entity | Attribute | Relationship | Extended ER (EER)

Physical Level × File

0

Pros	Cons
Space	*Time* *Hard Disk Drive (HDD), Sequential Access, Electromechanical* *Solid State Drive (SSD), Random Access, Expensive*
DURABLE Persistency | Long-term Retention	*Not Easy DML* DELETE | UPDATE
Portable (self-explanatory)	*Not Portable*
	No ACID Properties

Physical Level × File × ACID

Transaction
A group one or more operations (DDL|DML) into a single unit of work.

BEGIN TRANSACTION
    INSERT *Movie*
    INSERT *Movie's Director* IF NOT EXIST
    INSERT *Movie's distributor Company* IF NOT EXIST
END TRANSACTION

Physical Level × File × ACID

Atomicity /a·tuh·mi·suh·tee/

All-or-nothing execution of transaction

BEGIN TRANSACTION
INSERT *Movie*
INSERT *Movie's Director* IF NOT EXIST
INSERT *Movie's distributor Company* IF NOT EXIST
END TRANSACTION
Transaction Committed.
1 or 2 or 3 records affected.

Physical Level × File × ACID

Atomicity/a·tuh·mi·suh·tee/

All-or-nothing execution of transaction

BEGIN TRANSACTION

INSERT *Movie*

Error!

Partial Execution of Transaction!

ROLLBACK any changes, i.e., Movie record!

Physical Level × File × ACID

Atomicity/a·tuh·mi·suh·tee/

All-or-nothing execution of transaction

```
hfani@charlie:~/comp3150_f2022/Moviesion02/ca.uwindsor.cs.comp3150.pl$ ./Moviesion
1) Add Movie
2) Print Movie by Id
3) *Print Movie by Title
4) *Print All Movies
5) Print All Movies of a Director
6) *Edit Movie by Id
7) *Delete Movie by Id
0) Quit
Enter command:
```

Physical Level × File × ACID

Atomicity/a·tuh·mi·suh·tee/

All-or-nothing execution of transaction

Another example in online banking?

Physical Level × File × ACID

Atomicity/a·tuh·mi·suh·tee/

All-or-nothing execution of transaction

BEGIN TRANSACTION

Withdraw Money from Hossein's Saving Account

Deposit Money to Hossein's Checking Account

END TRANSACTION

Physical Level × File × ACID

Consistency

Respect constraints or expectations among data instances

CONSTRAINT#1: All Movies Must Have ReleaseDate ≥ 1890
CONSTRAINT#2: All Movies Must Have at Least One Director
CONSTRAINT#3: All Movies Must Have UNIQUE Title

Physical Level × File × ACID

Consistency

Respect constraints or expectations among data instances

BEGIN TRANSACTION
    INSERT *Movie*
    INSERT *Movie's distributor Company* IF NOT EXIST
END TRANSACTION

Transaction Conflicts with CONSTRAINT#2.
ROLLBACK any changes.

Physical Level × File × ACID

Consistency

Respect constraints or expectations among data instances

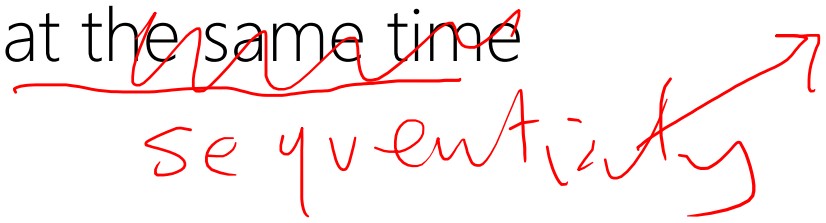
BEGIN TRANSACTION
    INSERT *Movie*
    Error: INSERT *Movie's Director* IF NOT EXIST
    INSERT *Movie's distributor Company* IF NOT EXIST
END TRANSACTION
Transaction Committed.
2 records affected.
Partial Execution of Transaction → Transaction Conflicts with CONSTRAINT#2

Inconsistency was due to lack of atomicity.
ACID properties are not orthogonal (they are dependent).

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time



Multi-user environment!

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

```
hfani@charlie:~/comp3150_f2022/Moviesion02/ca.uwindsor.cs.comp3150.pl$ ./Moviesion
1) Add Movie
2) Print Movie by Id
3) *Print Movie by Title
4) *Print All Movies
5) Print All Movies of a Director
6) *Edit Movie by Id
7) *Delete Movie by Id
0) Quit
Enter command:
```

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

Transaction Queue (Sequential Execution, One-at-a-time)		
T1	W1	T1.1. INSERT *2001: A Space Odyssey*
	W2	T1.2. INSERT *Stanley Kubrick* IF NOT EXIST
	W3	T1.3. INSERT *MGM* IF NOT EXIST
T2	W4	T2.1. INSERT *A Clockwork Orange*
	W5	T2.2. INSERT *Stanley Kubrick* IF NOT EXIST
	W6	T2.3. INSERT *Warner Bros* IF NOT EXIST
T3	W7	T3.1. UPDATE Director SET PlaceOfBirth = "United States of America" WHERE PlaceOfBirth = "USA"
T4	W8	...

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

Transaction Queue (Parallel Execution)		
W1	T1.1. INSERT *2001: A Space Odyssey*	T2.1. INSERT *A Clockwork Orange*
W2	T1.2. INSERT *Stanley Kubrick* IF NOT EXIST	T2.2. INSERT *Stanley Kubrick* IF NOT EXIST
W3	T1.3. INSERT *MGM* IF NOT EXIST	T2.3. INSERT *Warner Bros* IF NOT EXIST
W4	T3.1. UPDATE Director SET PlaceOfBirth = "United States of America" WHERE PlaceOfBirth = "USA"	

8 Works → 4 Works

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

Transaction Queue (Parallel Execution)		
W1	T1.1. INSERT *2001: A Space Odyssey*	T2.1. INSERT *A Clockwork Orange*
W2	T1.2. INSERT *Stanley Kubrick* IF NOT EXIST	T2.2. INSERT *Stanley Kubrick* IF NOT EXIST
W3	T1.3. INSERT *MGM* IF NOT EXIST	T2.3. INSERT *Warner Bros* IF NOT EXIST
W4	T3.1. UPDATE Director SET PlaceOfBirth = "United States of America" WHERE PlaceOfBirth = "USA"	

What's the Result?

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

Transaction Queue (Parallel Execution)		
W1	T1.1. INSERT *2001: A Space Odyssey*	T2.1. INSERT *A Clockwork Orange*
W2	T1.2. INSERT *Stanley Kubrick* IF NOT EXIST	T2.2. INSERT *Stanley Kubrick* IF NOT EXIST
W3	T1.3. INSERT *MGM* IF NOT EXIST	T2.3. INSERT *Warner Bros* IF NOT EXIST
W4	T3.1. UPDATE Director SET PlaceOfBirth = "United States of America" WHERE PlaceOfBirth = "USA"	

What's the Result? Duplicate *Stanley Kubrick*!

Physical Level × File × ACID

Isolation

Transaction appear to be executed as if no other transaction is executing at the same time

Transaction Queue (Parallel Execution)			
W1	T1.1. INSERT *2001: A Space Odyssey*	T2.1. INSERT *A Clockwork Orange*	
W2	T1.2. INSERT *Stanley Kubrick* IF NOT EXIST	T1.3. INSERT *MGM* IF NOT EXIST	T2.3. INSERT *Warner Bros* IF NOT EXIST
W3	T2.2. INSERT *Stanley Kubrick* IF NOT EXIST		
W4	T3.1. UPDATE Director SET PlaceOfBirth = "United States of America" WHERE PlaceOfBirth = "USA"		

Physical Level × File × ACID

Durability

Once a transaction has committed, the effect must never be lost

Function Exception
Program Crash
System Crash: files might not be durable at this level.


Apocalypse Now
MARLON BRANDO ROBERT DUVALL MARTIN SHEEN in APOCALYPSE NOW
FREDERIC FORREST ALBERT HALL SAM BOTTOMS LARRY FISHBURNE and DENNIS HOPPER Produced and Directed by FRANCIS COPPOLA
Written by JOHN MILIUS and FRANCIS COPPOLA Narration by MICHAEL HERR Co-Produced by FRED ROOS, GRAY FREDERICKSON and TOM STERNBERG Director of Photography VITTORIO STORARO
Production Designer DEAN TAVOULARIS Editor RICHARD MARKS Sound Design by WALTER MURCH Music by CARMINE COPPOLA and FRANCIS COPPOLA
United Artists AN OMNI ZOETROPE PRODUCTION


Starchild: Database Management System

Database Management Systems (DBMS) [25]

A Software System
Provides General Purpose, Efficient, Convenient, & Safe Multi-user Storage of and Access to Massive Amounts of Persistent Data

ACID Properties Guaranteed!

