

LSDF-Portal: Project administration for the Large Scale Data Facility

Felix Karg

Wednesday, February 16th, 2022

Steinbuch Centre for Computing (SCC)



Karlsruher Institut für Technologie

What is the LSDF-Portal

Introduction to Django

Project Results

What is the LSDF-Portal

Introduction to Django

Project Results

What is the LSDF-Portal

Overview

Project Creation

What is the Large Scale Data Facility?

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research
- About 18 PetaByte (1PB = 1000 TB)

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research
- About 18 PetaByte (1PB = 1000 TB)
- All commonly used Protocols

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research
- About 18 PetaByte (1PB = 1000 TB)
- All commonly used Protocols
- Direct Integration with e.g. HoReKa HPC-systems

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research
- About 18 PetaByte (1PB = 1000 TB)
- All commonly used Protocols
- Direct Integration with e.g. HoReKa HPC-systems
- Primarily for KIT, but also worldwide partners and Baden-Württemberg in general

What is the Large Scale Data Facility?

- Cloud-based storage for scientific research
- About 18 PetaByte (1PB = 1000 TB)
- All commonly used Protocols
- Direct Integration with e.g. HoReKa HPC-systems
- Primarily for KIT, but also worldwide partners and Baden-Württemberg in general
- LSDF-Portal: *administration* of storage projects

What is the LSDF-Portal?

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)
- Research groups can request storage for timeframes

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)
- Research groups can request storage for timeframes
- Internally available at <https://lsdf.kit.edu>

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)
- Research groups can request storage for timeframes
- Internally available at <https://lsdf.kit.edu>
- My test instance available at <https://lsdf.fkarg.de>

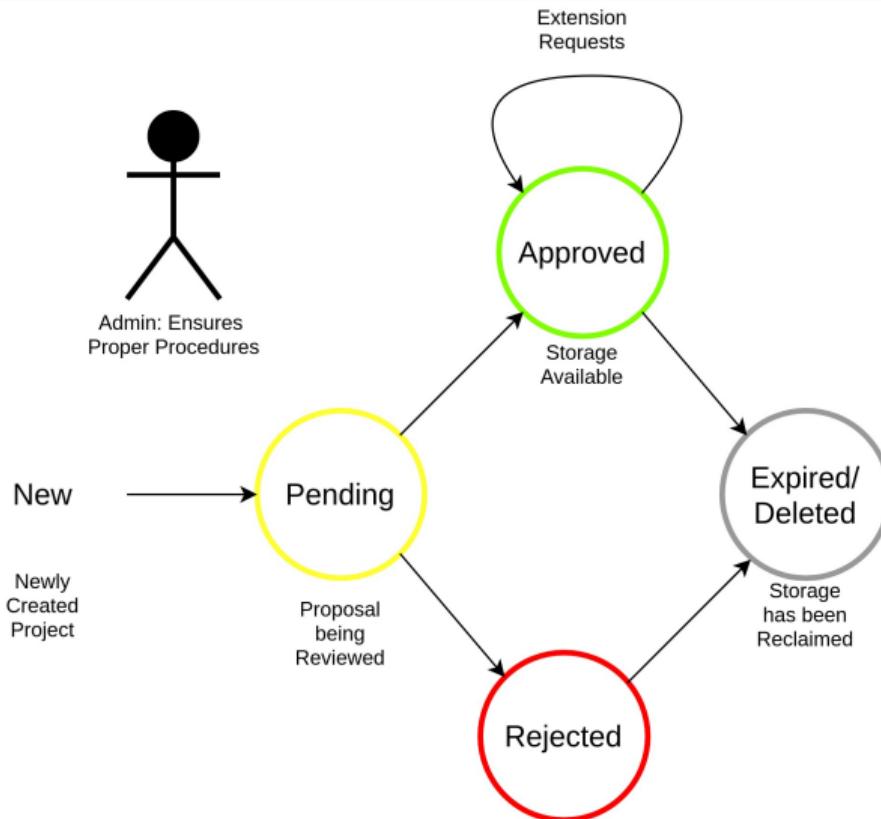
What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)
- Research groups can request storage for timeframes
- Internally available at <https://lsdf.kit.edu>
- My test instance available at <https://lsdf.fkarg.de>
 - Play around with user:test or admin4:lsdf (both admin)

What is the LSDF-Portal?

- LSDF-Portal: *administration* of storage projects
 - Based on Python/Django (more on that later)
- Research groups can request storage for timeframes
- Internally available at <https://lsdf.kit.edu>
- My test instance available at <https://lsdf.fkarg.de>
 - Play around with user:test or admin4:lsdf (both admin)
 - (Cannot properly route from within KIT for some reason)

Project Lifecycle



Live Demo!

What is the LSDF-Portal

Overview

Project Creation

First Login

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Storage projects



No requests are available.

Project Creation

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



[Home](#)

Create new project

Project information

Project name*

Contacts

Owner of the project

Firstname*

Lastname*

Project Creation: Add Contacts

Contacts

Owner of the project

Firstname*

Lastname*

Email*

Institute

Roles

- Head of the project
- Technical contact

Organization

Additional contact



Firstname

Lastname

Email

Institute

Available Fields

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

DFG Discipline*

DFG Review Board*

DFG Subject Area*

End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

Expected storage capacity in TB (1 TB = 1000 GB)

Directory name*

Fields filled out

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

DFG Discipline*

Natural Sciences (3)



DFG Review Board*

Mathematics (312)



DFG Subject Area*

Mathematics (312-01)



End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

15

Expected storage capacity in TB (1 TB = 1000 GB)

Submit Proposal

Protocols

- SSH, SFTP, SCP, HTTPS/Web are enabled for all storage projects
- CIFS
- NFS V3 (Client needs to be connected to KIT-IDM)

Access control

Owner name*

Who should be the owner the project directory? The owner can be a KIT user (e.g. ab1234) or a KIT service account (e.g. OE-ProjectName-0001). Please, contact your ITB to create a service account.

Group name*

Which group should get access to your project directory (e.g. OE-ProjectName-LSDF)? Please contact your ITB to create a group.

Group permission*

No permissions

▼

- Extended group permissions (ACLs)

Submit Storage Proposal

Successful Submission

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)



Your storage request has been successfully saved.

[Home](#)

Edit project



Project information

Project name*

new project

Contacts

Owner of the project

Project in List

The screenshot shows a web-based interface for managing storage projects. At the top, there's a header bar with the KIT logo, the text "Steinbuch Centre for Computing (SCC)", and navigation links for HOME, SITEMAP, ENGLISH, IMPRESSUM, DATENSCHUTZ, and KIT. Below the header is a banner featuring a green abstract background and the SCC logo.

The main content area is titled "Storage projects". It includes a search bar, a dropdown for selecting the number of entries (set to 10), and a "Search:" input field. A green button with a "+" sign is located in the top right corner of the main content area.

A table lists one storage project:

Project Name	Institute	Capacity	Last changes	
new project		Requested: 15 TB	6.2.2022 14:27	

At the bottom, it says "Showing 1 to 1 of 1 entries" and has navigation buttons for "Previous", "1", and "Next".

Page number 17 is visible in the bottom right corner.

Storage Use Histogram

(TODO: Screenshot: Histogram of storage usage.)

