Jozef Mišutka, Amir Kamran, Ondřej Košarko, Michal Sedlák, Pavel Straňák
LINDAT/CLARIN team, Charles University in Prague, Czech Republic
clarin-list@ufal.mff.cuni.cz

Framework integrated with DSpace 5.2

Sharing and reusing data legally

Guided License Selection

The necessity to share and preserve data and software is becoming more and more important. A research paper without underlying data and software is more of an advertisement for possible scholarship than scholarship itself. Without the data and the software, research cannot be reproduced and tested by the scientific community. Making data and software simply reusable and legally unequivocal requires choosing a license for data and software which is not a trivial task.

Sharing and reusing data

- different licenses for data sets and software (they cannot be mixed)
- no "fair use" in EU, research exceptions in some countries
- authors often want user names and signatures, but rarely need them
- recommend a license as free as possible (and as limited as needed)
- avoid proliferation for data - CC 4.0 BY *, CC0 , CC Public Domain

Licensing Framework

- one license per submission (can be meta license)
- easy management of licenses
- explicit download tracking– fairly detailed list of users
- part of DSpace 5.x
- https://github.com/ufal/lindat-dspace
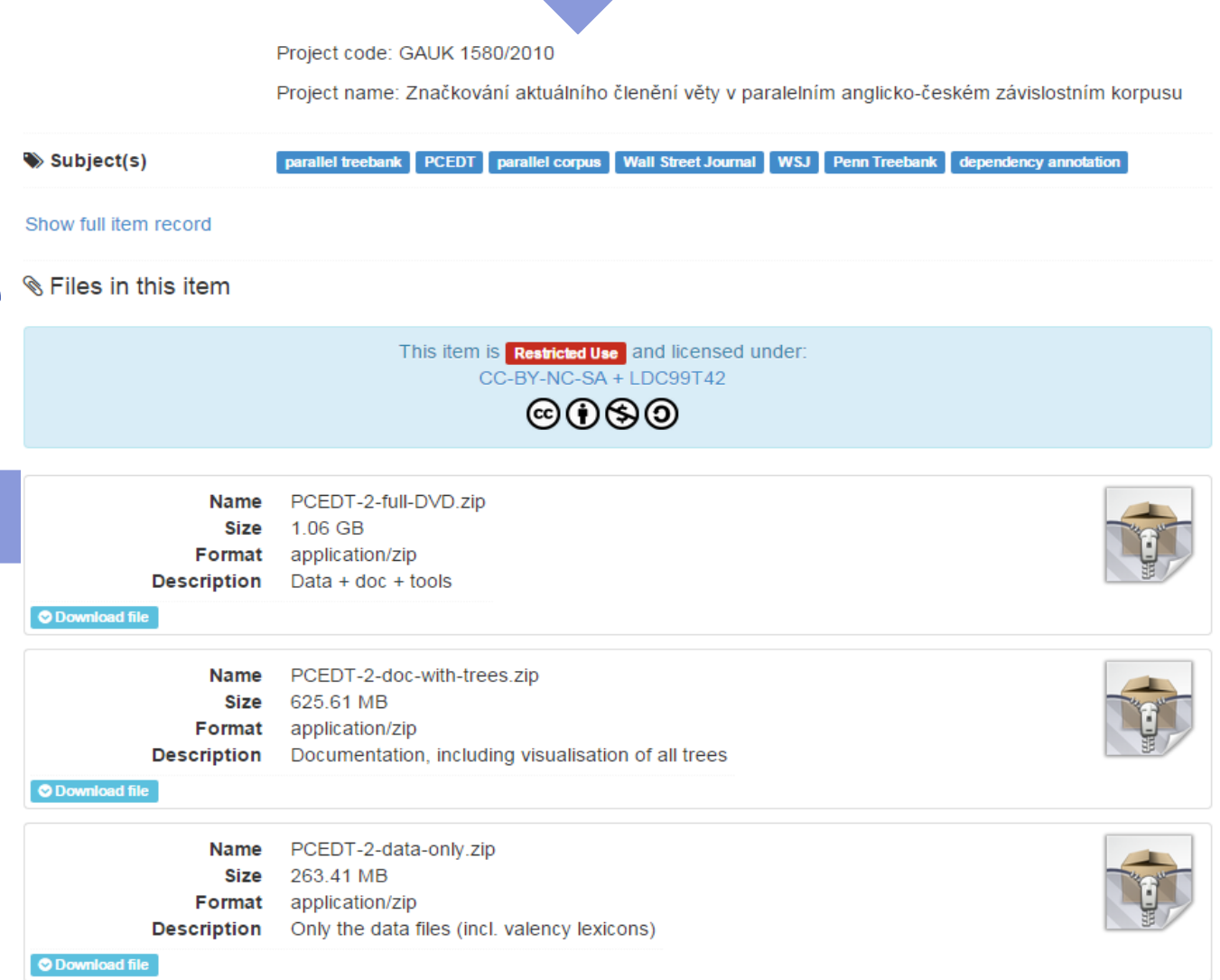
License Selector