DBMS

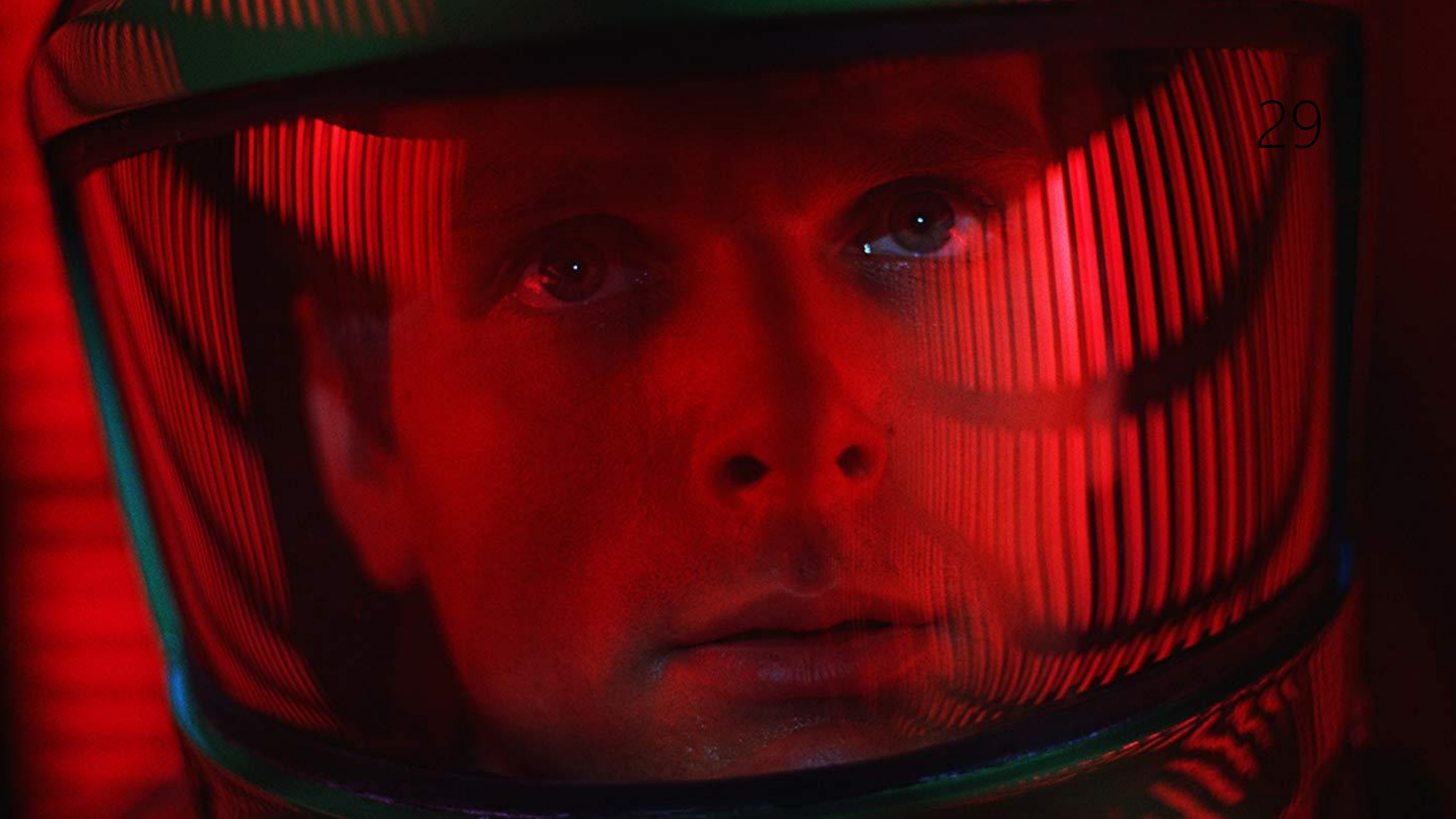
General Purpose	Different Applications: Moviesion, Library,
Safe	From Malicious Users
Convenient	Simple DDL & DML
Efficient	Don’t Search All Files in Order to Get a Record
ACID	Atomicity, Consistency, Isolation, Durability

DBMS

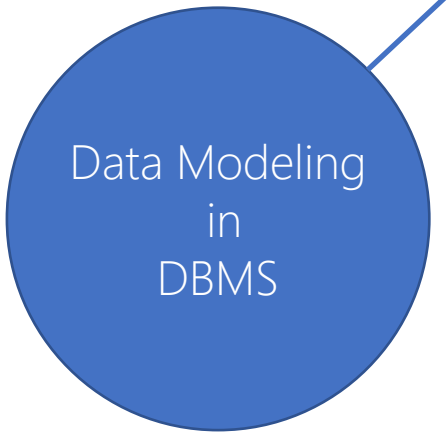
	People (Role)
Is this course about building a DBMS?	No! DBMS Designers
Is this course about maintaining a DBMS?	No! DB Administers (DBA)
Is this course about using a DBMS?	Yes. DB Designers

DBMS Products

Company	DBMS
Oracle	Oracle, MySQL
Microsoft	MySQLServer
IBM	DB2
OpenLink Software	Virtuoso Universal Server
Apache Software Foundation	CouchDB
This Course	SQLite



2019: A Data Odyssey × Real World



Real World Entity

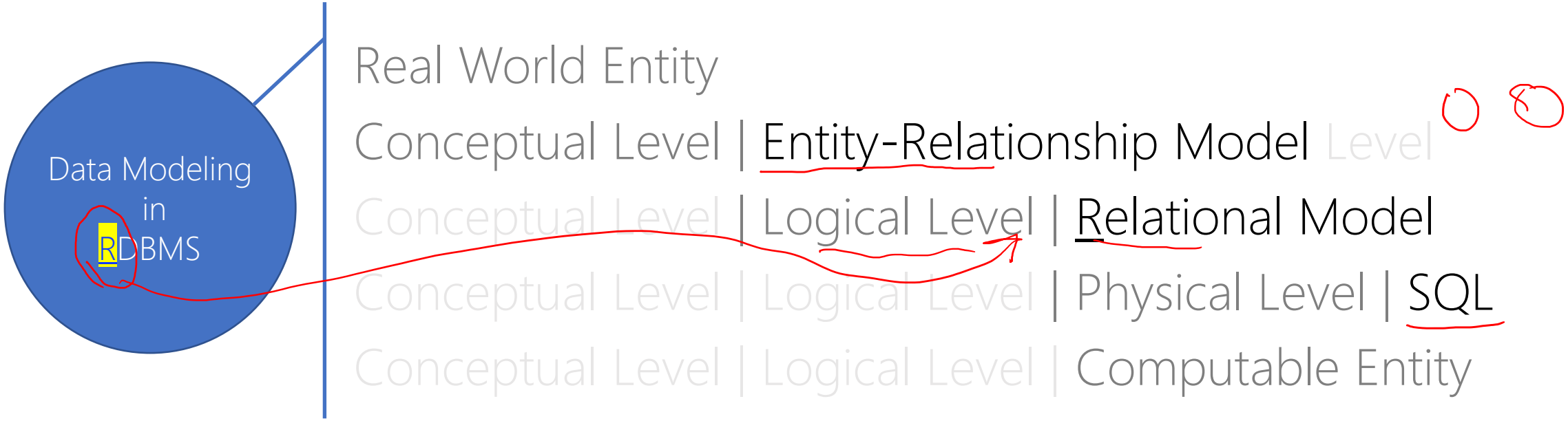
Conceptual Level | Logical Level | Physical Level

Conceptual Level | Logical Level | Physical Level

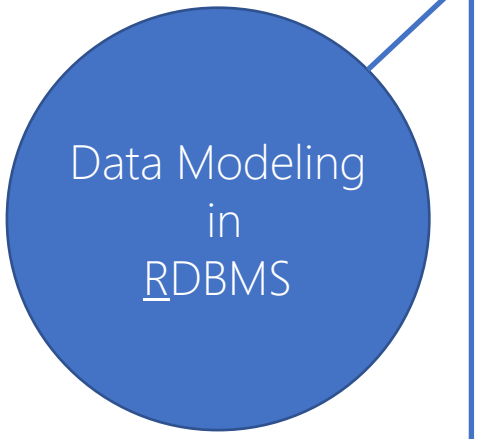
Conceptual Level | Logical Level | Physical Level

