

5.9.2024 Technical Computing Camp 2024

OpenScience s prostředím MATLAB



Michal Blaho

blaho@humusoft.sk

www.humusoft.cz

info@humusoft.cz

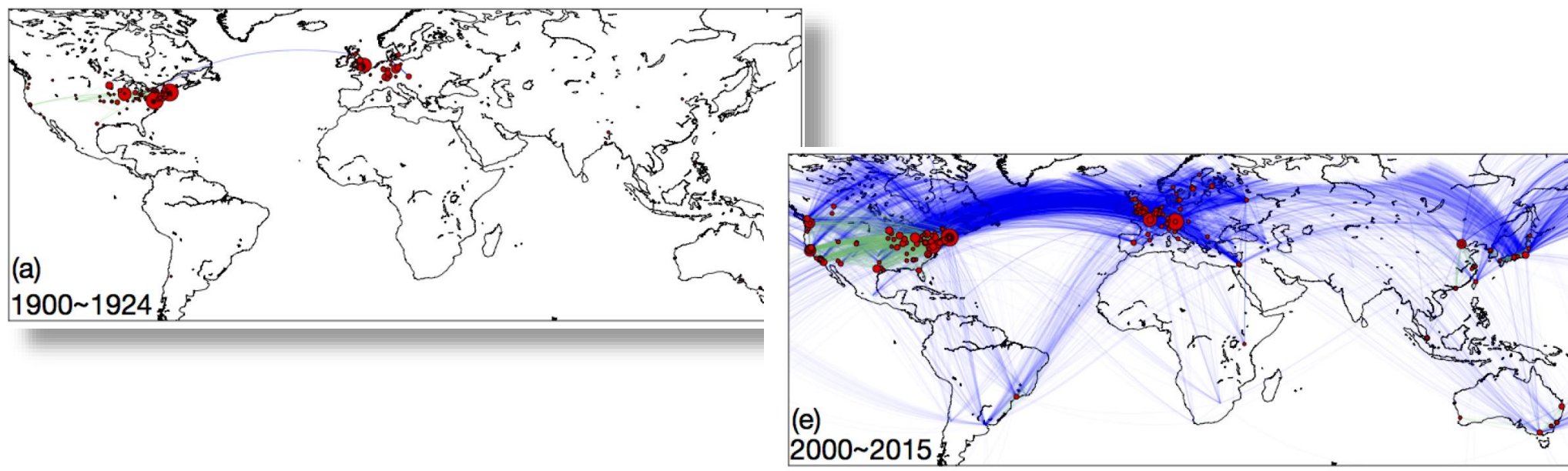
www.mathworks.com

Jan Daněk

danek@humusoft.cz

Výzvy vedeckého výskumu

Research is a collaborative effort – more than ever today



A Century of Science: Globalization of Scientific Collaborations, Citations, and Innovations

Yuxiao Dong, Hao Ma, Zhihong Shen, Kuansan Wang, Microsoft Research. <https://doi.org/10.1145/3097983.3098016>

Open Science

[Products](#)[Solutions](#)[Academia](#)[Support](#)[Community](#)[Events](#)

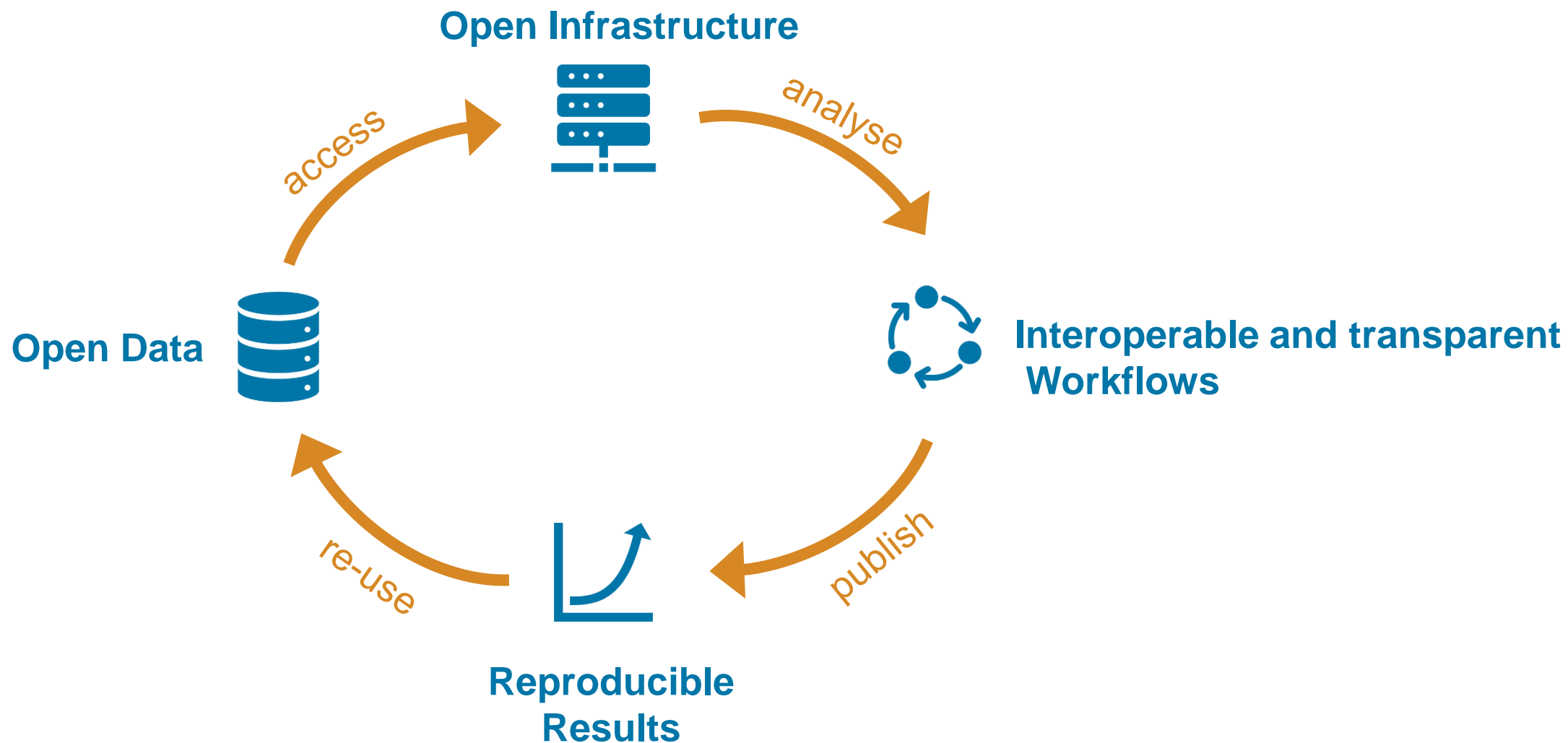
What Is Open Science?

What Is Open Science?

Open science is the practice of making scientific methodologies and output (such as publications, data, and software) transparent and broadly accessible. Open science maximizes the reuse of available data and code and enables scientists to build on the work of their peers.

Researchers using MATLAB® and Simulink® can practice open science to increase the scientific impact of their work.

Spoločné úsilie mnohých zúčastnených strán



Spoločné úsilie mnohých zúčastnených strán



Využitie údajov o pohybe zvierat na predpovede a plánovanie



for animal tracking data

Temp,
Wind,
Vegetation

Climate
Change

Roads,
Settlements

Urban
planning

Animal
Behavior,
Env. data

Behavior

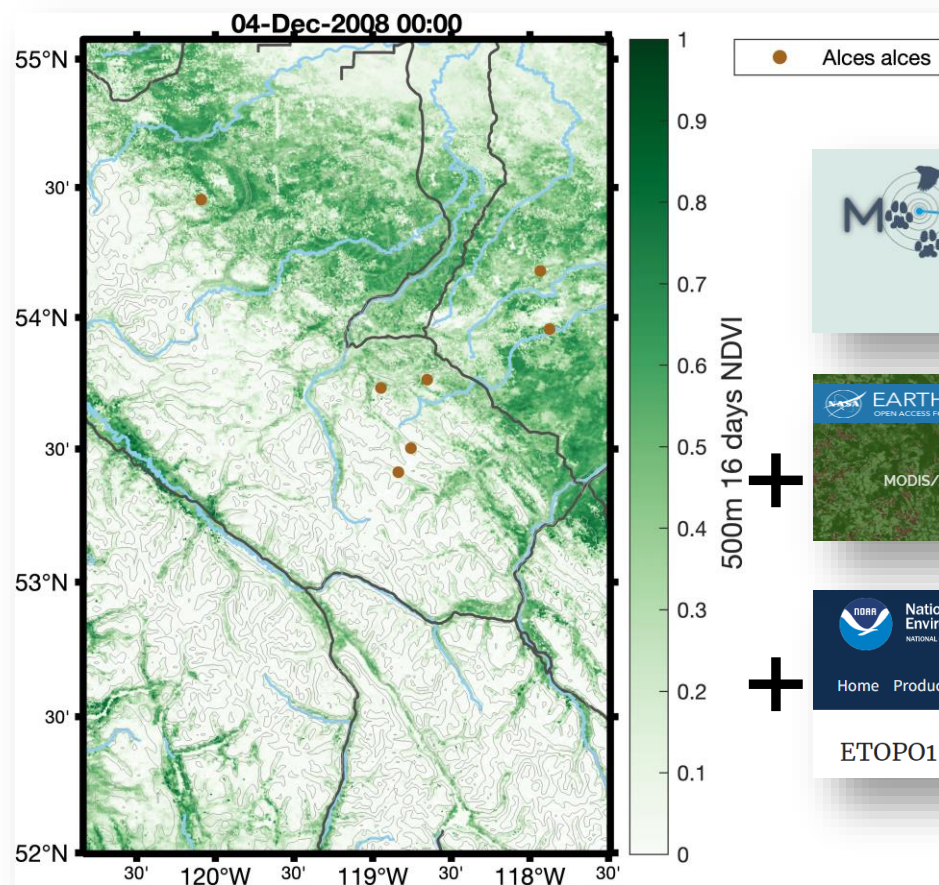
Policy
makers



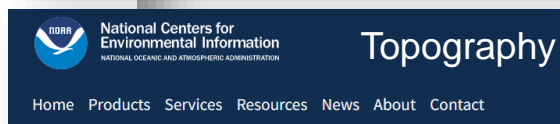
Prof. Gil Bohrer
Ohio State

Image: CC-BY-SA 3.0 License. [Wikimedia](#)

