LIBER QUARTERLY 19 (3/4), February 2010 – ISSN: 1435-5205. P259–274
http://liber.library.uu.nl/
Igitur, Utrecht Publishing & Archiving Services

Persistent Identifiers: the 'URN Granular' Project of the German National Library and the University and State Library Halle

Dorothea Sommer

Deputy Director, University and State Library Sachsen-Anhalt,
06098 Halle (Saale), Germany,
dorothea.sommer@bibliothek.uni-halle.de

This is a slightly revised version of the author's presentation at the LIBER-EBLIDA Digitisation workshop, The Hague, October 19–21, 2009. The slides can be found at http://www.libereurope.eu/files/Tue_Access_d)DS-DE.pdf.

Abstract

This paper describes a project carried out by the German National Library, the University and State Library Sachsen-Anhalt in Halle and Semantics GmbH Aachen to establish routines of persistent identification for individual pages of web publications in order to enable and facilitate reliable and long-term valid citation practices for the academic community.

The project originated in a pilot project to digitise approximately 10,000 German imprints from the seventeenth century comprising altogether about 600,000 pages, which had to be completed within two years. The material of the 'Ponickau Collection' had been catalogued in the German national bibliography of seventeenth-century imprints (VD 17), which was enriched and turned into a virtual library by adding the digitised texts. This article investigates the means of presentation and indexing of digitised imprints in order to ensure their usability. It also sketches the workflow among the various partners involved in the process. The article highlights the application of Visual Library, which contains various tools for automated creation of metadata, the implementation of persistent identifiers (URN) and the automated enrichment of catalogue entries by the regional cataloguing cooperative Gemeinsamer Bibliotheksverbund (GBV). Special emphasis is given to questions of quality management; the quality is guaranteed by a combination of automated tools and intellectual control at various stages of the digitisation process.

Key Words: persistent identifiers; uniform resource names; digitisation; citation practices

1. Project Context

A major challenge facing libraries in the next years is the creation of an integrated information environment by facilitating access to digitised materials. In order to provide broad access to digitised materials institutions increasingly rely on mass digitisation. In his book *Information Foraging Theory* Peter Pirolli, who researches human-computer interaction in order to assess how information environments can best be shaped for people, addressed one of the pitfalls of the interaction with information through technology when he said: 'In an information-rich world, the real design problem to be solved is not so much how to collect and distribute more information but rather how to increase the rate at which persons can find and attend to information that is truly of value to them.'[1] He also states: '…people prefer information-seeking strategies that yield more useful information *per unit*…People prefer, and consequently select, technology designs that improve returns on information foraging'.[2]

These two statements highlight the general context of the URN Granular Project. URN Granular was developed in order to provide an opportunity for users to address and quote electronic resources or web publications in an exact and persistent way. The necessity to deal with this issue arose during the course of the work on the Ponickau project, which is one of the first four mass digitisation projects which was supported by the German Research Foundation within the VD16/VD17 programme line. This programme line aims at enriching the national German bibliographies of 16th- and 17th-century prints with digital content.[3]

Between July 2007 and 2009 the University and State Library Sachsen-Anhalt digitised 10,147 imprints of the seventeenth century from the Ponickau Collection comprising around 600,000 pages, and it developed a web presentation for the collection.[4]

The original collection itself is a special collection which belonged to the old and famous Wittenberg University, but was transferred to Halle university library after the university in Wittenberg was closed by Napoleon Bonaparte in 1813. Wittenberg University, founded in 1502, and the 'Fridericiana', the Friedrichs University in Halle, founded in 1694, were merged in 1817. The collection was initially a donation to Wittenberg University by the Saxon nobleman Johann August von Ponickau, who lived from 1718 until 1802. With its precious literary and historical sources it covers not only the region of the former Prussian province of Saxony (approximately the region of Sachsen-Anhalt today), but also the regions of Saxony and Thuringia. The collection contains original editions of Martin Luther and his contemporaries, resources about the history of the universities of Leipzig, Wittenberg, Jena, Halle etc., and a lot of other regional occasional literature.

