

# **E-Periodica: The Platform for Digitised Swiss Journals**

#### **Regina Wanger**

ETH Zürich, ETH-Bibliothek, DigiCenter, Zürich, Switzerland regina.wanger@library.ethz.ch

#### **Michael Ehrismann**

ETH Zürich, ETH-Bibliothek, IT-Services, Zürich, Switzerland michael.ehrismann@library.ethz.ch



Copyright © 2016 by Regina Wanger and Michael Ehrismann. This work is made available under the terms of the Creative Commons Attribution 4.0 Unported License: <u>https://creativecommons.org/licenses/by/4.0/</u>

# Abstract:

*E-Periodica is an ETH-Bibliothek platform for digitised Swiss journals from the fields of science, history and culture. Since its launch in 2007, the service has been expanded continuously to include new publications. Due to the on-going growth in data and in order to keep the platform running in the long-term, E-Periodica's IT infrastructure underwent a complete overhaul. This makeover also included a redesign of its online presence with a focus on responsive design for mobile devices. This paper presents E-Periodica as a library's reliable open access service. It explains organisational and financial aspects and describes the work processes and central functions of the new-look website, based on the new design.* 

#### **Keywords:**

Digitised journals, journal platform, open access, responsive design, role based workflow

# Contents

1 Introduction	3
2 History	3
3 Organisation	4
3.1 ETH-Bibliothek's Overall Responsibility	4
3.2 Cooperation with the Swiss National Library	5
3.3 Contractual Partners	5
3.4 Development of the Platform	5
3.5 Inclusion Criteria	5
3.6 Legal Aspects	6
4 E-Periodica Processes	6
4.1 Negotiation	6
4.1.1 Cost Factors	7
4.1.2 Cost Model	7
4.2 Digitisation	7
4.3 Role-Based Workflow: Structuring	8
4.3.1 Data Provision	8
4.3.2 Recording the Journal	9
4.3.3 Loading the TIFF Files	9
4.3.4 Recording the Structure	10
4.4 Uploading	12
4.4.1 Internal Uploading	12
4.4.2 Uploading onto E-Periodica.ch	12
4.4.3 Communication, References, Searching	12
4.5 Creation of Archive Capsules	13
4.6 Architecture and Software	13
5 Online Presence	13
5.1 Homepage	14
5.2 Searching and View Options	15
5.2.1 Split View: Table of Contents and Page View	16
5.2.2 Full Screen and Zoom View	16
5.2.3 List of Hits	17
5.3 Using and Sharing Content	17
5.4 Referencing Embargoed Issues	18
6 Conclusion	18
7 Sources	19
8 Appendices	20

# **1** Introduction

E-Periodica<sup>1</sup> is the ETH-Bibliothek's<sup>2</sup> open access platform for digitised Swiss journals from the fields of science, history and culture. The timeframe of the uploaded publications stretches from the  $18^{\text{th}}$  century up to current issues. E-Periodica's core service is aimed at guaranteeing straightforward and free access to the journals. Frequently, this also results in issues and runs that were previously out of print being made available again. The platform focuses on publications by non-commercial publishing houses. At the beginning of 2016, 362 titles with a combined number of 4.2 million pages were available online.

As a basic principle, the full text of all journals is freely accessible and can be downloaded as a PDF; no usage fees are charged. However depending on agreements with respective copyright holders, more recent issues can be embargoed for up to a maximum of five years.

E-Periodica's target audience includes people from different fields of interest and age groups:

- Students and scientists from all over Switzerland and Europe who are conducting research for academic projects and studies
- Journalists on the lookout for background information
- Journal subscribers, often including association members who consult their association publication online
- Members of the general public in search of information for their personal interests and hobbies

In the following, E-Periodica will be presented as a specific practical example for the provision of digitised journals. Particular attention will be paid to two improvements, which were implemented in the scope of a technical refinement:

- The simplification of the work processes via workflow-controlled journal preparation
- A complete redesign to guarantee a contemporary appearance

Additionally, work processes and strategic and organisational issues will be examined for a comprehensive overview of E-Periodica's service.

# 2 History

Originally called retro.seals.ch, the service was conceived as part of the since completed innovation and cooperation project "Swiss electronic library" (e-lib.ch). A project sponsored by the Swiss University Conference (SUC), the ETH Board and the State Secretariat for Education, Research and Innovation (SERI). The aim of e-lib.ch was, among other things, to create "a national portal that offers improved availability of scientific information on a sustainable basis and facilitates research and access"<sup>3</sup>. Incidentally, this is also the origin of the original name retro.*seals*: Swiss Electronic Library Services.

