StudentWeb/FS integration with institutional repositories for Master's Thesis submission

Richard Jones, Cottage Labs, richard@cottagelabs.com
Kristian Salcedo, University of Oslo, k.r.salcedo@ub.uio.no

DUO
Digitale Utgivelser ved UiO

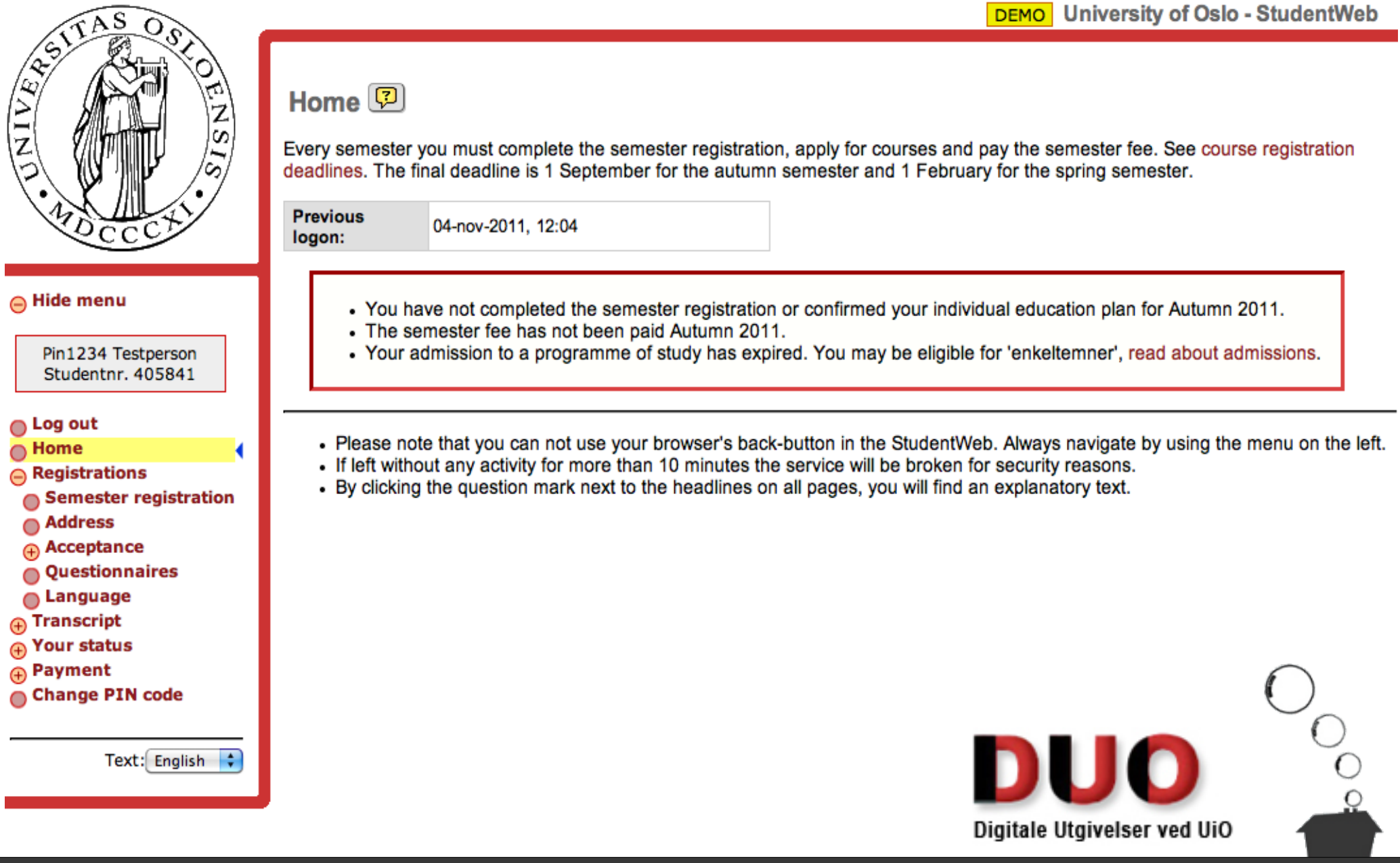
Background

- new IR system platform - DSpace
- old ideas become new: Integration of student info system (StudentWeb) with IR

- national and institutional scope:
  - re-use of authoritative metadata
  - StudentWeb = one-stop shop for students and staff
  - national solution
  - transition from print to electronic

DUO
Digitale Utgivelser ved UiO

What is StudentWeb?



