Evaluation of the semantic research data management system CaosDB in glaciology

Alexander Schlemmer[1,*], Henrik tom Wörden[1,*], Johannes Freitag[2], Timm Fitschen[1,*], Johanna Kerch[2], York Schlomann[2], Daniel Hornung[1,*], Ilka Weikusat[2], Ulrich Parlitz[1], Frank Wilhelms[2] and Stefan Luther[1]

MPI for Dynamics and Self-Organization, Göttingen (1)
AWI for Polar and Marine Research, Bremerhaven (2)
IndiScale GmbH i.G. (*)

2019-06-05

Caosdb

CaosDB

Important Requirements for Research Data Management

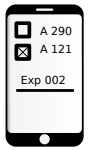
- Simple but expressive search functionality
- Ability to store every data format, at any file size
- Possibility to store, link and retrieve raw data, processed data, analysis results and documentation
- Support all kinds of data analysis software, from simple scripts to high-level software
- Minimally invasive workflow
- Scientific environments change often: Need for flexible data model

Research Data Management during Data Analysis

Data Aquisition

Data Publication

Electronic Lab Notebooks

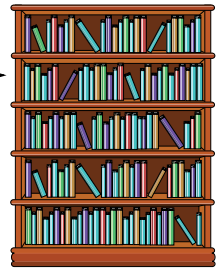


z.B. RSpace, IDBS E-WorkBook, Biovia

Data Repositories

ePIC / PID

Metadata



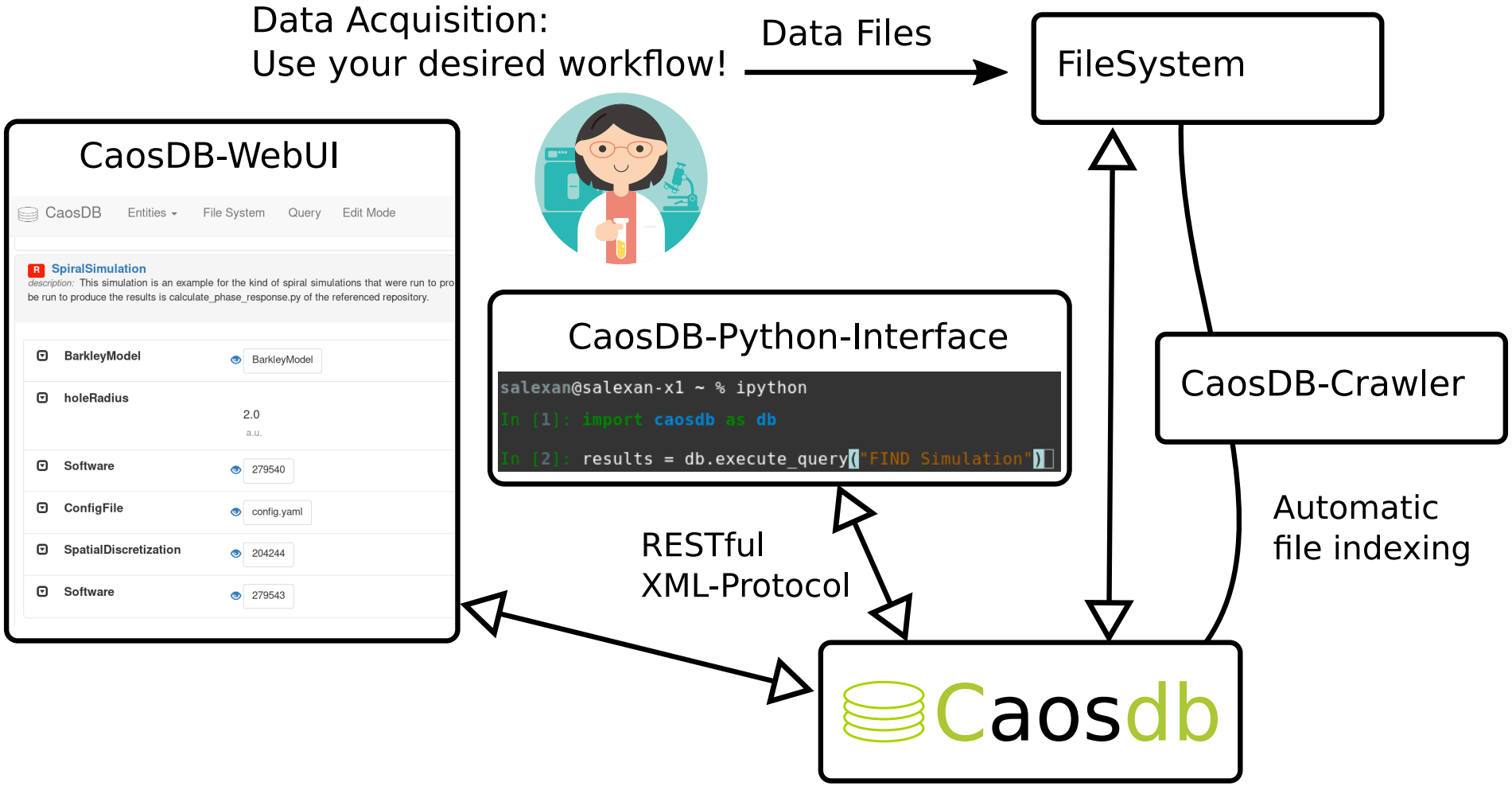
Data Analysis

Caosdb

Scientist picture: 201705 Scientist bench F.svg from commons.wikimedia.org/wiki/Category:Life_science_images_from_DBCLS, CC-BY 4.0

