Enhancing OJS journals with advanced online reading and viewing capabilities

EKT

PKP Conference 2011, FU Berlin 26/09/2011

Dr. Panagiotis Stathopoulos, **Dr. Nikos Houssos,**
Dr. Ourania Stathopoulou,George Stavrou,Alexandros Soumplis

email: pstath;nhoussos@ekt.gr

Hellenic National Documentation Center /N.H.R.F.

- The Hellenic National Documentation Centre

- Motivation

- Demo

- Components and OJS modifications

- Workflow

- Other extensions, conclusions and future work

The Hellenic National Documentation Centre-EKT

The Hellenic National Documentation Centre (EKT) is the national Infrastructure for collecting, aggregating and provisioning scientific documentation, online information and support services on research, science and technology. Acts as a hub for universities, research centres, enterprises, public and private sector bodies.

Means: Development of interoperable e-infrastructures at the national level for the aggregation, organization and dissemination of scientific content

Aim: Open Access to knowledge for all

Open access infrastructures
repositories, journals and aggregators

– >**4M** pages of content, **x100Ks** of items locally , **>400K** aggregated items

•Digital Library for Research and Technology

– Union catalogue and interloan system for journals, data bases, e-journals

– **15K** journals, **247** libraries

•Research and Technology Indicators and CRIS systems

– publications/citations of Greek institutions in scientific journals

•Library automation technologies

– ABEKT , Greek Edition of the Dewey classification system

– **500+** libraries subscribed at ABEKT helpdesk **2.5K** installations
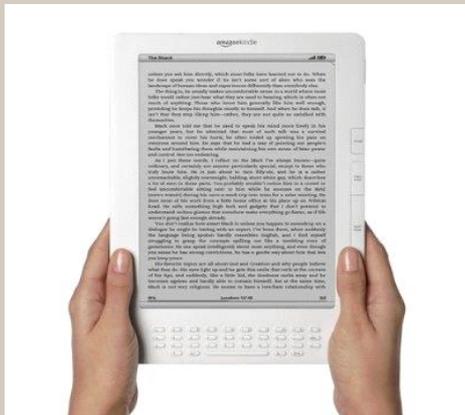
•Participation in European initiatives and projects
LIBER, OASPA, euro-CRIS, Driver, Open AIRE, DART Europe

- Publishing activity commenced at 2007
- 5 electronic journals, peer reviewed and newsletter, mainly in humanities
- Current issues: full electronic process
- Backissues available after digitisation
  - Generally as PDF files
- On the same period interesting developments on viewers and devices

•E-book readers

•Tablets

•Netbooks

•.. new innovative applications and user experiences

And new applications: Google books,

Slideshare, etc.

But also

- Open Library by Internet Archive
- Repositories etc.

•Initial motivation:

–User friendly and highly intuitive reading experience

•Page fliping, thumbnail view, thumbnail reuse, etc.

–Solution for delivery of digitised back issues

•Usually multiMB image PDF files with not intuitive browsing

•But also:

–Page level reference and page level identification

–Large images and datasets visualisation

Journal Management System	High capability viewers	Scalable image servers	Orchestration and conversion management
Editing workflow **Metadata** **Interoperability** **etc.**	**Enhanced intuitive user experience:** **Page be page online reading** **Arbitrary thumbnail views** **Arbitrary zoom in /zoom out levels** **Rotation** **and full text search** **Open standards compliant, Native XHTML no Flash**	**Dynamic generation of image quality and size levels** **Open std: TIFF,JP2000,..** **Highly scalable archs.**	**Batch conversion** **Exploit multicore and server clustering** **Interface with OJS**
OJS	**IA Open Library BookReader**	**IA Open Library Web and Datanode Server and/or djatoka**	**OJS, jp2kdistiller (over djatoka and/or Jasper)**

Djatoka kakadu JP2K library only for non commercial use, JASPER JP2K library is open source needs decoding performance improvements

Interworking best of breed open source systems

- Extensions to OJS galley publishing proccess - new data structures for online reader url and interface to jp2k-distiller

  – Shortly available at the forum

- jp2k distiller - workflow support for automatic batch conversion from PDF (main format) and TIFF to JP2000 and publishing to image server

