



ride. A review journal for digital editions and resources

published by the IDE

Review of “Papyri.info”

Papyri.info, Joshua Sosin (ed.), 2010. <http://papyri.info> (Last Accessed: 23.10.2018). Reviewed by Lucia Vannini (Institute of Classical Studies), lucia.vannini (at) postgrad.sas.ac.uk.



Abstract

Papyri.info, made available by the Duke University, is a text collection of over 50,000 documentary papyri, i.e., Greek and Latin documents, dating back to the IV century BC – VIII century AD, which constitute a fundamental body of evidence for ancient everyday life in the classical antiquity. The collection consists of transcriptions encoded in EpiDoc (a subset of TEI for the representation of ancient documents preserved in inscriptions and in papyri), metadata, links to related resources, and, for some records, images and translations. It also includes a platform for the scholarly community to contribute to the database, whether for the digitisation of already published material or for the proposal of new contributions, with an editorial vetting process. This review aims to illustrate the content of the collection, discussing whether the relevant information is presented in a clear way to the user; to assess the usability of the interface for searching and browsing the papyri; and to analyse the technical aspects of the resource, especially the integration of different databases and the possibility of downloading and reusing the data, while highlighting, for all these aspects, both strengths and features that could be improved.

Aims and content

1 *Papyri.info*, launched in 2010,¹ is an extensive digital text collection of Greek and Latin documentary papyri, that is, ancient documents (accounts, petitions, tax receipts,

letters, etc.) dating from the IV century BC to the VIII century AD, mostly written on papyrus, and mainly discovered in archaeological excavations in Egypt.

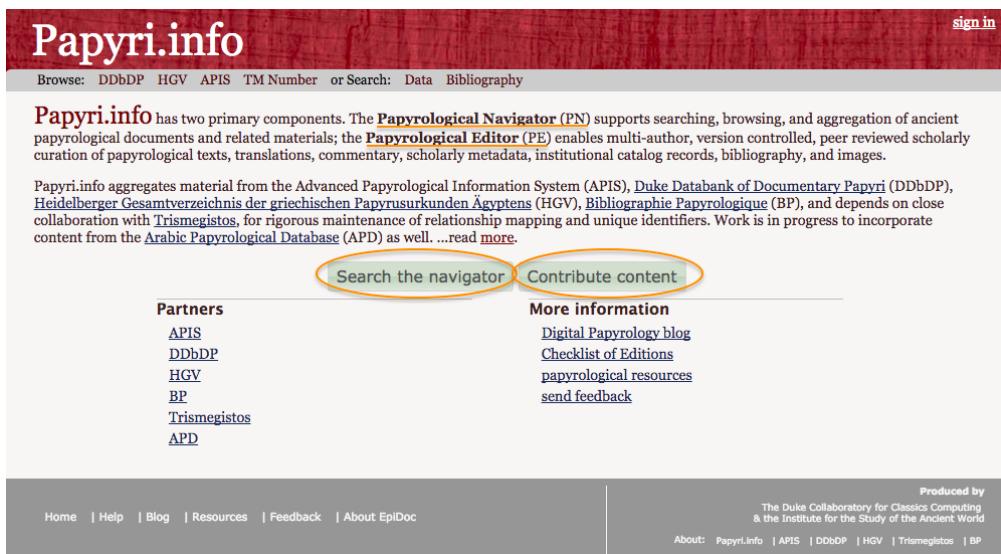


Fig. 1: Home page of *Papyri.info*, with the two components – the PN and the PE – highlighted. <https://web.archive.org/web/20170710143913/http://papyri.info>.

2 This resource is composed of two core elements: the *Papyrological Navigator* (PN), a tool to browse and search the texts (on which I will therefore focus my review), and the *Papyrological Editor* (PE), which enables users to contribute to the collection (fig. 1). The content, i.e., the text of the papyri and their metadata, is mostly obtained from the aggregation of pre-existing databases, but also from users' contribution: anyone can sign up and enter new texts and metadata, or edit those already existing, with a procedure that includes an editorial vetting in order to meet scholarly standards.²

3 The collection is a general purpose one, based on the principles of completeness and of representativeness, within its domain of documentary papyrology. In fact, it has become a point of reference for documentary papyrology, as it aggregates three major databases of documentary papyri: the *Duke Databank of Documentary Papyri* (DDbDP), a very large collection of texts of documentary papyri realised by the Duke University; the *Heidelberger Gesamtverzeichnis der griechischen Papyruskunden Ägyptens* (HGV), created at the University of Heidelberg, which provides the metadata of all documentary papyri; and the University of Michigan *Advanced Papyrological Information System* (APIS), focused on images, as well as metadata, of papyri from the North American institutions, plus a few small federated collections worldwide³; also important is the integration of an essential tool for papyrological research, namely its main

bibliographical database, the *Bibliographie Papyrologique* (BP). The aims of completeness and representativeness are not explicitly stated on the *Papyri.info* website. We find a mention of the collection content on the home page, but this presents no information on its extension: "The Papyrological Navigator (PN) supports searching, browsing, and aggregation of ancient papyrological documents and related materials". Notwithstanding, we can infer that completeness is a principle of the collection from the words of one of the founders of the DDbDP, John F. Oates (1993, 62): "We are aiming to construct a machine-readable data bank comprising all published Greek and Latin documentary papyri". Also, the contributions published by William H. Willis, co-founder of the DDbDP with Oates, on the plan of the entering of the texts and on the updates on this progress demonstrate the intention that the collection may be as complete as possible: firstly they digitised the papyri published in the more recent volumes of papyrological collections, secondly the remaining volumes "in reverse chronological order until all are done", thirdly "all texts published separately in journals", and lastly the most significant corrections gathered in the dedicated papyrological tool, the *Berichtigungsliste* (Willis 1984, 172; updates on the progress of the data entry are provided in Willis 1988, 1992, 1994). The collection currently amounts to 55,080 texts, which have been merged from the DDbDP, each of them provided with metadata exported from the HGV and also, as to papyri from American collections, from the APIS. In addition to these records with the papyrus texts, there are 31,936 records with metadata only, exported from the HGV and APIS; these records pertain to published papyri whose texts have not been digitised yet, or to unpublished papyri whose metadata and images have been put at disposal by the APIS researchers to enable the scholarly community to study them.⁴ Finally, as just mentioned, there are images of some papyri, exported from the APIS database, and therefore pertaining to fragments mostly housed in American institutions.⁵

4 The collection is meant to be used in research, especially by papyrologists or by ancient historians who need to consult papyrological primary and secondary sources. Besides the fields of papyrology and ancient history, the collection contributes to historical linguistics, helping scholars to examine the features of the language of the papyri – texts that constitute an important body of evidence for Greek and Latin in their postclassical phases, and for their contacts with other languages of the Mediterranean.⁶ *Papyri.info* intends to meet scholarly standards by submitting the texts and the changes entered by users to an editorial vetting, and by reporting the emendations and the

supplements (both already published and newly proposed by users) of the text transcriptions.

Participating institutions and staff, and their bibliographic identification

The screenshot shows the Papyri.info homepage. At the top, there's a red header bar with the site name 'Papyri.info' and a 'sign in' link. Below it is a grey navigation bar with links for 'Browse: DDbDP HGV APIS TM Number' and 'Search: Data Bibliography'. The main content area has two sections: 'Partners' (listing APIS, DDbDP, HGV, BP, Trismegistos, and APD) and 'More information' (links to a Digital Papyrology blog, Checklist of Editions, papyrological resources, and a feedback form). A large orange oval highlights the 'Produced by' section at the bottom right, which states 'The Duke Collaboratory for Classics Computing & the Institute for the Study of the Ancient World'. The footer contains links for Home, Help, Blog, Resources, Feedback, and About EpiDoc, along with a copyright notice for 2013.

Fig. 2: *Papyri.info*, home page. <https://web.archive.org/web/20170710143913/http://papyri.info>.

This screenshot shows the 'About papyri.info' page. It begins with a 'Prototype' section detailing the development of the Papyrological Navigator (PN) from 2006 to 2013, mentioning the Andrew W. Mellon Foundation and the National Endowment for the Humanities. It then describes the 'Release' in 2010, the 'Stewardship' by the Duke Collaboratory for Classics Computing (DC3), and the 'Moving Parts' architecture, which includes a RDF triple store, Apache Solr search, and a Git-based version control system. The footer is identical to Fig. 2, with the 'Produced by' note circled in orange.

Fig. 3: *Papyri.info*, "About" page. <https://web.archive.org/web/20170603015536/http://www.papyri.info/docs/about>.

5 Several institutions were involved in the development of Papyri.info. The Duke Collaboratory for Classics Computing (DC3) of the Duke University, to which Ryan Baumann, Hugh Cayless and Josh Sosin belong, and the Institute for the Study of the Ancient World (ISAW), in particular Roger Bagnall (formerly based at the Columbia University) are mentioned in this resource as the producers, on the home page and on the "About" page ([fig. 2](#) and [3](#)).

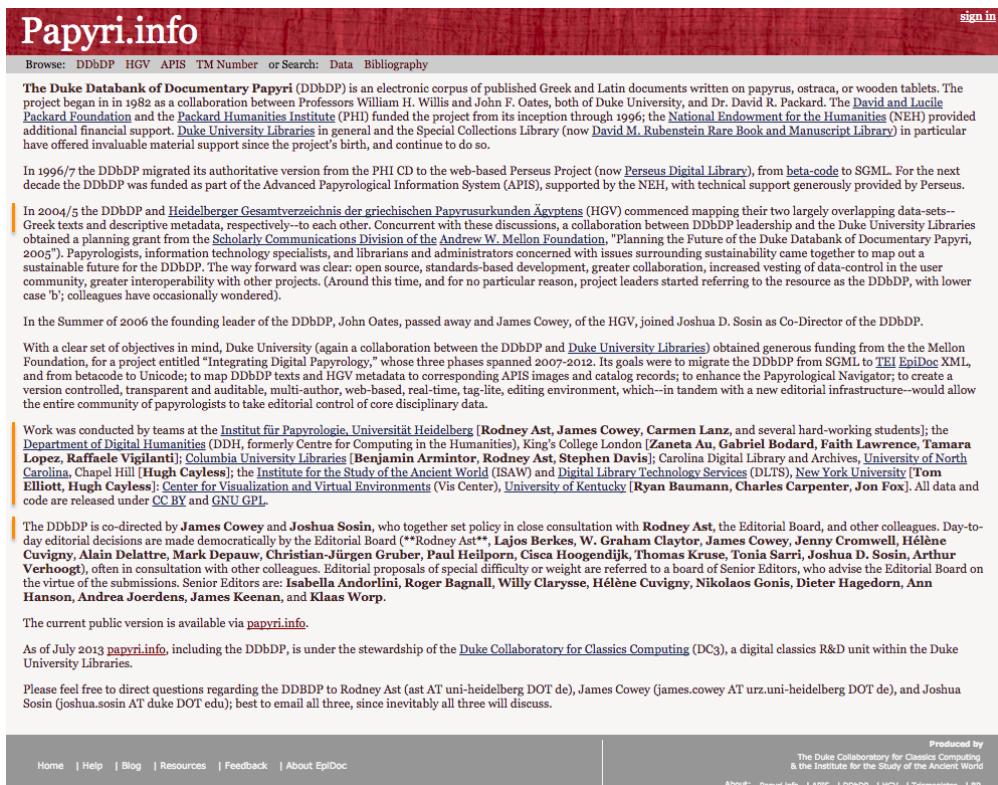


Fig. 4: *Papyri.info*, "The DDbDP". <https://web.archive.org/web/20170925122526/http://papyri.info/docs/ddbdp>.

6 As well as these, more institutions were responsible for the development of this resource. Their role is however acknowledged only on the page about the *Duke Databank of Documentary Papyri* (DDbDP, [fig. 4](#)); as their names do not appear on the home page, nor on the "About" page, they are therefore little visible.

7 The then Centre for Computing in the Humanities, now Department of Digital Humanities, at King's College London, was in charge of the first round of conversion from Beta Code into the Unicode character encoding, and from SGML into the EpiDoc XML standard (Bodard, Sosin, and Viglianti 2011). The Heidelberg Institut für Papyrologie mapped its HGV database with the other databases of papyri aggregated into *Papyri.info*, namely the DDbDP and APIS⁷, and, also currently, works for continuing the

integration of the HGV metadata of new papyri into the PN⁸; also, James Cowey in Heidelberg is co-director of the DDbDP together with Joshua Sosin, as reported on the information page of the DDbDP ([fig. 4](#)). All this information reported on the DDbDP page is also very relevant to *Papyri.info* as a whole, as the texts, which it contributed, constitute the core of the PN. Furthermore, this page contains details on the editorial board who approves users' contributions, and it thus concerns the current state of the collection as well. As King's College London and the Institut für Papyrologie, too, gave a contribution to *Papyri.info*, and the latter still continues to collaborate, their role should be put in more evidence, and be as visible as the mention of the DC3 and the ISAW. Hence, King's College London and the Institut für Papyrologie, too, should occur on the home page and in the history of the development of *Papyri.info* on the "About" page ([fig. 3](#)), and at least the HGV should be reported on the home page under the heading "Produced by" (at the bottom, on the right: see [fig. 2](#)). Other involved institutions mentioned on the "DDbDP" page, which should occur on the "About" page of *Papyri.info*, are the Columbia University Libraries, the University of North Carolina, the University of Kentucky and the New York University. Alternatively, the two pages about *Papyri.info* and the DDbDP could be merged into a single page.