2. The Need for Persistent Addressing of Web Resources

As the digitisation project was part of the digital information initiative of and financially supported by the German Research Foundation, the best practice requirements issued by the German Research Foundation[5] had to be complied with. These guidelines contain principles for planning digitisation projects, such as the selection of material, technical requirements for image production, full-text recognition, long-term preservation and metadata. They provide standards for the collection of structural metadata and the provision of electronic documents, including a statement on the importance of open access, and they and specify presentation standards and formats. Also, the guidelines recommend application of effective and cost-conscious methods which can be applied systematically to large amounts of material. One of the requirements stipulates that the usability of web publications must be furthered by providing opportunities for persistent citation of digital texts. The guidelines read:

> 'When digitisation was in its infancy, the issue of citability of digital resources was frequently underestimated. But it is exactly citability that makes Internet-based digitised resources suitable for academic research. Other than older secondary formats, such as microfilm or paper print-outs, an Internet resource is not just a copy of the original which can be treated and hence cited

> like the original, but rather an independent object in a dynamic integral research space. Unlike traditional photocopies, digital copies require special citation rules if they are on the Internet… When a copy is online, it needs a unique address so that other documents or databases can link to it…. Two functionalities in particular are important for online presentation: the addressability of a work as a whole, and the addressability of individual pages or double pages within a work… In the future, individual physical pages will also have to be reliably accessible and citable. Institutions should implement suitable mechanisms (PURL, URN, DOI, Handle etc.) to ensure the persistence and linkability of a resource in order to provide reliable resources for scientific research'.[6]

URN Granular thus aims to establish routines for persistent identification of individual pages of web publications in order to enable and facilitate reliable and long-term valid citation practices for the academic community.

The guidelines on digitisation published by the German Research Foundation are not the only policy document demanding long-term secured access to digitally generated data. Several other documents reflect the increased call for open access to the output of publicly funded research and permanent access to primary quality-assured research data in Germany:

- According to the German Research Foundation Guidelines 'Safeguarding Good Scientific Practice'[7], recommendation no. 7: primary data on which a publication is based shall be stored for at least ten years in a durable form in the institution of origin in order to guarantee that research results are reproduceable.
- The Alliance of the German Science Organisations launched its new priority *Initiative Digitale Information* in June 2008[8], which aims at the establishment of an integrated digital research environment in the period 2008–2012. In the target area number 4, it identified the necessity to establish structures that enable the collection, archiving and the subsequent reuse of primary research data in all applicable disciplines.
- The European Science Foundation (ESF)/European Heads of Research Councils (EUROHORCS)[9] expressed their commitment to a vision of a European Research Area (ERA) with permanent access to primary quality-assured research data (Vision point no. 8).

3. Persistent Addressing of Web Resources: Reality

As is generally known, digital texts and publications allow for forms of use different from printed media. While there are many advantages to online publications, new technologies raise altogether new issues that have positive as well as negative implications. In reality one has to face the fact that digital data are fragile and ephemeral. They are increasing in kind and number, but they are prone to loss. Finally, they are dependent on a reliably functioning technical environment, which has to last the whole life cycle of a digital object.

Message 404 — a broken link on the internet — is well-known to all users of web applications. URLs are not stable and vanish when their address changes. Thus URLs are not suitable tools for referencing services. According to a study by Robert Dellavalle in the journal *Science,* up to 13% of online resources are irretrievable 27 months after their publication.[10] Another telling example is the following complaint by the Chinese poet Tammy Ho Lai Ming on the demise of small online publications on her blog:

> 'My poem *'Covers and Spines'* was published in issue #2 of the *Foundling Review,* a brand-new weekly online literary publication (debuted two months ago in May). It publishes 2 stories, 3–4 poems and a B&W or sepia photograph in each issue. I visit the website at least once a week to read the newly-selected creative pieces. It is highly worrying that the website hasn't been updated for two weeks…. I hope this does not mean that the publication is "dead". I hate to see online publications go. *Lily*, *HiNgE*, *Softblow*, *Postal Poetry*… the list goes on. All of a sudden the published works in those publications become orphans, URL-less….your pieces are in a limbo: published yet not published.'[11]

There are a number of options and strategies to avoid loss of information and orientation in the dynamic digital environment. The challenge is both of an organisational and a technical nature. One of the key words in this context is *security*. The digital world has its own so-called 'safe places' or 'domains of trust' — to which libraries hopefully belong. In Germany, the Deutsche Nationalbibliothek and the regional libraries of the different states are responsible for the collection of not only printed and other physical media, but also of internet-based documents published in Germany. As a consequence, copies of digital documents should not only be kept at the institution of origin, but should also be forwarded to the DNB for collection and long-term preservation purposes. The DNB provides the necessary technical infrastructure by maintaining a domain of trust as well as tools and routines that are needed for scientists and researchers to transfer their working methods from the printed world to a digital research environment. One of these tools is the application of persistent identifiers, which requires both a common policy among the institutions involved as well as a reliable technical infrastructure.

Persistent identifiers (PIs) are distinct and unique definers of digital objects. They signify the name, not the address or location of a digital object. They describe the resource itself in a global and persistent way. In contrast to storage addresses, PIs remain valid in case of technical or organisational changes. Part of the concept is a resolving service, which administers the names and addresses of all registered digital objects. Thus a digital object shall be identifiable and addressable throughout its life-cycle across all system boundaries and changes. Persistence is achieved by means of a policy, rules and defined procedures. It is not a characteristic in itself, but has to be developed and maintained. There are a number of systems of PIs and one of them is the Uniform Resource Name (URN). Tim Berners-Lee developed the concept of the Uniform Resource Identifier (URI)[12] as the only web naming/addressing technology in 1994.