Bookshelf: https://openclipart.org/detail/289378/bookshelf-with-blue-books

CaosDB Overview



OpenSource! https://gitlab.gwdg.de/bmp-caosdb

State and Future of the Project

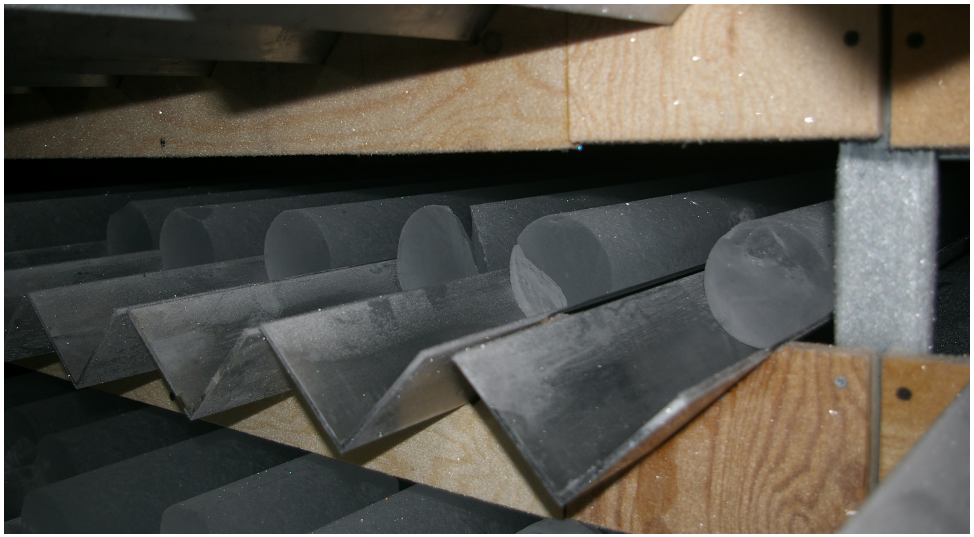
- CaosDB is a scientific project at the Research Group Biomedical Physics (MPI Göttingen)
- It is developed since ≈8 years and running stable since ≈2016
- CaosDB has been released as OpenSource software in 2018
- Service Company: IndiScale GmbH i.G.
- CaosDB is currently tested in other workgroups in and outside of Göttingen
  → Very positive feedback

CaosDB in Glaciology at the AWI in Bremerhaven

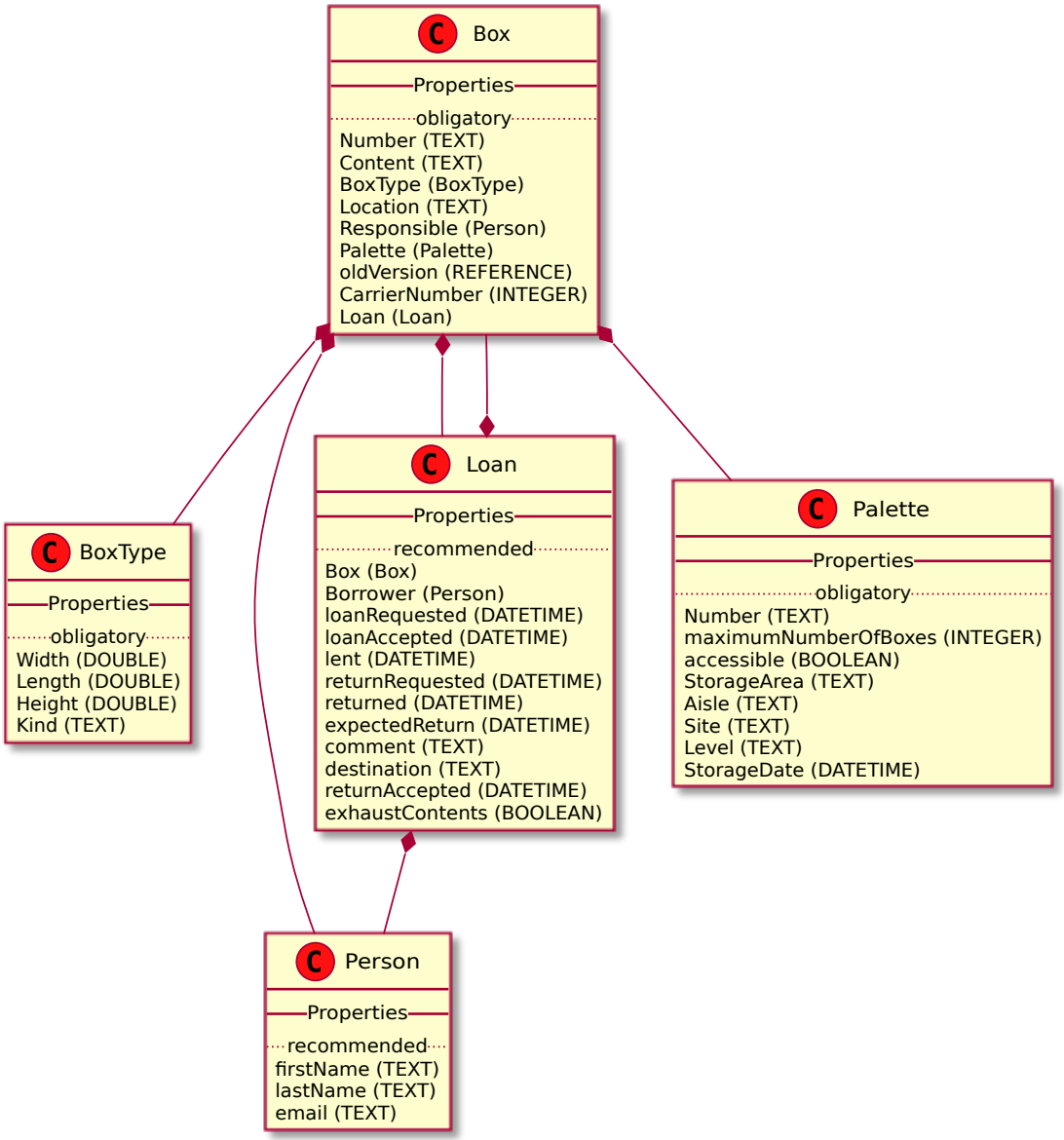
Aims of the Collaborative Project

- System for efficient management of ice core boxes
  $\rightarrow$ Future extension to sample management
- Testing of the flexible data model of CaosDB in the context of polar research
- Evaluation of CaosDB in a different scientific environment
- Feedback from new users