8 Funding was provided by the Mellon Foundation and the National Endowment for the Humanities (NEH), as reported on the "About" page. Furthermore, from the "DDbDP" page we learn that financial support was granted to the *Duke Databank* by the Packard Foundation and the Packard Humanities Institute (PHI). As already noted, since *Papyri.info* has its roots in the DDbDP and therefore their histories are closely connected, it would be fair to mention the Packard funding in the "About papyri.info" page, as well: without the initial support from the Packard institutions, who also gave technical assistance, many of the texts that now populate the database could not have been entered.⁹

9 As for the long-term archiving of the data and institutional curation, information is provided, on the "About" page (in the "Stewardship" section) about the Duke Collaboratory for Classics Computing, based at the Duke University Library, reported as the unit who has the task of the "maintenance and enhancement of the papyri.info toolset and community". More details on the way the sustainability of *Papyri.info* is pursued are available in the related bibliography: open source development, through the use of a Git repository to store and update the data (*Papyri.info IDP (Integrating Digital Papyrology) Data*, <https://web.archive.org/web/20180202152933/https://github.com/papyri/idp.data>;

standards-based formats, i.e., the Unicode character encoding, EpiDoc-conformant TEI markup, and the Resource Description Framework (RDF) model for the integration of the databases in the context of the Semantic Web.¹⁰

10 In conclusion, information about the creators and the participating institutions that provided expertise and funding is presented; on the other hand, it needs to be made more discoverable by integrating it into the information about the whole resource, rather than being referred to just in connection with one of its compounds. More details on the long-term data preservation, which are found in the bibliography on the project, could be added to the "About" page for any user interested in following up in the reasons of the choices made by the developers; as an alternative, users could be provided with the bibliographical references where this information can be found.

The integration of multiple papyrological resources

The different kinds of source databases

11 The Papyrological Navigator (PN) is formed by the aggregation of three pre-existing databases (as anticipated): while the *Duke Databank of Documentary Papyri* (DDbDP), which provided the texts of the papyri, is now merged into the PN, the two other resources that have been integrated into it, namely the *Heidelberger Gesamtverzeichnis der griechischen Papyruskunden Ägyptens* (HGV) and the *Advanced Papyrological Information System* (APIS), continue to exist independently. After the integration process of the HGV and APIS, updates in the HGV (i.e., metadata of newly published papyri or corrections to the metadata already existing) continue to be entered into the PN. James Cowey (Institut für Papyrologie of the University of Heidelberg), who is both editor of the HGV and co-director of the DDbDP, is the person responsible for updating the HGV and for exporting this data into GitHub, with the technical assistance of Carmen Lanz (of the Institut für Papyrologie as well). They report the changes to the researchers of the Duke Collaboratory for Classics Computing (Ryan Baumann, Hugh Cayless and Joshua Sosin), who are in charge for the maintenance of *Papyri.info*, so that they can merge the updates from GitHub into the *Papyri.info* branch. Since the data of the HGV are maintained and updated in a FileMaker database, whereas *Papyri.info* is an XML based project, the XML export possibility of FileMaker is used to crosswalk the data to XML, before adding them to GitHub.¹¹

Papyri.info

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

Refine Search New Search Search

Date on or after: 200 CE X

49269 hits. Records per

Identifier	Title	Location	Date
no. 14297	Receipt	Alexandria?	300 CE - 399 CE
basp 49 135	Fragment ...	Koptos	297 CE
O.Berenike 2 260	Lamp ... This lamp was discovered in the 1998 season.		500 CE - 599 CE
P.Tebt. 2 267	List ...		175 CE - 200 CE
P.Tebt. 2 268	Trojan ...		200 CE - 225 CE
P.Tebt. 2 270	Grammar		200 CE - 299 CE
P.Tebt. 2 271	Account		100 CE - 299 CE

Identifier Title Location Date

Series --- All values --- Vol.
or

Collection --- All values --- ID #

Provenance --- All values ---

Name --- All values ---

Date on or after 200 CE (49269) BCE CE
Date before --- All values --- BCE CE

Loose All values ---
Language 250 CE (19414)
300 CE (22211)
350 CE (28330)
400 CE (29258)
450 CE (31206)
500 CE (31578)
550 CE (39898)
600 CE (40739)
650 CE (46052)
700 CE (46390)
Has Transcription 750 CE (48585)

New publications Go

Fig. 5: Search form of *Papyri.info*. The fields for the search of a date are highlighted.

Fig. 6: Search form of the HGV; highlighted, the numerous date fields, which allow a detailed search. <https://web.archive.org/web/20171002153018/http://aquila.zaw.uni-heidelberg.de/search>.

12 As opposed to this, updates in the APIS are unfortunately not visible in the PN.¹² The fact that the PN receives updates from the HGV implies that we obtain the same results if we perform the same search on the two databases. However, more precisely, the number of hits of the HGV can be slightly higher, because a text may have been reported on two or three different records if its dating is uncertain or if a papyrus contains more than one date (the latter happens when, after being reused, a papyrus is composed by separate texts copied in different moments); in this case, the PN instead records the dates, or a time span, in the "Date" field of the papyrus record.¹³ Furthermore, if the same fragment contains two or more texts written by different scribes, whether with a date or not, the HGV, again, treats them as distinct items and uses a record for each of them, whereas the PN presents the texts in only one record.¹⁴ An opposite treatment is instead assigned to papyri that preserve two or more different texts, which are however written by the same hand and published together in the same edition: the PN considers them as a single item, assigning them to only one record, while the HGV stores each text in a

different record.¹⁵ One should then resort to the PN database rather than to the HGV, if one intends to obtain exact numerical data for statistical purposes. On the other hand, users may wish to resort to the HGV in some cases, as its user interface allows more possibilities of search and more specific searches within the dating field (see below).

The screenshot shows a detailed record page for a papyrus fragment. At the top, there's a navigation bar with links like 'Browse', 'DDBDP', 'HGV', 'APIS', 'TM Number', and 'Or Search'. Below the title 'p.cair.zen.1.59001 = HCV P.Cair. Zen. 1 59001 = Trismegistos 663 = oxford-ipap.apis.484 = psi.4.321' are links for 'Previous record' and 'Next record'. A red box highlights the 'TM data' section, which contains links to various databases: 'metadata', 'HGV data', 'TM data', 'APIS catalog record', 'text', 'transcription', 'open in editor', and 'Canonical URI'. The main content area is divided into sections: 'Publications', 'Post-Concordance BL Entries', 'Origin', 'Provenance', 'Material', 'Date', 'Print Illustrations', 'Subjects', 'Images', 'License', 'Trismegistos: 663 [source]', 'Catalog Record: oxford-ipap.apis.484 [xml]', and 'related resources'. The 'Trismegistos: 663 [source]' section is also highlighted with a red box. It lists publications, date ranges (274-273 BCE), subjects (Ptolemaic Egypt, Memphis), and provenance (Kharabat el Gerza). The 'Catalog Record: oxford-ipap.apis.484 [xml]' section provides a detailed description of the fragment, mentioning its size (17.8 x 6.3 cm), material (papyrus), and context (loan of money). The 'related resources' section includes a link to the GLRT.

Fig. 7: A record in the PN (P.Cair.Zen. I 59001, a loan of money). Highlighted, the TM section with its relative links to the TM databases of collections, archives, people and places.

<http://papyri.info/ddbdp/p.cair.zen;1;59001?>

http://papyri.info/ddbdp/p.cair.zen;1;59001?rows=2&start=0&fl=id%2Ctitle&fq=series_led_path%3Ap.cair.zen%3B1%3B*&sort=series+asc%2Cvolume+asc%2Citem+asc&p=0&t=124, accessed 24 October 2017.

13 The integration with the *Trismegistos* database (a platform with metadata of ancient texts, in a wide variety of languages and on different supports) allows users to access this resource through links in PN records. One can thus gain further information on the place of provenance of a papyrus, on the collection that currently houses it, on the people and the places mentioned in the text, and on the possible ancient archive the fragment belonged to originally.

Papyri.info

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

p.princ.2.97 = HGV P.Princ. 2 97 = Trismegistos 17374 = princeton.apis.p481
 << Previous record Next record >>

[metadata](#) [HGV data](#) [TM data](#) [APIS catalog record](#) [text](#) [transcription](#) [images](#) [translation](#) [open in editor](#) Canonical URI: <http://papyri.info/ddbdp/p.princ;2;97>

HGV: P.Princ. 2 97 [source] [xml]

Title Letter
 Publications P.Princ. 2 97 [More in series P.Princ.] [More in series P.Princ., vol. 2]
 Support/Dimensions Papyrus
 Post-Concordance BL Entries XII 164
 Origin unbekannt [More from unbekannt]
 Material Papyrus
 Date 326 - 327 [More from the period between 326 CE and 327 CE]
 Commentary Vgl. Comunicazioni 1997, S. 54.
 Mentioned Dates Z. 121:25 - 326
 Print Illustrations Plate IX
 Subjects Brief (privat); N.N. an Horion; Aufträge; Ausfertigung von Quittungen; Zahlung von Steuern
 Images <http://wwwapp.cc.columbia.edu/dpd/app/apis/item?mode=item&k...>
 License © Heidelberg Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens. This work is licensed under a [Creative Commons Attribution 4.0 License](#).

Trismegistos: 17374 [source]

Publications P. Princ. 2 97
 Inv. no. Princeton, University Library Garrett deposit 7619
 Date AD 326 - 327 (Diocletians), year (43)
 Language Greek
 Provenance Egypt [found & written]
 People mentioned people
 Places mentioned places

Catalog Record: princeton.apis.p481 [xml]

Title Letter
 Summary Letter in which the writer gives instructions to the recipient to issue receipts to certain farmers and vine-dressers who have delivered the required number of pigs in settlement of claims against them. He is instructed to further credit these deliveries on his accounts. In most cases one pig has been delivered. Deliveries are credited for the years 325/326 and 326/3227. The letter also contains an instruction for the settlement of taxes on the estate under the supervision of the addressee
 Citations P.Princ. II 97
 P.Princ. 2:97
 Inv. Id GD 7619
 Support/Dimensions 1 papyrus fragment; 23.5 x 9.5 cm
 Origin Provenance Unknown [More from Provenance Unknown]
 Language In Greek
 Date 326/327 AD [More from the period between 326 CE and 327 CE]
 Note (general) Written across the fibers; on the recto traces of an account
 Note (general) 30 lines; one line missing at the top
 Note (general) The text is faded in the center and at the right
 Note (general) The hand is a fluent cursive
 Subjects Papyr.; Letter
 License © Heidelberg Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens. This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 License](#).

Citations 66565. Nikolaos GONIS, Notes on Some Private Letters..

DDbDP transcription: p.princ.2.97 [xml]

AD26/7?

v Όρισμα []ττ. οι[...] , [άδελ-]
 φω χαίρεν.
 τῶν κοῖρων ἀπενεχθέντων
 υπό τῶν γεωργῶν καὶ αὐτοῖς-]
 λουργῶν τοῦ ἐποκιού, φρόν-]
 τιου λαβέντων τὰ γράμματα
 τα ἔκδος ἀποχώρα γράμμα-
 τα τοῖς ξῆς ίνα [επί]

Image [open in new window]

Fig. 8: Example of a record in the PN with an image of the papyrus (P.Princ. II 97, letter),

imported from APIS. <http://papyri.info/ddbdp/p.princ;2;97?rows=3&start=2828&fl=id%2Ctitle&fq=%28images-int%3Atrue%29&sort=series+asc%2Cvolume+asc%2Citem+asc&p=2829&t=3010>, accessed

24 October 2017.

14 In the records of the PN we can thus find texts, metadata, and images or links to images (provided that images are available on APIS or other databases, usually those of the physical collections).

