Achtung!
Dies ist eine Internet-Sonderausgabe des Kapitels
„Catalogues and Cataloguing of Oriental Manuscripts in The Digital Age“
von Jost Gippert (2014).
Sie sollte nicht zitiert werden. Zitate sind der Originalausgabe in
Comparative Oriental Manuscript Studies. An Introduction
Hamburg: COMSt 2015, 531–537
zu entnehmen.

Attention!
This is a special internet edition of the chapter
“Catalogues and Cataloguing of Oriental Manuscripts in The Digital Age”
by Jost Gippert (2014).
It should not be quoted as such. For quotations, please refer to the original edition in
Comparative Oriental Manuscript Studies. An Introduction
Hamburg: COMSt 2015, 531–537.

6. Catalogues and cataloguing of oriental manuscripts in the digital age (JG)

With the introduction of electronic means and methods into the cataloguing of manuscripts since the late 1980s, both the production of catalogues and their dissemination have begun to change considerably. As a matter of fact, the number of cataloguing projects concerning oriental manuscripts that are not digitally-based is steadily decreasing these days. A major difference in this context concerns the question whether the catalogues to be produced are still meant to be published in the 'traditional' way, that is in printed form, with digital means remaining restricted to preparatory functions, or whether the envisaged output is planned to be digital itself, that is online or via an electronic storage medium such as a DVD. If we leave the case of mere electronic typesetting (as a preliminary stage of printing) aside, both these aims usually build upon similar grounds in that they presuppose the conceptualization of database structures, but with different requirements concerning the scope, the granularity, and the retrievability of the data to be compiled, and with different prospects concerning their later usage. In the following sections, we will discuss the basics of database schemes and structures that are applicable to the cataloguing of manuscripts, the formats of electronic catalogues and their potential, and the challenges electronic cataloguing brings about.

6.1. Database schemes and structures

No matter whether a cataloguing project aims to result in book form or as an online website, databases that are meant to cover the information to be disseminated will have to address the basic issues of manuscript cataloguing as outlined in Ch. 4 § 1, namely (1) the manufacture of the manuscript and its physical aspects; (2) the contents of the manuscript; (3) the history of the manuscript after its making; and (4) bibliographical data related to both the physical and textual aspects of the manuscript. To what extent these data are compiled mostly depends on how detailed the catalogue is meant to be—anything between a mere 'inventory list' and a fully-fledged 'analytical' catalogue can be covered by the database structures of today, and as a matter of fact, database-like structures have even been underlying, explicitly or implicitly, many cataloguing projects that were initiated before the digital age. Thus, for example, the giant project of a 'Union Catalogue of Oriental Manuscripts in German Collections' of the Göttingen Academy of Sciences (*Katalogisierung der orientalischen Handschriften in Deutschland*, with the series *Verzeichnis* etc.) has been designed since its foundation in 1958 to follow a description pattern which comprises 26 sub-items, including shelf number, cover, material, state of preservation, page number, format, number of lines, writing style, decoration, scribe, date, origin, author, title of the work, quotations of the first and the last lines of the text as well as of colophons, further remarks and a number of registers (or concordances); and it should well be possible to continue the work 'by using the format of the newly developed database framework which was set up on the basis of the MyIHS system developed by Leipzig University in 2006 according to the conditions of GNU (General Public License)' as envisaged now (Raschmann 2012). The very fact that older catalogues usually possess an inherent database-like structure has been the reason why many digital cataloguing projects of today started by transforming the information stored in printed works or file cards into electronic data fitting into a database scheme, and 'digitally-born' cataloguing projects are not necessarily richer right from the beginning with respect to the data fields they contain, for example, the DOMLib / Ethio-SPaRe Manuscript Cataloguing Database project comprised a 'minimal data model' in its application phase (2009), consisting of 'Signature; Short title; Material; Measurements; Number of folia; Dating; Scribe; Author; Donor; Location original; Location current; Short contents; *Incipit*; Illuminations; *Additiones*; Further details; Bibliography' (the set has since been extended; see Nosnitsin 2012b).

