Open Source Software to Manage Digitization Projects – The Kitodo Example

Frank Ulrich Weber; Zeutschel GmbH; Tübingen, Germany

Abstract

Zeutschel today has more than fifty years of experience in the field of developing analog and digital exposure systems. The time goes by and now the digital business displaced the analog. Besides just producing digital images it is more and more interesting how to manage the digitization and how to make a valuable source of information out of the digital nuggets. Those questions arise in all sizes of institutions like libraries, archives, museums and service providers. To offer them, beside just deliver the scanners, a complete solution for the whole range of digitization we decided to support the open source software suite Kitodo as one important part in our solution portfolio. That solution generates a valuable surplus for the client.

Joining the Kitodo Association

The whole world went digital and linked. To master the first challenge we have to develop and manufacture digital scanners besides the analog microfilm. On that basis the second challenge arise. In a linked world of digital content we have to offer our clients more than just a scanner. We realized that the digitization process itself can be a very painful process without the right software tools to manage and monitor it. And at the end of that process the digitized objects should be available and searchable for the interested audience. It was nearly ten years ago that we join the development team of the open source software suite Goobi (now renamed to Kitodo) and later the "Kitodo – Key to digital objects" association. Using, developing and supporting open source products is part of our company strategy. Zeutschel is a member of the Kitodo association and represented in the association board.

A Little Bit of History

The development of Kitodo started over ten years ago within a huge digitization project of the State and University Library of Göttingen founded by the DFG (German Research Foundation). It was always the goal not to use the system just as an in-house solution. The need for an easy and flexible tool to manage digitization exists within other institutions, too. That was the reason why in 2010 four big libraries and the Zeutschel GmbH signed a Goobi Release Management Agreement and started regular meetings as a Goobi Community Board to push the distribution and further development. Because it is a community based and driven open source system it is vital to include the institutions that are using it by organizing regular meetings. That was the beginning of the half year open user meetings in 2011. To establish a stable environment for the system an association was founded by twelve libraries and private companies in 2012. With the association it is possible to achieve legal and financial stability and independence. The most important thing about this point is to run a reliable release management and a controlled ongoing development. To resolve some legal problems with the name

Goobi the 2016 annual meeting of members decided to rename Goobi to Kitodo.

Current Status

Kitodo is:

- an open source software suite for the production and presentation of digital objects and collections
- a growing community of users
- a legal entity and association "Kitodo. Key to Digital Objects e.V.

The fast growing international community of users (over 40 institutions and service providers), the Kitodo association and the release management offers some important advantages:

- support and contact via mailing lists
- documentation
- meetings of users
- coordinated development
- continuous development
- feedback from many different users and usage scenarios

The different institutions, service providers and usage scenarios lead to a huge supported set of different media types:

- address books
- archive material
- letters
- prints
- photos
- graphics
- manuscripts
- books
- incunabula
- maps
- inheritances
- school books
- drawings
- newspapers
- magazinesland registers

-

Flexible Architecture The two independent layer

The two independent layers of Kitodo offer a flexible setup of the system. If you just need a workflow tool or a tool for make your digital collections accessible it's up to you to choose. Both layers of Kitodo can be used completely independently. The interface between the two layers consists, simply spoken, solely of images and a METS file.



Figure 1: The two independent layers of Kitodo

A main focus within the development of Kitodo was and still is the idea of platform independence. On the client side only a newer web browser is required. On the server the basis is Java for Kitodo.Production and PHP for Kitodo.Presentation.

A major advantage of Kitodo.Presentation is the used Content Management System (CMS) TYPO3. You are free to either use a complete preconfigured website or use all the options of configuring and adding content offered by the CMS.

To fulfill the requirements of different users, projects and media types the system has to be very flexible and much more important support international standards like METS/MODS/ALTO or OAI-PMH2 to deliver digital objects to portals like Europeana.

The flexibility of the system can be measured on one hand by looking on the over forty international institutions successfully using Kitodo for different projects.

On the other hand the flexibility can be measured by the configurable integrated functionalities.

The Benefits of Workflows

To organize and manage the different tasks of digitization projects efficiently, the first important part is the support of the needed workflows.