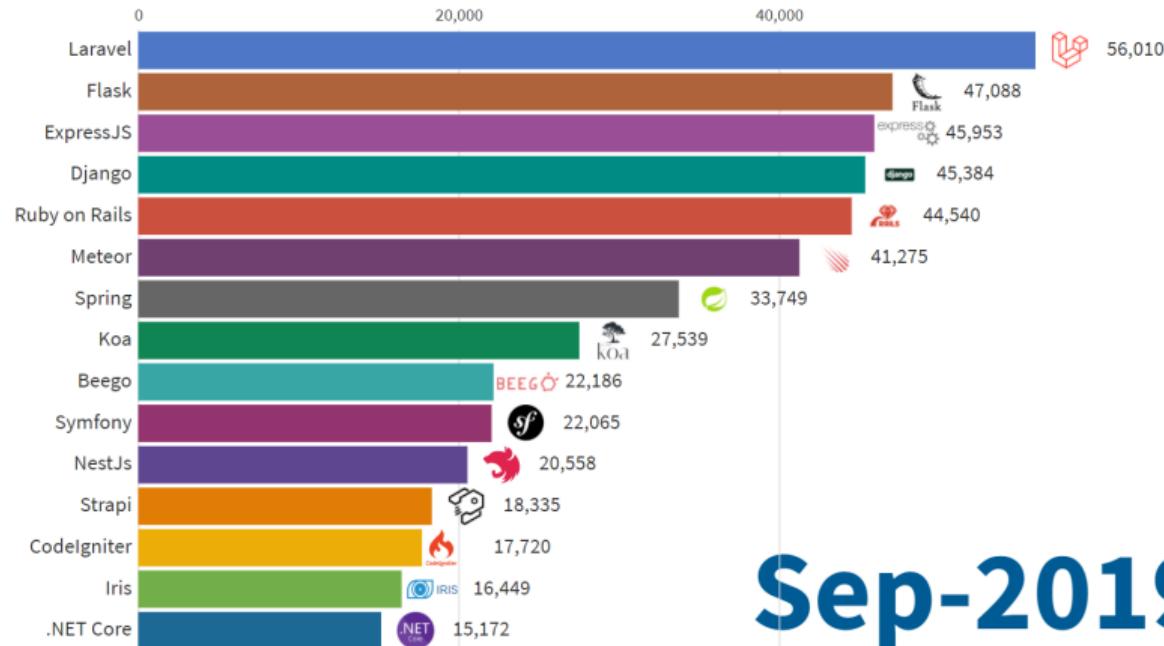
What is the LSDF-Portal

Introduction to Django

Project Results

Django is a Popular Framework

Most Popular Backend Frameworks



Sep-2019

Image source: [1]

What is Django?

Meet Django

Django is a high-level Python web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source.



Ridiculously fast.

Django was designed to help developers take applications from concept to completion as quickly as possible.



Exceedingly scalable.

Some of the busiest sites on the web leverage Django's ability to quickly and flexibly scale.

Django Overview

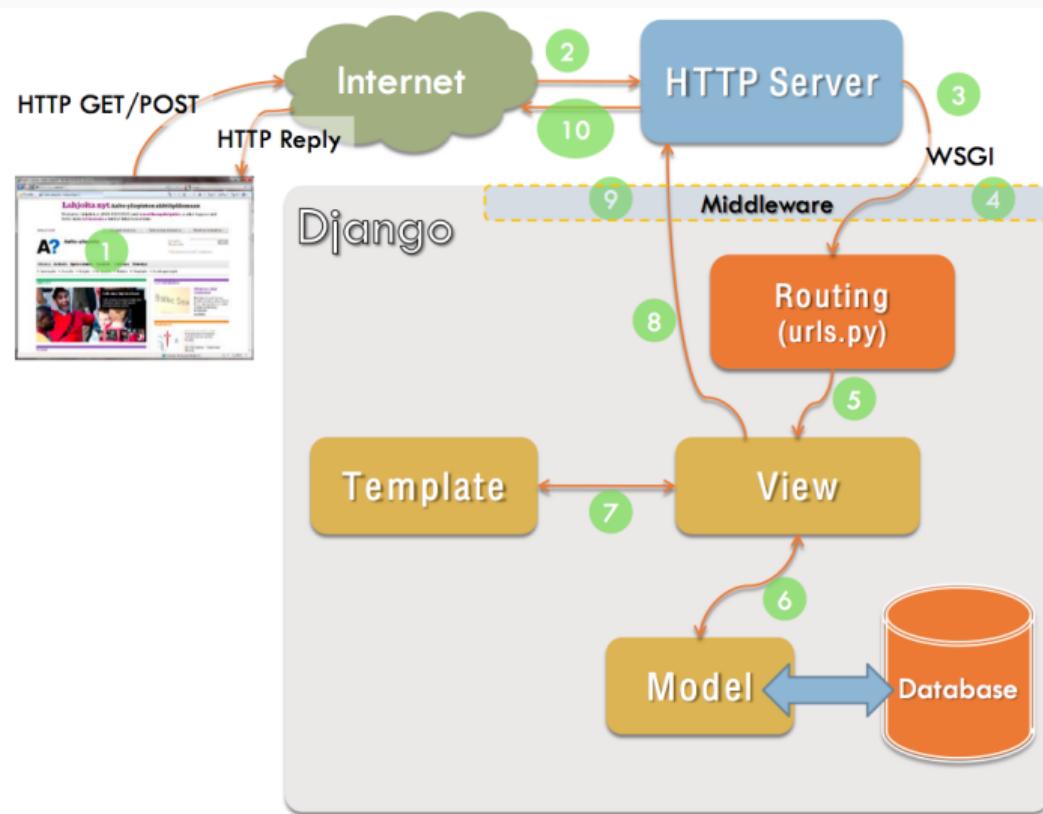


Image source: [2]

What is the LSDF-Portal

Introduction to Django

Project Results

Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

Changes to Logging

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements
- Good to get overview of Codebase and touch many files

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements
- Good to get overview of Codebase and touch many files
- Now: Properly defined Logging levels (debug, info, warn, error, critical)

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements
- Good to get overview of Codebase and touch many files
- Now: Properly defined Logging levels (debug, info, warn, error, critical)
- Now: Log messages are sent to console, syslog and saved in a file locally

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements
- Good to get overview of Codebase and touch many files
- Now: Properly defined Logging levels (debug, info, warn, error, critical)
- Now: Log messages are sent to console, syslog and saved in a file locally
- Now: file and console formatting differs (timestamps, colors, ...)

