



národní  
úložiště  
šedé  
literatury

### **ARCLib – LTP řešení pro knihovny**

Pavlásková, Eliška; Vašek, Zdeněk  
2018

Dostupný z <http://www.nusl.cz/ntk/nusl-387443>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Licence Creative Commons Uveďte původ-Zachovejte licenci 4.0

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 19.05.2022

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní [nusl.cz](http://nusl.cz) .



komplexní řešení pro dlouhodobou archivaci  
digitálních (knihovných) sbírek

## **ARCLib – LTP solution for libraries**

This presentation is licensed under the Creative Commons: [CC-BY-SA-4.0](https://creativecommons.org/licenses/by-sa/4.0/), via <http://repozitar.techlib.cz/record/1297>

Mgr. Eliška Pavlásková, Ph.D.

PhDr. Zdeněk Vašek, Ph.D.

Library of Czech Academy of Sciences

11th Conference on Grey Literature and Repositories

Prague, October 24, 2018



# Basic information

- Research project funded by Ministry of Culture of the Czech Republic (NAKI program)
- 2016-2020
- 23 mil. CZK
- Collaboration of Library of Czech Academy of Sciences, National Library of Czech Republic, Moravian Library, and Masaryk University
- **Development of open source long-term digital preservation solution ARCLib**

[www.arclib.cz](http://www.arclib.cz)



# Objectives of the Project

- **Open source** solution ARCLib for long term preservation of digital data
- The methodology for **logical** preservation of digital data
- The methodology for **bit-level** preservation and proposal of a storage solution
- Test of the solution in **pilot**



# Expected Properties of the Solution

- **Interoperability** with National Library of Czech Republic
  - Use of NL CR standards
- Open solution with possibility of further development
  - **Additional types of data** – Producer SIP profile
  - **System development** – customization for different fields – e.g. archiving
- **Open source** alternative to commercial software
- Focused on **national, special and regional libraries**
  - Supported by methodologies for system implementation and management



# The methodology for logical preservation of digital data

- General/theoretical part
  - LTP in general, core standards, strategies and best practices
- Application of theoretical knowledge in ARCLib implementation
  - Implementation of long term preservation requirements into design of concrete SW solutions.
  - Description of the system with regard to implementation.
- Implementation part
  - Recommendation for ARCLib users with emphasis on specific data types.
- Certified 2017 – will be updated periodically



# The methodology for bit-stream preservation

- General/theoretical part
  - Long term preservation strategies
  - Legislative requirements on storage of digital archival materials
  - Best practices and certification requirements
- Application of theoretic knowledge in ARCLib
- September 2018 – submitted to certification

