

# UNIT 1

## CONCEPTUAL MODELING (EER)

**BASES DE DATOS 2023/2024**  
CFGs DAW

### WORKSHOP A: FIRST STEPS WITH ER

**Reviewed by:**

Sergio Badal

**Author:**

Paco Aldarias

Date: 09/21/23

License: Creative Commons



**Acknowledgment - NonCommercial - ShareAlike (by-nc-sa):** A commercial use of the original work or possible derivative works is not allowed, the distribution of which must be done with a license equal to that which regulates the original work.

**GENERAL RECOMMENDATIONS ON ER MODELING:**

1. Be consistent with the type of notation you choose (filled diamonds or max N, double lines or min 1, ...).
2. Include EVERY **participation (a,b)** and **cardinality (c:d)**.
3. [from Workshop B] Justify and describe EVERY **weakness** and every **ternary** relationship.
4. [from Workshop B] Do not create an **existence weakness** unless it is stated explicitly.
5. Do not cross lines or make a small bend in every cross if you can't avoid them.
6. Include just the **attributes explicitly mentioned** on the exercise and an **identifier** for each entity.
7. If you are not sure of which identifier attribute to choose, just set cod\_xxx or id\_xxx.
8. Name all entities in plural or singular (do not mix both), all relationships using verbs, and all attributes in singular.
9. An attribute should become an entity when it's multivaluated or it has more than one attribute.
10. Two binary relationships become one ternary when the three entities are always participating.

Although we all love to type and tinker with our computer, for this section I advise everyone to have a pen and a piece of paper nearby to make drafts until we get the model that seems to be the most appropriate.

It is also very important to keep in mind that when the E-R model has already some complexity, surely the solution is not unique. It is possible that one may be better than the other, but this is like programming, two different programs can solve the same problem, one may be more optimal than the other, but both work.

**Draw now the most suitable ER diagram for every context of the following.**

**YOU MUST HAND ONLY EXERCISES: 2, 3, 4, 5, 6**

**EXERCISE 1 (solved)**

Suppose Andrea writes computer books as a hobby and she would like to be able to store the information from the books she writes and publishes on her own. Of every book, we want to store ISBN, Title, Number\_Pages, Publication\_Date, Publisher.

CHECK THE SUGGESTED SOLUTION AT THE END OF THIS DOCUMENT

**EXERCISE 2 (solved)**

Andrea has been telling her colleagues at the secondary school where she teaches that she has created a database with the books she has written. Other teachers tell her that they have also written books, sometimes on paper and sometimes in digital format.

For this reason, Andrea has decided to modify her database so that a book **MUST** be written by an author and an author **CAN** write several books. From each author, their name, idcard (DNI) and specialty are requested.

Draw a new ER model (starting from the previous one) to reflect this new context.

CHECK THE SUGGESTED SOLUTION AT THE END OF THIS DOCUMENT

**EXERCISE 3**

Andrea and her colleagues know that the average time it takes an author to write a textbook is a couple of years, so they have decided to allow several authors work on the samebook.

Does the model we have from the previous example work for this new context? If you had to change something, what would it be?

Create a new ER model (based on the previous one) to reflect this new context.

**EXERCISE 4**

Several teachers from other schools have seen Andrea's database and have liked the initiative very much, interested in including their **works** in that database. On the other hand, Andrea and the other authors of **publications** have realized that there can be

several **texts** for the same subject and they want to reflect in some way which books **can** be used for each of the subjects taught, so that every book **is** suitable for a single subject. For each author, we need to know just the name his/her school (center). Each subject requires hours per week, code, name, studies and year (1st, 2nd, 3rd...).

*Be careful: the statement talks about **books, publications, works and texts**. If you notice, they refer to the same entity. This will often happen and you have to be careful not to create new entities.*

Draw a new ER model (starting from the previous one) to reflect this new context.

### EXERCISE 5

After having elaborated the previous model, when reviewing it we have realized that there are books that can be used for several subjects.

Does the model proposed in the previous image consider this possibility? What would need to be changed?

Draw a new ER model (starting from the previous one) to reflect this new context.

### EXERCISE 6

The initiative of Andrea and his colleagues has been very well received and many publishers have wanted to participate. Now we have that the authors of the books have been contracted by various publishers for the publication of their books. Each book has a contract with its authors and is exclusively with a publisher.

Create a new ER model (based on the previous one) to reflect this new context.

## SUGGESTED SOLUTIONS

## SOLUTION EXERCISE 1

The first thing we will do is to look for entities. We could think of "Authors" and "Books". However, how many authors are we going to store? We would only have one occurrence (also called instance): Andrea.

Is it worth saving in this case? Probably not. She is the one who is going to manage it and she already knows herself, she does not need to store information about herself. Is the entity "Books" needed? Of course it is, it is where she is going to store the information of the books she writes.

Secondly, look for the relationships. In this case we only have one entity and it doesn't seem we need reflexive relationships, so in this case we have no relationships and therefore we don't have to look for either participation or cardinality.