Changes to Logging

- Used to be done in inconsistent format over individual `print` statements
- Good to get overview of Codebase and touch many files
- Now: Properly defined Logging levels (debug, info, warn, error, critical)
- Now: Log messages are sent to console, syslog and saved in a file locally
- Now: file and console formatting differs (timestamps, colors, ...)
- Now: Rotating files: Keep last five days, overwrite after

Logging Now: Example

```
INFO [basehttp.log_message]: "GET / HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /admin/login?next=/ HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /admin/login/?next=/admin/login%3Fnext%3D/ HTTP/1.1" 200 2245
INFO [basehttp.log_message]: "GET /static/admin/js/nav_sidebar.js HTTP/1.1" 200 1360
INFO [basehttp.log_message]: "GET /static/admin/css/nav_sidebar.css HTTP/1.1" 200 2271
INFO [basehttp.log_message]: "GET /static/admin/css/login.css HTTP/1.1" 200 939
INFO [basehttp.log_message]: "GET /static/admin/css/base.css HTTP/1.1" 200 19513
INFO [basehttp.log_message]: "GET /static/admin/css/responsive.css HTTP/1.1" 200 18545
INFO [basehttp.log_message]: "GET /static/admin/css/fonts.css HTTP/1.1" 200 423
INFO [basehttp.log_message]: "GET /static/admin/fonts/Roboto-Light-webfont.woff HTTP/1.1" 200 85692
INFO [basehttp.log_message]: "GET /static/admin/fonts/Roboto-Regular-webfont.woff HTTP/1.1" 200 85876
WARNING [log.log_response]: Not Found: /favicon.ico
WARNING [basehttp.log_message]: "GET /favicon.ico HTTP/1.1" 404 7000
INFO [basehttp.log_message]: "POST /admin/login/?next=/admin/login%3Fnext%3D/ HTTP/1.1" 200 2405
INFO [basehttp.log_message]: "GET /static/admin/fonts/Roboto-Bold-webfont.woff HTTP/1.1" 200 86184
INFO [basehttp.log_message]: "GET /admin/login/?next=/ HTTP/1.1" 200 2205
INFO [basehttp.log_message]: "POST /admin/login/?next=/ HTTP/1.1" 200 2365
INFO [basehttp.log_message]: "GET /admin/login/ HTTP/1.1" 200 2204
INFO [basehttp.log_message]: "POST /admin/login/ HTTP/1.1" 200 2364
INFO [basehttp.log_message]: "GET / HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /admin/login?next=/ HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /admin/login/?next=/admin/login%3Fnext%3D/ HTTP/1.1" 200 2245
INFO [basehttp.log_message]: "POST /admin/login/?next=/admin/login%3Fnext%3D/ HTTP/1.1" 200 2405
INFO [basehttp.log_message]: "GET /persons/ HTTP/1.1" 200 13898
```

Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

Why attempt automatic documentation?

Situation:

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase
- Especially on available routes, endpoints, internal dependencies

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase
- Especially on available routes, endpoints, internal dependencies

Expectations:

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase
- Especially on available routes, endpoints, internal dependencies

Expectations:

- Automatic, Complete, Up-to-date Overview

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase
- Especially on available routes, endpoints, internal dependencies

Expectations:

- Automatic, Complete, Up-to-date Overview
- Creation of a frequently used reference

Why attempt automatic documentation?

Situation:

- Still only rudimentary grasp on Codebase
- Especially on available routes, endpoints, internal dependencies

Expectations:

- Automatic, Complete, Up-to-date Overview
- Creation of a frequently used reference
- Provides a documentation pipeline and default

Automatically Generate Documentation

The screenshot shows a Swagger UI interface with the following sections:

- traveler** (Travellers)
 - GET /api/traveler - Retrieve all travelers.
 - POST /api/traveler - Create a traveler.
 - GET /api/traveler/{traveler} - Display the specified traveler.
 - PUT /api/traveler/{traveler} - Update the specified traveler.
- trip** (Trips)
 - GET /api/trip - Retrieve all trips.
 - POST /api/trip - Create a trip.
 - GET /api/trip/{trip} - Display the specified trip.
 - PUT /api/trip/{trip} - Update the specified trip.
- quote-request** (Quote Requests)
- purchase-request** (Purchase Requests)
- Schemas**

Failed, because:

(Success from different Project)

Was worth a try.

Automatically Generate Documentation

traveler	Travelers	▼
GET	/api/traveler	Retrieve all travelers.
LOCK		
POST	/api/traveler	Create a traveler.
LOCK		
GET	/api/traveler/{traveler}	Display the specified traveler.
LOCK		
PUT	/api/traveler/{traveler}	Update the specified traveler.
LOCK		
trip	Trips	▼
GET	/api/trip	Retrieve all trips.
LOCK		
POST	/api/trip	Create a trip.
LOCK		
GET	/api/trip/{trip}	Display the specified trip.
LOCK		
PUT	/api/trip/{trip}	Update the specified trip.
LOCK		
quote-request	Quote Requests	▶
purchase-request	Purchase Requests	▶
Schemas		▼

(Success from different Project)

Failed, because:

- Autogeneration usually used to generate documentation for *JSON endpoints*

Was worth a try.

Automatically Generate Documentation

traveler	Travelers	▼
GET	/api/traveler	Retrieve all travelers.
LOCK		
POST	/api/traveler	Create a traveler.
LOCK		
GET	/api/traveler/{traveler}	Display the specified traveler.
LOCK		
PUT	/api/traveler/{traveler}	Update the specified traveler.
LOCK		
trip	Trips	▼
GET	/api/trip	Retrieve all trips.
LOCK		
POST	/api/trip	Create a trip.
LOCK		
GET	/api/trip/{trip}	Display the specified trip.
LOCK		
PUT	/api/trip/{trip}	Update the specified trip.
LOCK		
quote-request	Quote Requests	▶
purchase-request	Purchase Requests	▶
Schemas		▼

(Success from different Project)

Failed, because:

- Autogeneration usually used to generate documentation for *JSON endpoints*
- 'Primitive' Views lacking important information for automatic generation

Was worth a try.

Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

What Research is the Storage used for?

Engineering Sciences (131 Members)		
RB-Nr.	Review Board / Subject Area	Subject Areas
401	Production Technology	□
402	Mechanics and Constructive Mechanical Engineering	□
403	Process Engineering, Technical Chemistry	□
404	Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering	□
405	Materials Engineering	□
406	Materials Science	□
407	Systems Engineering	□
408	Electrical Engineering and Information Technology	□
409	Computer Science	□
	409-01 Theoretical Computer Science	
	409-02 Software Engineering and Programming Languages	
	409-03 Security and Dependability	
	409-04 Operating, Communication, Database and Distributed Systems	
	409-05 Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation	
	409-06 Information Systems, Process and Knowledge Management	
	409-07 Computer Architecture and Embedded Systems	
	409-08 Massively Parallel and Data-Intensive Systems	
410	Construction Engineering and Architecture	□

What Research is the Storage used for?

Engineering Sciences (131 Members)		
RB-Nr.	Review Board / Subject Area	Subject Areas
401	Production Technology	□
402	Mechanics and Constructive Mechanical Engineering	□
403	Process Engineering, Technical Chemistry	□
404	Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering	□
405	Materials Engineering	□
406	Materials Science	□
407	Systems Engineering	□
408	Electrical Engineering and Information Technology	□
409	Computer Science	□
	409-01 Theoretical Computer Science	
	409-02 Software Engineering and Programming Languages	
	409-03 Security and Dependability	
	409-04 Operating, Communication, Database and Distributed Systems	
	409-05 Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation	
	409-06 Information Systems, Process and Knowledge Management	
	409-07 Computer Architecture and Embedded Systems	
	409-08 Massively Parallel and Data-Intensive Systems	
410	Construction Engineering and Architecture	□

- For reporting and analysis Purposes

What Research is the Storage used for?

Engineering Sciences (131 Members)		
RB-Nr.	Review Board / Subject Area	Subject Areas
401	Production Technology	□
402	Mechanics and Constructive Mechanical Engineering	□
403	Process Engineering, Technical Chemistry	□
404	Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering	□
405	Materials Engineering	□
406	Materials Science	□
407	Systems Engineering	□
408	Electrical Engineering and Information Technology	□
409	Computer Science	□
	409-01 Theoretical Computer Science	
	409-02 Software Engineering and Programming Languages	
	409-03 Security and Dependability	
	409-04 Operating, Communication, Database and Distributed Systems	
	409-05 Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation	
	409-06 Information Systems, Process and Knowledge Management	
	409-07 Computer Architecture and Embedded Systems	
	409-08 Massively Parallel and Data-Intensive Systems	
410	Construction Engineering and Architecture	□

- For reporting and analysis Purposes
- We want to know more about our users

What Research is the Storage used for?