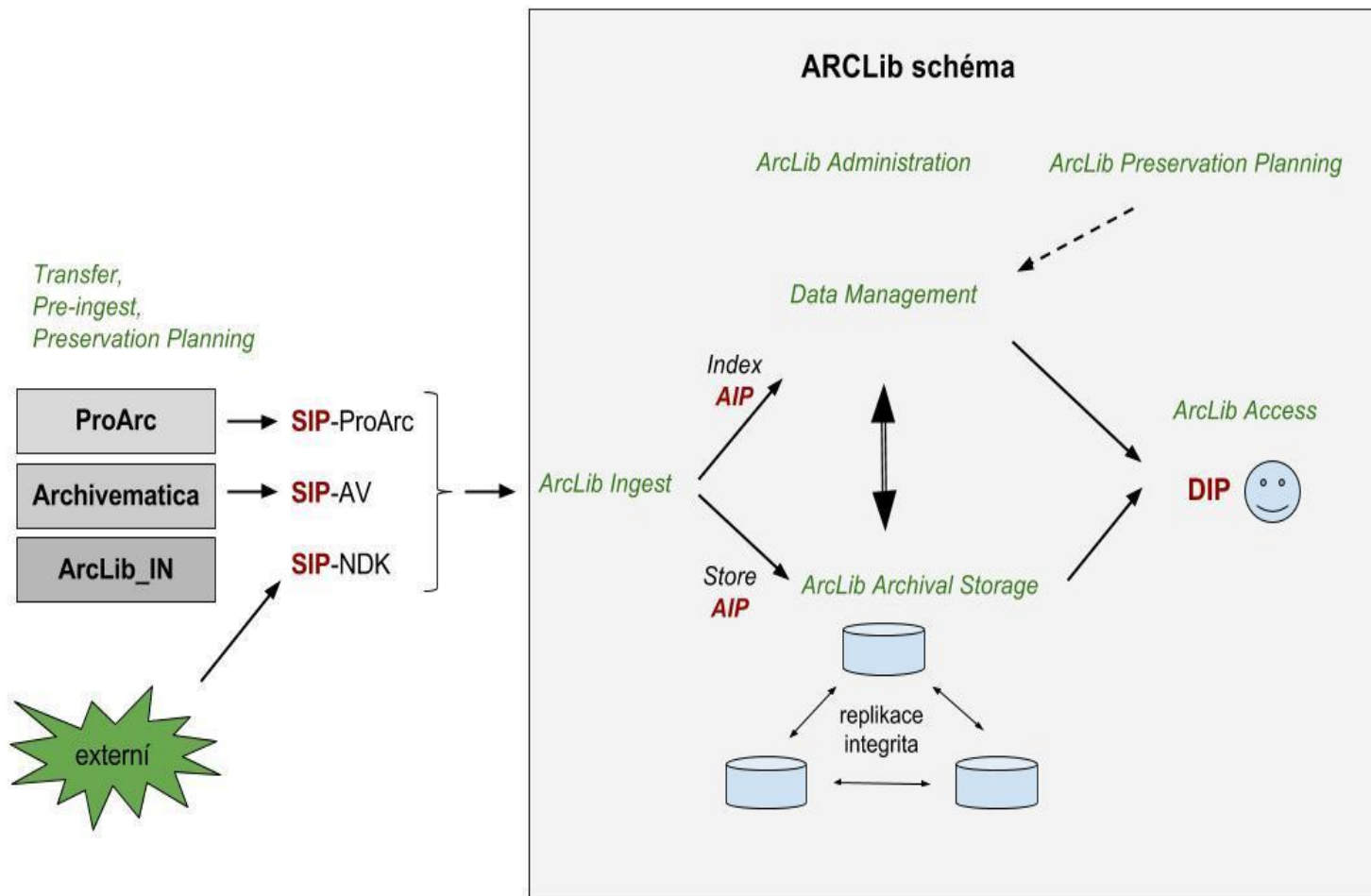


# Description of ARCLib

- Implementation of OAIS (ČSN ISO 14721)
- Do not replicate a function of ProArc or Archivematica (these tools are used for SIP creation).
- Development is focused on core modules:
  - ARCLib Ingest
  - ARCLib Data management
  - ARCLib Archival storage
  - ARCLib Administration
  - ARCLib Access



## Scheme





# ARCLib Ingest

- Requires input of full SIP
- System is able to process :
  - ProArc NDK monographs
  - ProArc NDK periodicals
  - ProArc native monographs and periodicals
  - ProArc audio documents
  - NDK monographs and periodicals
  - Archivemata DSpace
  - Archivemata General
  - NDK electronics documents
- Functions
  - SIP validation
  - Metadata extraction from SIP
  - Metadata creation
  - Processing according producer profile requirements
- ARCLib AIP XML
  - Metadata profile based on METS, PREMIS



# ARCLib Data management

- Management of AIP
- Search and indexation (also as an API)
  - Descriptive metadata
  - Technical metadata
  - Administrative metadata
- Display of AIP content
- Metadata editing
- Reporting

