1. **[Home Page](https://geniza.princeton.edu/en/) (Arabic only)**

Angry letters from a [wife](https://geniza.princeton.edu/en/documents/4033/) to her absent husband. Complaints about [business travel](https://geniza.princeton.edu/en/documents/9690/). Memo about an [opium](https://geniza.princeton.edu/en/documents/12442/) deal. Accounts of a [silk](https://geniza.princeton.edu/en/documents/2548/) merchant. Graphic descriptions of [gastrointestinal](https://geniza.princeton.edu/en/documents/2925/) problems. Complaints about [bedbugs](https://geniza.princeton.edu/en/documents/2935/).

The Cairo Geniza, a cache of texts preserved in a medieval Egyptian synagogue, is a snapshot of everyday life over the past [thousand](https://geniza.princeton.edu/en/documents/3092/) years, a portrait of the Jewish communities of the [Mediterranean](https://geniza.princeton.edu/en/documents/2505/) and [Red Sea](https://geniza.princeton.edu/en/documents/9675/) regions in an era when the vast majority of the world’s Jews lived in the Islamic Middle East.

Thousands of [letters](https://geniza.princeton.edu/en/documents/?doctype=Letter), [legal deeds](https://geniza.princeton.edu/en/documents/?doctype=Legal%20document), [accounts](https://geniza.princeton.edu/en/documents/?q=account) and [doodles](https://geniza.princeton.edu/en/documents/?q=doodle&sort=relevance) are fueling a new history of the global Middle Ages.

Try your luck with a [random fragment](https://geniza.princeton.edu/en/documents/). Explore the database by keyword in English. Search the transcriptions for specific terms in Judaeo-Arabic, Hebrew, Arabic and other languages. Or use the advanced search function. We encourage you to look for connections between people, places and objects as you engage with the corpus.

Follow us on [Twitter](https://twitter.com/GenizaLab).

Miss the [old PGP website](https://genizaprojects.princeton.edu/pgpsearch/)? It's still available for now; please let us know what you miss about it that you can't do in the new site.

## [**How to Cite**](https://geniza.princeton.edu/en/about/how-to-cite/)

## Anyone is welcome to use our transcriptions and descriptions, but please cite them accordingly.

Some general principles, with examples:

### *Always* mention the shelfmark when citing a geniza fragment. Example:

* Letter from Barhūn b. Mūsā al-Tahirtī to Nahray b. Nissim. Jewish Theological Seminary, ENA NS 18.24.

### Transcriptions that come from print editions should acknowledge the original publication. Example:

* Cambridge University Library, T-S 13J23.3, ed. Miriam Frenkel,*The Compassionate and Benevolent: The Leading Elite in the Jewish Community of Alexandria in the Middle Ages*‎, in Hebrew (Jerusalem, 2006), doc. 75, pp. 545–51; transcription available through the Princeton Geniza Project.

### Transcriptions that PGP lists as unpublished should be cited with credit to the editor(s) who transcribed the text and to PGP. Examples:

* Cambridge University Library, T-S 8J17.15, S. D. Goitein’s unpublished edition available online through the Princeton Geniza Project.
* Field-guide to taxation. ENA 2747.16, ed. Marina Rustow and Naïm Vanthieghem, available online through the Princeton Geniza Project.
* Letter from Mardūk b. Mūsā, Alexandria, to Nahray b. Nissim, Fustat. T-S 12.373, S. D. Goitein’s unpublished edition, with minor emendations by Alan Elbaum, available through the Princeton Geniza Project.

# [Glossary](https://geniza.princeton.edu/en/about/glossary/)

## data ontology

* **document:** a discrete unit of text
* **fragment:** a discrete piece of material support

*A fragment can have more than one document written on it; a document can be comprised of more than one fragment.*

## manuscript structure

* **folio:** a page, including both front (recto) and back (verso)
* **bifolio:** a page folded once to yield four writing surfaces (two rectos, two versos)
* **recto:** the front of the page
* **verso:** the back of the page

*In theory, recto should be the side of first inscription. In practice, conservators make the call before the cataloguers and scholars get to it; sometimes the side of first inscription is the one catalogued as verso.*