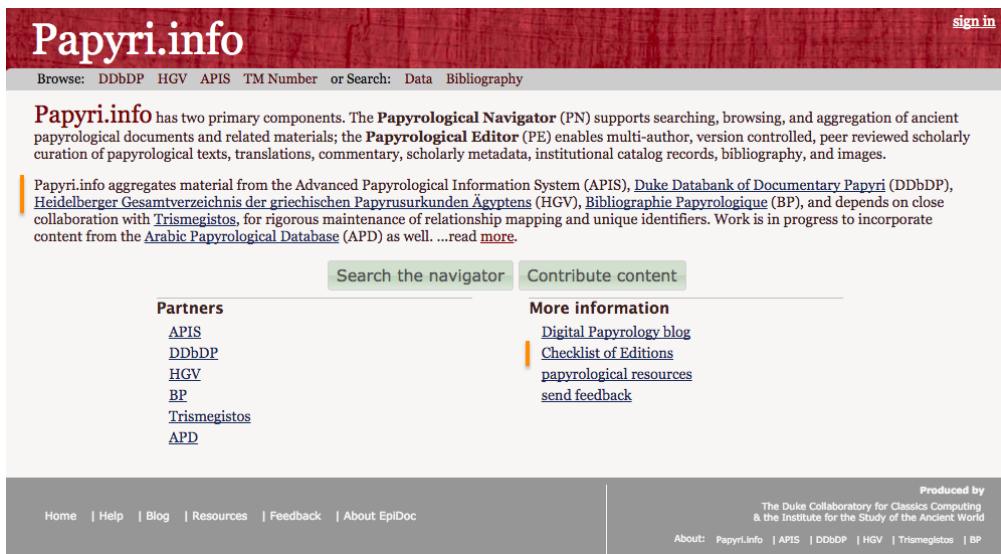


Fig. 9: Home page of *Papyri.info*, with its aggregated databases highlighted: <https://web.archive.org/web/20170710143913/http://papyri.info>.

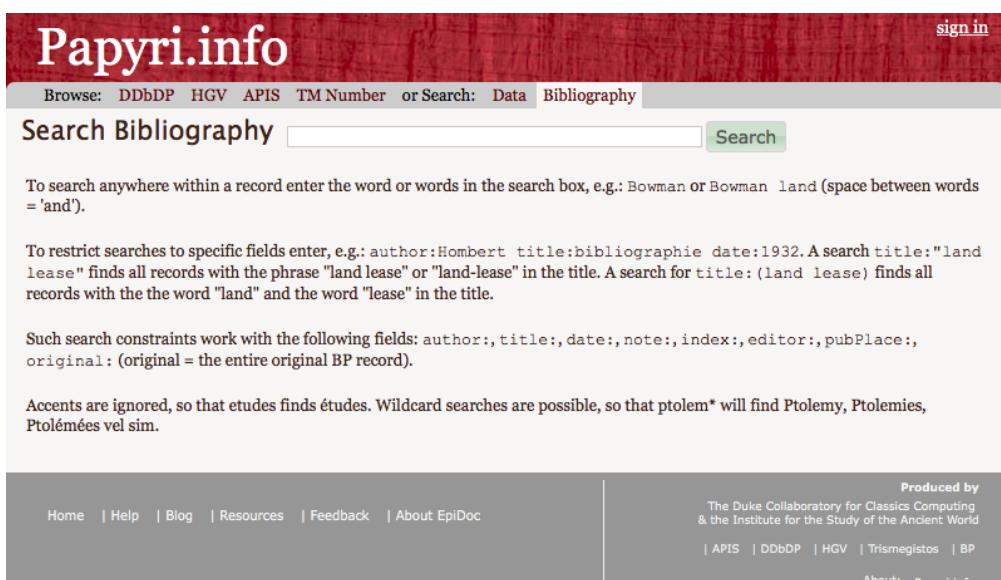


Fig. 10: "Search Bibliography" form: <https://web.archive.org/web/20170603030657/http://www.papyri.info/bibliosearch>.

15 In addition to this data, the PN includes the most important database of bibliographical information for papyri, the *Bibliographie Papyrologique* (BP). It should be noted, however, that this source is not clearly stated: the aggregation of the BP, though pointed out on the home page within the information about the related projects (fig. 9), is however not included in the "Search Bibliography" form (fig. 10). Given that the bibliographical database is indicated as "Bibliography" both in the "Search Bibliography" form and in the menu bar, there is no way for the user to know that it is formed by the BP except reading the information about *Papyri.info* on the home page; they may however

skip this step to go directly to the "Search Bibliography" form or to the "Search" form. A user I interviewed for my research, who stated that he prefers to consult the papyrological bibliography on *Papyri.info* rather than on the BP, was in fact not aware of the relationship between the two databases. Regarding this bibliographical database, users should also be informed that the BP records later than 2012 are not available in the PN, as one can only learn from the *Bibliographie Papyrologique* website.¹⁶ This issue is indicative of the advantages, but also of the risks, entailed by the integration of different sources into one platform. On one hand, it is possible to create links between the record of a papyrus and the bibliographical record, to allow adding new bibliographical information or editing the existing records, and creating stable URLs to indicate a bibliographical entry – features that are not possible on the BP database.¹⁷ On the other hand, there is a need for transparency for the provenance of the data: the source databases should be clearly indicated and regularly updated, or at least their updating status should be stated.

16 Finally, the *Papyrological Navigator* includes a *Checklist of Editions*, a list of papyrological editions and instruments also used as a reference for standard abbreviations. The presence of the *Checklist*, however, needs to be highlighted more on the home page: as a research tool, it should be distinguished from the other, as much as useful, resources indicated in the "More information" column of the home page, namely the *Digital Papyrology blog*, a list of papyrological resources, and indications for sending feedback ([fig. 9](#)). The link to the Checklist could be positioned next to the "Bibliography", in the menu bar on top. Also, the presence of the *Checklist* should be mentioned in the information about the material aggregated into *Papyri.info*, at least on the "About" page¹⁸, and possibly on the home page as well.

17 Regarding the mentioned *Digital Papyrology blog*, since its last update dates back to 2012, it seems more appropriate to include its link on the *Papyri.info* "About" page, indicating that this blog provides further details on the development of the resource: most posts concern updates on the functionalities of *Papyri.info* and the related TM database, and thus constitute a source of information for the history of the project.¹⁹

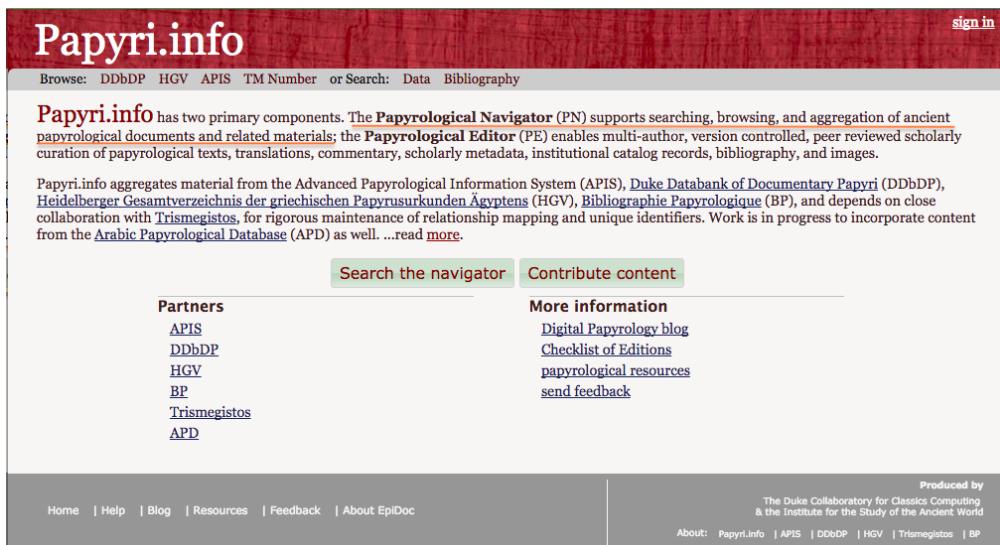


Fig. 11: Home page of *Papyri.info*, with the information about its content highlighted: <https://web.archive.org/web/20180110131138/http://papyri.info>.

18 To sum up, the material merged, integrated, or linked with the *Papyrological Navigator* provides a very extensive collection of texts of the chosen kind, i.e., documentary papyri, along with comprehensive information and aids for papyrological research. However, it would be necessary to inform users in a clearer and more precise way about the characteristics of its material, as noted above about the source of the "Bibliography" and the status of its update, and the presence of the *Checklist*. But, even more important, a key information that is not provided concerns the kind of papyri included in the collection, as there is no clear indication that *Papyri.info* concerns *documentary papyri* and not literary ones. The home page contains the information "The Papyrological Navigator (PN) supports searching, browsing, and aggregation of ancient papyrological documents and related materials" (fig. 11); however, the phrase "ancient papyrological documents and related materials" rather than "documentary papyri" may indicate, more in general, all kinds of papyri, including literary and paraliterary ones. Since the name of the resource does not contain a reference to the documentary character of the collection either, this essential information ought to be stated at the very beginning of the home page.²⁰

Issues in the collection content

The screenshot shows the homepage of the DCLP (Duke Collaborative for Classics Computing) project. The header features the text "DCLP" in large letters, followed by "Editor not available." A navigation bar includes links for "Browse: DDbDP HGV APIS DCLP", "browse by series", "browse by TM number", "browse by authors and works", "TM Number", and "Search: Data Bibliography". Below this, a paragraph describes the project's goal of peer-reviewed curation of ancient literature preserved on papyri, mentioning partners like LDAB, BP, and Trismegistos. A "Partners" section lists API, DDbDP, HGV, BP, Trismegistos, APD, and Anagnosis. A "More information" section links to the Digital Papyrology blog, Checklist of Editions, papyrological resources, and a feedback form. The footer contains links for Home, Help, Blog, Resources, Feedback, and About EpiDoc, along with a "Produced by" note for The Duke Collaboratory for Classics Computing & the Institute for the Study of the Ancient World, and a "About" link for Papyri.info, APIS, DDbDP, HGV, Trismegistos, and BP.

Fig. 12: The DCLP home page, which shows the same design as *Papyri.info*. [https://web.archive.org/web/20171025150016/http://ltpap.info>](https://web.archive.org/web/20171025150016/http://ltpap.info) with the results of a search for [magic papyri](#).

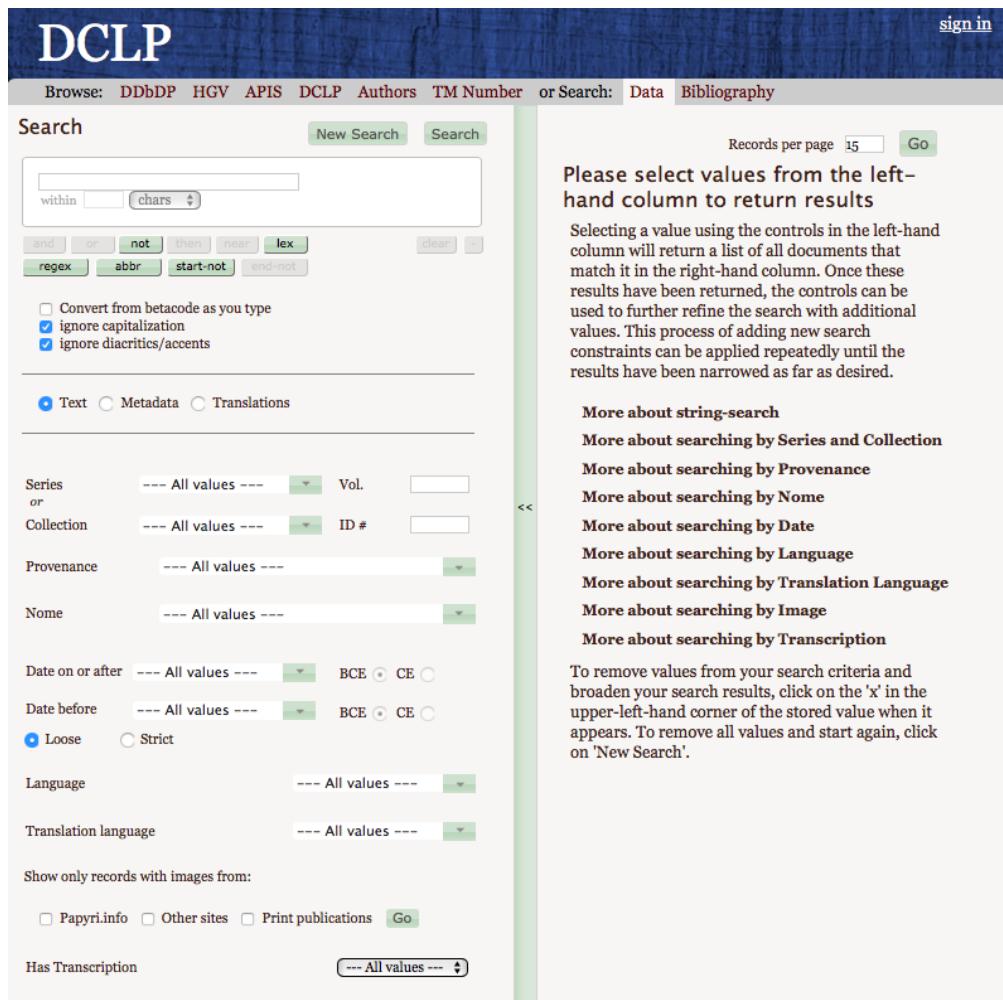


Fig. 13: The *Papyrological Navigator* of the DCLP. <https://web.archive.org/web/20170107212926/http://www.litpap.info:80/search>.