Why integrate StudentWeb/FS with institutional repository

- Ensure all Master's Theses are archived, irrespective of grade
- Ensure quality of data
- Be able to deliver the relevant theses on to other systems such as Riksarkivet
- Be able to make those Master's Theses which make a pass grade available Open Access

DUO
Digitale Utgivelser ved UiO

Technical solution proposal - It began ...

DSpace workshop, Oslo, September:

- StudentWeb/FS - *Systems owner*
- University of Oslo - *Nye DUO project*
- Cottage Labs - *Domain Expert, Independent 3rd Party*
- University of Bergen - *Stakeholder*
- Cristin - *Interested Party*
- University College Oslo - *Stakeholder*

DUO
Digitale Utgivelser ved UiO

Technical Approach

- SWORD 2.0 for transport protocol
  - http://www.swordapp.org/
- BagIt for content packaging
  - http://tools.ietf.org/html/draft-kunze-bagit-06
- Open Source Software for DSpace
  - https://github.com/nye-duo

DUO
Digitale Utgivelser ved UiO

What is SWORD 2.0

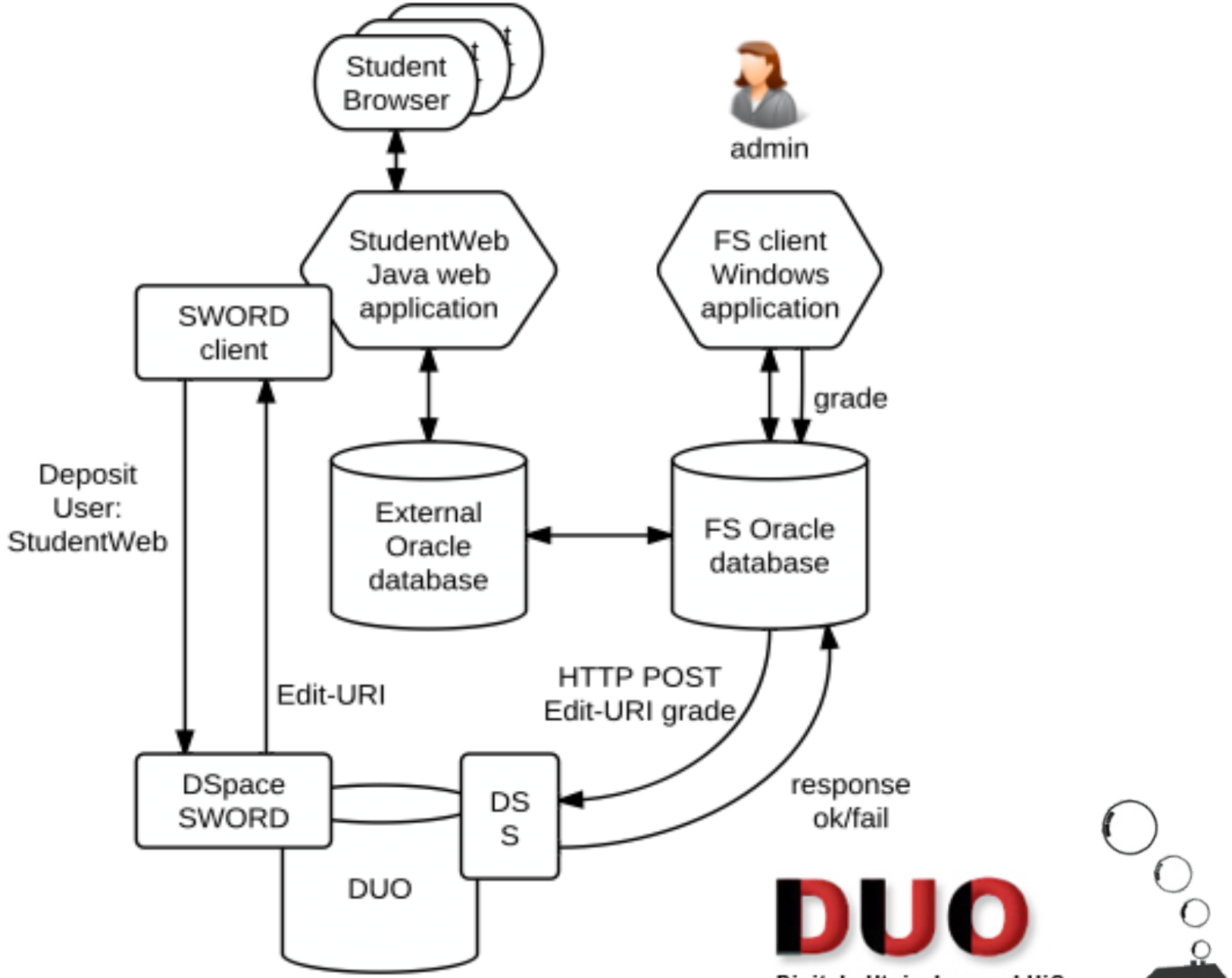
- 1.x developed by JISC in the UK for basic repository deposit
- 2.0 developed to support full CRUD for a range of scholarly systems
- Client implementations: Java, Python, PHP, Ruby, (C#)
- Server implementations: DSpace, EPrints, Fedora, Connexions