4. Aim of the URN Granular Service

In the scholarly community the ability to cite resources is of critical importance. Whereas in the printed world various rules and standards have long since been applied, digital resources lack practically validated working practices that allow not only for citation of a work as a whole, but also for citing its individual pages. Correct citation implies that a work and its pages can be addressed reliably, long-term and in a sustainable and persistent way.

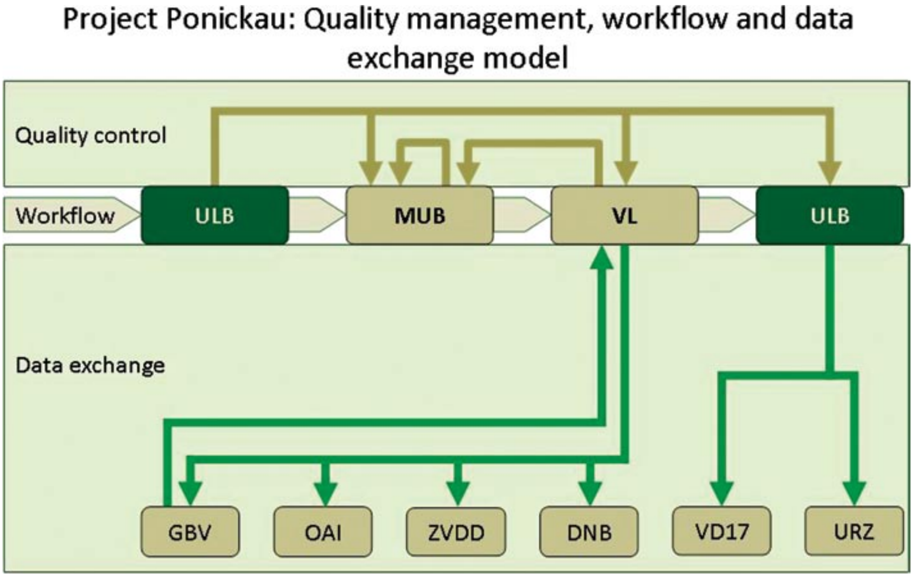
Such citations should also apply to text fragments and logical entities of texts such as entries of an encyclopaedia or single articles within newspapers. In fact, the whole document and all its constituent parts, including numbered and unnumbered pages, cover etc. should be included in the modus of citation. This is of particular importance in the field of retro-digitisation of prints from the 16th until the 18th century (or later), since digitisation not only enables researchers from all over the world to access the material, but also protects the locally kept originals from further damage.

Like the Deutsche Nationalbibliothek, the University and State Library Halle has decided to use Uniform Resource Names (URNs) as persistent identifiers for addressing their web publications. These web publications include not only the digital copies from the Ponickau Project, but also digitised prints from other projects of the 16th or 18th century, regional deposit literature and retro-digitised papers (including their articles) from our special subject collection on the Middle East & North Africa. The wealth of the material has substantial implications for the technical implementation processes in terms of scale and diversity.

5. The Ponickau Project: Workflow Implications and Quality Management

An important prerequisite for the URN Granular Project was the establishment of a clearly defined workflow within the digitisation project itself — in particular with regard to time-slots for the harvesting processes — and rigid quality management. It should be mentioned here that the project could only

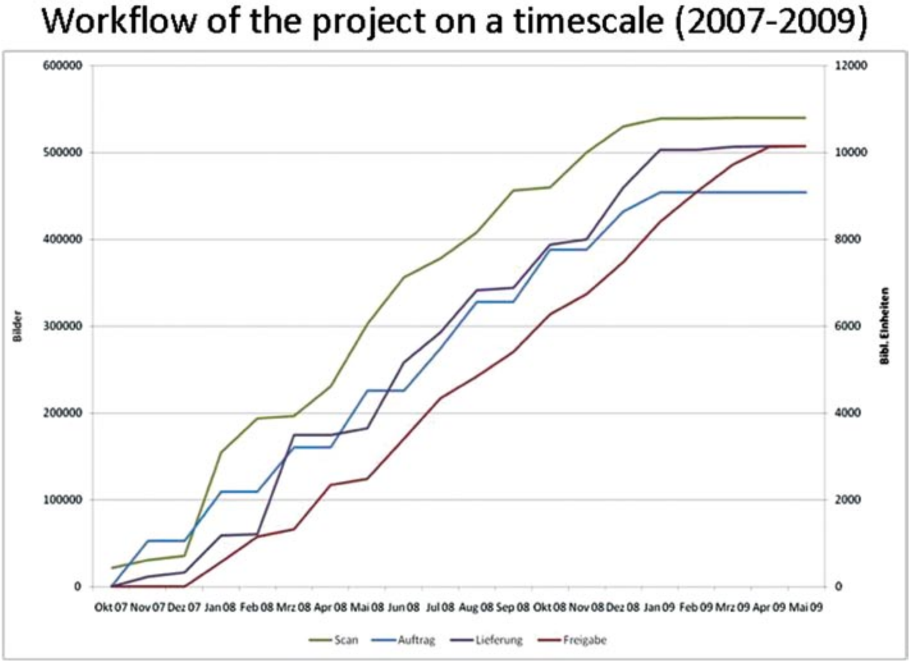
Fig. 1: Shows the workflow of the project with the various partners involved.