19 Searches among the metadata show that there are also a few non-documentary papyri, i.e., literary, medical and magical texts (they can be found by entering "literary", "magical" and "medical" in the Search form and selecting the "Metadata" option), even though these categories are now being systematically published in a separate corpus, the *Digital Corpus of Literary Papyri* (DCLP) developed at the University of Heidelberg.²¹ It seems that, in part, these extra texts belong to one of the source databanks included in *Papyri.info*, i.e., the *Advanced Papyrological Information System* (APIS), since many of them are housed in American collections. As to the medical papyri, they were first included into the *Papyrological Navigator* (within the *Digital Corpus of the Greek Medical Papyri* project of the University of Parma²²) and only in a second moment they have been exported to the DCLP, where they now are being added;²³ the first digitised medical papyri seem therefore to have remained in the PN. It would be opportune to remove these texts, which are duplicates of DCLP records, and whose presence can affect the results of searches in *Papyri.info* made for statistical purposes. At least, their

existence in *Papyri.info* should be pointed out to users, who should also be suggested to search the DCLP for specific queries among those categories, for more complete results. Finally, it would be at any rate useful to add the information about the DCLP on the *Papyri.info* home page, since they are two separate but similar projects, both with regards to content and technical implementation.



Fig. 14: The "Browse" menu bar on the home page (also available on the "Search" page). <https://web.archive.org/web/20161203142342/http://papyri.info/search>.



Fig. 15: Browsing the DDbDP: list of the digitised papyrological series and collections. <https://web.archive.org/web/20161201145231/http://papyri.info:80/browse/ddbdp>.

20 While the content types are not precisely stated, there is a clear classification of the content from the point of view of the print series and collections in which the texts have been published: one can find them by browsing the three aggregated databases (DDbDP, HGV and APIS) from the menu bar (figg. 14, 15). However, this classification, which is useful to find a papyrus by its edition abbreviation, cannot be used to find types of content, as these cannot be inferred from the publication series. It would be helpful to add, in the search form, a "content type" or "subject" field, by which to search the "subject" field present in the papyri records (derived from the "Inhalt" field of the HGV source database, as discussed below).

Usage and reuse scenarios of the collection

Usage for word-searches

21 When it was created, the goal of the collection was to speed word searches in order to find textual parallels and documents of the same type more quickly: this helps scholars to propose supplements, carry out a linguistic analysis, and study documents according to their different typologies. Such objectives, set out by the founders of the *Duke Databank*, John Oates and William Willis,²⁴ still today represent the main type of usage of the collection: the users I have interviewed for my research so far²⁵ have mentioned these purposes as the reasons why they avail themselves of the *Papyrological Navigator* in their everyday scholarly activities.

Papyri.info

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

Refine Search New Search Search

within chars

and or not then near lex clear

regex abbr start-not end-not

Convert from betacode as you type
 ignore capitalization
 ignore diacritics/accents

Text Metadata Translations

Series --- All values --- Vol.
 Collection --- All values --- ID #
 Provenance --- All values ---
 Nome --- All values ---
 Date on or after --- All values --- BCE CE
 Date before --- All values --- BCE CE
 Loose Strict
 Language --- All values ---
 Translation language --- All values ---
 Show only records with images from:
 Papyri.info Other sites Print publications
 Has Transcription

Substring: #γι°
 Target: text
 No Caps: On
 No Marks: On

2434 hits.

Identifier	Title	Location	Date
bgu 19	Revenue ...	Ptolemais Euergetis (Arsinoites)	276 CE
bgu 121	keiner	Prektis (Hermopolites)	340 CE
7.	γι(νεται) όμοι [. . .] λήμ(ματα) (τάλαντα) [Ζ]τλζ .		101 CE
bgu 183	keiner	unbekannt	- 300 CE
4.	αὐλ(ῆς) γι(τονες) (*) νό(του) βύμη βασιλ(ική),		185 CE
bgu 1186	keiner	Arsinoites	- 186 CE
16.	γι(τονες) (*) νότ(ου) και λιβός [-ca.? -]		101 CE
bgu 1217	keiner	Arsinoites	- 300 CE
18.	ό δὲ [Τού]ρβων (ἄρουραι) ι, και Στοτοῆτ(ις) (ἄρουρ		
bgu 1304	keiner	Herakleopolites	647 CE
14.	γι(νονται) σί(του) (ἀρτάβα) κ, και ἔτοιμως		601

Fig. 16: The results of the search for an abbreviation with the PN, performed by typing the letters that have not been abbreviated by the scribe (here, γι); then, clicking on the "abbr" button allows limiting the search to abbreviated words, as indicated by the small superscript circle. Additionally, to look up the abbreviation γι() for γι(νεται) or γι(νονται), which occurs in accounts with the meaning of "equal", I allowed a hashtag to precede the letters, so as to furtherly narrow down the search to the word-initial position (as explained in the directions, in "More about string-search").

22 Other kinds of searches for which the collection can be used, as users point out, are those of scribal mistakes and abbreviations. The ability to search for mistakes is useful to check the validity of a conjecture when one suspects that, in one's own editorial work with documentary texts, a damaged line contained, in the gap, a misspelled word: since the original form employed by the scribe is always reported in the texts in *Papyri.info* (as well as the normalised one), the user can look up the incorrect word they have conjectured, to verify whether it is a mistake commonly attested in papyri. As regards the search for abbreviations, this is useful to understand how an abbreviation found in one's text could be expanded; this search is possible because both the

abbreviated word reported by the scribe and the expanded form restored by the editor are indicated in the texts in *Papyri.info*, and because the search form includes the possibility to specifically search for abbreviations ([fig. 16](#)).

The EpiDoc encoding standard

23 Another function, which instead could not have been foreseen in the age when the *Duke Databank* was founded, is to allow the reuse of its data for new purposes, enabled by the introduction of standards-based encoding formats. Namely, the PN database also works as a textual repository, in that it can provide text data for further collection-external research. In fact, its texts are encoded in the standard EpiDoc markup, a set of guidelines for the digital representation of ancient texts of inscriptions and papyri, based on TEI elements and attributes, which has been employed in a variety of projects.²⁶ The texts have been converted to EpiDoc when the source databanks were integrated into *Papyri.info*, in the *Integrating Digital Papyrology* project.²⁷

