

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

Felix Karg

February 16th, 2022

Steinbuch Centre for Computing (SCC)



What is the LSDF-Portal

Introduction to Django

Project Results

Next Project?

What is the LSDF-Portal

Introduction to Django

Project Results

Next Project?

Live Demo!

What is the LSDF-Portal

Overview

Project Lifecycle

Project Creation

Project Approval

Requesting More Storage

Storage management for 18PB for research projects

Based on Python/django (more on that later)

available at: lsdf.kit.edu

(my test server: lsdf.fkarg.de with user:test and admin4:lsdf)

What is the LSDF-Portal

Overview

Project Lifecycle

Project Creation

Project Approval

Requesting More Storage

- Project Creation -> NEW
- Admins take a look -> PENDING
- granted -> APPROVED
- not granted -> REJECTED
- done -> EXPIRED/DELETED

What is the LSDF-Portal

Overview

Project Lifecycle

Project Creation

Project Approval

Requesting More Storage

Storage projects



No requests are available.

[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*

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)

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

Your storage request has been successfully saved.

[Home](#)

Edit project



Project information

Project name*

new project

Contacts

Owner of the project

Project in List



Steinbuch Centre for Computing (SCC)

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



Storage projects



Show entries

Search:

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

Showing 1 to 1 of 1 entries

Previous

1

Next

What is the LSDF-Portal

Overview

Project Lifecycle



Project Creation

Project Approval

Requesting More Storage

Project Pending

HOME | SITEMAP | ENGLISH | IMPRESSUM | DATENSCHUTZ | KIT



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

contact your ITB to create a service account.

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](#)



Steinbuch Centre for Computing (SCC)



SCC

Storage projects



Show 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"



What is the LSDF-Portal

Overview

Project Lifecycle

Project Creation

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

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](#)



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)*

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

Reason*

Filled out Capacity Extension

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



Steinbuch Centre for Computing (SCC)

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



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



Projects Admin View

Storage projects

+

Publications

Extension requests

Show

10

 entries

Search:

Project Name	Institute	Capacity	Last changes
name		Requested: 66 TB	11.1.2022 00:00
new project		Requested: 15 TB	6.2.2022 14:33
sghg		Requested: 1000000000000000000 TB	6.2.2022 14:37
some other project		Requested: 16 TB	23.1.2022 02:43

Extension Request Overview

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



Steinbuch Centre for Computing (SCC)



[Home](#)

Open Capacity Requests

Show 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	Approve Reject
some other project	2022-02-06	PENDING	some weird reason	New: 18 TB Current: 16 TB	Approve Reject

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

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

Request Timeframe Extension

Capacity*

20

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

Request Capacity Extension

Storage Use Histogram

(TODO: Screenshot: Histogram of storage usage.)

What is the LSDF-Portal

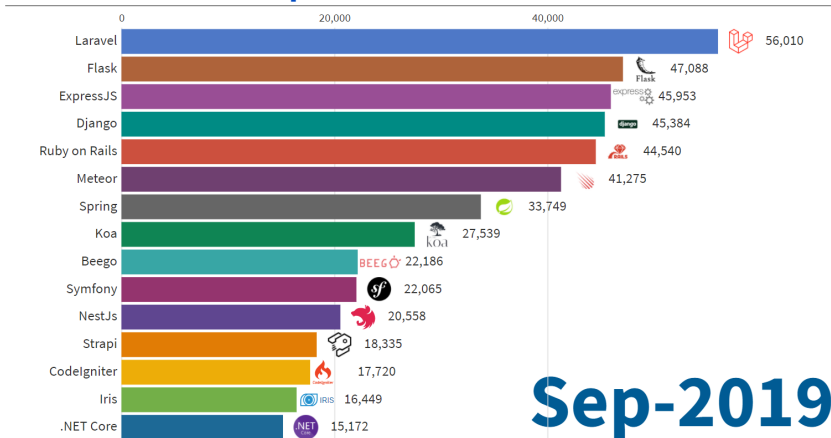
Introduction to Django

Project Results

Next Project?

Django is a Popular Framework

Most Popular Backend Frameworks



Sep-2019

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.

