CALD
Comparing Arabic Legal Documents: a database

How to use the CALD database

by Christian Müller

The database CALD provides a new approach to the study of legal documents in Islam. It was developed by the ERC-AdG-project Islamic Law Materialized (ILM). Its purpose is to study how the rules of Muslim legal thinking (fiqh) became "applied law", as its rulings "materialized" in the documentary relics of judicial practice. A particular attention is paid on documents that bear signs of use from the cadi-courts of law. CALD contains the meta-data and full Arabic texts, next to images of the originals, of documents from the 7th to the 16th century C.E. from all over the Muslim world. It allows a search for legal concepts and formula, in combination with a typology of documents and their dates.

CALD’s typology uses the function of documents as analytical approach. This may not be obvious from the outset and does not correspond to the titles given in printed text editions. You can find a document-type by its description and by a number, which fall into four parts, a frame, legal content, circumstances, and judicial certification. For more details and ongoing modifications on this, see "What’s new" on the website. Document typology and the identification of textual “sequences” within documents as well as the attribution of “sequence numbers” (SQN) for particular elements are the result of an ongoing research on the structure of these texts - and should be treated as such. CALD provides a hitherto unknown method of analysing legal documents. Its publication online is considered as one of the biggest achievements of the ILM project and makes CALD one of the most powerful tools of work on legal documents.

CALD describes each document by meta-data of 1) its physical support (manuscript-level) and 2) its notarized texts (item-level). The full Arabic text is divided into different relevant sequences (Arabic text-level). The complete Arabic text of an item forms the sequence-unity “complete text”, other partial regroupings are possible.

The following details technical aspects about how to use CALD for research.

Approaching legal documents in 11 steps:

Discover a document in CALD (step 1) and how to search for it as an object or by its Arabic text (steps 2-4) with the help of structured references (step 5). Go further with a login-account (step 6) in order to see the images (step 7), to save your personal research parameters (step 8) and your results with adapted export options (step 9), or to study document groups as a freakylining up of abstract numbers (step 10). Finally, step 11 explains technical choices on how to label documents, Arabic spelling conventions and document’s typology.

CALD published (version 2.1.3 - 3/11/2020) - ERC AdG 2008, n° 230261- ILM -
Step (1) What is a “legal document” in CALD?

As an example, the following screen shot shows the meta-data of the document Berlin, papyrus No 7514, here referred to with its “short inventory number” Ber_7514 (On this approach to cite documents see below, step 11: further explanations). The left blue box with several elements leads to the “manuscript description” – with information on “textual item” on the right side, especially its typology and date. For more details on this textual item you click on the icon with the magnifying glass and pencil (“details”).

The click on “Arabic text” brings you to the document’s text.

The text reads from top to the end, split into meaningful sequences labelled by numbers (SQN-definition and SQN-position, see step 5 “Reference tables”)

Step (2) Search for documents or Arabic text:

CALD offers two query-forms
1) Search via meta-data (dates, type of notarization, etc.) gives a list of documents and its items (see Step 3)
2) Search for Arabic texts and its components provides the Arabic text directly (see Step 4)

With these forms, you can query for
- meta-data
- Arabic words or expressions
- specific parts of a document (via SU or “sequence unities”)
- limit your search to various parameters, i.e. a time-period, types of documents, manuscript collections, etc.

Search forms combine criteria for advanced search options. Whatever forms you may use, the results will always link to the Arabic text and meta-data of your results.

Step (3) Search for manuscripts or textual items: meta-data

The click on “Search manuscript or textual item” on meta-data displays the interface below. You can search several fields combined with others. There is no option “or”.

Object/Descriptor/Value: this search concerns additional information, so called “descriptors”, searchable with “Object”, on two levels, “Manuscript” and “Manuscript’s item”.

In a second section, you can also search other specificities, according to “Item typology”, dates, Inventory numbers, etc.

In Bibliography you can search documents according to the study where it was published.

The results appear in a list of documents:

In this example, 17 parchments (manuscripts) contain 22 documents (textual items).

The results (headed in the grey line) could be organized by “Inv_number_short”, date and by type of the document (from left to right).
Use the green arrows in the grey headline for changing the order. A click on the icon next to Inv_number_short brings you to the “Manuscript”, a click on the same icon on the right side (next to Typology) opens up the “Textual Item”, and with the “نص” (in blue), you switch to the “Arabic text” (Step 1).

The Manuscript Search screen is an easy way to work with several documents at the same time, and you do not need to search for them individually.

**Step (4) Search of a text in Arabic:**
The form “Search Arabic text” opens as follows:

Three sections, marked by a horizontal grey bar, divide this screen:
1. **Section on the Arabic text.** By this, you can directly search for Arabic terms and/or sentences in Arabic characters, combined, when you need it, with SQN-positions and SQN-definitions.
   Without the help of a key word in Arabic, you can also search for textual items by different aspects as Item typology (see step 12: typology), date and inventory number. The following criteria on SU (sequence unities that combine various text sequences) groups of Arabic texts and the comment on Arabic text are rarely informed. “Arabic text – comment” mainly refers to data from paper editions, to
either variations or signs of a papyrological edition that indicate additions, holes or scribal errors.