24 The EpiDoc model has been created to represent in digital form the texts that are traditionally transcribed according to the 1932 Leiden editorial conventions.²⁸ These texts, inscriptions and papyri, are ancient documents of direct tradition, i.e., discovered in archaeological excavations rather than handed down to us through medieval tradition. This implies the advantage of their chronological closeness to the ancient historical events and literary works, but also the problem of their severely damaged state of preservation. Hence the necessity to develop such a system as the Leiden Convention: this allows accounting for the physical appearance of the texts and for the editorial decisions taken in the restoration of the missing words and in the interpretation of the signs employed by the ancient scribes. Examples from the Leiden transcription style are the square brackets that surround the letters lost in a lacuna and restored by the editor, the dots under the line to indicate letters of uncertain reading, and round brackets to expand a word written in abbreviation by the scribe.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <?xml-model href="http://www.stoa.org/epidoc/schema/latest/tei-epidoc.rng" schematypens="http://relaxng.org/ns/structure/1.0"?>
3  <?xml-model href="http://www.stoa.org/epidoc/schema/latest/tei-epidoc.rng" schematypens="http://purl.oclc.org/dsdl/schematron"?>
4  <TEI xmlns="http://www.tei-c.org/ns/1.0" xml:space="preserve" xml:lang="en">
5    <teiHeader>
6      <fileDesc>
7        <title>title of document</title>
8        </titleStmt>
9        <publicationStmt>
10       <authority></authority>
11       <idno type="filename"></idno>
12       </publicationStmt>
13     <sourceDesc>
14       <msDesc>
15         <msIdentifier>
16           <repository>museum/archive</repository>
17           <idno>inventory number</idno>
18           </msIdentifier>
19         <physDesc>
20           <objectDesc>
21             <supportDesc>
22               <support>description of object/monument (likely to include <material/>
23                 and <objectType/> information, <dimensions/>, etc.)</support>
24             </supportDesc>
25             <layoutDesc>
26               <layout>description of text field/campus</layout>
27             </layoutDesc>
28           </objectDesc>
29         <handDesc>
30           <handNote>description of letters, possibly including <height>letter-heights</height>
31         </handDesc>
32       </physDesc>
33       <history>
34         <origin>
35           <origPlace>Place of origin</origPlace>
36           <origDate>Date of origin</origDate>
37         </origin>
38         <provenance type="found"> Findspot and circumstances/context
39         </provenance>
40         <provenance type="observed">Modern location(s) (if different from repository, above)
41         </provenance>
42       </history>
43     </msDesc>
44   </sourceDesc>
45 </fileDesc>
46 </teiHeader>
47 <facsimile>
48   <graphic url="photograph of text or monument"/>
49 </facsimile>
50 <text>
51   <body>
52     <div type="edition" xml:space="preserve">
53       <ab>
54         <lb n="1"/>Greek or Latin (etc.) text here
55       </ab>
56     </div>
57     <div type="apparatus">
58       <p>external apparatus criticus (if applicable)</p>
59     </div>
60     <div type="translation">
61       <p>translation(s)</p>
62     </div>
63     <div type="commentary">
64       <p>commentary</p>
65     </div>
66     <div type="bibliography">
67       <p>bibliography of previous editions, discussion, etc.</p>
68     </div>
69   </body>
70 </text>
71 </TEI>
72
73
74

```

Fig. 17: Basic structure of an EpiDoc file. [https://web.archive.org/web/20170710143913/
http://papyri.info.](https://web.archive.org/web/20170710143913/http://papyri.info.)

25 By customising the TEI guidelines for the needs of epigraphists and papyrologists, the EpiDoc markup aims at representing the information conveyed by the Leiden signs, as well as at encoding the features of the history and of the materiality of the object. For encoding the item description, the `<teiHeader>` element of an EpiDoc file contains a `<sourceDesc>` element, which marks information about the institution that houses the papyrus or inscription, the inventory number, the support that bears the writing, the layout, the hand, the place of origin and the date (see fig. 17 for the template with the basic structure of an EpiDoc file; this is downloadable from <https://sourceforge.net/p/epidoc/wiki/Examples/> (accessed 11 January 2018). After the `<teiHeader>`, the `<text>` element contains `<div>` elements for the different parts of the edition: text (enclosed by a further element, `<ab>`), apparatus, translation, commentary and bibliography.

Papyri.info

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

p.warr.2 = HGV P.Warr. 2 = Trismegistos 13701 = leidenpapinst.apis.0002
<< Previous record Next record >>

[metadata](#) [HGV data](#) [TM data](#) [APIS catalog record](#) [text](#) [transcription](#) [images](#) [translation](#) [Open in editor](#) Canonical URI: <http://papyri.info/ddbdp/p.warr.2>

HGV: P.Warr. 2 ([source](#)) [[xml](#)]

Title Notice of birth
Publications P.Warr. 2 [[More in series P.Warr.](#)]
Support/Dimensions Papyrys
Post-Concordance BL Entries VIII 196
Origin Arsinoites [[More from Arsinoites](#)]
Provenance Arsinoites Ägypten
Material Papyrus
Date 11. Juli 72 [[More from the period between 72 CE and 73 CE](#)]
Mentioned Dates Z. 12:59 - 70
Subjects Anzeige (Geburt); Namen
License © Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens. This work is licensed under a [Creative Commons Attribution 3.0 License](#).

DDbDP transcription: p.warr.2 [xml]

AD72 Arsinoite

1 [] [- ca.10 - ἀμφ6-]
[δοῦ] Αρέβιο[:: ca.13 -]
παρά Ἡρακλεου τοῦ Σαμβίδη
τοῦ Σοκέως μπρός Ηρα-
5 κλοῦτος(*) τῆς Διψιλοῦ ὀνογρα(φομένου)
ἔπι ἀμφόδου Αρέβιου, ἀπογρά(φομαι)
τὸν ἐπιγενθέντα μο[ν]
ἐκ τῆς συνούσης μοι —
γυναικός Ταμέριουν
10 τῆς Ἡρακλεού(*) οἰσν —
Ἡρακληνή(*) γεγενημένη(νον)
τού, β έτη Αὐτοκράτορος
Καισαρος Οὐστησανού
Σεβαστος καλ θντη (Ἐπον) γ
15 Επιδισσωμά τη(ν) ἀπογ(ραφήν). —
(Hand 2) Απολόνιος θυμοβ(άρχης) σεσήμεωμα
περὶ ἐπιγενημένου. Έπος
τετάρτου Αὐτοκράτορος Καισαρος
Οὐσιαστανού Σεβαστος Έριπ
20 Ιε-

Apparatus

Α 4-5. BL 8.196 : Ἡροκλίτος prev. ed.
Δ 10. I. Ἡρακλεῖον, BL 8.196 : Ἡρακλεῖον prev. ed.
Δ 11. I. Ἡρακλίδην, BL 8.196 : Ἡρακλεῖδην prev. ed.

Editorial History: All History: ([detailed](#))

© Duke Databank of Documentary Papyri. This work is licensed under a [Creative Commons Attribution 3.0 License](#).

Images [[open in new window](#)]



leidenpapinst.apis.pwarren02.r §

Notice: Each library participating in APIS has its own policy concerning the use and reproduction of digital images included in APIS. Please contact the [owning institution](#) if you wish to use any image in APIS or to publish any material from APIS.

APIS Translation (English)

(From P.L.Bat. I 2) [To Apollonios, the amphodarch of the] Arabs-[quarter] – from Herakleios son of Sambas, the son of Sokeus, his mother being Herakleis (5) daughter of Diphilos, registered in the Arabs-quarter. I return my son Herakleides born to me by my wife Tamaron, (10) daughter of Herakleios, born in the 2nd year of Imperator Caesar Vespasianus Augustus and now in his 3rd year. (15) I deliver the return. (Subscription of the amphodarch. Date).

Fig. 18: Record with the text of a papyrus in *Papyri.info*: P.Warr. 2, notice of birth. [https://web.archive.org/save/http://papyri.info/ddbdp/p.warr.2?fq=PHRASE:\(birth\)&rows=2&start=0&fl=id%2Ctitle&fq=%28images-int%3Atrue%29&fq=translation_language%3Aen&fq=metadata%3A%28birth%29&sort=series+asc%2Cvolume+asc%2Citem+asc&p=1&t=1](https://web.archive.org/save/http://papyri.info/ddbdp/p.warr.2?fq=PHRASE:(birth)&rows=2&start=0&fl=id%2Ctitle&fq=%28images-int%3Atrue%29&fq=translation_language%3Aen&fq=metadata%3A%28birth%29&sort=series+asc%2Cvolume+asc%2Citem+asc&p=1&t=1)

Fig. 19: EpiDoc file of P.Warr. 2, downloaded from *Papyri.info*.

26 As we can see in the example of an XML file of a papyrus text in *Papyri.info* (fig. 18 and 19), the square brackets that indicate a lacuna in the text, e.g. of five letters in line 1, are rendered with the XML tag `<gap reason="lost" quantity="5" unit="character"/>`; an uncertain letter, e.g. ε in line 3, is enclosed within the `<unclear>` tag, and will be displayed in *Papyri.info* with a dot under the letter, as according to the Leiden convention; abbreviations are rendered by surrounding the whole expanded word with the `<expan>` tag, and adding a further tag, `<ex>` around the letters omitted by the scribe.

27 The XML files of *Papyri.info* have a Creative Commons Attribution 3.0 Licence, which allows free adaptation of the material for scholarly reuse.²⁹ The collection has recently started being exploited in this sense, even though there are still few examples of this use, which obviously requires higher technical skills. One needs to download the XML files of the texts, which can be done both for single papyri and for the collection as a whole. Texts can be downloaded individually by clicking on the "xml" link in the transcription heading of each papyrus record. Otherwise, one can download the entire collection from the GitHub data repository of *Papyri.info* (*Papyri.info IDP (Integrating Digital Papyrology) Data*, <https://web.archive.org/web/20180202152933/https://github.com/papyri/idp.data>). Here the XML files of the texts are contained in the

"DDB_EpiDoc_XML" folder, grouped in bundles according to the publication volumes; a bundle with the texts provided with translations in English or German is also available from the GitHub repository, by the name "HGV_trans_EpiDoc". While the decision to make this and all the other data of *Papyri.info* publicly available ensures transparency and allows a variety of scholarly reuses, the possibility to download of the texts could be more flexible and more clearly indicated. Ideally, it should be possible to download a bundle of selected files as well, besides those grouped by their publication volume and those that include translations. Also, a link to the GitHub repository should be provided, as there is currently no way to know about the possibility of downloading all the texts in one bundle or the translations (as well as the files of the metadata, the bibliography, some Arabic papyri, and the RDF data).³⁰

Projects that reuse the content of *Papyri.info*

28 The University of Helsinki *Semantia* project³¹ is utilising texts downloaded from *Papyri.info* to build up a corpus of treebanked Greek documentary papyri. After being downloaded, the texts are enriched with morpho-syntactic annotation, carried out with the *Arethusa* tool on the *Perseids* platform. The corpus will enable the analysis of the linguistic structures of the texts on an extensive scale, and will facilitate comparisons between the scribes' original text (with mistakes, peculiar spellings, and non-expanded abbreviations) and the standard Greek, and between different scribes' spellings. Another example of reuse of the collection for the needs of one's research is a project that is being carried out for a doctoral dissertation in the Department of Linguistics of the KU Leuven, by Alek Keersmaekers. Like in the *Semantia* project, he is using the material downloaded from the PN to create a corpus of morphologically annotated texts of documentary papyri; however, in *Semantia* the linguistic annotation is carried out manually, whereas Keersmaekers performs it automatically on the basis of machine learning techniques.³²

29 The possibility of downloading the EpiDoc XML files with the texts and metadata and to repurpose the collection to other researches' ends obviates the issue of the lack of linguistic or semantic annotation, sometimes pointed out by users.³³ In fact, the texts of *Papyri.info* have been used in the *Trismegistos* (TM) platform to create databases of person and place names, respectively *TM People* and *TM Places*: a computer-aided extraction of these categories of nouns was applied to the downloaded texts, starting from the automatic filtering of capitalised words.³⁴ Recently, *Trismegistos* added to its

platform a new database, *TM Words*³⁵, developed in cooperation with Alek Keersmaekers, again created on the basis of the XML version of the *Papyri.info* texts available on GitHub: it contains all the words attested in papyri in the texts on *Papyri.info*, and they are searchable by lemmata and conjugated or declined forms.

The screenshot shows the Papyri.info website with a red header bar containing the site's name. Below the header, a navigation bar includes links for 'Browse' (DDhDP, HGV, APIS, TM Number) and 'Search' (Data, Bibliography). The main content area is titled 'Text Leiden+ Documentation'. A sidebar on the left lists various documentation categories with their sub-links:

- Formatting - Word and Document
 - Document Division
 - Line Number
 - Vacat
 - Marginalia
 - Special Formatting
- Preservation
 - Letters Lost
 - Lines Lost
 - Illegible
 - Vestiges
- Leiden Conventions
 - Leiden Angle Brackets < >
 - Leiden Braces { }
 - Leiden Double Square Brackets []
 - Leiden Parentheses ()
 - Leiden Square Brackets []
 - Leiden Underline
- Other Editorial Conventions
 - Handshift
 - Editorial Note
 - Quotation
- Apparatus
 - Orthographic Regularization
 - Alternate Reading
 - Scribal Correction
 - Spelling Correction
- Corrections to Published Texts
 - Berichtigungliste Corrections
 - Editorial Correction
 - PN/PE Correction
 - Complicated Corrections