Retro.seals.ch was initiated in collaboration with the ETH-Bibliothek project E-Archiving (2005-2008). As part of E-Archiving print journals were retro-digitised and a suitable presentation platform was set up. After the completion of E-Archiving, e-lib.ch then launched

<sup>&</sup>lt;sup>1</sup> www.E-Periodica.ch (until 2015: retro.seals.ch)

<sup>&</sup>lt;sup>2</sup> ETH-Bibliothek: http://www.library.ethz.ch/

<sup>&</sup>lt;sup>3</sup> Swiss electronic library: http://www.e-lib.ch/

the retro.seals.ch project (2008-2013). The project was initially realised by the Consortium of Swiss Academic Libraries<sup>4</sup> in collaboration with the ETH-Bibliothek (2008-2012). The latter assumed full responsibility for the platform in 2013 and has been in charge of organising and running it ever since.

When the platform was launched in 2007, the service focused on Swiss architectural journals – the so-called "memory of Swiss construction" – and mathematical and scientific publications. Gradually, new subjects were added and the service blossomed into the broad range of topics and contents that are available today.

Retro.seals.ch was not originally designed for such a large and constantly growing range of journals. The enormous data volume imposed increasingly high demands on the IT infrastructure. In order to guarantee the long-term operation of the platform, ETH-Bibliothek redesigned the application extensively in collaboration with the operator SRZ Berlin<sup>5</sup>. This overhaul was centred on a newly designed Workflow Client, which allows the work processes to be controlled centrally (see Chap. 4.3f). retro.seals.ch was also redesigned within the scope of this enhancement (see Chap. 5) and renamed E-Periodica.

# **3** Organisation

## 3.1 ETH-Bibliothek's Overall Responsibility

All the digitisation and data structuring, the operation of the platform's IT infrastructure and data storage is conducted at the ETH-Bibliothek or by ETH Zurich. ETH-Bibliothek's DigiCenter is in charge of organising and coordinating E-Periodica. The infrastructure is maintained by the ETH-Bibliothek's IT staff. Students assist with operative tasks like scanning, quality control and structuring.



Organisational chart/organisation of E-Periodica (ETH-Bibliothek 2016)

<sup>&</sup>lt;sup>4</sup> Consortium of Swiss Academic Libraries: http://lib.consortium.ch/

<sup>&</sup>lt;sup>5</sup> SRZ Berlin: http://www.srz.de/

# 3.2 Cooperation with the Swiss National Library

The platform is operated in collaboration with the Swiss National Library (NL). The NL selects journals from its archives, clarifies the rights and outsources the digitisation of journals to external service-providers. The data, however, is indexed in-house with the same XML editor used for E-Periodica, and supplies the fully prepared datasets to ETH-Bibliothek for uploading. Usually, these datasets focus on certain subject areas, which are documented with a collection of publications (e.g. "History of Education" or "Ageing").

## **3.3 Contractual Partners**

The addition of a new journal to E-Periodica is governed with a written contract by the involved parties – including copyright holders, publishers and editors (see Chap. 4.1). (Exceptions are only made for older orphan works, for which the copyright holders or responsible parties can no longer be determined.) As of the end of 2015, more than 100 contracts had been signed with publishers, as well as umbrella organisation or responsible Swiss libraries, within the scope of E-Periodica.

## **3.4** Development of the Platform

The platform is constantly being expanded with new content. This involves adding new journals and supplementing existing journals with current issues. Requests for new journal projects are submitted to E-Periodica through various channels:

- Directly from the responsible parties or copyright holders of a journal, such as publishers and editors who are looking to digitise their publication and to offer it on E-Periodica.
- Indirectly by libraries; mostly academic libraries that would like to use this channel to make their own holdings available to the public. Often, these involve publications from the respective university environment.
- Or E-Periodica might have a specific interest in a journal. This usually happens following a suggestion from users, in order to supplement an existing subject or to explore a new topic with a view to completing the service.

### 3.5 Inclusion Criteria

While the platform focused exclusively on scientific journals initially, this restriction was soon relaxed in favour of a broader service. The decision on whether to include a new journal is made based on the following criteria:

- Journal-like character of the publication (regular publication frequency)
- Consent of the copyright holders (unless it is an orphan publication)
- Ties to Switzerland (e.g. thematically, headquarters of the publisher)
- Importance of material for research
- Of general interest
- Non-commercial character
- No equivalent online service available elsewhere

Moreover, E-Periodica endeavours to offer a complete run of a journal. As a result, complete holdings are needed for digitisation These may be obtained from several different sources if necessary.