- for both data and software
- standalone and extensible
- designed by an international lawyer and data expert
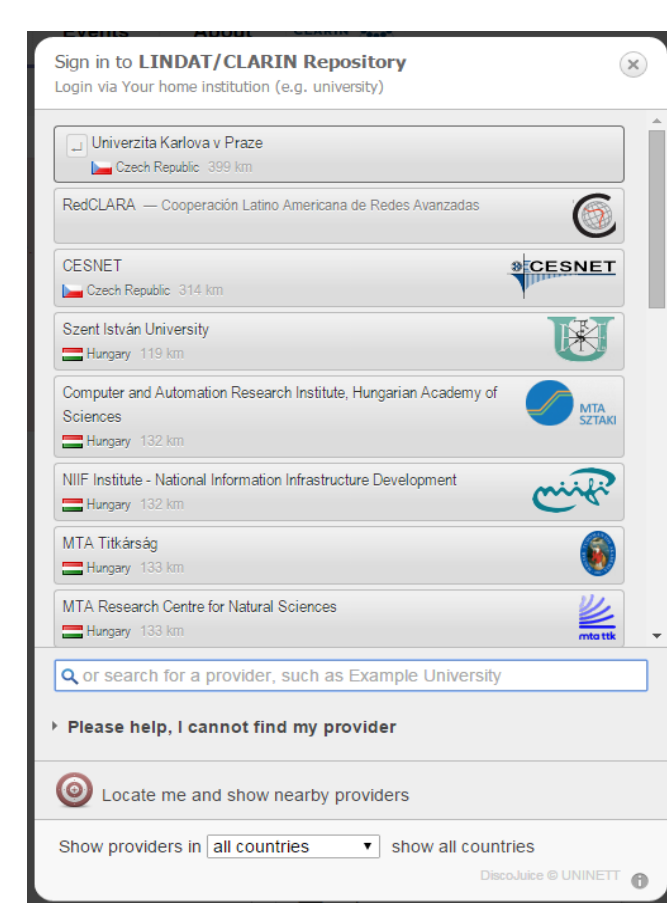- set of questions where answers specify the requirements
- https://github.com/ufal/lindat-license-selector