Engineering Sciences (131 Members)		
RB-Nr.	Review Board / Subject Area	Subject Areas
401	Production Technology	■
402	Mechanics and Constructive Mechanical Engineering	■
403	Process Engineering, Technical Chemistry	■
404	Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering	■
405	Materials Engineering	■
406	Materials Science	■
407	Systems Engineering	■
408	Electrical Engineering and Information Technology	■
409	Computer Science	■
	409-01 Theoretical Computer Science	
	409-02 Software Engineering and Programming Languages	
	409-03 Security and Dependability	
	409-04 Operating, Communication, Database and Distributed Systems	
	409-05 Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation	
	409-06 Information Systems, Process and Knowledge Management	
	409-07 Computer Architecture and Embedded Systems	
	409-08 Massively Parallel and Data-Intensive Systems	
410	Construction Engineering and Architecture	■

- For reporting and analysis Purposes
- We want to know more about our users
- This includes affiliation of research projects

What Research is the Storage used for?

Engineering Sciences (131 Members)		
RB-Nr.	Review Board / Subject Area	Subject Areas
401	Production Technology	■
402	Mechanics and Constructive Mechanical Engineering	■
403	Process Engineering, Technical Chemistry	■
404	Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering	■
405	Materials Engineering	■
406	Materials Science	■
407	Systems Engineering	■
408	Electrical Engineering and Information Technology	■
409	Computer Science	■
	409-01 Theoretical Computer Science	
	409-02 Software Engineering and Programming Languages	
	409-03 Security and Dependability	
	409-04 Operating, Communication, Database and Distributed Systems	
	409-05 Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation	
	409-06 Information Systems, Process and Knowledge Management	
	409-07 Computer Architecture and Embedded Systems	
	409-08 Massively Parallel and Data-Intensive Systems	
410	Construction Engineering and Architecture	■

- For reporting and analysis Purposes
- We want to know more about our users
- This includes affiliation of research projects
- Good existing classification from 'Deutsche Forschungsgesellschaft'

Selection of DFG Subject Area upon Project Creation

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

DFG Discipline*

DFG Review Board*

DFG Subject Area*

End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

Expected storage capacity in TB (1 TB = 1000 GB)

Directory name*

Selection of DFG Discipline

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

Category*

-
-
- Humanities and Social Sciences (1)
- Life Sciences (2)
- Natural Sciences (3)
- Engineering Sciences (4)**

Field*

-

End of the project*

15.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

12

Expected storage capacity in TB (1 TB = 1000 GB)

Selection of DFG Subject Area without Board

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

Category*

Engineering Sciences (4)

Board*

Field*

End of the project*

15.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

12

Expected storage capacity in TB (1 TB = 1000 GB)

Selection of DFG Review Board

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

Category*

Engineering Sciences (4)

Board*

Field*

End of the project*

15.02.2023

How long you need the storage. You will be able to request an extension

Capacity*

12

Expected storage capacity in TB (1 TB = 1000 GB)

Selection of DFG Subject Area

Please Specify where your project is mainly located according to the [DFG Fachsystematik](#):

Category*

Engineering Sciences (4)



Board*

Computer Science (409)



Field*

Theoretical Computer Science (409-01)

Software Engineering and Programming Languages (409-02)

Security and Dependability (409-03)

Operating, Communication, Database and Distributed Systems (409-04)

Interactive and Intelligent Systems, Image and Language Processing, Computer Graphics and Visualisation (409-05)

Information Systems, Process and Knowledge Management (409-06)

Computer Architecture and Embedded Systems (409-07)

Massively Parallel and Data-Intensive Systems (409-08)

Expected storage capacity in TB (1 TB = 1000 GB)

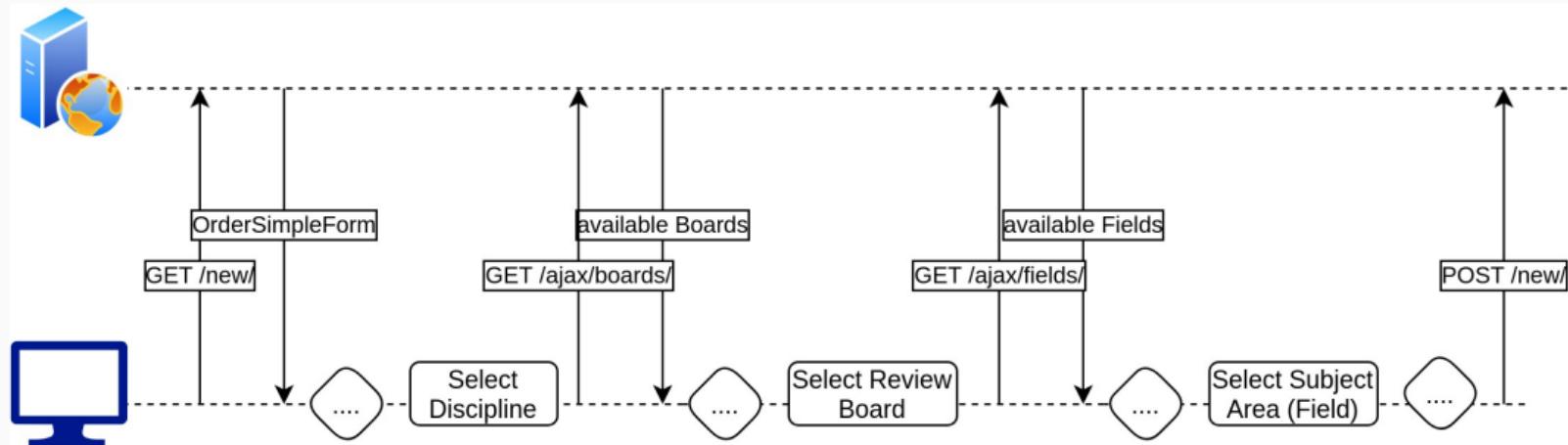
Website: Requesting Fields for Boards

```
1  $("#id_board").change(function () {
2      /* on change of `review board`, request new data for `subject area` field */
3      var url = $("#form").attr("data-fields-url"); // <host>/ajax/fields/
4      var board = $(this).val(); // Which board got selected
5
6      $.ajax({ // request available fields based on selected board
7          url: url, // resolved to <host>/ajax/fields/
8          data: { 'board': board }, // send board as part of request
9          success: function (data, textStatus, jqXHR) {
10              $("#id_field").html(data); // fill in available fields
11          }
12      });
13  });
```

Backend: answering with available Fields

```
1 ## urls.py
2 path('ajax/fields/', views.view_science_fields, name="ajax_load_fields"),
3
4 ## views.py
5 def view_science_fields(request):
6     b_pk = request.GET.get('board') # we get the pk of the selected board
7     if b_pk: # select available Fields from this Board
8         fields = Science_Field.objects.filter(board__pk=b_pk)
9         return render(request, 'dropdown_list_options.html',
10                      {'options': fields})
11
12     return render(request, 'dropdown_list_options.html',
13                  {'options': Science_Field.objects.none()})
14
```