Ice Cores



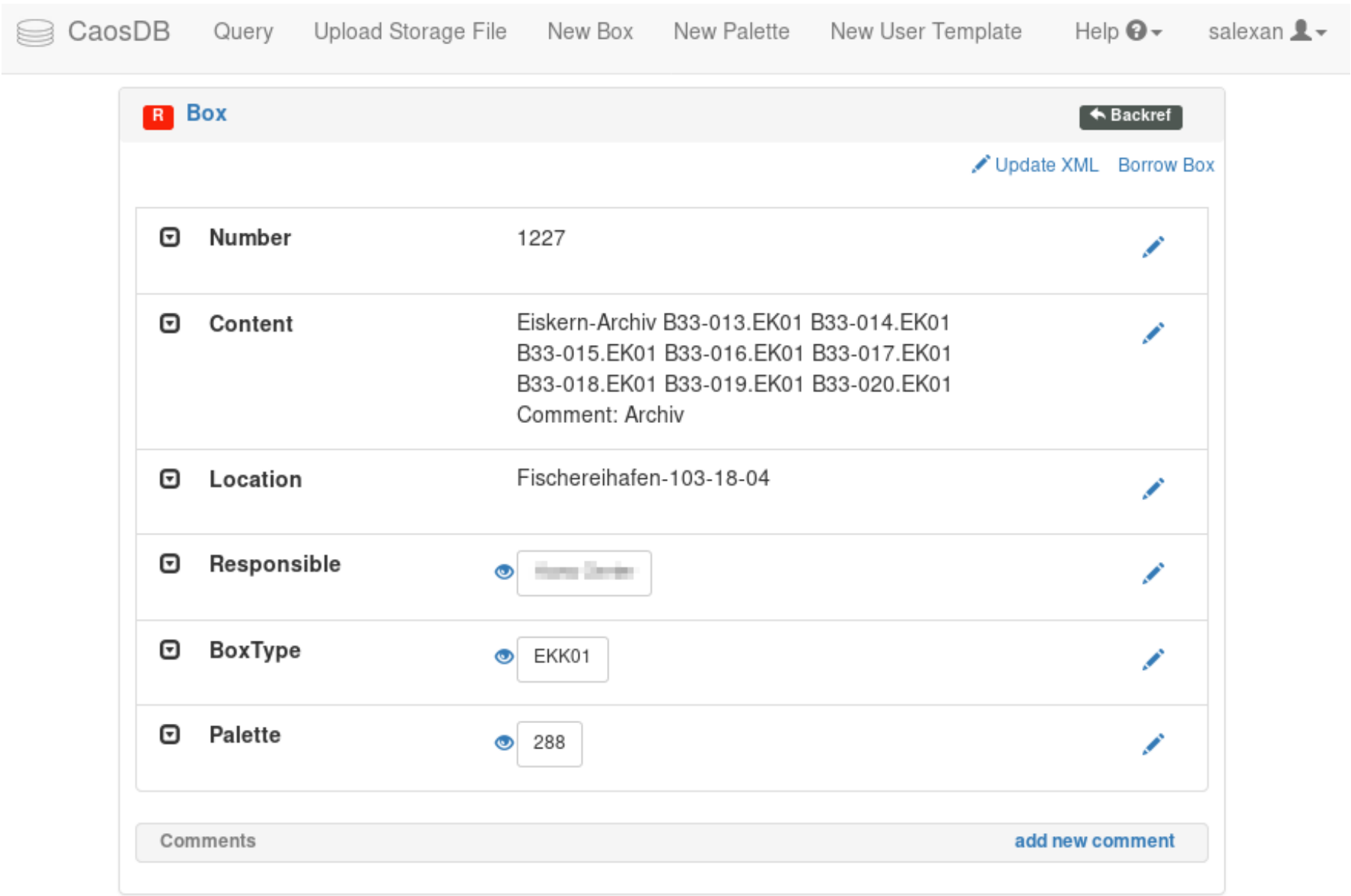
UML diagram of the data model



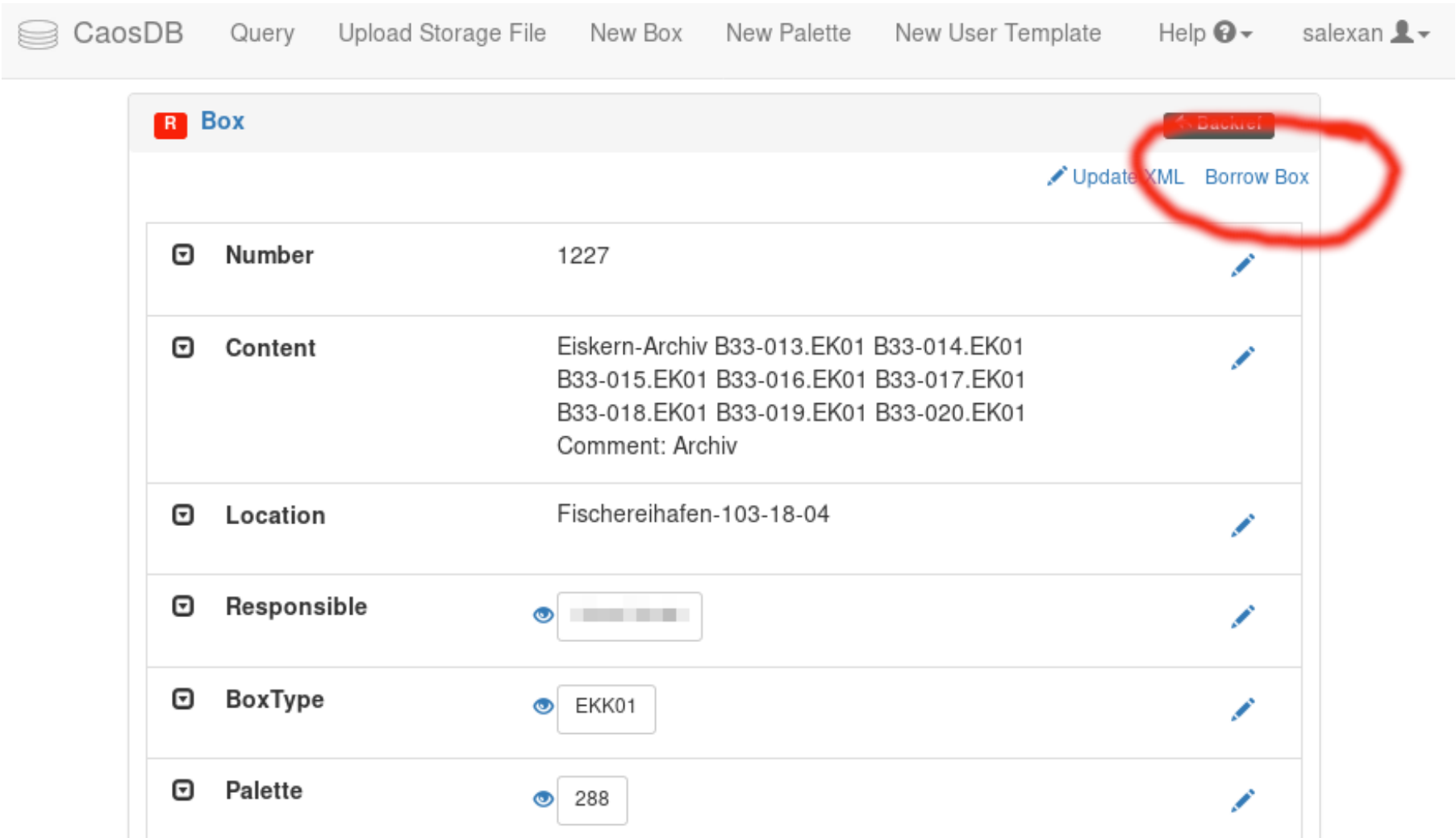
Querying the Data Model

- `FIND Record Box with Number = "0258"`
- `SELECT Number, Content from Box with Content like "ice"`
- `FIND Record Box with palette with number = "12"`
- ```
  FIND Fabric
  which references Section
  which references Bag
  which references PPStrip
  which references Core with name = "EGRIP"
  ```
  ```

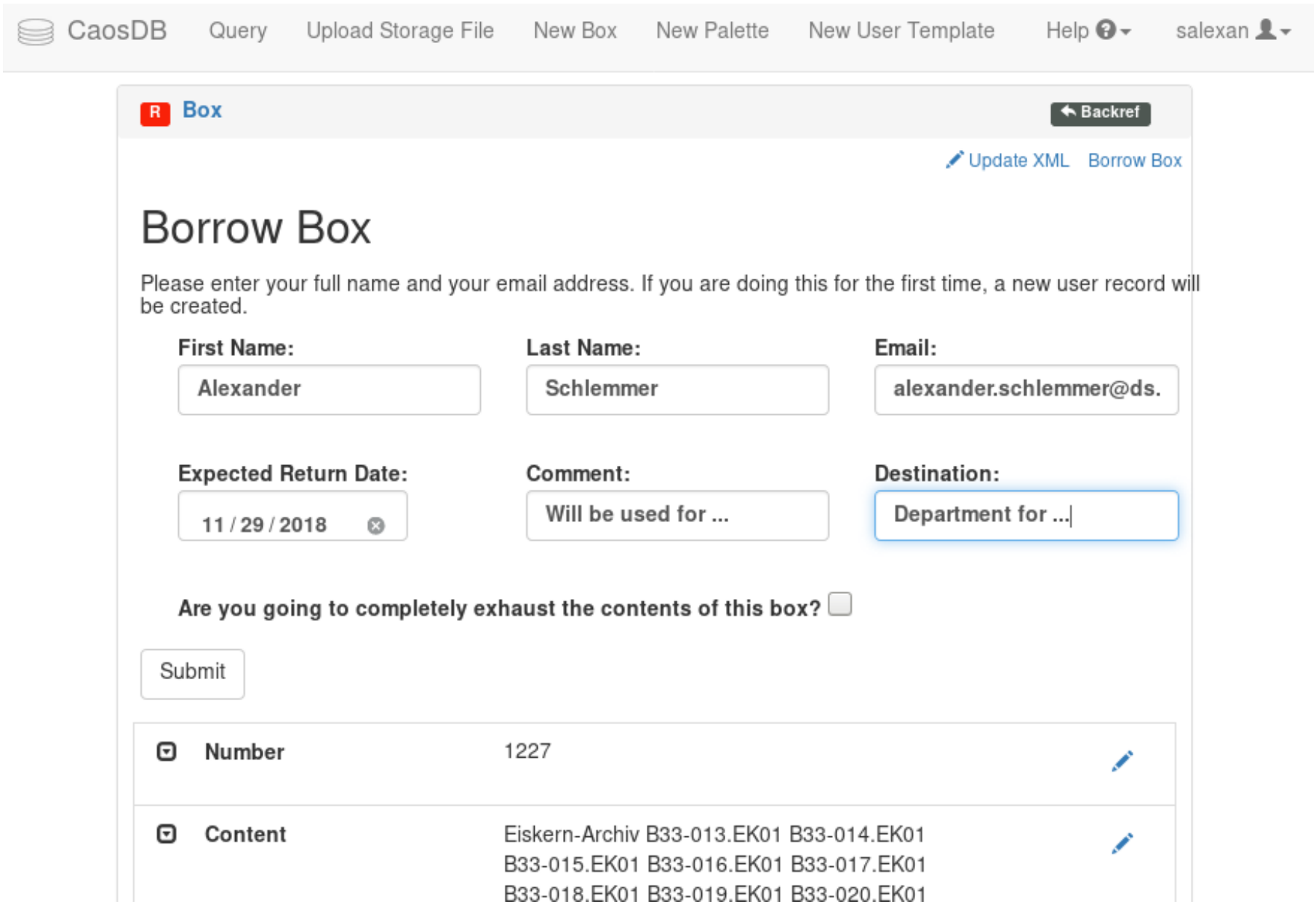
WebUI


CaosDB
Query
Upload Storage File
New Box
New Palette