* **quire:** several bifolios nested like a pamphlet
* **codex:** several quires sewn together to form a book
* **rotulus:** a vertical scroll (text is written parallel to the short side of the page)
* **volumen:** a horizontal scroll (text is written parallel to the long side in columns)
* **writing support:** the material substrate to which the scribe applies ink. In the geniza, rag paper is the most common support, followed by parchment; a few pieces are written on cloth or papyrus

## physical condition

* **join:** two or more fragments—once part of the same page, quire, or codex—now separated, and sometimes even housed in different collections—which a scholar has matched. The fragments remain physically separate and retain their old shelfmarks, but the scholar who has matched them publishes the join with a “+” between the shelfmarks
* **lacuna:** a gap in the text where the writing support has torn

## scholarship

* **discussion:** a passage in a book or article in which a scholar has offered an interpretation of a document or part of a document
* **transcription:** copy of a document, whether typed or handwritten
* **edition:** a transcription that renders the document from manuscript to print
* **paleography:** decipherment; also the study of old scripts, often for the purpose of determining when an undated manuscript was written or identifying the scribe who wrote it
* **unpublished edition:** an edition from a scholar's corpus not previously published before their inclusion and display in our database; these editions were not found formally printed before digital publication by PGP and are not peer-reviewed as other modern editions might be when published in a book or journal

## genres of text

* **legal deed:** a contract, testimony or other document produced by or for a legal court, often with probative or dispositive value
* ***piyyut*:** liturgical poetry (Hebrew; pl. *piyyutim*)
* ***responsum*:** answer to a legal query by a rabbinic authority, usually made in writing (Latin; pl. *responsa*)

## dating and calendar systems

Geniza documents use a range of calendar systems depending on the document’s use and provenance. Many documents are not explicitly dated. For those that are, the dates are included in each entry, both as given on the document and as converted to common era dating. You can also convert dates yourself using this [calendar converter](https://www.muqawwim.com/).

The following is a non-exhaustive list of the most common dating systems you will encounter in geniza documents, as well as their relationship to common era dates.

* **Common Era (CE):** This is the western calendar that we use everyday, originally the Julian calendar, established in 45 BCE (before the common era), then, after the papal calendar reform of 1582 CE, the Gregorian calendar. Converted dates in PGP entries are Julian before 1582, Gregorian after 1582 (regardless of whether the community in question adopted the Gregorian calendar in 1582 or later; in the latter case, this is known as the *proleptic* Gregorian calendar, and is a convention among modern historians).
* **Anno Mundi (AM):** The Jewish calendar, based on the biblical accounts of the world’s creation, calculated from sunset on 6 October 3761 BCE (according to the proleptic Julian calendar). The Jewish calendar as it developed over the first millennium CE is lunisolar: the months are lunar and change with the new moon, but the years must remain in line with the agricultural (solar) calendar. The Jewish months don’t slip backwards as the Islamic months do (see Hijri calendar below). The Jewish year consists of either twelve or thirteen months, alternating in a regular pattern over a nineteen-year cycle, and beginning on 1 Tishri, which falls in late August or September. The months are Tishri (or Tishrei), Ḥeshvan (or Marḥeshvan), Kislev, Ṭevet, Shevaṭ, Adar, Nisan, Iyyar, Sivan, Tammuz, Av and Elul. In leap-years, there is a second month of Adar before Nisan.
* **Hijrī** **calendar (AH)**: The Islamic calendar, a lunar calendar of twelve months. The lunar year is 11 days shorter than the solar year, so the months slip backwards relative to the solar calendar (so Ramaḍān can fall in any season). The hijrī is calculated from 622 CE, the year when Muḥammad and his followers migrated from Mecca to Medina.
* **Seleucid calendar (Sel. or SE; also Anno Graecorum or AG)**: Though this calendar is now obsolete in Jewish practice, for the numbering of years, it is the most common calendar found in geniza documents. It is used only for years, and calculated from Seleucus I Nicator’s reconquest of Mesopotamia in 312–11 BCE. There were variants of the Seleucid era used in the Middle Ages. The Seleucid era in Jewish documents was deemed to begin on 1 Tishri 312 BCE. In practice, to convert a Seleucid date to CE, if the day falls before 1 Tishri, subtract 311 from the Seleucid year; if after, subtract 312.
  + ***Examples:***
    - *29 Elul 1311 Sel. = 29 Elul 1000 CE = 1 September 1000 CE*
    - *1 Tishri 1312 Sel. = 1 Tishri 1000 CE = 2 September 1000 CE*

# [Resources](https://geniza.princeton.edu/en/about/resources/)

Working with geniza fragments takes practice, but it can be mastered with a practice and patience. The first thing you'll need to do is to develop technical skills.

