



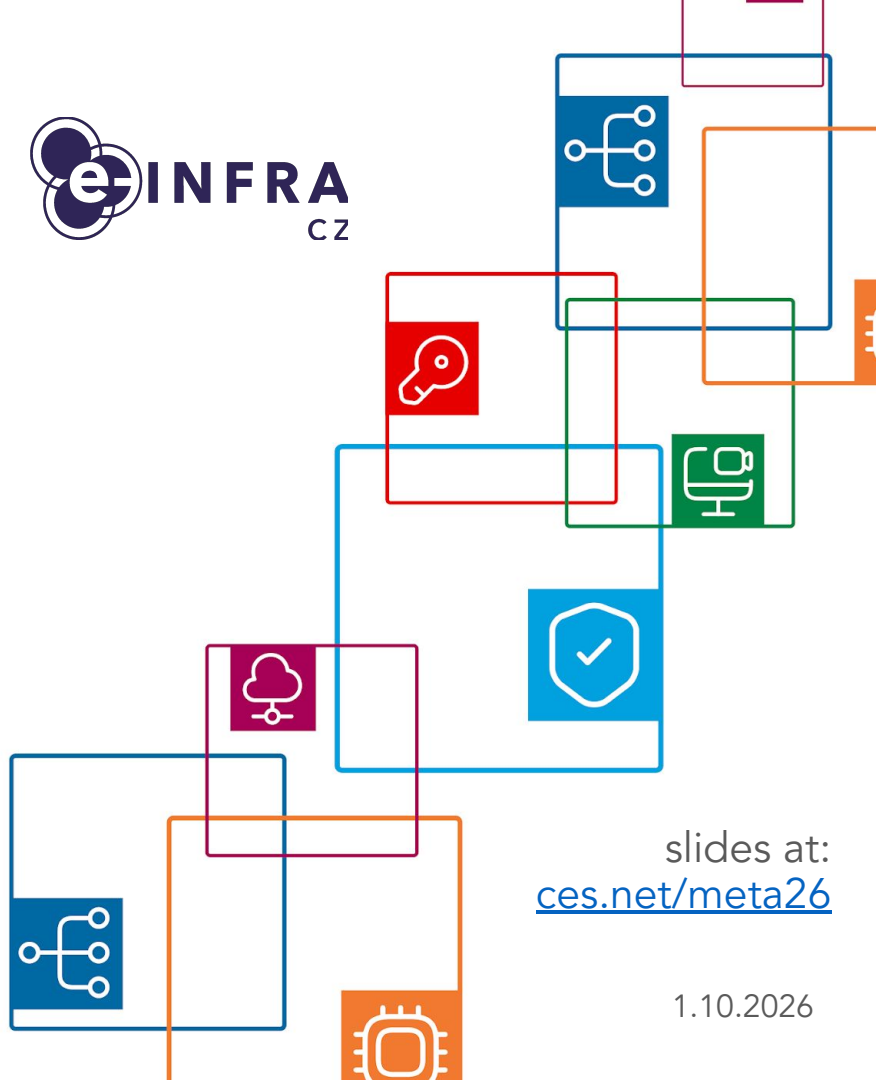
 **useGalaxy.cz**

aneb MetaCentrum v prohlížeči

Martin Čech

slides at:
ces.net/meta26

1.10.2026

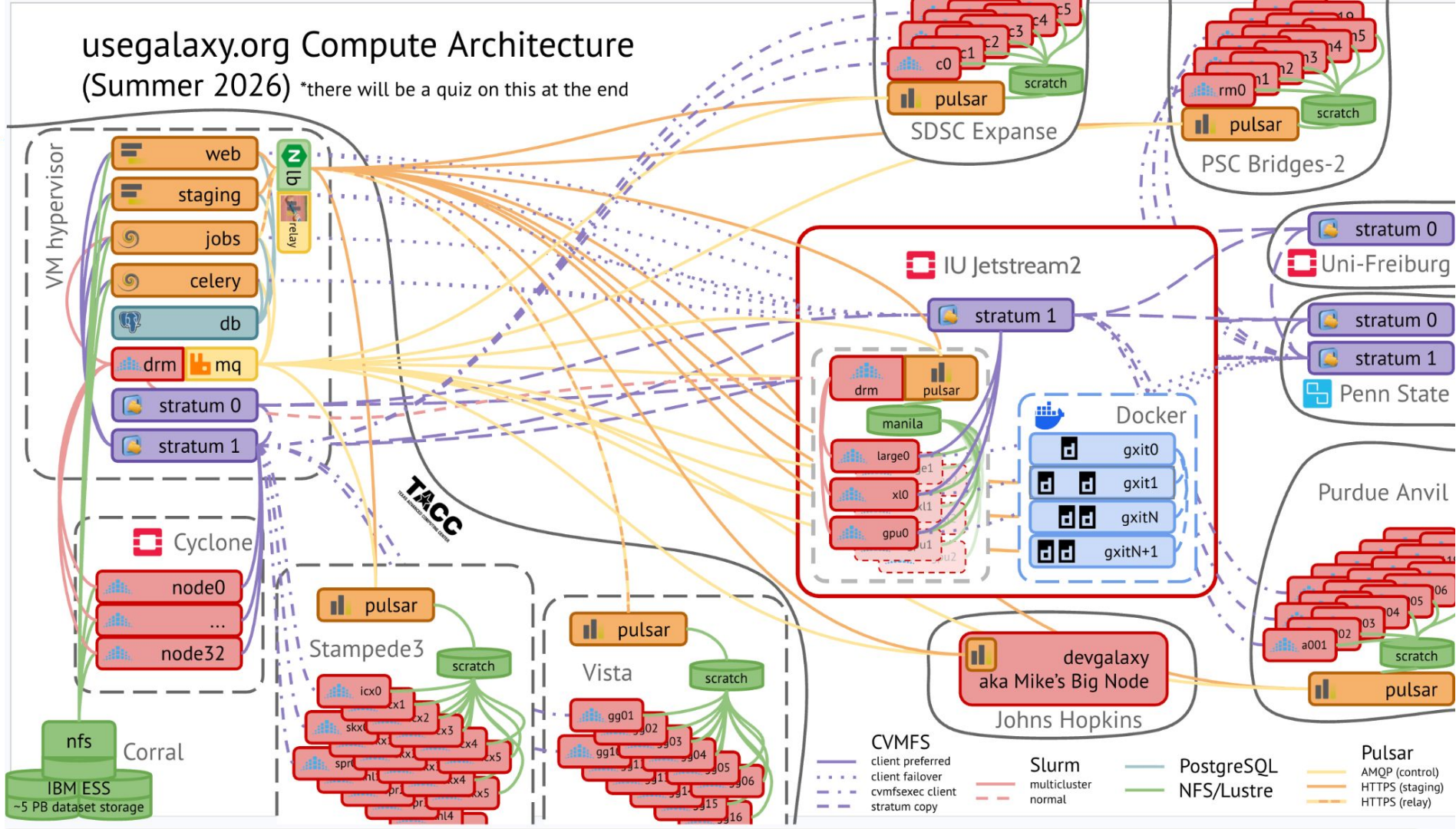


Co je Galaxy?

Jednoduchá aplikace v prohlížeči umožňující přístup ke komplexní výpočetní infrastruktuře a robustnímu software.

usegalaxy.org Compute Architecture

(Summer 2026) *there will be a quiz on this at the end



usegalaxy
(Summe

useGalaxy.cz

Galaxy

EDAM Topics

Discover 3k+ Tools

search tools

Tools

Upload

Beta Upload

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Recent Exports

GalaxyAI

Notifications

More

Admin

Systems biology

Ecology

Environmental science

Chemistry

Computational chemistry

Computational biology

Mapping

Probes and primers

Sequence composition, complexity and repeats

Sequence analysis

Sequence assembly

Nucleic acid sites, features and motifs

Metatranscriptomics

Structure prediction

Computer science

Data visualisation

Experimental design and studies

Informatics

Data architecture, analysis and design

Bioinformatics

Workflows

Laboratory techniques

Sequencing

RNA-Seq

Proteomics experiment

Mathematics

useGalaxy.cz

Welcome to the Czech Galaxy instance of E-Infra CZ and ELIXIR CZ hosted at MetaCentrum. We aim to enable accessible, reproducible, and transparent computational research in Czechia. We support thousands of documented and maintained tools that are free to use and backed by a robust public infrastructure.

Please visit our [Documentation](#) for details about using this service. In case of any further questions, please contact us at galaxy@cesnet.cz.

Have a look at one of our [trainings](#) or -- if you are logged in -- you can start with an [interactive tour](#):

[Galaxy UI](#) [Galaxy History](#) [Workflow Manager](#) [Deferred Datasets](#)

Galaxy is an open platform for supporting data intensive research developed by its many contributors.

History

search datasets

DONE: megadetector

META26

1.24 GB

66: Predictions_recap_CSV

65: Predictions_CSV

64: Images with bounding boxes

a list with 30 jpg datasets

33: zviratka