Why use SWORD 2.0 here?

- Supports full CRUD over a RESTful web service, so is functionally rich enough to provide the solution
- Is an emerging standard in our sector
- Has implementations for the major open source repository platforms
- Has existing client libraries already written
- No obvious competitor (WebDAV, CMIS, AtomPub?)

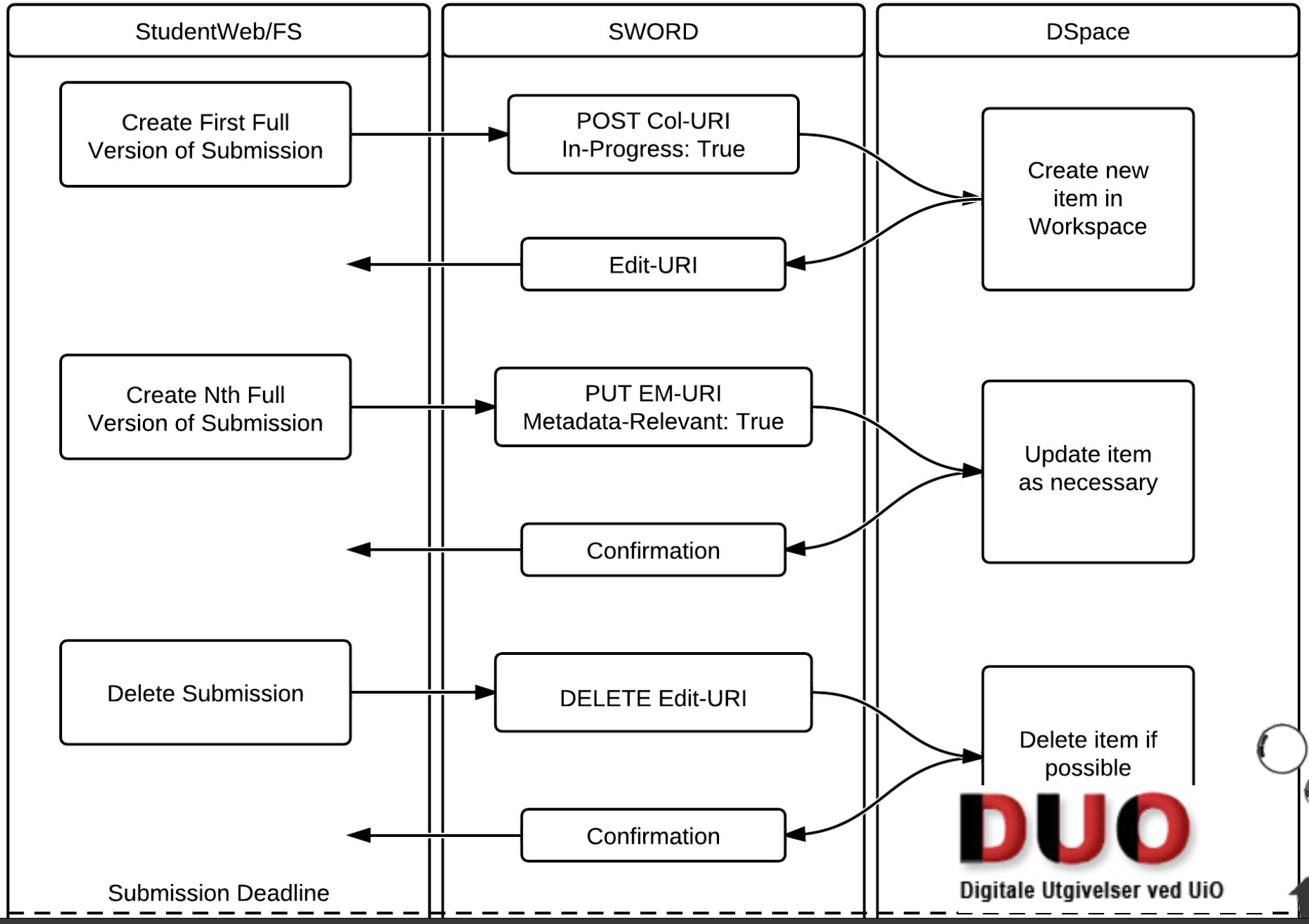
DUO
Digitale Utgivelser ved UiO

Usage of SWORD 2.0



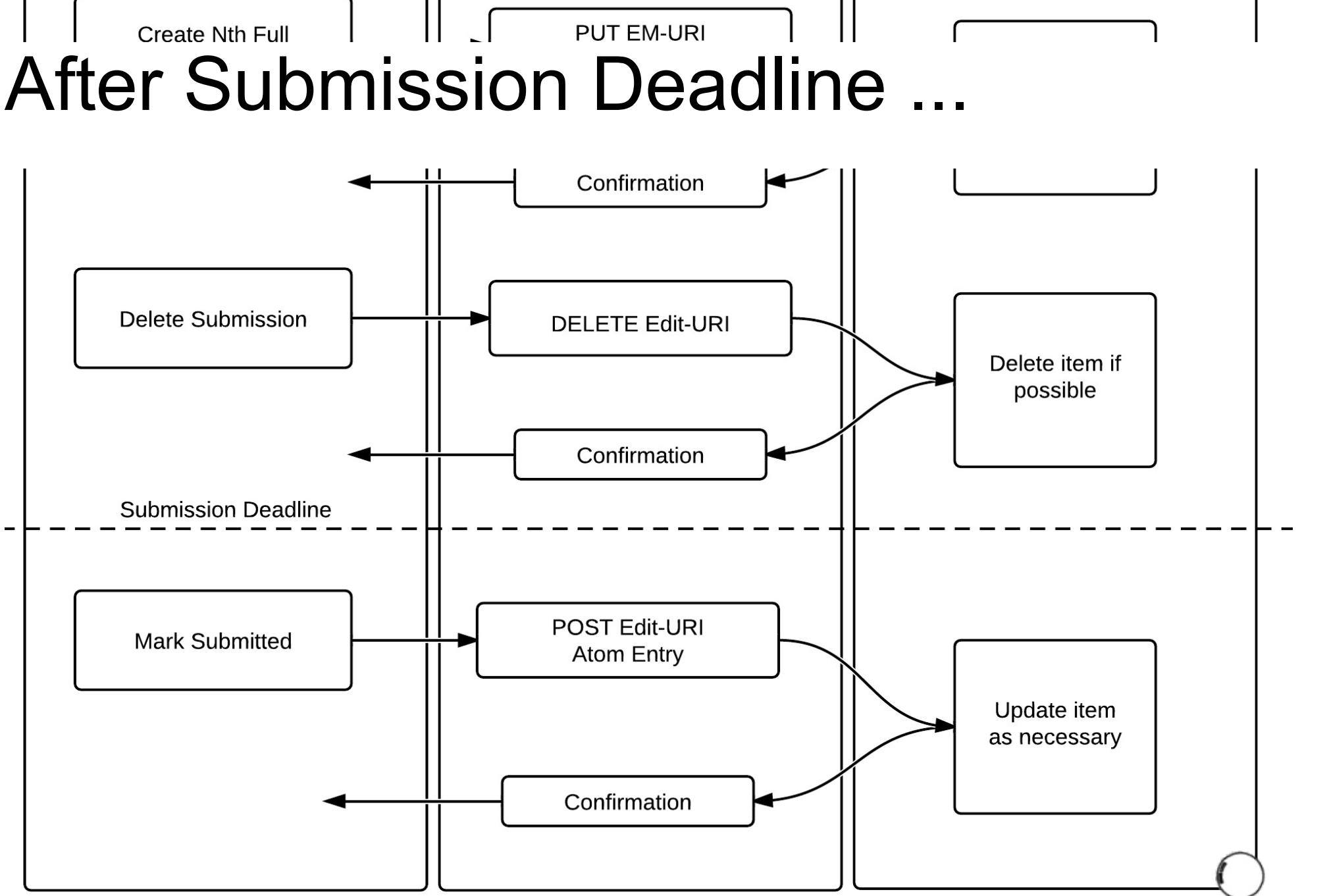
Digitale Utgivelser ved UiO

Before Submission Deadline ...



DUO
Digitale Utgivelser ved UiO

After Submission Deadline ...



Digitale Utgivelser ved UiO

BagIt Packaging

- Primary/Final Content Files (PDFs + Attachments)
- Secondary and Source Files (Any Formats)
- Licence (Plain Text)
- Metadata (XML)

```