Visualisation



Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

The DFG schema changes frequently

The DFG schema changes frequently

- The DFG schema changes every four years

The DFG schema changes frequently

- The DFG schema changes every four years
- It was last changed in 2020

The DFG schema changes frequently

- The DFG schema changes every four years
- It was last changed in 2020
- So it'll change again in two years

The DFG schema changes frequently

- The DFG schema changes every four years
- It was last changed in 2020
- So it'll change again in two years
- Not clear how much (probably not a whole lot)

The DFG schema changes frequently

- The DFG schema changes every four years
- It was last changed in 2020
- So it'll change again in two years
- Not clear how much (probably not a whole lot)

So I implemented a command to compare any csv to what is currently in the database: `manage.py dfg_schema_diff`

Usage of dfg_schema_diff

```
usage: manage.py dfg_schema_diff [-h] [--locale LOCALE] [--columns COLUMNS]
```

```
...
```

```
FILE
```

Show difference from given file schema to DFG schema in database. By default, ignores the now deprecated hierarchy level 1. ...

positional arguments:

FILE	Path to dfg_systematic.csv
------	----------------------------

options:

-h, --help	show this help message and exit
------------	---------------------------------

--locale LOCALE	Set the language of the name column to select. Can correctly select both '<locale>' and 'prefLabel@<locale>' columns. (Default: 'en')
-----------------	---

--columns COLUMNS	Dictionary mapping columns (numbers) to expected values "level" (in the hierarchy, category: 0, deprecated/ignored: 1, board: 2, field: 3), "notation" (e.g. 101-27), and locale translations, e.g. "en" (double quotes are important!). Defaults to auto.
-------------------	--

```
...
```

```
...
```

Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

```
1 class OrderSimpleForm(forms.ModelForm):
2     p1_email = forms.EmailField(max_length=100, label='Email')
3     p1_institute = forms.CharField(max_length=300, required=False, label='Institute')
4     p1_organization = forms.CharField(max_length=300, required=False, label='Organization')
5     p1_firstname = forms.CharField(max_length=100, required=True, label='Firstname')
6     p1_lastname = forms.CharField(max_length=100, required=True, label='Lastname')
7     p1_roles = forms.MultipleChoiceField(widget=forms.CheckboxSelectMultipleWithDisabledOption, label='Role')
8                     choices=(
9                         ("ROLE_HEAD", "Head of the project"),
10                        ("ROLE_TECH", "Technical contact"),
11                    ))
12
13     p2_email = forms.EmailField(max_length=100, required=False, label='Email')
14     p2_institute = forms.CharField(max_length=300, required=False, label='Institute')
15     p2_organization = forms.CharField(max_length=300, required=False, label='Organization')
16     p2_firstname = forms.CharField(max_length=100, required=False, label='Firstname')
17     p2_lastname = forms.CharField(max_length=100, required=False, label='Lastname')
18     p2_roles = forms.MultipleChoiceField(widget=forms.CheckboxSelectMultiple, required=False, label='Role')
19                     choices=(
20                         ("ROLE_HEAD", "Head of the project"),
21                         ("ROLE_TECH", "Technical contact"),
22                     ))
23     ...
```

Intermediate Results

```
1  class PersonForm(forms.ModelForm):
2      roles = forms.MultipleChoiceField(
3          label="Roles", required=False,
4          choices=(
5              ("ROLE_HEAD", "Head of the project"),
6              ("ROLE_TECH", "Technical contact"),
7          ))
8  class Meta:
9      model = Person
10     fields = [ "first_name", "last_name", "email",
11                 "institute", "roles", "organization",
12             ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

Intermediate Results

```
1 class PersonForm(forms.ModelForm):
2     roles = forms.MultipleChoiceField(
3         label="Roles", required=False,
4         choices=(
5             ("ROLE_HEAD", "Head of the project"),
6             ("ROLE_TECH", "Technical contact"),
7         ))
8     class Meta:
9         model = Person
10        fields = [ "first_name", "last_name", "email",
11                   "institute", "roles", "organization",
12                 ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

- Abstraction to one dedicated PersonForm

Intermediate Results

```
1 class PersonForm(forms.ModelForm):
2     roles = forms.MultipleChoiceField(
3         label="Roles", required=False,
4         choices=(
5             ("ROLE_HEAD", "Head of the project"),
6             ("ROLE_TECH", "Technical contact"),
7         ))
8     class Meta:
9         model = Person
10        fields = [ "first_name", "last_name", "email",
11                   "institute", "roles", "organization",
12                   ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

- Abstraction to one dedicated PersonForm
- PersonFormSet can have arbitrarily many Persons (was only five)

Intermediate Results

```
1 class PersonForm(forms.ModelForm):
2     roles = forms.MultipleChoiceField(
3         label="Roles", required=False,
4         choices=(
5             ("ROLE_HEAD", "Head of the project"),
6             ("ROLE_TECH", "Technical contact"),
7         ))
8     class Meta:
9         model = Person
10        fields = [ "first_name", "last_name", "email",
11                   "institute", "roles", "organization",
12                   ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

- Abstraction to one dedicated PersonForm
- PersonFormSet can have arbitrarily many Persons (was only five)
- Simplifies implementation in View

Intermediate Results

```
1 class PersonForm(forms.ModelForm):
2     roles = forms.MultipleChoiceField(
3         label="Roles", required=False,
4         choices=(
5             ("ROLE_HEAD", "Head of the project"),
6             ("ROLE_TECH", "Technical contact"),
7         ))
8     class Meta:
9         model = Person
10        fields = [ "first_name", "last_name", "email",
11                   "institute", "roles", "organization",
12                   ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

- Abstraction to one dedicated PersonForm
- PersonFormSet can have arbitrarily many Persons (was only five)
- Simplifies implementation in View
- Difficult to propagate ValidationErrors properly

Intermediate Results