Project Ponickau: Quality management, workflow and data exchange model
Quality control
Workflow
ULB
MUB
VL
ULB
Data exchange
GBV
OAI
ZVDD
DNB
VD17
URZ

be realised within the given time-frame because of highly automated workflows. The basis for the processes was provided by high-quality data from the National Bibliography VD 17[13], which the document management system Visual Library automatically converted into formats such as Dublin Core[14], METS[15]; MODS[16] and MARC XML[17] so that an OAI-interface could be served. In the course of the project more than 10,000 additional bibliographic entries were automatically generated for the online resources, in co-operation with the Common Library Network (GBV).

Quality management was of crucial importance for the project. The URN can only be assigned when all quality-assured processes, which were carried out both in an automated mode as well as by intellectual control, are fin-

Fig. 2: Depicts the digitisation process on a time scale. Four distinct processes are identified. They include the preparations for scanning (University and State Library/ULB, in-house), the scanning process (MikroUnivers Berlin/MUB, first service provider), conversion and data processing (Visual Library by Semantics, second service provider), and the final steps of structuring, quality control and online publication (ULB, in-house).



ished. A variety of problems which are likely to occur have to be addressed before the actual start of the harvesting process. They can have their origin in incorrect sequences of images, insufficient colour management and incorrect data marking resulting from the structuring processes of web publications. Therefore it is necessary to implement a number of quality assurance procedures during the various stages of the workflow. The various measures of quality assurance encompass preceding specifications and regulations. They include continuous quality control measurements and a detailed check of the single steps and various partners of the digitisation process. Finally, feedback by the users has to be taken into account. Quality control during the project was carried out by using both automatically generated procedures as well as intellectual control measurements. Automated quality control was a very essential asset since a substantial amount of data had to be handled.

In the end all the URN-links to the digitised objects were integrated into the VD 17 and the databases of the Common Library Network (GBV)[18] as well as the Central Registry of Digitised Prints (zvdd)[19] in an automated procedure. Moreover, via the OAI-pmh interface they are visible for other search engines such as Google or Yahoo on the internet.

As statistics show, usage of the digital collection is very good. While in 2008 21,994 searches and 10,706 users were registered, in 2009 a step increase of these figures was noted. During the first six months 46,087 searches were carried out by 29,056 users who accessed 1,143,797 pages of web publications.

6. URN Strategy of the German National Library: Project Application

The initial approach to the URN Granular Project was to define a digital object as a unit that can be assigned a URN.[20] This unit currently refers to a static publication, like an online resource in monographic form.

The syntax of a URN is laid down in several requirements articulated by the Internet Engineering Task Force (IETF) in the Request For Comments (RFC) 1737 in 1994.[21] According to RFC 1737 the requirements for URN namespaces are: global scope, global uniqueness, persistence, scalability, legacy support, extensibility, independence and resolution. The document also mentions that the URN scheme is designed to facilitate the integration of existing number systems such as National Bibliography Numbers, ISSN or ISBN into the definition of a URN. At the same time, the URN syntax is open to the integration of standard number systems that are newly introduced by particular interest groups and communities. Following an initiative of the Conference of European National Libraries (CENL)[22] in 2001, the German National Library registered a namespace with the Internet Assigned Number Authority IANA[23], where namespaces are filed. The syntax of URNs depends on the choice of the number system:

- International Serial Number (RFC 3044) urn:issn[24]
- International Standards Book Number (RFC 3187) urn:isbn[25]
- National Bibliography Number (RFC 3188) urn:nbn[26], etc.

Like other national libraries, the German National Library can assign other trusted organisations their own subnamespaces. Trusted organisations in the context of the project are the Common Library Network (GBV) and the University and State Library Sachsen-Anhalt. Thus all digitised prints of the Ponickau Collection are registered with the German National Library as legal deposit materials. The syntax is as follows:

urn:nbn:de:gbv:3 — In this sequence of characters:

- urn is the scheme name,
- nbn is the name space identifier (nid) pointing to various systems of national bibliography numbers, registered by IANA,
- de is the country code based on ISO 3166, which signifies the German origin of the respective nbn,
- gbv is the first subnamespace identifier (snid) and stands for the Common Library Network (GBV), the library in Halle belongs to,
- 3 as second subnamespace identifier is the identification number of the University and State Library within the German library system.