ECODATA Animate: vizualizácia údajov o pohybe zvierat



App: download and install for all + open code on GitHub



ETOPO1 1 Arc-Minute Global Relief Model

Welcome to ECODATA-Animate's documentation!

[ECODATA - Animate](#)

ECODATA-Animate

ECODATA-Animate is a MATLAB® program for creating customized animated maps of animal movements. The program creates image frames that can be animated using the [ECODATA-Prepare Movie Maker App](#). Define track visualization options and include additional layers from raster files, shapefiles, an elevation model and label lists. See [ECODATA-Prepare](#) for additional tools for preparing input data. Development is supported by MathWorks® and the NASA Earth Science Division, Ecological Forecasting Program, as part of the [Room to Roam: Y2Y Wildlife Movements](#) project.

Check out the documentation! [App documentation](#) [Developer guide](#)

Note: ECODATA-Animate is in the early stages of development, and any feedback is welcome. We have any suggestions or feature requests, encounter any bugs, or come across places where the documentation is unclear, please [submit a GitHub issue](#).



"I think MathWorks really got it right. Whatever we make can be compiled into an app with an executable. That app can be shared and does not require you to be a licensed user. For those who might want to work on the code further, we will publish the MATLAB code base." – Gil Bohrer, PI, Ecodata/Prof. Ohio State University



Brain Observatory Toolbox: rozhranie pre verejné dáta

Allen Brain Observatory - Visual Coding AWS Public Data Set

electrophysiology image processing imaging life sciences Mus musculus neurobiology neuroimaging signal processing

Description

The Allen Brain Observatory – Visual Coding is a large-scale, standardized survey of physiological activity across the mouse visual cortex, hippocampus, and thalamus. It includes datasets collected with both two-photon imaging and Neuropixels probes, two complementary techniques for measuring the activity of neurons in vivo. The two-photon imaging dataset features visually evoked calcium responses from GCaMP6-expressing neurons in a range of cortical layers, visual areas, and Cre lines. The Neuropixels dataset features spiking activity from distributed cortical and subcortical brain regions, collected under analogous conditions to the two-photon imaging experiments. We hope that experimentalists and modelers will use these comprehensive, open datasets as a testbed for theories of visual information processing.

Update Frequency

Annually

License

<http://www.alleninstitute.org/>

Documentation

Open in MATLAB Online  File Exchange

Brain Observatory Toolbox




A MATLAB toolbox for accessing and using the neural recording public datasets from the **Allen Brain Observatory**¹.

■ Get oriented and get started with **3 lines of code**. You can:

- **Open in MATLAB Online**
- Enter `>>bot . README` on your own local/cloud installation

Either will orient you to several **live script examples** available to guide new users, including **demos** of neural data analysis

You can also individually view (👁️) or run (▶️) these examples on MATLAB Online:

Example Type	Data Type	View	Run	Data Type	View	Run
 Quickstart	Calcium Imaging (Ophys)	👁️	▶️	Neuropixels Probe (Ephys)	👁️	▶️
 Demo	Calcium Imaging (Ophys)	👁️	▶️	Neuropixels Probe (Ephys)	👁️	(*)
 Tutorial	Calcium Imaging (Ophys)	👁️	▶️	Neuropixels Probe (Ephys)	👁️	(*)

Resources on AWS

Description

Project data files in a public bucket

Resource type

S3 Bucket

Amazon Resource Name (ARN)

`arn:aws:s3:::allen-brain-observatory`

AWS Region

`us-west-2`

AWS CLI Access (No AWS account required)

- Open Source
- Cloud workflow – no downloads
- Uses the AWS Open Data Registry

Obtain Session of Interest

From the table of filtered items, one or more specific items can be obtained for further examination using the `bot.get()` functions:

```
sessionUnderStudy = bot.getSessions(filteredSessions(1,:));
```

Downloading file: [https://allen-brain-observatory.s3.us-west-2.amazonaws.com/visual-coding-2p/ophys_experiment_data/503412730 to cache location: /external/neuralcoding/prod6/specimen_495727000/ophys_experiment_503412730/503412730.nwb...

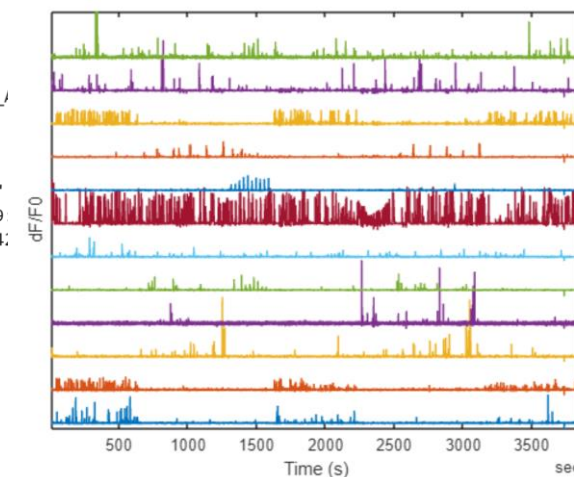
sessionUnderStudy.info

ans = struct with fields:

```

    id: 503412730
  experiment_container_id: 511510695
        stimulus_name: three_session_
targeted_structure_acronym: VISal
    fail_eye_tracking: 0
    imaging_depth: 175
        cre_line: "Cux2-CreERT2"
  date_of_acquisition: 23-Feb-2016 19
                name: "20160223_2224:
specimen_id: 495727000
  experiment_container: [1x1 struct]
        specimen: [1x1 struct]
targeted_structure: [1x1 struct]
    well_known_files: [3x1 struct]

```



Dátové formáty podporované MATLABom

Scientific data formats

NetCDF, HDF, FITS, and CDF formats

NetCDF Files

Network Common Data Form

HDF5 Files

Hierarchical Data Format, Version 5

HDF4 Files

Hierarchical Data Format, Version 4

FITS Files

Flexible Image Transport System

Band-Interleaved Files

Band-interleaved data

Common Data Format

CDF files

Reading in Data into MATLAB

Discipline specific data formats



Neuroscience

- BIDS (`bids.`)
- NIFTI (`niftiread`)
- NWB (`matnwb`)



Biology

- EDF (`edfread`, `edfinfo`)
- DICOM (`dicomread`)
- FASTA (`fastaread`)
- SAM (`samread`)



Earth, Ocean, Atmosphere

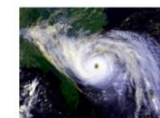
- IRIS (`irisfetch.m`)
- NetCDF (`netcdf.open`)
- HDF5 (`h5read`)

Community-driven data standards

NetCDF/GRIB reader

version 1.1.0.0 (5.57 MB) by Klaus Wyser

A collection of functions to read GRIB and NetCDF files



readHurdat2

version 1.1.1 (2.35 KB) by Joshua Port

Reads in all data for a specified storm from the NHC HURDAT2 file



bids-matlab

version 0.1.0 (2.42 MB) by Remi Gau

MATLAB / Octave tools for BIDS datasets

<https://bids-matlab.readthedocs.io>



Add-Ons

File Exchange

Access data on the cloud

```
setenv("AWS_ACCESS_KEY_ID",id)
setenv("AWS_SECRET_ACCESS_KEY",secid)
setenv("AWS_REGION","us-east-1")
fileLoc = "s3://mw-s3-tests-us/datasets/FoodPic.jpg";
```



```
setenv("HADOOP_HOME",hadoopPath)
fileLoc = "hdfs://hadoop01g1nxa64:123/datasets/FoodPic.jpg";
```



```
img = imread(fileLoc);
```