```
1 class PersonForm(forms.ModelForm):
2     roles = forms.MultipleChoiceField(
3         label="Roles", required=False,
4         choices=(
5             ("ROLE_HEAD", "Head of the project"),
6             ("ROLE_TECH", "Technical contact"),
7         ))
8     class Meta:
9         model = Person
10        fields = [ "first_name", "last_name", "email",
11                   "institute", "roles", "organization",
12                   ]
13
14 PersonFormSet = formset_factory(PersonForm)
15
16 class OrderSimpleForm(forms.ModelForm):
17     """ Form for Project requests. """
18     owner = OwnerForm() # main contact responsible
19     additional_contacts = PersonFormSet({
20         'form-TOTAL_FORMS': '0',
21         'form-INITIAL_FORMS': '0',
22     })
```

- Abstraction to one dedicated PersonForm
- PersonFormSet can have arbitrarily many Persons (was only five)
- Simplifies implementation in View
- Difficult to propagate ValidationErrors properly
- Would have deleted about 300 lines of boilerplate

Reasons for Failure

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms
- Manually register fields for ValidationErrors

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms
- Manually register fields for ValidationErrors
- Manually put in values for Validation (from PersonFormSet)

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms
- Manually register fields for ValidationErrors
- Manually put in values for Validation (from PersonFormSet)
- Manually take values out to save new Models

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms
- Manually register fields for ValidationErrors
- Manually put in values for Validation (from PersonFormSet)
- Manually take values out to save new Models

Nothing impossible, but requires a lot of intricate details to get right.

Reasons for Failure

- Nesting of Forms is **not** supported (owner within OrderSimpleForm)
- Particularly with the many assumptions and automated parts of ModelForms
- Manually register fields for ValidationErrors
- Manually put in values for Validation (from PersonFormSet)
- Manually take values out to save new Models

Nothing impossible, but requires a lot of intricate details to get right. We decided to instead implement something else.

Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

Motivation

It happens frequently that a project needs more space than initially requested, or isn't done by the time their timeframe runs out (initial timeframe was 4 Years). In these cases, timeframe and capacity can be extended manually by an admin. To automate/improve this process, we implemented it directly.

Timeframe Extension Request View

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)

Timeframe Extension Request

Project name*

new project

Capacity*

15

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

New End of Project*

2024-02-06

Extension request for one year

Reason*

Publications*

Capacity Extension Request View

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Capacity Extension Request

Project name*

Capacity*

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

How long you need the storage. You will be able to request an extension

Request Capacity (New Total)*

Storage capacity you would like to get in TB (1 TB = 1000 GB)

Reason*

Filled out Capacity Extension

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)



Capacity Extension Request

Project name*

new project

Capacity*

15

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

Request Capacity (New Total)*

20

Storage capacity you would like to get in TB (1 TB = 1000 GB)

Reason*

Artifacts got bigger than expected

Extension System Message

The screenshot shows a software interface with a light gray header bar. Below it is a white main area containing three teal-colored horizontal bars, each representing a system message. The first message [6.2.2022 14:32] Changed project state to "PENDING" is followed by a second message [6.2.2022 14:33] Changed project state to "APPROVED". The third message [6.2.2022 14:36] Requested capacity extension to 20 TB is at the bottom. In the top right corner of the main area, there is a green rectangular button with a white speech bubble icon containing a plus sign, and next to it is the text "Save changes". On the far right edge of the main area, there is a vertical gray scroll bar.