New User Template
Help
salexan
R
Box
Backref
Update XML
Borrow Box
Number
1227
Content
Eiskern-Archiv B33-013.EK01 B33-014.EK01
B33-015.EK01 B33-016.EK01 B33-017.EK01
B33-018.EK01 B33-019.EK01 B33-020.EK01
Comment: Archiv
Location
Fischereihafen-103-18-04
Responsible
BoxType
EKK01
Palette
288
Comments
add new comment

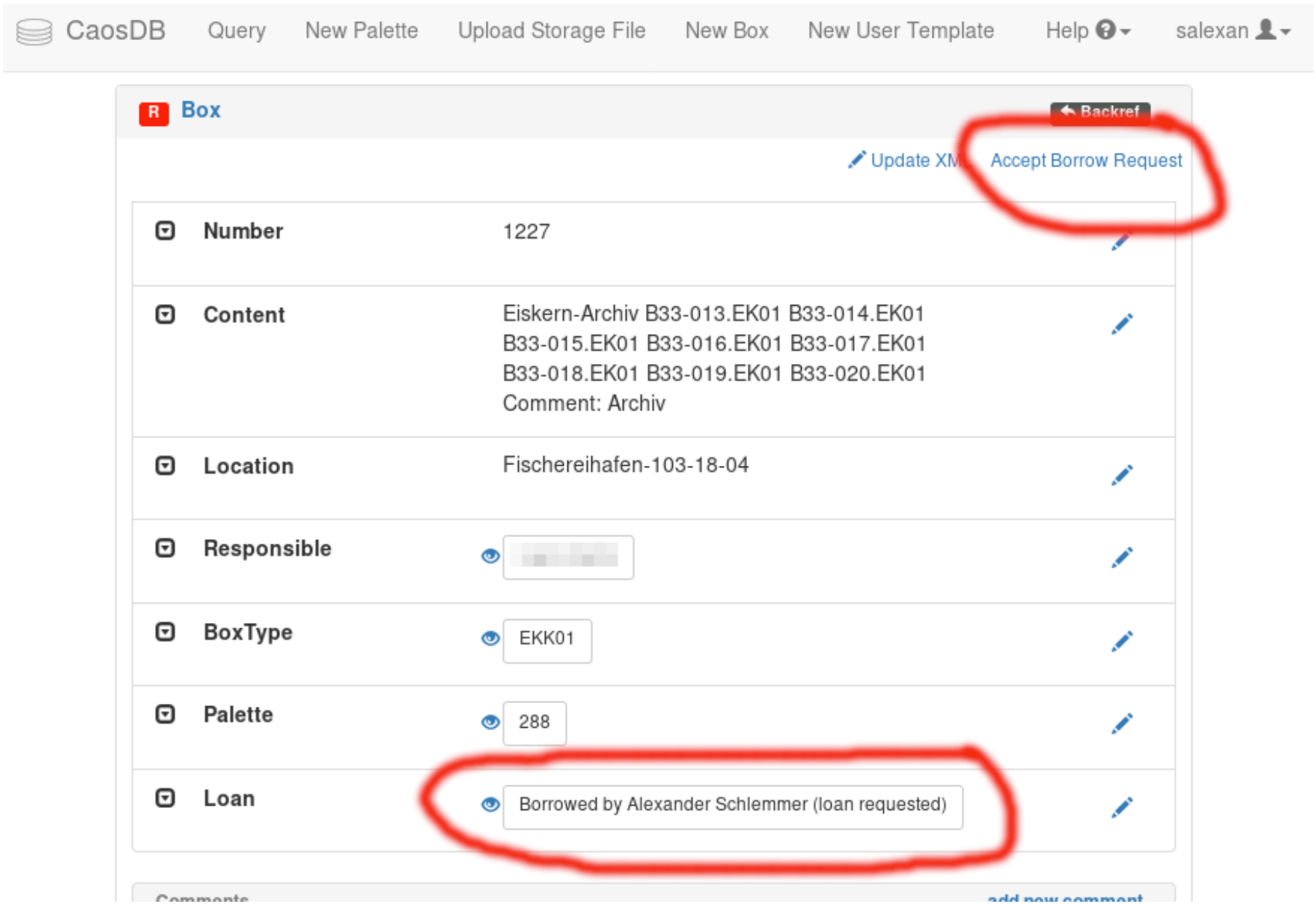
Implementation of the Loan Workflow



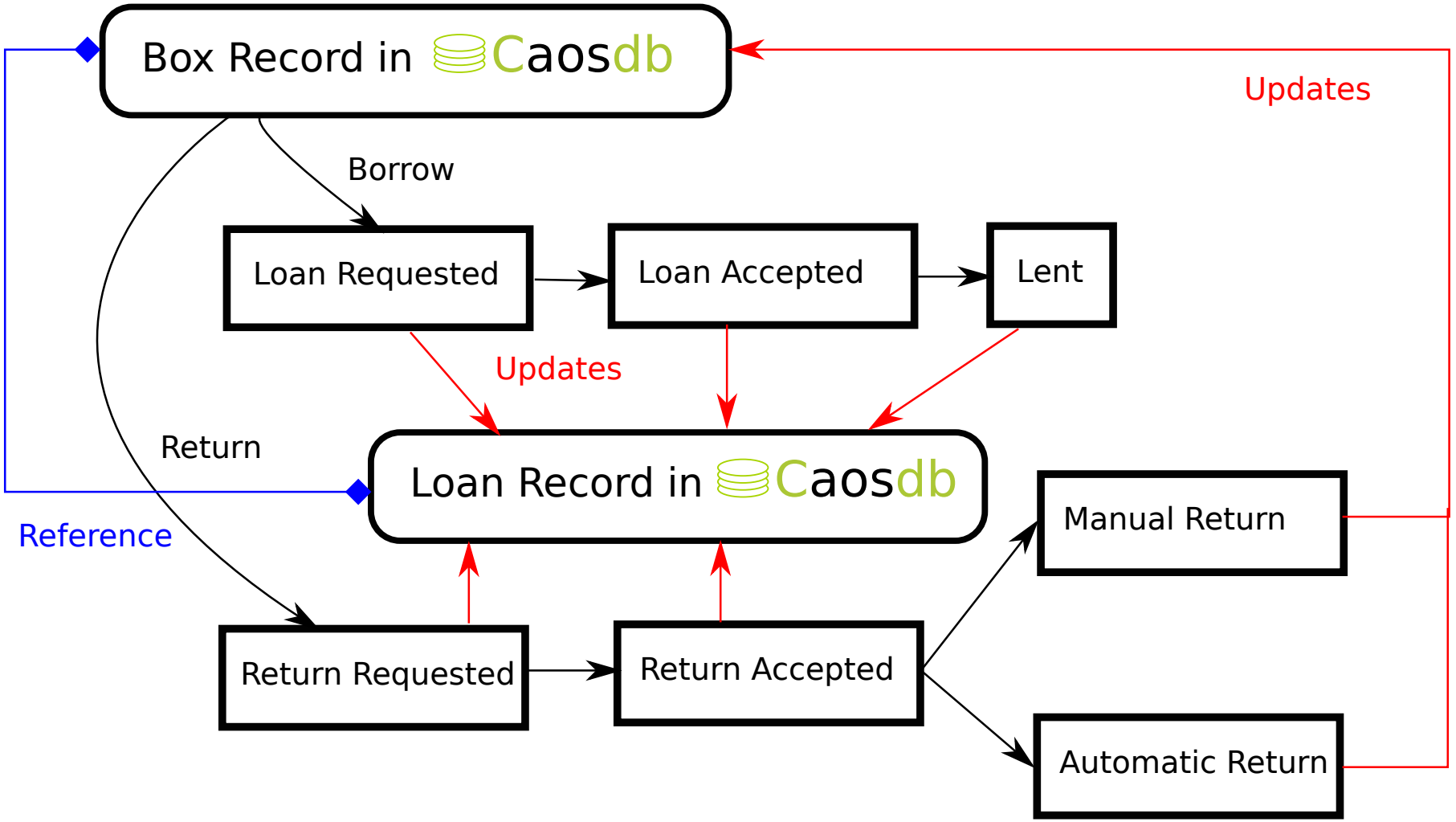
Implementation of the Loan Workflow



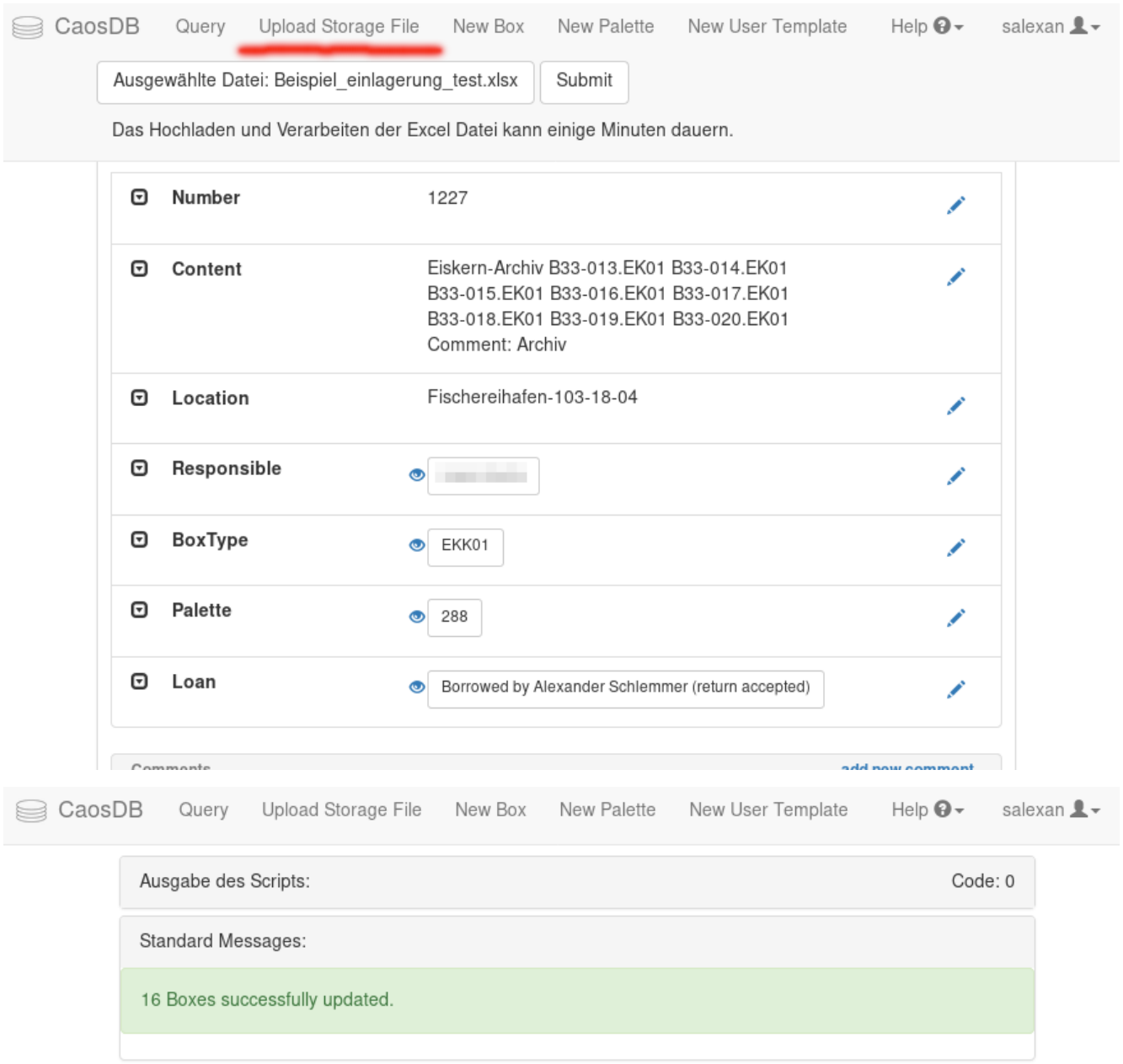
Implementation of the Loan Workflow



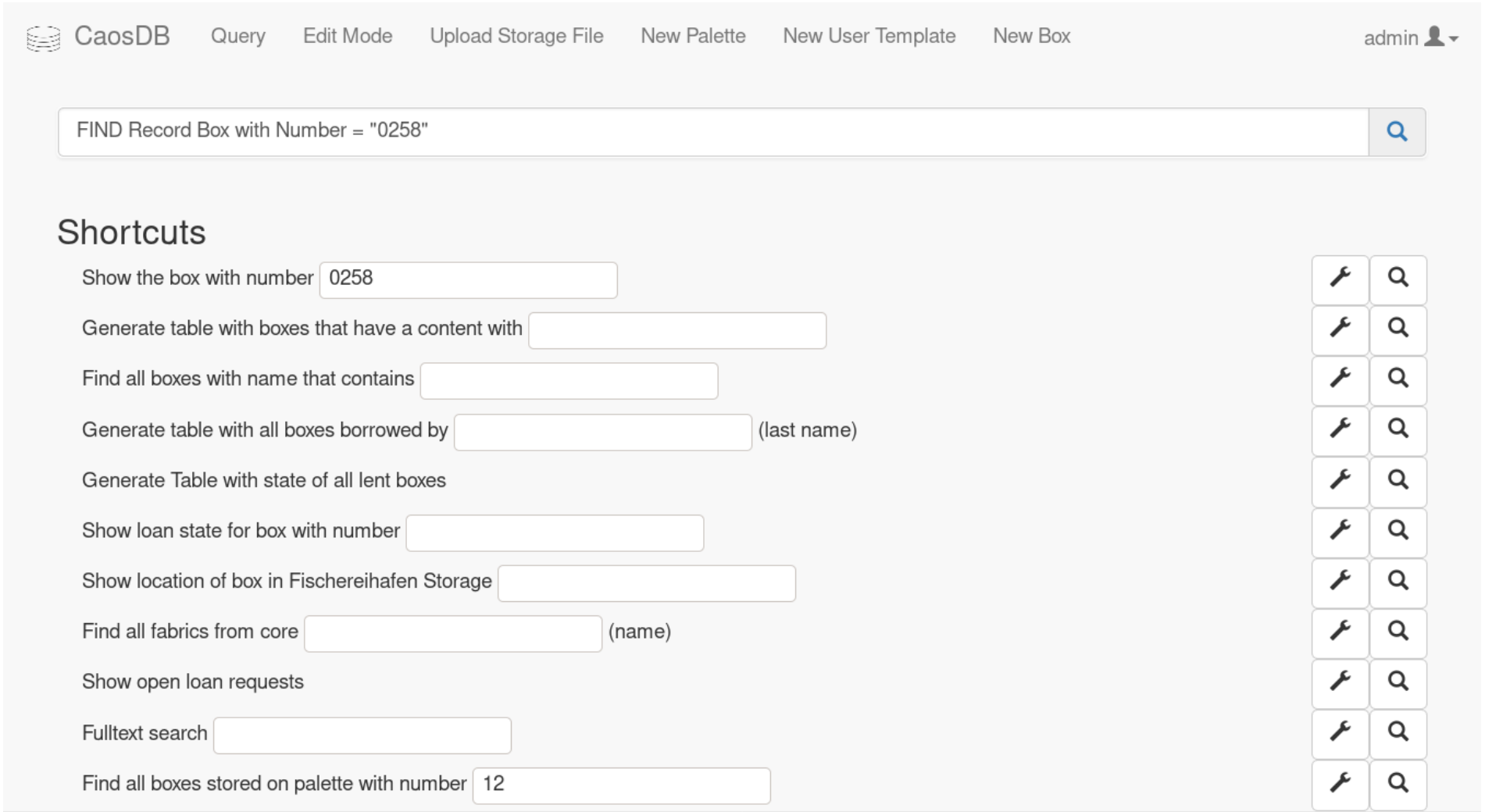
Full Ice Core Box Loan Workflow



Automatic Return



Searching


CaosDB
Query
Edit Mode
Upload Storage File
New Palette
New User Template
New Box
admin
FIND Record Box with Number = "0258"
Shortcuts
Show the box with number 0258
Generate table with boxes that have a content with
Find all boxes with name that contains
Generate table with all boxes borrowed by (last name)
Generate Table with state of all lent boxes
Show loan state for box with number
Show location of box in Fischereihafen Storage
Find all fabrics from core (name)
