

Between legal restrictions, technological possibilities and user needs – the acquisition strategy of German Open Access Repository for East European Studies “OstDok



- background:
 - why a repository for East European studies at the Bavarian State Library?
 - research on Eastern Europe in Germany
- Very short demonstration
- content acquisition of the repository
 - digitized publications under copyright
 - Digitized publications out of copyright
 - born digital publications

The Bavarian State Library



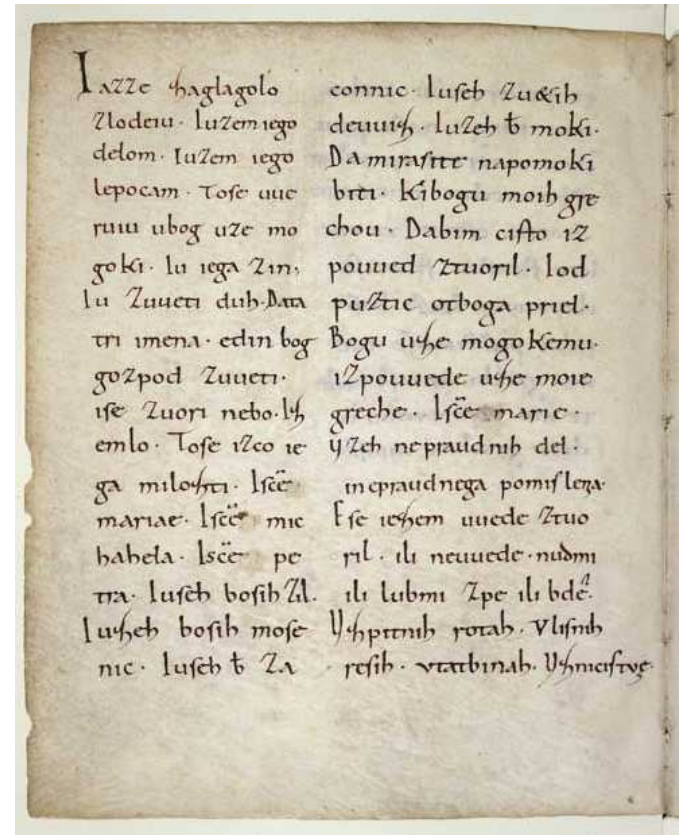


The Bavarian State Library

- founded 1558
- central regional and archival library of Bavaria (legal deposit since 1668)
- universal library
- 9.8 million volumes, 59,700 current periodicals, 1.1 million E-Books
 - 93,600 manuscripts, 20,000 incunabula
- ca. 800 employees
- acquisition per year: 138,000 volumes
- special collection fields (with function of a National Library):
 - 16th Century printed books (140,000 items)
 - history, classical antiquity, Eastern Europe, music, information sciences (supported by the German Research Association)

The East European Collection of the Bavarian State Library

- 1,3 million volumes
- 6.000 current print- and electronic journals
- 30.000 maps
- about 300 databases
- a large microfilm collection of archival materials
- precious manuscripts and old imprints
- New acquisitions: 20.000 monographical volumes a year

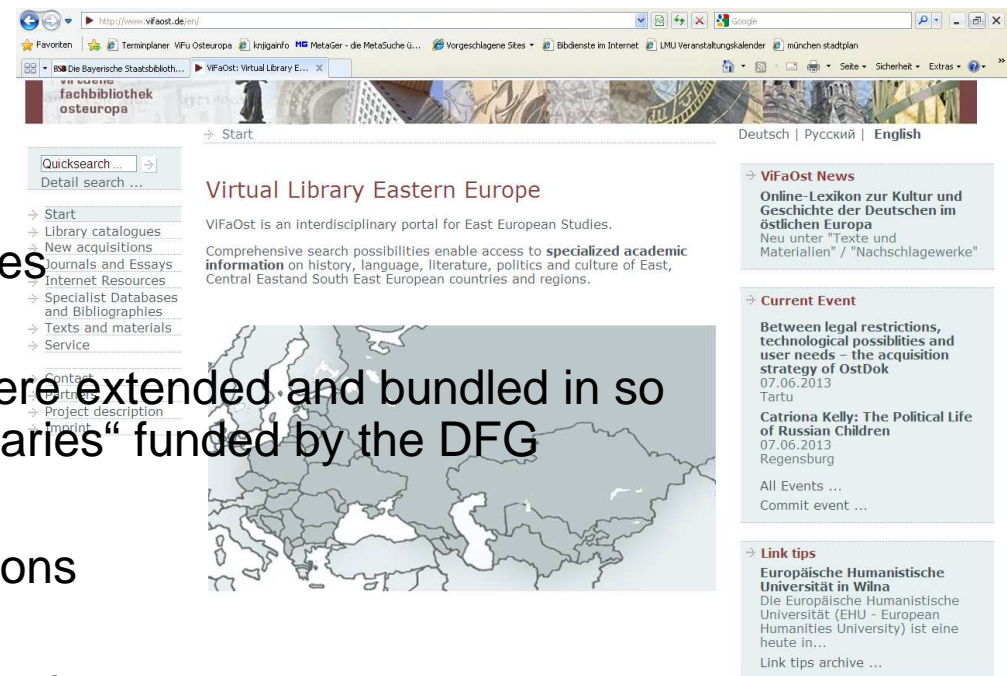


East European Research in Germany

- about 40 universities teaching subjects related to Eastern Europe (history, slavistics, politics , folk-culture, art ...)
- more than 20 non-university research institutions dealing with Eastern, Southeastern and Central Eastern Europe (mostly historical research)