## Scripts

* Learning to read handwriting from an era that isn't your own (specialists call this palaeography) is challenging.
* Most geniza fragments are in Hebrew or Arabic script.
* If you're interested in Hebrew script, this [aleph-bet chart](https://github.com/judaicadh/cairogeniza/tree/master/_docs/Eckstein%20Alephbet%20Chart) (by Laura Newman Eckstein) contains the range of scripts you're likely to encounter in real-world manuscripts.
* If you're interested in Arabic script, check out this [Arabic palaeography resource page](https://guides.lib.umich.edu/islamicmsstudies/arabicscriptpalaeography) from Evyn Kropf at the University of Michigan. But due warning: most guides to Arabic palaeography are geared toward codices.
* Documentary Arabic is a little different, and also varies considerably over time and place. Take a look at the [Arabic Papyrology School](https://www.apd.gwi.uni-muenchen.de/aps/home/) from our friends at the [Arabic Papyrology Database](https://www.apd.gwi.uni-muenchen.de/apd/show_new.jsp).
* Advanced palaeographers may be interested in [these guides to Coptic (or zimām) numerals](https://genizalab.princeton.edu/resources/instrumenta-studiorum/greek-coptic-zim%C4%81m-alphanumerals), the numeric ciphers most frequently used in medieval geniza fragments.
* For more on palaeography, see our [FAQ](https://geniza.princeton.edu/en/about/faq/) under Skills.

## Languages

* A language you'll frequently encounter in geniza fragments is Judaeo-Arabic. That's a modern name for the Arabic language written in Hebrew script, often with some code-switching to Hebrew and Aramaic thrown in (especially when the writer is trying to be formal or sound erudite).
  + The best way to learn Judaeo-Arabic is to learn Arabic first.
  + If your Arabic is good enough to read real-world prose (as distinct from the graduated prose in textbooks), you're ready for this primer: [How to read Judaeo-Arabic manuscripts in five easy steps if you already know Arabic](https://docs.google.com/document/d/1hcwWwBuPEmxc_FHpSk1JWoBZvs5Y1LOn8v6PIqaP9mA/edit).
* There are also geniza documents in Hebrew, Arabic (in Arabic script), Judaeo-Persian, Ladino and other romance languages.
* You'll even find some Aramaic, but mainly in legal deeds and/or in frozen phrases. By the tenth century, Jewish Aramaic was no longer a spoken language outside a geographic region roughly contiguous with modern Kurdistan.
* Perplexities may, however, remain. When deciphering letters, for instance, you may be able to read every letter and translate every word, but you may still have trouble understanding what's going on. This is especially true with frequent correspondents, who have built up a context over time but aren't, alas, telling you what's going on. The key is knowing when the difficulty is merely yours or inherent in the text. Distinguishing between the two circumstances takes some experience; it can be easier to get there if you work with a study-partner or a group.
* Princeton regularly offers graduate-level seminars in Judaeo-Arabic, Arabic and other geniza-related subjects through the [Department of Near Eastern Studies](https://nes.princeton.edu/courses#Graduate%20Courses).

## Dating and calendar systems

* Sometimes documents are agreeable and explicitly give you their dates, but usually in calendars other than the common era (CE) such as the Anno Mundi, Hijrī and Seleucid. For more about the calendar systems in use in geniza documents (and how to convert those dates to common era dates) head over to our [Glossary](https://geniza.princeton.edu/en/about/glossary/).

## Geography

* Even though most of the documents in the PGP survived in Cairo, they come from a vast swath of the planet stretching from Iberia (modern Spain and Portugal) to Sumatra.
* Check out this [interactive map](https://pumagic.maps.arcgis.com/apps/webappviewer/index.html?id=2e9a88772d5a459bbf71f2cbdf118514), which contains some common points of origin for geniza fragments. The bigger the circle, the more documents started out there.

