

Computer Aided Archaeology

04 - Database 2

Martin Hinz

Institut für Archäologische Wissenschaften, Universität Bern

27/10/21

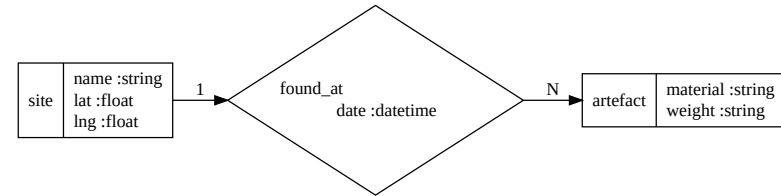
Designing a data model

Entity - Relationship (ER)

Entity:* Real-world object, distinguishable from other objects. An entity is described using a set of attributes.

Relationship:* Association among two or more entities. E.g., a fibula was found at Münsingen.

- relationships can have their own attributes.



***Entity Set:** A collection of similar entities. E.g., all employees.

- All entities in an entity set have the same set of attributes. (Basically)
- Each entity set has a key (!).
- Each attribute has a domain, that means, a range of possible values.

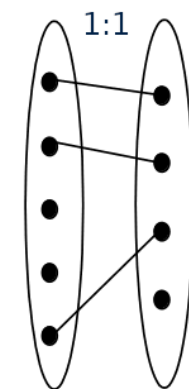
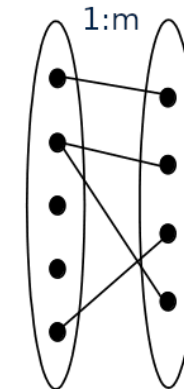
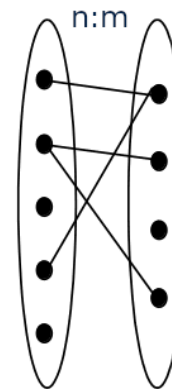
***Relationship Set:** Collection of similar relationships.

Types of Relationships

1:1, 1:n, n:m

Examples

- potsherds and features (n:m)
 - the sherds of one pot can be found at 1:n features
 - a feature can contain 1:n potsherds
- sample and measurements (1:n)
 - 1 sample has 1:n measurements
- artefact and find lable (1:1)
 - 1 Artefact has 1 find label



(primary) keys

Each record must be uniquely identifiable.

Primary key!

either

- a set of attributes that are already there and make the record unique
 - example: Lab Code and Lab Number identify a radiocarbon date

or

- is an explicit (artificial) attribute that is a sequential number
 - example: an id number from $1 \dots \infty$

The latter is not pure dogma, but most of the time more practical

(primary & foreign) keys

If a record is uniquely identifiable, this can be used in relation to other entities:

sites
Münsingen
Worb

burials
Burial 1
Burial 2
Burial 3
Burial 1
Burial 2
Burial 3

id	site
1	Münsingen
2	Worb

id	burial	site_id
1	Burial 1	1
2	Burial 2	1
3	Burial 3	1
4	Burial 1	2
5	Burial 2	2
6	Burial 3	2

The identifier of a record is the **primary key**.

The identifier of another record in relation to this one is the **foreign key**.

Normalisation

Database normalization is the process of structuring a relational database in accordance with a series of so-called normal forms in order **to reduce data redundancy** and **improve data integrity**. -- wikipedia

1NF

To satisfy 1NF, the values in each column of a table must be **atomic**. (Meaning one information at the time)

2NF

Each data record represents only one fact. If there is data in a table that does not represent only 1 fact, this data is subdivided into individual thematic tables.

or more formal:

It does not have any non-prime attribute that is functionally dependent on any proper subset of any candidate key of the relation. A non-prime attribute of a relation is an attribute that is not a part of any candidate key of the relation.

3NF

No data in a record should automatically follow from other data in the same record.

That's enough

There also exists the 4th, 5th and 6th Normal Form (not to mention the Boyce–Codd normal form (BCNF))...

In practise, normalising to the 3th Normal Form is absolutely enough.

Informally, a relational database relation is often described as "normalized" if it meets third normal form. Most 3NF relations are free of insertion, update, and deletion anomalies. --
wikipedia

Most of that comes naturally if you think about your relations as objects in the 'Real World'tm.