2. Section (Object/Descriptor/Value): combines descriptors of the four levels: Arabic text or SUs, manuscript and textual item levels.

3. Section (Bibliography): renders all occurrences linked to the specified book or article. You may choose among various categories of bibliographical references.

Search results appear as follows (with the Arabic term highlighted in yellow):

Again in this search, with “نص” you might switch to the full Arabic text of the document (Step 1).

Step (5) Reference tables

The lists called “Reference tables” help you for defining your search, according to values already defined in the database by the collaborators of ILM.

It is recommendable to look through these lists in order to have an idea of the possibilities of CALD.

You may access the reference tables, after having opened the query-form by clicking on “Back to search menu” that will bring you to a search menu with the two forms and a third entry “Reference tables”.

Reference tables in CALD are always a work in progress and therefore may change and expand. See the webpage for major developments.
Step (6) Create an account and log-in (for Steps 7 to 10)
A CALD-account opens up several possibilities: access to images (Step 7), saving research parameters (Step 8), export results (Step 9) and SQN-analysis of documents (Step 10). All you need is to give your name and internet address. This is necessary for legal reasons but will never be used for commercial or data processing purposes.

Step (7) Access to images

The images of the documents in CALD are located at the level of manuscripts and you can access them from the blue box at the left (see step 1) - uploaded images are marked by an asterisk.

Once the image shows up, you may click on the magnifying glass (at the right) to zoom in or out - a new window will pop up and, therefore, you can easily juxtapose the edited Arabic text and its manuscript image in two windows.

Step (8) Save your research parameters for future use

As logged-in user, you might save the searched parameters of your previous work in CALD. For the same query, you only have to choose from the list of the drop-down menu (see below). You may rename the query or modify the order of the list by a click on “Saved searches”.

This feature will be only for your individual CALD-account.
Step (9) Export your search results
As a logged-in user, the head of your search-result-form will look as follows:

```
Manuscript and textual item search results

Search criteria
Item typology : (423126) real estate: sale
Number of results (manuscript) : 32
Number of results (manuscript's item) : 33

Save this search under this name : [text box]
Save

Export these manuscripts
SQN sequences analysis

<table>
<thead>
<tr>
<th>Inv. number short</th>
<th>Date</th>
<th>Typology</th>
</tr>
</thead>
</table>
```

If you click on “Export these manuscripts” and you can transform the results of any search into a simple-text format.

Export options

- inventory number
- inventory number (short)
- institution (library/museum)
- collection
- publication notice
- contributors
- Arabic text
- Arabic text comment
- Arabic text glossary
- Arabic text descriptors
- SQN explanations
- images
- images uploaded (yes/no)
- document dimension (width – height)
- descriptors
- keywords (publication)
- item typology
- bhl-date
- gregorian-date
- locus (of item)
- bibliography
- all other SmS
- Su descriptors
- incipit
- text layout
- ma_content (commentary)
- item_content (Arabic text)
- groups
- search criteria

Export all manuscripts

Step (10) SQN-analysis of a search result:
As a logged-in user, you might also find the “SQN sequences analysis” the head of your search-result-form. Use the SQN-analysis for in-depth comparison of document groups. SQN-analysis displays the documents by their SQN-number independently from the Arabic phraseology. This abstraction allows for discerning structure-patterns in documents’ sub-groups.
The following example shows the documents inventory numbers (SIN) at left, sided to its right by each document’s content as SQN-numbers (succession from left to right). Such text abstraction facilitates the orientation within larger document groups. The colour of each SQN-number refers to its occurrence within the selected group according to frequency.

A SQN existing in more than 80% of the selected documents turns yellow, 50% to 80% green, 25% to 50% blue, 10% to 25% violet, 5% to 10% pink, less than 5% red, and finally a single appearance is orange.

Different lengths of the SQN-chains reflect to shorter/longer or simple/complex documents. The order of documents depends on the order by the search-result form, with the possibility of changing (by date, inventory number or typology).

**Step (11) Further explanations**

- **Short inventory number (SIN)**

The SIN combines information on place and holding institution, eventually on particular collections within an institution, and the inventory number of each specimen. It consists of three letters for the city (Jer=Jerusalem, Gra=Granada), eventually a forth upper score letter for each institution within that city, the fifth underscore letter may refer either to different collections within an institution or helps to distinguish several institutions. The underscore sign marks the beginning of
the inventory number; additions should be shortened as much as possible. By adopting this shortened way of annotation, the SIN usually does not exceed 6-10 letters and can easily be cited (for abbreviations used for places, institutions and collections, see Reference tables).