## Libraries

* Today, geniza fragments are housed in more than sixty libraries and private collections, mostly in Europe, North America, Egypt and Israel.
* This [graphic](https://docs.google.com/spreadsheets/d/e/2PACX-1vQIvpq9lTcdtfwV96u3MtuBm_k2jX-RRmbYd2rdTKW6xq1eCkHgvsOn2BmDVGN0Gmj3rXDSjE_OjXt0/pubchart?oid=1365174636&format=interactive) shows the distribution of fragments in some of the major collections.
* You'll find all the geniza collections worldwide listed in the [Friedberg Genizah Project Database](https://fjms.genizah.org/) (login required), together with images, bibliographic information and helpful functionality (like computer vision–based join-finding algorithms).
* Some libraries have open digital repositories of their geniza fragments, notably:
  + The [Genizah Collections of the Cambridge University Library](https://cudl.lib.cam.ac.uk/search?keyword=genizah&page=1&x=0&y=0), roughly half of all geniza fragments worldwide. The library's [Taylor-Schechter Genizah Research Unit](https://www.lib.cam.ac.uk/collections/departments/taylor-schechter-genizah-research-unit) publishes a newsletter called [Genizah Fragments](https://www.lib.cam.ac.uk/genizah-fragments), a column for new discoveries called [Fragment of the Month](https://www.lib.cam.ac.uk/collections/departments/taylor-schechter-genizah-research-unit/fragment-month), and [audio-visual material](https://www.lib.cam.ac.uk/collections/departments/taylor-schechter-genizah-research-unit/audio-and-video) made by the Unit's staff.
  + The [Genizah Collections of the John Rylands Institute and Library](https://luna.manchester.ac.uk/luna/servlet/ManchesterDev~95~2) at the University of Manchester
  + The [Genizah Collections of the Bodleian Library](https://genizah.bodleian.ox.ac.uk/?f%5Bms_collection_s%5D%5B%5D=Genizah;f%5Btype%5D%5B%5D=manuscript) at the University of Oxford. (The catalogue information is more than a century old; it's a good idea to cross-reference it in PGP.)
  + All the [geniza fragments at the Jewish Theological Seminary Library](https://dpul.princeton.edu/catalog?f%5Breadonly_collections_ssim%5D%5B%5D=Cairo+Geniza) are online by agreement with the Princeton University Digital Library (DPUL).

## Bibliography

* The website of our home unit, the [Princeton Geniza Lab](https://genizalab.princeton.edu/), contains plenty of other [resources](https://genizalab.princeton.edu/resources), including [datasets](https://genizalab.princeton.edu/resources/datasets) and a [bibliography](https://genizalab.princeton.edu/resources/bibliography) with links to articles, catalogues and dissertations.
* If you're looking more specifically for references listed by individual shelfmark, your best bet is the [Cambridge Bibliography Editor — Cairo Genizah Collection](https://cudl.lib.cam.ac.uk/bibliographies/genizah/search), which is extremely thorough, but of utility only for fragments housed at Cambridge, or the [Friedberg Genizah Project Database](https://fjms.genizah.org/) (login required), which is complete for editions but less thorough for passing references to fragments.

## Videos and Podcasts

* We've posted a plethora of [geniza-related videos and podcasts](https://genizalab.princeton.edu/resources/videos-podcasts) on the Princeton Geniza Lab website.

### Looking for a resource but not finding it here, or have a resource to suggest? [Contact us](https://geniza.princeton.edu/en/contact/).

# [How to Search](https://geniza.princeton.edu/en/about/how-to-search/)

Whether you’re new to the PGP or an experienced researcher of our corpus, understanding our search tools will improve your ability to find specific entries, easily browse the PGP corpus, and familiarize yourself with geniza materials. Though you can start by typing in a keyword of interest, sometimes that will bring up too many results to sift through. Fortunately, there are a few ways you can refine the results to get to the information that will most directly help your research.

## Search at a Glance

* You can search using a **keyword** or a **phrase** in English or the original language of the document
* You can search by **field**, such as shelfmark or PGPID by prepending your search with the field and a colon (e.g. **shelfmark:**)
* You can filter by scholarship records and document types to further refine your results

## Search Within a Specific Field

If you want to search within a specific field instead of across everything, you can specify a field name using the syntax field:term or field:“search term”.

For example, if you wanted to search for items where Avraham Maimonides is mentioned in the descriptions, you would use description:avraham.