establishment of virtual services for East European Studies in the BSB

- Since the 1990ies virtual services for the Special Collections were established and partly funded by the German Research Foundation (DFG), e.g.
 - new acquisitions databases
 - digitization of printed books
 - born digital projects
 - catalogues of internet resources
 - ...
- Beginning 2000 virtual services were extended and bundled in so called „Subject oriented virtual libraries“ funded by the DFG
ViFaOst online since 2003
(including a few digitized publications and born digital publications)
 - > more and more digitized books
 - > more and more requests for digital publication



OstDok – the project

- funding: German Research Foundation (since 2009-2012 + 2013-2014)
- open Access subject repository (disciplinary repository)
- content: partners for all countries of Eastern Europe
 - Collegium Carolinum München: Tshechia and Slovakia
 - Herder-Institut Marburg: Central Eastern Europe and the Baltic States
 - Inst. f. Ost- und Südosteuropaforschung Regensburg: Eastern Europe and Southeastern Europe
- academic board (historians, publishers and information specialists)



goals and principles

- for users / academics worldwide:
 - provide them with fast and direct access to (important) academic publications online
 - provide full text search facilities for these publications
 - provide the possibility of Online-Publishing
- for participating research Institutions
 - increase the visibility of their publications
 - increase the visibility of German Research in the field
- promote open access (in East European Studies)
- close interconnection to the library catalogue
- only machine-readable text, no documents with images only

Herzlich Willkommen bei "Osteuropa-Dokumente online" (OstDok), der zentralen Plattform für elektronische Volltexte der Osteuropaforschung mit derzeit über 2400 Publikationen.

OstDok ist ein Gemeinschaftsprojekt der Bayerischen Staatsbibliothek, des Collegium Carolinum München, des Herder-Instituts Marburg sowie des Instituts für Ost- und Südosteuropaforschung Regensburg. Seit 2009 wird es von der Deutschen Forschungsgemeinschaft gefördert.

OstDok bietet:

- Wissenschaftliche Arbeiten der Osteuropaforschung (Monografien, Schriftenreihen, Zeitschriften, Hochschulschriften u.a.)
- Suchmöglichkeiten in Volltexten und Metadaten aller Dokumente
- freien Zugriff auch auf urheberrechtlich geschützte Werke für private und wissenschaftliche Nutzung
- eine Plattform für die elektronische Publikation wissenschaftlicher Abhandlungen

online since 2011

Schnellsuche

[Neue Suche]

Schlagwort

Autor / Herausgeber

Abrüstung Anthologie Arbeitsmarkt Aspekt <Linguistik> Ästhetik **AUFSATZSAMMLUNG** Außenhandel **AUSSENPOLITIK** Außenwirtschaft Auswanderung Bibliographie Biographie Demokratisierung Deutsche Frage Deutschlandbild Entspannungspolitik Eurokommunismus Europäische Gemeinschaften Europäische Integration **Geschichte** Gesellschaft Glasnost Gorbačëv, Michail S. Grammatik Ideologie Informationspolitik Innenpolitik Internationale Politik Katholische Kirche Kirche Kirchenpolitik Kommunismus Kommunistische Partei Kommunistische Partei der Sowjetunion Kongress Kongreß Kultur Landwirtschaft **Literatur** Lyrik Marxismus Marxismus-Leninismus Militär Militärdoktrin Militärpolitik Nachfolgestaaten Nationalbewusstsein Nationalismus Nationalität Nationalitätenfrage Nationalitätenpolitik NATO Ost-West-Beziehungen Ost-West-Handel Partei Perestroika Philosophie Planwirtschaft Poetik Politik Politische Elite Politische Führung Politischer Wandel Politisches System Prosa Quelle Rat für Gegenseitige Wirtschaftshilfe Reform **Reformpolitik** Revolution Rezeption Roman Russisch-Orthodoxe Kirche Russlandbild Rüstungspolitik Semantik Sicherheitspolitik Sozialer Wandel Sozialismus Sozialistische Einheitspartei Deutschlands Sprache Staat Strategie Syntax Übersetzung Umweltschutz Verb Verfassung Warschauer Pakt Wirtschaft Wirtschaftsbeziehungen Wirtschaftsentwicklung Wirtschaftskooperation Wirtschaftskrise Wirtschaftsplanung Wirtschaftspolitik **Wirtschaftsreform** Wirtschaftssystem Wörterbuch Wortschatz



Herzlich Willkommen bei "Osteuropa-Dokumente online" (OstDok), der zentralen Plattform für elektronische Volltexte der Osteuropaforschung mit derzeit über 2400 Publikationen.