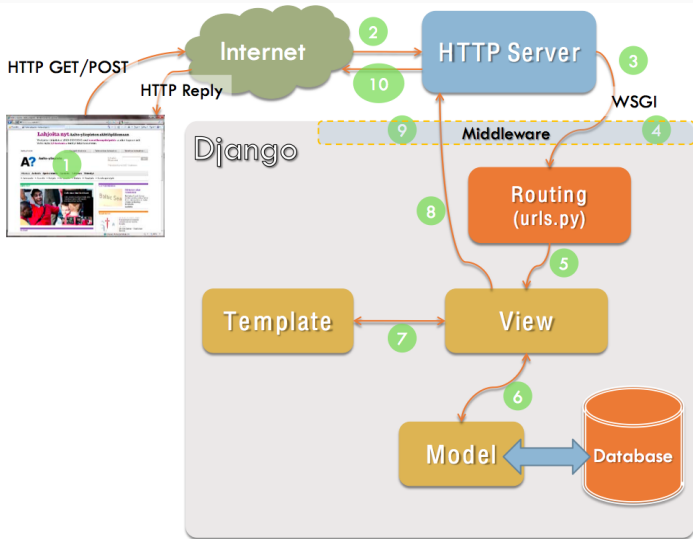
Django was designed to help developers take applications from concept to completion as quickly as possible.



Exceedingly scalable.

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

Django Overview



Source: [2]

What is the LSDF-Portal

Introduction to Django

Project Results

Next Project?

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

From Django 2 to Django 3

(Django 4 released during the project, has been decided to go to 3 in production first)

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

Logging Now

```
INFO [basehttp.log_message]: "GET /extensions/capacity/8/reopen/ HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /extensions/ HTTP/1.1" 200 19287
INFO [basehttp.log_message]: "GET /edit/8/ HTTP/1.1" 200 56667
INFO [basehttp.log_message]: "GET /edit/9/ HTTP/1.1" 200 57407
INFO [utils.save_system_msg]: System Message on new project: Changed project capacity from "20" to "15" TB
INFO [basehttp.log_message]: "POST /edit/9/ HTTP/1.1" 302 0
INFO [basehttp.log_message]: "GET /edit/9/ HTTP/1.1" 200 58029
INFO [basehttp.log_message]: "GET /edit/9/ HTTP/1.1" 200 56545
INFO [basehttp.log_message]: "GET /edit/9/ HTTP/1.1" 200 53654
INFO [basehttp.log_message]: "GET /edit/9/ HTTP/1.1" 200 56545
INFO [basehttp.log_message]: "GET /extensions/ HTTP/1.1" 200 19287
INFO [utils.save_system_msg]: System Message on new project: Changed status of requested capacity extension to APPROVED.
INFO [utils._notify]: extension-update new project
```

- Properly defined Logging levels (debug, info, warn, error, critical)
- Logging to console, syslog and file
- Differing formatter for console and other (timestamps, ...)
- Rotating files: Keep last five days, overwrite after

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests


Automatically Generate Documentation

traveler Travelers ▼

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 ▼

(Example)

Failed, because:

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

Was worth a try.

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

Excerpt of DFG subject areas

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	⊞


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*

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](#):

DFG Discipline*

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

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 Subject without Board

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

DFG Discipline*

Natural Sciences (3)



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 Subject

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)

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*

(TODO: code of change function with request) (TODO:
maybe from other side as well?)

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

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: (TODO: code listing)

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

Explain the problem (show exemplary code) of redundancy Managed to get quite far, very good (deep) learning example, and absolutely possible. just requires some more time than expected. Nags me that it's not modular yet. Some day.

Project Results

Upgrade Dependencies

Proper Logging

Attempt to automatically generate documentation

Selecting a DFG board

Diff to schema in database

Attempt to properly modularize PersonForm

Extension Requests

Extension Requests

Much-requested feature. (TODO: Overview of newly introduced routes, views, forms, models and relations)

What is the LSDF-Portal

Introduction to Django

Project Results

Next Project?

Previously barely any documentation existed. Now a bunch of stuff has comments. Show tokei stats?

LOC stats

=====

Language	Files	Lines	Code	Comments	Blanks
----------	-------	-------	------	----------	--------

=====

CSS	28	19162	17248	284	1630
Dockerfile	1	11	7	0	4
JavaScript	27	63499	43096	13725	6678
Python	48	3078	2485	159	434
SVG	14	14	14	0	0
YAML	1	27	22	1	4

----- HTML 15 1239

1111	43	85	—	CSS	1	312	282	28	2	—	JavaScript	7	248	204	23	21	(Total)	1799
1597	94	108	—															

----- Markdown 1 55 0 36 19 — BASH 1 4 4 0 0 (Total) 59 4 36 19

=====

Total	135	87085	63983	14248	8854
-------	-----	-------	-------	-------	------

=====

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

- change to new kit design
- upgrade to dynamic personForms
- login with non-unique identifiers
- user administration: registered/unregistered/...
- automation of project creation by accessing the storage api directly
- ...

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