Note: single word terms (e.g. avraham) don’t need to be in quotes, but any term with a space in it (such as “avraham maimonides” or a shelfmark like “T-S 10J12.16”) must be in quotes, or the search function will treat it as an OR search.

### You can search the following fields (note that the field names are case sensitive):

* **pgpid** – the PGPID is each document’s unique identifier in our database
* old\_pgpids – this is an older version of each document’s unique identifier (relevant for joins or documents that were previously entered multiple times into the database and have since been cleaned)
* **shelfmark** – a shelfmark is a locator; indicating where the physical manuscript is held
* collection – collections, often indicated in shelfmarks, tell us more about where an institution holds the physical item (some libraries have multiple collections)
* **description** – the description of each document
* **transcription** – transcription means a digital rendering of the document text in the original language (transcription:אברהם will pull all of the documents where someone named Avraham is mentioned in the text)
* **tags** – tags are keywords added by our scholars to describe attributes about a document (such as #india or #complaints). Note that our tagging system is inconsistent and incomplete.
* **language\_code** – this indicates the language of the document (language\_code:he will pull up all the Hebrew language documents, ar=Arabic, jrb=Judaeo-Arabic)
* **input\_year** – this indicates the year the document was added to the PGP

You can combine searching fields with other search syntax, including quotation marks (description:“avraham maimonides”), or Boolean operators (description:“avraham maimonides” AND tag:petition).your results.

## Search for an exact phrase

If you want to look for a particular phrase, wrap your search terms in quotation marks..

For example, if you wanted to search for records that reference Avraham Maimonides, son of Moses Maimonides, you could search “avraham maimonides*”* to distinguish him from his father, or “nagid avraham” to distinguish him from other people named Avraham in the database.

You don’t have to use quotation marks for single words, only for phrases.

## Combine Terms

[Boolean operators](https://libguides.princeton.edu/c.php?g=1215083&p=8887984) are connectors that let you control how your search terms are combined.

**The default behavior is OR — to match any term.**

* If you search for avraham maimonides, the search will return all results that contain the words “avraham” or “maimonides”.
* OR is an inclusive operator, meaning that it returns results containing any of your terms.

**Using AND allows you to combine search terms — to match all terms.**

* For example, if you want to find all results for responsa written by Avraham Maimonides, you could search “avraham maimonides” AND responsum*.*

**Using NOT excludes certain results.**

* For example, if you wanted to see all documents written by Avraham Maimonides **except** for petitions, you could search “avraham maimonides” NOT petition.

## Search and Filter Results

With or without a search term, you are now able to filter your search results by types of scholarship records available and document type.

Note that you do have to hit enter on your keyboard or click the magnifying glass for the numbers next to the filters to accurately capture how many documents fall under that category.

The filters are also additive: if you select Has Transcription and hit enter, the filters will update to show how many documents have transcriptions AND translations, descriptions, etc.

If the field is filled in (green on light mode and berry on dark mode) that means that filter is still operating and impacting the number of search results you see.

## Search and Sort

The PGP has three main ways to sort the results you get:

1. **Relevance** - This should try to give you the most relevant documents for your search term
2. **Scholarship Records** (Most/Least) - This will show your results in order of how many scholars have discussed, edited, and translated the document. By filtering for Most-Least, you are more likely to find documents that have been edited and translated into English. Least-Most will help you find documents that have received less scholarly attention. Note that the PGP has not systematically added scholarship records for any category other than editions that appear in our database.
3. **Random** - This generates a page of random documents to help you find new and interesting material!

# [FAQ](https://geniza.princeton.edu/en/about/faq/)

## THE CAIRO GENIZA

### What is the Cairo Geniza?

* The Cairo Geniza is a cache of roughly 400,000 manuscript fragments that survived in the [Ben Ezra Synagogue](https://egymonuments.gov.eg/monuments/ben-ezra-synagogue/) in Cairo.
* For more, see [What Is the Cairo Geniza?](https://genizalab.princeton.edu/about/what-cairo-geniza)