OstDok ist ein Gemeinschaftsprojekt der Bayerischen Staatsbibliothek, des Collegium Carolinum München, des Herder-Instituts Marburg sowie des Instituts für Ost- und Südosteuropaforschung Regensburg. Seit 2009 wird es von der Deutschen Forschungsgemeinschaft gefördert.

OstDok bietet:

- Wissenschaftliche Arbeiten der Osteuropaforschung (Monografien, Schriftenreihen, Zeitschriften, Hochschulschriften u.a.)
- Suchmöglichkeiten in Volltexten und Metadaten aller Dokumente
- freien Zugriff auch auf urheberrechtlich geschützte Werke für private und wissenschaftliche Nutzung
- eine Plattform für die elektronische Publikation wissenschaftlicher Abhandlungen

Schnellsuche

Schlagwort

Autor / Herausgeber

Abrüstung Anthologie Arbeitsmarkt Aspekt <Linguistik> Ästhetik **AUFSATZSAMMLUNG** Außenhandel **AUSSENPOLITIK**
 Außenwirtschaft Auswanderung Bibliographie Biographie Demokratisierung Deutsche Frage Deutschlandbild Entspannungspolitik Eurokommunismus
 Europäische Gemeinschaften Europäische Integration **Geschichte** Gesellschaft Glasnost Gorbačëv, Michail S. Grammatik Ideologie Informationspolitik
 Innenpolitik Internationale Politik Katholische Kirche Kirche Kirchenpolitik Kommunismus Kommunistische Partei Kommunistische Partei der Sowjetunion
 Kongress Kongreß Kultur Landwirtschaft **Literatur** Lyrik Marxismus Marxismus-Leninismus Militär Militärdoktrin Militärpolitik Nachfolgestaaten
 Nationalbewusstsein Nationalismus Nationalität Nationalitätenfrage Nationalitätenpolitik NATO Ost-West-Beziehungen Ost-West-Handel Partei Perestroika
 Philosophie Planwirtschaft Poetik Politik Politische Elite Politische Führung Politischer Wandel Politisches System Prosa Quelle Rat für Gegenseitige
 Wirtschaftshilfe Reform **Reformpolitik** Revolution Rezension Roman Russisch-Orthodoxe Kirche Russlandbild Rüstungspolitik Semantik



Schnellsuche

Q tartu

[Neue Suche]

117 Bände in 3.408 Sek.

1 - 10 > >>

Die Wirtschaft des Baltikums; 56 [Relevanz: 100%]

Autor / Herausgeber: Götz, Roland | **Verlagsort:** Köln | **Erscheinungsjahr:** 1990 | 36 S.

Zeitschrift / Sammelband: Bundesinstitut für Ostwissenschaftliche und Internationale Studien <Köln>: Berichte des Bundesinstituts für Ostwissenschaftliche und Internationale Studien; 1990

Kontext: Suche nach *tartu* in diesem Band - 2 Treffer Q

In der Hauptstadt Tallin (Reval)² leben knapp eine halbe Million Menschen und damit etwa 30% der Bevölkerung. Tartu (Dorpat) hat 114 000 Einwohner, die Städte Narwa, Kohtla-Järve und Pärnu

[Suche im Band] [PDF-Download]

Die militär-politische Erziehung der Komsomolzen : Pläne und Wirklichkeit; 8 [Relevanz: 94.28%]

Autor / Herausgeber: Lewytskij, Borys | **Verlagsort:** Köln | **Verlag:** Bundesinst. für Ostwiss. und Internat. Studien | **Erscheinungsjahr:** 1968 | 18 S.

Zeitschrift / Sammelband: Bundesinstitut für Ostwissenschaftliche und Internationale Studien <Köln>: Berichte des Bundesinstituts für Ostwissenschaftliche und Internationale Studien; 1968

Kontext: Suche nach *tartu* in diesem Band - 1 Treffer Q

nisationen ablesen, die in den letzten Monaten des vergangenen Jah-



Blättern | Land / Region | Baltikum, Finnland | Estland

10 Bände in 4.465 Sek.

Ētnosociolingvistika "Perestrojki" v SSSR : antologija zapečatlenogo vremeni [Relevanz: 100%]

Autor / Herausgeber: Duličenko, Aleksandr D. | **Verlagsort:** München | **Verlag:** Sagner |

Erscheinungsjahr: 1999 | 591 S.