Let's get practical

We want to design a data base for finds of your site.

What **Informations** do we like to record?

What **Entities** and **Relations** do we have?

What **Attributes** will the **Entities** have?

How can we transform that into tables (this usually comes naturally than)?

Table:

site	literature	link
Birmensdorf-Rameren	Mäder, Andreas, Die mittelbronzezeitlichen Gräber von Birmensdorf-Rameren. Zürcher Archäologie, Heft 24. 2008	link
Chables	Leyla Duvanel, Henri Vigneau, Michel Guélat et Michel Mauvilly, La nécropole de l'âge du Bronze de Châbles/Les Biolleyres 1. Fribourg 2018	link
Singen	Rüdiger Krause: Die endneolithischen und frühbronzezeitlichen Grabfunde auf der Nordstadtterrasse von Singen am Hohentwiel (= Forschungen und Berichte zur Vor- und Frühgeschichte in Baden-Württemberg. Bd. 32 = Die Grabfunde von Singen. Bd. 1)	link
Murten/Löwenberg	Archäologie und Autobahn A1 : 25 Jahre Ausgrabungen im Murtenbiet	link
Prag-Miškovice	Ernée , Michal. Prag-Miskovice Archäologische und naturwissenschaftliche Untersuchungen zu Grabbau, Bestattungssitten und Inventaren einer frühbronzezeitlichen Nekropole Römisch-Germanische Forschungen Band 72 2016	link
Spiez, Einigen, Holleeweg 3	Gubler, Regula, Spiez-Einigen, Holleeweg 3. Gräber am Übergang zwischen Früh- und Mittelbronzezeit . Archäologie Bern/Archéologie bernoise – 2010, 147	link

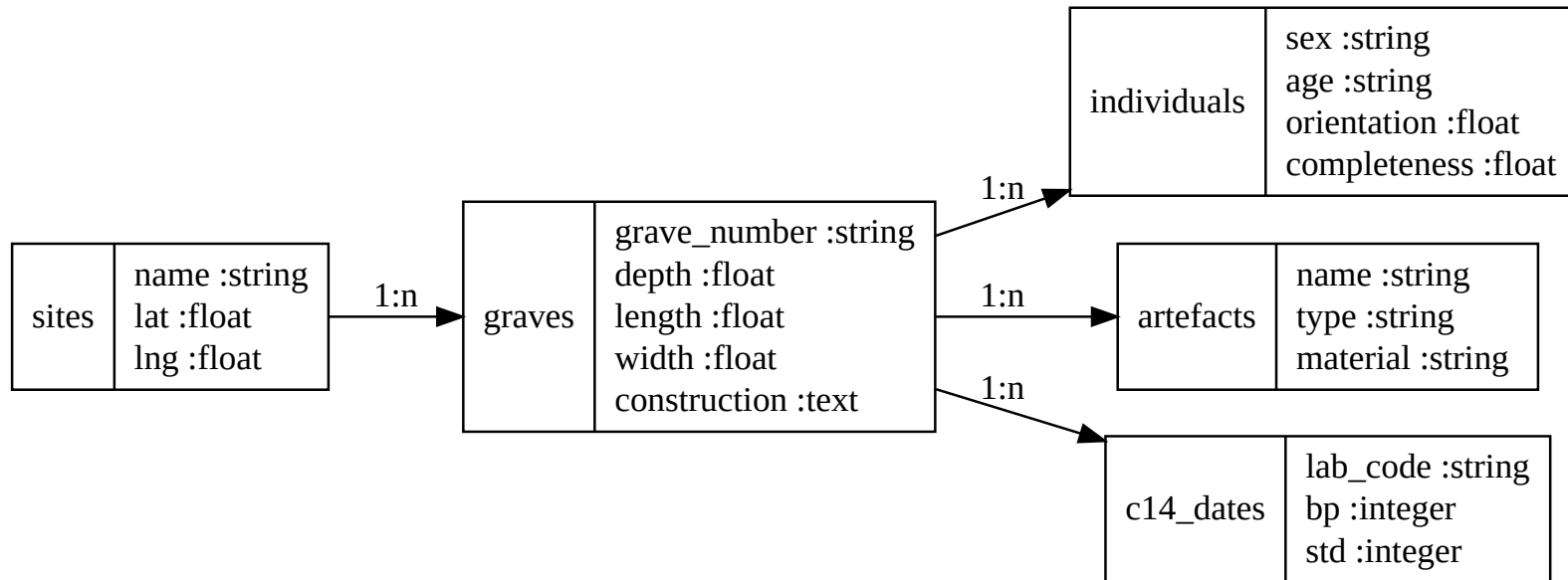
Structure of your Database

Draft

You can start as you like, paper, blackboard, mindmap, actual ER design tool...