### What kinds of texts did the geniza preserve?

* An estimated ninety percent of geniza fragments come from long-form literary texts, including [liturgy](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=liturgy&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=), [Hebrew Bible](https://cudl.lib.cam.ac.uk/search/advanced/results?FacetCollection=Cairo%20Genizah&author=&excludeText=minute&fullText=masora%20bible&keyword=&language=&location=&page=1&place=&recallScale=1&shelfLocator=&subject=&tagging=1&textJoin=or&title=&yearEnd=&yearStart=), [rabbinic literature](https://cudl.lib.cam.ac.uk/search/advanced/results?FacetCollection=Cairo%20Genizah&author=&excludeText=minute&fullText=rabbinic%20mi%C5%A1nah%20talmud%20gemara&keyword=&language=&location=&page=1&place=&recallScale=1&shelfLocator=&subject=&tagging=1&textJoin=or&title=&yearEnd=&yearStart=), [philosophy](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=philosophy&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=), [medicine](https://cudl.lib.cam.ac.uk/search/advanced/results?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=medical&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=), [astronomy](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=astronomy&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=), [astrology](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=astrology&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=), [Jewish law](https://cudl.lib.cam.ac.uk/search?keyword=Hala%E1%B8%B5a&page=1&x=0&y=0), [lexicography](https://cudl.lib.cam.ac.uk/search/advanced/results?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=lexicography&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=or&title=), [poetry](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=poetry&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=) and [theology](https://cudl.lib.cam.ac.uk/search?author=&excludeText=&expandFacet=&facetCollection=Cairo%20Genizah&fileID=&fullText=&keyword=theology&language=&location=&page=1&place=&recallScale=0.0&shelfLocator=&subject=&tagging=1&textJoin=and&title=).
* The remaining ten percent are documents:[letters](https://geniza.cdh.princeton.edu/en/documents/?q=&doctype=Letter&sort=scholarship_desc), [legal deeds](https://geniza.cdh.princeton.edu/en/documents/?q=&doctype=Legal+document), [lists, accounts](https://geniza.cdh.princeton.edu/en/documents/?q=&doctype=List+or+table&sort=scholarship_desc), [state documents](https://geniza.cdh.princeton.edu/en/documents/?q=&doctype=State+document&sort=scholarship_desc) and other everyday writings and ephemera.

### Where were geniza texts written?

* While all the texts survived in Egypt, many were written in other [places](https://arcg.is/1C5jz10) across the Mediterranean and Indian Ocean basins, including the Iberian peninsula, North Africa, Sicily, Syria, Yemen, Iraq, Iran and India.

### When is the material from?

* The majority of geniza documents date to the period between 950 and 1250 CE.
* There are also sizable clusters from the sixteenth and nineteenth centuries.

### Where are the contents of the geniza now?

* Geniza documents can now be found in more than sixty libraries and private collections in Europe, North America and the Middle East.

### For more, see [Resources](https://geniza.princeton.edu/en/about/resources/).

## THE PRINCETON GENIZA PROJECT

### What is the Princeton Geniza Project?

* The PGP is a database devoted to the documentary geniza fragments, which make up ten percent of the whole Cairo Geniza cache.

### How did the project begin?

* The PGP was founded in 1986 by [Mark R. Cohen](https://nes.princeton.edu/people/mark-cohen) and [A. L. Udovitch](https://nes.princeton.edu/people/abraham-udovitch) to digitize transcriptions of geniza documents. The first 2,200 transcriptions they uploaded were those of their teacher [S. D. Goitein](https://genizalab.princeton.edu/about/history-princeton-geniza-lab/goitein-and-his-lab) (1900–85).
* Over the decades, the PGP has come to include transcriptions by many other researchers, as well as descriptions and research aids.
* Version 4.x of the PGP, launched in 2022, includes high-resolution images.

### What is the Princeton Geniza Lab? Is it the same as the PGP?

* The PGP is a longstanding project of the Princeton Geniza Lab, but the Lab now houses other projects.
* For more, see [History of the Princeton Geniza Lab](https://genizalab.princeton.edu/about/history-princeton-geniza-lab) and the Lab’s [Projects](https://genizalab.princeton.edu/projects) page.

### For whom is the Princeton Geniza Project intended?

* The PGP began as a resource for professional geniza researchers, but we’ve recently overhauled it to make it more accessible to non-specialists, including students and the public.
* Anyone interested in the social and economic history of the medieval and early modern Middle East and its Jewish communities can benefit from the PGP.
* Other disciplines can make use of our perspectives, resources, and approaches, e.g., other digital humanities projects, linguists, open source advocates, librarians, archivists and software engineers.