Spoločné úsilie mnohých zúčastnených strán

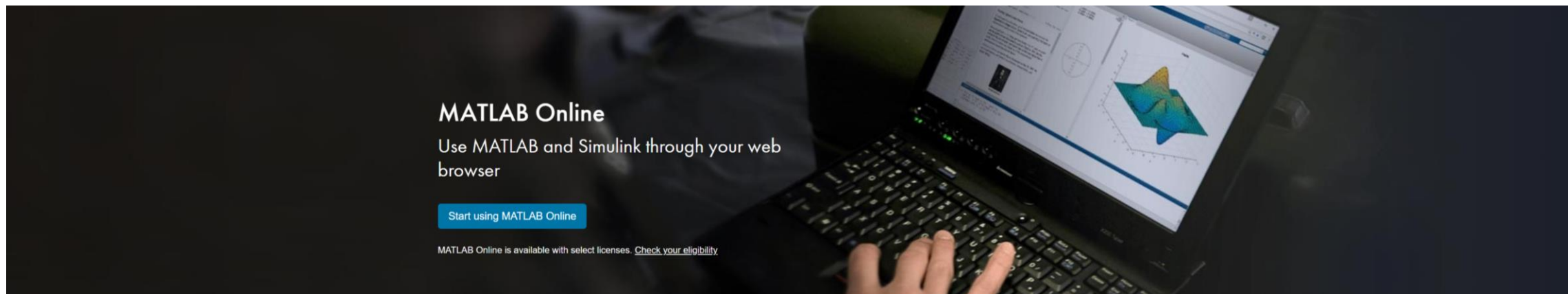
Open Infrastructure

- Diverse data formats
- Public open data sets
- Big Data

Open Data



MATLAB & Simulink Online + MATLAB Drive

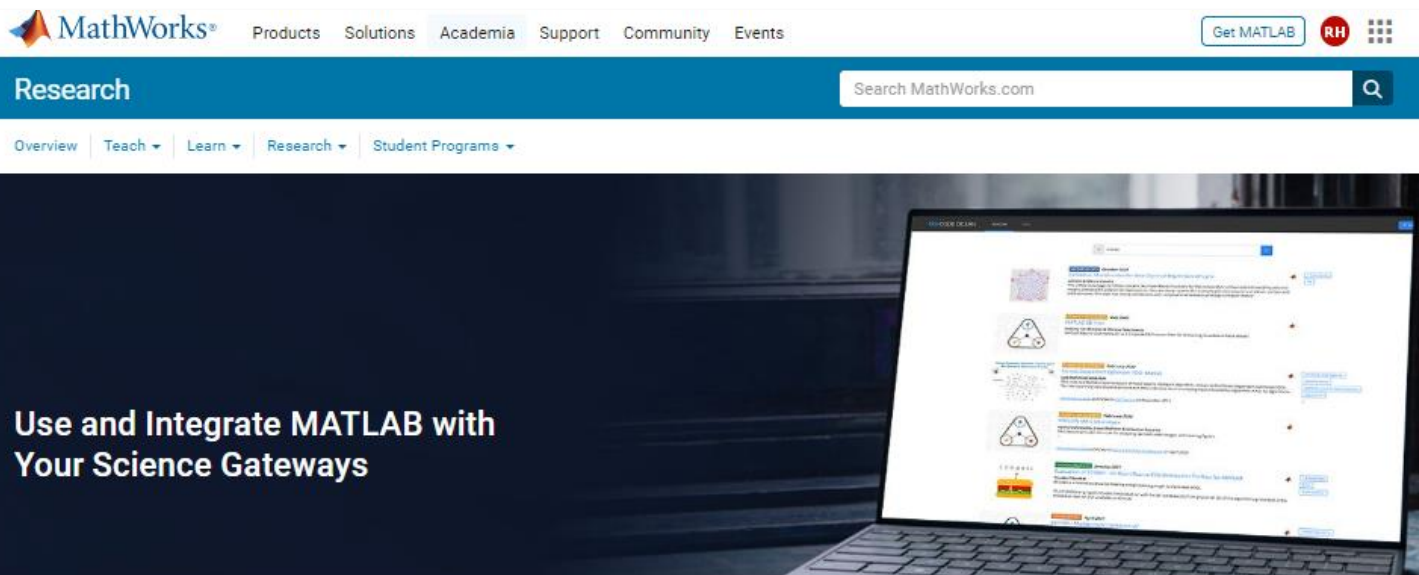


- No installation
- Collaborate over your browser
- Synchronize with cloud data (MATLAB Drive)
- Host your own MATLAB Online

<https://www.mathworks.com/products/matlab-online-server.html>

<https://www.mathworks.com/products/matlab-online.html>

MATLAB a Scientific Gateways



Science gateways are online portals where researchers access shared resources. They focus on community sharing for a specific academic discipline, such as neuroscience; HPC and research centers hosting compute; and curriculum tools that support teaching. In addition to targeting the physical sciences, many science gateways serve researchers and educators in engineering, math, and the social sciences.

The number and uses of science gateways have increased recently in response to a push for Open Science and transparent research, data and code sharing, and sustainable platforms and projects. As part of this trend, science gateway developers are incorporating MATLAB and related community-contributed tools based on MATLAB into their online portals.

Learn about Open Science with MATLAB

<https://mathworks.com/academia/research/science-gateways.html>

Examples of Science Gateways with MATLAB

Allen Brain Observatory

The Allen Brain Observatory provides publicly available neuroscience data.

- » Access data using the Brain Observatory Toolbox for MATLAB

BioLib

BioLib is a portal for creating and running biology-related software applications.

- » Build MATLAB tools on BioLib

bio.tools

bio.tools aggregates tools for biological sciences.

- » Explore MATLAB tools on bio.tools

CDS

Copernicus's Climate Data Store hosts climate data.

- » Visualize and process CDS climate data with MATLAB

Code Ocean

Code Ocean enables researchers to upload and share code associated with their published research.

- » Learn how to upload and share reusable MATLAB code on Code Ocean

Compute Canada

Compute Canada provides computing resources to researchers across Canada.

- » Use MATLAB on Compute Canada resources

CUAHSI/HydroShare

HydroShare is CUAHSI's online collaboration environment for sharing data, models, and code.

- » Use MATLAB Online on CUAHSI to run HydroShare resources

EGI

The European Grid Infrastructure provides advanced computing and data access for research.

- » Analyze data on the EGI with MATLAB

nanoHUB

nanoHUB provides nanotechnology tools for teaching and research.

- » Find nanotechnology tools for use with MATLAB

SERC

The Science Education Resource Center offers educator-contributed curricula and teaching best practices.

- » Explore SERC resources for teaching computation with MATLAB

SGCI

The Science Gateways Community Institute provides resources for developing and supporting science gateways.

- » Partner with MathWorks to support MATLAB on your science gateway

SNIC

The Swedish National Infrastructure for Computing is an HPC consortium that provides hardware and software access to Swedish researchers.

- » Use MATLAB on SNIC resources

Whole Tale

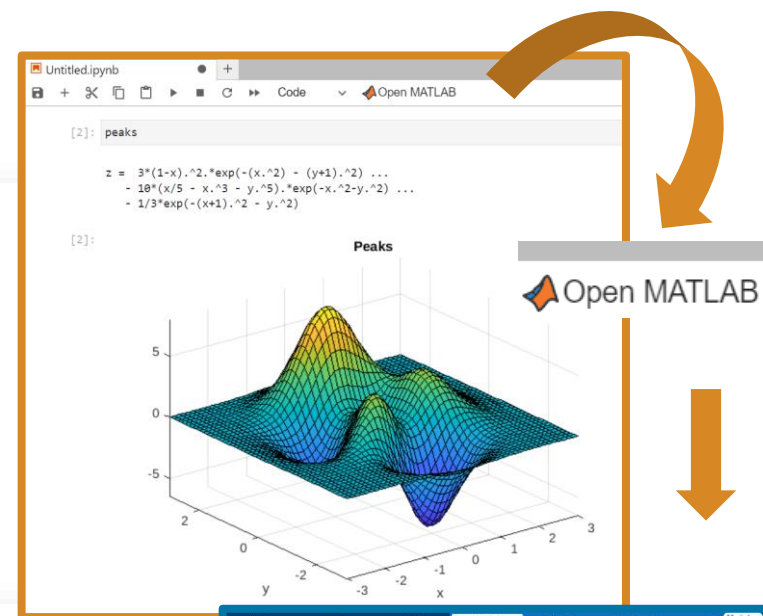
Whole Tale enables the creation, publication, and execution of tales, or executable research objects.

