Deutsche Nationalbibliothek

Realizing a Preservation Strategy

Format Migration with koLibRI

Tobias Steinke, German National Library

Overview

- The German Project kopal
- The open source software koLibRI
- The Migration Manager

kopal
Data into the future

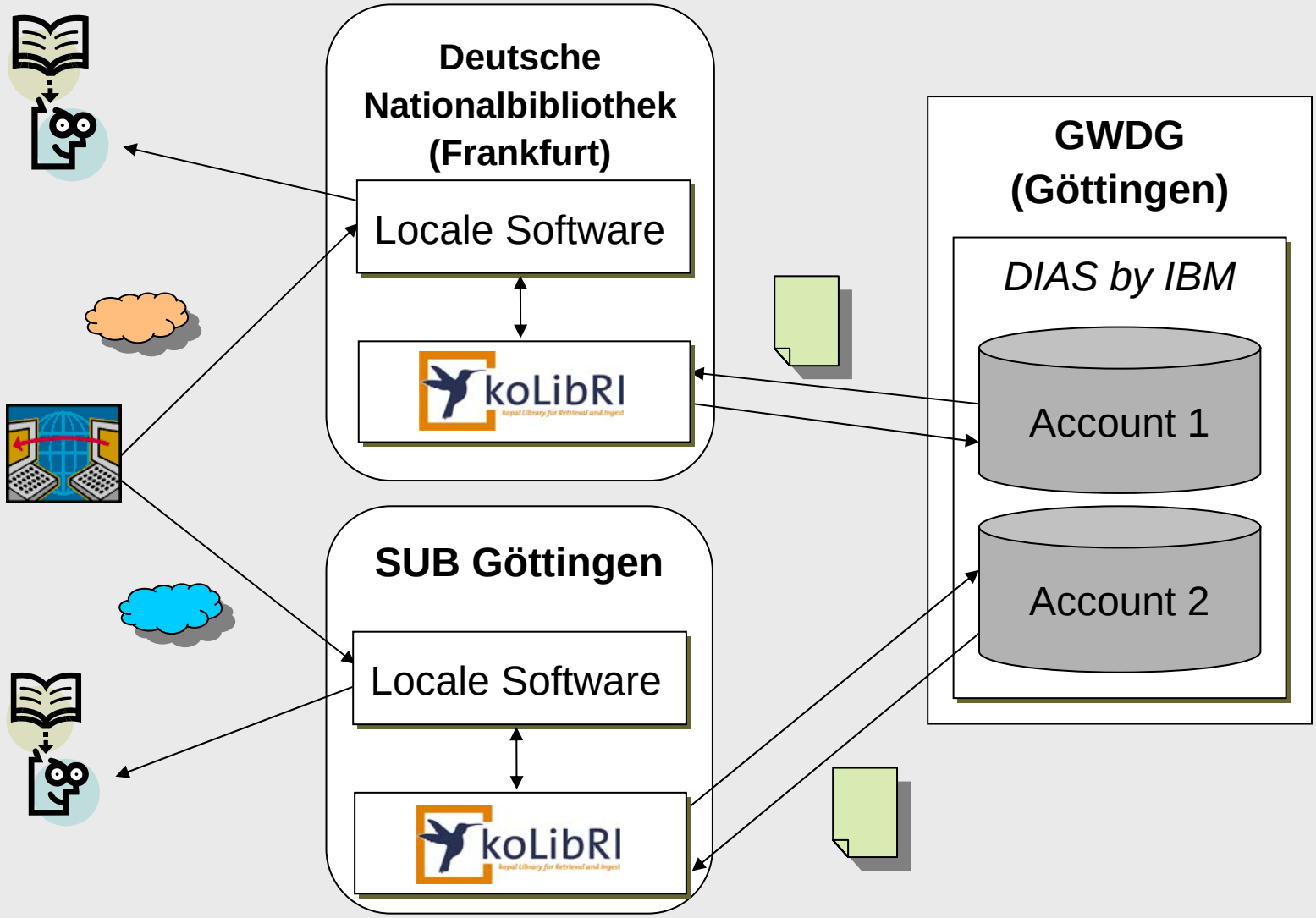
- Co-operative Development of a Long-Term Digital Information Archive
- German project (2004-2007)
- Partners: German National Library (DNB), Göttingen State and University Library (SUB Göttingen), IBM Germany, Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen
- Based on OAIS reference model
- Core: DIAS by IBM
- Software tools: koLibRI (open source)

kopal
Data into the future

- Multi-user support
- Universal Object Format: METS, Long-term preservation Metadata for Electronic Resources (LMER)
- No restrictions on file formats for ingest in archival system
- Preservation based on file format migration

kopal
Data into the future

Deutsche Nationalbibliothek



- kopal Library for Retrieval and Ingest
- Open source Java software library by SUB Göttingen and DNB
- Flexible and generic adaptation of the integration of DIAS by the partners
- Configurable and expansible
- Available at

  http://kopal.langzeitarchivierung.de/

- Unique interfaces for different systems
- Modular workflows for different objects
- Extraction of existing metadata
- Generation of technical metadata with JHOVE
- Generation of the METS file and the submission information package (Universal Object Format)
- Capsulated communication with DIAS for ingest, access and administrative functions
- Migration Manager: Management and execution of file format migration

Migration Manager

- Modul of koLibRI 1.0
- Uses other koLibRI modules for workflow handling, extraction and generation of metadata
- Fetches requests for necessary migrations
- Perform the migration by querying the archival system and executing a suitable conversion tool

Migration Scenario (1)

- Event listening service reacts on certain migration recommendations by an external source
- Requirements for the migration activity are transformed in a suitable query for the archival system (e. g. transfer all non-animated black & white GIF files to JPEG2000)
- Every object of the response list by DIAS is accessed as DIP in the Universal Object Format

Migration Scenario (2)

- METS file within the UOF is analysed, dependencies of the files are resolved (e. g. GIF - HTML)
- Chosen tool for converting is executed for every affected file
- New METS file is generated including history data about the migration
- New SIP is packed and ingested in the archival system

Conclusion

- Migration Manager as part of koLibRI is a starting point for further developments, but works already
- International co-operation is needed to build up a network for recommendations on migrations and conversion tools

Thank You!

- kopal: http://kopal.langzeitarchivierung.de/