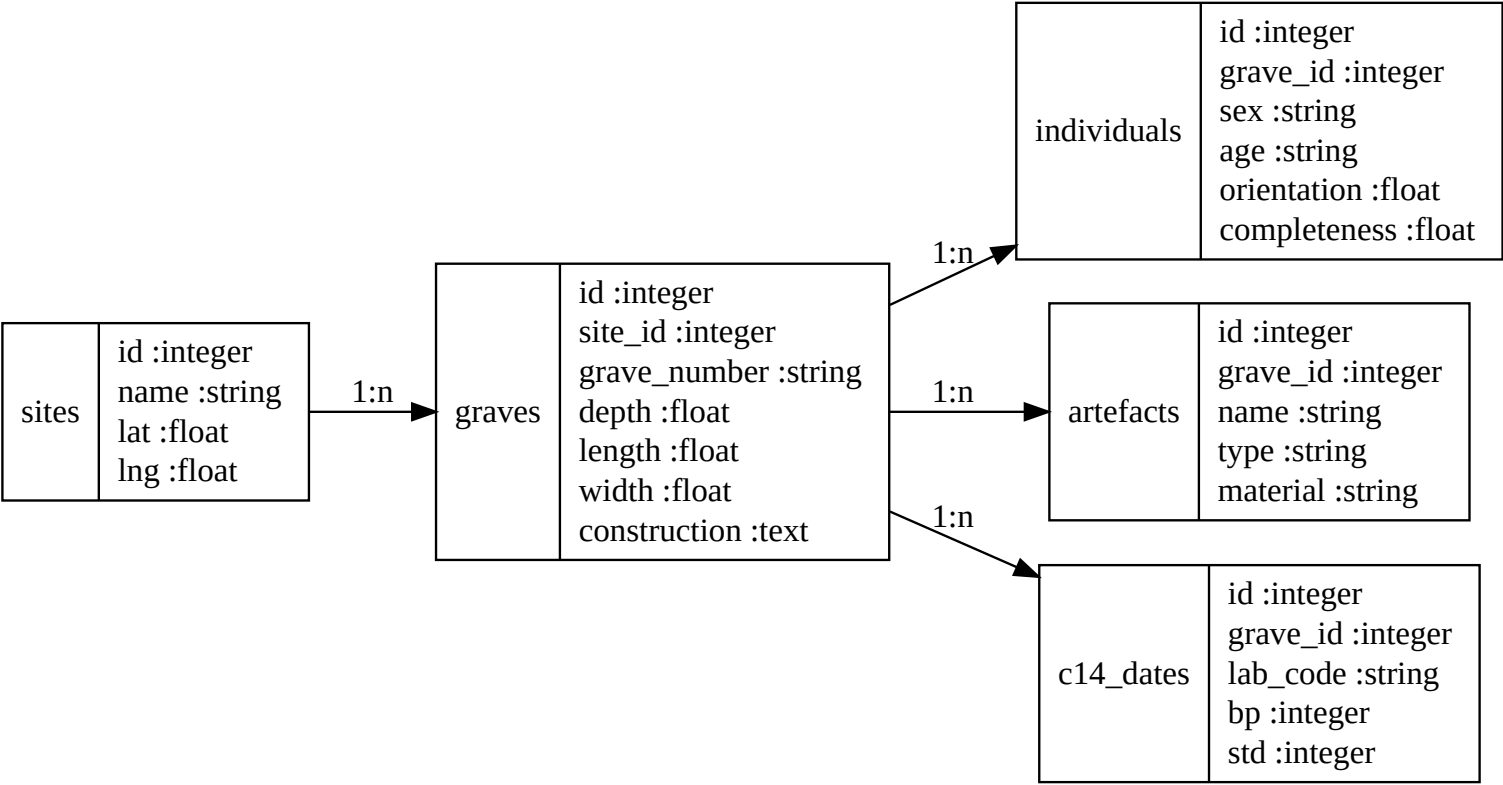
Structure MiscoviceDB

ER-Diagramm



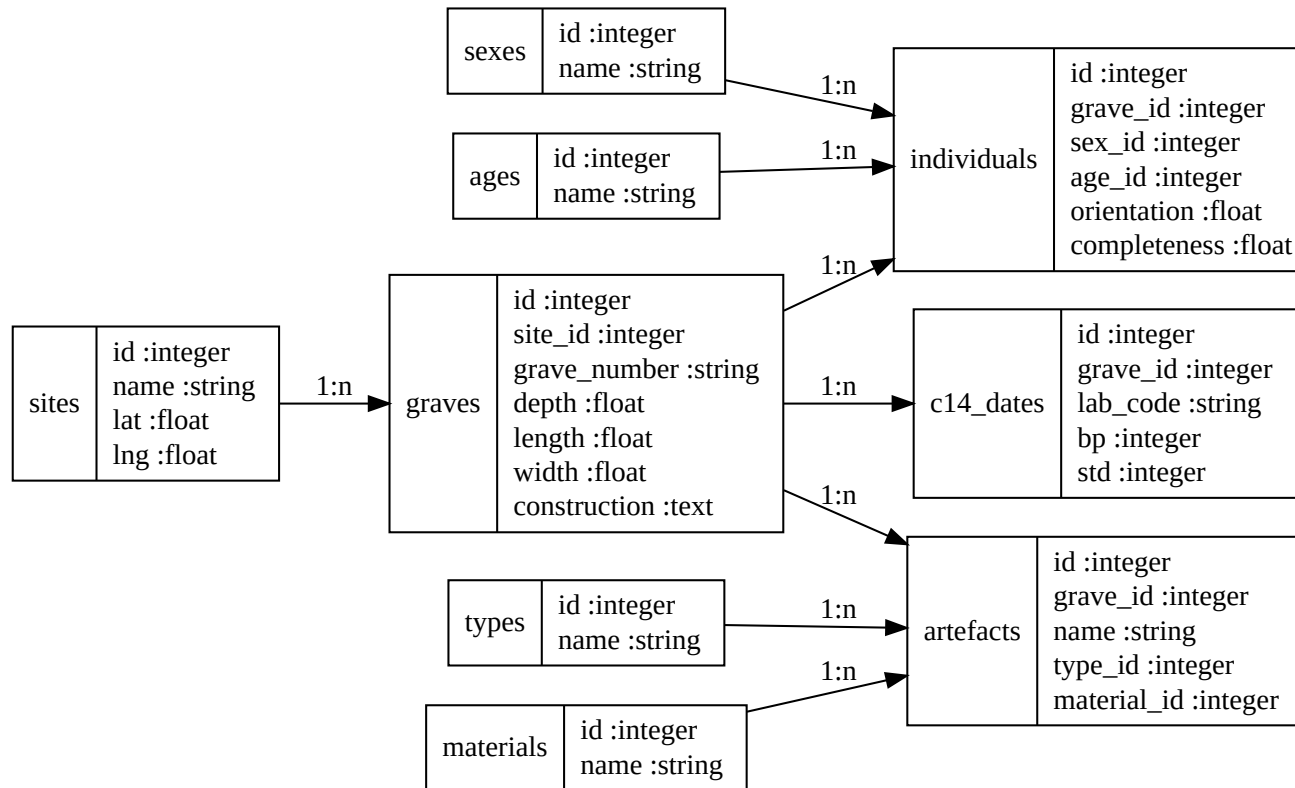
Structure MiscoviceDB

Primary and foreign keys