Conceptual Level | Logical Level | Computable Entity

2019: A Data Odyssey × Real World



Today



Real World Entity

Conceptual Level | Entity-Relationship Model (E/R) Level

Conceptual Level | Logical Level | Relational Model

Conceptual Level | Logical Level | Physical Level | SQL

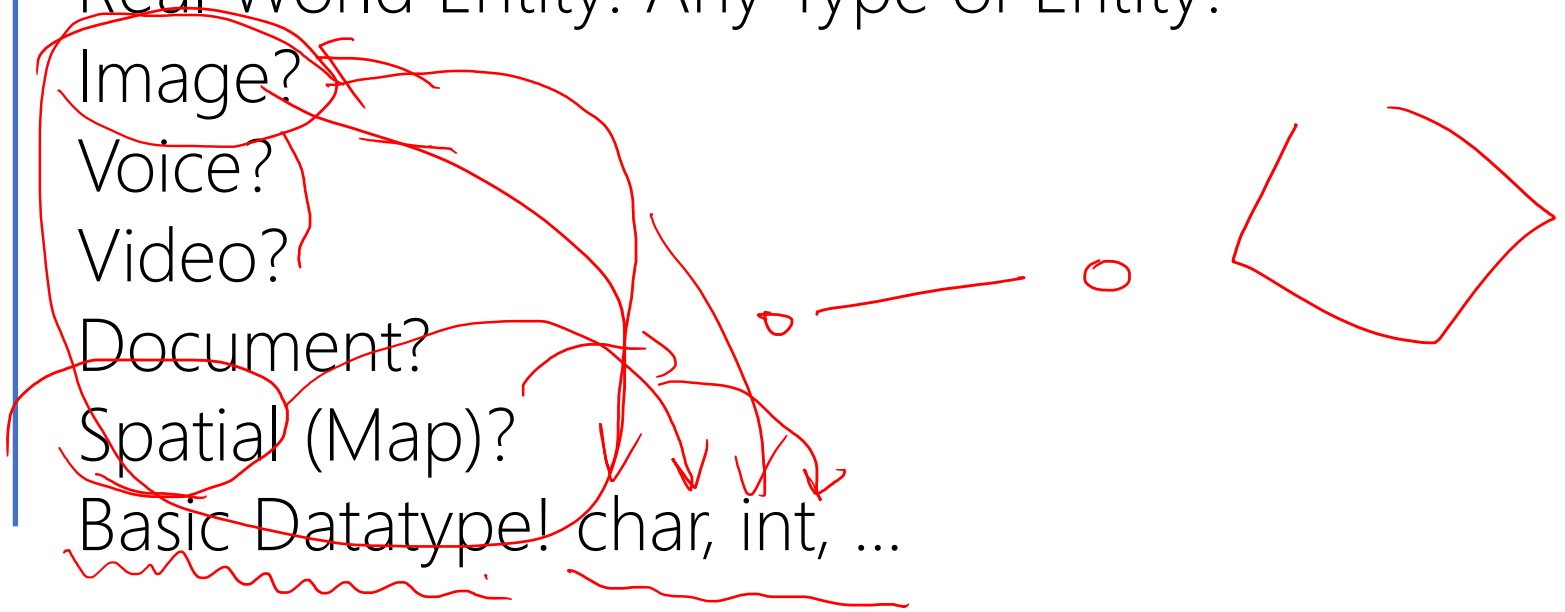
Conceptual Level | Logical Level | Computable Entity

Last Week | DBMS | Entity | Attribute | Relationship | Extended ER (EER)

Oops!

Data Modeling in RDBMS

Real World Entity: Any Type of Entity? No!
Image? No!
Voice? No!
Video? No!
Document? No!
Spatial (Map)? No!
Basic Datatype! char, int, … Yes.

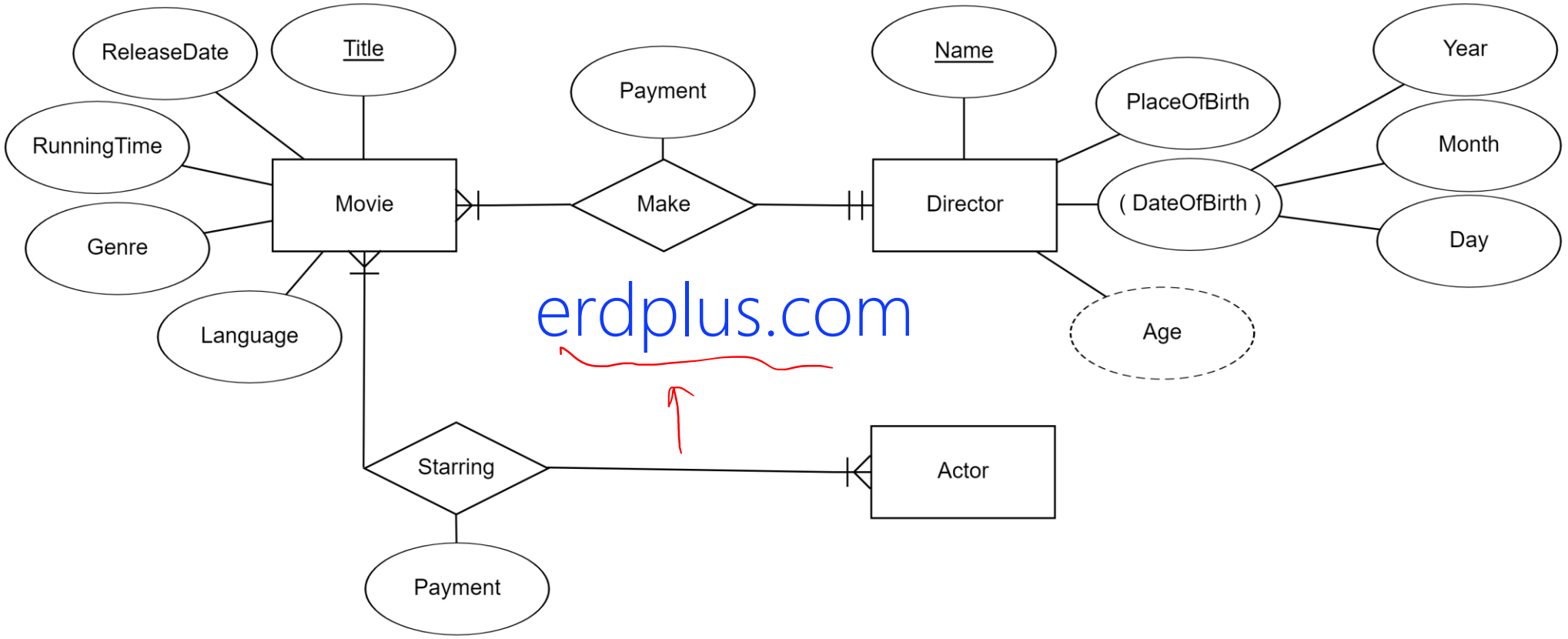


Entity-Relationship Model (E/R)

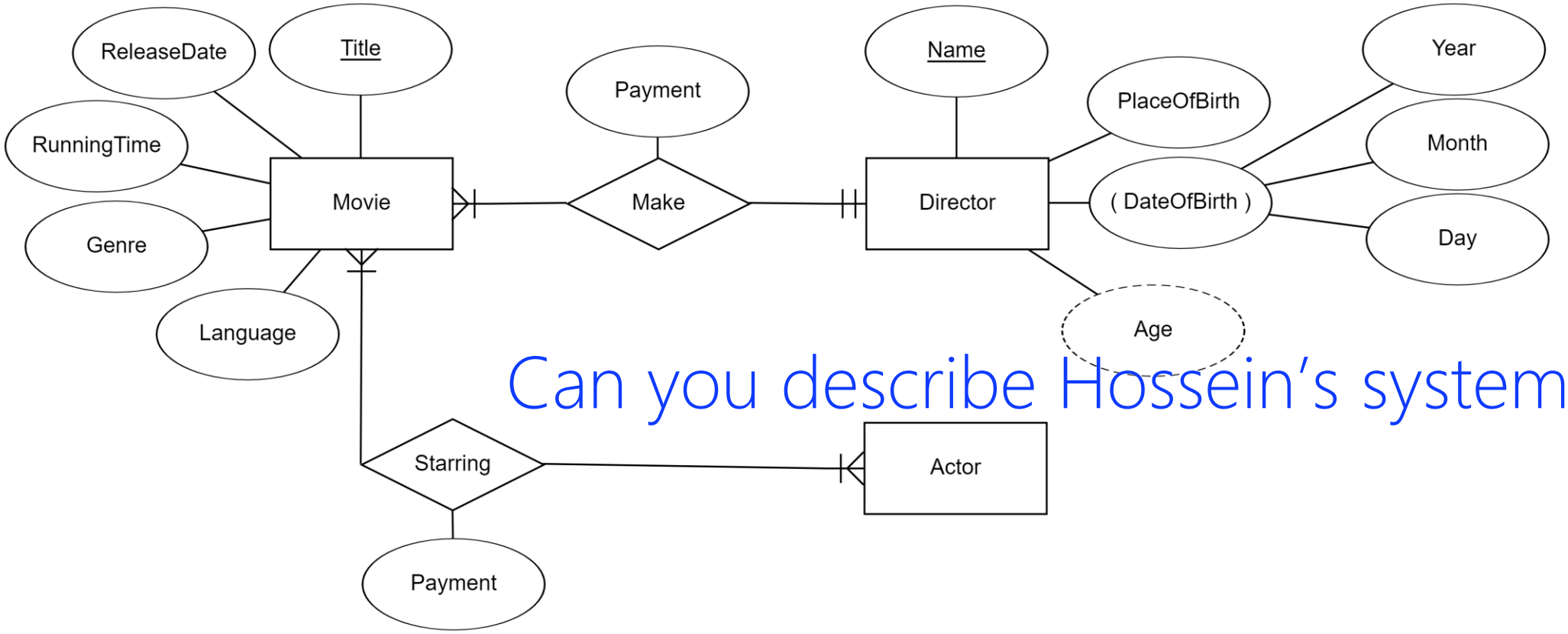
1. Identify Real World Entities, Attributes, Relationships
2. Create Graphical Schema: ER Diagram (ERD)

A picture is worth a thousand words!
Standard: All data engineers all over the world understand ERD!

E/R × ERD


ReleaseDate
Title
RunningTime
Genre
Language
Movie
Payment
Make
Name
PlaceOfBirth
(DateOfBirth)
Year
Month
Day
Director
Age
erdplus.com
Starring
Actor
Payment

E/R × ERD



Can you describe Hossein's system?