However, database schemes developed for the cataloguing of manuscripts are anything but identical even today, and there are still some major differences in the database structures proper. First of all, it may be crucial for the database in question whether (1) the main concern of the compilers is the manuscript as an object of codicology (for example in establishing a catalogue of manuscripts of the same provenance); whether (2) they are rather interested, as philologists, in the texts contained in the manuscripts they intend to describe (for example in cataloguing all manuscripts containing the works of a given author, which may result in individual manuscripts being described only partially—an example of this is the online catalogue of the 'Handschriftencensus' project, <http://www.handschriftencensus.de/>, which aims at surveying the

manuscript transmission of German texts of the Middle Ages and which does not always mention texts in other languages present in the codices under consideration); or whether (3) they are primarily engaged in providing additional information ('metadata') concerning digitized images of the manuscripts they intend to make available. Differences may further be due to the scope of a given catalogue, depending, for example, on whether the object of cataloguing is the complete manuscript collection of a given country or repository, a special collection of a country or a repository determined by language, script, writing support, or age of the manuscripts, or manuscripts across different collections that share certain features. Noteworthy differences between databases may also arise from the level of consistency they require, especially with a view to how many fields must effectively be filled in, and, more importantly, how they have to be filled in; the existence of 'descriptive standards' (a good example—though in a poor design—for a descriptive standard usable for manuscript cataloguing is the Swiss 'Verbundkatalog Handschriften – Archive – Nachlässe (HAN)', <http://aleph.unibas.ch/F?con_lng=GER&func=file&local_base=DSV05&file_name=verbund-han>) may be crucial for later interoperability of the database (in the sense of data linkage and integration into portals and hypercatalogues, cf. below).

The most crucial question determining the structure of a database from the codicological point of view is to what extent and how it is able to reflect the different 'production units' that may be present in a given manuscript codex (cf. Ch. 4 § 4). To give but one example (kindly provided by P. Andrist): let us assume that part 1 of a given manuscript contains a text of Plato copied in the twelfth century, while part 2 of the same manuscript contains a text of Aristotle copied in the sixteenth century. It must be guaranteed in this case that the manuscript will not be 'hit' when executing a query for codices containing twelfth century copies of Aristotle's texts. The problem may even be more crucial if we take into account smaller production units such as reader and owner notes, as well as restoration and bindings, each with their own dating, and for the time being, no database (or search engine) seems to guarantee the contiguity of individual production units with the chronological information pertaining to them. This is due to the fact that relational database structures that were developed in the beginning of the digital age were mostly not flexible enough to adequately cover the structure of 'mixed' manuscripts consisting of several parts of different age, provenance, and/or content; but future generations, be they based upon SQL ('Structured Query Language', a widespread programming language to be used with relational databases) or XML ('eXtensible Markup Language', a more freely adaptable encoding system, cf. General introduction § 2.1) are likely to be able to overcome the problem. This is especially true for tree-like database structures using the XML-based 'MS Description' recommendation of the Text Encoding Initiative (TEI P5). However, it should be clear that this is not primarily a question of technology (relational databases vs. XML schemes) but of methodology, i.e. it depends on the overall design of the database.

6.2. Electronic catalogues and their potentials

With the establishment of digital databases containing the data to be catalogued, and especially with the development of the World Wide Web offering online access to them, it has become more and more tempting to disseminate the catalogued information on manuscripts electronically instead of or alongside a printed form, and the number of 'electronic catalogues' that are accessible online is steadily increasing today. Among the main advantages of this type of catalogues, we may mention, (a) that the data they contain can be retrieved dynamically via interactive search entry forms or the like, instead of static indexes (we may neglect here the mere electronic reproduction of a printed catalogue, for example, in PDF form, as this does not imply peculiar retrieval facilities other than sequential searches in the text); (b) that they can be linked to any other kind of electronic content in the web such as, for example, other catalogues, image repositories, or bibliographical materials; (c) that they can be steadily corrected, updated, extended and expanded. A few examples may suffice to illustrate these advantages.

(a) Search facilities and human interfaces

To access and retrieve the information contained in databases, it is necessary to provide human interfaces that allow for interactive communication between the user and the database in terms of search queries. Depending on the complexity of the database structure, search interfaces can be more or less sophisticated; in the maximum case, the search form may offer as many fields for the entry of query items as the database structure contains. This is true, for example, for the query forms offered by *SfarData*, the codicologi-

cal database of the Hebrew Paleography Project run by the Israel Academy of Sciences and Humanities, which provides, among others, seven check boxes to choose a given subset of data (styled 'corpus': for example, 'documented dated', 'documented undated with identified scribe'); nine select boxes to choose a 'major area', 'specific region', 'country', 'city', etc.; several input boxes to enter search terms like 'catalog no.', 'microfilm at IMHM', or 'words in colophon'; and a large set of additional entry points for searching the names of scribes, illustrators, and others (note that the site <http://sfardata.nli.org.il> is still under construction as of October 2014).