Structure MiscoviceDB

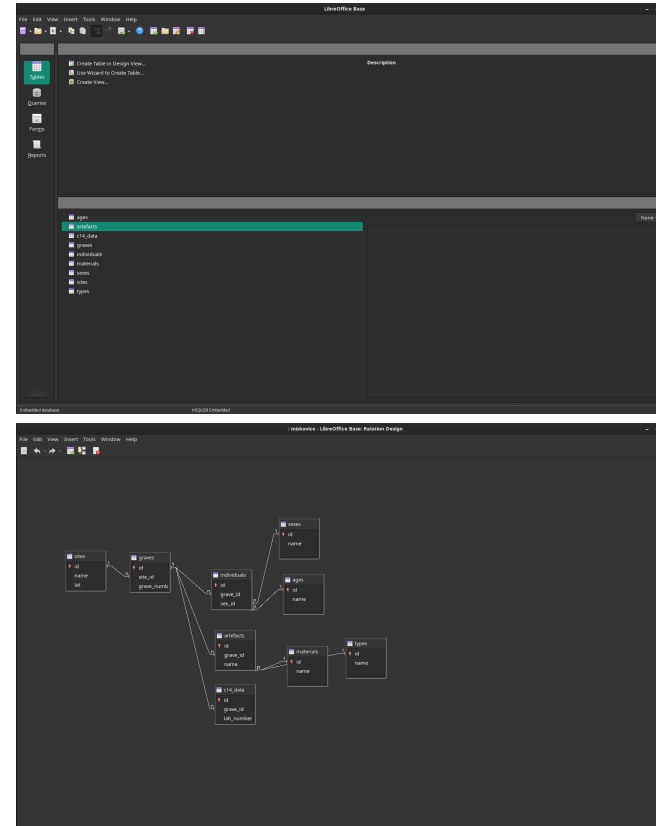
Lookup tables



Structure reflected in the DB-Program

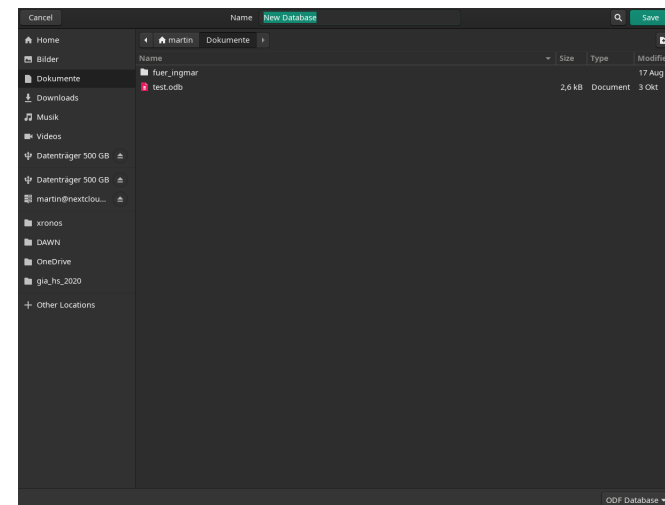