The University and State Library Halle itself has defined various groups of subnamespace identifiers describing various projects and collections that are introduced and registered with the German National Library. The snid *1* stands for the digitised volumes of the Ponickau collection and other retro-digitisation projects, regional deposit literature receives snid *2*, snid *3* is assigned to materials originating from in-house digitisation on demand, snid *4* denotes electronic dissertations and snid *5* marks documents digitised for the National Middle East and North Africa collection at the University and State Library Halle. Consequently, the URN of a digitised copy of a document belonging to the Ponickau Collection carries the following string: urn:nbn:de: gbv:3:1. To this sequence a namespace specific string (niss) is added for each document: urn:nbn:de:gbv:3:1:xxx [niss].

The URNs are assigned to the digital objects in an automated process. Apart from the opportunity to cite web publications persistently, another advantage can be noted: that of redundancy. Access to a digital object is still guaranteed when the location of its storage changes. The transfer of the URNs from the University and State Library Halle to the DNB is accomplished through the OAI 2.0. protocol.

7. The Principle of Granularity

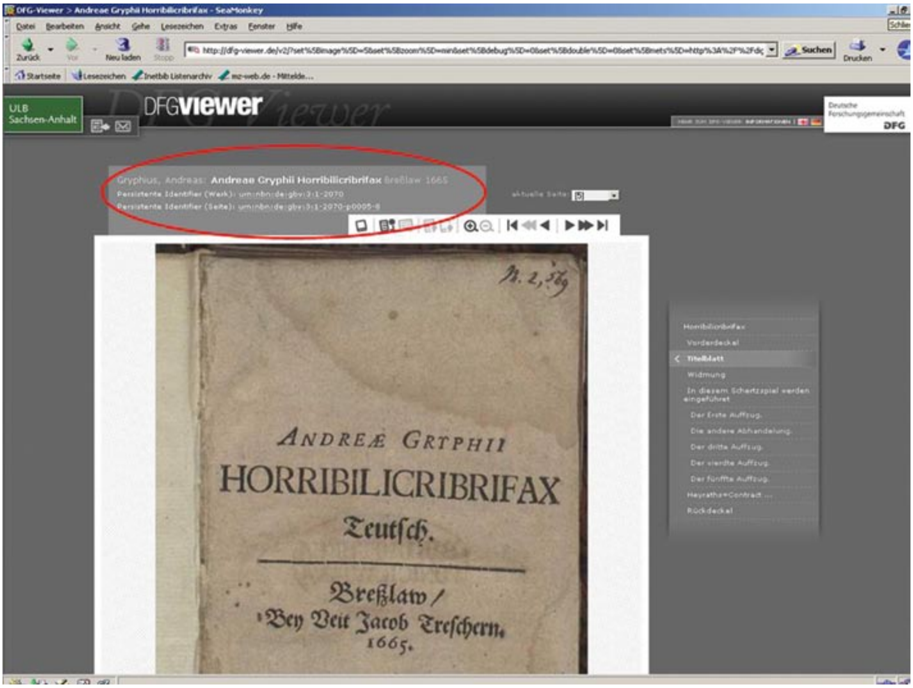
In order to facilitate addressing of single pages within a digital copy of a print[27], the logical structure of the syntax of the URN had to be followed and further extended. Following the principle that URNs can be assigned to complete digital objects or parts or versions of them[28], the units of the digital objects were divided into their granular parts on the basis of the X-Epicur Format.[29] The syntax of the URN now looks even more detailed, as it contains a reference to the pages (without semantic content):

- urn:nbn:de:gbv:3:1-xxxx-**p00001-3**
- <urn of the digital print>-p<physical pagination, four numbers|>-<check digit>

For example, the electronic version of the title *Andreae Gryphii Horribilicribifax, Breßlaw: Trescher, 1665* carries the URN urn:nbn:de:gbv:3:1-2070.[30] Its title page, which is preceded by four unnumbered pages, is identified as: urn:nbn: de:gbv:3:1-2070-p0005-8[31] (Figure 3).