### 3.6 Legal Aspects

ETH-Bibliothek only holds the right to digitise and upload the journals within the scope of E-Periodica. Other copyrights or usage rights remain fully with the publication's copyright holders. As far as authors' rights are concerned, a pragmatic solution was adopted. Obtaining the rights from every author would potentially cause a disproportionate effort or might be impossible in some cases. This could cause many journal projects to fail from the outset. Consequently, this approach was abandoned. If an author disagrees with his or her paper being uploaded onto E-Periodica, however, it is deleted or replaced with a filler text.

The contents on offer are available for non-commercial purposes in teaching and research, as well as for private use. Files or printouts from the service can be passed on with usage instructions and the correct designations of origin.

The publication of images in print and online publications is permitted only with the prior consent of the copyright holders. ETH-Bibliothek forwards any corresponding requests to the journals' copyright holders. As soon as consent is given, ETH-Bibliothek provides a high-resolution master file (TIFF) of the page or image in question upon request. A processing fee is charged for providing the data.

Storing parts of the electronic service on other servers also requires written consent from the copyright holders. Including links to pages from E-Periodica's service, however, is possible at all times and encouraged.

# 4 E-Periodica Processes

The following sections illustrate E-Periodica's processes, standards and technical solutions. The individual steps are discussed in detail based on a process model for the inclusion of a new journal:



# 4.1 Negotiation

If a journal meets the criteria and is included on E-Periodica, the most important legal, organisational and financial issues are clarified with those responsible for the journal and set down in a contract. This includes the following points:

- Legal situation: ETH-Bibliothek only holds the right to digitise and upload the journal. All other rights remain with the copyright holders.
- Organisation of holdings: It is evaluated whether complete holdings are available for processing.
- Saving the data: the generated image, OCR and XML files are saved on ETH-Zurich servers. Upon request, the contracting parties can also obtain a copy of the data.

- Uploading new issues with or without an embargo (provided that the journal is still being published).
- Costs: the agreement contains a cost estimate and stipulates who will cover which costs.

The actual project launch takes place once all parties have signed the contract. Information on the intended digitisation is published on  $digicoord^6$ , an information platform for Swiss digitisation projects.

## 4.1.1 Cost Factors

The following cost factors form the basis for the cost model used to calculate a journal project:

- Infrastructural costs: scanners (maintenance, repairs, new acquisitions), as well as PC stations for quality control and structuring
- Personnel costs
- Maintenance of the platform's IT infrastructure
- Data hosting
- Licence costs for OCR and image processing

### 4.1.2 Cost Model

The costs for a journal project are calculated based on the total number of pages of all the runs published.

The costs are divided into:

- One-off project costs: these include the digitisation of retro-holdings up to the latest issue, quality control with automated image processing, the creation of OCR files (Antiqua and Fraktur), and the structuring of the data for online presentation.
- Annual costs: these comprise hosting the data and processing and uploading the latest issues (provided a corresponding update has been agreed).

For very extensive publications, the costs can also be correspondingly high. Consequently, E-Periodica's cost model stipulates that the majority of the project costs for processing of the retro-holdings is co-funded by ETH-Bibliothek and the annual costs – which are considerably lower – are passed on to the contractual partners. There are also umbrella organisations<sup>7</sup> or other libraries<sup>8</sup> that cover the one-off project costs and/or annual costs partially or in full. This cost model prevents an excessive financial burden on the contractual partners and enables interested parties with smaller budgets to participate. This in turn leads to a greater variety of journals on E-Periodica.

# 4.2 Digitisation

Should the contractual partners have complete journal holdings at their disposal, these will be used for the digitisation. If doublet holdings that are no longer required are available, they are

<sup>&</sup>lt;sup>6</sup> Digicoord: information platform for Swiss digitisation projects: https://www.digicoord.ch/

<sup>&</sup>lt;sup>7</sup> Swiss Academy of Humanities and Social Sciences: http://www.sagw.ch/sagw.html

<sup>&</sup>lt;sup>8</sup> University Library, Bern: http://www.unibe.ch/universitaet/dienstleistungen/universitaetsbibliothek/, The Library am Guisanplatz, Bern: www.guisanplatz.ch/, Kantonsbibliothek Graubünden, Chur: www.kantonsbibliothek.gr.ch

cut up to render the work process more efficient and disposed of once processing is complete. If needed, missing volumes or issues are loaned from other libraries. Depending on the type and suitability of the master copy, the journals are digitised on a reflective scanner, by a scanning robot or with a document scanner in TIFF format, 300 dpi, colour or grey scale at the DigiCenter of the ETH-Bibliothek. Once digitised, the scans undergo a standardised quality control, which includes verifying their completeness, colourfastness and sharpness. If necessary, automated image processing is performed (for the run) using two different software programmes – PageImprover<sup>9</sup> and ScanTailor<sup>10</sup>. The TIFF files are then converted into compressed derivatives in JPEG format for data structuring and online presentation.