### What are the aims of the PGP?

* to provide access to geniza documents and the interim products of scholarship on the geniza, including unpublished notes and transcriptions of scholars in the field;
* to facilitate access to documents in order to fuel research into premodern global history;
* to capture all the documentary texts from the Cairo Geniza and related caches, such as the Ottoman and modern Jewish community archives in Cairo.

### How are PGP entries structured?

* PGP records include five kinds of information:
  + **Classifications.** Each document is titled with a shelfmark (the call-number in the collection where it’s housed) and classed into one of six [types](https://docs.google.com/document/d/e/2PACX-1vQP8TvNH7b8bysPwN64A3_oUu4zCoHtEHUWdSXQYrx0_RZAU9QNdcof_pm-cF-rH7ObEa5W8be2Zl7b/pub): legal document, letter, list or table, paraliterary text, or state document.
  + **Descriptive information.** Two-thirds of our entries have detailed descriptions of the document’s contents. Many also have #tags, but tags aren’t comprehensive; they merely represent the interests of the researchers who have done the tagging.
  + **Images.** We currently display images from two collections: Cambridge University Library and the Jewish Theological Seminary. The images are displayed in conformity with the International Image Interoperability Framework (IIIF). As more geniza-holding institutions adopt the IIIF, we will add their images to our site.
  + **Transcriptions.** Because it can be challenging to read the handwriting of medieval scribes, scholars produce typewritten copies that can be read easily and searched digitally. Transcriptions are also referred to as scholarly editions. The PGP currently has 3,707 transcriptions, with more to come.
  + **Scholarship records.** Our records list who has transcribed the document (as well as whether the transcription has been published, and if so, where). They also list the published books and articles or unpublished notes from which we have derived the information in our descriptions.

### Can you briefly describe your data model?

* First: what is a data model? A data model is a way to organize different types of data (some examples in our case: documents, fragments, images and descriptions), and to standardize how they relate to each other.
* At the core of our data model is a many-to-many relationship between physical *fragments* and the textual units that we call *documents*. A single *fragment* can contain multiple *documents*, as when a scribe used the blank back of a page to write another text). Conversely, a single *document* can be written across multiple *fragments*, as when a text was torn and the pieces now have different classmarks, and/or are in different libraries.

### Which philological transcription conventions do you follow?

* The PGP has followed varying sets of transcription conventions over its long history, sometimes reflecting the choices of the text-editors whose editions we have digitized.
* We instituted a transcription reform in 2021, but we've decided not to apply the new conventions retroactively, at least for the moment.
* Here are both the old and new [PGP transcription conventions](https://docs.google.com/document/d/1pyFajrGRYTzRsb2yFjfX0XmNqpz1kUO9mhlISGwfpxk/edit#heading=h.iwnb9i3748yu).

### How do I get involved with the PGP?

* We have a team of dedicated and talented researchers who have come to us from many directions. They include undergraduates, graduate students, postdocs and faculty at Princeton and other institutions, as well as teachers, librarians and other professionals interested in Judaic and Islamic studies.
* If you would like to contribute information to the PGP, we’ll soon be adding links to document records through which you can add suggestions. In the meantime, please [contact us](https://geniza.princeton.edu/en/contact/).
* If you would like to work for the [Princeton Geniza Lab](https://genizalab.princeton.edu/), write to us. We welcome inquiries from students and researchers, as well as data specialists, software developers and machine learning experts with ideas for specific research projects or modules.  
    
  SKILLS

### What languages do I need to know to read geniza documents?

* We are including more and more English translations in the PGP.
* If you want to read documents in the original, the main language you need is Judaeo-Arabic (Arabic written in the Hebrew alphabet).
* There are also many documents in Arabic, Hebrew, Aramaic, Ladino and Persian.
* In many medieval documents, there are [numbers written in Coptic](https://docs.google.com/document/d/1DVAGujbJbnaRODiBqV2ZdZW1rXYQN59XTiqkc3WuKLQ/edit?usp=sharing).
* For more on the languages that are useful for studying geniza fragments, see [Resources](https://geniza.princeton.edu/en/about/resources/).