a list with 30 jpg datasets

2: weights

1: config



Co-funded by
the European Union



National Czech
Programme



nfs

IBM ESS

~5 PB dataset storage

Corral

CVMFS

client preferred
client failover
cvmfsexec client
stratum copy

Slurm
multiclusterv
normal

PostgreSQL
NFS/Lustre

Pulsar

AMQP (control)
HTTPS (staging)
HTTPS (relay)

Co je Galaxy?

The screenshot displays the Galaxy web interface in a browser. The address bar shows the URL: `usegalaxy.cz/?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fclimate%2Fclds_essential_variability%2Fclds_essential_variability%2F0.2.0&...`. The Galaxy header includes a logo, a grid icon, a help icon, a storage status indicator ("Using 4% of 268.4 GB"), and a user profile "martincech".

The left sidebar contains navigation links: Activities, Tools (with a "Discover 3k+ Tools" button and a search bar), Upload, Beta Upload, Workflows, Workflow Invocations, Visualization, Histories, History Multiview, Recent Exports, GalaxyAI, Notifications, and Admin.

The main content area is titled "Copernicus Essential Climate Variables for assessing climate variability (Galaxy Version 0.2.0)". It features a "Tool Parameters" section with the following options:

- Variable(s)** - optional: A dropdown menu with selected items "Surface air temperature", "Surface air relative humidity", and "Precipitation".
- Select type of data**: A dropdown menu with "Climatology" selected.
- Select Climate reference period**: A dropdown menu with "1991-2020" selected.
- Select month(s)** - optional: A dropdown menu with selected items "July", "August", "September", "October", "November", and "December".
- Advanced Options**: A section with a "Resolution" dropdown set to "high".
- Additional Options**: A section with "Email notification" set to "No" and "Attempt to re-use jobs with identical parameters?" set to "No".
- Output Tags**: A section with "Add tags" and a description: "Enter tags to apply to the output datasets in your history (e.g., 'sample!'). Tags starting with '#' (e.g., '#sample!') will propagate to datasets derived from these outputs. Tags help you organize and search your history."

The bottom of the sidebar shows a "Chemistry" section with links to Computational chemistry, Computational biology, Mapping, Probes and primers, Sequence composition, complexity and repeats, Sequence analysis, and Sequence assembly.

Co je Galaxy?

Galaxy

Activities

Discover 3k+ Tools

search tools

Upload

Beta Upload

Workflows

Workflow Invocations

Visualization

Histories

History Multiview

Datasets

Recent Exports

GalaxyL1

Computational biology

Mapping

Probes and primers

Sequence composition, complexity and repeats

Sequence analysis

Sequence assembly

Nucleic acid sites, features and motifs

Metatranscriptomics

Structure prediction

Computer science

Data visualisation

Experimental design and studies

Informatics

Data architecture, analysis and design

Bioinformatics

Workflows

Laboratory techniques

Sequencing

RNA-Seq

Proteomics experiment

Mathematics

Pure mathematics

Statistics and probability

Medicine

Public health and epidemiology

Omics

Genomics

Comparative genomics

Functional genomics

Proteomics

Metabolomics

Physics

Uncategorized

13: Alphafold 2 on : PDB ranked 0

Preview Raw Visualize Details Edit

State Tree

Sequence of https://usegalaxy.cz/api/datasets/6c11a2c2d0746c2 Chain 1: Polymer 1 [...]

Model 1

Model 449 elements

Polymer 449 elements

Cartoon

Structure Tools

Structure

https://usegalaxy.cz/api/datasets/6...

Type Model

Nothing Focused

Measurements

+ Add

Quick Styles

Apply Representation

Default Cartoon Spacefill Surface

Apply Style

Default Illustrative

Components

Preset + Add

Polymer Cartoon

Export Models

Export Animation

Export Geometry

History

search datasets

alphafold

1.87 MB

15: UCSC Main on Human: knownGene (chr7:155,607,379-155,612,663)

14: Alphafold 2 on : Visualization

13: Alphafold 2 on : PDB ranked 0

Add Tags

449 atoms and 6 HET-atoms chain_ids: A format pdb, database ?

Templating HTML file... Validated FASTA sequence(s):

ATOM	1	N	ALA	A	1	15.824
ATOM	2	H <th>ALA</th> <th>A</th> <th>1</th> <th>16.096</th>	ALA	A	1	16.096
ATOM	3	H2 <th>ALA</th> <th>A</th> <th>1</th> <th>15.015</th>	ALA	A	1	15.015
ATOM	4	H3 <th>ALA</th> <th>A</th> <th>1</th> <th>16.593</th>	ALA	A	1	16.593
ATOM	5	CA <th>ALA</th> <th>A</th> <th>1</th> <th>15.518</th>	ALA	A	1	15.518

12: Alphafold 2 on : PDB ranked 1

11: Alphafold 2 on : PDB ranked 2

10: Alphafold 2 on : PDB ranked 3

9: Alphafold 2 on : PDB ranked 4

8: UCSC Main on Human: knownGene (chr7:155,607,379-155,612,663)

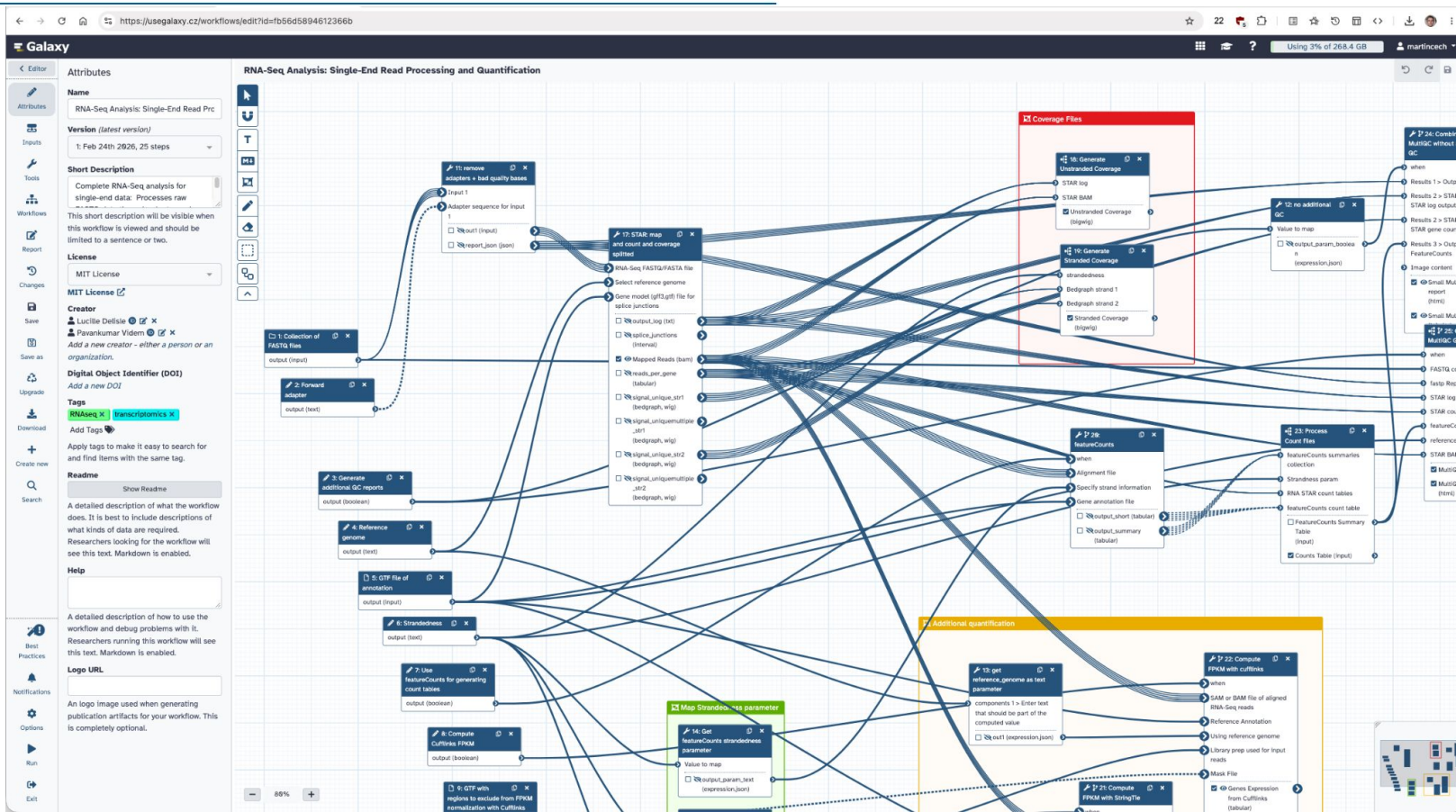
7: Pasted Entry

16:54:13 Created Polymer in 0ms.

16:54:13 Created Cartoon in 20ms.

16:54:13 Updated Structure Focus Representation in 1ms.

Co je Galaxy?



1. Galaxy je pouze pro práci s genomickými daty.
2. Galaxy není užitečná pro informatiky a lidí co používají nástroje napřímo.
3. Galaxy neškáluje pro použití na velkých nebo komplexních problémech.
4. Používat Galaxy je složité.
5. Hodí se jen k výuce, není užitečná pro opravdovou práci.
6. Nelze ji použít pro data vyžadující ochranu soukromí.
7. Galaxy je zdarma, proto má spoustu chyb a nemůže být dostatečně použitelná a spolehlivá.
8. Open source projekty jako Galaxy mají špatnou dokumentaci a nemají podporu, na kterou se můžou obrátit s vlastním problémem.
9. Investovat čas do nasazení Galaxy není rozumné ani udržitelné, protože za ní nestojí dostatečně silná organizace.
10. Nástroje a workflows v Galaxy jsou zastaralé.





W. A. Bacon et al., Ten Common Misconceptions About Galaxy (and Why They Are Wrong!)
<https://doi.org/10.1371/journal.pcbi.1013869>



Jaké Galaxy spravujeme?

- **useGalaxy.cz**
 - all purpose instance (and we'll support your tools)
 - aims to make the resources of National Grid Infrastructure more accessible
- **RepeatExplorer**
 - specialized instance for graph-based clustering and characterization of repetitive sequences in NGS
 - run for the Institute of Plant Molecular Biology in Ceske Budejovice
 - one of the ELIXIR CZ services
- **UMSA**
 - Untargeted Mass Spectrometry Analysis
 - operated on the behalf of MUNI RECETOX
 - publish tools, workflows, and other research results
 - provide pipeline as a service for collaborators

Galaxy a NRP

useGalaxy.cz

- Pilotní provoz integrace s repozitáři (test a catch-all)
- Import dat (včetně "deferred")
- Export dat (RO-Crate archivy s metadaty - historie, workflows, wf runs)

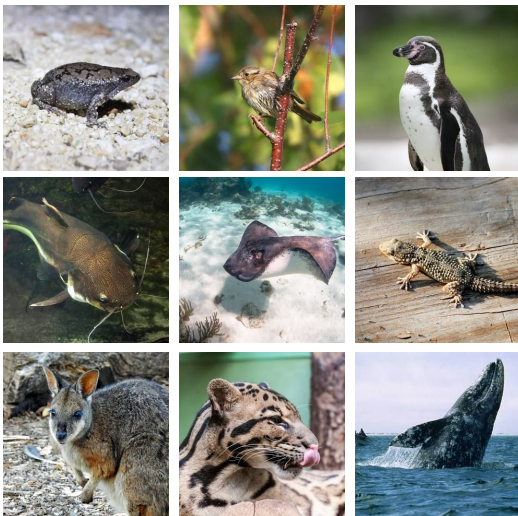
<https://docs.metacentrum.cz/en/docs/graphical/galaxy-nrp-integration>

The screenshot shows the Galaxy web interface at usegalaxy.cz. The main panel is titled 'Import Data Beta / Choose Remote Files'. It displays a list of remote file sources with columns for 'Name' and 'Details'. The sources are categorized by 'single-cell' and include various RNA sequencing and disease-related datasets.

Name	Details
Single-cell RNA sequencing of cerebral organoid-glioblastoma co-cultivation models	No details available
Single-cell RNA sequencing atlas of Alzheimer's disease cerebral organoid development	No details available
Single-cell RNA sequencing of cerebral organoid-glioblastoma co-cultivation models	No details available
Single-cell RNA sequencing atlas of Alzheimer's disease cerebral organoid development	No details available
Single-cell RNA sequencing of isogenic cerebral organoid models of early Alzheimer's disease development	No details available
Single-cell RNA sequencing of isogenic cerebral organoid models of early Alzheimer's disease development	No details available
Raw image data for article "Role of casein kinase 1 (CK1) in the amoeboid migration of B cell leukemic and lymphoma cells"	No details available



A PROJECT OF THE G10K CONSORTIUM



The Plan:

- Publish genome assemblies of 12 species every week
- Scale until all the 77,321 extant vertebrate genomes are known
- Maybe a decade?

Galaxy provides:

- Integration of the Genome Ark on public Galaxy servers.
- A Galaxy platform with toolkits tailored for assembly
- Workflows using the most up-to-date VGP pipelines.
- A list of publicly-available histories for each assembly completed on Galaxy as they are generated.
- Extensive training materials used to spread the load and accelerate

Using Galaxy helps make the assemblies reproducible and enables scaling.

<https://galaxyproject.org/projects/vgp>



Vertebrate genome assembly using HiFi, Bionano and Hi-C data - Step by Step

Authors: Delphine Larivière Alex Ostrovsky Cristóbal Gallardo Anna Syme Linelle Abueg Brandon Pickett Giulio Formenti Marcella Sozzoni

Overview

Questions:

- What combination of tools can produce the highest quality assembly of vertebrate genomes?
- How can we evaluate the quality of the assembly in a reference-free way?

Objectives:

- Learn the tools necessary to perform a de novo assembly of a vertebrate genome
- Evaluate the quality of the assembly

Requirements:

- [Introduction to Galaxy Analyses](#)
- [Slides: Quality Control](#)
- [Hands-on: Quality Control](#)

Time estimation: 5 hours

Level: Intermediate

Supporting Materials:

[Datasets](#) [Workflows](#) [FAQs](#) [Recordings](#)

Published: Jun 4, 2021

Last modification: Sep 27, 2024

License: Tutorial Content is licensed under [Creative Commons Attribution 4.0 International License](#). The GTN

PURL: <https://gxy.io/GTN:T000039>

Rating: 5.0 (1 recent ratings, 3 all time)

Revision: 54

bit.ly/vgp-training

Galaxy VGP



Workflow

Visualize

Data

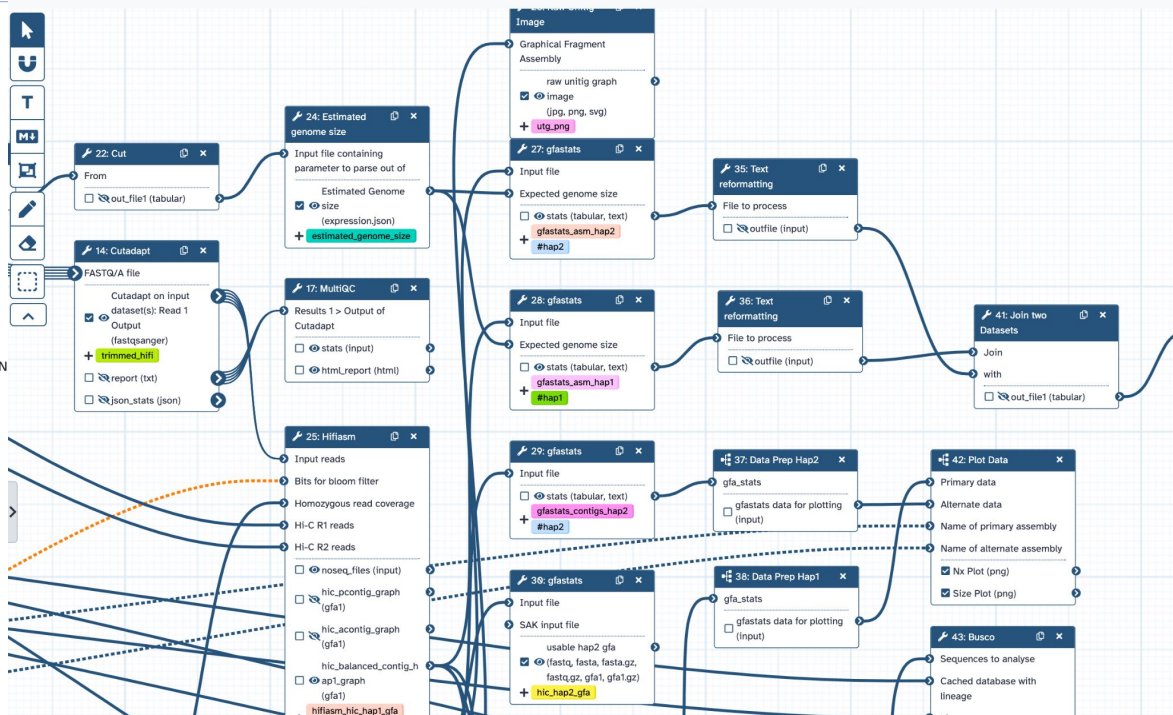
Admin

Help

User



Copy of Assembly-Hifi-HiC-phasing-VGP4 (release v0.2.2) shared by user iwc



Intrahost allelic variant SARS-CoV-2 analysis for COG-UK (2021/09/01 - 2022/02/28) samples

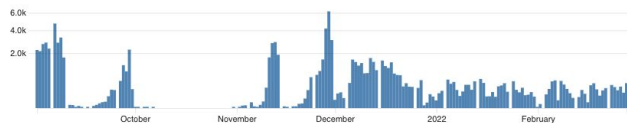
Select the data set to view

COG-UK (2021/09/01 - 2022/02/28) ▼

Results computed using the global platform for SARS-CoV-2 analysis

Poisson modeling estimated the global error rate of 0.00129149/base; 0 allelic variants (AV) occurring in fewer than 0 samples are deemed to be unreliable and have been removed.

All result processing takes place in the browser. You can select a date range to restrict the analysis to using the widget below. Current date range : 2021/09/01 - 2022/02/28



Recompute for 2021/09/01 : 2022/02/28

93550 samples analyzed	62586 unique allelic variants (AV)	63.44 AV per sample
49.42 fixed AV per sample	13.39 (0.00:64.00) low frequency AV per sample (95% range)	9.64 : 4 mean Syn : mean

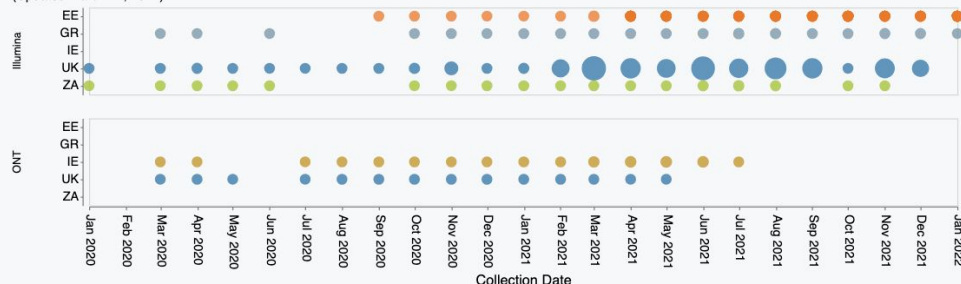
<https://galaxyproject.org/projects/covid19/>

Galaxy provided:

- Continuous analysis of within-host sequence variants in high quality public read-level datasets.
- Maintenance of curated workflows for the analysis of SARS-CoV-2 sequence data and free powerful infrastructure to execute them.
- Providing access to all results in raw and aggregated form for immediate use.

Timeline for 314,941 samples

(Updated March 17, 2022)



(když wi-fi bude přát)

44: LLM Hub (kimi) on dataset 6 error

Details Edit Error

Dataset Error Report

An error occurred while running the tool `toolshed.g2.bx.psu.edu/repos/bgruening/llm_hub/llm_hub/2.53.0+galaxy0`.

Possible Causes

We can use AI to analyze the issue and suggest possible fixes. Please note that the diagnosis may not always be accurate.

Let our Help Wizard Figure it out!

Details

Execution resulted in the following messages:

Fatal error: Exit code 1

Tool generated the following standard error:

```
Traceback (most recent call last):
  File "/srv/galaxy/var/shed_tools/toolshed.g2.bx.psu.edu/repos/bgruening/llm_hub/3834a8bfaf2/llm_hub/llm_hub.py", line 217,
in <module>
    answer, finish_reason = stream_completion()
    ^^^^^^^^^^^^^^^^^^^^^
  File "/srv/galaxy/var/shed_tools/toolshed.g2.bx.psu.edu/repos/bgruening/llm_hub/3834a8bfaf2/llm_hub/llm_hub.py", line 199,
in stream_completion
    with client.chat.completions.create(**api_params) as stream:
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/srv/galaxy/venv/lib/python3.11/site-packages/openai/_utils/_utils.py", line 298, in wrapper
    return func(*args, **kwargs)
    ^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/srv/galaxy/venv/lib/python3.11/site-packages/openai/resources/chat/completions/completions.py", line 1215, in create
    return self._post(
    ^^^^^^^^^^^^^
  File "/srv/galaxy/venv/lib/python3.11/site-packages/openai/_base_client.py", line 1332, in post
    return cast(ResponseT, self.request(cast_to, opts, stream=stream, stream_cls=stream_cls))
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/srv/galaxy/venv/lib/python3.11/site-packages/openai/_base_client.py", line 1105, in request
    raise self._make_status_error_from_response(err_response) from None
openai.BadRequestError: Error code: 400 - {'error': {'message': 'litellm.ContextWindowExceededError: litellm.BadRequestError: ContextWindowExceededError: Hosted_vllmException - ({'object': 'error', 'message': 'The input (1787719 tokens) is longer than the model\'s context length (1048576 tokens).', 'type': 'BadRequestError', 'param': null, 'code': 400})\nmodel=kimi. context_window_fallbacks=None. fallbacks=None.\n\nSet \'context_window_fallback\' - https://docs.litellm.ai/docs/routing#fallbacks. Received Model Group=kimi\n\nAvailable Model Group Fallbacks=None', 'type': None, 'param': None, 'code': '400'}}}
```

<https://usegalaxy.cz/datasets/0a12c5655c72b0d5/details>
<https://usegalaxy.cz/datasets/dca7ea37df1b4591/details>

44: LLM Hub (kimi) on dataset 6 error

Details Edit Error

Dataset Error Report

An error occurred while running the tool `toolshed.g2.bx.psu.edu/repos/bgruening/llm_hub/llm_hub/2.53.0+galaxy0`.

Possible Causes

We can use AI to analyze the issue and suggest possible fixes. Please note that the diagnosis may not always be accurate.

Likely cause: The LLM request sent to the kimi model contained 17,877,719 tokens, which exceeds the model's maximum context length of 1,048,576 tokens, so LiteLLM/VLLM rejected the request with HTTP 400.

Solution:

Reduce the prompt/input size before calling the model: 1) inspect the input dataset/file and remove or truncate unnecessary content; 2) if the tool has a max lines/characters/tokens option, lower it to keep the request well under 1,048,576 tokens; 3) split the input into smaller chunks (at least ~18 chunks for 17.8M tokens, preferably more to leave room for system prompt and output) and process each chunk separately; 4) alternatively select a model with a larger context window if available, but the current input is ~17x the limit so chunking or truncation is required.

Was this answer helpful?



This feedback helps us improve our responses.

Galaxy a LLMs

LLM Hub Call any LLM
(Galaxy Version 2.53.0+galaxy0)

Tool Parameters

Choose the model

Multimodal models

Multimodal models are capable to have image and text as input.

Model *

kimi: Latest Kimi, strong coding performance

Select the model you want to use.