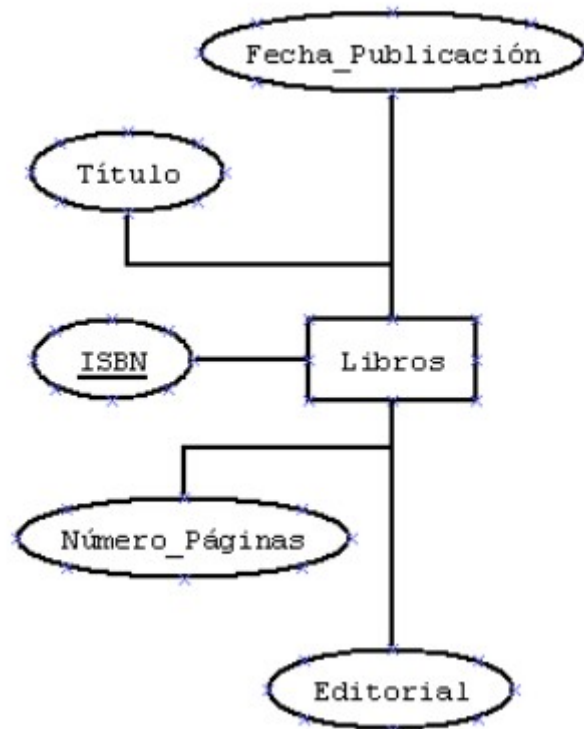
Let's go now for the attributes (also called fields or properties) of the entity. As the statement does not specify anything special we will place the most usual ones for this entity. However, this information should be checked with Andrea (the client), in case she wants to add something else or remove any that does not interest her.

Some of these attributes could be: ISBN, Title, Number\_Pages, Publication\_Date, Publisher. Remember you are only entitled to set an identification/main key attribute if no more are stated.

Now that we have of all the attributes, which one do you think can uniquely identify each of the books?

For a single author (as it is the case here) we could think that the "Title" could be a good candidate because it does not seem logical to publish two books with the same title, however what happens if you review the book three years later and make a new reprint? You could not use the same title, you would have to change it, since the main key cannot be repeated.

That's why we consider that the ISBN is the main key field, since it is a set of characters that you are assigned when you publish a book and it is unique for each book. If the book were to be republished, it would get a new ISBN when it is published, so there would be no problem.



At this point, maybe you think we could split the Publication\_Date into day, month and year but **AS A COMMON RULE** we don't split the dates into fields.

## SOLUTION EXERCISE 2

The context of our problem has been modified and, with it, our E-R model will also be modified. We perform the same process as in the previous exercise:

First, we look for the entities. Now we will be interested in having the entity "Author" since there are several authors that will be stored in that entity (it will have several occurrences) and on the other hand we also need the entity "Book" to store the information of the published books.

We look now for the relationships. The relationship will be between the entity "Author" and the entity "Book" and we can call it "has\_written".

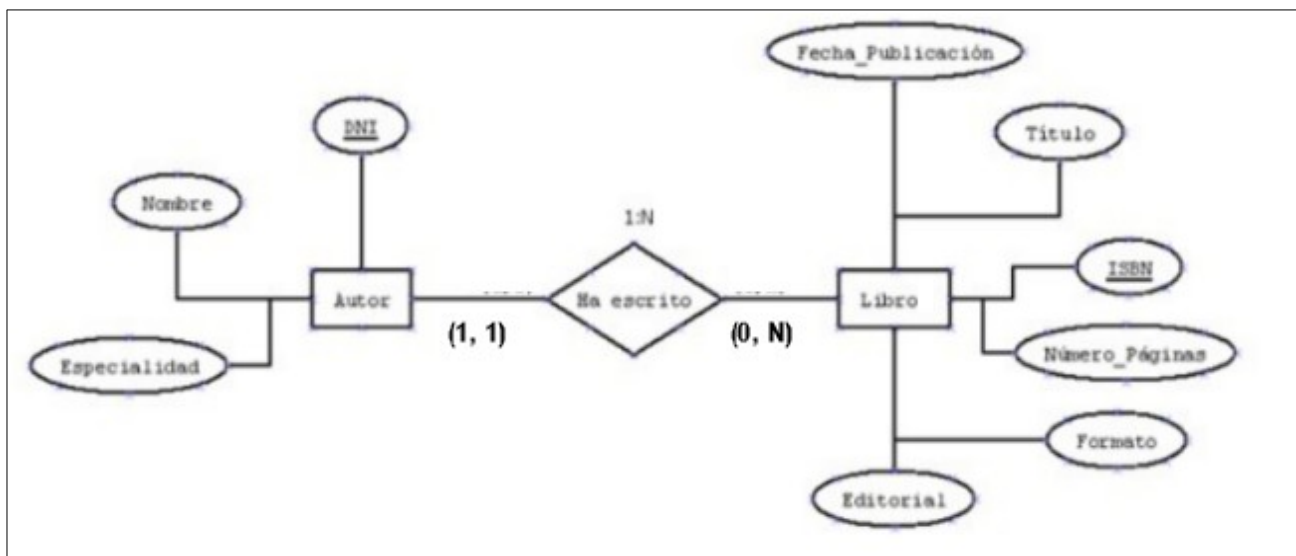
We go now for the participation of this relationship. An "Author" CAN write several "Books", then it would be (0,n). On the other hand, a "Book" has always been written by an "Author" and at most by only one, since the statement does not say that a "Book" could have been written by several authors at the same time (in case of doubt **we would have to ask the client** or confirm it by looking at the physical occurrences of the "Books"), then it would be (1,1). Then the cardinality (the maximum of each participation) would be 1:n.

Next, we look for the attributes of each entity. Some of the attributes of the entity "Author" could be: ID, Name and Specialty. Remember you are only required to set an identification attribute if no more are stated.

As a key for the entity "Author" we will use the field DNI (id\_card) which is unique for each of the occurrences/instances of the entity.

Some of the attributes of the entity "Book" can be: ISBN, Title, Page\_Number, Publication\_Date, Publisher, Format.

The primary key (or identification field) of the entity "Book" will be ISBN, as we show before.



**AS A COMMON RULE**, we will write all entities in singular "Subject", "Book"... or all in plural. Either option is irrelevant as long as the same criteria is applied to ALL entities in the same diagram.