Home page	Home page >> Process management					
- General - Technical background Current users - Workflow -	Process management Current processes Create new process					
Ny taoka			Select pre-defined filter:	proj:Digi		
Search for volume	Hts: 161		iter:			
Current processes	Process title 🌝 🛦	Status VA	Project VA	Ac	Activity	
Process templates	B DzieBe_1980_10_01		DigiZED	B + +	L METS POP	
- Administration -	@ DzieBe_1980_11_01		DigiZED	B	L BHETS FOR	
Users	■ DzieBe_1980_12_01		DigiZED	E	E BALLS HO	
Jeer groups	® evenunepe_PPN04171235X	-	DigiZED	Eb 10 4 1 1	S I HETS HE	
Projects	@ grypchri_330102311		DigiZED	25, 17 A 2 10	L BHETT PER	
Rule sets Dockets	@ HildTasc_12345		DigiZED	80, 10 A 1 1	E. B. HETS HE	
DAP groups	8 horsendsi_264428472	-	DigiZED	E	S. III HETS HE	
Task manager	@ InstfidiB_9789949765225_20161103	-	DigiZED	B 1 10	EL METS HE	
Module manager	iii MrkMa_1965_1	-	DigiWunschbuch	E	E BALTS HO	
- User data -	8 NLE_test_001	-	Ajalehed enne 1945	E 1 1 *	MINISTER PLA	
User configuration Change password	Create new process		H]			

Figure 2: Workflows

Kitodo at any time allows authorized users to have a survey of all running tasks. Beside that the granular assignment of users to tasks, projects and user groups enables that a user sees only what he is able to do. That makes it very easy for the user to work with the system without interrupting the work of others or getting confused by too much irrelevant and needless information and functionalities.



Figure 3: Metadata

The powerful METS-Editor allows to paginate, to structure and to enrich the document with additional metadata. The basic metadata can be either fetched from a catalog system by using a standardized interface like Z39.50/SRU or typed in manually. The system allows adding additional functionalities like using OCR results for metadata creation or using templates for the basic structure of a document. The required meta and structure data fields are free configurable.

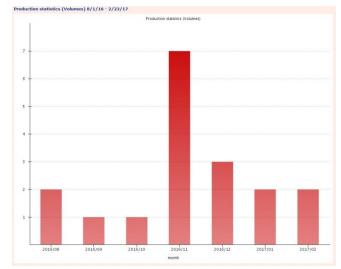


Figure 4: Statistics

To monitor the progress or analyze a finished project, Kitodo comes with comprehensive statistics, like:

- production throughput
- duration of tasks
- storage calculator
- error tracking
- number of images

The digitization is not a self-purpose. The main aims of digitization as we know them are conservation and accessibility. The fist aim you can achieve by using Kitodo.Production in combination with a long term preservation system like Rosetta (ExLibris). Procedures for the transfer to Rosetta are available for

the system. The second aim you can achieve by using Kitodo.Presentation.

Accessibility

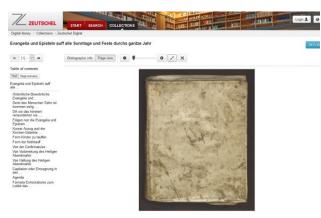


Figure 5: Kitodo.Presentation

Kitodo. Presentation provides a wide variety of functionalities:

- granular access control
- collection management
- fast search index (Lucene)
- visualization of structure, metadata and images
- based on the powerful CMS TYPO3
- free configurable metadata fields used for:
 - indexing
 - o tag clouds
 - o sorting
 - o facetted browsing
 - extended search
- OAI-PMH2 interface
- ALTO Support (for indexing, highlighting and displaying full text)
- Basket feature

Current Projects and Developments

The flexible architecture allows a great variety of different users and usage scenarios. Frequently we are confronted with the belief that Kitodo is only for the big institutions like state or national libraries. Sure, these institutions can effectively use the software to manage huge digitalization projects, but there are institutions only one person uses the system to organize its daily work, too.

Within the first years of Kitodo (Goobi) there was a strong focus on the requirements of libraries. As time goes by many institutions with different material, but the same need to do a fast, well organized and effective digitization, came to the point that Kitodo fits to their need on the whole. So it is fact that today archives, land record offices, museums and service providers can use the system effectively, too.

Conclusions

To work with open source software you have to pass some challenges of a cooperative development community:

- development planning an controlling
- establish programming guidelines
- coordinate the different views of public institutions, private companies and free developers
- cooperation between public institutions, private companies and free developers
- sustainability (support and maintenance)

If the challenges are mastered you gain of the huge experience of all the different members of the community to create a sustainable environment of future development and success. A good example for that is the 1.3 Mio. Euro DFG project to enhance the possibilities of Kitodo.Production. This project started in summer 2016 and will be finished in 2018. Main focus is

- new system architecture
- improved usability
- advanced import and export routines
- more standardized interfaces and formats

References

- [1] http://www.kitodo.org/en/
- [2] http://www.zeutschel.de/en/produkte/workflowloesungen/Kitodo.htm1
- [3] https://github.com/kitodo

Author Biography

Frank Ulrich Weber studied business sciences with focus on business informatics at the University of Hohenheim. At this time he and his colleagues developed, as part of a study project, a document delivery system for the library of the university. Since over ten years now he is product and project manager of the Zeutschel GmbH. His focus is on creating the best solution for every client concerning digitization.