In accordance with recent practice developed for the retrieval from electronic catalogues of libraries, there is a strong tendency today to facilitate the entry of queries by enabling 'full text' searches, i.e. searches across many or even all the fields of the database without having to specify the fields in question. This has been implemented, for example, by *e-codices*, the Virtual Manuscript Library of Switzerland, which offers a facility to 'search in manuscript descriptions' right on its start page, thus allowing the user, for example, immediately to find 164 catalogue entries containing the word *parchment* among the 981 manuscripts catalogued in the database (<http://www.e-codices.unifr.ch/>). Even with this comfortable 'Google-like' search facility, users will have to be aware of possible differences in the interpretation of their queries; this is true, for example, for the entry of query strings consisting of more than one word: a search for *parchment codex* may be taken by the search engine to represent the two words *parchment* and *codex* individually, delivering all instances of descriptions that contain at least one of them (*parchment* 'or' *codex*), or it may expect that both words must both be present in a given description (*parchment* 'and' *codex*; this is the assumption of the *e-codices* search engine, which delivers 68 occurrences for the 'collocation' of the two words). In order to search for *parchment codex* as a two-word unit, it may be necessary to enter the string in quotation marks, again in accordance with 'Google-like' practices (this is true for *e-codices*, which delivers 6 occurrences of 'parchment codex' in its database).

b) Data linkage, portals and hypercatalogues

The possibility of linking data up with materials residing elsewhere in electronic form is with no doubt a major advantage of the digital age. In the case of electronic catalogues of manuscripts, this opens a wide range of hitherto unavailable functions that may be styled revolutionary indeed. For example, while printed catalogues normally contain but a few sample images of the manuscripts they describe or even no images at all, online catalogues may offer links to digital images of every single page of a given codex, provided they are accessible via an internet address (URL = Uniform resource locator). In the same way, the content description of a given manuscript can be linked up with a full representation of the text(s) it contains within an online edition or an electronic corpus, and bibliographical data in a catalogue can be linked up with digital representations of the works in question (for example, a first account of a manuscript in a traveller's report, a former catalogue, or a scientific treatise concerning the codex). It goes without saying that all these links can be established in both directions, thus opening wide perspectives of cooperation for cataloguers with philologists and other people dealing with manuscripts.

Another great advantage of the linkability of online data consists in the possibility to put the information contained in various catalogues together, for example with respect to manuscripts of the same provenance that are stored in different repositories or even countries, or codices with similar textual content spread over the world. In this connexion, we can distinguish two different approaches which may be covered by the terms 'portal' and 'hypercatalogue', albeit the two terms are not always sharply differentiated. Portals in the proper sense are web sites that offer more or less structured access to a set of related web sites. Different from simple link-lists which contain nothing but a collection of links to internet addresses (cf. <http://titus.fkidg1.uni-frankfurt.de/curric/comst/links.htm> for the COMSt-internal link-list as of February 2010, last accessed June 2014; another example concerning manuscripts is <http://palaeography_training.bangor.ac.uk/paleo.php>, last accessed June 2014), portals may be able to combine the search facilities of the sites they are linked to, either by collecting their data via 'data harvesting' and storing them locally, i.e. on the portal site itself, or without doing so; this can be achieved via so-called protocols which allow for inter-machine communication, distributing a query across the sites involved and gathering the query results 'on the fly'; an example of a portal working with both methods is the portal of the *Consortium of European Research Libraries* (CERL; <http://www.cerl.org/web/en/resources/cerl_portal/organization>; the CERL portal operates with the OAI-PMH protocol for data harvesting

