Using SQL: The Basics for the Getty Vocabularies
Relational Tables data model

1 January 2023
Using SQL to access the Getty Vocabulary data in relational tables

- SQL stands for ‘Structured Query Language’
- SQL lets you access and manipulate databases
- For the Getty Vocabularies, SQL allows you to query Relational Tables database that contains the Vocabulary data in VCS
- SQL is a standard of the American National Standards Institute (ANSI) and of the International Organization for Standardization (ISO)

- Read about SQL here: https://www.w3schools.com/sql/sql_intro.asp

- See also OpenRefine: https://www.getty.edu/research/tools/vocabularies/obtain/openrefine.html
- See also SPARQL end point: https://vocab.getty.edu/queries#Finding_Subjects

 Getty Vocabularies: SQL 1 January 2023
• Boolean operators **and**, **or**, **not**
• Plus the use of parentheses and other punctuation to place phrases in logical order

baumann and (lydia or anna or daniel)

bauman% and (lydia or ann% or daniel)

(bauman% and (lydia or ann% or daniel)) or (bauer and (carl or karl or johan%))
Select * from Term where Name like (bowman or bauman%) and (lydia or anna)

1. Bauman, Lydia
   (British painter of Polish birth, born 1955) [500155374] (Persons, Artists)
   Lydia Bauman

2. Baumann, Elisabeth
   (Danish painter and writer, 1819-1881) [500009695] (Persons, Artists)
   Anna Maria Elisabeth Jerichau Baumann
   Baumann, Anna Maria Elisabeth Jerichau
   Elisabeth Maria Anna Jerichau-Baumann
   Jerichau-Baumann, Anna Maria Elisabeth
   Jerichau-Baumann, Elisabeth Maria Anna
To create SQL queries, use Relational Tables diagram to understand what are relationships between tables.

Each field is a column in a table (here horizontal); (each instance is a row).

The tables are linked (often using subject_id).

Easiest queries are done within a single table, but you may join tables for more complex queries.
Select distinct subject_id from revision_history
where user_name like 'PHARPRING'
and date_time > '8 Apr 2023
Select distinct subject_ID
from language_rels
where language_code like '70431'
or language_code like '70432'

What does this retrieve?

Select distinct subject_ID
from biography
where preferred like 'P'
and biography like '%&%'

What does this retrieve?
If you need to search on two or more columns in different tables, you need a “join”

```sql
select count(distinct term.term_id)
from term, subject, language_rels
where language_rels.language_code not like '70051'
and subject.candidate_stat like 'N'
and term.subject_id=subject.subject_id
and term.subject_id=language_rels.subject_id
```