- » Create a tale based on MATLAB using Whole Tale

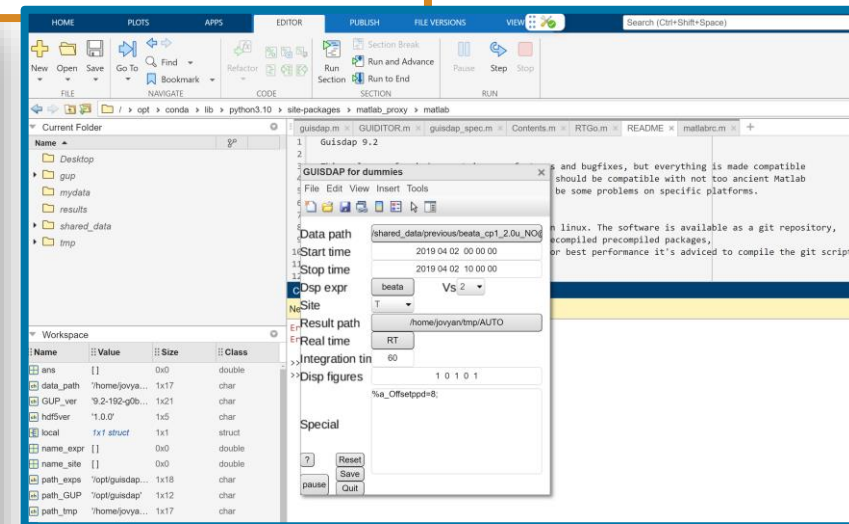
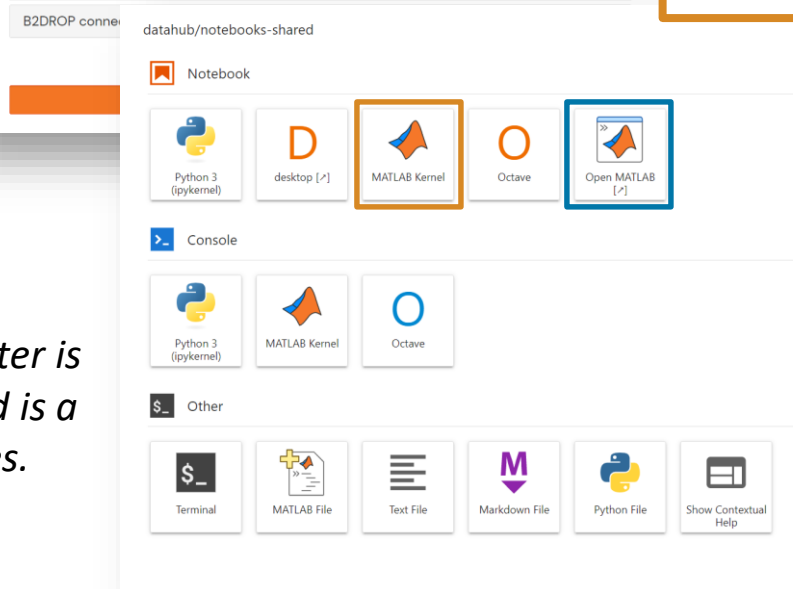
MATLAB kernel v JupyterHub



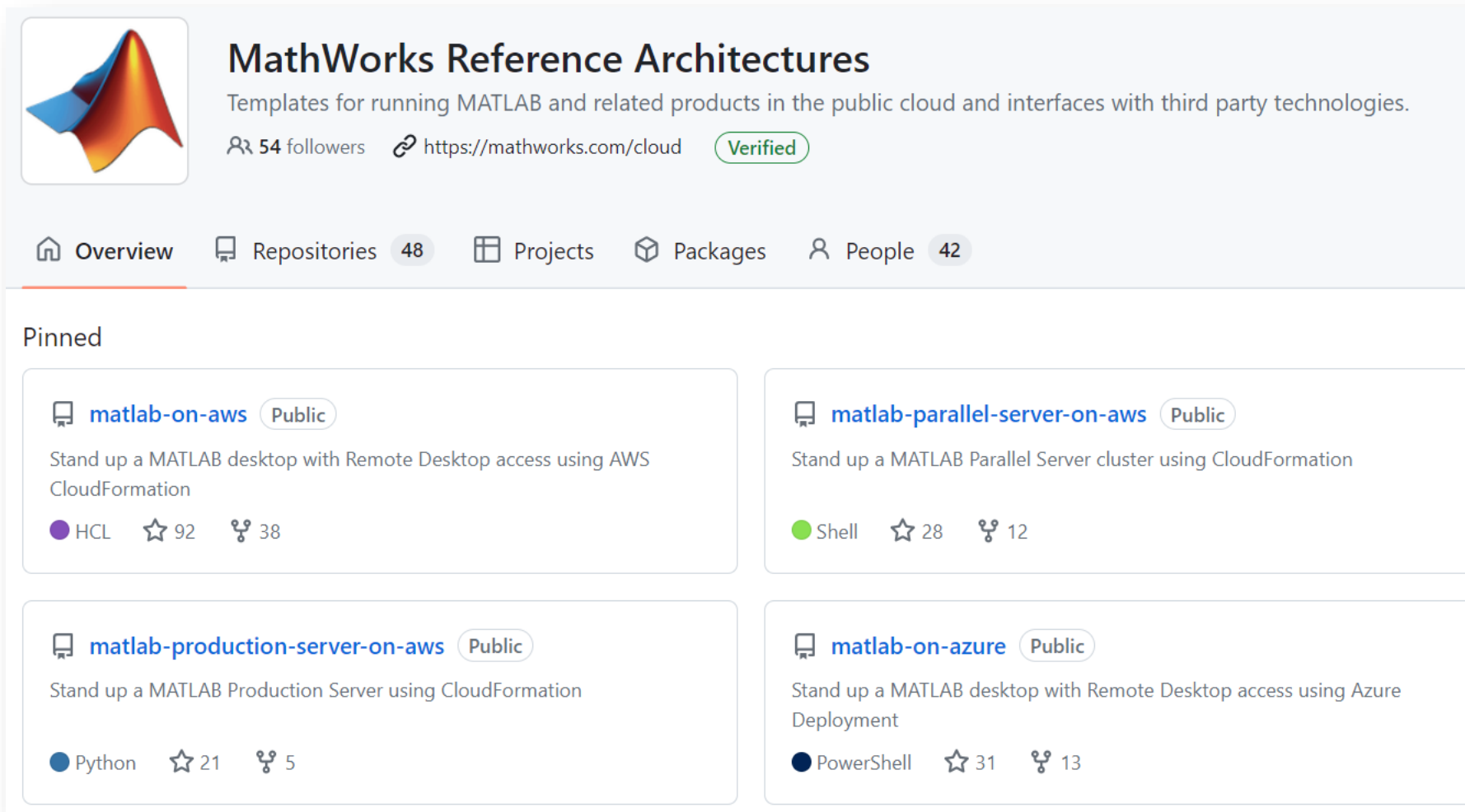
- ### Server Options
- Default EGI environment – 6 GB RAM / 2 core
The Default notebook environment includes Python, R, Julia and Octave kernels
 - MATLAB Environment (Basic) – 4GB RAM / 4 cores
The MATLAB environment (requires a valid license), includes Python and MATLAB kernels
 - MATLAB Environment (Full) – 4GB RAM / 4 cores
The MATLAB environment (requires a valid license), includes Python, MATLAB kernels and additional MATLAB packages
 - EISCAT environment – 4 GB RAM / 2 cores
The EISCAT environment.



The EISCAT integration of Matlab into Jupyter is the perfect tool for new (and old) users and is a window to Open Science and FAIR principles.
- Ingemar Häggström, Head of Operations



Vlastná integrácia



MathWorks Reference Architectures
Templates for running MATLAB and related products in the public cloud and interfaces with third party technologies.
54 followers <https://mathworks.com/cloud> Verified

Overview Repositories 48 Projects Packages People 42

Pinned

- matlab-on-aws** Public
Stand up a MATLAB desktop with Remote Desktop access using AWS CloudFormation
HCL ☆ 92 🍴 38
- matlab-parallel-server-on-aws** Public
Stand up a MATLAB Parallel Server cluster using CloudFormation
Shell ☆ 28 🍴 12
- matlab-production-server-on-aws** Public
Stand up a MATLAB Production Server using CloudFormation
Python ☆ 21 🍴 5
- matlab-on-azure** Public
Stand up a MATLAB desktop with Remote Desktop access using Azure Deployment
PowerShell ☆ 31 🍴 13

[MathWorks reference architectures](#)

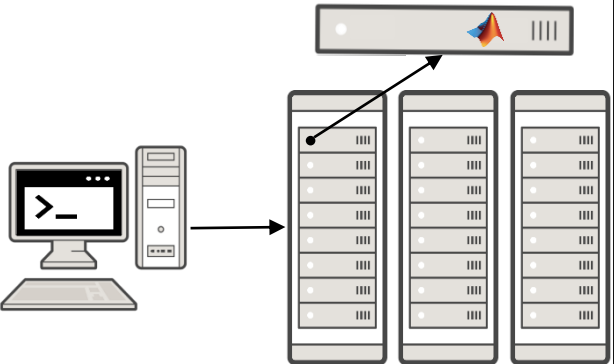
Práca v HPC infraštruktúre

More Similar Likeness to traditional HPC workflows*

Run MATLAB on the cluster via batch or interactively using the scheduler

Required Knowledge:

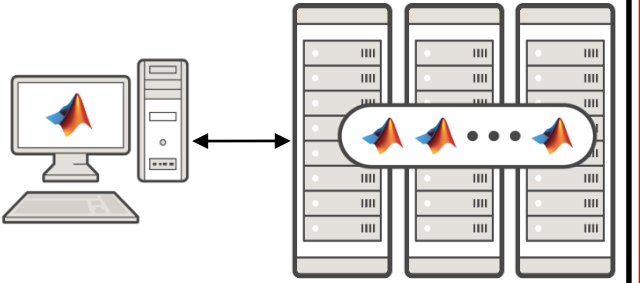
- MATLAB
- Shell
- Scheduler
- File transfer



Submit work from MATLAB client to cluster running MATLAB Parallel Server

Required User Knowledge:

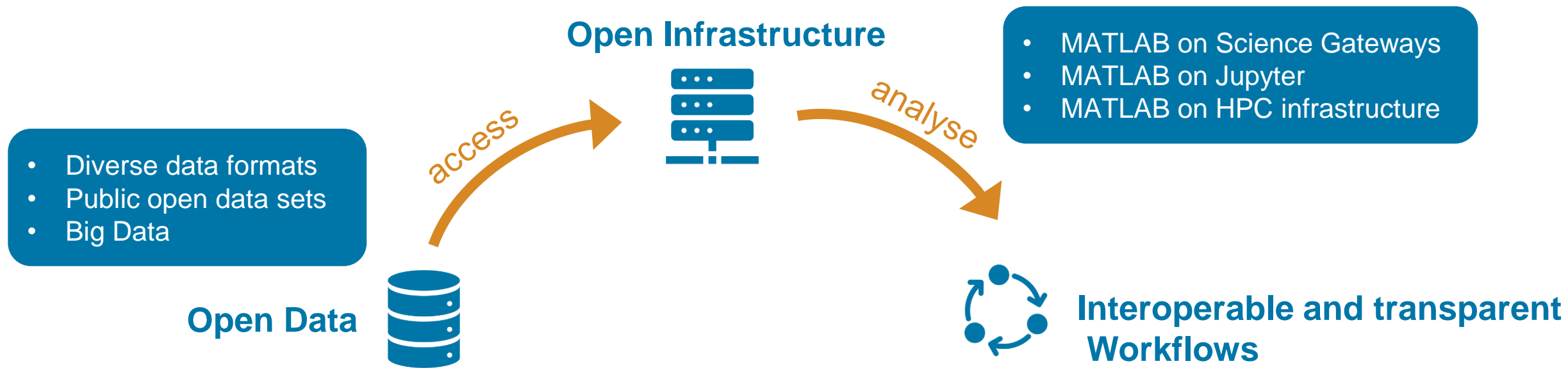
- MATLAB



Ease of use for MATLAB users Less Difficult

MATLAB Parallel Server

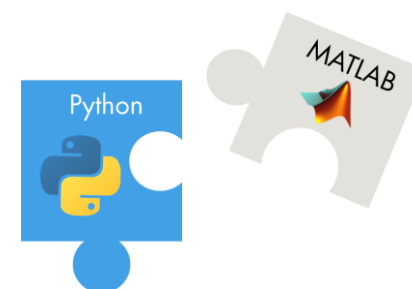
Spoločné úsilie mnohých zúčastnených strán



Práca v rôznych prostrediach

Why researchers often need more than one tool

- Use functionality of one in another
- Run code available in another language (hybrid workflows)
- Collaborate



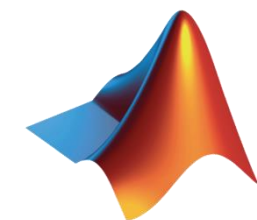
Call MATLAB from Python

```
$ pip install matlabengine
>>> import matlab.engine
>>> m = matlab.engine.start_matlab()
>>> x = m.sqrt(42)
```



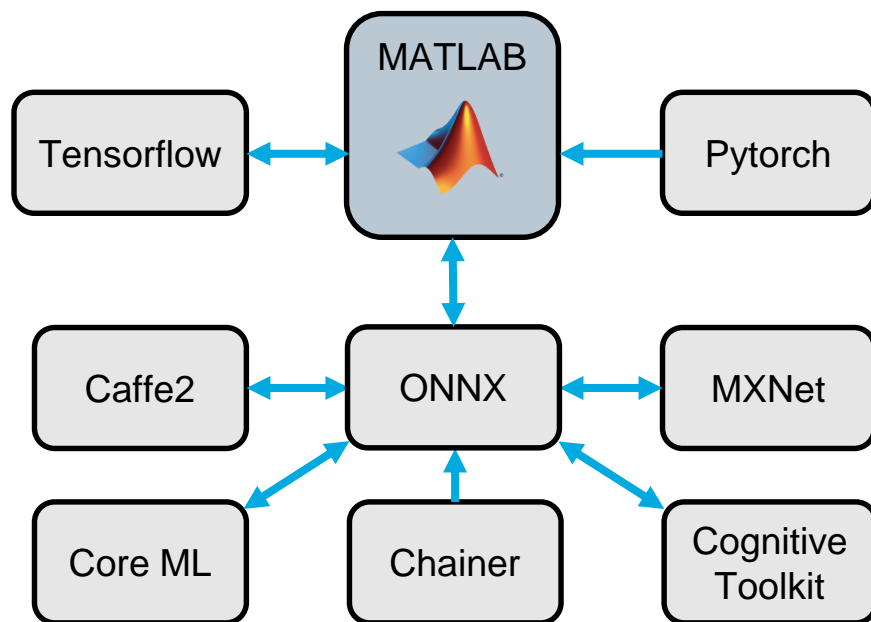
Call Python from MATLAB

```
>>> py.math.sqrt(42)
>>> py.importlib.import_module()
```



Práca v rôznych prostrediach

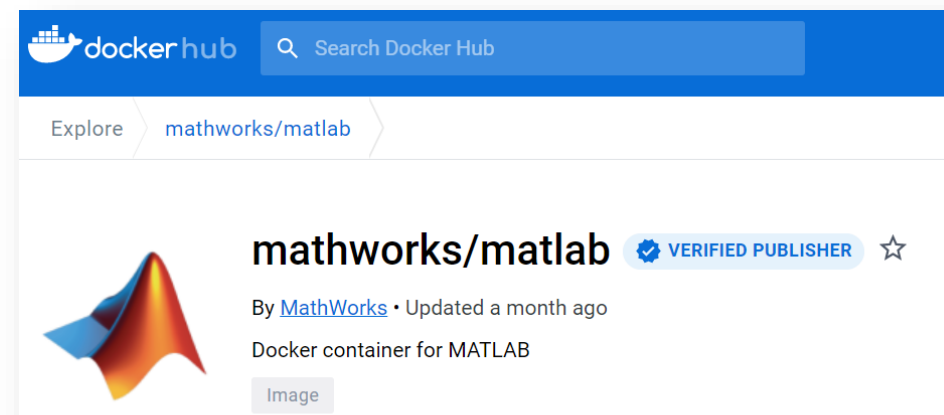
Import DL models from other frameworks



```
importTensorFlowNetwork
importNetworkFromPytorch
importONNXNetwork
```

[Interoperable with DL frameworks](#)

Pre-built containers on Docker-Hub

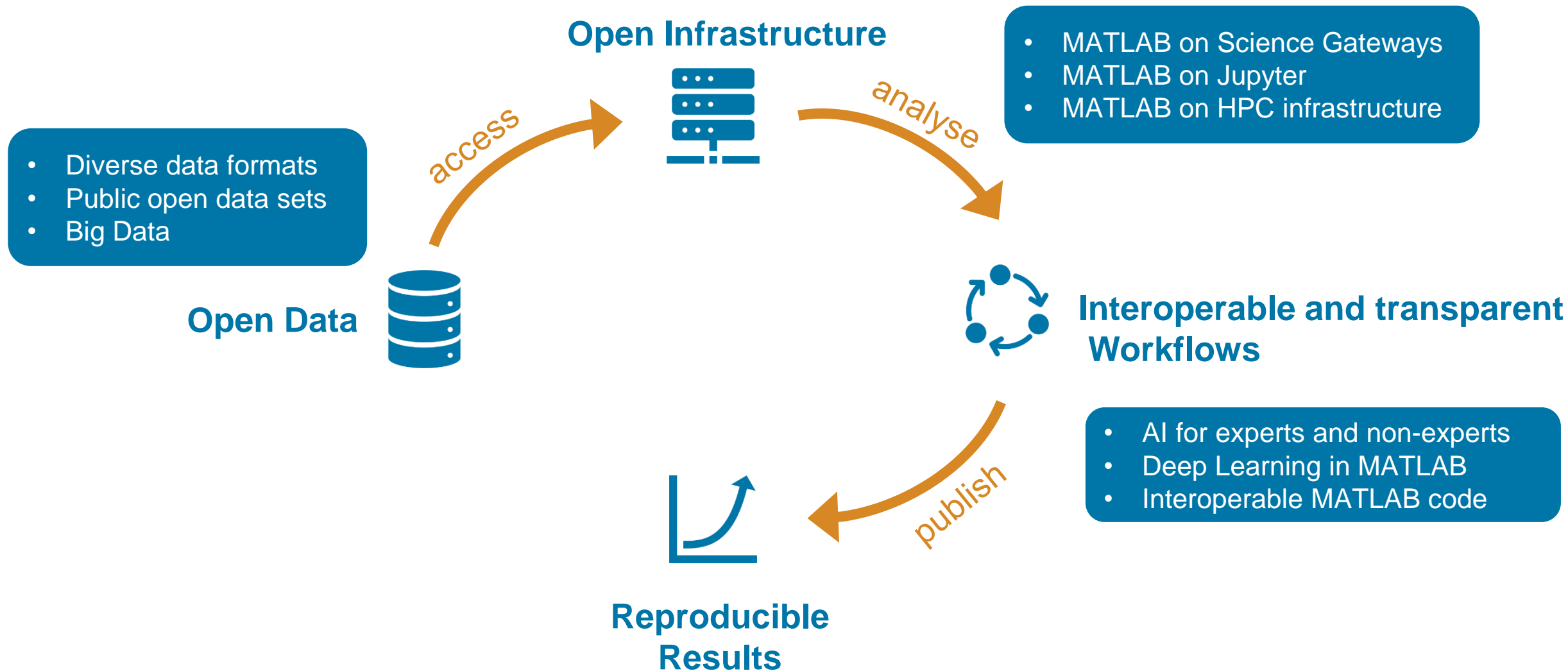


MATLAB Container on Docker Hub

Access MATLAB® on the cloud or in server environments by using the MATLAB container available on [Docker Hub](#). The MATLAB container provides a simple and flexible solution to run MATLAB in cloud environments such as AWS® or Microsoft® Azure®. For more information on containers, see [What is a Container?](#).