Zeitschrift / Sammelband: Slavistische Beiträge; 378

[Suche im Band] [PDF-Download]

Akten zur deutschen auswärtigen Politik / B.2,1 : 1925 - 1933. Deutschlands Beziehungen zur Sowjet-Union, zu Polen, Danzig und den baltischen Staaten ; Dezember 1925 bis Juni 1926 [Relevanz: 100%]

Autor / Herausgeber: Deutschland <Deutsches Reich> | **Verlagsort:** Göttingen | **Verlag:** Vandenhoeck & Ruprecht | **Erscheinungsjahr:** 1967 | 594 S.

[Suche im Band] [PDF-Download]

Migrationsströme im sowjetisierten Baltikum und ihre Nachwirkung auf die baltischen Staaten nach Wiederherstellung der Selbständigkeit; 9 [Relevanz: 100%]

Autor / Herausgeber: Saar, Ellu; Titma, Mikk Ch. | **Verlagsort:** Köln | **Erscheinungsjahr:** 1992 | 60 S.

Zeitschrift / Sammelband: Bundesinstitut für Ostwissenschaftliche und Internationale Studien <Köln>: Berichte des Bundesinstituts für Ostwissenschaftliche und Internationale Studien; 1992

[Suche im Band] [PDF-Download]

Akten zur deutschen auswärtigen Politik / B.2,2 : 1925 - 1933. Deutschlands Beziehungen zur Sowjet-Union, zu Polen, Danzig, und den baltischen Staaten ; Juni bis Dezember 1926 [Relevanz: 100%]

Autor / Herausgeber: Deutschland <Deutsches Reich> | **Verlagsort:** Göttingen | **Verlag:** Vandenhoeck & Ruprecht | **Erscheinungsjahr:** 1967 | 599 S.

[Suche im Band] [PDF-Download]

Die lebenden Puppen : Prosa in Auswahl. Aus d. Estn. ins Dt. übers. v. Barbara u. Friedrich Scholz. Mit e. Einl. v. Friedrich Scholz [Relevanz: 100%]

Autor / Herausgeber: Tammsaare, Anton H.; Scholz, Barbara; Scholz, Friedrich | **Verlagsort:** München | **Verlag:** Fink | **Erscheinungsjahr:** 1979 | 218 S.

[Suche im Band] [PDF-Download]

Suchergebnis einschränken

Autor / Herausgeber

- Deutschland <Deutsches Reich>
- Bammesberger, Alfred (1)
- Duličenko, Aleksandr D. (1)
- Götz, Roland (1)
- Mečkovskaja, Nina B. (1)
- Saar, Ellu (1)
- Scholz, Barbara (1)
- Scholz, Friedrich (1)
- Mehr...

Schlagwort

- NICHT ANGEGEBEN (1)
- Quelle (3)
- Außenpolitik (2)
- Diplomatischer Dienst (2)
- Verzeichnis (2)
- Abstraktum (1)
- Aufsatzsammlung (1)
- Ethnolinguistik (1)
- Mehr...

Thema

- NICHT ANGEGEBEN (1)
- Historische Demographie
- Volkskunde, Brauchtum (1)
- Geschichte allgemein (3)
- Geographie, Heimat- und Ortskunde
- Recht- und Verfassungsgeschichte
- Literatur (1)
- Politikgeschichte (4)
- Mehr...

lawien

en

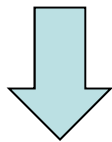
staaten

g
ect cataloguing and
ation for all documents

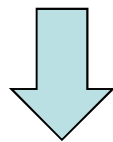
Content acquisition

1. documents under copyright

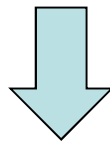
- monographs, book series and conference-proceedings edited by the three partner-institutions in the project



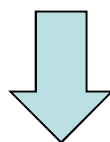
institutions with their own publishing house (HI)



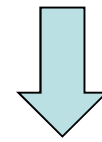
no negotiations needed



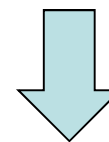
institutions working with publishers on the basis of commission (all rights with author and institution) (CC)



in principle no negotiations needed, but ...
publishers were informed **



institutions working with (several) publishers on a „regular“ basis (IOS)



negotiations with publishers

about 1.000 publications now,
about 1.000 to come

Content acquisition

1. documents under copyright

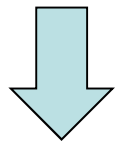
- monographs, book series and conference-proceedings edited by other research institutions

about 700 publications,
further 1.000 – 2.000 to come

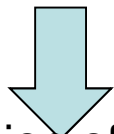
Content acquisition

1. documents under copyright

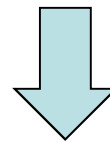
- Journals edited by the three partner-institutions in the project



institutions owning the
exclusive exploitation
rights



information of the authors
(with the possibility of
contradiction)



institutions owning no rights
(exploitation rights: publisher)



negotiations with publisher,
(moving wall) Information of
the authors

five major journals

not integrated yet for technical reasons

>> up to now no journal articles separately

Content acquisition

1. documents under copyright

- Individual authors: monographs published between 1966 and 1994 in Germany
 - legal possibility for authors to claim digital exploitation rights for these publications up to the end of 2008
 - In November / December 2008 the library contacted 100 authors of important publications
 - › asking them for permission to include their works in OstDok
 - › supplying them with letters
 - » informing the publisher that they claim digital exploitation rights
 - » granting the Bavarian state library non exclusive right of exploitation in digital form

80 authors answered and agreed !!