Show open loan requests
Fulltext search
Find all boxes stored on palette with number 12

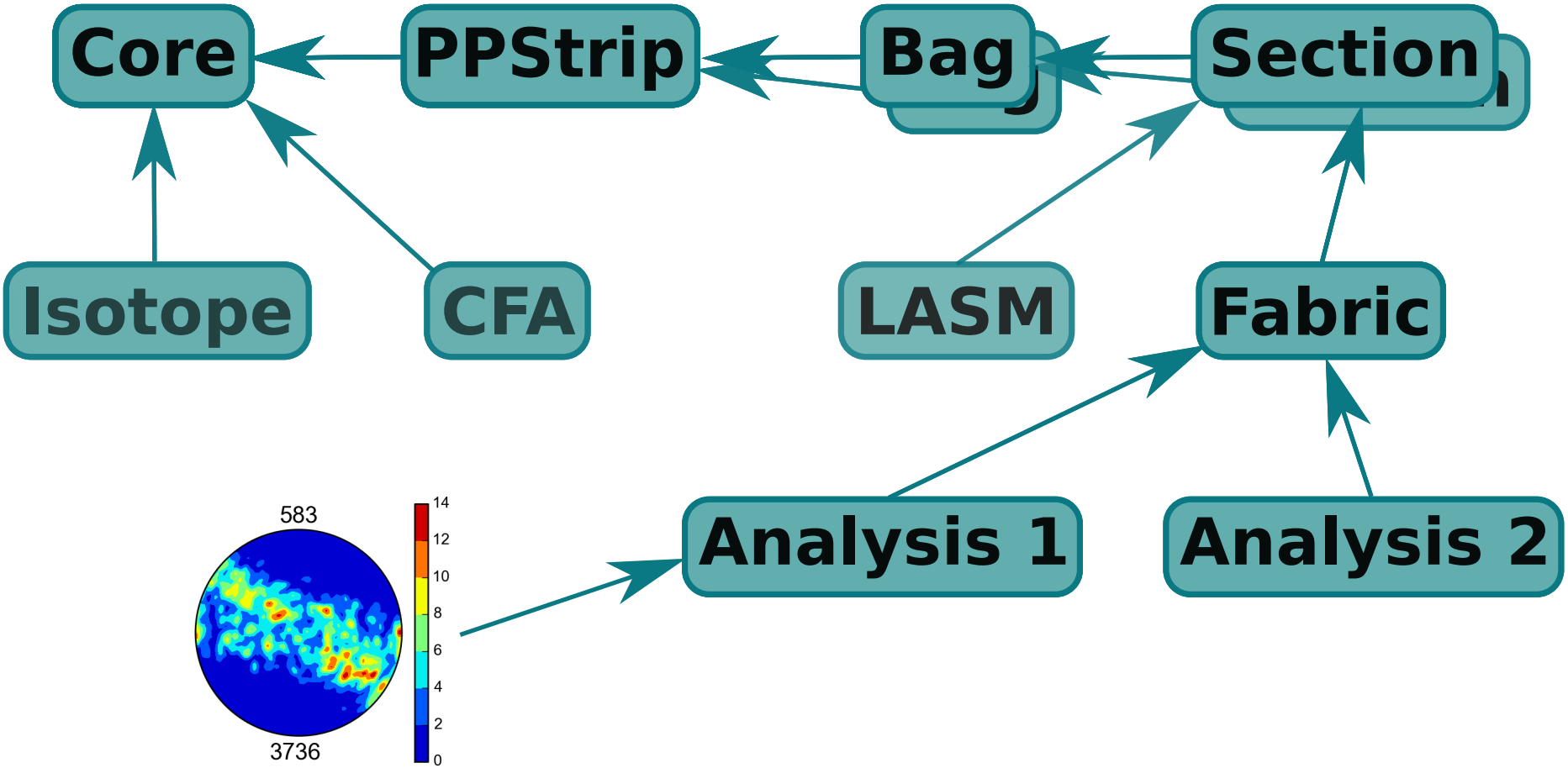
Listings

CaosDB Query Upload Storage File New Box New Palette New User Template Help salexan

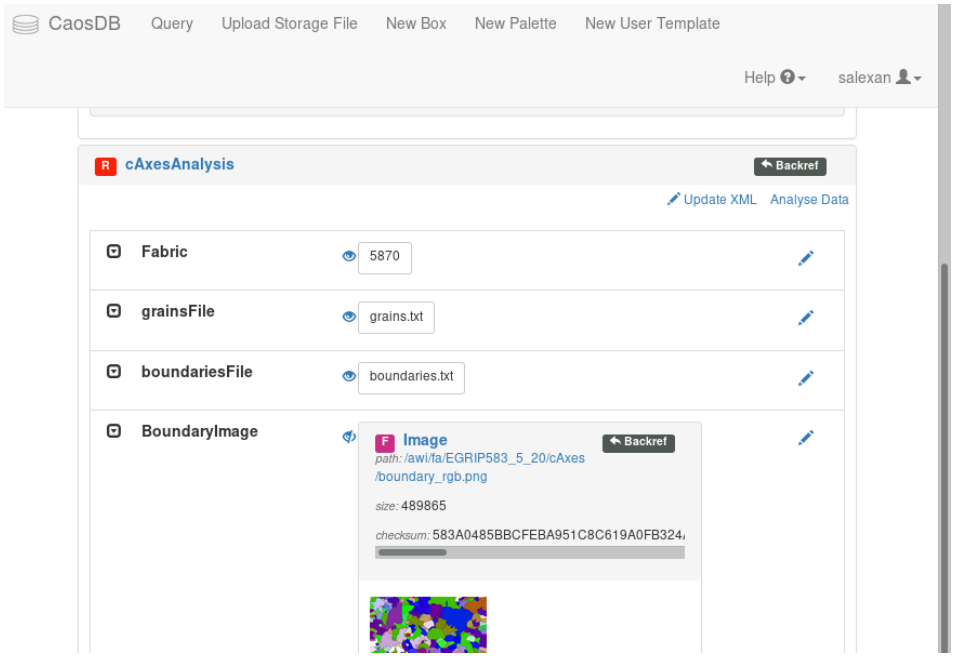
Table of selected fields Download TSV

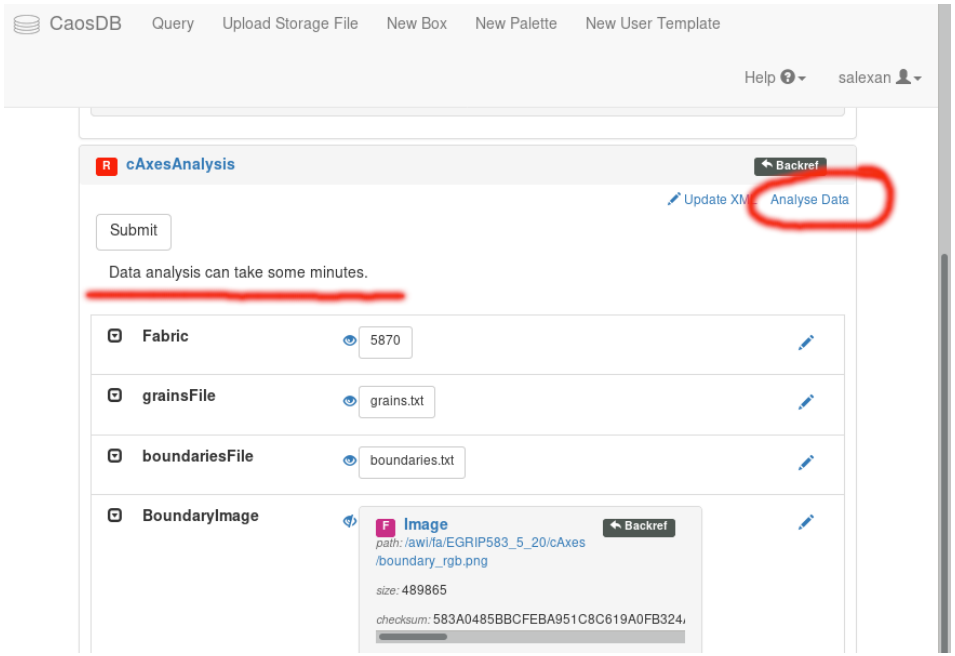
Borrower	expectedReturn	lent	returned
Alexander Schlemmer	2018-11-28	2018-11-19T22:00:03+0100	
Alexander Schlemmer	2018-11-21	2018-11-19T21:52:06+0100	
Alexander Schlemmer	2018-11-27	2018-11-19T21:14:58+0100	