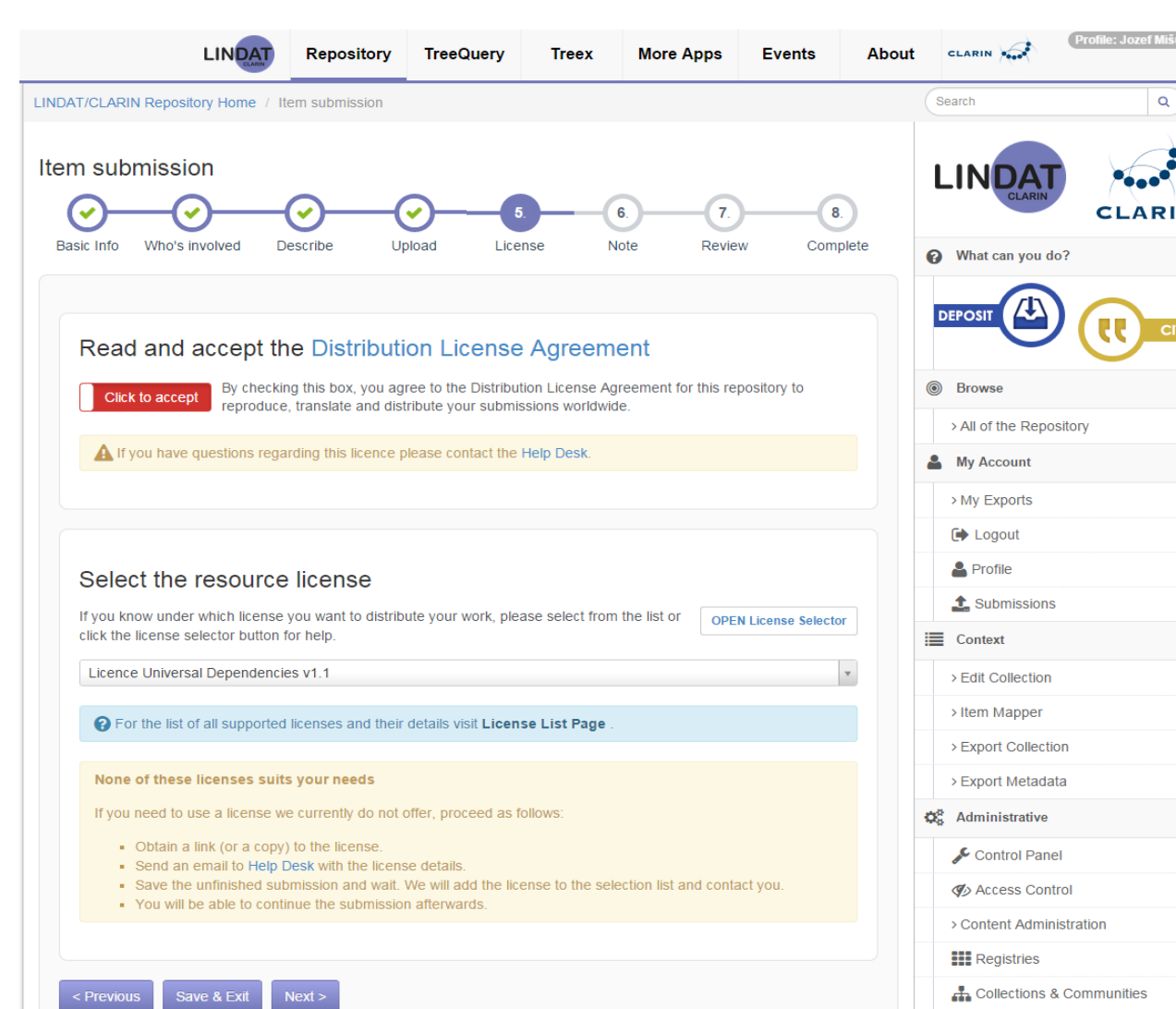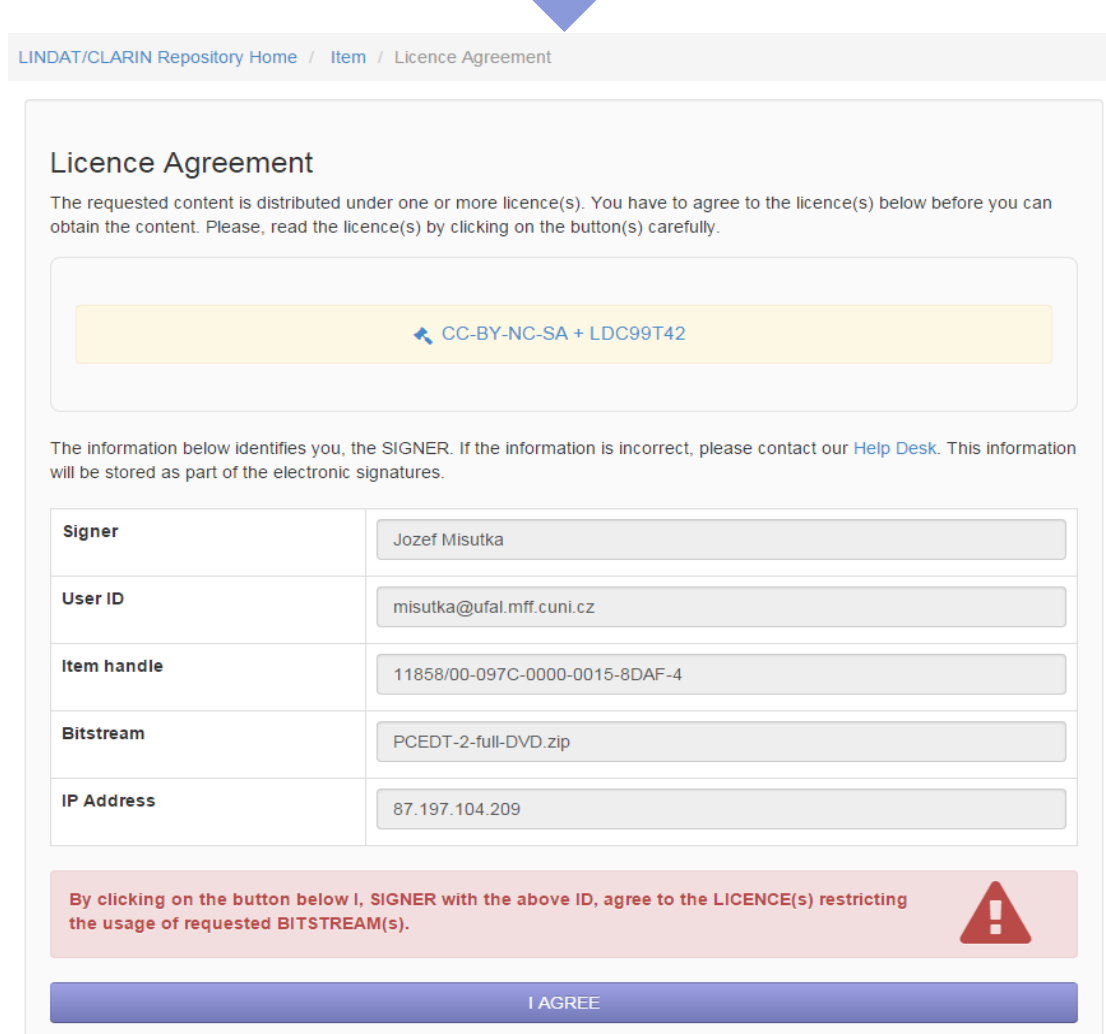
submitter selects license