E/R × Entity (e_i)

Real world thing that exists & is distinguishable from other things

e_1=The Birds, e_2=Rosemary's Baby, e_3=L.A. Confidential
e_4=Alfred Hitchcock, e_5=Roman Polanski, e_6=James Cameron
e_7=Tippi Hedren, e_8=Mia Farrow, e_9=John Cassavetes
e_{10}=Soroush Ziaenejad, e_{11}=Hossein Fani, …
e_{12}=Bruno Mars, e_{13}=Taylor Swift, …
e_{14}=19Q4, e_{15}=Brothers Karamazov, …

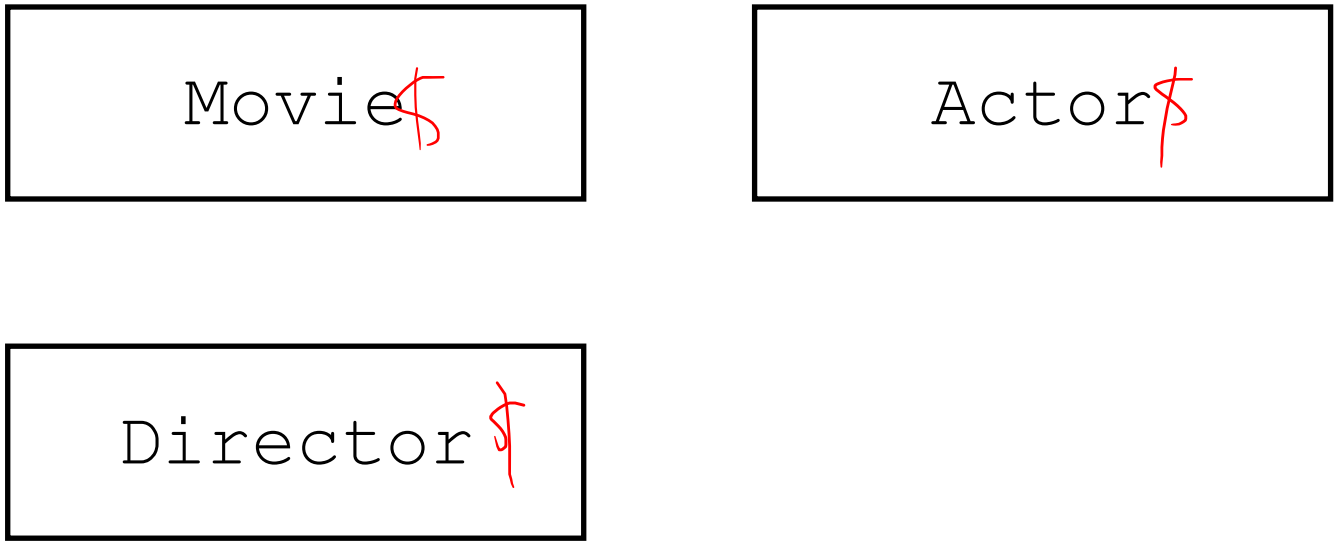
E/R × Entity Set (E)

A set of entities of the same type that share the same properties

Movie={e_1=The Birds, e_2=Rosemary's Baby, e_3=L.A. Confidential, …}
Director={e_4=Alfred Hitchcock, e_5=Roman Polanski, e_6=James Cameron, …}
Actor ={e_7=Tippi Hedren, e_8=Mia Farrow, e_9=John Cassavetes, …}
Instructor={e_{10}=Soroush Ziaenejad, e_{11}=Hossein Fani, …}
Singer={e_{12}=Bruno Mars, e_{13}=Taylor Swift, …}
Novel={e_{14}=19Q4, e_{15}=Brothers Karamazov, …}

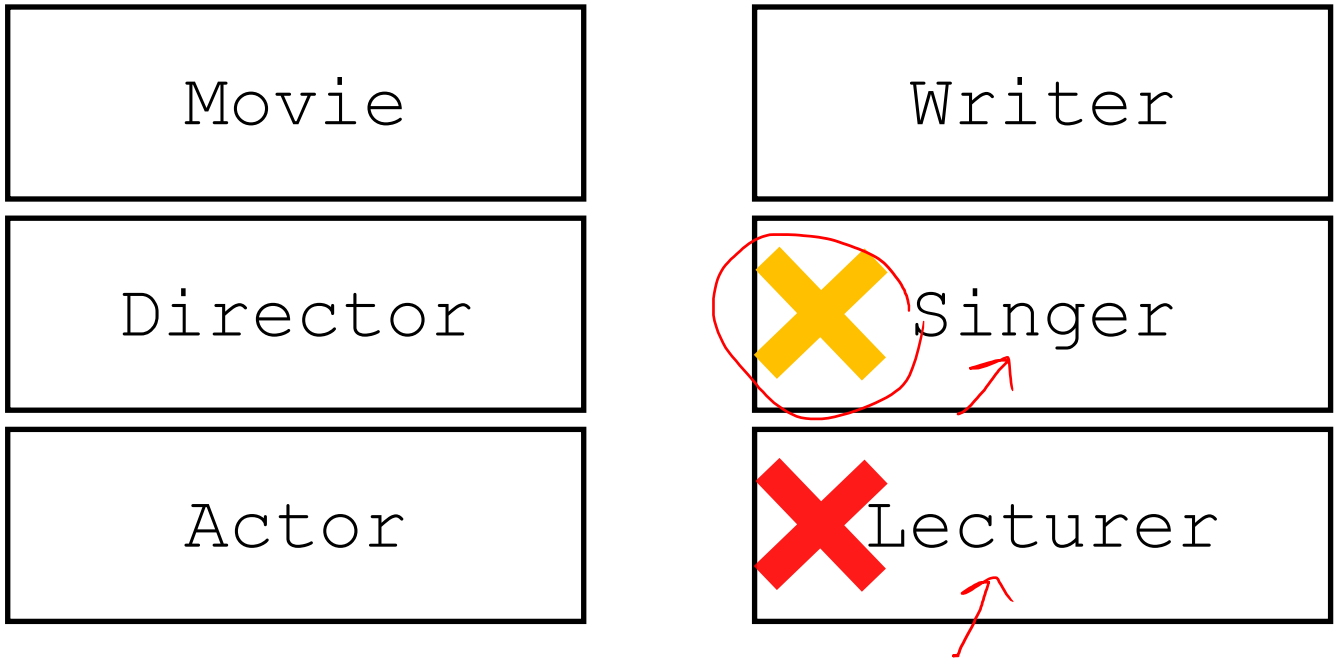
E/R × Entity Set (E)