### Wait. This Hebrew doesn’t look like the Hebrew I learned how to read. What gives?

* Modern, printed Hebrew is derived from square script, a formal register of Hebrew that geniza scribes wrote when they were getting paid to copy a text or trying to impress their readers. Scribes writing for everyday purposes tended to use more informal handwriting, with varying degrees of cursiveness.
* Learning to read premodern handwritten Hebrew is a skill that can be mastered with time and motivation. (You might want to start with this [aleph-bet chart](https://github.com/judaicadh/cairogeniza/blob/7c30bfe3276fad9c92e62aa432272a3f4a8c71c5/_docs/Eckstein%20Alephbet%20Chart/Alephbet.pdf) created by Laura Newman Eckstein for a [Zooniverse](https://www.zooniverse.org/) project we helped with, [Scribes of the Cairo Geniza](https://www.scribesofthecairogeniza.org/).)

### Why are there so few dots in the Arabic documents, and how on earth do you expect me to make sense of these scribbles?

* Documentary Arabic hands are notoriously difficult to decipher. These are some of the challenges:
  + a dearth of canonical dots, which renders many letter-shapes ambiguous; the sporadic dot phenomenon inspired this [classic study](https://www.dropbox.com/s/2kzd09zm5qs5pdb/Kaplony_2008_What%20are%20those%20few%20dots%20for.pdf?dl=0)
  + scribes’ reluctance to lift the pen, which created abusive ligatures, or strokes connecting letters that in standard Arabic writing should remain unconnected
  + *Verschleifung* (literally, “slurring”) pen-strokes so that letters are skipped or subsumed into other letters; for instance, اربع is often written لع (with *Verschleifung* of the letters and an abusive ligature after the *alif*)
* The good news is that most geniza documents are in Hebrew script, which is rarely as cursive as Arabic (but there are exceptions, among them the handwriting of Yehuda Halevi and Moses Maimonides).

### Which additional skills do I need to study the geniza documents?

* Patience and spreadsheets.
* Patience is essential because the texts are dispersed across dozens of institutions and they aren’t all cataloged.
* Good record-keeping is essential because there are thousands of documents in no particular order. If you have worked in an archive, you’ve had the luxury of someone else creating order before your arrival. If you work with geniza fragments, you’re often assembling your own archives (or dossiers, which is the technical term to use when the material you’re assembling wasn’t actually archived). We use lots of spreadsheets.
* Other technical skills one can expect to pick up include understanding how legal documents are structured, recognizing scribes by their handwriting, recognizing the names of coins and units of weights and measurement and learning the patterns of shelfmarks in dozens of library collections.

For more, check out our [Resources](https://geniza.princeton.edu/en/about/resources/) page!

# [Contact Us](https://geniza.princeton.edu/en/contact/)

For questions not addressed in our[FAQ](https://geniza.cdh.princeton.edu/en/about/faq/), or if you’re interested in contributing to the Princeton Geniza Project, please email us at [geniza@princeton.edu](mailto:geniza@princeton.edu).

For document specific questions, consider asking us on [Twitter](https://twitter.com/GenizaLab).

For questions about the site architecture or code, please refer to the technical documentation or email [cdhdevteam@princeton.edu](mailto:cdhdevteam@princeton.edu).

For general questions about the Center for Digital Humanities at Princeton, please email [cdh-info@princeton.edu](mailto:cdh-info@princeton.edu).

For issues related to the site’s accessibility, please use the [accessibility assistance form](https://accessibility.princeton.edu/help) to request assistance.

# [Technical](https://geniza.princeton.edu/en/about/technical/)

Princeton Geniza Project v4.x is implemented as a custom Python/Django web application; the documented and tested [source code is available on GitHub](https://github.com/Princeton-CDH/geniza).

The project codebase includes a [change log of features by version](https://github.com/Princeton-CDH/geniza/blob/main/CHANGELOG.md); [database diagrams and documentation](https://dbdocs.io/princetoncdh/geniza) are auto generated using [dbdocs](https://dbdocs.io/).

See [How to Cite](https://geniza.princeton.edu/en/about/how-to-cite/) for information on citing the project as a whole or (eventually) this codebase in particular.

## Software developed

This project makes use of several locally developed open source packages; many were developed, enhanced, and released during the course of this project.

## Platforms, Libraries, and Technologies

Princeton Geniza Project v4.x relies on an ecosystem of existing open and free software packages, libraries, and languages. The team is grateful for the many organizations and individual developers who share our commitment to creating well-documented, open-source software solutions: