



Compute, Storage, and Hosting

Discover services for data-driven life science

May 28, 2024

Johan Alfredéen and Hamza Imran Saeed
SciLifeLab Serve team

Overview



Compute

Storage

Hosting

SciLifeLab Serve

SciLifeLab Data Platform

Q&A





Compute resources

Computations vs hosting/publishing

Traditional HPC

GPU clusters

Sensitive vs non-sensitive data

NAISS (SNIC)

Data support: SND, InfraVis

Many universities provide own compute resources

Apply at SUPR or European Prace portals

<https://data.scilifelab.se/resources/e-infrastructure/>

Compute Resource	Description	Example uses
SENS, NAISS	Life science tools, sensitive data	Bioinformatics
Rackham, NAISS	Life science tools, no sensitive data	Bioinformatics
NAISS Science Cloud	Cloud system, Nvidia GPUs	AI and ML
Alvis, NAISS	GPU cluster	AI and ML
Tetralith, NAISS	Large CPU partition, GPUs	HPC
Dardel, NAISS	HPC, AMD GPUs	HPC
Berzelius	GPU-rich NVIDIA DGX	AI/ML research, etc
AIDA	Secure, GPUs	Medical imaging diagnostics AI
LUMI, EuroHPC	Larger than Dardel	Share work with researchers in Europe using LUMI
EuroHPC	Multiple resources in Europe	



Storage

Check with your host university

NAISS, Swestore, Swedish Science Cloud

SciLifeLab FAIR Storage

Projects that advance Swedish data-driven life science

Yearly calls, open now

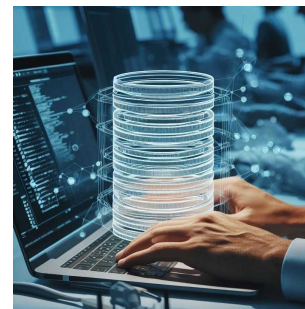
<https://data.scilifelab.se/services/fairstorage/>

SciLifeLab Data Platform

European Open Science Cloud

SND, Swedish National Data Service

Other storage options, see <https://data.scilifelab.se/resources/storage/>



Hosting of applications and tools



SciLifeLab Serve

SciLifeLab Data Platform - general hosting

Long-term sustainability

Need for apps as supplement to articles



nature communications



Article

<https://doi.org/10.1038/s41467-022-33050-0>

An online atlas of human plasma metabolite signatures of gut microbiome composition

Received: 8 February 2022

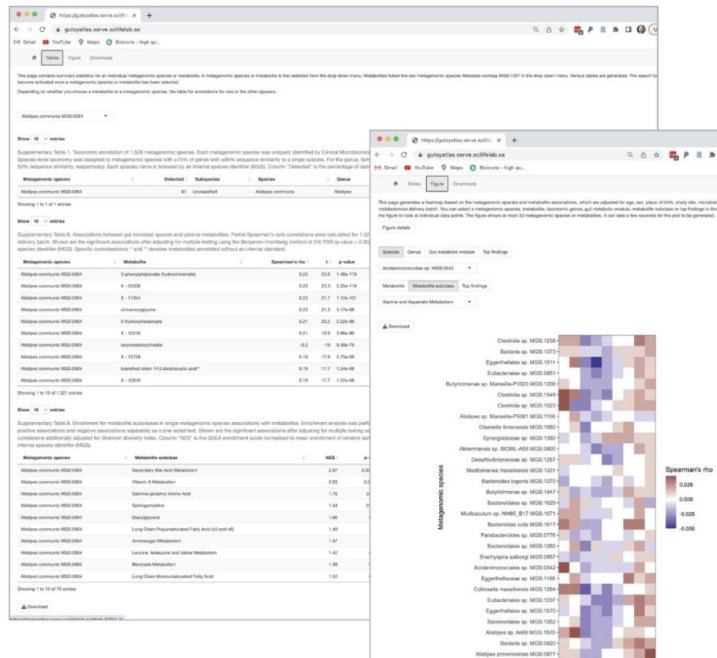
Accepted: 26 August 2022

Published online: 23 September 2022

Check for updates

Koen F. Dekkers^{1,13}, Sergi Sayols-Baixeras^{1,13}, Gabriel Baldanzi¹, Christoph Nowak², Ulf Hammar³, Diem Nguyen³, Georgios Varotsis¹, Louise Brunkwall⁴, Nynne Nielsen⁵, Aron C. Eklund⁶, Jacob Bak Holm⁶, H. Bjørn Nielsen⁶, Filip Ottosson⁶, Yi-Ting Lin⁷, Shafqat Ahmad⁸, Lars Lind⁶, Johan Sundström^{7,8}, Gunnar Engström⁹, J. Gustav Smith^{9,10,11}, Johan Årnlöv¹², Marju Orho-Melander^{9,13} & Tove Fall^{9,13}✉

Human gut microbiota produce a variety of molecules, some of which enter the bloodstream and impact health. Conversely, dietary or pharmacological compounds may affect the microbiota before entering the circulation. Characterization of these interactions is an important step towards understanding the effects of the gut microbiota on health. In this cross-sectional study, we used deep metagenomic sequencing and ultra-high-performance liquid chromatography linked to mass spectrometry for a detailed characterization of the gut microbiota and plasma metabolome, respectively, of 8583 participants invited at age 50 to 64 from the population-based Swedish Cardio-Pulmonary Biomarker Study. Here, we find that the gut microbiota explain up to 58% of the variance of individual plasma metabolites and we present 997 associations between alpha diversity and plasma metabolites and 546,819 associations between specific gut metagenomic species and plasma metabolites in an online atlas (<https://gutsyatlus.serve.sciifelab.se/>). We exemplify the potential of this resource by presenting novel associations between dietary factors and oral medication with the gut microbiome, and microbial species strongly associated with the uremic toxin *p*-cresol sulfate. This resource can be



SciLifeLab Serve



Hosting of apps and trained ML models for scientists

Aim to be quick and easy; self-service

FAIR principles

Own subdomain name:

my-project.serve.scilifelab.se

Available free of charge to all life science researchers affiliated with a Swedish research institution

The screenshot shows the SciLifeLab Serve website. At the top is a navigation bar with the SciLifeLab logo, the text 'SciLifeLab Serve', and links for 'Home', 'Apps', 'Models', 'User guide', 'Register', and 'Log in'. Below the navigation bar is a hero section with the SciLifeLab logo and the text 'SciLifeLab Serve'. A paragraph describes the platform: 'SciLifeLab Serve (beta) is a platform offering machine learning model serving, app hosting, web-based integrated development environments, and other tools to life science researchers affiliated with a Swedish research institute.' Below this are three feature cards: 'Machine Learning model serving' (with a robot icon), 'ML model demo apps' (with a server icon), and 'Data and analysis apps' (with a bar chart icon). Each card contains a brief description of its functionality. Below the feature cards is a 'Recent updates' section with two columns: 'Apps' and 'Models'. The 'Apps' column lists 'Dictyostelium discoideum RNA...' (a Shiny App updated 21 May 2024) and 'Vector oligo search' (a Custom App updated 21 May 2024). The 'Models' column lists 'PTP: Predictive Target Profile' (an ML Model updated 21 Dec 2023) and 'ALLIUM (ALL subtype Identific...)' (an ML Model updated 19 Mar 2024). A third card, 'Ames Mutagenicity Prediction', is partially visible at the bottom.

SciLifeLab Serve



- In beta testing, open to all users
- Constantly changing and improving based on user feedback
- Stable service



Machine Learning model serving

Deploy your trained models and obtain API endpoints for inference requests. We support TorchServe, TensorFlow Serving, MLFlow.



ML model demo apps

Share web apps with a graphical user interface to allow users to make inferences with custom input. We support apps built using Gradio, Streamlit, Flask, FastAPI, etc.



Data and analysis apps

Share web apps with interactive visualisations, analyses, text, tables, etc. We support apps built using R Shiny, Plotly Dash, etc.

Primary focus: platform for deployment of ML models and data science apps

Secondary use: browser-based notebooks, including for use in teaching

Public apps and ML models in Serve



Public apps, tools and ML models relating to life science

Currently 80 **public** apps

Available for everyone to use

Source code of public apps

SciLifeLab Serve

Home Apps Models User guide My projects Profile

Public apps 78

starbase

starbase is a database and toolkit for exploring of large transposable elements in Fungi

Owner: Adrian Forsythe

Open

Custom App Last updated: 26 Apr 2024

KG Dashboard

Dashboard for the R4A Expertise KG

Owner: Yojana Gadiya

[Custom App](#) [Dashboard](#) [KG](#)

Open

Custom App Last updated: 23 Apr 2024

United Kingdom Flower Classification App

This app allows users to input an image (in jpeg format) of a flower and get a prediction on what type of flower it is from a machine learning model. Top three most likely flower types according to the model's prediction are shown.