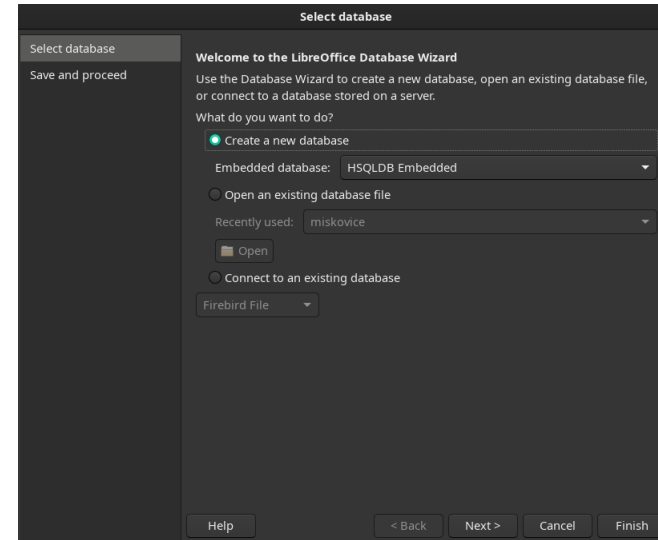
- Tables -> Tables
- Relationships -> primary and foreign keys, relationship using the relationship tool