In exceptional cases, digital copies generated by third parties are also accepted – provided they meet E-Periodica's basic standards. This data is also subject to a quality control and optimised, if necessary.

## 4.3 Role-Based Workflow: Structuring

Following the digitisation and quality control phase, the journal is structured manually. As part of the overhaul of E-Periodica's IT infrastructure, a tailored Workflow Client was designed, which communicates directly with the workflow server and enables the journals and data processing to be controlled and managed centrally. The following chapters outline the work processes supported on this Workflow Client. A diagram of the complete structuring workflow is provided in the appendices (Figure 1).

### 4.3.1 Data Provision

The RGB-TIFF files created at the DigiCenter (or, in exceptional cases, digital copies generated by third parties) with a resolution of 300dpi, form the starting point for uploading a journal.

The TIFF files waiting to be imported are filed in a predetermined file structure on a network attached storage (NAS) system. The structure of this file system indicates the journal, run and volume to which the files belong.



This example shows the file structure created for import on the NAS system with the first five TIFF files of the journal *Appenzellisches Monatsblatt* ("apm" is the abbreviation used for the journal) starting from the 1840 run, issue number 016.

<sup>&</sup>lt;sup>9</sup> PageImprover, 4digitalbooks: http://www.4digitalbooks.com/\_soft\_imaget.html

<sup>&</sup>lt;sup>10</sup> ScanTailor: http://scantailor.org/

#### 4.3.2 Recording the Journal

Before the images can be imported, the journal needs to be entered into the workflow system. This step is performed with the application "Workflow Client". This programme, a practicable Java application, is stored centrally on a NAS area, with which it communicates via a REST<sup>11</sup> interface. It can be accessed via an EXE file from anywhere on the server (providing NAS access was granted). Once the administrator has logged into the Workflow Client, he or she can record different metadata from the journal, such as the journal's title, ZDB and ISSN numbers, publisher, year of publication and volume numbers, in a form.

Journal jsk-001 (56)		
Kollektion	gui	
UniqueId	jsk-001	
dbid	3025	
Version WF	03.02.2016 16:06:53	
Version Staging	03.02.2016 16:04:20	
Version Prod	03.02.2016 16:06:51	
Klassifikation	ddc:793, ddc:796	
ZDB	600235-3	
ISSN	1421-0770	
Titel	Ski : Jahrbuch des Schweizerischen Ski-Verbandes = Annuaire d	
Herausgeber	Schweizerischer Ski-Verband	
Erscheinungsjahr (von/bis)	1908-1932	
Bandnummer (von/bis)	4-28	
Link		
Zugriffschutz	0	
Logo		
pdf-generation-issue	Nein	

Example of a journal's recorded metadata

### 4.3.3 Loading the TIFF Files

Once the journal is recorded in the system and the TIFF files are ready for import, the administrator can load them in the Workflow Client. For this, he or she opens the import dialogue, which displays all the issue orders of the journal in question for selection on the NAS system. Selecting one or several files for import triggers various processes in the background:

• A structure XML file is created, which contains the journal's metadata recorded in the previous step, supplemented with the issue run information from the underlying file structure.

<sup>&</sup>lt;sup>11</sup> REST – Representational State Transfer: programming paradigm for distributed systems

- Another process concerns the generation of various JPEG derivatives from the TIFF files, which are required for structuring in the XML editor and for the frontend. The image server responsible for this initially creates the JPEG derivatives with a width of 800 pixels, as this is the size required by the XML editor. This ensures that the structuring process does not experience any unnecessary delays. JPEG derivatives are then generated in different sizes (100, 200, 400, 600, 1200 and w0, with the latter referring to the original size) in the background for the zoom view on the webpage.
- In parallel, loading the TIFF also initiates the OCR server, which creates the full texts based on the TIFF files. Depending on the configuration, Antiqua, Fraktur or mixed OCR is generated.

Once all the 800-pixel derivatives of a journal run have been created, this run of issues is marked as "Ready for indexing" and can be structured.

4.3.4 Recording the Structure

The XML editor, another Java application, is used to record the structure. Like the Workflow Client, this application resides centrally on a NAS and is launched via an EXE file. After the member of staff has signed into the XML editor with his or her user data, it establishes a connection to the workflow server and communicates with it analogously to the Workflow Client via a REST interface. The journal to be processed can now be selected via a pop-up menu. This can be a journal that has to be re-structured, is still in the works or needs correcting. Depending on which structuring workgroup the member of staff belongs to, journals in various statuses are available for selection.