Text Context - optional



UCSC Main on Human: knownGene (genom

accepted formats

switch to column select

Image Context - optional



Nothing selected

accepted formats

Prompt *

what type of data is in the column 4 of the provided dataset?

Prompts or tasks you want the LLM to perform.

Advanced Options

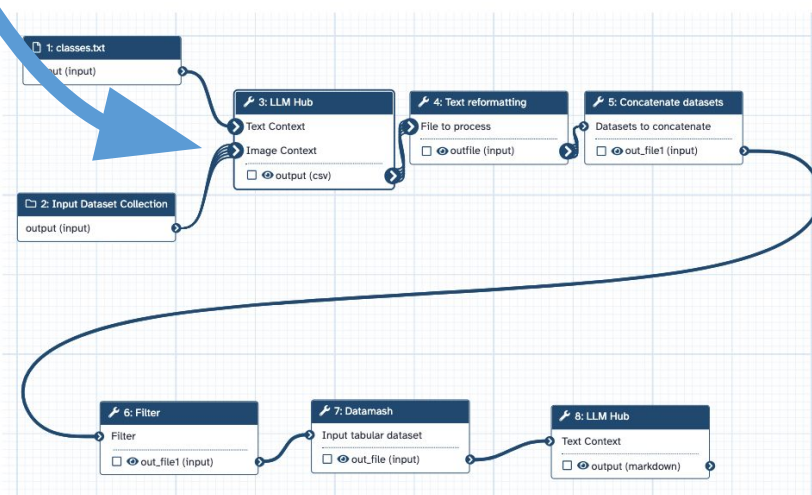
Temperature - optional

Controls randomness in the output. Lower values make output more focused and deterministic, while higher values make it more creative and random. If not set, the model's default temperature will be used.

Autor: Aleš Křenek

Blogpost: <https://blog.e-infra.cz/blog/galaxy-llm/>

Workflow Preview



Download Import Run

About This Workflow

Implementation of 'Galaxy LLM e-Infra CZ blog - Version 7

Author

ljocha_ics



All published Workflows by ljocha_ics

Description

This Workflow has no description.

Tags

License

No License specified

Last Updated

Wednesday Mar 4th 16:28:59 2026 GMT+1

Sharing

Use the following link to share preview of this workflow:
<https://usegalaxy.cz/published/workflow?id=ff5944efdb2c4f18>
Manage sharing settings here.

Galaxy MCP Server

- Model Context Protocol server for interacting with the Galaxy platform.
- Enables AI assistants and other clients to connect to Galaxy instances, search and execute tools, manage workflows, and access other features of the ecosystem.
- “API for Agents”, ships with Galaxy

Galaxy agentic plugins

- Claude Code, Codex, Gemini, whatever.

Orbit: Galaxy's own agent

- API key for Galaxy + API key for LLM
- <https://galaxyproject.org/agents/orbit/>

FILES

data3

genomes48

alignments312

AlignmentViewer2.016

foldseek_work4

fs_clusters24

ortho15

phylo_q126

phylo_q227

mm_tmp1

trees30

fs001_3di_trim...88 KB

fs001_3di.fa277 KB

fs001_aa_trim...88 KB

fs001_aa.fa277 KB

fs001_combin...169 KB

fs001_db112 KB

fs001_db_h5.3 KB

fs001_db_h.dbtype4 B

fs001_db_h.in...3.8 KB

fs001_db_ss112 KB

fs001_db_ss.dbty...4 B

fs001_db_ss.l...4.5 KB

fs001_db.dbtype4 B

fs001_db.index4.5 KB

fs001_db.look...6.9 KB

fs001_db.source33 B

fs001_subclu...420 KB

fs001_subclust...35 KB

fs001_subclu...117 KB

fs001.nw5.6 KB

plot_and_test.R3.4 KB

q2_3di_tree.pdf40 KB

q2_aa_tree.pdf39 KB

q2_part_tree.pdf40 KB

CHAT

yes, but show formal plan

PLAN DRAFT — AWAITING YOUR APPROVAL

Plan E: Potential Orthologs of FS_001 Members between PvW1 and PvPAM [local]

Identify potential 1-to-1 orthologs of FS_001 VIR genes between the two long-read assemblies (PvW1, Thailand; PvPAM, Peruvian Amazon) using reciprocal best hits, microsynteny preservation, and consistency with the FS_001 phylogeny from Plan C Q2. PvP01 is excluded because its Illumina assembly fragments subtelomeric VIRs across many small contigs, defeating microsynteny analysis.

Steps

■ 1. Extract FS_001 protein sequences for PvW1 and PvPAM {#plan-e-step-1} — Filter the combined VIR FASTA to FS_001 members from these two genomes and write per-genome FASTAs.

◦ Routing: local

◦ Tool: python / seqkit

◦ Verification: 349 PvW1 sequences + 448 PvPAM sequences extracted; both per-genome FASTAs non-empty.

■ 2. Reciprocal best hits with MMseqs2 {#plan-e-step-2} — Run mmseqs easy-rbh PvW1 vs PvPAM with mutual coverage 0.5 and E-value 1e-5. Output is the set of RBH pairs along with their bitscore, alignment length, and sequence identity.

◦ Routing: local

◦ Tool: mmseqs easy-rbh

◦ Verification: RBH table contains pairs with all required columns; pair count is sensible (expect 100-300 pairs).

■ 3. Build flanking-gene context tables {#plan-e-step-3} — Parse the full PlasmoDB proteome FASTAs for PvW1 and PvPAM (already on disk). For every FS_001 RBH gene, record the 2 nearest non-VIR genes upstream and downstream on the same chromosome.