E/R × Entity Set



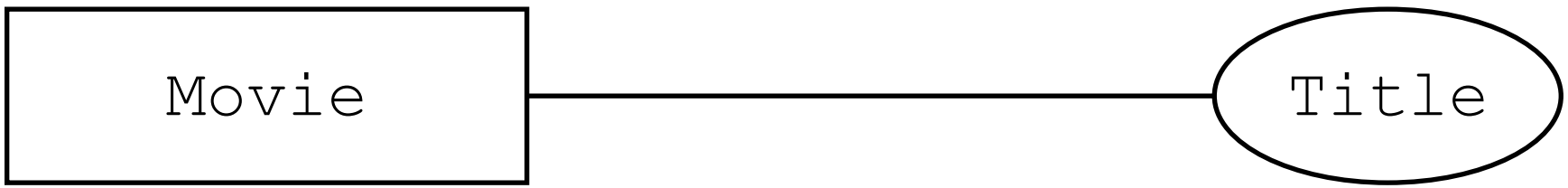
E/R × Entity Set × Faithfulness

Design should be faithful to the specifications of the application



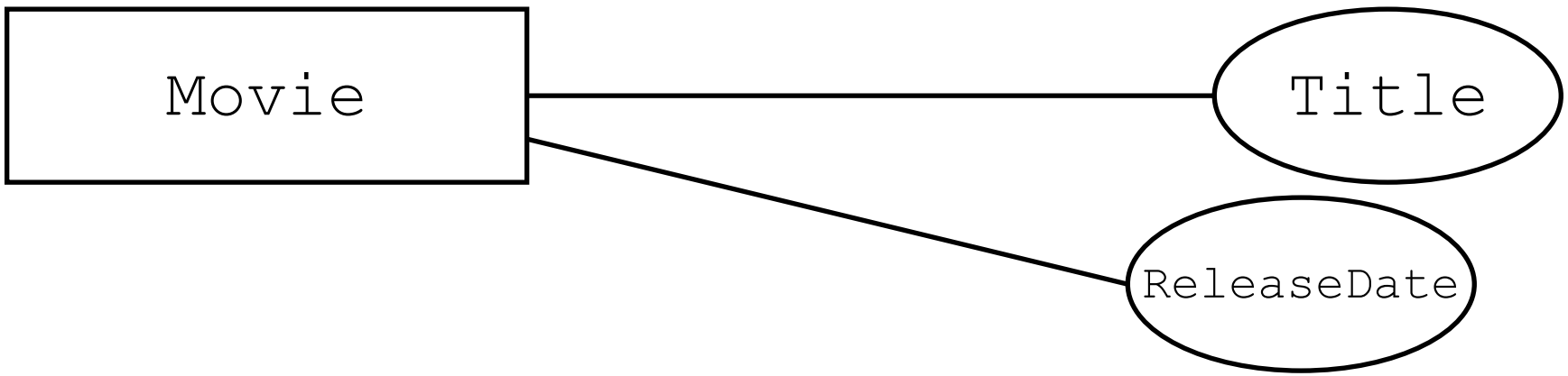
E/R × Attribute

Properties of entities in entity set



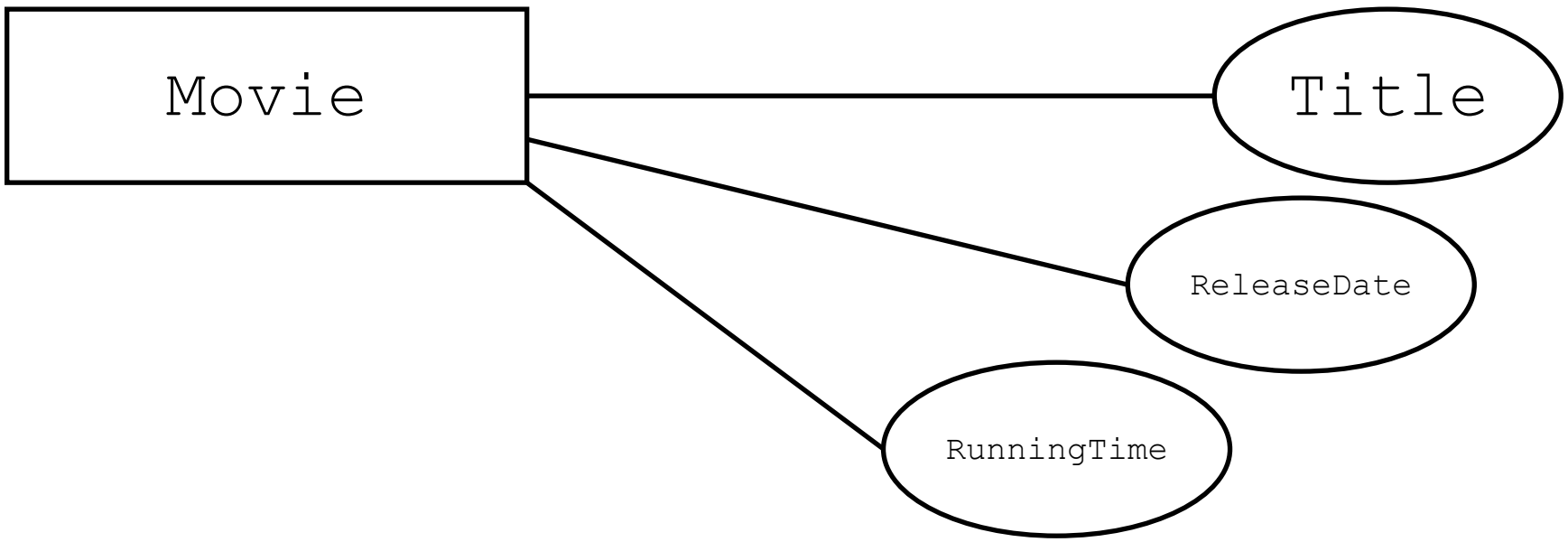
E/R × Attribute

Properties of entities in entity set



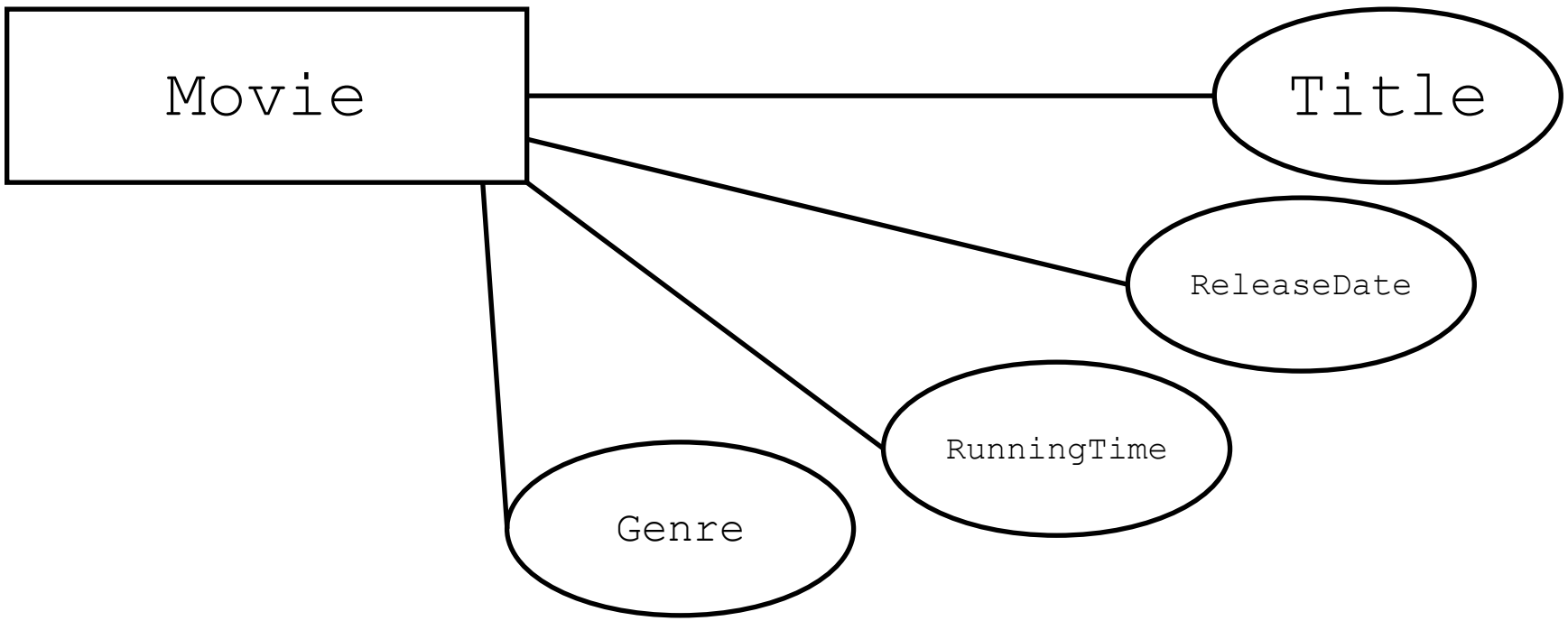
E/R × Attribute

Properties of entities in entity set



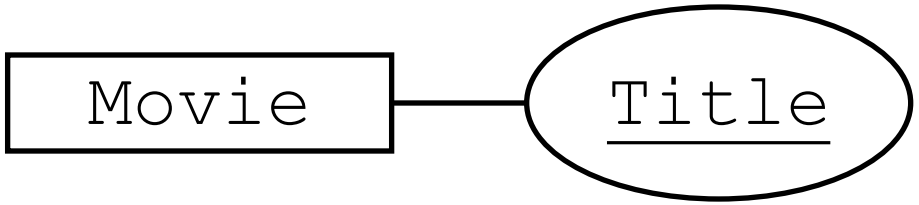
E/R × Attribute

Properties of entities in entity set



E/R × Attribute × Key

An attribute or a set of attributes uniquely identify an entity in entity set



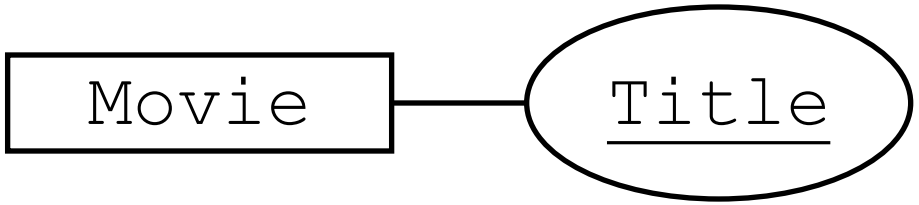
A movie is identified by its `Title`

Two movies cannot have same value for Title

The value in Title identifies one and only one movie!

E/R × Attribute × Key

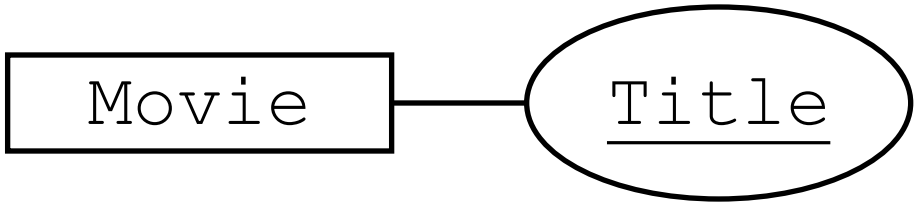
An attribute or a set of attributes uniquely identify an entity in entity set



`Title` by itself as a key?

E/R × Attribute × Key

An attribute or a set of attributes uniquely identify an entity in entity set

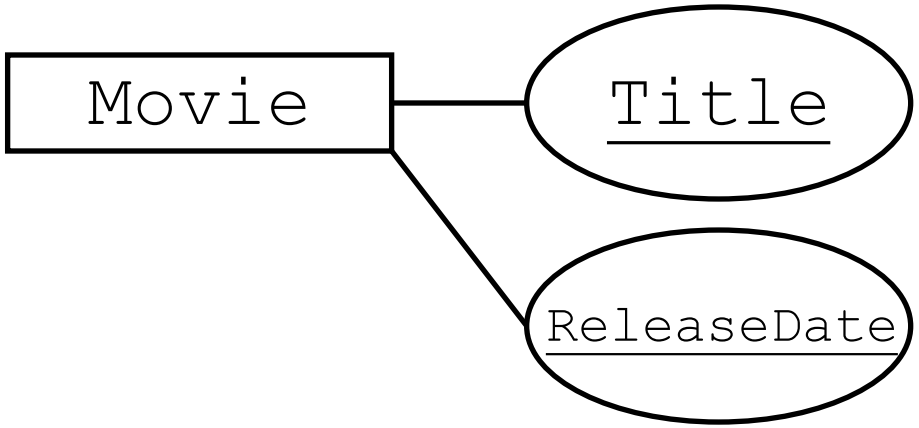


`Title` by itself as a key?

1. Rosemary's Baby (1968) by Roman Polanski
2. Rosemary's Baby (2014) (TV Mini-Series) by Agnieszka Holland
3. Rosemary's Baby (2015) by Dane Kissel

E/R × Attribute × Key

An attribute or a set of attributes uniquely identify an entity in entity set



{Title, ReleaseDate} as the key?