DocumentId	
hvg-001:1865:1	
hvg-001:1866:2	
hvg-001:1867:3	
hvg-001:1868:4	
hvg-001:1869:5	
hvg-001:1870:6	
hvg-001:1871:7	
hvg-001:1873:9	
hvg-001:1874:10	
hvg-001:1875:11	-

Once a journal has been selected for processing, the XML editor appears as follows:

Datei Bearbeiten Ansicht Bild Hilfe		
hvg-001:1872:8 (online)	hvg-001_1872_008_0031.jpg	Element
a 🗄 Zeitschrift		<ul> <li>Artikel</li> </ul>
a 🥁 Band		Attribute
a 🔚 Titelseiten		
1		Titel Der Kanton Glarus unter der Helvetik : di
a 🔚 Inhaltsverzeichnis		Autor Heer I
2		
🔺 🔚 Erratum		Bildanhang
📰 3	E Contraction of the Contraction	Textanhang
a 🔚 Artikel		
📰 4	Der Kanton Glarus unter der Helvetik.	GlueID
5 - 2	Dritter Zeitraum: Herbst 1799 bis August 1802.	DOI
📻 6 - 3	Von Di. J. Heen.	
7 - 4	a subset	
🔺 🔚 Artikel		
8 - 5	Der letzte Theil unserer Darstellung (Jahrbuch, VI. Heft, S.	
9 - 6	13 ffg.) hat die Geschichte des Kantons bis zu dem Zeitpunkte	
<b>10 - 7</b>	verfolgt, wo die russische Armee unter Suworoff über den Panixer-	
🔺 🔚 Artikel	pass aus unseren Grenzen entwichen war und damit der Feldzug	
11 - 8	der Coalition gegen Frankreich, soweit er seinen Schauplatz in der	
🗐 12 - 9	Schweiz land, sein Ende erreicht hatte. Der Krieg selbst war damit	
13 - 10	neutra noch lange nicht zum Abschlusse gelangt und es dauerte	
<b>14 - 11</b>	1001) denselben definitiv beendigte. Die gum Frühichn 1900 iedeeb	
15 - 12	ubten im Wegentlichen die Waffen und nun versingelte Kömpfe	
16 - 13	von untergeordnater Wiebtigkeit welche em untern Nelder von	
17 - 14	Zeit zu Zeit vorfielen erinnerten daran dass nene Waffengänge	
18 - 15	abermale bevorstehen	
<b>19 - 16</b>	Die Schweiz blieb indessen nach wie vor von den Französen	
20 - 17	besetzt: unter dem Oberhefehl des Generals Lecourbe stand eine	
21 - 18	fränkische Armee von annähernd 30.000 Mann in einer weitlänfigen	
22 - 19	Aufstellung, welche von Lauffenburg bis Ragatz reichte und so den	
23 - 20	deutschen und italienischen Schauplatz des Krieges, so weit möglich.	
24 - 21	mit einander in Verbindung setzte.	
25 - 22	Für den Augenblick mochte es für die neu eintretenden Behörden	
25 22	des Ct. Linth ein Trost sein, dass sie diese militärischen Streitkräfte zur	
20 - 25	Hand hatten und sich der Beihülfe derselben versichert halten durften; die	Page Info
27 - 24	Lage, in welcher sie sich befanden, wäre sonst, dem schwergeprüften	
20-25	und missgestimmten Volke gegenüber, noch schwieriger gewesen,	Erfassung Bemerkung:
29 - 20	als sie es wirklich war. Auch so indessen waren die Verhältnisse	
50 - 21	so beschaffen, dass es muthiger und fester Herzen bedurfte, um an	
	die Aufgaben heranzutreten, welche sich aus der Lage der Dinge,	
31 - 28	selbst ganz abgesehen von dem Widerstreben der Gemüther gegen	Korrektur Bemerkung:
32 - 29		
33 - 30		
34 - 31		
🚍 35 - 32	- I	

The tree structure of the content is visible on the left-hand side. In this example, the structuring is already underway, as is evident from the structural entries and pagination. The middle window displays the 800-pixel JPEG derivative of the page selected on the left. And the structural metadata is recorded in the section on the right-hand side – in this case, the title and author of the article.

The manual structuring contains:

- Entry of authors and article titles
- Pagination
- Indication of the type of content with so-called "elements", e.g. article, foreword, book review, obituary, advertisement, etc.