  – http://code.google.com/p/jp2k-distiller/

- Assumes an image server infrastructure

ΒΥΖΑΝΤΙΝΑ ΣΥΜΜΕΙΚΤΑ

BYZANTINA SYMMEIKTA

ΙΝΣΤΙΤΟΥΤΟ
ΒΥΖΑΝΤΙΝΩΝ
ΕΡΕΥΝΩΝ

INSTITUTE FOR
BYZANTINE
RESEARCH

Galley

Edit a Layout Galley

Galley File Information

Label*	TEXT
	The galley label is used to identify the item's file format (e.g. HTML, PDF, etc.).
Language*	English
File name	963-1341-1-PB.pdf
Original file name	ZIVKOVIC.pdf
File type	application/pdf
File size	409KB
Date uploaded	09-12-2010
	☐ Create the ebook for this galley file.
Replace File	Επιλογή αρχείου Δεν έχει …α αρχείο Use Save to upload file.

Save   Cancel

* Denotes required field

http://byzsym.org/index.php/bz/editor/editGalley/963/920

ΒΥΖΑΝΤΙΝΑ ΣΥΜΜΕΙΚΤΑ

BYZANTINA SYMMEIKTA

ΙΝΣΤΙΤΟΥΤΟ
ΒΥΖΑΝΤΙΝΩΝ
ΕΡΕΥΝΩΝ

INSTITUTE FOR
BYZANTINE
RESEARCH

HOME   ABOUT   USER HOME   SEARCH   CURRENT   ARCHIVES   ANNOUNCEMENTS

Home > User > Editor > Submissions > #963 > Editing > **Galley**

More

Galley

Edit a Layout Galley

Galley File Information

Label*	TEXT
	The galley label is used to identify the item's file format (e.g. HTML, PDF, etc.).
Language*	English
File name	963-1341-1-PB.pdf
Original file name	ZIVKOVIC.pdf
File type	application/pdf
File size	409KB
Date uploaded	09-12-2010
	☐ Create the ebook for this galley file.
	The galley file is being processed in order to create the ebook
Replace File	Επιλογή αρχείου Δεν έχει …α αρχείο Use Save to upload file.

Save   Cancel

* Denotes required field

The reader – thumbnail browsing

depuis les 400 dernières années environ[1]. Le passé le plus ancien des Serbes et des Croates, relaté dans ces chapitres[2], est au cœur d'un débat

———————

1. L'ouvrage a été publié pour la première fois en 1611, en ayant comme base le manuscrit du XVIème siècle (*Codex Vaticanus Palatinus gr. 126*), copié par Antoine Eparchos, érudit humaniste grec venant de l'île de Corfu. À cette occasion, l'éditeur I. Meursius a défini le titre de l'ouvrage, préservé en historiographie jusqu'à nos jours; *Constantini Imperatoris Porphyrogeniti, De administrando imperio, ad Romanum f.*, éd. I. MEURSIUS, Lugduni Batavorum 1611. Le plus ancien manuscrit du *DAI*, datant du XIème siècle (*Codex Parisinus gr. 2009*), n'a été publié qu'en 1711, par A. Banduri, habitant de Dubrovnik; *Imperium Orientale sive Antiquitates Constantinopolitanae in quatuor partes distributae*, éd. A. BANDURI, Parisiis 1711. I. Bekker a republié le *DAI* en 1840, en s'appuyant sur Meursius et Banduri; *Constantinus Porphyrogenitus De thematibus et De administrando imperio*, éd. I. BEKKERUS, I-II, Bonnae 1840 (désormais: BEKKER). La meilleure édition critique de cette oeuvre est, *Constantine Porphyrogenitus De administrando imperio*, éd. R. J. H. JENKINS – Gy. MORAVCSIK (CFHB 1), Washington, D.C. 1967 (désormais: *DAI* I). Le texte a été commenté par F. DVORNIK – R. J. H. JENKINS – B. LEWIS – Gy. MORAVCSIK – D. OBOLENSKY – S. RUNCIMAN, *Constantine Porphyrogenitus De administrando imperio*, II: *Commentary*, London 1962 (désormais: *DAI* II). L'édition la plus récente est, *Konstantin Bagrùnorodnóy*