- Numbers and Special Characters
 - Numbers
 - Diacriticals
 - Special Characters
- Language
 - Non-default Language
 - Omitted
 - Orthographic Regularization, for Language

Below the sidebar, there are two sections: 'Formatting - Word and Document' and 'Document Div, ab (anonymous block)'. Each section contains examples of PN preview, Leiden+, XML, and example PN links, along with notes about the encoding of text blocks.

Fig. 20: The "Leiden+ Documentation" on the *Papyrological Editor*: the index and the beginning of this list of markup for encoding the texts. https://web.archive.org/web/20171025163741/http://papyri.info/docs/leiden_plus.

30 A minor problem encountered in downloading the texts, experienced by Alek Keersmaekers for his research and by myself when importing a text to *Arethusa* to carry out the linguistic annotation for *Sematia*, is the necessity to sometimes split up words manually, as the space for the word division can be missing in the XML files of the texts. This issue occurs whenever the end of a word coincides with the end of a line in the papyrus text. Regarding this problem, it can be mentioned that there are currently no guidelines in *Papyri.info* on how to enter texts with the *Papyrological Editor*, as much as the "Text Leiden+ Documentation" (the complete list of markup signs for encoding the texts)³⁶ is of great help, it would be convenient to add guidelines so as to provide users with indications on the whole encoding procedure of a text, especially those who have not attended an EpiDoc workshop or might need to recollect the procedure after some time. Within these directions, it could be suggested to contributors to add a space after the word endings that occur at the end of a line: this detail, which entails no difference in the text publication, would however ensure a correct word division in the XML files whenever they are downloaded.

The user interface of the Papyrological Navigator (PN)

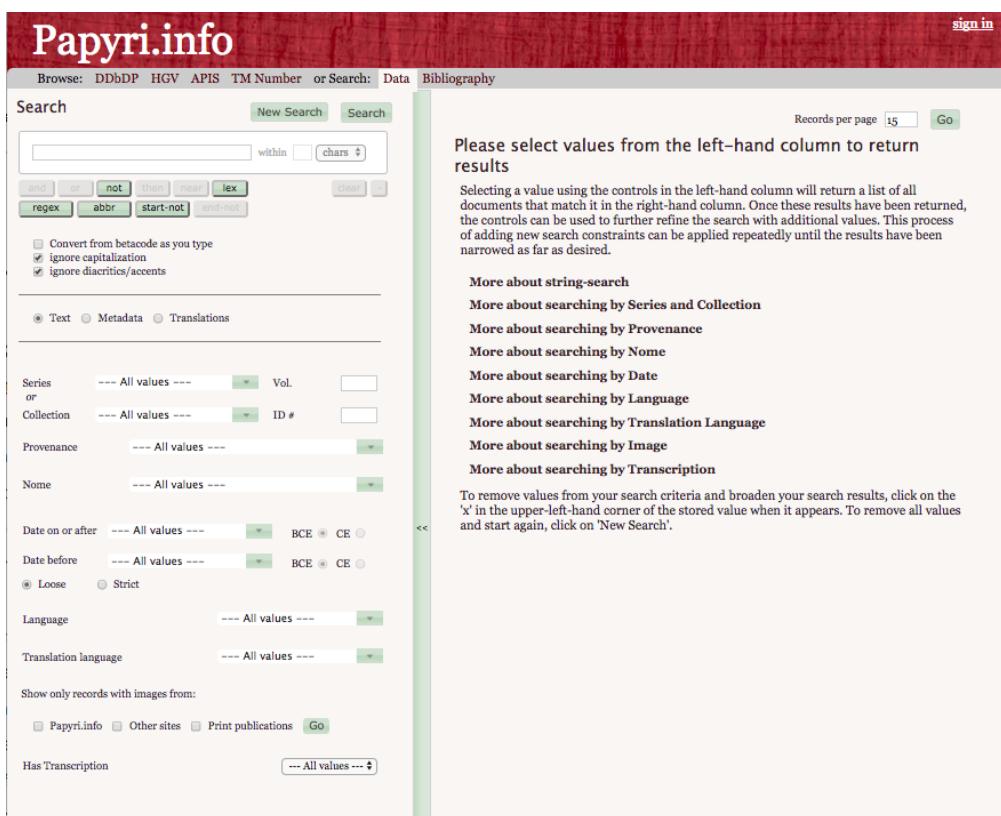


Fig. 21: The PN search interface. <http://papyri.info/search>.

The screenshot shows the Papyri.info search interface. On the left, there is a 'Refine Search' sidebar with various filters like 'Series or Collection', 'Provenance', 'Name', 'Date on or after', 'Language', and 'Translation language'. Above the sidebar is a search bar with operators like 'and', 'or', 'not', 'then', 'near', 'lex', 'regex', 'abbr', 'start-not', and 'end-not'. Below the search bar are checkboxes for 'Convert from betacode as you type', 'ignore capitalization', and 'ignore diacritics/accents'. The main area shows a list of 7455 hits. The first few results are:

Identifier	Title	Location	Date	Languages
basp 49 135	Fragment	Koptos	297 CE	grc
bgu 1 15	keiner	Arsinoites	197 CE	grc
bgu 1 33	keiner	Arsinoites	101 CE - 300 CE	grc
bgu 1 37	keiner	Arsinoites	50 CE	grc
bgu 1 38	keiner	Arsinoites	101 CE - 300 CE	grc
bgu 1 48	keiner	Arsinoites	101 CE - 300 CE	grc

Fig. 22: The PN search interface after performing a search.

31 The user interface is divided into two columns, which contain the search form, on the left-hand side, and detailed directions for its use, on the right-hand side (fig. 21), respectively. A small drawback of the directions' display is that, if we narrow the browser window even only slightly, they move to the bottom of the page, sometimes becoming completely invisible. We also cannot see the directions after performing a search, as the right column gets occupied by the list of the results (fig. 22). Hence it would be useful to always have a link to the directions, in the result page, too, as one interviewed user highlighted. In fact, she could not recall where the directions were, as they are not available on the result page and she did not notice them on the PN home page, since she was focused on the left part of the screen to enter her data. The link to the search directions could be added to the grey menu bar on top of the page, after the "Bibliography" link, thereby being visible on every page and regardless of the width of the browser window.

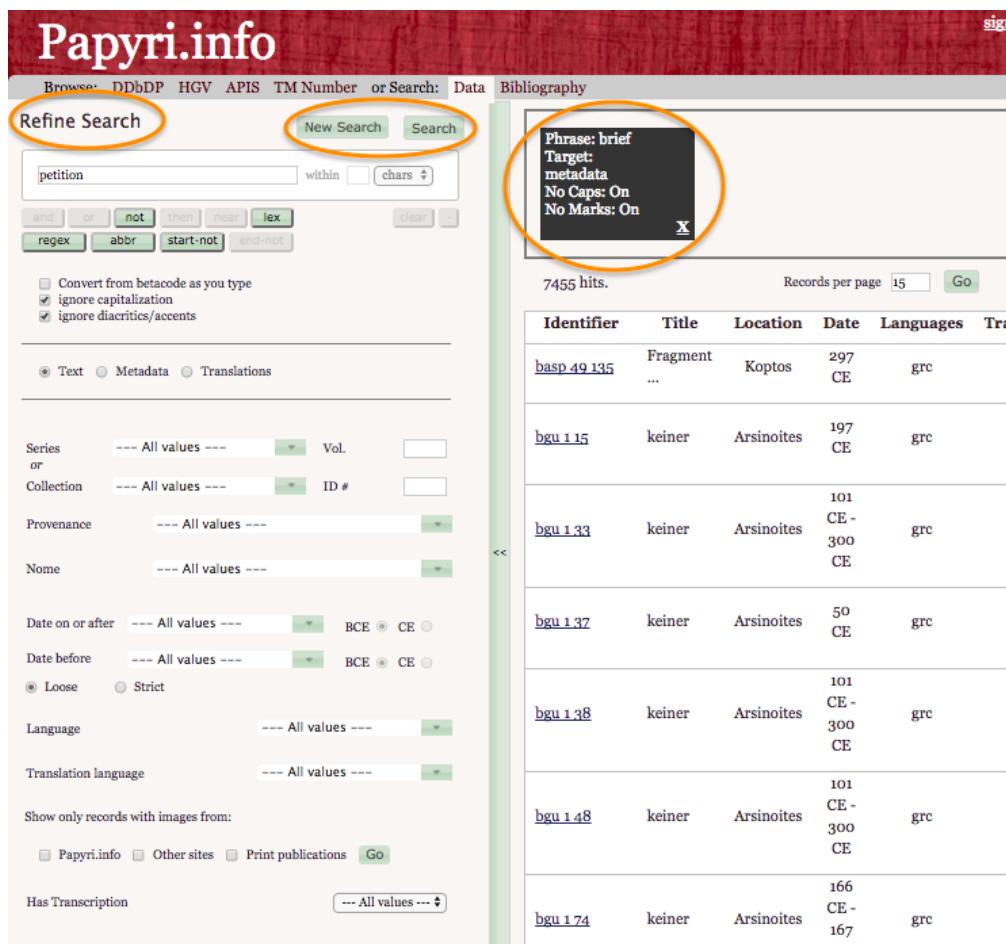


Fig. 23: The PN search interface after performing a search. The buttons and the information to perform new searches are highlighted.

32 Users can choose to search the texts, the metadata or the translations by selecting the respective buttons. It would be a small improvement to facilitate a new search on the result page of a previous query, rather than on the home page.

33 On the result page, one may expect that entering a new word and hitting the "Search" button will perform a new search; the result is, instead, a refined search. Users should, in fact, click on the x-button in the black box on top of the results to remove the previous search value; however, this button is not evident, as was also pointed out by one of the interviewed users. As an alternative, one can perform a new search by clicking on the "New Search" button; users, though, need to remember that they should first click on "New Search", and only then enter the new value (and finally perform their search by hitting "Search"): if one types the new word first, and then hits "New Search", as it may happen, this will remove the new value. This aspect could be improved by allowing the "New Search" button to perform a search with the new value, rather than to go back to the PN home page. Moreover, the "Search" button could be renamed "Refine

"Search" to explain its function more clearly; even though this phrase is already present, on the top left, it would be more evident if it was positioned on the button, where users actually need to click.

34 Other aspects of the PN user interface can be better illustrated by way of a comparison with the interface of the HGV, the Heidelberg database with records of metadata. As we have seen, the PN offers more accurate results than the HGV because it does not contain multiple records for the same papyrus, as the HGV does whenever a dating is uncertain. On the other hand, a reason why one may still wish to utilise the HGV database is that its interface offers more possibilities of search than that of the PN. The HGV allows searching for a date according to a single year, and even a month or a day; moreover, one can search for any range of time (by entering any year into the "ChronMinimum" and "ChronMaximum" fields). Instead, on the PN one can perform a search by date exclusively by selecting a predefined value in the "Date on or after" and/or "Date before" fields, within a fifty-year range. Surprisingly, if one tries searching for a single year among the PN metadata (by entering it in the search field and selecting the "Metadata" option), no result is retrieved, although the help information (in the right-hand pane of the search) states that this search is possible, as one in fact expects.³⁷

sign in

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

chla.41.1195 = HGV P.Cair. Masp. 1 67030 = Trismegistos 18994 = oxford-ipap.apis.154
 = p.cair.masp.1.67030

<< Previous record Next record >>

[metadata](#) [HGV data](#) [TM data](#) [APIS catalog record](#) [text](#) [transcription](#) [open in editor](#) Canonical URI: <http://papyri.info/ddbdp/chla:41;1195>

HGV P.Cair. Masp. 1 67030 [source] [xml]

Title	Ordre
Publications	P.Cair. Masp. 1 67030 [More in series P.Cair. Masp.] [More in series P.Cair. Masp., vol. 1] ChLA XLI 1195
Support/Dimensions	Papyrus
Post-Concordance	IX 41; X 33; XI 52
BL Entries	
Origin	Aphrodites Kome (Antaiopolites) [More from Aphrodites Kome (Antaiopolites)]
Provenance	Aphrodites Kome Antaiopolites Ägypten
Material	Papyrus
Date	531 (?) [More from the period between 531 CE and 532 CE]
Date	Anfang 546 (?) [More from the period between 546 CE and 547 CE]
Commentary	Gehört zum Dioskoros - Archiv. Alternativdatierung: Anfang 546. Zur Datierung vgl. PLRE III A, S. 607; BL IX, S. 41 und Zuckerman, Du village à l'empire, S. 190. Griechisch - lateinisch. Pl. XVIII - XX, ChLA XLI, S. 38, 39, 41; Fournet, Les archives de Dioscore, S. 361, Planche IX
Print Illustrations	
Subjects	Anordnung; praeses; Verschiffung; Embole
Images	http://ipap.csad.ex.ac.uk/4DLINK/4DACTION/IPAPwebquery?vPub...
Images	http://www.misha.fr/papyrus_bipab/pages_html/P_Cair_Masp_1_6...
License	© Heidelberg Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens. This work is licensed under a Creative Commons Attribution 3.0 License .

Trismegistos: 18994 [source]

Publications	P. Cairo Masp. 1 67030 = ChLA 41 1195
Inv. no.	Cairo, Egyptian Museum JdE 40773
Date	AD 546? (cf. PLRE III A, p. 607; BL 9, p. 41); or AD 531?
Language	bilingual: Greek / Latin
Provenance	Egypt, U15 - Antinoopolis (El-Sheikh Ibada) [written] ; Egypt, U10a - Aphroditopolis (Kom Ishgau) [find place & destination]
Archive	Dioskoros
Places	mentioned places

Fig. 24: The initial part of the record of a papyrus, P.Cair.Masp. 1 67030, in the PN ([http://papyri.info/ddbdp/chla:41;1195/?q=PHRASE:\(fournet\)&rows=3&start=0&fl=id%2Ctitle&fq=metadata%3A%28fournet%29&sort=series+asc%2Cvolume+asc%2Citem+asc&p=1&t=137](http://papyri.info/ddbdp/chla:41;1195/?q=PHRASE:(fournet)&rows=3&start=0&fl=id%2Ctitle&fq=metadata%3A%28fournet%29&sort=series+asc%2Cvolume+asc%2Citem+asc&p=1&t=137), accessed 4 October 2017). The HGV section and its "Subjects" field are highlighted. In whatever language the reference edition is, and therefore the title of the record (here in French: "Ordre"), the keywords of the content type are reported in German, among the "Subjects" ("Anordnung", etc.). They are the same as the keywords reported in the HGV "Inhalt" field (fig. 21).

sign in

Browse: DDbDP HGV APIS TM Number or Search: Data Bibliography

chla.41.1195 = HGV P.Cair. Masp. 1 67030 = Trismegistos 18994 = oxford-ipap.apis.154
 = p.cair.masp.1.67030

<< Previous record Next record >>

[metadata](#) [HGV data](#) [TM data](#) [APIS catalog record](#) [text](#) [transcription](#) [open in editor](#) Canonical URI: <http://papyri.info/ddbdp/chla:41;1195>

HGV P.Cair. Masp. 1 67030 [source] [xml]

Title	Ordre
Publications	P.Cair. Masp. 1 67030 [More in series P.Cair. Masp.] [More in series P.Cair. Masp., vol. 1] ChLA XLI 1195
Support/Dimensions	Papyrus
Post-Concordance	IX 41; X 33; XI 52
BL Entries	
Origin	Aphrodites Kome (Antaiopolites) [More from Aphrodites Kome (Antaiopolites)]
Provenance	Aphrodites Kome Antaiopolites Ägypten
Material	Papyrus
Date	531 (?) [More from the period between 531 CE and 532 CE]
Date	Anfang 546 (?) [More from the period between 546 CE and 547 CE]
Commentary	Gehört zum Dioskoros - Archiv. Alternativdatierung: Anfang 546. Zur Datierung vgl. PLRE III A, S. 607; BL IX, S. 41 und Zuckerman, Du village à l'empire, S. 190. Griechisch - lateinisch. Pl. XVIII - XX, ChLA XLI, S. 38 - 39 - 41; Fournet, Les archives de Dioscore, S. 361, Planche IX
Print Illustrations	
Subjects	Anordnung; praeses; Verschiffung; Embole
Images	http://ipap.csad.ex.ac.uk:42100/x4/4ACTION/IPAPwebquery?vPub...
Images	http://www.misha.fr/papyrus_bipab/pages_html/P_Cair_Masp_1_6...
License	© Heidelberg Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens. This work is licensed under a Creative Commons Attribution 3.0 License .

Trismegistos: 18994 [source]

Publications	P. Cairo Masp. 1 67030 = ChLA 41 1195
Inv. no.	Cairo, Egyptian Museum JdE 40773
Date	AD 546? (cf. PLRE III A, p. 607; BL 9, p. 41); or AD 531?
Language	bilingual: Greek / Latin
Provenance	Egypt, U15 - Antinoopolis (El-Sheikh Ibada) [written] ; Egypt, U10a - Aphroditopolis (Kom Ishgau) [find place & destination]
Archive	Dioskoros
Places	mentioned places

Fig. 25: The HGV record of P.Cair.Masp. 1 67030. Circled, the keywords, which are exported to the PN "Subjects" field. <https://web.archive.org/web/20171004154817/http://aquila.zaw.uni-heidelberg.de/hgv/18994>.

35 Another frequent kind of search is the one by content type, as scholars need to find documents of the same type (e.g., a petition or a will), to study their features and to verify supplements in lacunous passages. This kind of search is possible in the PN, too, like in the HGV, but it is less clear to find out how this search works. While in the Heidelberg database there is a specific field for the search by kind of content, in the PN this information needs to be searched for in the "Search" field, ticking the "Metadata" option, as it does not have a dedicated field, unlike, for example, "Provenance" or "Date". In this way, the ability to do this search may not be evident: one of the users I interviewed was not aware of that possibility, even though he had interest in finding documents of a specific type on which his research focuses. Above all, users are not made aware that, because the metadata is imported from the HGV, the term referring to the content type needs to be entered in German (whereas in the HGV search interface this is self-evident as German is the language of all the information throughout the database). In fact, in

every PN record the content type of a papyrus is indicated by a term in German, in the "Subjects" field of the HGV section ([fig. 25](#)). Because this indication is not reported, one of the users I interviewed, to make sure to find all the documents of a specific type in whatever language they have been published, enters the term of her interest in several languages (English, German, French and Italian), which is however unnecessary.