Things look different in eg. MS Access, but work the same



To start

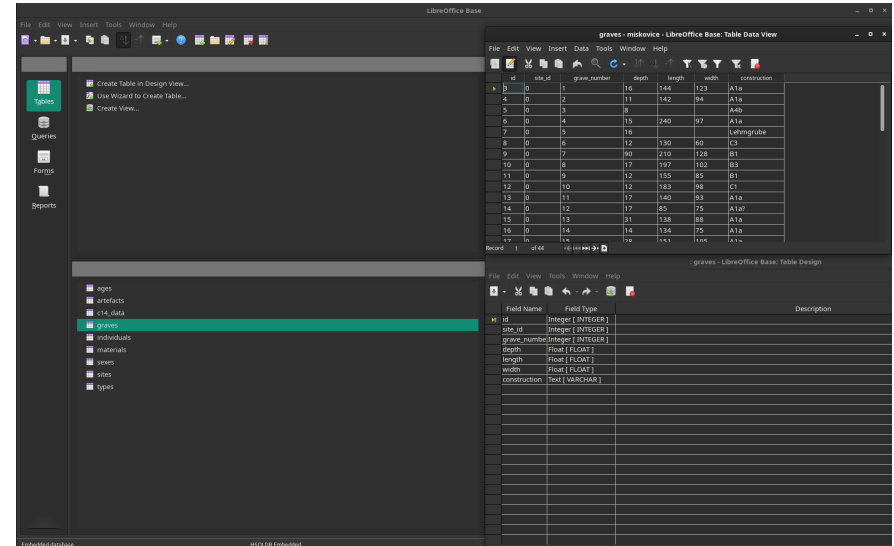
- Open LibreOffice Base
- Create a new Database
- Click on Finish
- Save your new Database to a reasonable folder



Tables

Store your information on specific items (rows) with defined fields (columns)

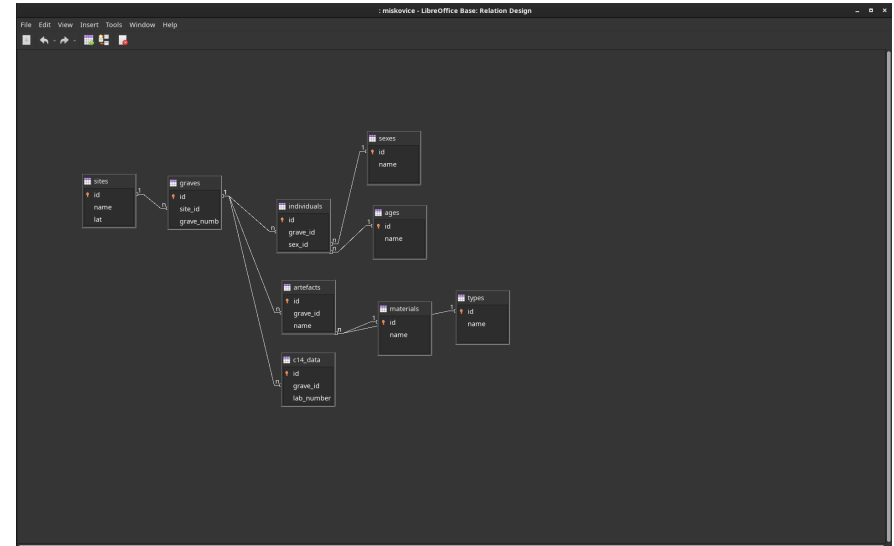
- Double click shows the stored information
- Right Click > Edit lets you change the field definitions (dangerous, if you have already data!!!)



Relationships

Define the relationships between your tables

- most important is the reasonable naming of the primary key ('id'!?) and the foreign keys ('site_id'!?)
- Defining the relationships in the "Relationship tool" helps to ease things later on (it provides information for the database tool)