◦ Routing: local

◦ Tool: python (HMM-filter to identify non-VIR genes; sort by chromosome+position)

◦ Verification: per-RBH-gene flanking table with 4 neighbor IDs each; gene IDs resolve in the proteome FASTAs.

■ 4. Build housekeeping-gene reciprocal-best-hit table {#plan-e-step-4} — Run mmseqs easy-rbh on the HMM

Ask for a plan, drop a file path, or type / for commands

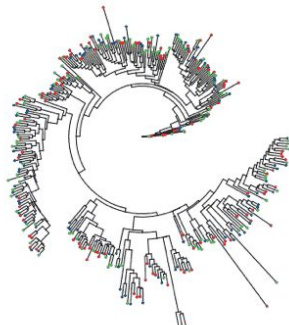
Type / to see available commands

NotebookActivityFile x

data/genomes/alignments/phylo_q2/q2_aa_tree.pdf39 KB

≡ R 1 / 1 — 90% + ↺ ↻ ↶ ↷ ↸ ↹ ↻ ↶ ↷ ↸ ↹

Q2: FS_001 AA tree (n=338)



genome • PvP01 • PvPAM • PvW1

ReadyGaaxyReportLocalCloudGPT-5.50 tok<\$0.01Users/anton/.loom/analyses/phist2/

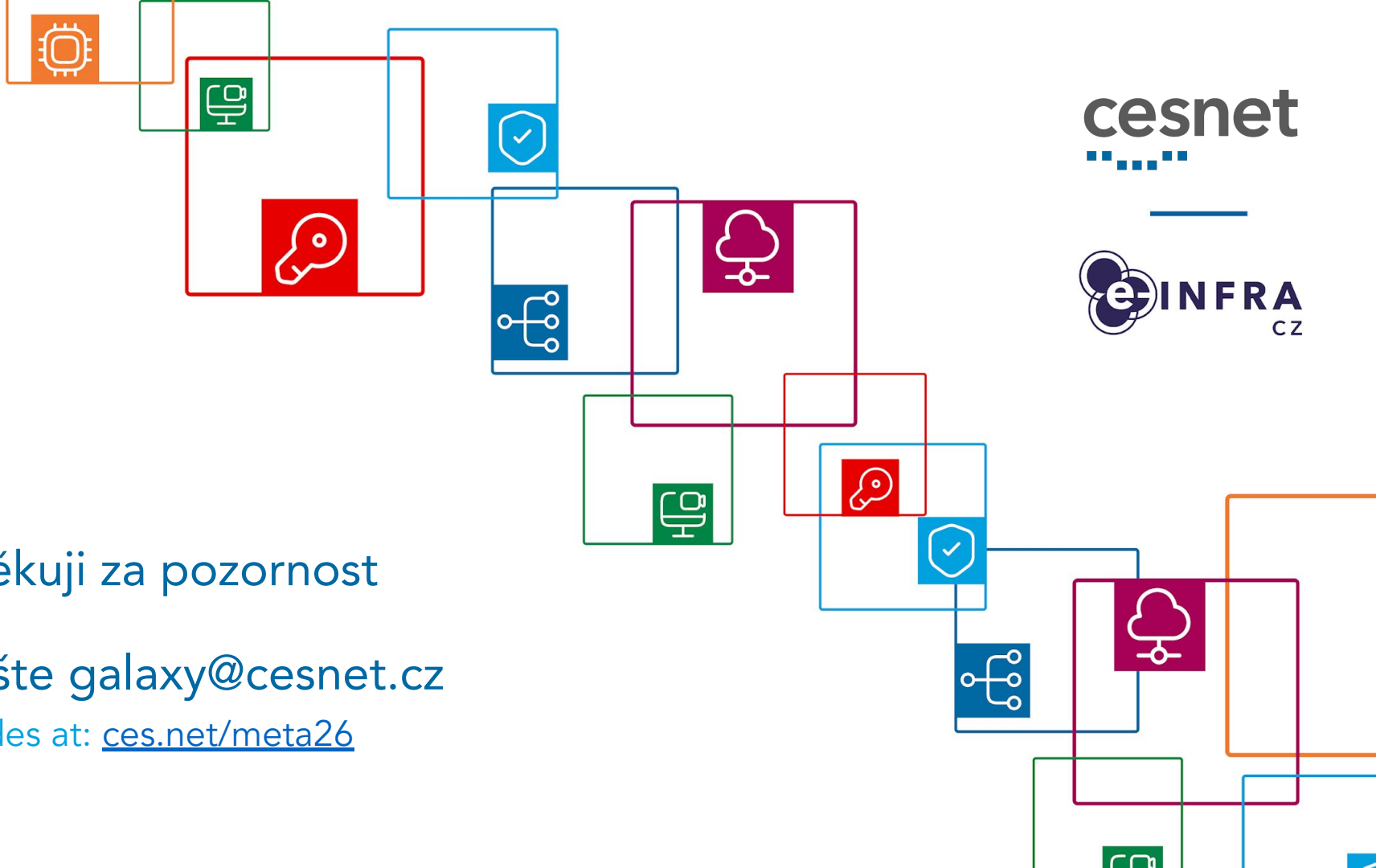
A free service running on usegalaxy.cz that provides users with dedicated resources to ensure educational exercises can be completed promptly.

Includes dashboard to monitor student usage.

<https://usegalaxy.cz/tiaas>

<https://training.galaxyproject.org/training-material/topics/instructors/tutorials/setup-tiaas-for-training/tutorial.html>





cesnet
.....

eINFRA
CZ

Děkuji za pozornost

Pište galaxy@cesnet.cz

slides at: ces.net/meta26