[MATLAB containers on DockerHub](#)

Spoločné úsilie mnohých zúčastnených strán



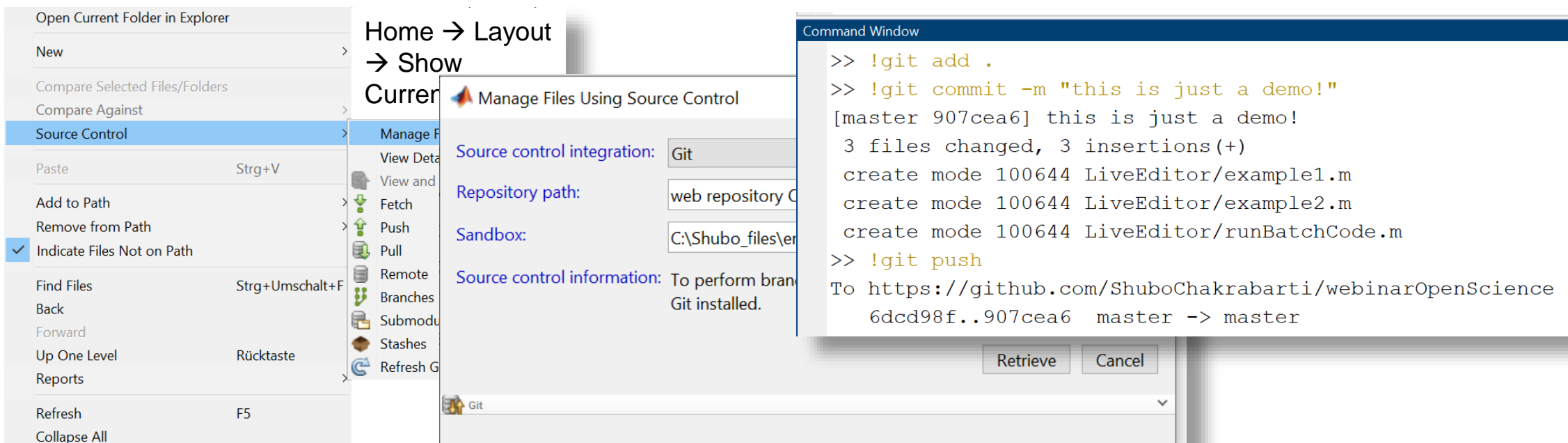
Efektívny vývoj softvéru – správa verzií

Source Control

- Keep track of changes
- Collaborate efficiently



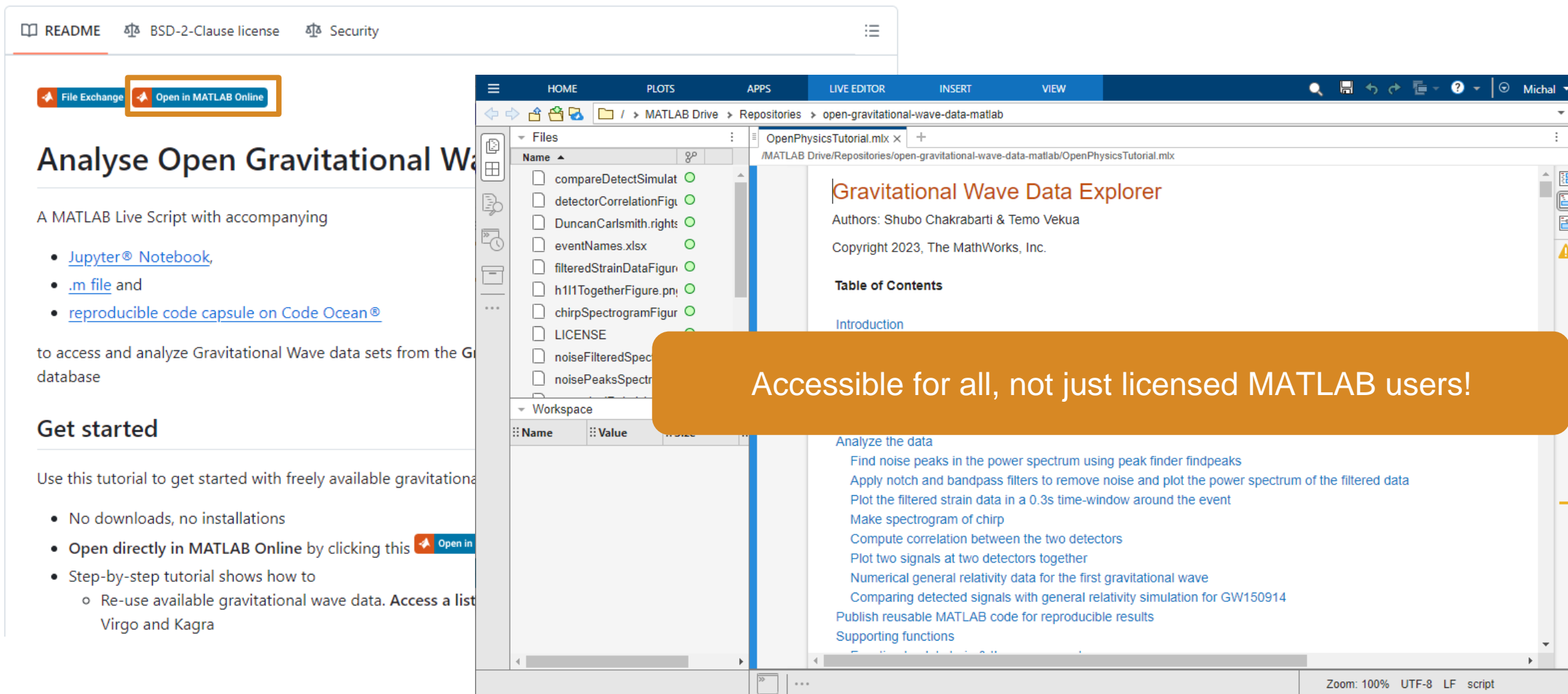
- Integrate git & svn with your MATLAB code...access from IDE & command line !git



The screenshot shows the MATLAB IDE interface. On the left, the 'Source Control' menu is open, showing options like 'Manage Files Using Source Control', 'View Details', 'Fetch', 'Push', 'Pull', 'Remote Branches', 'Submodules', 'Stashes', and 'Refresh'. The 'Manage Files Using Source Control' dialog box is open, showing 'Source control integration: Git', 'Repository path: web repository C...', and 'Sandbox: C:\Shubo_files\er...'. The 'Command Window' is also open, showing the following commands and output:

```
>> !git add .
>> !git commit -m "this is just a demo!"
[master 907cea6] this is just a demo!
3 files changed, 3 insertions(+)
create mode 100644 LiveEditor/example1.m
create mode 100644 LiveEditor/example2.m
create mode 100644 LiveEditor/runBatchCode.m
>> !git push
To https://github.com/ShuboChakrabarti/webinarOpenScience
6dcd98f..907cea6 master -> master
```


Zverejňovanie reprodukovateľných výsledkov – GitHub



File Exchange **Open in MATLAB Online**

Analyse Open Gravitational Wave

A MATLAB Live Script with accompanying

- [Jupyter® Notebook](#),
- [.m file](#) and
- [reproducible code capsule on Code Ocean®](#)

to access and analyze Gravitational Wave data sets from the G database

Get started

Use this tutorial to get started with freely available gravitational

- No downloads, no installations
- **Open directly in MATLAB Online** by clicking this **Open in**
- Step-by-step tutorial shows how to
 - Re-use available gravitational wave data. **Access a list** Virgo and Kagra

Gravitational Wave Data Explorer

Authors: Shubo Chakrabarti & Temo Vekua
Copyright 2023, The MathWorks, Inc.

Table of Contents

- Introduction
- Analyze the data
 - Find noise peaks in the power spectrum using peak finder findpeaks
 - Apply notch and bandpass filters to remove noise and plot the power spectrum of the filtered data
 - Plot the filtered strain data in a 0.3s time-window around the event
 - Make spectrogram of chirp
 - Compute correlation between the two detectors
 - Plot two signals at two detectors together
 - Numerical general relativity data for the first gravitational wave
 - Comparing detected signals with general relativity simulation for GW150914
 - Publish reusable MATLAB code for reproducible results
 - Supporting functions

Accessible for all, not just licensed MATLAB users!