define new license if not available
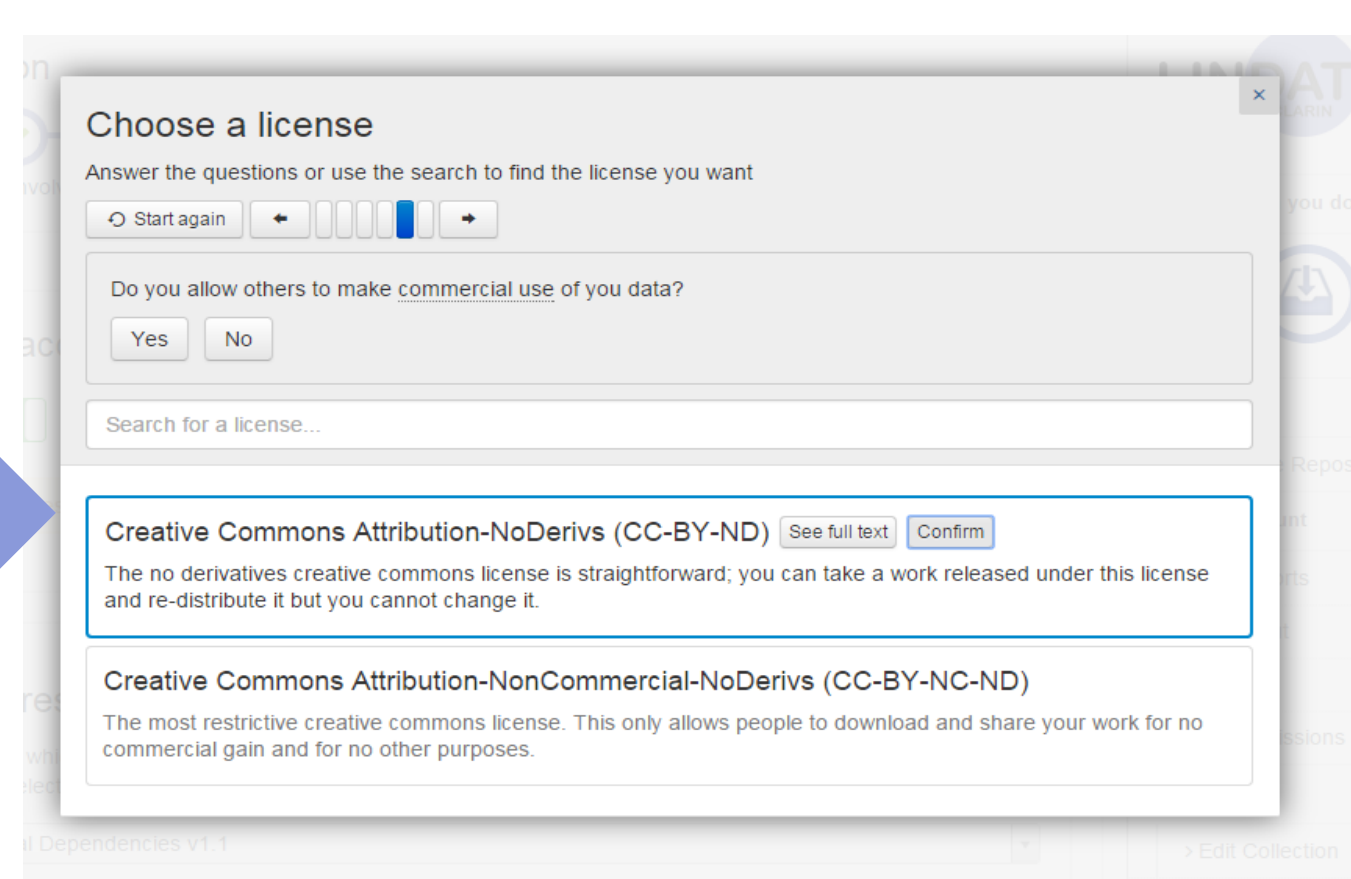