In addition, further functions are available, which enable content to be labelled more effectively and, if necessary, to be merged:

• Position frame

Several pieces of content on a page (e.g. several individual articles) are being differentiated with the use of a manually set position frame. This enables the corresponding metadata, e.g. article title and author, to be recorded separately for every section.

• Page identifier

Page identifiers (PID) are used to link elements such as boards located in a volume's appendix or a series of images that stretches throughout a volume to the relevant article.

• GlueID

The so-called GlueID uses a similar principle. This identifier can be used to join

together an article that is interspersed with several pages of advertising or other content.

The added value of these three linking and differentiation options becomes apparent when downloading content. On the one hand, the PDF in question automatically contains the complete article with any appendices, without having to search for and download them separately. On the other hand, the PDF contains solely the desired article, as any inserted pages that do not belong to the article are being hidden automatically.

Once the structuring process is complete, the journal run is marked with the status "Indexed" and can be checked in the next workflow step. Here, the dual control principle is applied, i.e. another member of staff checks the structured journal and either approves it for internal uploading (status: "Indexing checked") or requests a revision (status: "Revision requested").

## 4.4 Uploading

## 4.4.1 Internal Uploading

Before a journal goes online, it is checked in a test environment. Once again, the administrator is responsible for this step and uploads the journal runs with the status "Indexing checked" to a so-called "staging server". This server contains a test environment that encompasses the complete online service of E-Periodica. Here, the new journals to be uploaded are tested for several factors, including the following views and functions:

- Are the pages being displayed correctly?
- Is the table of contents being displayed correctly?
- Can the PDFs of the individual structure nodes be generated?
- Is the content of the journal run searchable?

If errors are discovered during quality control, they are being documented in the Workflow Client and the status of the journal run is set to "Correction requested". The XML editor then makes the corrections.

### 4.4.2 Uploading onto E-Periodica.ch

Once no more errors have been detected on the staging server, the journal run can be uploaded to the productive server, at which point it becomes visible on www.E-Periodica.ch.

### 4.4.3 Communication, References, Searching

Information on new uploads is provided via ETH Zurich's news channels and mailing lists. Usually, the publishers make an announcement themselves, reporting on the uploading of their journal onto E-Periodica on their websites and/or in a current journal issue.

E-Periodica's journal holdings can be searched via various catalogues and databases:

- ETH-Bibliothek's Knowledge Portal: http://www.library.ethz.ch/en/
- The NEBIS catalogue: Network of Libraries and Information Centres in Switzerland: http://www.nebis.ch/eng/
- swissbib.ch: the catalogue of all Swiss university libraries, the Swiss National Library, several cantonal libraries and of other institutions: https://www.swissbib.ch/

- The Electronic Journals Library of the University of Regensburg: http://rzblx1.uni-regensburg.de/ezeit/
- International providers (e.g. EBSCO DiscoveryServices) also harvest the data via an OAI interface: http://www.ebsco.com/

## 4.5 Creation of Archive Capsules

The final step in the entire workflow process concerns the creation of archive capsules. For every complete journal run, an uncompressed ZIP file comprising the original TIFF files, the OCR files and the structure XML is created. This ZIP capsule is subsequently transferred to the archival system overnight. Two weeks after the successful archiving, the original TIFF files on the NAS are deleted. This timeframe allows any undiscovered errors in the TIFF files still on the NAS server to be rectified without having to retrieve the TIFF files from the archive. If corrections are made to a journal run – whether it is a modification of the structural or bibliographical metadata, the images or the OCR files – a delta archive capsule containing solely the altered data is generated.

### 4.6 Architecture and Software

The productive and staging servers (web application) both run under Tomcat 6.0.24 on a virtualised Linux server with Red Hat 4.4.7-16 as the operating system with 12GB RAM and 4 CPUs. The workflow server, image server and OCR engine run on a virtualised Windows server 2008 R2 Enterprise, with 8GB RAM and four CPUs. Figure 2 (in the appendices) shows the server architecture as it is currently in use.

The backend of the web application is programmed completely in Java. The frontend is based on HTML5 and CSS3, and used as a framework bootstrap.