At the work level each object receives a random consecutive number which does not have any correlation to already existing denominations, and this principle is repeated at the level of the pages. Thus the electronic version of the title *Verzeichnis und Zeigung des Hochlobwirdigen Heiligthumbs der StifftKirchen*

Fig. 3:



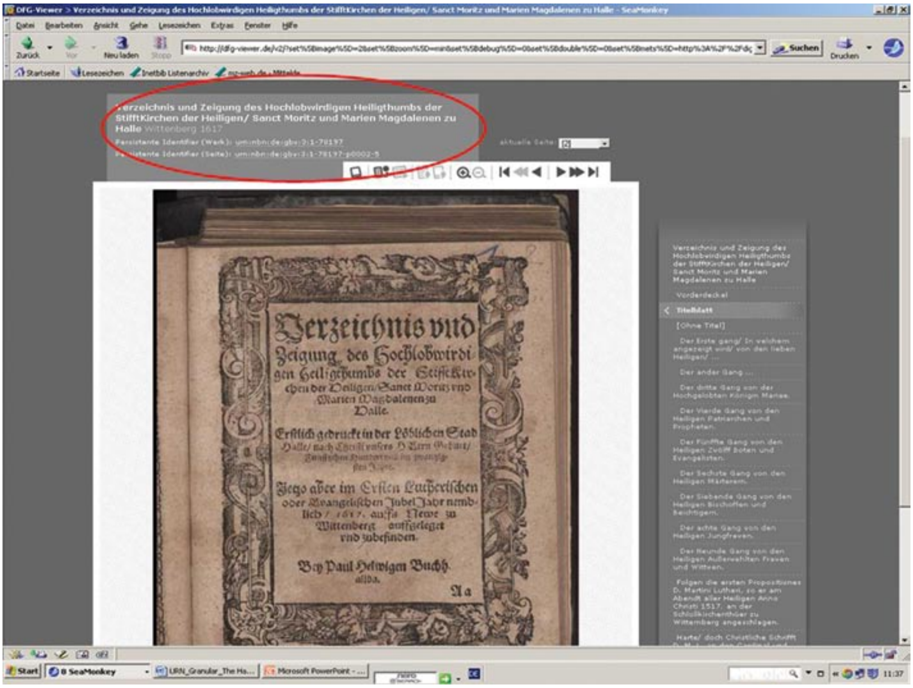
Gryphius, Andreas: **Andreae Gryphii Horribilicribrifax** Breßlaw 1665
Persistente Identifier (Werk): urn:nbn:de:gbv:3:1-2070
Persistente Identifier (Seite): urn:nbn:de:gbv:3:1-2070-p0005-6

der Heiligen/ Sanct Moritz und Marien Magdalenen zu Halle: Erstlich gedruckt in der Löblichen Stad Halle/ nach Christi unsers Herrn Geburt/Funfftzehen Hundert und im zwantzigsten Jahre Wittenberg: Helwig, 1617 (urn:nbn:de:gbv:3:1-78197)[32] contains a section on *'die ersten Propositiones D. Martini Lutheri, so er am Abendt aller Heiligen Anno Christi 1517. an der Schloßkirchenthüer zu Wittemberg angeschlagen'*.[33] In the print edition this section starts at page 100. The citation link to the online publication is urn:nbn:de:gbv:3:1-78197-p0101-5. Intendedly the page element does not correlate to the pagination of the original (Figure 4).

8. The Harvesting Process

Because a substantial flow of data had to be handled, an automated harvesting process was designed. Both during the project and presently, after other projects have been added, the harvester of the DNB directs queries to the data provider via OAI-PMH-2.0 twice a day. It automatically reg-

Fig. 4:



DFG-Viewer > Verzeichnis und Zeigung des Hochlobwirdigen Heiligthumbs der StifftKirchen der Heiligen/ Sanct Moritz und Marien Magdalenen zu Halle - SeaMonkey

http://dfg-viewer.de/v2/?set%5Bimage%5D=2&set%5Bzoom%5D=min&set%5Bdebug%5D=0&set%5Bdouble%5D=0&set%5Bmets%5D=http%3A%2F%2Fdc

Verzeichnis und Zeigung des Hochlobwirdigen Heiligthumbs der StifftKirchen der Heiligen/ Sanct Moritz und Marien Magdalenen zu Halle Wittenberg 1617

Persistente Identifier (Werk): urn:nbn:de:gbv:3:1-78197

Persistente Identifier (Seite): urn:nbn:de:gbv:3:1-78197-p0002-5

Verzeichnis und Zeigung des Hochlobwirdigen Heiligthumbs der StifftKirchen der Heiligen/ Sanct Moritz und Marien Magdalenen zu Halle

Vorderdeckel

Titelblatt

[Ohne Titel]

Der Erste gang/ In welchem angezeigt wird/ von den lieben Heiligen/ ...

Der ander Gang ...

Der dritte Gang von der Hochgelobten Königin Mariae.

Der Vierde Gang von den Heiligen Patriarchen und Propheten.

Der Fünffte Gang von den Heiligen Zwölff Boten und Evangelisten.

Der Sechste Gang von den Heiligen Märterern.