1. POST: galley_id, article_id, journal,url to pdf galley file,article metadata

jp2kdistiller

OJS

4. Image server Update database/fs updates

3. Parallel thread groups

djatoka

djatoka

6. IA viewer instance

5. OJS callback: galley_id, article_id, journal
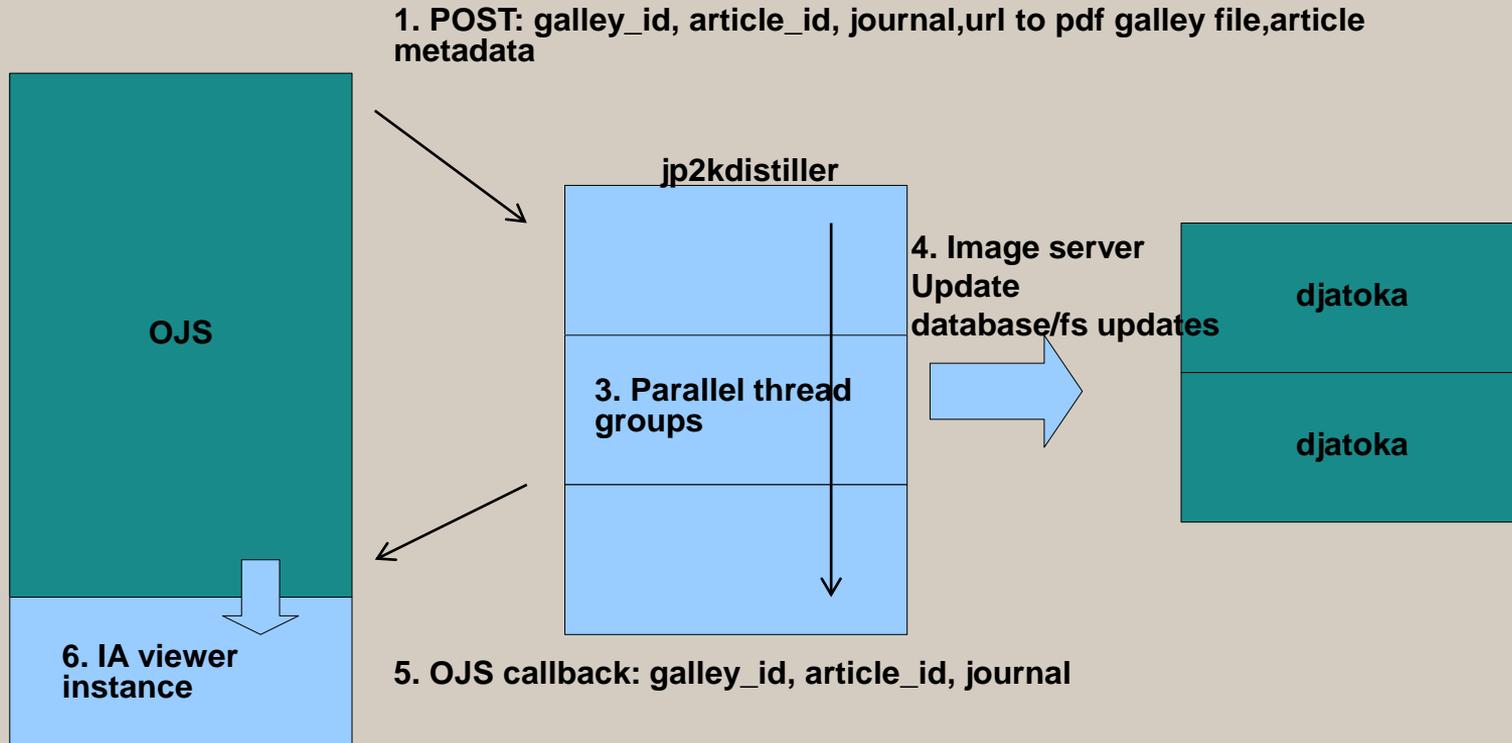
OJS modifications:

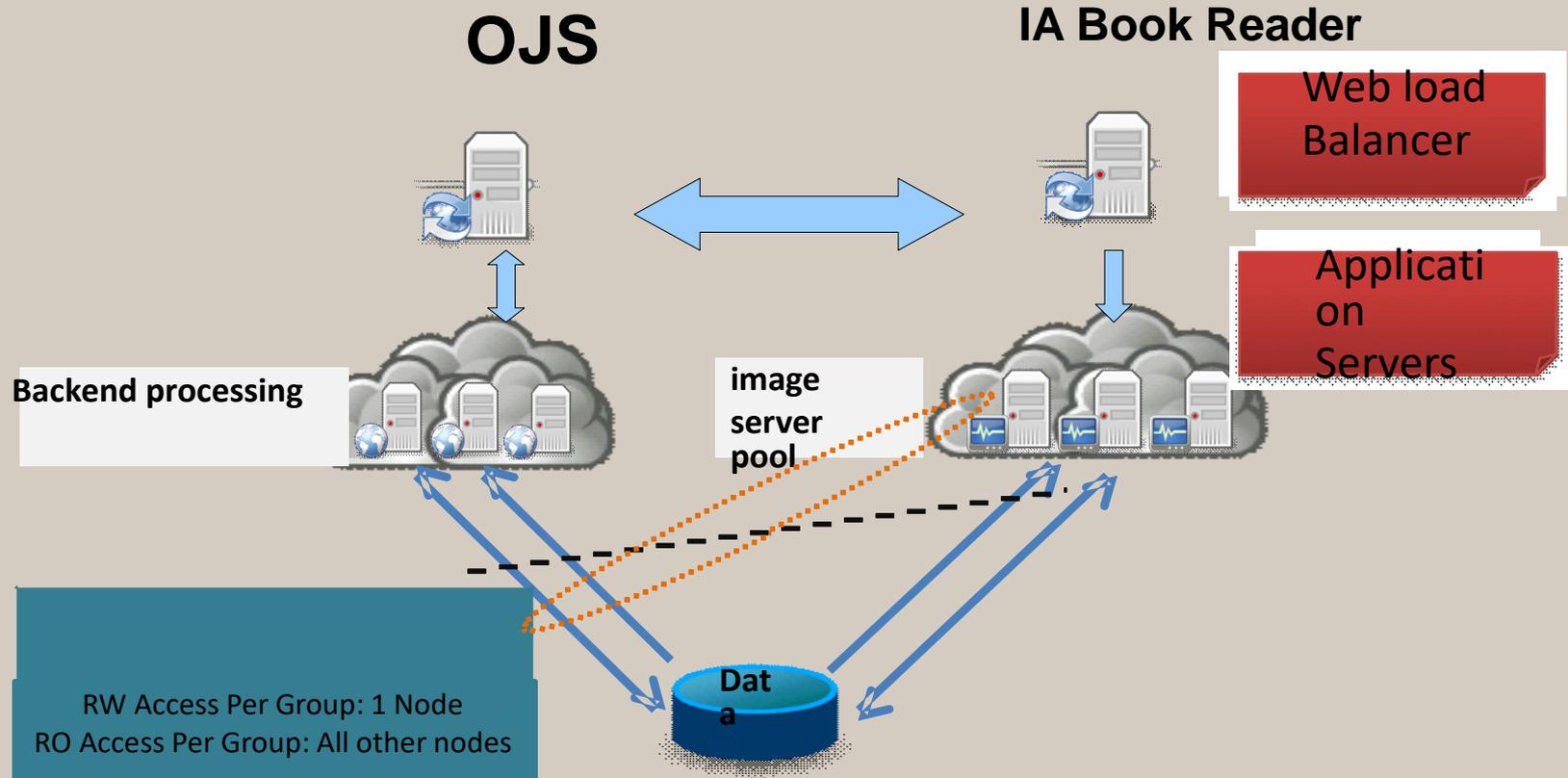–New data stored: url of reader and flag for reader status (no reader, ready, pending)

–APIs for communication with jp2kdistiller

•Editor request a galey file conversion to online reader

•J2pkdistiller retrieves the pdf and converts in batch and in parallel

–When all threads finish calls back to OJS

•OJS sets new flag and page by page online reading activated

•Python script for batch conversion of PDF/TIFFs/JPGs to JPEG2000,publish to djatoka

–Multithreaded application, tunable number of threads x5-x10 speed in comparison with single thread batch convert

•Interworking with OJS

–Notification from OJS for conversion

–Callback function to OJS in order to activate reader

–Transparent for editor

Scalable multiered architecture

OJS

IA Book Reader

Web load Balancer

Applicati on Servers

Backend processing

image server pool

Dat a

RW Access Per Group: 1 Node
RO Access Per Group: All other nodes

• Configurable browsing by any metadata field

– Configuration through the OJS GUI by the Journal Manager

– Distributed as an OJS plug-in

• Automatic creation of cover pages for PDFs

– Insert article metadata into cover page template

– Developed using a separate FOSS library created by EKT

• New browsing capabilities by interfacing with external systems

•Transparent to editors - highly scalable – automated conversion infrastructure

•Lots to wish about:

–Full text search/highlighting is the next step

–Linking with an enhanced repository based backed

–Interfacing to other format conversion apps,e.g. Doc to pdf, odt to pdf, etc

–API specifications

–Jasper enhancements

Thank you for your attention!