# **5** Online Presence

As the original online presence of retro.seals.ch was modified and honed only marginally during the platform's existence, it gradually ceased to meet the contemporary standards.

retrodigit	ized journals	
НОМЕ	٩	
Search		
Advanced search Last hitlist	open all   close all	
Browse by classification by collection by title by author by year of publication	DDC-000: Computer science, knowledge & systems     DDC-020: Library & Information sciences     DDC-060: Associations, organizations & museums     DDC-070: News media, journalism & publishing     DDC-100: Philosophy	
About retro.seals.ch History Usage information Digitisation projects	DDC-200: Religion     DDC-200: Christiantly & Christian theology     DDC-230: Christiantly & Christian theology     DDC-200: Social sciences, sociology & anthropology     DDC-300: Social sciences, sociology & anthropology     DDC-300: Social sciences, sociology & anthropology	
Actual repository-content Journals: 374 Volumes: 12449 Articles: 534781 Pages: 4572616	DDC-340: Law     DDC-350: Public administration & military science     DDC-350: Social problems & social services     DDC-370: Education     DDC-380: Commerce, communications & transportation     DDC-390: Ethnology, folklore     DDC-400: Language     DDC-400: English & Old English languages     DDC-400: Other language     DDC-400: Other language     DDC-400: Other languages     DDC-400: Dther languages     DDC-400	
	• DDC-500: Sciences	

Previous entry page retro.seals.ch (status: December 2015)

Within the scope of the technical refinement of the platform, the retro.seals.ch website was therefore redesigned and renamed E-Periodica. The main new features of the comprehensive new look – designed and realised by the web design company  $Cando^{12}$  – will now be explained in more detail.

The core concept is the preparation of the content in a responsive design. This concept takes the small size of the displays on mobile terminal devices – like tablets and smartphones – compared to desktop monitors into consideration and contains a flexible layout grid for different screens (Zillgens, 2013 p. 15). Content-wise, the focus lies on core functions and the central statements of the website (ibid. p. 231f). Further information about the website and its contents is available in so-called "off-canvas elements", which can be faded in if desired. This allows all the necessary information to be provided without compromising design and legibility.

Moreover, E-Periodica's new online presence is characterised by two central criteria and requirements:

- The original concept of a more text-oriented presence was replaced completely with the new design and visual access to the content. This should do more justice to the content and vividness of the journals.
- Accessibility: the images are labelled with alternative titles and can be displayed in the form of list elements. Users with special needs can have the texts read to them.

The compatibility of the design realisation with the platform's existing requirements was a particular challenge. Ways had to be found to combine design standards and ideas with the technical and content-related structure of the platform.

The core features of the new design are outlined below.



# 5.1 Homepage

<sup>12</sup> Cando Image GmbH: https://www.cando-image.com/

The homepage provides a very visual entry point: every journal is presented with its latest cover. In order to avoid performance problems while loading the homepage (there are a considerable number of images), so-called "lazy loading" is used, an optimised loading method:

"So-called 'lazy loading' offers a possibility to speed up the page set-up. In doing so, certain components are deliberately excluded from the initial main loading process. Instead, they are loaded 'lazily', i.e. passively, as soon as the main loading process is complete and the resulting activity tower has settled." (Zillgens, p. 345).

This means that the images or journal covers on E-Periodica are only loaded if they are displayed during scrolling.

However, there is also the option to switch to a list view with an alphabetical list of titles without images.

=	e-periodica	Search in full text	Search P+	ETH BIBLIOTHEK
		Schweizer Zeitsch	iften online. Revues suisses en ligne. Riviste svizzere online. Swi	iss journals online.
	> Alter und neuer grosse	er Staats-, Kriegs- und Friedens Appenzeller-Cak	nder, oder, Der hinkende Bott	NEW
	Achi : nista svizzera G archiettura, ingegneria e utranistica = Swiss review of architecture, engineering and utran planning			
	Archiv des Historischen Vereins des Kantons Bern			NEW
	> Abhandlungen des His	Abhandtungen des Historischen Vereins des Kantons Bern		
	> Abhandlungen und Be	Abhandlungen und Beobarthlungen durch die Ökonomische Gesellschaft zu Bern gesammet		
	Acta Tropica			
	Action : Zivilschutz, Be	Action : Zivitschutz, Bevölkerungsschutz, Kulturgisterschutz = Protection ofwile, protection de la population, protection des biens cultures = Protezione della populazione, protezione della populazione, protezione dei beni culturali		
	Algemeine schweizensche Militärzetischnit - ASMZ			
	> Allgemeine schweizen	Algemeine schweizerische Militärzefung = Journal militäre suisse = Gazetta militäre svizzena		
	Alpexpress. Ticino : la rivista di AlpTranut San Gottardo SA			
	AlpTransit in Ticino	AlpTranst in Ticino		
	An die zürcherische Ju	An die zürcherische Jugend auf das Jahr		

#### 5.2 Searching and View Options

The website offers various forms of access to the full text of a journal: (a) simple search, (b) advanced search, and (c) list of journals.



