Data management

published datasets

data collection

data model
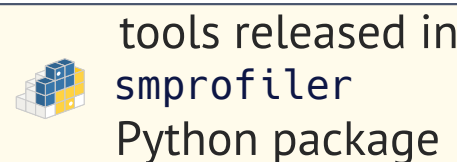

database

80m cells

1500 samples

>100 proteins

source code

data curation and ETL

preprocessing images

cell segmentation and marker quantification

efficient/normalized dataset archiving

tools released in smprofiler Python package

Analysis

SMProfiler

cloud resources

Infrastructure-as-Code

FastAPI — API framework

TS — client web application

quantitative analysis

squidpy — cell-level spatial metrics

KEDA — workload autoscaling

histocartography — GNN framework

ML framework

threshold determination / gating

identification/classification of cell phenotypes

exploratory data analysis

metrics computation
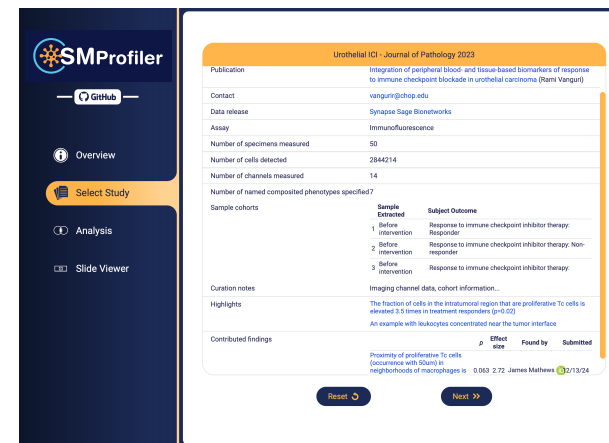
metric computation in realtime

model training

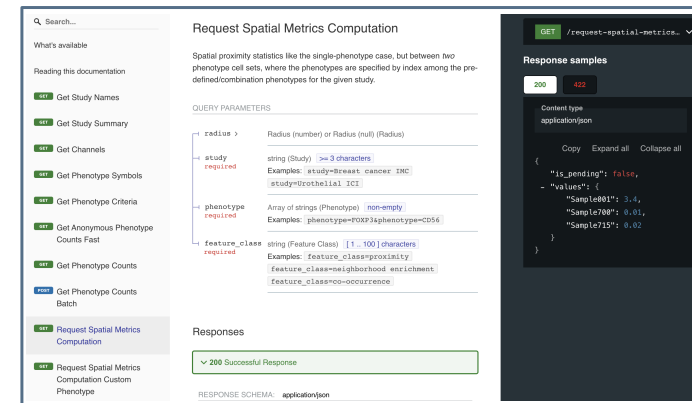evaluating models/features

study overview

feature matrix stats

The average value of the proximity score for phenotype(s) CD38+, CD15+ is 1.43 times higher in cohort 3 than in cohort 1. (p=0.015, t-test)

API service

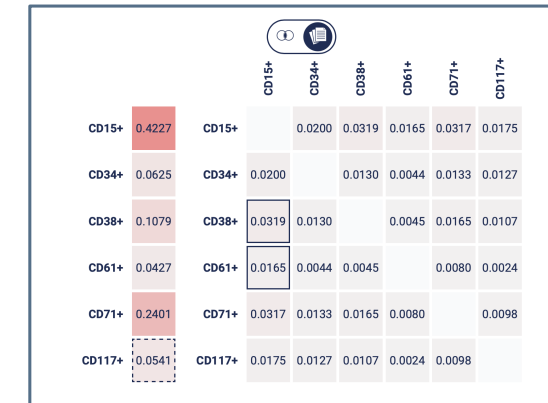### heatmaps

virtual slide viewer
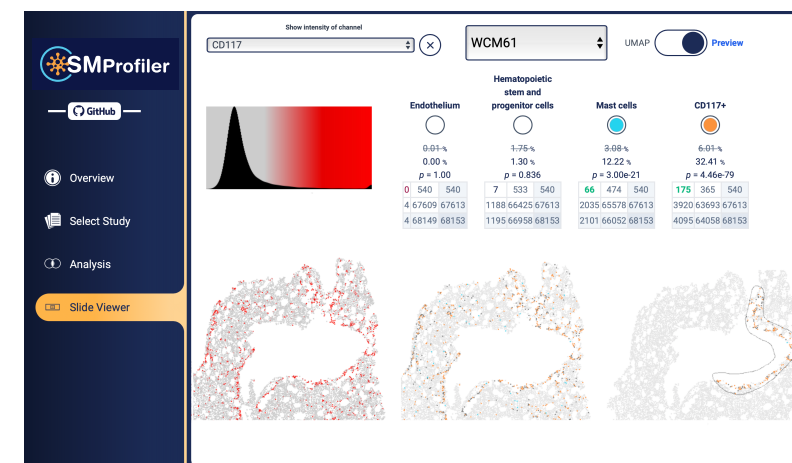
record significant results

Interpretation and results

identify salient results

refine statements

statistical assessment

determine significance for treatment/outcomes

publication

attribution with ORCID researcher profiles