36 It is also true that users look for types of documents not only by means of a metadata search, but also by searching the texts for significant words that are typical of certain documents; they are therefore able to find texts of a certain category with this other method. Two users I interviewed combine this kind of search with the previous one about the metadata: for example, they look up formulas or keywords, like " $\epsilon\kappa\alpha\pi\eta\zeta$ " ("at once"), a word typically occurring in summons, in order to retrieve all the texts of this category. Indeed, the HGV warns users that "there is no guarantee of completeness in the case of searching according to content as we did not develop a system of categorisation"³⁸; it would be useful to point out this limitation in the PN search form, as well, in the help information in the right-hand part of the search page (among the information about "Metadata", in the "More about string-search"), so as to suggest users to use both search methods.

37 In conclusion, there are a few circumscribed problems that slightly limit the usability of the search interface, especially for first users, sometimes causing them to spend more time than necessary when performing a new search, looking for typologies of documents, and seeking for texts within a limited time-span.

Conclusion

38 The process of integration of the three main databases, accomplished in the *Integrating Digital Papyrology* project, has allowed the integration of different kinds of data, in different formats: the texts, coming from the *Duke Databank* and encoded in Beta Code and SGML, were converted into Unicode and EpiDoc TEI; they were integrated with the metadata and the images, on records encoded in FileMaker and exported to XML, from the Heidelberg HGV and the Michigan APIS databases. The merging of texts and metadata was facilitated by the addition, to the texts, of a unique identifier, i.e., the *Trismegistos* number. A great effort was put into this process, in the *Integrating Digital Papyrology* project which consisted in three phases and lasted six years, from 2007 to 2012 (Sosin 2010, Baumann et al. 2011). Papyrologists can thus now access their main resources from only one extensive platform. The conversion to EpiDoc, an open markup

standard which is a subset of TEI, ensures that *Papyri.info* has a high degree of compatibility with other projects in the Humanities and a long-term sustainability.

39 To summarise the strengths and weaknesses of *Papyri.info*, we might highlight that the comprehensiveness of the material, the integration of the different databases (both of texts and metadata, and of research tools) and the ability of an integrated search across them explain why this is an indispensable resource for the study of documentary papyri. However, precisely to valorise the collection, that is, to enable users to take advantage of all its features, there should be a better communication about the characteristics of its content and tools, and about the procedure to contribute new data. Other useful additions would be pointing to resources where similar content can be found, as happens for the literary and medical papyri on the *Digital Corpus of Literary Papyri*, and more accuracy in the documentation relative to the involved institutions and staff.

Notes

1. The launch took place at the XXVI International Congress of Papyrology (Geneva), with a talk by Joshua Sosin (Sosin 2010); this information is reported in the "About" page of *Papyri.info*, <https://web.archive.org/web/20170603015536/http://www.papyri.info/docs/about>.

2. This collaborative model of publishing, which also includes an editorial board to ensure the quality of the texts, is illustrated in Baumann 2013, 92-93; Bagnall 2010, 5-6; Sosin 2010.

3. For the physical collections that have been digitised in the APIS database, see the overview of this project outlined in Bagnall-Gagos 2007, in part. 66-67. Roger Bagnall and Traianos Gagos, founders of APIS, soon realised the potential for digitisation to bring to light almost forgotten collections; this led them to digitise not only the papyri of the Michigan and other American collections, but also some unpublished papyri of European institutions (Bagnall-Gagos 2007, 66-67).

4. Gagos 1996, 19-20; Gagos 1997, 155.

5. These figures are drawn from the search form on the PN home page: at the bottom, the drop-down menu of the "Has Transcription" selection criterion shows the values "true

(55080)" and "false (31936)". As one can infer, this indicates the records complete with texts and those which so far exclusively contain metadata.

6. Cf. Evans-Obbink 2010, 1-3, for an overview on the significance of the papyri for the study of the ancient languages. The authors also explain how the research on the language of the papyri is one of the areas that mostly benefits from technological advances. Through the citations of the DDbDP in this volume (found by searching for "DDbDP" in the index) one can find instances of linguistic researches conducted with the aid of this database, now merged into *Papyri.info*.

7. Sosin 2010; *Papyri.info*, "The DDbDP"; *Integrating Digital Papyrology* 3, 10-11, 18-21.

8. J. Cowey and C. Lanz, personal communications.

9. On the decisive contribution of the Packard Foundation and the PHI to the creation of the DDbDP, see the history of this database outlined in Oates 1993, 62-64, 69-71.

10. On the conversion of the texts in Beta Code and SGML into Unicode and EpiDoc XML, see Baumann et al. 2011, esp. "Integrating Digital Papyrology" and "Lessons from the conversion of the Duke Databank of Documentary Papyri from legacy formats into EpiDoc TEI XML"; Baumann (2013, 93-100) explains the choice of Git and its functioning; Baumann et al. 2011, esp. "The Papyrological Navigator: Project Integration with RDF", and Baumann 2013, 100 deal with the use of an RDF triplestore for integrating the source databases.

11. I was provided with the above information on the updating of the HGV and the PN by James Cowey and Carmen Lanz, in personal communications.

12. Joshua Sosin, personal communication.

13. James Cowey, personal communication. The presence of multiple records for one text, in order to represent different possibilities of dating, has also been reported in Hagedorn 1994, 227, and in Bagnall 1998. For example, a search for petitions retrieves 2251 results on the PN and 2285 on the HGV (it is possible to search for the content type by entering a term in German, e.g. "Eingabe" for "petition", both in the HGV, in the "Inhalt" field, and in the PN selecting the "Metadata" option).

14. This feature has been noted by Reggiani (2017, 74-75).

15. James Cowey, personal communication. The presence of multiple records for one text, in order to represent different possibilities of dating, has also been reported in Hagedorn 1994, 227, and in Bagnall 1998. For example, a search for petitions retrieves 2251 results on the PN and 2285 on the HGV (it is possible to search for the content type by entering a term in German, e.g. "Eingabe" for "petition", both in the HGV, in the "Inhalt" field, and in the PN selecting the "Metadata" option).

16. We find this information on the *Bibliographie Papyrologique* home page: 'Une version électronique simplifiée de la «BP rétrospective» (provisoirement arrêtée en 2012) est également accessible via la plate-forme papyri.info'. This problem is pointed out in Reggiani (2017, 20-21).

17. Cf. Reggiani 2017, 20-22 for this discussion on the advantages of the integration between the PN and the BP.

18. <https://web.archive.org/web/20170603015536/http://www.papyri.info/docs/about>, see below, fig. 3.

19. Reggiani (2017) has in fact used the information available on the *Digital Papyrology blog* to show the modifications added to *Papyri.info* over the years; see, e.g., Reggiani 2017, 236, n. 119.

20. A user I interviewed proved not to be aware of the scope of the collection, as she remarked that she cannot find, on the PN, the literary papyri she needed; she had been attending a module of Papyrology a few years ago, but either was not given this indication or was not able to recollect it as she had then not carried out papyrological research for a few years.

21. Cf. the DCLP home page: "DCLP offers information about and transcriptions of Greek and Latin literary and subliterary papyri". On the addition of literary and medical papyri to this database, see Reggiani 2017, 253, 274-275.

22. On the development of this project of the University of Parma, and on a discussion of the issues involved in the digitisation of texts with different features from the documentary ones, see Reggiani 2016 and Reggiani 2017, 251-254, 274-275.

23. Cf. Reggiani 2017, 275: "At the time, digitizations took place in the *Papyrological Editor*, in a special community called 'ParmaMed', which avoided the phase of submitting

the edited texts to the *Papyri.info* board; when the DCLP challenge started, the Parma medical project became one of the earliest partners to contribute content".

24. Oates 1993, 64-66; Willis 1984, 167-169.

25. The sixteen users I have interviewed to date include PhD students and early career researchers of different European universities. Interviews and other data will be analysed in my PhD dissertation with the title *The role of Digital Humanities in Papyrology: practices and user needs in papyrological research*, at the Institute of Classical Studies, London. This analysis aims to gain insight into the scholarly use of Digital Classics methods and tools, in particular by papyrologists and other classicists who need to consult papyrological resources.

26. EpiDoc guidelines and material are available at <http://epidoc.sf.net> (accessed 11 January 2018). A list of projects that use the EpiDoc standards is available at *The Digital Classicist Wiki*, "Category:EpiDoc", <https://web.archive.org/web/20170930144706/> <https://wiki.digitalclassicist.org/Category:EpiDoc>.

27. Cf. *Integrating Digital Papyrology* 3; Sosin 2010.

28. For more details on the history of EpiDoc, I refer to Bodard 2010, 101-104, and Cayless et al. 2009, 1-2.

29. Creative Commons, "Attribution 3.0", <https://web.archive.org/web/20170930110422/> <https://creativecommons.org/licenses/by/3.0/>. As well as the XML files of the texts, also the code of its core components is open to reuse, as it is freely available online, stored in Git (Baumann 2013).

30. The information on the GitHub repository, where the complete data of *Papyri.info* can be accessed, is only reported, as far as I am aware, in the bibliography where the GitHub repository is mentioned (Baumann 2013, 93-100), and on *The Digital Classicist*. *The Digital Classicist Wiki*. "Greek and Latin texts in digital form". https://wiki.digitalclassicist.org/Greek_and_Latin_texts_in_digital_form#Texts_with_translations. Accessed 2 February 2018.

31. *Semavia*. University of Helsinki. <https://semavia.hum.helsinki.fi/docs/>. Accessed 5 October 2017. This project and the issues of linguistic annotation of fragmentary papyrus texts are discussed in Reggiani 2017, 182-184.

32. KU Leuven, *Research Portal*, "Variety and change in the Ancient Greek papyri: a corpus-linguistic study of the use of tense, aspect and modality in the Greek complementation system" by A. Keersmaeker, <https://www.kuleuven.be/onderzoek/portaal/#/projecten/3H160543> (accessed 10 October 2017).

33. One of the interviewees mentioned the fact that there are no morpho-syntactic annotations in the texts in the PN, which therefore cannot be queried in this way, for example to search for all the infinitives. By the same token, Clarysse (2010, 44) remarks the absence of syntactical annotation, which would permit to search for different kinds of clauses.

34. The procedure for extracting person names from the texts of the DDbDP is described in Depauw-van Beek 2009, 33-37.

35. *TM Words* (<https://web.archive.org/web/20180313155651/https://www.trismegistos.org/words/index.php>), unlike the others TM databases, is not indicated on the TM platform homepage, maybe because it is still in its beta phase. It is anyway known to papyrologists as it was announced by Mark Depauw, the TM project director, on the *Papy-list*, the papyrological discussion list, on 9 January 2018.

36. *Papyri.info*, "Text Leiden+ Documentation", https://web.archive.org/web/20171010170456/http://papyri.info/docs/leiden_plus.

37. Cf. the following passage in the "More about string-search" section, about the metadata search: "Metadata: Searches the metadata associated with the document - for example, the identification number, location, and dating of the document".

38. HGV, "Search Tips and other Information", <https://web.archive.org/web/20171004170724/http://aquila.zaw.uni-heidelberg.de/help/search/en>.

References

APIS – Advanced Papyrological Information System, University of Michigan Library,
<https://web.archive.org/web/20160810083015/http://www.lib.umich.edu:80/papyrology-collection/advanced-papyrological-information-system-apis>.

Arethusa.

<https://web.archive.org/web/20171009155602/http://www.perseids.org/tools/arethusa/app/#/>.

Bagnall, R. S. 1998. "Review of Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens". Accessed 24 October 2018. *Bryn Mawr Electronic Resources Review*.

<http://bmcr.brynmawr.edu/bmerr/1998/BagnaHeideAug.html>.

Bagnall, R. 2010. "Integrating digital papyrology". In *Online humanities scholarship: the shape of things to come*, edited by F. Moody & B. Allen. New York: Mellon Foundation. Accessed 30 September 2017.

<http://hdl.handle.net/2451/29592>.

Bagnall R. S. and T. Gagos. 2007. "The Advanced Papyrological Information System: Past, Present, and Future." In *Proceedings of the 24th International Congress of Papyrology, Helsinki 2004*, edited by J. Frösén, T. Purola, and E. Salmenkivi, vol. 1, 59-74. Helsinki: Societas Scientiarum Fennica.