('Open Archives Initiative' Protocol for Metadata Harvesting) and the Z39.50 protocol for data access 'on the fly'), which is mostly dedicated to manuscripts and early printed material of western provenance and which assembles the data from twelve online catalogues (or, to be more correct, the databases underlying them) to date. Unlike the 'portal' approach of this type, a 'hypercatalogue' would be more restricted thematically, bringing together the contents of different individual catalogues that have common objects such as, for example, a certain language, a script, a country, a time-span, or the like; one example of this is the *Pinakes* database hosted at the Institut de recherche et d'histoire des textes in Paris, which aims at providing a 'census of all Greek texts from the beginning up to the end of the sixteenth century which are contained in manuscripts described in printed catalogues, with the exception of papyri' ('*Pinakes* a pour objet principal le recensement de tous les textes grecs, des origines à la fin du XVI^{e} siècle, contenus dans les manuscrits décrits dans les catalogues imprimés de bibliothèques à l'exception des papyrus', cf. <http://pinakes.irht.cnrs.fr/>; the project was initiated as early as 1971 at the Pontifical Institute of Mediaeval Studies in Toronto and moved to Paris in 1993). In contrast to this project, which is explicitly based upon printed information digitized manually, hypercatalogues that are based upon electronic catalogues presuppose the ability of intercommunication between servers and databases, with at least a small common set of metadata elements distinguishing, for example, fields containing information on textual content, authors, scribes, provenance, or the like; for this purpose, database developers may refer to several international standards that have mostly been designed for library catalogues: cf. the so-called 'Dublin Core' metadata set, which consists of 15 elements (Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, Rights; cf. <http://dublincore.org/documents/dcmi-terms/>, last accessed June 2014). More recently accepted standards are METS (Metadata Encoding and Transmission Standard), an XML schema for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, and MODS (Metadata Object Description Schema), a schema for a bibliographic element set that may be used for a variety of purposes.

c) Dynamic data

Unlike printed catalogues, electronic catalogues can be steadily corrected, updated, extended and expanded, provided they are built upon a database. As a matter of fact, many cataloguing projects of today make their data accessible long before they are completed, not only for better visibility but also in order to enable users to provide feedback. For portals and hypercatalogues, the resulting flexibility of the data they rely upon has a strong impact in that it forces them continuously to keep track of changes. This strongly speaks in favour of accessing the data 'on the fly' instead of data harvesting, because the latter method yields a snap-shot of a given moment that may then be perpetuated longer on the harvesting server than in its source.

6.3. Challenges and problems of electronic catalogues

The problem of the 'dynamicity' of digital data as outlined above is not the only challenge that compilers and users of electronic catalogues have to cope with. A few other problems that may be crucial will be treated below; they are related to the reliability of query results, to questions of authorship, and to the maintenance of data.

a) Reliability of query results

As was stated above, one advantage of electronic catalogues consists in the fact that the data constituting them are accessible from various perspectives via more or less sophisticated search engines. However, it may be questionable whether or not the data retrieved in this way are reliable, especially in the case of hypercatalogues and portals. The reason is that the more complex a database structure is, the more it has to be consistent in the data it contains, and the more its developers have to take care for users not to be misled. To give but two examples: a) as was shown above, the *e-codices* project offers easy access to the information contained in the manuscript descriptions it has stored, via a 'Google-like' full-text search. However, the user should be aware that the manuscript descriptions that are accessible in this way are written in several languages (German, French, English) so that a search for *parchment* will yield different results (164 hits) from a search for *Pergament* (452 hits) or *parchemin* (129 hits), and even the abbreviated form *Parch.* gives 60 hits. To find the information that the *e-codices* database comprises a total of 791 manuscripts whose support is parchment (including 2 mixed codices with papyrus, and 6 with paper), the user