>>problems with second authors, publishers and images

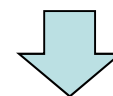
Content acquisition

1. documents under copyright

- individual authors: other monographs with regular copyright



exploitation right with
publishers



publisher no longer exists



does another publisher own
the exploitation rights?

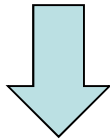
- contact authors directly (sometimes the other way around!)
 - › asking them for permission to include their works in OstDok
 - › supplying them with information on their rights (>>contract)
 - › how to contact the publisher and what to ask for
- spreading information via mailing lists

about 20 publications until now,
further 100 – 200 publications
(dissertations)

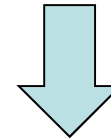
Content acquisition

1. documents under copyright

- individual authors: monographs out of print



exploitation right with
publishers



publisher no longer exists

- contact authors directly
 - › asking them for permission to include their works in OstDok
 - › supplying them with information on their rights (depending on the contract)
 - › how to contact the publisher and what to ask for
- Spreading information via mailing lists

about 20 publications until now,
further 100 – 200 publications
(dissertations)

Content acquisition

1. documents under copyright

- Individual authors: orphan works
 - › waiting for the (German) law / „diligent search“?

about 100 publications identified

Content acquisition

1. documents under copyright

- individual authors and institutions – new monographs and congress proceedings: advisement for digital publishing
 - In order to secure secondary rights for digital publishing (with moving wall)

Content acquisition

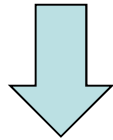
1. documents under copyright

- Publishers (back lists): monographs and journals
 - negotiations with publishing houses
 - › permission to digitize and make available via open access
 - › certain fee for each book paid (about 40 to 50 Euro)
 - › Four publishing houses agreed, one of them specialized in Eastern Europe
 - » Verlage Vandenhoeck & Ruprecht und V & R unipress
 - » Wilhelm Fink Verlag
 - » Verlag Ferdinand Schöningh
 - » **Verlag Otto Sagner**
 - offer (!) of several journals by small publishers

about 800 titles

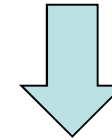
Content acquisition

2. documents out of copyright

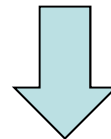


- publications before 1873

ca. 80.000 already digitized items
belonging to the East European
collection,
(not integrated in OstDok)



- publications between 1874 and 1943, where rights are not obvious
 - authors, authors or prefaces, illustrations



What should we include
in the repository?

Content acquisition

2. documents out of copyright

- Considerations
 - Include everything available >> interface with a default search for 20th and 21st century publications <> extended search 1600-1900
 - Include a selection >> who does the selecting?
 - › Researchers:
 - » advisory board
 - » smaller institutions with „new libraries“
 - › Users:
 - » „missing anything in OstDok“?



Already digitized
ant BSB?
>include in OstDok



not digitized but in
possession of one of
the partners
> digitize and
include



digitized in Russian
Presidential library?
> exchange, OCR
and include



available via eoD?
½ price funded by
DFG
> include

Content acquisition

3. born digitals

- born digitals published by the partners (for OstDok):
 - „book series“: very good master-theses (Collegium Carolinum)
 - conference proceedings and monographs edited by the partner-institution
 - conference proceedings, dissertations, other monographs
 - › quality selection: editorial team consisting of all partner institutions
 - » what license? Creative commons?
 - » commercial exploitation of possible print-version/ print on demand
 - › offer of technical support for born digital journals

Content acquisition

3. born digitals

- Publications on certain historical events on which research is being done because of anniversaries (thematic portals):
 - › 1913-2013: Balcan wars
 - › 1914- 2014 First world war
 - › 1415 – 2015 Jan Hus
- existing born digitals from other institutions and institutional repositories (serial publications and monographs)
- research data (primarily from German non-university-institutions)