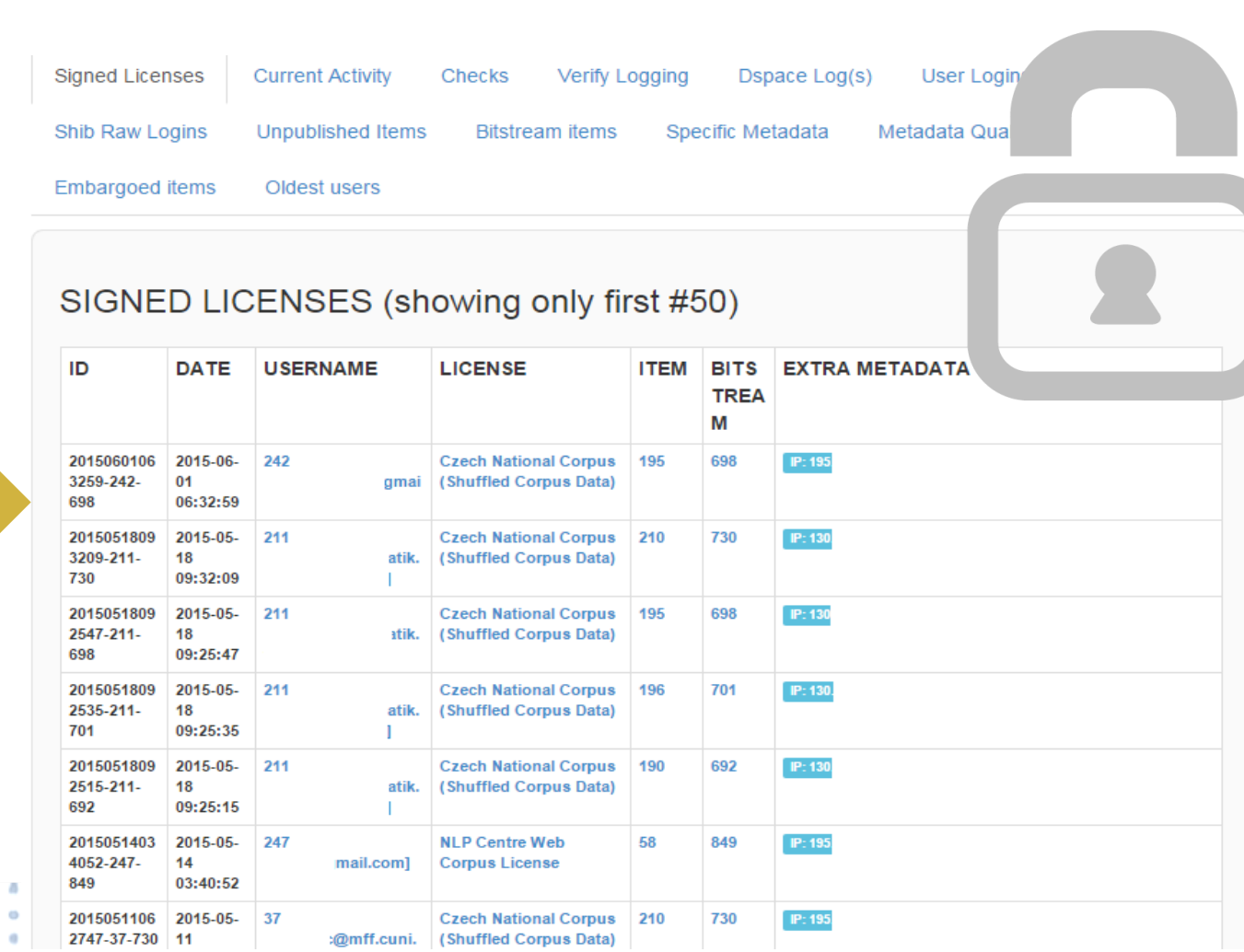
submission has license