- Arabic spelling conventions (Arabic text)

  CALD aims to produce comparable text units, not to reproduce of the exact aspect of the handwriting as in papyrological editions (being eventually referred to in the field of Arabic text “comment”). Therefore, the following conventions apply for “Arabic text”:

  - “yā” is written with punctuation,
  - “hamza” is given according to modern spelling (شئ/شئ نائب and not شئ/شئ \\
  - proper names with diacritical signs in the documents may be reproduced as such, otherwise no short vowels are used. Regarding personal names as ابرھيم إسحق أحمد اسمعيل, we write them as such, without alifs but only with hamza under alif.
  - Similarly, the names of Arabic and Coptic months (شعین جمدى) and numbers like سبعمائة/
  - the common spelling for Ibrāhīm is ابرھيم and not ابرھيم.
  - Arabic words with uncertain reading are followed by\[\] . In order to reproduce the rasm as accurately as possible without proposing a definitely wrong word – without reasonable reading – we sometimes, but rarely, use letters without dots, like 

- Typology

  Document typology is an important criterion in CALD. According to the overall research question of the ILM-project “how do legal documents reflect the praxis of Islamic law (fiqh)?”, we sharpen the definition of a “legal document” as being distinct from “documents with a legal content” in a wider sense, like administrative decrees or receipts, containing tax payments – even the Islamic tax zakāt, etc. The CALD corpus substantiates that subjective rights of individuals were notarized in Islam since early on in the presence of witnesses – informal notes on debts excluded. Witness attestation, the requirement of legal evidence in Islamic law, thus emerged as a major indicator of a legally valid document. All “legal documents” therefore should bear witness attestations or refer to an attestation, with some specific exceptions. The search of an item-typology could be done following the brief descriptions we gave to each category in the typology list (see reference table "Item typologies"). But the search might be better by using the numbers given to each type. These numbers are mainly divided in three parts (see table in "What's new" on the webpage):

  First part:
  - The first two digits show the functional or formal frame given for the document.

  Second part:
  - The third and fourth digits reflect the basic subject matters dealt with in the documents.

---

- Some specific subjects need a subdivision (fifth and sixth digits) for reflecting analytical transversal legal aspects.

Third part:
- The seventh and eighth digits of the type show, only when present in the document, judicial certifications and further legal proceedings.

Forth part:
- Exceptionally, ninth and higher digits indicate further legal specificities.

Several examples will give a better sense:
For example, the sale of Real Estate has the typology-number 423126. The first two digits, 42...... point to the form of the document as narrative of a legal action, ...31... point to the content as being a Real Estate and the final ......26 reflects the fact that this is a contract between two parties (2), The number 26 is given to a sale form. Another example might be also useful. A legacy has the typology-number 42351. Again, the first two digits show the narrative of a legal action, however ...35. in the third and forth digits reflect an aspect of property being transferred after death (heritage) and, finally, the fifth digit refers to a unilateral action.

**Extended search possibilities (wildcards):**
In general, you may use the underscore (_), as a wildcat to replace one number also in the middle of a typology, i.e. 4_35 for all types of documents concerning heritage (its perspicacity depends on nomenclature of typology).
*0: only types with a zero at the end are included in the search (that is all fragmentary documents). The same is possible for *1, *2, etc.
_ (underscore) represents any number (0-9)
You may combine “underscore” with “asterisk ”:
*__0 (asterisk, two underscores and zero): this will give you all documents with three numbers and zero at the end, i.e. 410, 480, 920, but 00 or 40730 will be excluded.
*___0 (asterisk, three underscores and zero) will choose all four number types, etc.
The operator exclamation mark “!” creates the negation of a typology.

In order to find all texts that are not Arabic legal documents (starting with 4), the search criteria is “4!” – with the exclamation mark excluding all 4 and its sub-categories.
In order to restrict the search to one type, it is possible to add a hash mark at the end of the number, so that, for example, 45924# will only give the "court case: šahāda (Andalus)" and exclude documents on court cases with additional legal certifications that are expressed in additional digits (e.g. 4592406 "court case: record, cadi ḥukm authentication").

In order to find all contracts of sale, for example, you may use 4___26 and you get sales of real estate (423126), sales of animals (423426) - also non-attested ones (403426), or those real estate-sales that were notarized in a cadi court of law (453126), etc. Concerning the latter, if you only search for the huğga-documents, use the hash mark # (453126#) that excludes other forms of judicial certification.
How to use the CALD database: Table of content

Approaching legal documents in 11 steps: ................................................................. 1
Step (1) What is a “legal document” in CALD? .......................................................... 2
Step (2) Search for documents or Arabic text: .......................................................... 2
Step (3) Search for manuscripts or textual items: meta-data .................................... 3
Step (4) Search of a text in Arabic: ........................................................................ 4
Step (5) Reference tables .......................................................................................... 5
Step (6) Create an account and log-in (for Steps 7 to 10) ........................................ 6
Step (7) Access to images ........................................................................................ 6
Step (8) Save your research parameters for future use ............................................ 6
Step (9) Export your search results .......................................................................... 7
Step (10) SQN-analysis of a search result: ............................................................. 7
Step (11) Further explanations .................................................................................. 8
  - Short inventory number (SIN) ............................................................................ 8
  - Arabic spelling conventions (Arabic text) .......................................................... 9
  - Typology ........................................................................................................... 9