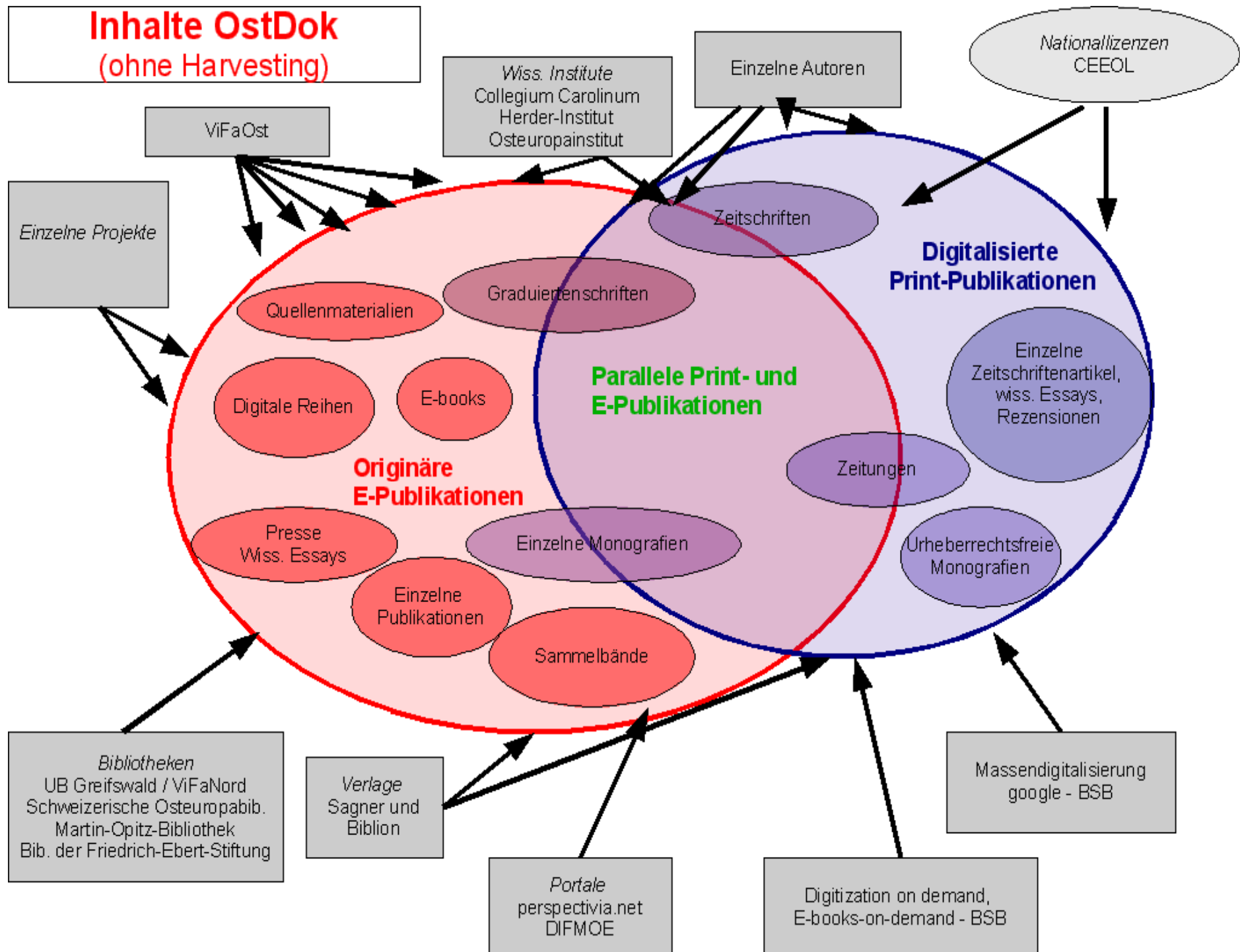
conclusions

- success of the project:
 - usage (statistics: about 200 visits /150 visitors daily)
 - closer cooperation with research institutes (and publishers)
 - but: content: 90% of the content under copyright is from Germany (>> partners, legal knowledge ...)
- >> cooperation with repositories or research institutions from East European countries

thank you for your attention




gudrun.wirtz@bsb-muenchen.de


Dr.Gudrun Wirtz




- Founded in 2001, [Creative Commons](#) (CC) is a non-profit organization^[37] which aims to facilitate the legal sharing of creative works. To this end the organization provides a number of copyright license options to the public, free of charge. These licenses allow copyright holders to define conditions under which others may use a work and to specify what types of use are acceptable.^[37] Terms of use have traditionally been negotiated on an individual basis between copyright holder and potential licensee. Therefore, a general CC license outlining which rights the copyright holder is willing to waive enables the general public to use such works more freely. Six general types of CC licenses are available. These are based upon copyright holder stipulations such as whether he or she is willing to allow modifications to the work, whether he or she permits the creation of derivative works and whether he or she is willing to permit commercial use of the work.^[38] As of 2009 approximately 130 million individuals had received such licenses