Baumann, R. 2013. "The Son of Suda On-Line". In *The Digital Classicist 2013* (= BICS Supplement 122), edited by S. Dunn and S. Mahony, 91-106. London: Institute of Classical Studies, School of Advanced Study, University of London.

<https://web.archive.org/web/20171025174938/https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/8414/SUPP-122-07-baumann-offprint%20%281%29.pdf?sequence=1>.

Baumann, R., G. Bodard, H. Cayless, J. Sosin, and R. Viglianti. 2011. "Panel: Integrating Digital Papyrology", in *Digital Humanities 2011: June 19-22, Conference of the Alliance of Digital Humanities Organisations, hosted by the Stanford University Library*.

<https://web.archive.org/web/20170710154928/http://dh2011abstracts.stanford.edu/xtf/view?docId=tei/ab-193.xml&brand=default>.

Berichtigungsliste der griechischen Papyrusurkunden aus Ägypten, voll. 1-13. 1922-2017.

Bibliographie Papyrologique. Association Égyptologique Reine Élisabeth, Centre de Papyrologie et d'Épigraphie grecque de l'Université libre de Bruxelles. Accessed 25 September 2017.

<http://www.aere-egke.be/BP.htm>.

Bodard G. 2010. "EpiDoc: Epigraphic Documents in XML for Publication and Interchange". In Feraudi-Gruénais, F. 2010. *Latin on stone: Epigraphic Research and Electronic Archives*. Lanham, MD: Lexington Books, 101-118.

https://web.archive.org/web/20180111113342/http://www.stoa.org/wordpress/wp-content/uploads/2010/09/Chapter05_EpiDoc_Bodard.pdf.

Cayless H., C. Roueché, T. Elliott, and G. Bodard. 2009. "Epigraphy in 2017". *Digital Humanities Quarterly* 3.1.

<https://web.archive.org/web/20170624164625/http://www.digitalhumanities.org:80/dhq/vol/3/1/000030/000030.html>.

Clarysse, W. 2010. "Linguistic Diversity in the Archive of the Engineers Kleon and Theodoros". In Evans-Obbink 2010, 35-50.

Digital Corpus of Literary Papyri (DCLP). University of Heidelberg.

<https://web.archive.org/web/20171024181128/http://litpap.info>.

Digital Papyrology blog.

<https://web.archive.org/web/20170927110954/http://digitalpapyrology.blogspot.co.uk>.

Depauw M. and B. Van Beek. 2009. "People in Greek Documentary Papyri. First Results of a Research Project". *JJP* 39, 31-47.

EpiDoc: Epigraphic Documents in TEI XML. Elliott T., Bodard G., Cayless H., et al. (2006-2016), Online material, available, accessed 30 September 2017:
<http://epidoc.sf.net>.

Evans, T.V. and D.D. Obbink. 2010. *The language of the papyri*. Oxford: Oxford University Press.

Gagos, T. 1996. "Scanning the Past: A Modern Approach to Ancient Culture." *Library Hi Tech*, 14 (1): 11-22. Accessed 13 March 2018.
<http://dx.doi.org/10.1108/eb047974>.

Gagos, T. 1997. "Advanced Papyrological Information System (APIS): The Michigan Experience." *LLC* 12 (3): 155-157. Accessed 13 March 2018.
<https://doi.org/10.1093/lrc/12.3.155>.

Hagedorn D. 1994. "Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens" In *Datenbanken in der Alten Geschichte* (= Computer und Antike 2), edited by M. Fell, C. Schäfer, L. Wierschowski, 226-231. St. Katharinen.

HGV – Heidelberger Gesamtverzeichnis der griechischen Papyruskunden Ägyptens, Institut für Papyrologie. Universität Heidelberg.
<https://web.archive.org/web/20170713104205/http://aquila.zaw.uni-heidelberg.de/start>.

Oates J. F. 1993. "The Duke Data Bank of Documentary Papyri". In *Accessing Antiquity: The Computerization of Classical Studies*, edited by Jon Solomon and Robert D. Novak, Tucson: University of Arizona Press, 62-72.

Papyri.info.

<https://web.archive.org/web/20170925124125/http://papyri.info>.

Reggiani N. 2016. "The Corpus of Greek Medical Papyri and Digital Papyrology: new perspectives from an ongoing project". In *Altertumswissenschaften in a Digital Age: Egyptology, Papyrology and beyond. Proceedings of a conference and workshop in Leipzig, November 4-6, 2015*, edited by M. Berti and Naether. Leipzig.

<https://web.archive.org/web/20171024182110/http://www.qucosa.de/fileadmin/data/qucosa/documents/20172/1.9.%20DHEgypt15%20-%20Reggiani,%20Greek%20Medical%20Papyri%20-%20Text.pdf>.

Reggiani N. 2017. *Digital Papyrology*, vol. 1, *Methods, Tools and Trends*. Berlin-Boston: De Gruyter. Accessed 13 March 2018.

<https://www.degruyter.com/viewbooktoc/product/486978>.

Semavia. University of Helsinki. Accessed 5 October 2017.

<https://semavia.hum.helsinki.fi/docs/>.

Sosin J., 26 October 2010. "Digital Papyrology". Blog post on *The Stoa Consortium*.
<https://web.archive.org/web/20170408110552/http://www.stoa.org/archives/1263>.

TEI: Text Encoding Initiative.

<https://web.archive.org/web/20170711100834/http://www.tei-c.org/index.xml>.

The Digital Classicist. *The Digital Classicist Wiki*.

https://web.archive.org/web/20170715110130/http://wiki.digitalclassicist.org/Main_Page.

Trismegistos (TM). KU Leuven.

<https://web.archive.org/web/20170701160733/http://www.trismegistos.org/index.html>.

Willis W. H. 1984. "The Duke Data Bank of Documentary Papyri", in *Atti del XVII Congresso Internazionale di Papirologia, Napoli 1983*, vol. 1, 167-173. Napoli: Centro internazionale per lo studio dei papiri ercolanesi.

Willis W. H. 1988. "The Duke Data Bank of Documentary Papyri". In *Proceedings of the XVIII International Congress of Papyrology, Athens 1986*, edited by B. Mandilaras, vol. 2, 15-20. Athens: Greek Papyrological Society.

Willis W. H. 1992. "The new mode of access to the Duke Data Bank of Documentary Papyri". In *Proceedings of the XIX International Congress of Papyrology, Cairo 1989*, edited by A. H. S. El-Mosalamy, vol. 1, 125-131. Cairo: Ain Shams University.

Willis W. H. 1994. "The New Compact Disk of Documentary Papyri". In *Proceedings of the 20th International Congress of Papyrology, Copenhagen 1992*, edited by A. Bülow Jacobsen, 628-631. Copenhagen: Museum Tusculanum Press, University of Copenhagen.

Factsheet

Resource reviewed	
Title	Papyri.info
Editors	Joshua Sosin
URI	http://papyri.info
Publication Date	2010
Date of last access	23.10.2018

Reviewer	
Surname	Vannini
First Name	Lucia
Organization	Institute of Classical Studies
Place	London
Email	lucia.vannini (at) postgrad.sas.ac.uk

General Information		
Bibliographic description	Can the text collection be identified in terms similar to traditional bibliographic descriptions (title, responsible editors, institution, date(s) of publication, identifier/address)? (cf. Catalogue 1.1)	yes
Contributors	Are the contributors (editors, institutions, associates) of the project documented? (cf. Catalogue 1.3)	yes
Contacts	Is contact information given? (cf. Catalogue 1.4)	yes
Aims		
Documentation	Is there a description of the aims and contents of the text collection? (cf. Catalogue 2.1)	yes
Purpose	What is the purpose of the text collection? (cf. Catalogue 2.2)	Research
Kind of research	What kind of research does the collection allow to conduct primarily? (cf. Catalogue 3.1.8)	Qualitative research

Self-classification	How does the text collection classify itself (e.g. in its title or documentation)? (cf. Catalogue 2.3)	other: aggregation of papyrological documents and related materials
Field of research	To which field(s) of research does the text collection contribute? (cf. Catalogue 2.2)	History, Linguistics
Content		
Era	What era(s) do the texts belong to? (cf. Catalogue 2.5)	Classics
Language	What languages are the texts in? (cf. Catalogue 2.5)	Arabic, Greek, Latin, other: Coptic
Types of text	What kind of texts are in the collection? (cf. Catalogue 2.5)	Private documents, Files/Records, other: Papyri
Additional information	What kind of information is published in addition to the texts? (cf. Catalogue 2.5)	Bibliography, other: metadata on provenance date, etc.
Composition		
Documentation	Are the principles and decisions regarding the design of the text collection, its composition and the selection of texts documented? (cf. Catalogue 3.1.1-3.1.3)	yes
Selection	What selection criteria have been chosen for the text collection? (cf. Catalogue 3.1)	Genre
Size		
Texts/records	How large is the text collection in number of texts/records? (cf. Catalogue 3.1.4)	> 1000
Tokens	How large is the text collection in number of tokens? (cf. Catalogue 3.1.4)	unknown
Structure	Does the text collection have identifiable sub-collections or components? (cf. Catalogue 3.1.5)	yes
Data acquisition and integration		
Text recording	Does the text collection record or transcribe the textual data for the first time? (cf. Catalogue 3.1.6)	no

Text integration	What kind of material has been taken over from other sources? (cf. Catalogue 3.1.6)	Full texts, Metadata, Annotations, other: images (when they are available on the APIS database)
Quality assurance	Has the quality of the data (transcriptions, metadata, annotations, etc.) been checked? (cf. Catalogue 3.1.7)	yes
Typology	Considering aims and methods of the text collection, how would you classify it further? For definitions please consider the help-texts. (cf. Catalogue 3.1.8)	General purpose collection, Corpus, Reference corpus
Data Modelling		
Text treatment	How are the textual sources represented in the digital collection? (cf. Catalogue 3.2.1)	Normalized transcription, Diplomatic transcription, Edited text, Translated text
Basic format	In which basic format are the texts encoded? (cf. Catalogue 3.2.4)	XML
Annotations		
Annotation type	With what information are the texts further enriched? (cf. Catalogue 3.2.2)	Editorial annotations
Annotation integration	How are the annotations linked to the texts themselves? (cf. Catalogue 3.2.2)	Embedded
Metadata		
Metadata type	What kind of metadata are included in the text collection? (cf. Catalogue 3.2.3)	Descriptive, Structural, Administrative
Metadata level	On which level are the metadata included? (cf. Catalogue 3.2.2)	Individual texts
Data schemas and standards		
Schemas	What kind of data/metadata/annotation schemas are used for the text collection? (cf. Catalogue 3.2.4)	Customized standard schema
Standards	Which standards for text encoding, metadata and annotation are used in the text collection? (cf. Catalogue 3.2.4)	other: EpiDoc, a subset of TEI
Provision		

Accessibility of the basic data	Is the textual data accessible in a source format (e.g. XML, TXT)? (cf. Catalogue 4.1)	yes
Download	Can the entire raw data of the project be downloaded (as a whole)? (cf. Catalogue 4.2)	yes
Technical interfaces	Are there technical interfaces which allow the reuse of the data of the text collection in other contexts? (cf. Catalogue 4.2)	SPARQL endpoint
Analytical data	Besides the textual data, does the project provide analytical data (e.g. statistics) to download or harvest? (cf. Catalogue 4.3)	no
Reuse	Can you use the data with other tools useful for this kind of content? (cf. Catalogue 4.4)	yes
User Interface		
Interface provision	Does the text collection have a dedicated user interface designed for the collection at hand in which the texts of the collection are represented and/or in which the data is analyzable? (cf. Catalogue 5.1)	yes
User Interface questions		
Usability	From your point of view, is the interface of the text collection clearly arranged and easy to navigate so that the user can quickly identify the purpose, the content and the main access methods of the resource? (cf. Catalogue 5.3)	yes
Access modes		
Browsing	Does the project offer the possibility to browse the contents by simple browsing options or advanced structured access via indices (e.g. by author, year, genre)? (cf. Catalogue 5.4)	yes
Fulltext search	Does the project offer a fulltext search? (cf. Catalogue 5.4)	yes
Advanced search	Does the project offer an advanced search? (cf. Catalogue 5.4)	yes
Analysis		

Tools	Does the text collection integrate tools for analyses of the data? (cf. Catalogue 5.5)	no
Customization	Can the user alter the interface in order to affect the outcomes of representation and analysis of the text collection (besides basic search functionalities), e.g. by applying his or her own queries or by choosing analysis parameters? (cf. Catalogue 5.5)	yes
Visualization	Does the text collection provide particular visualizations of the data? (cf. Catalogue 5.6)	no visualization
Personalization	Is there a personalisation mode that enables the users e.g. to create their own sub-collections of the existing text collection? (cf. Catalogue 5.7)	no
Preservation		
Documentation	Does the text collection provide sufficient documentation about the project in general as well as about the aims, contents and methods of the text collection? (cf. Catalogue 6.1)	yes
Open Access	Is the text collection Open Access? (cf. Catalogue 6.2)	yes
Rights		
Declared	Are the rights to (re)use the content declared? (cf. Catalogue 6.2)	yes
License	Under what license are the contents released? (cf. Catalogue 6.2)	CC-BY-NC
Persistent identification and addressing	Are there persistent identifiers and an addressing system for the text collection and/or parts/objects of it and which mechanism is used to that end? (cf. Catalogue 6.3)	Persistent URLs
Citation	Does the text collection supply citation guidelines? (cf. Catalogue 6.3)	no
Archiving of the data	Does the documentation include information about the long term sustainability of the basic data (archiving of the data)? (cf. Catalogue 6.4)	no

Institutional curation	Does the project provide information about institutional support for the curation and sustainability of the project? (cf. Catalogue 6.4)	yes
Completion	Is the text collection completed? (cf. Catalogue 6.4)	no
Personnel		
Editors	Joshua Sosin	
Contributors	James Cowey Hugh Cayless Ryan Baumann Gabriel Bodard Roger Bagnall	