E/R × Attribute × Key

An attribute or a set of attributes uniquely identify an entity in entity set

How about {`Title, ReleaseDate, RunningTime, Genre`} as the key?

At worst case, all attributes together are the key!

No duplicate entity in the entity set!

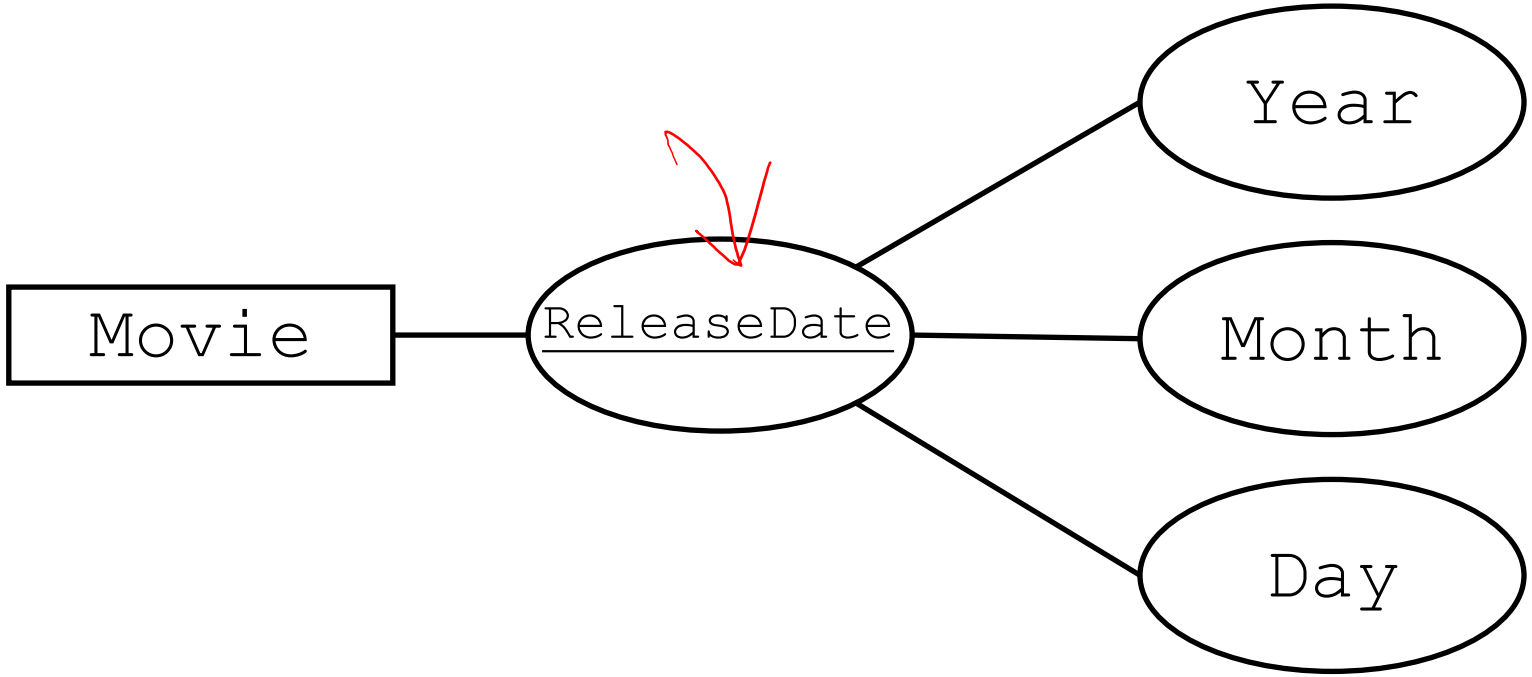
E/R × Attribute × Key

An attribute or a set of attributes uniquely identify an entity in entity set

Keys MUST have value.

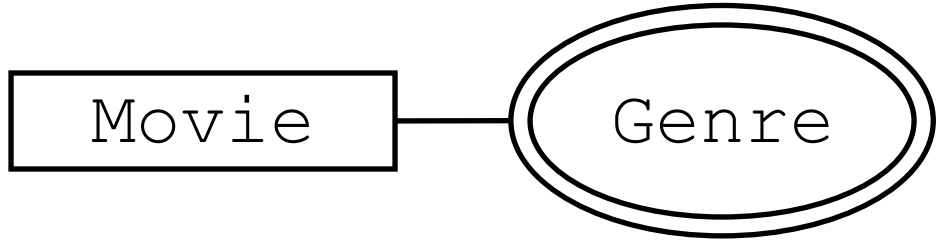
E/R × Attribute × Composite

Properties of entities in entity set



E/R × Attribute × Multivalued

Properties of entities in entity set

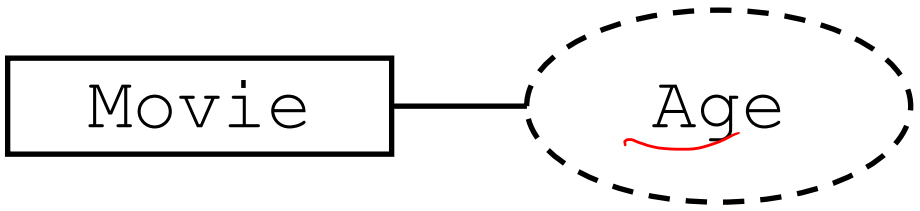


e_1.Title=`*The Birds*'
e_1.ReleaseDate=*March 28, 1963*
e_1.RunningTime=*119*
e_1.Genre={`*Drama*', `*Horror*', `*Mystery*'}

e_2.Title=`*Rosemary's Baby*'
e_2.ReleaseDate=*June 12, 1968*
e_2.RunningTime=*136*
e_2.Genre={`*Drama*', `*Horror*'}

E/R × Attribute × Derived

Properties of entities in entity set



e_1.Title=`The Birds'

e_1.ReleaseDate=March 28, 1963

e_1.RunningTime=119

e_1.Genre={`Drama', `Horror', `Mystery '}

e_1.Age=YEAR(e_1.ReleaseDate-NOW)

e_2.Title=`Rosemary's Baby'