Die Technik von OstDok

Adresse  <https://www.esdoc.org/>   Wechseln zu

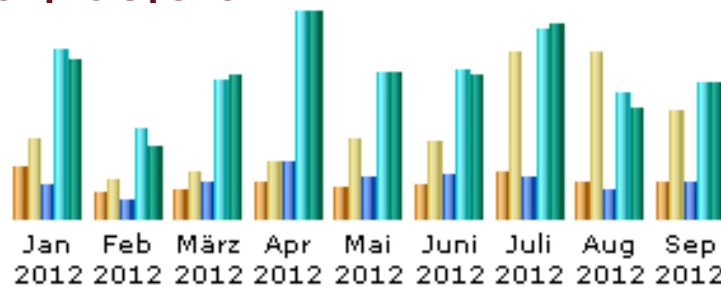


Startpage	Startpage	Search
Overview	eSciDoc Project Aim	Page Actions
General Concepts	eSciDoc is an eResearch environment developed specifically for use by scientific and scholarly communities to collaborate globally and interdisciplinary. It comprises core functionality including a Fedora repository (eSciDoc Infrastructure), a set of complementing services (eSciDoc Services), and application build on top of the infrastructure and the services (eSciDoc Solutions) that enable innovative eScience scenarios. Scientists, librarians, and software developers can work with research data, create novel forms of publications, and establish new ways of scientific and scholarly communication. more...	+ My Prefs
What is eSciDoc for?		+ Print Page
System Architecture		+ Log in
Content Model		<i>This page last changed on 17:12 01-Sep-2009 by swa.</i>
Publications		
Software		
eSciDoc Solutions		
eSciDoc Services		
eSciDoc Infrastructure		
Community	eSciDoc Days 2009	
Support	FIZ Karlsruhe and Max Planck Digital Library held the eSciDoc Days 2009 on Monday, June 15, and Tuesday, June 16, 2009 at FIZ Karlsruhe, Germany. The eSciDoc Days 2009 were targeted at both existing eSciDoc users and people interested in digital libraries, repository systems, e-Research infrastructures, and scholarly knowledge management.	
Sustainability	<ul style="list-style-type: none">eSciDoc Days 2009 Presentations (Karlsruhe, June 15-16, 2009)eSciDoc Days 2008 Presentations (Berlin, June 9-10, 2008)	
Sponsors		
Getting Started		
Download		
Installation		
Hints		
License		
Imprint		

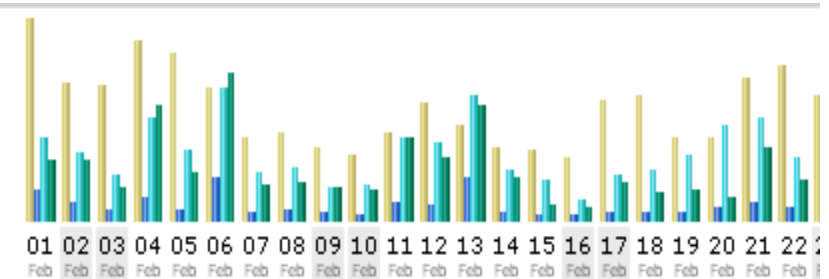


News
September 2, 2009 9:06 AM
Version 1.1.2 of eSciDoc Infrastructure is available
The eSciDoc Core Services (Infrastructure) version 1.1.2 is released. The new release includes the following new features and enhancements:

Statistika posjeta



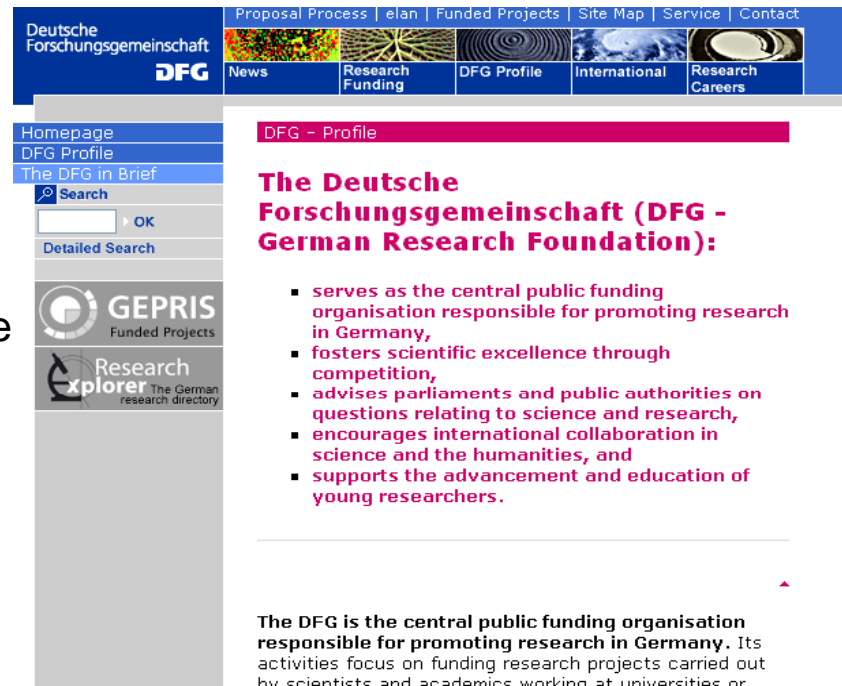
Monat	Unterschiedliche Besucher	Anzahl der Besuche	Seiten	
Jan 2012	1,498	2,386	25,516	
Feb 2012	755	1,190	15,492	
März 2012	856	1,350	27,601	
Apr 2012	1,048	1,654	43,128	
Mai 2012	893	2,302	32,766	
Juni 2012	976	2,197	34,092	
Juli 2012	1,375	4,805	32,585	
Aug 2012	1,057	4,829	21,105	
Sep 2012	1,107	3,140	26,838	
Okt 2012	1,276	5,982	30,577	
Nov 2012	1,781	4,359	31,005	
Dez 2012	1,650	3,594	31,644	
Total	14,272	37,788	352,349	