Zoom: 100% UTF-8 LF script

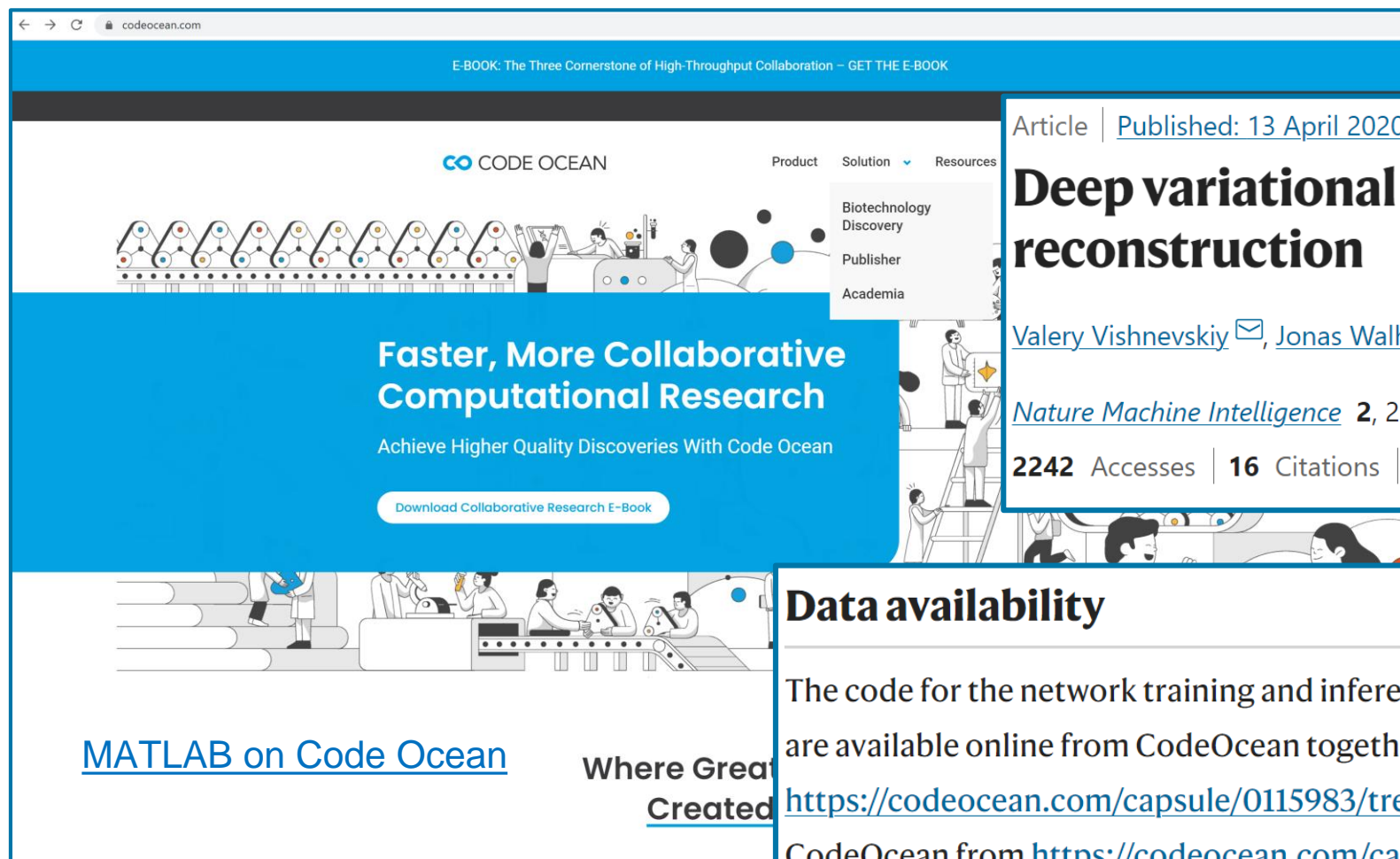
<https://www.mathworks.com/matlabcentral/fileexchange/154551-analyse-open-gravitational-wave-data-in-matlab>

<https://github.com/mathworks/open-gravitational-wave-data-matlab>

Portál pre reprodukovateľnosť s podporou MATLABu



CODE OCEAN



codeocean.com

E-BOOK: The Three Cornerstone of High-Throughput Collaboration – GET THE E-BOOK

CODE OCEAN

Product Solution Resources

Biotechnology
Discovery
Publisher
Academia

Faster, More Collaborative Computational Research

Achieve Higher Quality Discoveries With Code Ocean

Download Collaborative Research E-Book

[MATLAB on Code Ocean](#) Where Great [Created](#)

Article | [Published: 13 April 2020](#)

Deep variational network for rapid 4D flow MRI reconstruction

[Valery Vishnevskiy](#) , [Jonas Walheim](#) & [Sebastian Kozerke](#)

Nature Machine Intelligence **2**, 228–235 (2020) | [Cite this article](#)

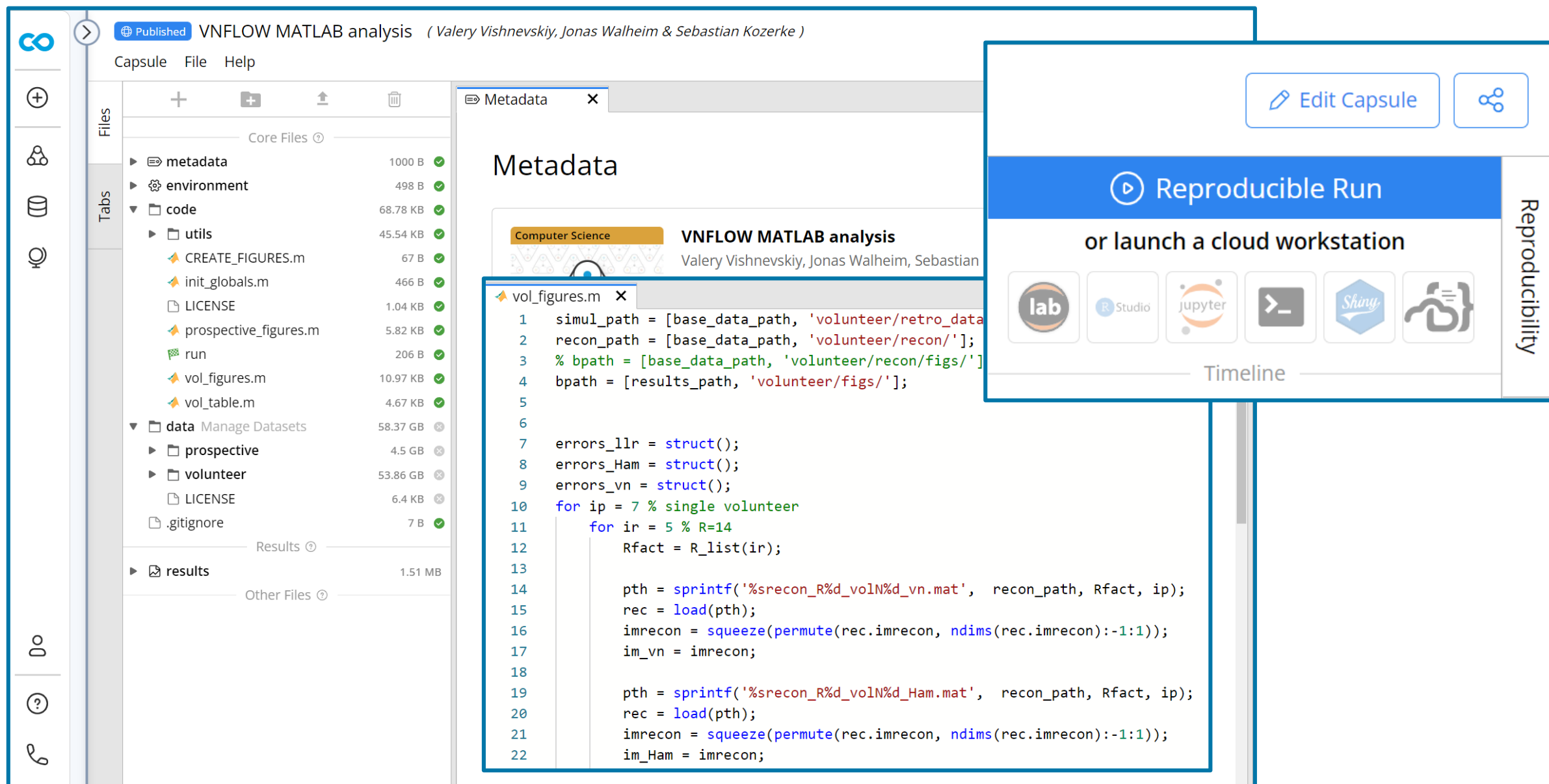
2242 Accesses | **16** Citations | **58** Altmetric | [Metrics](#)

Data availability

The code for the network training and inference used in this study as well as network weights are available online from CodeOcean together with volunteer data:

<https://codeocean.com/capsule/0115983/tree>⁴⁸. The code for analysis is available on CodeOcean from <https://codeocean.com/capsule/2587940/tree>⁴⁹.