Der Siebende Gang von den Heiligen Bischoffen und Beichtigern.

Der achte Gang von den Heiligen Jungfrawen.

Der Neunde Gang von den Heiligen Außerwehlten Frawen und Witwen.

Folgen die ersten Propositiones D. Martini Lutheri, so er am Abendt aller Heiligen Anno Christi 1517. an der Schloßkirchenthüer zu Wittemberg angeschlagen.

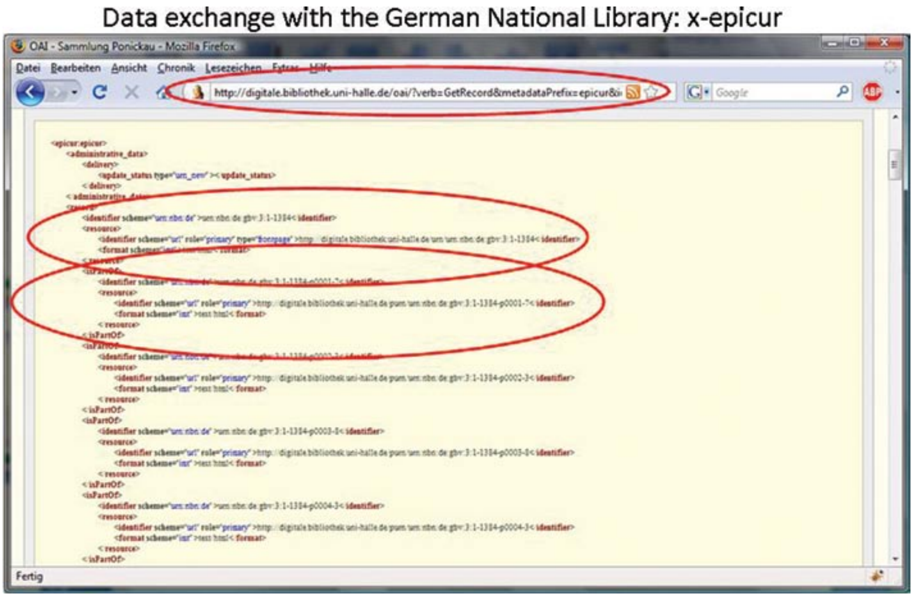
Harte/ doch Christliche Schrifft

isters those digital objects (including their pagination) which have been produced up to that point of time. There is no manual interference, neither by the German National Library nor by the library in Halle. Also the creation of URN-numbers has been fully automated. The resolver of the German National Library also allows for checks for consistency at specific time intervals (Figure 5).

9. Results and Outlook

The results are quite telling: within the Ponickau Project 10,147 digital copies were produced, which contain 582,293 pages. Altogether, 581,347 URNs were assigned both at the work level and at the level of single pages. The harvesting procedures operate regularly and are quite table. The process has now become a routine operation and is used for other projects as well, such as newspaper digitisation.

Fig. 5:



Data exchange with the German National Library: x-epicur

OAI - Sammlung Ponickau - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
http://digitale.bibliothek.uni-halle.de/oai/?verb=GetRecord&metadataPrefix=epicur&i
Google
Fertig

```
<epicur:epicur>
  <administrative_data>
    <delivery>
      <update_status type="urn_new" ></update_status>
    </delivery>
  </administrative_data>
  <record>
    <identifier scheme="urn:nbn:de" >urn:nbn:de:gbv:3:1-1384</identifier>
    <resource>
      <identifier scheme="url" role="primary" type="frontpage" >http://digitale.bibliothek.uni-halle.de/urn/urn:nbn:de:gbv:3:1-1384</identifier>
      <format scheme="imt" >text/html</format>
    </resource>
    <isPartOf>
      <identifier scheme="urn:nbn:de" >urn:nbn:de:gbv:3:1-1384-p0001-7</identifier>
      <resource>
        <identifier scheme="url" role="primary" >http://digitale.bibliothek.uni-halle.de/purn/urn:nbn:de:gbv:3:1-1384-p0001-7</identifier>
        <format scheme="imt" >text/html</format>
      </resource>
    </isPartOf>
    <isPartOf>
      <identifier scheme="urn:nbn:de" >urn:nbn:de:gbv:3:1-1384-p0002-3</identifier>
      <resource>
        <identifier scheme="url" role="primary" >http://digitale.bibliothek.uni-halle.de/purn/urn:nbn:de:gbv:3:1-1384-p0002-3</identifier>
        <format scheme="imt" >text/html</format>
      </resource>
    </isPartOf>
    <isPartOf>
      <identifier scheme="urn:nbn:de" >urn:nbn:de:gbv:3:1-1384-p0003-8</identifier>
      <resource>
        <identifier scheme="url" role="primary" >http://digitale.bibliothek.uni-halle.de/purn/urn:nbn:de:gbv:3:1-1384-p0003-8</identifier>
        <format scheme="imt" >text/html</format>
      </resource>
    </isPartOf>
    <isPartOf>
      <identifier scheme="urn:nbn:de" >urn:nbn:de:gbv:3:1-1384-p0004-3</identifier>
      <resource>
        <identifier scheme="url" role="primary" >http://digitale.bibliothek.uni-halle.de/purn/urn:nbn:de:gbv:3:1-1384-p0004-3</identifier>
        <format scheme="imt" >text/html</format>
      </resource>
    </isPartOf>
```