Alexander Schlemmer	2018-11-28		
Alexander Schlemmer	2018-11-15	2018-11-19T22:03:30+0100	
Henrik tom Wörden	2018-11-21	2018-11-20T09:27:04+0100	
Alexander Schlemmer	2018-11-22	2018-11-20T10:27:00+0100	
Alexander Schlemmer	2018-11-22	2018-11-20T10:42:07+0100	2018-11-20T11:43:51.824019+
Alexander Schlemmer	2018-11-29		2018-11-20T12:24:59.828999+

From Box Management to Sample Management



Data Analysis in CaosDB





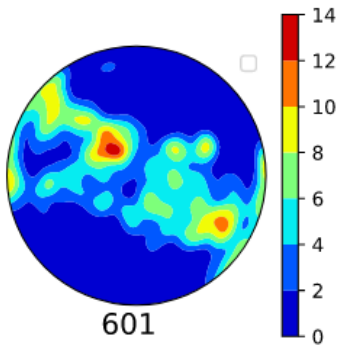
Data Analysis in CaosDB

Ausgabe des Scripts:

Errors:

No handles with labels found to put in legend.

Standard Messages:



5873

R stereoAnalysis Backref

Update XML

stereoFile stereo.txt

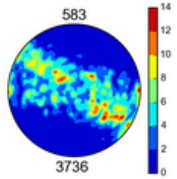
AnalysisResult

F Image AnalysisResult Backref

path: /awi/fa/583-stereo_py_con0.png

size: 40217

checksum: 01C8F63CDE2328EAAE649BBBB9931E0



Comments add new comment

More Information

Caosdb

- OpenSource Project: `https://gitlab.gwdg.de/bmp-caosdb`
- Paper on ArXiv: `https://arxiv.org/abs/1801.07653`, under review in Data[1]

Thank You!

[1] Addendum: The paper is now published, `https://doi.org/10.3390/data4020083`