<base directory>/
|   bagit.txt
|   manifest-md5.txt
|   tagmanifest-md5.txt
\--- data/
      \--- final
            |   [final version files]
      \--- supporting
            |   [supporting files]
      \--- licence/
            |   licence.txt
      \--- metadata/
            |   metadata.xml
\--- tagfiles/
      |   supporting.access.txt
      |   formats.txt
      |   final.sequence.txt
      |   supporting.sequence.txt
```

DUO
Digitale Utgivelser ved UiO

Available Metadata

- Student:
  - Name
  - Student Number
  - Institutional Username
  - Fodselsnummer
  - Postal Address
  - Telephone Number
  - Email Address
- Subject
  - Code
  - Name
- Unit Code
- Title
- Abstract
- Language
- Embargo Date
- Grade

Encoded as a custom XML format, due to lack of clear schema for the data

Where are we up to?

- **Requirements** have been drafted for both **StudentWeb/FS** side and **DSpace** side developments
- **Development** of the **code libraries** to deploy a **pilot** service
- **Integrated** the code libraries into **DSpace** and **StudentWeb/FS**
- Just finished running the **pilot**

DUO
Digitale Utgivelser ved UiO

After the Pilot?

Once the pilot completes, samarbeidstiltaket FS has indicated that there will be a national meeting organised by StudentWeb/FS.

It will be possible for all Norwegian institutions to hook their StudentWeb instance up to a SWORD 2.0 compliant Digital Repository.

Thanks for listening!

Richard Jones, Cottage Labs, richard@cottagelabs.com
Kristian Salcedo, University of Oslo, k.r.salcedo@ub.uio.no

DUO
Digitale Utgivelser ved UiO