UNIVERSITY INSTITUTIONAL REPOSITORY
AND ITS PLACE IN THE UNIVERSITY
INFORMATION INFRASTRUCTURE
Outline

• CTU and its repository
• Place of the repository within university information system
  – technical and process solution of the interconnection of the repository with subsystems of the university IS
  – design of the optimized data flow and data management between the components
• Assessing the solution's benefits
Our background

• Czech Technical University in Prague (CTU)
  – the biggest and oldest technical university in the Czech Republic (since 1707)
  – 8 faculties, 5 specialized institutes
  – 28,000 active users in university system (22,000 students / 2,200 academic staff / 3,800 other staff)

• University repository
  – DSpace platform (open source solution, most common in CZ – large user community)
  – current version: 1.7 with Manakin user interface
Phases of repository implementation

• 1st phase ("beginnings")
  – legislative change regarding ETDs → repository launch in 2008 (DSpace 1.4)
  – workflow for ETDs (SIS to DSpace)
  – re-use of data from SIS, automated loading of university organizational structure (OSSS)

• 2nd phase ("foreseeing future")
  – 2010 – workflow for research outputs (CRIS to DSpace)
  – current trends; connection to international infrastructures
  – 2014 – university demands the solution for research outputs (conditions of Horizon 2020) → we are ready

• 3rd phase ("adding value")
  – 2013 – Open Journal Systems for university research journals (pilot with Acta Polytechnica
  – assigning DOI to university publications (library as a coordinator)
Subsystems of the university information system involved

• Source systems
  – CRIS (Current Research IS): research outputs
  – SIS (Study Information System): Electronic Theses and Dissertations
  – OJS (Open Journal Systems): university journals / conference proceedings

• Supporting systems
  – AAI (Authentication and Authorization Infrastructure)
  – OSSS (Organizational Structure Source System)
  – DMS (Document Management System)

• Main target systems
  – Repository (DSpace; records and full texts from various subsystems)
  – Library System (Aleph; library entities’ records, incl. ETDs’ records, users’ records)
Overview of main subsystems

- **Source systems** – CRIS (Current Research IS): research outputs, SIS (Study Information System): Electronic Theses and Dissertations, OJS (Open Journal Systems): university journals / conference proceedings

- **Supporting systems** – AAI (Authentication and Authorization Infrastructure), OSSS (Organizational Structure Source System); DMS (Document Management System)

- **Main target systems** – Repository (DSpace; records and full texts from various subsystems), Library System (Aleph; library entities’ records, incl. ETDs’ records, users’ records)
Identity management & Authentication

- Support unified
  - Identity management
  - Identification and authentication

- User identifiers
  - Personal number
  - UID
Organizational structure synchronizing

- Source – OSSS
- Automatic transfer to
  - Repository
  - Library system
- Problems with mapping
  - Deleting department
  - Splitting and joining of department
SIS & CRIS details

Diagram showing the interaction between SIS, CRIS, Library System (Aleph), Repository (DSpace / HS), and DMS via ESB (Enterprise Service BUS) services interface.
Summary of the whole solution

- Authentication System
- Identity Management
- Full Automatic Processes
- SIS
- CRIS
- Export of Journals
- Authentication
- Structure & Identity sync
- Files & Metadata
- 『sql』
- «sql:select»
- Handle-ID
- «sql:select»
- Handle-ID / BitStream-ID
- DMS
- ESB farm
- Logger
- Append
- TABG
- Merge
- ALEPH Interface
- ESB (Enterprise Service BUS)
- Reporter
- Logs export
- «api»
- Transform
- DSpace Interface
- Handle-ID
- «soap»
- «sql:select»
- WS
- «sword»
- [n-threads]
- «jms»
- [n-threads]
- [n-threads]
- [n-threads]
- OSSS
- OJS
- Repository (DSpace / HS)
Final assessment and benefits of the solution

• Data reuse
  – student / researcher works with a familiar environment
  – data is inserted only once

• Automation of processes
  – e.g. automatically managed repository structure

• On-line solution
  – supporting manual, scheduled, full event data entry

• Scalability
• Expandability
  – simple connection / inclusion of another system
Current state & perspectives

• New RESTfull API of DSpace
  – Based on DSpace system RESTfull (Jersey)
  – Full writing operation (community, collection, items, bitstreams)

• Incremental update of organizational structure through ESB

• OJS change export mechanism similar to SIS/CRIS
Practical example

• Organizational structure synchronization
  – Example of full synchronization from scratch
  – Remote execution on university servers

• Steps
  – Pure installation of DSpace 4.0 (community-list)
  – Example of full synchronization (remote cmd)
  – Check the results (community-list)
THANK YOU FOR YOUR ATTENTION!

QUESTIONS?

Věra PILECKÁ
vera.pilecka@uk.cvut.cz