has to move to the main search page instead (<http://www.e-codices.unifr.ch/en/search/all>); b) under the title of *e-corpus*, the Centre de Conservation du Livre (CCL) in Arles is establishing, together with several partners, 'a collective digital library that catalogs and disseminates numerous documents: manuscripts, archives, books, journals, prints, audio recordings, video, etc.' (<http://www.e-corpus.org/>). Structured in the form of a portal, it provides access to several 'virtual collections', among them the complete collection of digitized microfilms of the Georgian manuscripts from Mount Sinai hosted at the university library of Louvain-la-Neuve in Belgium, an invaluable tool indeed. The collection can be accessed by entering *géorgien* in the 'Google-like' query field on the start page ('search in all collections'); however, in the query result, the 97 codices in question are arranged without any ordering rule being discernible, beginning with Sin.georg. 6, 8, 11, 16, 15 (<http://www.e-corpus.org/search/search.php?search=search&page=1&q=g%C3%A9orgien&search=Search>). Another result will be achieved by selecting *Bibliothèque de l'Université Catholique de Louvain, Institut orientaliste* as the 'Location' and *Georgian* as the 'Language' in the (main) search form (<http://www.e-corpus.org/search/index.php>): this, however, yields only 21 items, again starting with Sin.georg. 6 but continuing with Sin.grec 230 and 566. Searching for *Georgian* instead of *géorgien* via the 'Google-like' form yields no results from the collection at Louvain at all but, among others (of a total of 74 'hits'), a quotation from Tommaso Vallauri's *Latinae exercitationes grammaticae et rhetoricae studiosis propositae* (1869) about Virgil's *Bucolica et Georgica*. Note that a more or less numerical arrangement of the 97 items can be accessed via <http://www.e-corpus.org/eng/ref/96559/Sinai_Mf_UCL_Georgiens/> (last updated 7 September 2011), but this is not searchable.

b) Authorship and authors' rights

As with all other kinds of digital publications that are accessible online, compilers of electronic catalogues may be concerned about their authorship being protected well enough to prevent theft or misuse of their work. As everyone knows, not every change in a database is correct, and there are cases where the modified data is less correct than the previous one. Therefore it is very important that the name of the person who enters a given item of information into the database and who takes the 'scientific responsibility' for it to be secured, even across several stages of development of the database; in other words, if the compilers want to keep their author's rights guaranteed when being read and quoted by others, they should take care to place their authorship 'seal' on every single description that might be accessed and displayed separately (for example via a search function); depending on the retrieval engine used, this might also be generated automatically. This implies that regulations must be agreed upon between the compilers of electronic catalogues and the institutions hosting and publishing them concerning the preservation of the former's rights and concerning the question of whether, how and to what extent the institution may have the data modified or altered by others. And users of electronic catalogues should be sure to name the authors of the catalogued information they refer to, provided this is correctly 'signed'.

In addition, authors of electronic catalogues may have to face problems with 'scholarly referencing'. If, for example, in an article, someone refers to information found in a printed catalogue, the information is easily verifiable in the sense that anyone can check whether the person referring to it did, or did not, quote it correctly. When referring to information given in an electronic catalogue, there may be a discrepancy between a given quotation and what is found in the catalogue online at the time someone else accesses it; in this case it remains unclear if the quoting person did not excerpt it correctly or if the catalogue (output) was changed after the quotation was executed. It is therefore essential that quotations from online catalogues contain an indication of the exact date when the information was retrieved (in quite the same way as we should do with other online publications). It would further be desirable that future generations of databases to be used for electronic cataloguing provide a function for 'tracking changes' (in a similar way as Wikipedia does).

The safeguarding of authors' rights may also be crucial for the compilers of electronic catalogues themselves. In general, authors' rights are much harder to maintain in the globalized digital world than in the universe of printed matters. One reason is that the rights differ considerably from country to country, a fact that can easily be illustrated by the policy of the 'Google Books' project, which aims at digitizing all printed books of all times and places. Users of the project web site (<http://books.google.com/>) will realize easily that it is mostly older books that are accessible to them for full download (in PDF form), in accordance with copyright regulations. However, the terms of accessibility are different for users from the