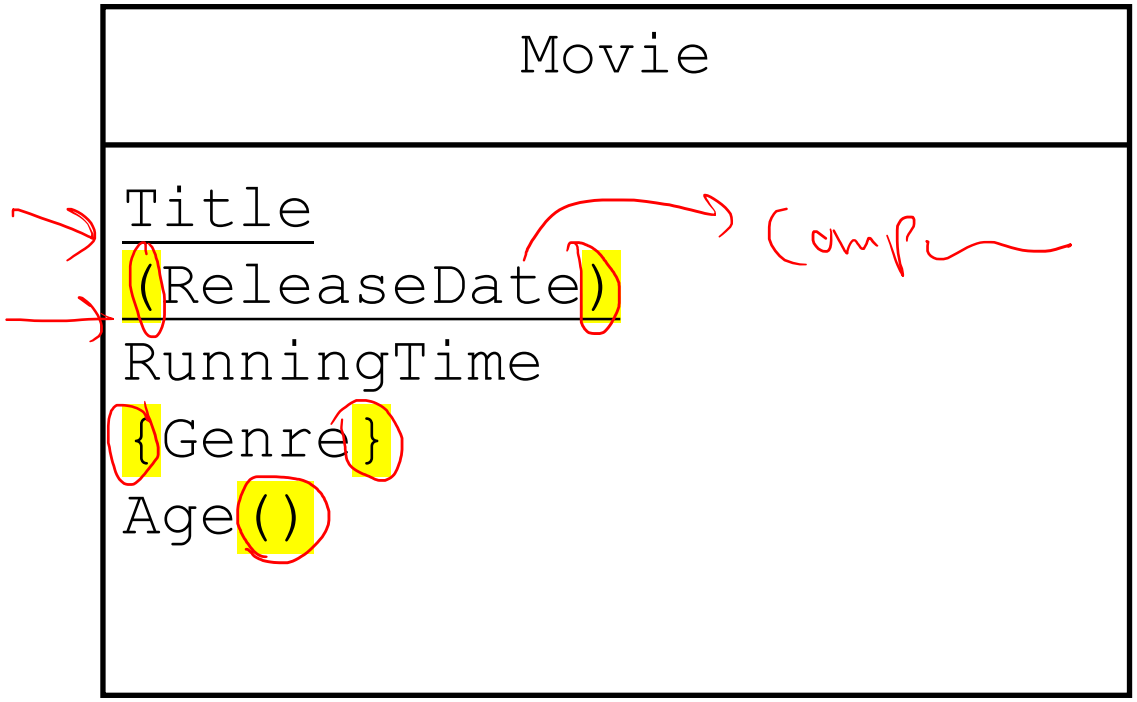
e_2.ReleaseDate=June 12, 1968

e_2.RunningTime=136

e_2.Genre={`Drama', `Horror'}

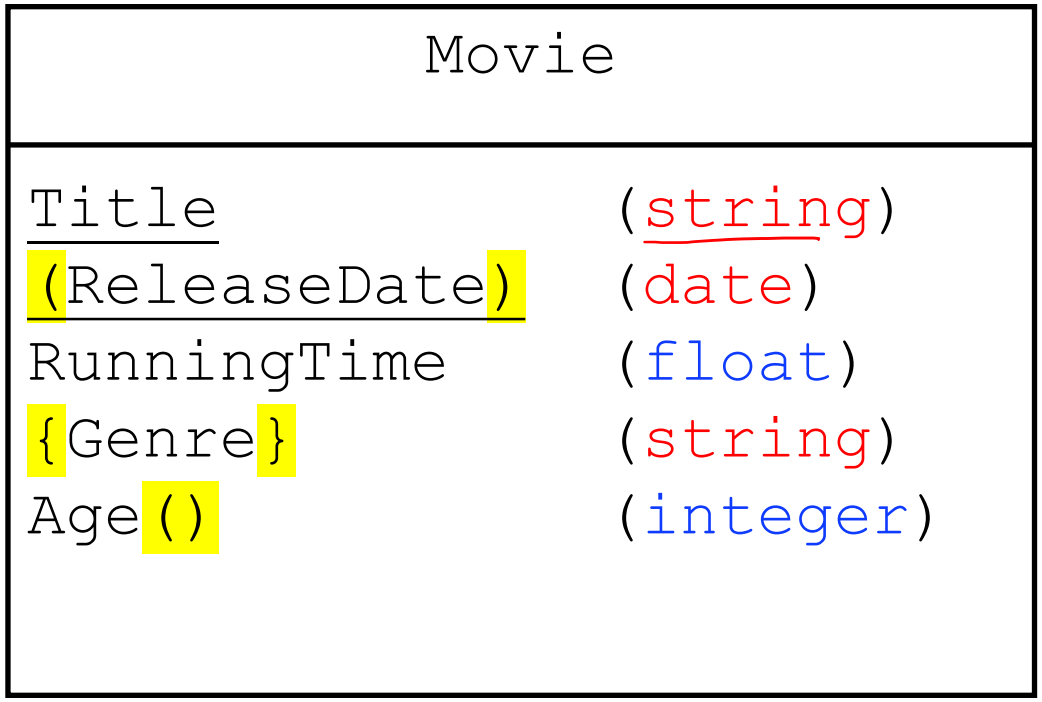
e_2.Age=YEAR(e_2.ReleaseDate-NOW)

E/R × Attribute



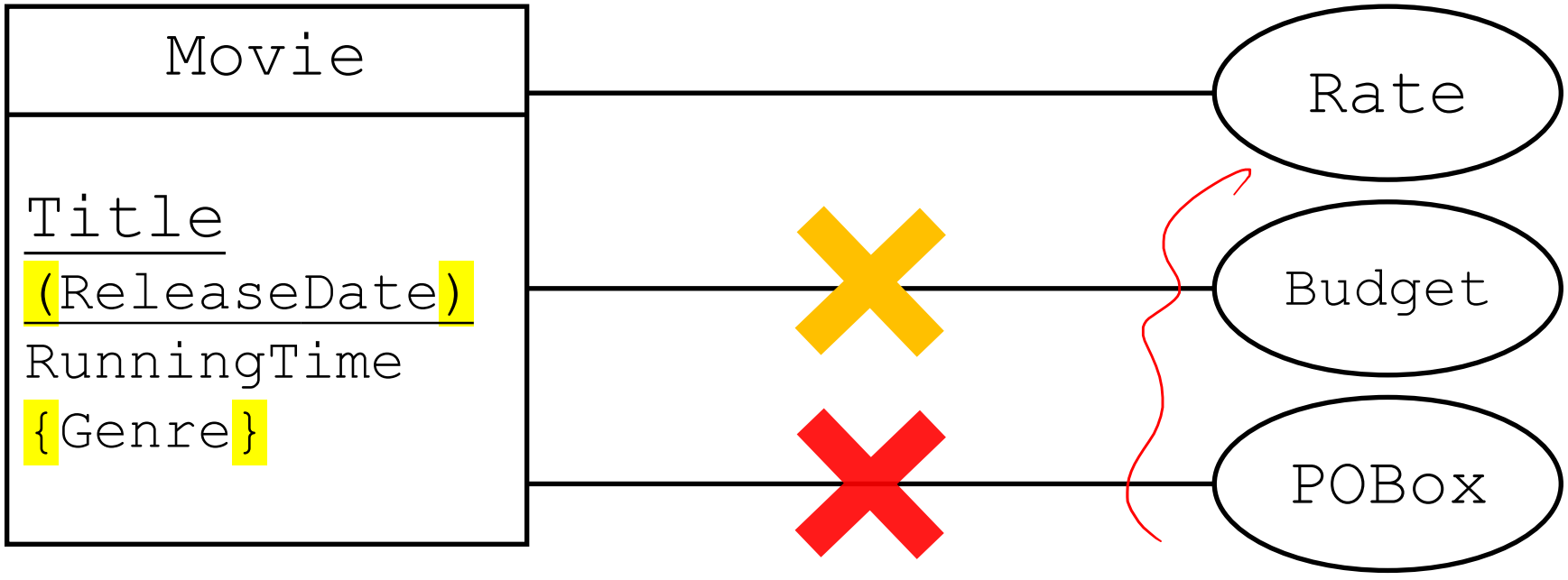
E/R × Attribute × Domain (Data Type)

Standard E/R model does not have!



E/R × Attribute × Faithfulness

Design should be faithful to the specifications of the application

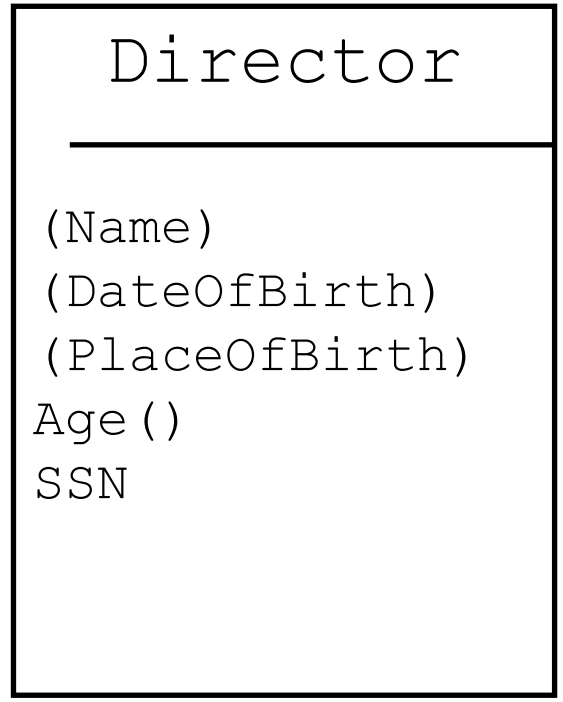


E/R × Attribute × Multiple Keys

Standard E/R model does not have!

K_1={Name, DateOfBirth}
K_2={SSN}

Choose?

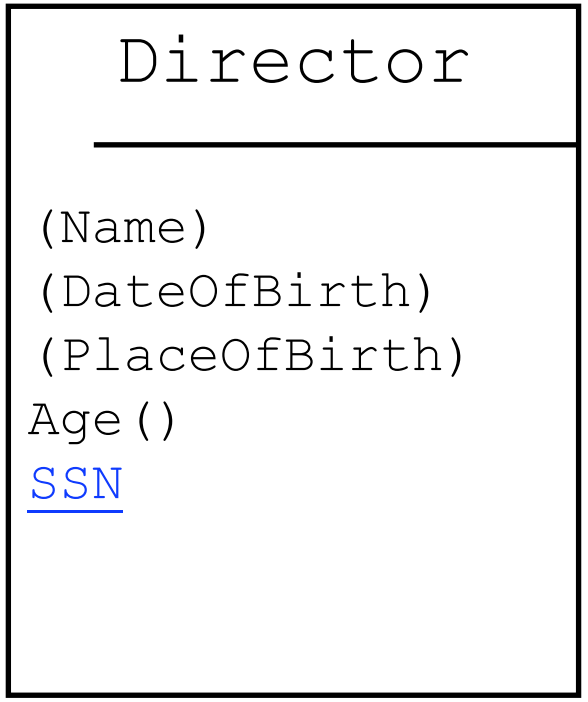


E/R × Attribute × Multiple Keys

Standard E/R model does not have!

K_1={Name, DateOfBirth}
K_2={SSN}

Choose?
Simplicity counts!
K_2 is called Primary Key (PK)