Forms

Help to structure your inputs

- Give the user (you?) a nice entry form
- Enable to select values easily with dropdown fields and lookup tables
- limit and structure the possible entry values
- enable with subforms to enter data to different tables at once

graves - LibreOffice Base: Database Form

File Edit View Insert Format Styles Table Form Tools Window Help

site Miscovice grave_number 1

depth 16 length 144 width 123

construction A1a

name	type	material

Record: 1 of 41

sex	age	orientation	completeness
	matur II	163	

Record: 1 of 41

lab_number	bp	std

Record: 1 of 41

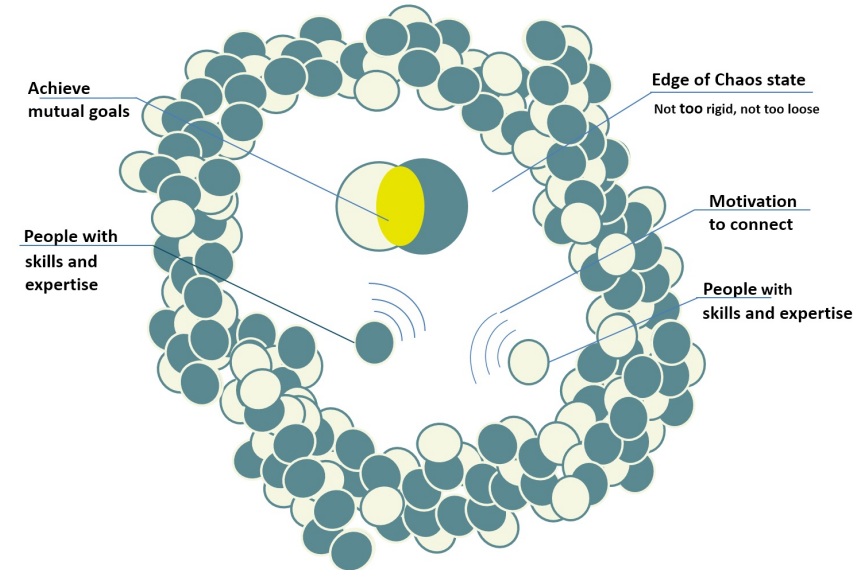
Page 1 of 1

Data Entry

Your task within the next two weeks:

- Enter data from your sites into the database with the structure you designed
- Distribute the work, share your workload!!!

Social collaboration happens at the Edge of Chaos

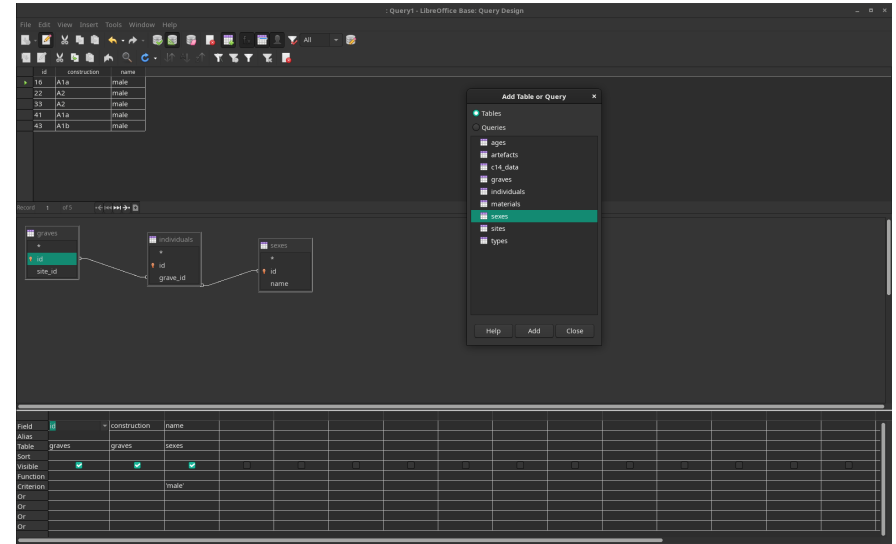


source: https://commons.wikimedia.org/wiki/File:Social_collaboration_at_edge_of_chaos.jpg

Queries

...are structured ways to receive information with specific conditions

- Most desktop DB systems offer "Wizards" and Tools to design a query
- In the background it will (nearly always) be SQL
- SQL ['si:kwəl]: "Structured Query Language"



```
SELECT "graves"."id", "graves"."construction", "sexes"."name"
FROM "individuals", "graves", "sexes"
WHERE "individuals"."grave_id" = "graves"."id"
AND "individuals"."sex_id" = "sexes"."id"
AND "sexes"."name" = 'male';
```

Any questions?

You might find the course material (including the presentations) at

<https://berncodalab.github.io/caa>

You can contact me at

martin.hinz@iaw.unibe.ch