USA and users from Europe: for US citizens (or, rather, users with a US-American IP address), Google will provide free access to books that were published before 1923; for Europeans (i.e. users with a European IP address), the borderline is the year 1871. This implies, for example, that the *Catalogue of the Hebrew Manuscripts Preserved in the University Library, Cambridge* (I) published by Salomon Marcus Schiller-Szinessy in 1876 is freely downloadable from the USA but not from Europe (<http://books.google.de/books?id=GEMPAAAAIAAJ>). This procedure is not only disturbing but also far from being legally exact, for no European country has a copyright regulation that would determine 1871 as a borderline year. Instead, copyright regulations usually prescribe that a given work be protected for a certain amount of time after the death of the author(s); in Germany, this is a period of 70 years. Obviously, Google deems it unnecessary to investigate authors' lives in order to determine the exact end of the protection period of a given work. I may quote here from an e-mail by Mr Jon Orwant (Google Inc.) of 2 May 2011: 'Each nation has their own copyright rules. Within the US, we are usually able to use 1923 as a cutoff date for determining whether books are in copyright, and so some of the seven books you identified are fully readable and downloadable inside the US. Outside the US we have to use the rules of the appropriate country. … We are able to make the PDFs available within a country when the books are out of copyright in that country, and when we do they're available directly from Google Book Search; if you had a US IP address you could just visit Google Book Search and download the PDF with no involvement from us. But because copyright status is often hard to determine, we have settled on the following rule for countries that don't have a cutoff like the US: either the book must have been published before 1871, or there must be clear and convincing evidence that all authors of the book died more than 70 years ago…' (in the given example, this would have been as early as 1960 as Salomon Marcus Schiller-Szinessy died in 1890, cf. <http://en.wikipedia.org/wiki/Solomon_Marcus_Schiller-Szinessy>, last access May 2014; by the way, the book can be freely downloaded from the 'Internet Archive' at <http://www.archive.org>, a Canadian site that does not block European users, see < https://archive.org/details/cataloguehebrew01schigoog>, last access October 2014).

Be that as it may, compilers of electronic catalogues who wish to include data from other works, either as quotations or via links to online resources, should treat these in the same way as they would do with quotations in a printed book.

c) Maintenance and longevity of data

In contrast to printed books, digital media are often regarded as unstable and therefore uncertain. It is true, for example, that most data carriers of today (for example, CDs, DVDs, hard disks) have only a limited 'lifetime' during which the data they contain remain both unaltered and retrievable, and the question whether data compiled and stored today will still be interpretable in, say, fifty years' time also depends on the availability of matching software. The developers of electronic catalogues should therefore consider thoroughly how and where to publish them, and they should be aware of the necessity to maintain the data by regularly adapting them to upgraded soft- and hardware and the like.

One important prerequisite for this is the usage of international standards. In the field of oriental languages, this is first of all a matter of the encoding of scripts and characters. We are in the lucky position today to be able to apply the Unicode standard for this (<http://www.unicode.org>; cf. General introduction § 2.1), which comprises nearly all scripts and characters of the languages involved, both modern and ancient, thus enabling us even to present original scripts and Roman transcriptions side by side in one document. In comparison with the chaos produced by the mapping of proprietary fonts in the 1980s and 1990s, this is a huge achievement leading towards long-time interpretability of our data.

Another important issue is the standards to be applied in the structure of databases and their output. Here, too, proprietary formats should be avoided right from the beginning. Any solution that is based upon XML structures will be preferable to relational databases as they provide greater flexibility and, what is more, transformability into usual output schemes, for example as HTML-based web pages; however, relational databases still tend to be more homogeneous than XML bases, and the quality of XML bases largely depends on the skills of the person who tags the data. Given the huge amount of data that is already available online, it is anything but probable that the basic features of web output will change dramatically in the future; in other words, both the use of Unicode character encoding and of 'markup languages' of the HMTL / XML type are likely to survive long enough so that electronic catalogues of today should build upon them.

References

Andrist 2010, 2014; Nosnitsin 2012b; Raschmann 2012. Web sources: CERL <http://www.cerl.org>, last access November 2014; Dublin Core <http://dublincore.org/documents/dcmi-terms>, last access November 2014; *e-codices* <http://www.e-codices.unifr.ch>, last access November 2014; *e-corpus* <http://www.e-corpus.org>, last access November 2014; Google Books <http://books.google.com/>, last access November 2014; Handschriftencensus <http://www.handschriftencensus.de/>, last access November 2014; Internet Archive <https://archive.org/>; Palaeography Training <http://palaeography_training.bangor.ac.uk/paleo.php>, last access November 2014; *Pinakes* <http://pinakes.irht.cnrs.fr>, last access November 2014; *SfarData* <http://sfardata.nli.org.il>, last access November 2014; TEI; TITUS <http://titus.fkidg1.uni-frankfurt.de/curric/comst/links.htm>, last access November 2014; Verbundkatalog Handschriften – Archive – Nachlässe (HAN) <http://aleph.unibas.ch/F?con_lng=GER&func=file&local_base=DSV05&file_name=verbund-han>, last access November 2014; Unicode <http://www.unicode.org>, last access November 2014.