Owner: Arnold Kochari

[Image classification](#)

Open

Custom App Last updated: 18 Apr 2024

BAGS-Shiny

Biome-specific gene catalogues have been recovered for many environments using shotgun metagenomics, followed by assembly and gene calling on the assembled contigs. We recently proposed a novel mix-assembly strategy, combining individual and co-assembly approaches, and used this approach to assemble an extensive non-redundant gene set from 124 Bal... [see more]

Owner: Luis Fernando Delgado Zambrano

Open

Custom App Last updated: 18 Apr 2024

Lee_et_al_2024_supplementary_data

All the browsable datasets from Lee et al 2024. "Open source, high throughput, targeted in situ transcriptomics"

Owner: Marco Grillo

Open

Shiny App Last updated: 18 Apr 2024

ADHD medication among 0-19 year olds in Sweden

Dashboard plotting the number of people in Sweden aged 0-19 who have have filled a prescription for ADHD medication each year.

Owner: Arnold Kochari

[Public health](#) [shiny](#)

Open

Shiny App Last updated: 18 Apr 2024

COVID-19 in Sweden

This dashboard visualises the spread of COVID-19 in Sweden: confirmed cases, ICU admissions, and deaths between February 2020 and March 2023.

Open

Custom App Last updated: 18 Apr 2024

cpLogD-v2.0

Conformal regression model for predicting LogD for drug-like molecules, trained on ChEMBL version 33 with over 2.2M compounds using the CPSPgn

Open

Custom App Last updated: 18 Apr 2024

Ames Mutagenicity Prediction

Predicts Ames Mutagenicity for a chemical structure using an SVM model with Signatures features via the CDSim mutation tool

Open

Custom App Last updated: 18 Apr 2024

Teaching using Serve



Small courses, 20-30 students

Participants can use JupyterLab notebooks or RStudio to work on tasks

Teacher and students do not need to worry about setup and resources

Easy setup through teacher-provided code repository or docker image

Application form



My projects > Example project

Example project

Settings

Project owner: admin@test.com

Develop

No instances.

Jupyter Lab

Create



RStudio

Create



VS Code

Create





Serve : limitations and requirements for hosting apps

“Simple” apps (e.g. no database or complex requirements)

Public code, published docker image

Default resources

- 2 vCPU, 4 GB RAM per user app
- 1 GB storage per user project
- Can request more resources

No sensitive user data

Intent is for all apps and models to be made public

Pre-trained ML models

Currently no new ML models - reworking model hosting function

App types in Serve



Custom App

Apps built with Gradio, Streamlit, Flask, etc.



Create

Dash App



Create

Shiny App



Create

TissUUmeps App

App to visualise and explore data using TissUUmeps.



Create

Jupyter Lab



Create

RStudio



Create

VS Code



Create

Additional app types through the *Custom app* type, e.g. Flask, Streamlit etc

Using Serve for hosting apps



To use Serve:

1. Build docker image and push it to a public registry
2. Create user account on Serve
3. Create a Serve project
4. Create the app

Refer to the user guide

<https://serve.scilifelab.se/docs/>

Dedicated team offers consultations, support and training

New project

Choose a template for your new project.

Default project



Use this template if you intend to only deploy apps or use integrated development environments (IDEs).

Create

Project with ML serving



Use this template if you intend to deploy machine learning models using the serving functionality.

Create

Create Shiny App

This form allows you to start hosting a Shiny app at SciLifeLab Serve. Please read our [documentation page on Shiny apps](#) for step-by-step instructions.

Name ①

Shiny Viz App

Description ①

A great Shiny visualization app.

☐ Activate Proxy

① Check this box if your app requires lots of compute power. This will enable a proxy to handle the load.

Subdomain ①

shiny-viz

The subdomain is available

Hardware ①

2 vCPU, 4 GB RAM

Permission ①

Private

Port ①

3838

Image ①

registry/repository/image:tag

Tags

shiny microbes

Enter a comma-separated tag string

Cancel

Submit



Serve : coming soon

- Better ML model serving
- GPU resources
- FAIR: metadata, model cards, DOI

Link to Serve: <https://serve.scilifelab.se>

General hosting - SciLifeLab Data Platform



More flexible than Serve

More complex apps, such as requiring a database

If requires more compute and storage than Serve provides

Free to use; requirements: open source code, security limitations, etc.

Not accepting new applications at this time, but will in the future. In case of questions, contact us.

<https://data.scilifelab.se/services/hosting/>

DATA

- Research projects
- Infrastructures
- Data sharing platform
- Databases
- Data acquisition technology



USERS

- Researchers
- Bioinformaticians
- Software engineers
- Health care
- Industry partners

E-INFRASTRUCTURES

Primarily through collaborations

THANK YOU



SciLifeLab



Vetenskapsrådet

*Knut and Alice
Wallenberg
Foundation*

Serve support:

serve@scilifelab.se

FAIR Storage

FAIRStorage@scilifelab.se

Other questions:

datacentre@scilifelab.se