We hope that this practice of persistent identification contributes towards general efforts to establish new methods in representation and standardisation of academic citation practices in the digital world. Furthermore, it serves to assure free of charge, long-term, detailed access per *unit* — as required by Peter Pirolli — to important digital resources. Last but not least, we hope that the URN granular tool can be used as a model for other applications in the future.

Notes

[1] Peter Pirolli, *Information Foraging Theory*, Oxford University Press, 2007, p. 13.

[2] ibid, p. 14.

[3] http://www.dfg.de/forschungsfoerderung/formulare/download/12_152.pdf; http://www.vd17.de/sonst.html.

[4] http://digitale.bibliothek.uni-halle.de/pon.

[5] http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/praxisregeln_digitalisierung_en.pdf.

[6] ibid p. 17 u. 26.

[7] http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf, p. 55.

[8] http://www.mpg.de/bilderBerichteDokumente/dokumentation/pressemitteilungen/2008/pressemitteilung20080612/: 'Alle Wissenschaftsorganisationen sehen einen dringenden Handlungsbedarf für die systematische Sicherung von im Forschungsprozess erzeugten Daten. Auch hier gilt es, in den jeweiligen Disziplinen angemessene Strukturen und Prozesse zu etablieren, um Forschungsprimärdaten zu sichern, zu archivieren und für eine Nachnutzung bereit zu stellen.'

[9] http://www.esf.org/ext-ceo-news-singleview/article/eurohorcs-esf-commit-to-play-key-role-in-shaping-a-competitive-era-456.html. Vision point no. 8: 'Open access to the output of publicly funded research and permanent access to primary quality assured research data.'

[10] Dellavalle Robert P., Hester Eric J., Heilig Lauren F., Drake Amanda L., Kuntzman Jeff W., Graber Marla, Schilling, Lisa M.: 'Information science. Going, going, gone: lost Internet references'. *Science.* 2003 Oct 31; 302 (5646): 787–8, http://www.sciencemag.org/cgi/content/full/302/5646/787.

[11] http://tammyholaiming.com/2009/07/07/the-demise-of-small-online-publications/.

[12] Tim Berners-Lee (2005): Uniform Resource Identifier (URI) — Generic syntax. http://www.ietf.org/rfc/rfc3986.txt.

[13] http://www.vd17.de.

[14] http://dublincore.org/.

[15] http://www.loc.gov/standards/mets/.

[16] http://www.loc.gov/standards/mods/.

[17] http://www.loc.gov/standards/marcxml/.

[18] http://www.gbv.de.

[19] http://www.digitalisiertedrucke.de/search?p=ponickau.

[20] Christa Schöning-Walter: 'Der Uniform Resource Name'. In: *nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Version 2.0. Hrsg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, M. Jehn, Kapitel 9.4.1. Boizenburg: Verlag Werner Hülsbusch, 2009.

[21] http://www.w3.org/Addressing/rfc1737.txt; http://www.ietf.org/rfc/rfc2141.txt.

[22] http://www.cenl.org/.

[23] http://www.iana.org/assignments/urn-namespaces.

[24] http://tools.ietf.org/html/rfc3044.

[25] http://tools.ietf.org/html/rfc3187.

[26] http://tools.ietf.org/html/rfc3188.

[27] Dorothea Sommer, Christa Schöning-Walter, Kay Heiligenhaus: 'URN granular: Persistente Identifizierung und Adressierung von Einzelseiten digitalisierter Drucke', *ABI Technik* 28/2 (2008), p. 106–114.

[28] http://www.persistent-identifier.de/?link=3352&lang=en.

[29] See: http://www.persistent-identifier.de/?link=210&lang=en; The abbreviation EPICUR stands for Enhancement of Persistent Identifier Services.

[30] http://nbn-resolving.de/urn:nbn:de:gbv:3:1-2070.

[31] http://nbn-resolving.de/urn:nbn:de:gbv:3:1-2070-p0005-8.

[32] http://nbn-resolving.de/urn:nbn:de:gbv:3:1-78197.

[33] http://nbn-resolving.de/urn:nbn:de:gbv:3:1-78197-p0101-5.