E-Periodica's search model

The search tool offers a combined search and filter option. In other words, during a simple search the filter responds dynamically by continuously filtering and arranging the selection of the journals found while the search terms are entered in the search field ("masonry style"<sup>13</sup>). When a journal is selected, the user is taken to its volume overview (a volume generally corresponds to a run), and from there to the manually created table of contents for the run in question with the corresponding page view.

#### $\equiv$ e-periodica Search in full text P. Search BIBLIOTHEK < Back to the volume index i Detailed inform Display > Du : die Zeitschrift der Kultur Table of contents Page thumbnails Information on the journal Heading Page PDF Volume 68 (2008-2009) Issue 783: Die Zeit Front matter DIE ZEIT DA. Advertising A D Table of Contents Advertising 2 B Preface: Zu dieser Ausgabe 3 A R 3 Advertising Article: Kulturtipp Wien : Coop Himmelb(I)au MAK P 6 R Advertising 7 Article: Kulturtipp Wien : Max Ernst : Albertina R R Advertising Article: Kulturtipp Wien : Arcimboldo : Kunsthistorisches Museum 10 2 11 R Pag Article: Kulturtipp Wien : Matthew Barney : Kunsthalle 12

## 5.2.1 Split View: Table of Contents and Page View

The two-part "split view" with a page view and table of contents has the advantage that the content overview of the volume or run in question remains visible while scrolling in the display. This allows quick and easy access to other articles. All the pages of a volume can be displayed as thumbnails and general information on the respective journal, such as the publisher or contact possibilities, is available in another register.

# 5.2.2 Full Screen and Zoom View

The full screen view is reached from the simple page view using the extendable OpenSource Viewer *OpenSeaDragon*<sup>14</sup>. The zoom function is performed smoothly and page turning is also possible in the zoomed detailed view. In the case of mobile devices, pages can be turned in both the simple page view and the zoom view by swiping.



<sup>13</sup> Masonry style: http://masonry.

<sup>&</sup>lt;sup>14</sup> OpenSeaDragon: https://openseadragon.github.io/

# 5.2.3 List of Hits

The list of hits in the simple or advanced search is displayed simultaneously with an image preview. The search terms are marked in both the list of hits and the preview. Facets (authors, timescales, publishers, document types and collections) are available to narrow down the number of hits.



# 5.3 Using and Sharing Content

#### PDF Download

The full texts can be downloaded as PDFs. The PDF is created from the respective JPEG files available directly during the download. (For reasons of memory space, no PDF files are deposited on the platform). Additionally, a cover sheet containing the most important metadata and conditions of use is generated during the download and attached to the PDF as the first page.

#### DOI

The contents designated as *articles* during the structuring process receive a digital object identifier (DOI) for permanent traceability. The DOIs are allocated when the data is imported into the platform's infrastructure. These are registered via ETH Zurich's DOI Desk.<sup>15</sup>

#### Social Media

The articles can be shared via Twitter, Facebook, Google+ and Pinterest.



### 5.4 Referencing Embargoed Issues

As mentioned earlier, new issues can be embargoed for a maximum of five years. The embargo serves to avoid a competitive situation between online and print versions if the print version of the latest issue carries a fee. Although these embargoed issues are uploaded onto E-Periodica, access to the full text is not activated. Despite this restriction, added value can still be offered to users in that they can learn about new papers from the table of contents and, if they are interested, fall back on the print edition.

Embargoed issues are indicated with a lock icon and corresponding message. Once the embargo runs out, the full texts are activated automatically.



# 6 Conclusion

E-Periodica has proven to be a successful platform in the field of journal digitisation. The growing access figures and feedback from numerous users just go to show that the service is frequently used and appreciated. However, due to the platform not having been designed for

<sup>&</sup>lt;sup>15</sup> DOI-Desk, ETH Zurich: https://www.library.ethz.ch/de/Dienstleistungen/Publizieren-registrieren-verwalten/DOI-Desk-der-ETH-Zuerich

such a large scope of journals originally, measures had to be taken to accommodate the constantly increasing volume of data. The successful overhaul means that the platform can now be operated on a sustainable basis. The new, role-based workflow enables an efficient realisation of all the work steps – from adopting the scan, generating the derivatives and full texts, via structuring to uploading and activating the data. Ultimately, the responsive design provides modern and contemporary access to the content.

# 7 Sources

Zillgens, Christoph. Responsive Webdesign. Munich: Hanser, 2013.

# **8** Appendices



Figure 1: Workflow Client, structuring workflow (ETH-Bibliothek 2015)



Figure 2: E-Periodica's server architecture (ETH-Bibliothek, 2015)