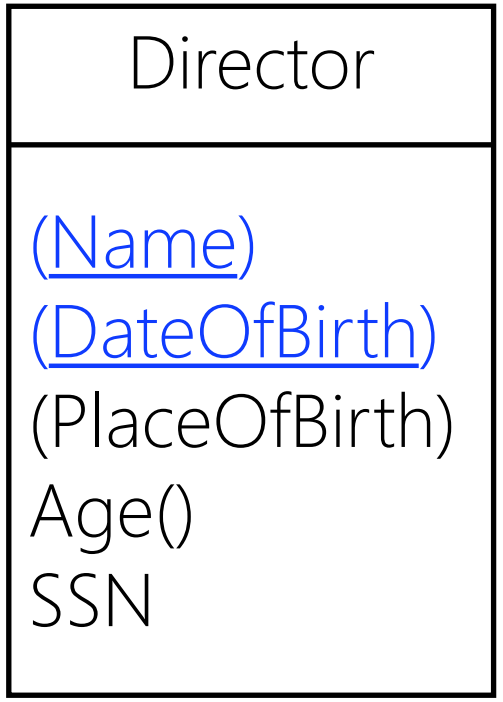


E/R × Attribute × Multiple Keys

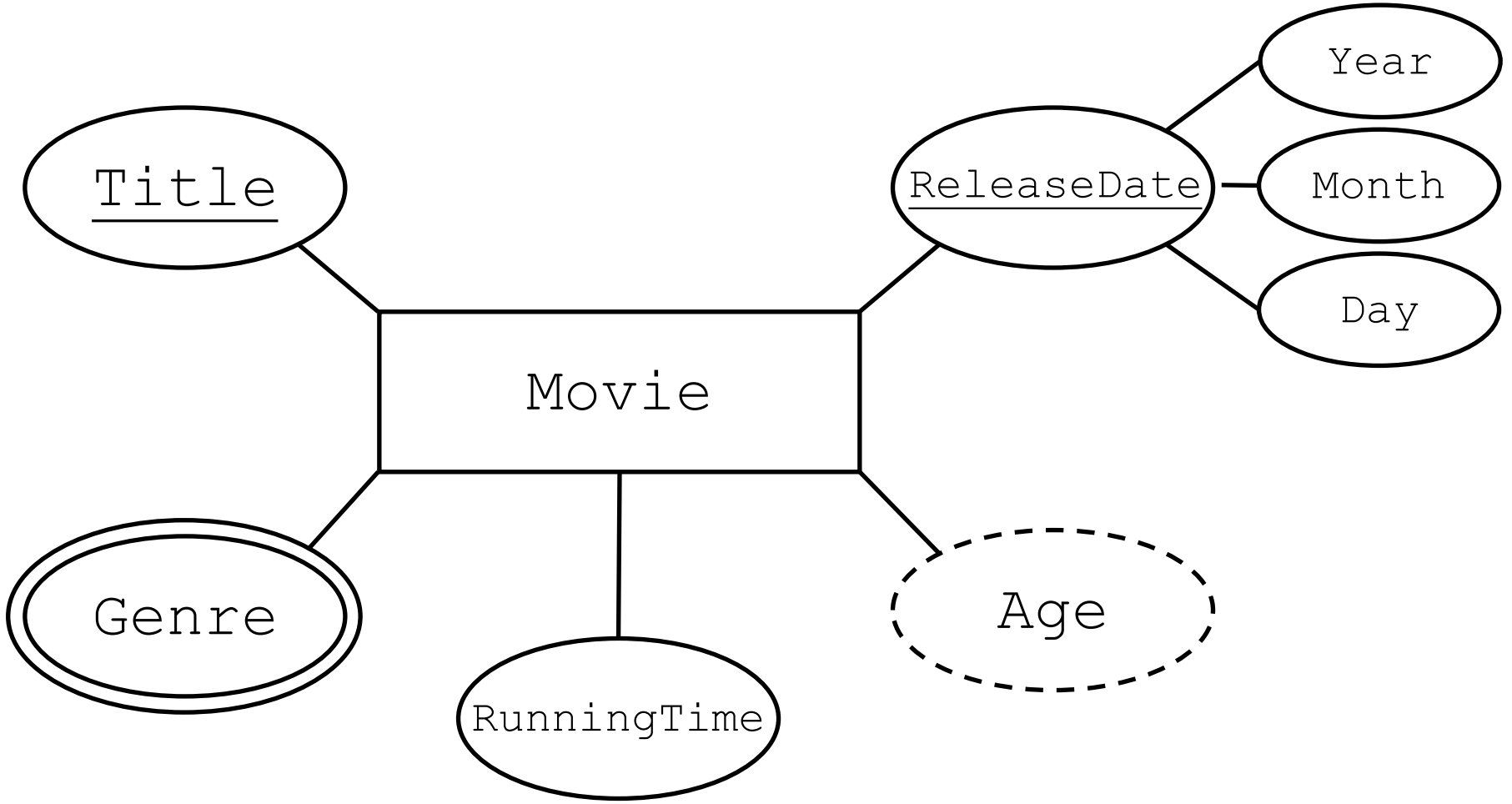
Standard E/R model does not have!

K_1={Name, DateOfBirth}
K_2={SSN}

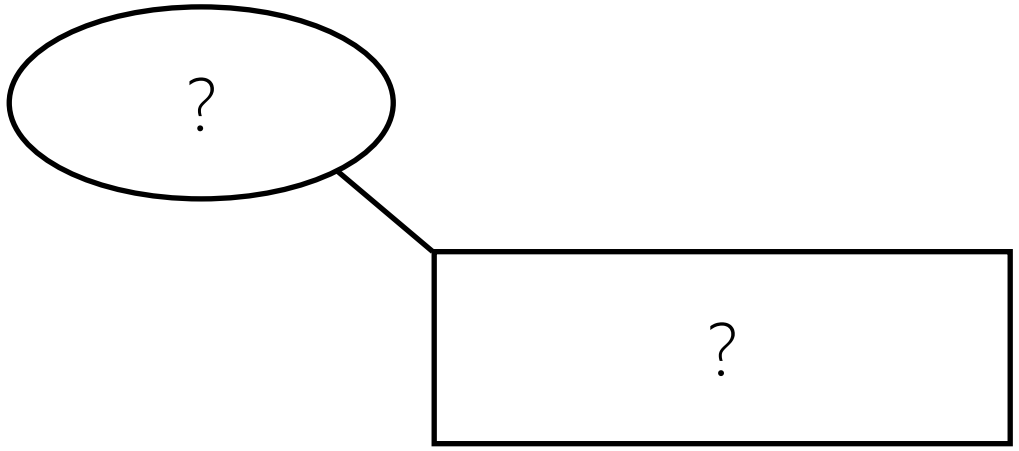
Choose?
Real world matters as well!
SSN is not available for all Directors!
K_1 is called Primary Key (PK)

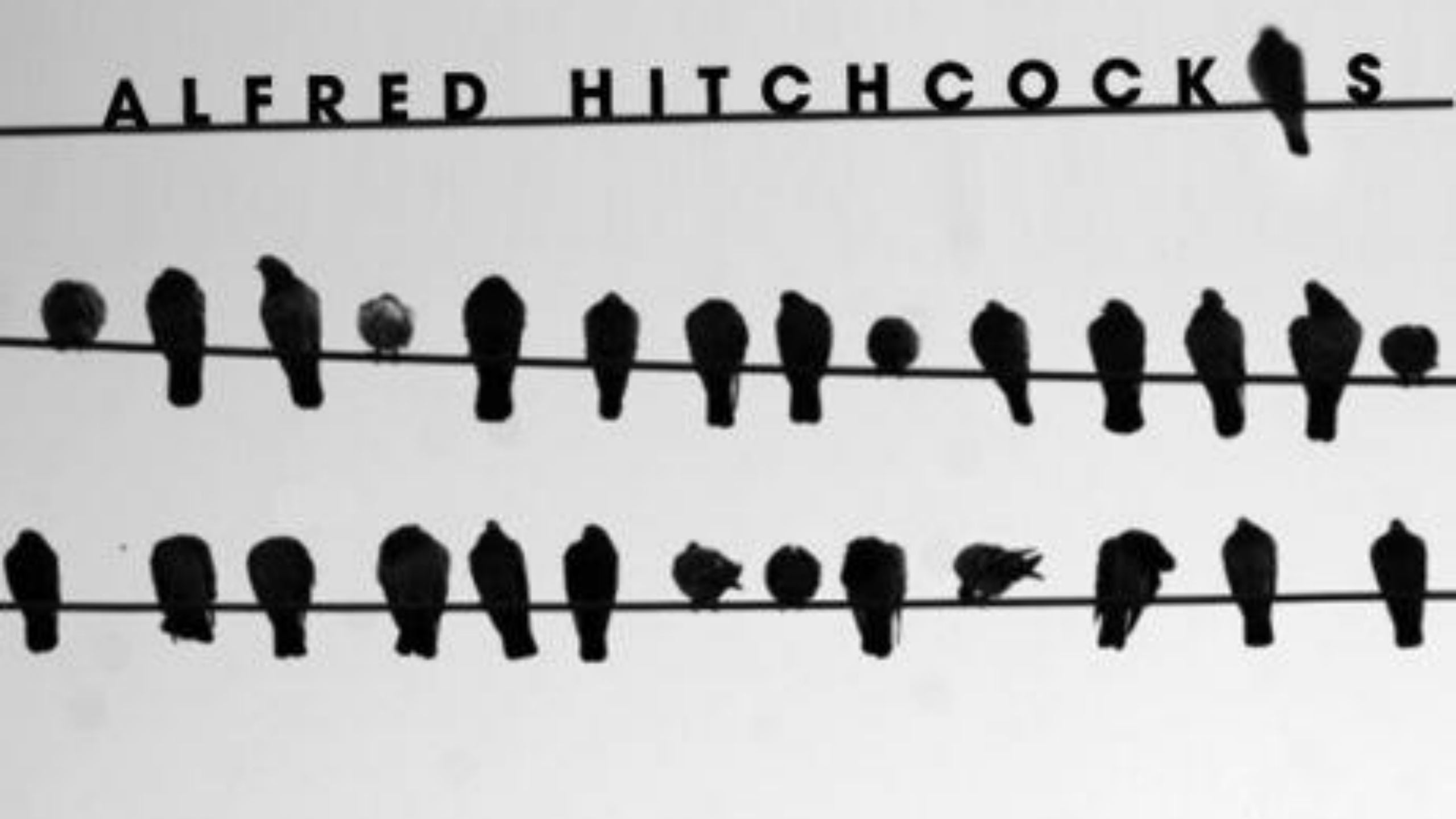


E/R × Movie Schema



E/R × Your System (15mins)




ALFRED HITCHCOCKS