[6.2.2022 14:32] Changed project state to "PENDING"

[6.2.2022 14:33] Changed project state to "APPROVED"

[6.2.2022 14:36] Requested capacity extension to 20 TB

System Message showing Extension Approval



Save changes

[6.2.2022 14:32] Changed project state to "PENDING"

[6.2.2022 14:33] Changed project state to "APPROVED"

[6.2.2022 14:36] Requested capacity extension to 20 TB

[6.2.2022 14:59] Changed status of requested capacity extension to APPROVED.



Increased Capacity from User View

Mathematics (312-01) ▾

End of the project*

06.02.2023

Request Timeframe Extension

How long you need the storage. You will be able to request an extension

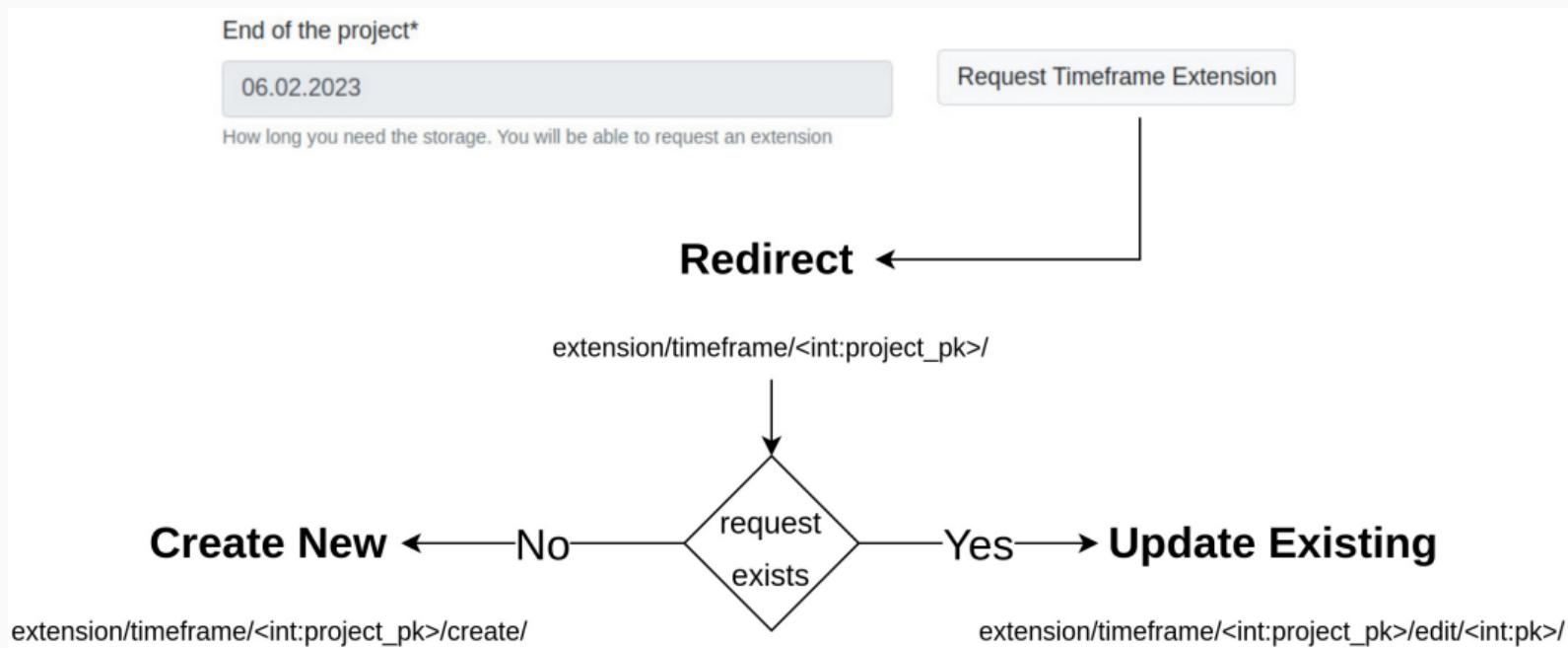
Capacity*

20

Request Capacity Extension

Expected storage capacity in TB (1 TB = 1000 GB)

Routing for Timeframe Requests



Project Results

Proper Logging

Automatically Generate Documentation

What Research is the Storage used for?

Diff CSV to DFG Schema in Database

Attempt to properly modularize PersonForm

Extension Requests

Conclusion

Next Steps

There's a bunch of stuff still left to do:

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design
- Finish dynamic PersonForms implementation

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design
- Finish dynamic PersonForms implementation
- Login with non-unique identifiers (EPPN)

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design
- Finish dynamic PersonForms implementation
- Login with non-unique identifiers (EPPN)
- Advanced User Administration of Accounts

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design
- Finish dynamic PersonForms implementation
- Login with non-unique identifiers (EPPN)
- Advanced User Administration of Accounts
- Realizing Projects through Storage API-Access

Next Steps

There's a bunch of stuff still left to do:

- Deployment of the newly implemented features
- Adapt design to current KIT corporate design
- Finish dynamic PersonForms implementation
- Login with non-unique identifiers (EPPN)
- Advanced User Administration of Accounts
- Realizing Projects through Storage API-Access
- ...

Sources i

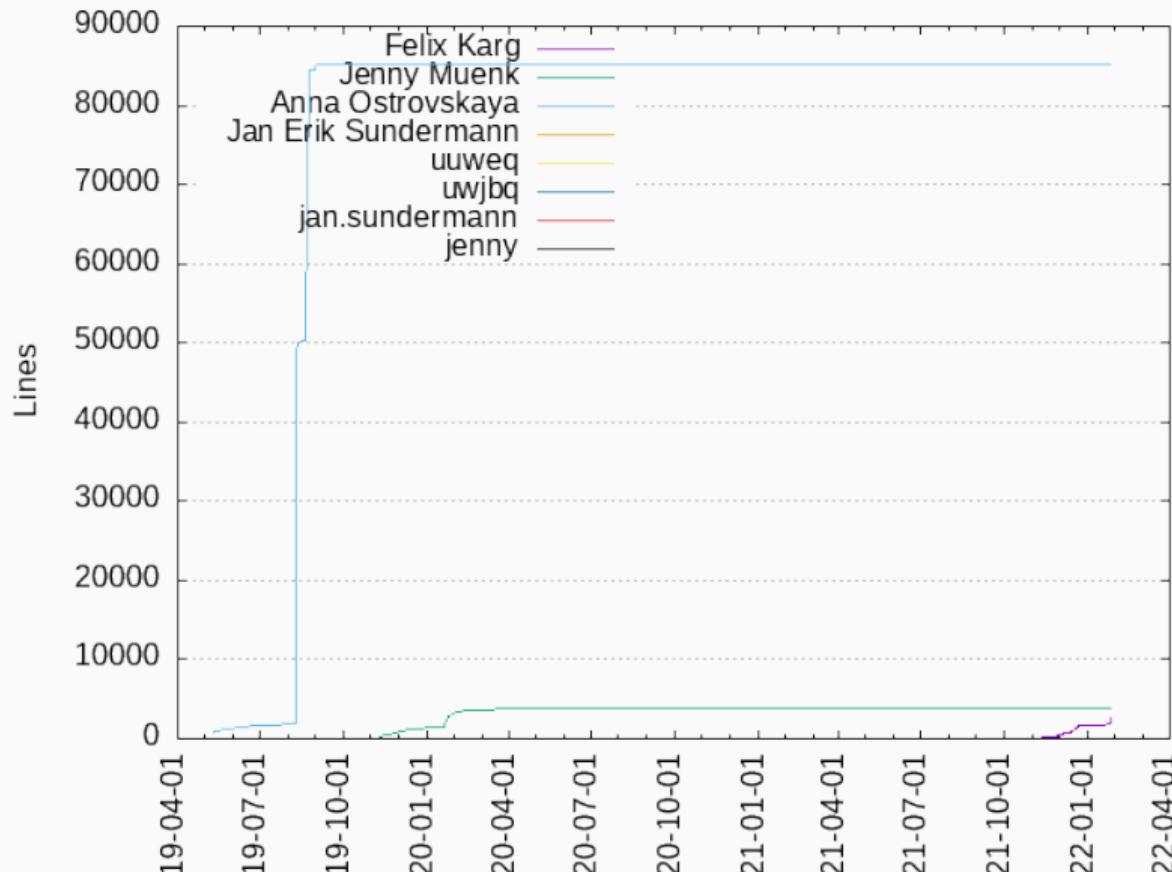
-  "Most popular backend frameworks – 2012/2019."
[https://statisticsanddata.org/data/
most-popular-backend-frameworks-2012-2021/, 2019.](https://statisticsanddata.org/data/most-popular-backend-frameworks-2012-2021/)
accessed 2022-02-07.
-  "Devopedia, "django"." <https://devopedia.org/django>, 2020.
accessed 2022-02-07.

End

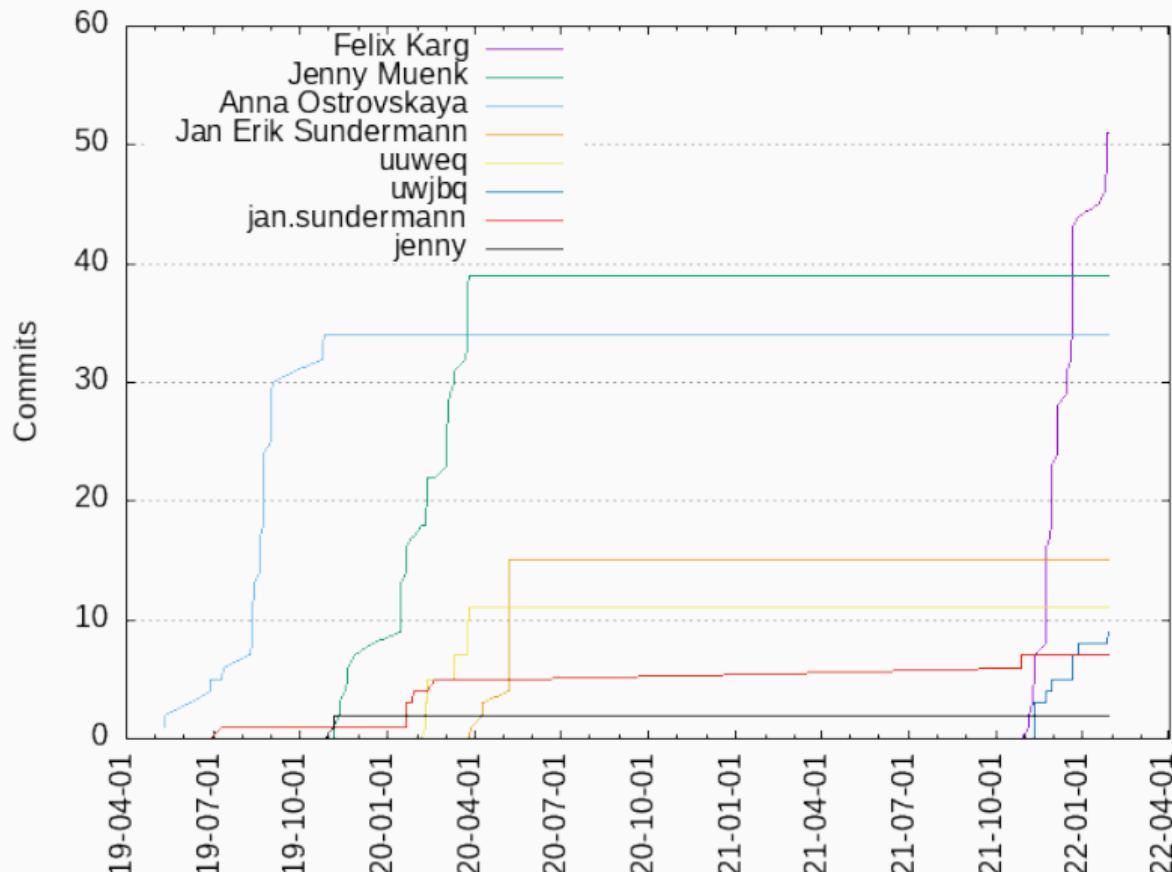
Statistics

More Screenshots

LOC by Author



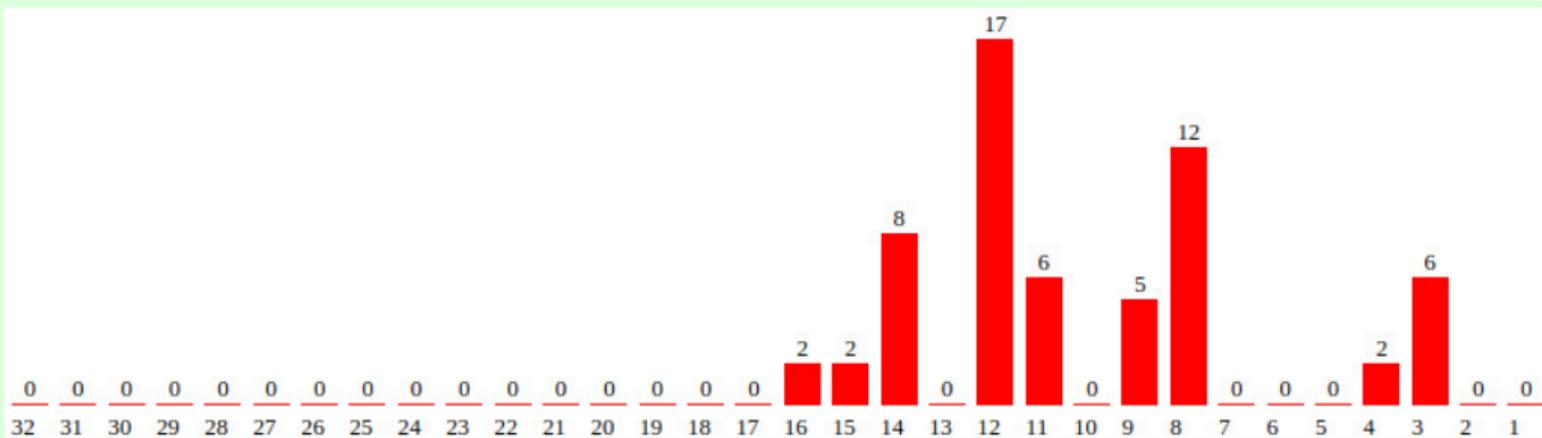
Commits by Author



Recent Activity (Commits by Week)

Weekly activity

Last 32 weeks



LOC stats: Before

Language	Files	Lines	Code	Comments	Blanks
CSS	28	19144	17235	284	1625
JavaScript	27	63499	43096	13725	6678
Python	34	1812	1510	56	246
SVG	14	14	14	0	0
HTML	10	938	838	30	70
— CSS	1	312	282	28	2
— JavaScript	5	155	126	13	16
Total	122	86008	62693	14627	8688

(The numbers don't quite add up, because I omitted e.g. Configs, Markdown, ...)

LOC stats: After

Language	Files	Lines	Code	Comments	Blanks
CSS	28	19162	17248	284	1630
JavaScript	27	63499	43096	13725	6678
Python	48	3078	2485	160	434
SVG	14	14	14	0	0
HTML	15	1231	1107	39	85
— CSS	1	312	282	28	2
— JavaScript	7	244	202	21	21
Total	134	87055	63960	14245	8850

(The numbers don't quite add up, because I omitted e.g. Configs, Markdown, ...)

LOC stats: Diff

Language	Files	Lines	Code	Comments	Blanks
CSS	0	+18	+13	0	+5
JavaScript	0	0	0	0	0
Python	+14	+1266	+975	+104	+188
SVG	0	0	0	0	0
HTML	+5	+293	+269	+9	+15
— CSS	0	0	0	0	0
— JavaScript	+2	+89	+76	+8	+5
Total	+12	+1047	+1267	+382	+162

(The numbers don't quite add up, because I omitted e.g. Configs, Markdown, ...)

Statistics

More Screenshots

More Screenshots

Project Approval

Requesting More Storage

Project Pending

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)



Storage projects



Show entries

Search:

Project Name	Institute	Capacity	Last changes	
new project		Requested: 15 TB	6.2.2022 14:32	

Showing 1 to 1 of 1 entries

Previous **1** Next