Tag	Anzahl der Besuche	Seiten	Zugriffe
01.02.2013	178	1,696	4,599
02.02.2013	122	966	3,747
03.02.2013	119	508	2,535
04.02.2013	160	1,315	5,812
05.02.2013	147	613	3,911
06.02.2013	117	2,396	7,470
07.02.2013	74	422	2,689
08.02.2013	79	557	2,975
09.02.2013	65	418	1,834
10.02.2013	59	371	1,994
11.02.2013	79	897	4,527
12.02.2013	104	871	4,333
13.02.2013	84	2,387	6,881
14.02.2013	64	485	2,789
15.02.2013	62	297	2,237
16.02.2013	57	257	1,234
17.02.2013	105	499	2,559
18.02.2013	110	417	2,866
19.02.2013	72	432	3,629
20.02.2013	74	709	5,316
21.02.2013	125	962	5,769

The German Special Subject Collections System

- established after World War II (1949)
- funded by the German Research foundation (DFG)
- 23 participating libraries fulfill functions of a National Library in these fields
 - complete collection
 - cataloguing
 - archiving
 - guarantee supraregional availability
 - other services



Special Collection Fields – Eastern Europe

	History, government, politics, anthropogeography	Information sciences, education	Language literature, folk-culture	Music	art	religio	Philosophy	Social Sciences	Law
<i>All countries</i>									
Russia	Bavarian State Library	Berlin State Library (1949-1992 BSB)			State and University Library Dresden Zentralinstitut für Kunstgeschichte München a.o.	University Library Tübingen	University Library Erlangen	University Library Cologne ZBW Kiel a.o.	Berlin State Library
Ukraine									
Belorussia									
Poland									
Tcheque Republic									
Slovak Republic									
Slovenia									
Croatia									
Serbia									
Bosnia-Hercegovina									
Montenegro									
Macedonia									
Bulgaria									
Kosovo									
Roumania									
Albania									
Moldova									
Greece									
Hungary	SUB Göttingen								
Baltic States	University Library Greifswald (1949-1997 BSB)								

- Welche Formate stellen wir bei der Digitalisierung her, was wird wozu genau benötigt (tiffs für LZA)?
Tiffs ist Ausgangsmaterial, diese werden langzeitarchiviert und auch für die Erzeugung der OCR verwendet. Aus den tiffs werden jpgs für die Anzeige im Internet erzeugt. Im OCR-Workflow erzeugen wir xml mit Buchstabenkoordinaten (diese Abbyy-XML werden dann noch weiter verarbeitet und konvertiert, bis zu einem sog. "MDZ-TEI") und durchsuchbare PDF: der Nutzer sieht ein Bild der gescannten Seite, dahinter ist der durchsuchbare Text hinterlegt. Aktuell werden tiffs langzeitarchiviert. Langfristig sollen auch die OCR-Daten archiviert werden.

Ich denke, dass wir auch mit JPEG2000 umgehen könnten (da die google-Daten auch in diesem Format sind), aber das nur als Info, falls jemand direkt danach fragt.

Wir nutzen den Abbyy Recognition Server 2.0. Für manuelle Korrekturen verwenden wir die dort integrierte Korrekturstation. (s. Folie 3 in technisches.pptx).

>>Etwas kurz zu Escidoc - was das ist, worauf die Suche basiert ...

Wir nutzen die EsciDoc Infrastruktur, die aus verschiedenen Schichten besteht, wobei PubMan dabei die eigentliche Repository-Software-Komponente ist. Für die Indexierung wird SOLR/Lucene verwendet.

>>Welches Format werden wir konvertieren für die mobilen Endgeräte?

Wir haben ja im Antrag stehen, dass wir EPub anbieten werden. Ob dann aber die Images oder die durchsuchbaren PDF in EPub konvertiert werden sollen, das weiß ich (und vermutlich auch niemand in DB) nicht. EPub scheint mir ohnehin ein Format zu sein, mit dem wir hier im Haus sehr wenig bis gar keine Erfahrung haben.

>>und alles, was Ihnen technisch sonst interessant scheint...

im angehängten PDF, das gut versteckt im Intranet abgelegt ist, sind die einzelnen Digitalisierungsschritte aufgeführt, auch für die OCR-Erzeugung. Vielleicht ist das auch ganz interessant.