Users

Producers

Producer profiles

Ingest

Ingest routines

Ingest batches

Validation profiles

Sip profiles

Storage administration

Workflow definitions

Deletion requests

Search queries

AIP search

## AIP search

## Sort

Producer ID  

## Producer ID

Contains 

## User ID

Contains 

## State

Contains 

## SIP version number

Greater or equal  Less or equal 

## ID of previous SIP version

Contains 

## XML version number

Greater or equal  Less or equal 

## ID of previous XML version

Contains 

## Document

Contains 

## Root

## Label

Contains 

## Type

Contains 

## Sip ID

Contains 

## Header

## Created

from  to 

## XML ID

Contains 

## Authorial ID

Contains 

## Descriptive metadata

## Generic Dublin Core

Contains 

## Specific sets of Dublin Core

Contains 

## Aggregated extracted technical metadata

## Formats

## Date created by application

from  to 

## File format

Contains 

## Format registry key

Contains 

## Format registry name

Contains 

## Creating application name

Contains 

## Creating application version

Contains 

## Preservation level value

Contains 

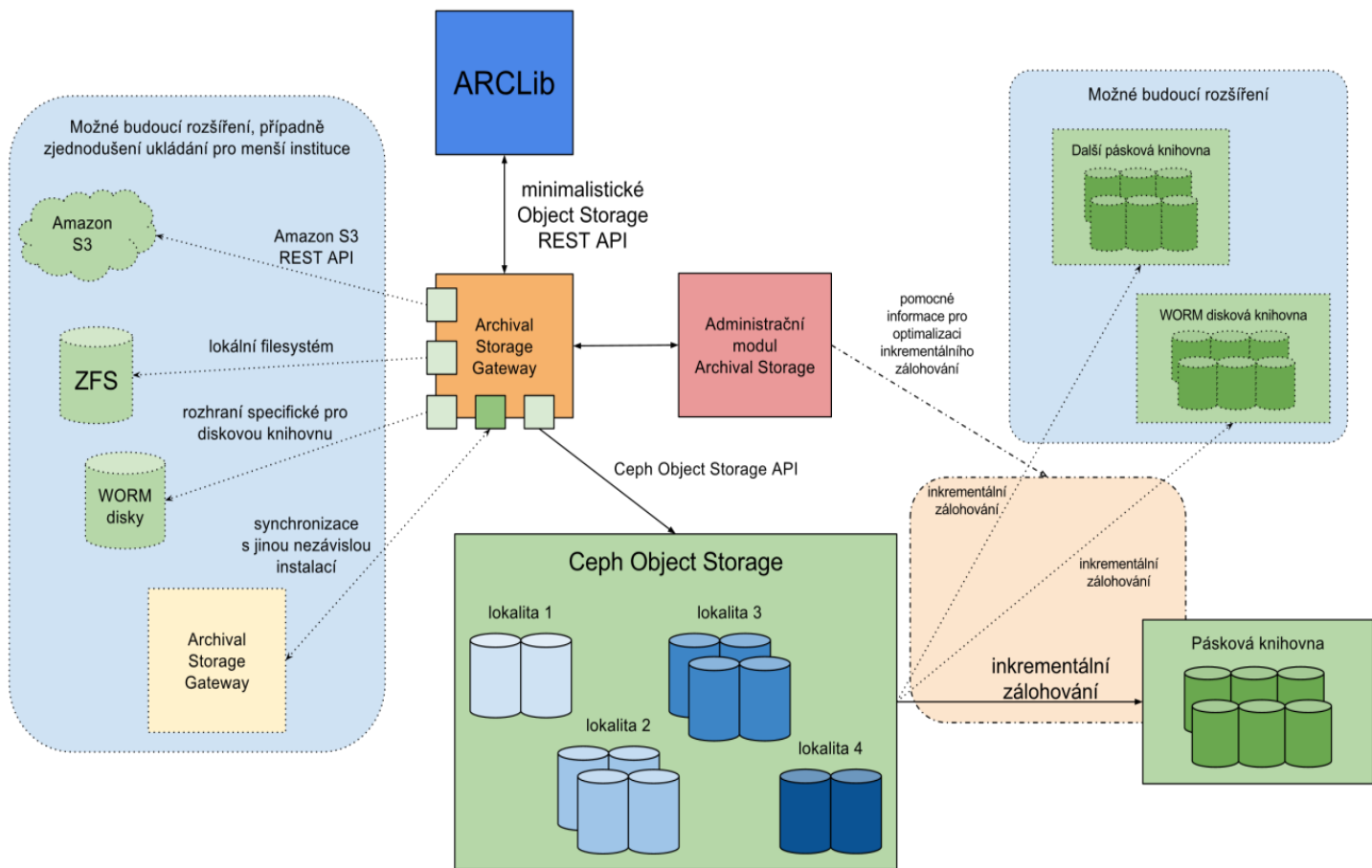
## Scanner model serial no

Greater or equal  Less or equal 

## File count

Greater or equal  Less or equal

# ARCLib Archival storage



Users

Producers

Producer profiles

Ingest

Ingest routines

Ingest batches

Validation profiles

Sip profiles

Storage administration

Workflow definitions

Deletion requests

Search queries

AIP search

## Storage administration / local storage



Delete

Storage

[Synchronization information](#)**Name**

local storage

**Host**

localhost

**Port**

0

**Priority**

10

**Storage type**

FS

**Configuration file**

```
{  
  "rootDirPath": "/opt/archival-storage/data"  
}
```

**Poznámka** Write only Reachable

Storno

Save and close



# ARCLib Administration

- Ingest workflow configuration
  - Relevant registers – validation profiles register, script register...)
- Management of third-party tools
- User management and authentication



# ARCLib Access and Preservation Planning

- **Dark archive** – not intended for end users
- Limited access options – only export functions, **DIP equals AIP**
- Format registry – connected to **PRONOM**
- Other functions of preservation planning are out of the scope of the system (supervision by **National Library CR**)





# Current state of the project (October 2018)

- **The methodology for logical preservation of digital data** (certified and published) - <http://hdl.handle.net/11104/0282107>
- **The methodology for bit-stream preservation** – submitted to certification
- **Test version of ARCLib** running in Library of Czech Academy of Sciences
- **2020** – assumption of existence of the whole ARCLib solution and pilot in Library of Czech Academy of Sciences



**Thank you for attention!**

Mgr. Eliška Pavlásková, Ph.D.  
[eliska.pavlaskova@ruk.cuni.cz](mailto:eliska.pavlaskova@ruk.cuni.cz)

PhDr. Zdeněk Vašek, Ph.D.  
[zdenek.vasek@ruk.cuni.cz](mailto:zdenek.vasek@ruk.cuni.cz)