instantly available

download can require signature and additional data

user must be traceable/identifiable

signature and other information stored
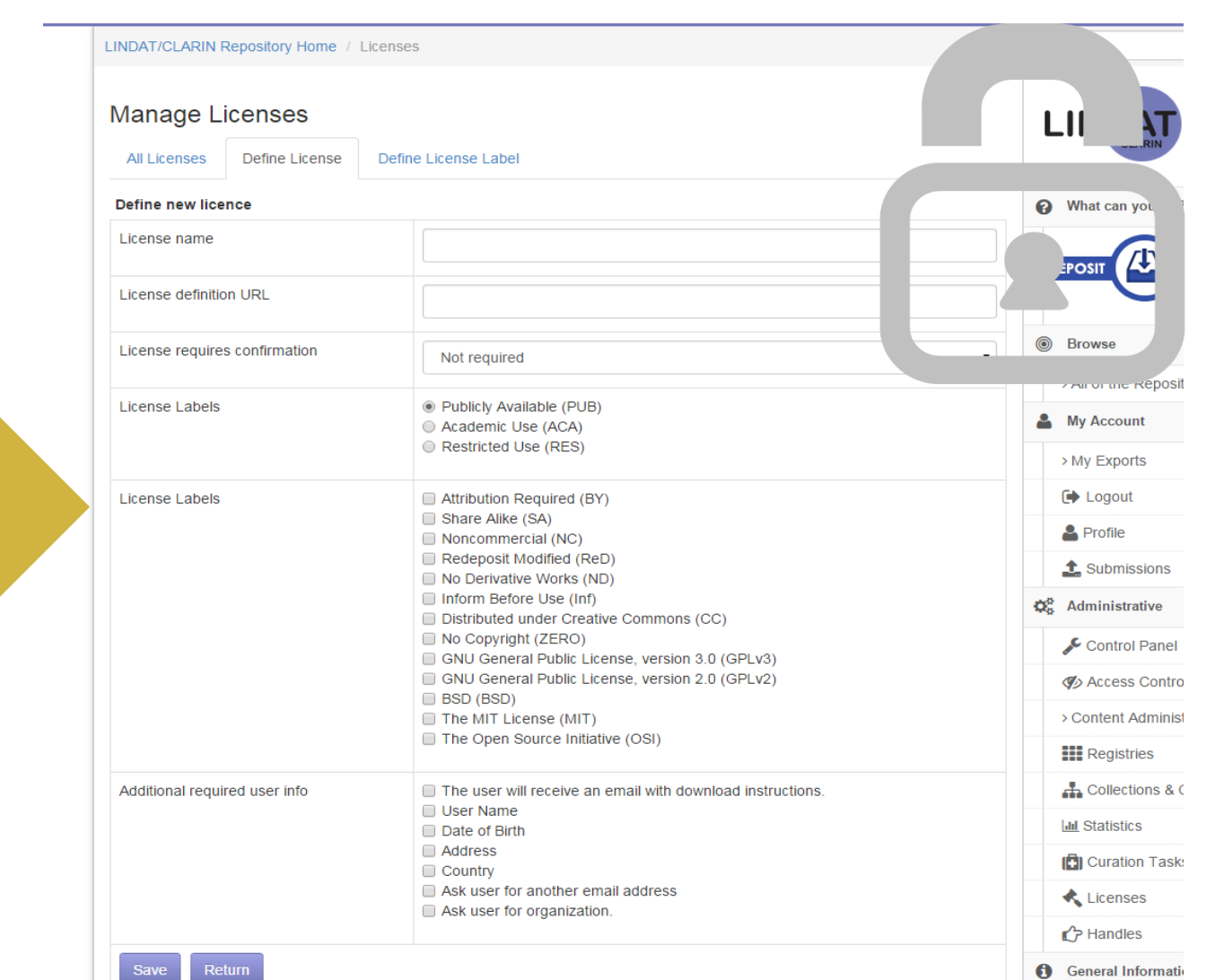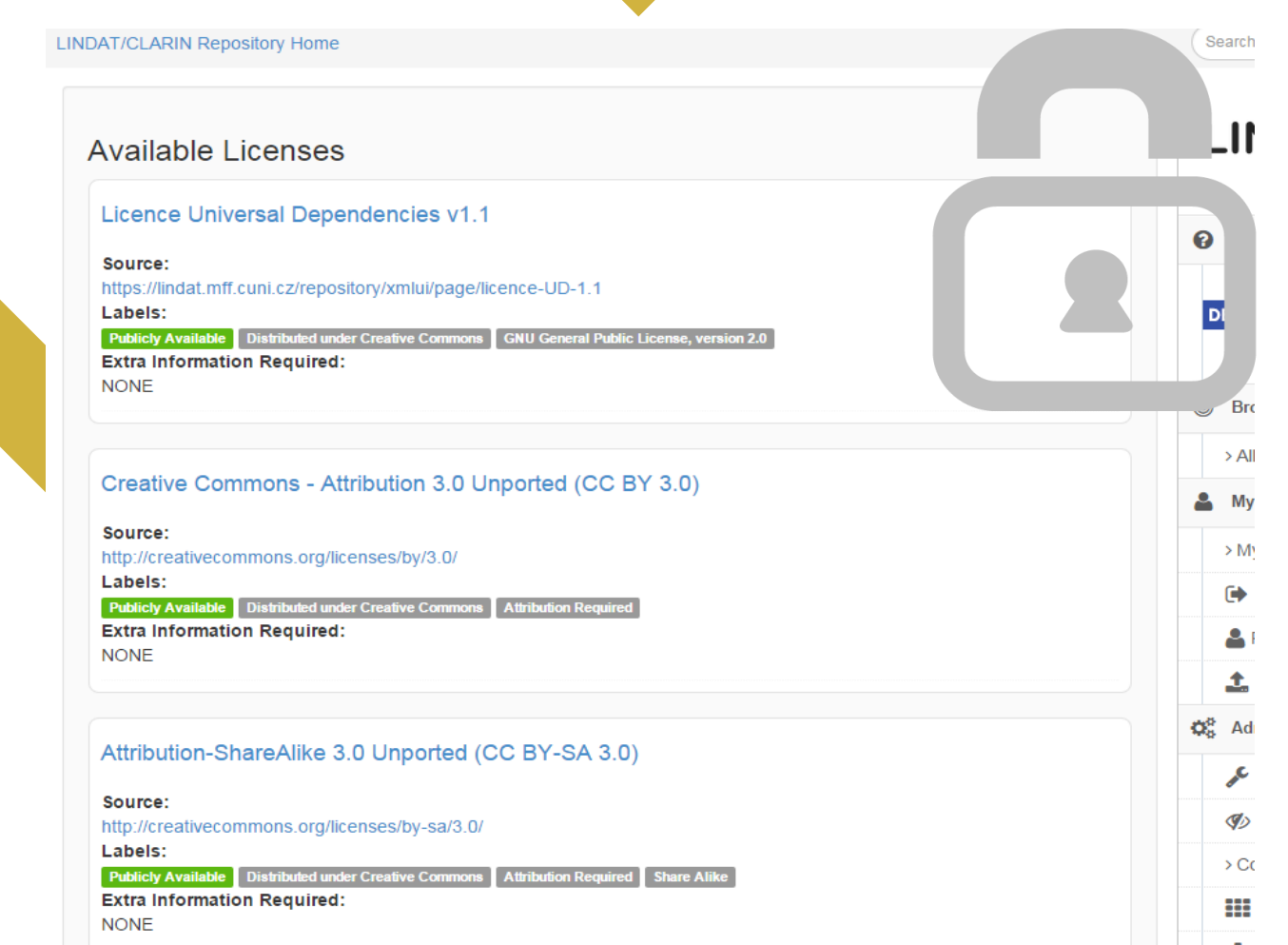
License Selector Questions - Data
- Is your data within the scope of copyright and related rights?
- Do you own copyright and similar rights in your dataset and all its constitutive parts?
  - Do you allow others to make derivative works?
  - Are all the elements of your dataset licensed under a public license or in the Public Domain?
- Do you require others to share derivative works based on your data under a compatible license?
- Do you allow others to make commercial use of you data?
- Do you want others to attribute your data to you?

...

License Selector Questions - Software
- Is your code based on existing software or is it your original work?
  - Do you want others who modify your code to be forced to also release it under open source license?
- Is your code used directly as an executable or are you licensing a library (your code will be linked)?

...

http://lindat.cz

clarin-list@ufal.mff.cuni.cz

follow us on twitter @lindatclarin

https://github.com/ufal

LINDAT/CLARIN

LINDAT CLARIN

ÚFAL

ÚSTAV FORMÁLNÍ A APLIKOVANÉ LINGVISTIKY