State Change Message

Want to add more service accounts? Contact your ITB to request access.

Group permission*

No permissions



Extended group permissions (ACLs)



Save changes

[6.2.2022 14:32] Changed project state to "PENDING"



Project Approved

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)

Storage projects



Show 10 entries

Search:

Project Name	Institute	Capacity	Last changes	
new project		Requested: 15 TB	6.2.2022 14:33	

Showing 1 to 1 of 1 entries

[Previous](#) [1](#) [Next](#)

State Change Message

Group permission*

No permissions

Extended group permissions (ACLs)



Save changes

[6.2.2022 14:32] Changed project state to "PENDING"

[6.2.2022 14:33] Changed project state to "APPROVED"



More Screenshots

Project Approval

Requesting More Storage

Extension Requests: The Project Needs to be Approved

DFG REVIEW BOARD

Mathematics (312) ▾

DFG Subject Area*

Mathematics (312-01) ▾

End of the project*

06.02.2023

Request Timeframe Extension

How long you need the storage. You will be able to request an extension

Capacity*

15

Request Capacity Extension

Expected storage capacity in TB (1 TB = 1000 GB)

Directory name*

projectdirectory

What should be the name of the project directory? We recommend using a short project acronym. Allowed characters are "a-z 0-9 _ -"

Timeframe Extension Request View

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Timeframe Extension Request

Project name*

new project

Capacity*

15

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

06.02.2023

New End of Project*

2024-02-06

How long you need the storage. You will be able to request an extension

Extension request for one year

Reason*

Publications*

Capacity Extension Request View

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)



Capacity Extension Request

Project name*

Capacity*

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

How long you need the storage. You will be able to request an extension

Request Capacity (New Total)*

Storage capacity you would like to get in TB (1 TB = 1000 GB)

Reason*

Filled out Capacity Extension

[HOME](#) | [SITEMAP](#) | [ENGLISH](#) | [IMPRESSUM](#) | [DATENSCHUTZ](#) | [KIT](#)



Karlsruher Institut für Technologie



Steinbuch Centre for Computing (SCC)



Capacity Extension Request

Project name*

new project

Capacity*

15

Expected storage capacity in TB (1 TB = 1000 GB)

End of the project*

06.02.2023

How long you need the storage. You will be able to request an extension

Request Capacity (New Total)*

20

Storage capacity you would like to get in TB (1 TB = 1000 GB)

Reason*

Artifacts got bigger than expected

Extension System Message

[6.2.2022 14:32] Changed project state to "PENDING"

[6.2.2022 14:33] Changed project state to "APPROVED"

[6.2.2022 14:36] Requested capacity extension to 20 TB



Save changes



Projects Admin View

Storage projects



Show 10 entries

Search:

Extension Request Overview

HOME | SITEMAP | ENGLISH | IMPRESSUM | DATENSCHUTZ | KIT



Karlsruher Institut für Technologie



Home

Open Capacity Requests

Show 10 entries

Search:

Project	Created	State	Reason	Capacity	Decision
new project	2022-02-06	PENDING	Artifacts got bigger than expected	New: 20 TB Current: 15 TB	<button>Approve</button> <button>Reject</button>
some other project	2022-02-06	PENDING	some weird reason	New: 18 TB Current: 16 TB	<button>Approve</button> <button>Reject</button>

Approving Request

Closed Capacity Requests

Show 10 entries

Search:

Project	Created	State	Reason	Capacity	Decision
new project	2022-02-06	APPROVED	Artifacts got bigger than expected	New: 20 TB Current: 20 TB	<button>Reopen</button>

System Message showing Extension Approval



Save changes

[6.2.2022 14:32] Changed project state to "PENDING"

[6.2.2022 14:33] Changed project state to "APPROVED"

[6.2.2022 14:36] Requested capacity extension to 20 TB

[6.2.2022 14:59] Changed status of requested capacity extension to APPROVED.



Increased Capacity from User View

Mathematics (312-01)



End of the project*

06.02.2023

Request Timeframe Extension

How long you need the storage. You will be able to request an extension

Capacity*

20

Request Capacity Extension

Expected storage capacity in TB (1 TB = 1000 GB)