Portál pre reprodukovateľnosť s podporou MATLABu



Published VNFLOW MATLAB analysis (Valery Vishnevskiy, Jonas Walheim & Sebastian Kozerke)

Capsule File Help

Files

Core Files

- ▶ metadata 1000 B ✓
- ▶ environment 498 B ✓
- ▼ code 68.78 KB ✓
 - ▶ utils 45.54 KB ✓
 - ▶ CREATE_FIGURES.m 67 B ✓
 - ▶ init_globals.m 466 B ✓
 - ▶ LICENSE 1.04 KB ✓
 - ▶ prospective_figures.m 5.82 KB ✓
 - ▶ run 206 B ✓
 - ▶ vol_figures.m 10.97 KB ✓
 - ▶ vol_table.m 4.67 KB ✓
- ▼ data Manage Datasets 58.37 GB ✗
 - ▶ prospective 4.5 GB ✗
 - ▶ volunteer 53.86 GB ✗
 - ▶ LICENSE 6.4 KB ✗
 - ▶ .gitignore 7 B ✓

Results

- ▶ results 1.51 MB

Other Files

Metadata

Computer Science VNFLOW MATLAB analysis
Valery Vishnevskiy, Jonas Walheim, Sebastian

vol_figures.m

```

1  simul_path = [base_data_path, 'volunteer/retro_data
2  recon_path = [base_data_path, 'volunteer/recon/'];
3  % bpath = [base_data_path, 'volunteer/recon/figs/'];
4  bpath = [results_path, 'volunteer/figs/'];
5
6
7  errors_llr = struct();
8  errors_Ham = struct();
9  errors_vn = struct();
10 for ip = 7 % single volunteer
11     for ir = 5 % R=14
12         Rfact = R_list(ir);
13
14         pth = sprintf('%srecon_R%d_vol%d_vn.mat', recon_path, Rfact, ip);
15         rec = load(pth);
16         imrecon = squeeze(permute(rec.imrecon, ndims(rec.imrecon):-1:1));
17         im_vn = imrecon;
18
19         pth = sprintf('%srecon_R%d_vol%d_Ham.mat', recon_path, Rfact, ip);
20         rec = load(pth);
21         imrecon = squeeze(permute(rec.imrecon, ndims(rec.imrecon):-1:1));
22         im_Ham = imrecon;

```

Reproducible Run
or launch a cloud workstation

lab R Studio jupyter >_ Shiny

Timeline

Reproducibility

MathWorks: zdieľanie a financovanie projektov

File Exchange: Hosting > 40,000 open source community toolboxes

PIVlab - particle image velocimetry (PIV) tool with GUI

version 2.36.5 (9.91 MB) by William Thielicke

Easy to use, GUI based tool to analyze, validate, postprocess, visualize and simulate (micro) PIV data. <http://pivlab.blogspot.com/>

★★★★★ 117 Ratings
559 Downloads
Updated 14 Oct 2020
[View Version History](#)
[view license on GitHub](#)

[+ Follow](#) [Download from GitHub](#)

Welcome to ECODATA-Animate's documentation!

ECODATA-Animate

ECODATA-Animate is a MATLAB® program for creating customized animated maps of animal movements. The program creates image frames that can be animated using the ECODATA-Prepare Movie Maker App. Define 1 input data. Development is Forecasting Program, as pi



William Thielicke • 2nd

+ Follow

Check out the documentat How cool is this: MathWorks pays professional programmers to improve my open- source code!

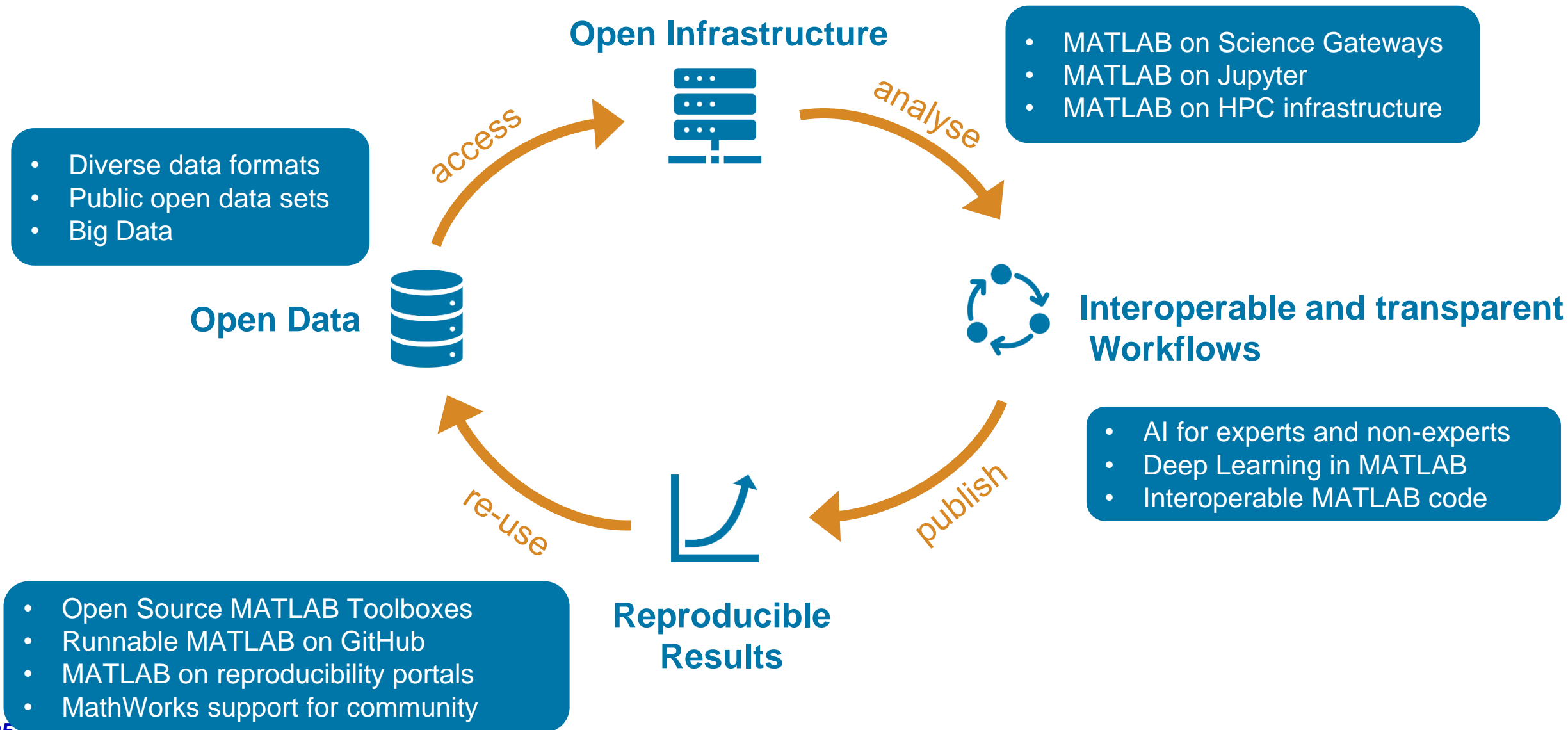
Note: ECODATA-Animate i have any suggestions or feature requests, encounter any bugs, or come across places where the documentation is unclear, please submit a GitHub issue.



Summer project collaboration between INCF and Mathworks

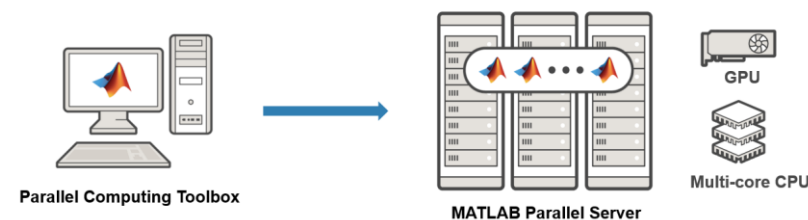
30 AUGUST 2021

Spoločné úsilie mnohých zúčastnených strán



HPC / Big Data v ČR a SR

- Použivatelia – akademická obec
 - CWL Univerzity
 - Akadémie Vied

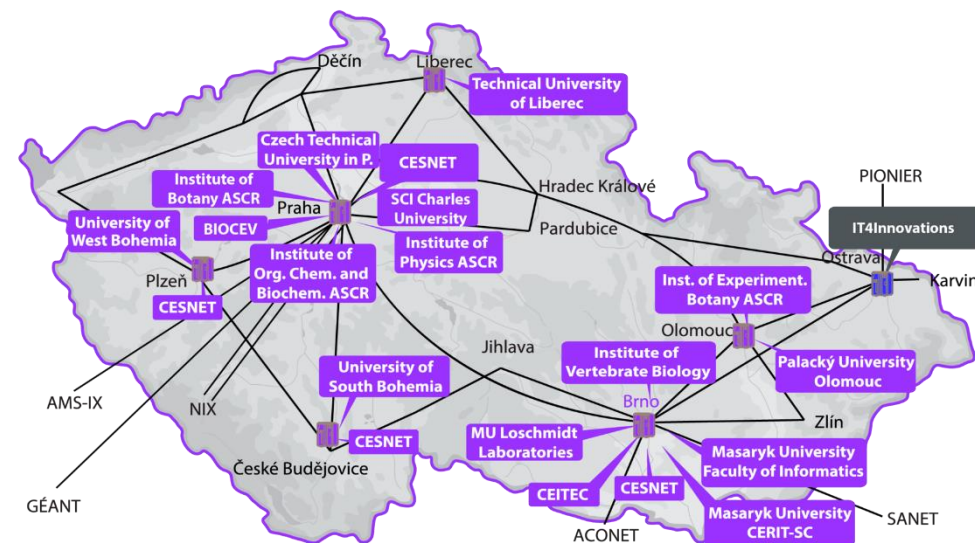


- HPC Infraštruktúra
 - CZ: e-INFRA, EOSC CZ
 - SK: SAV

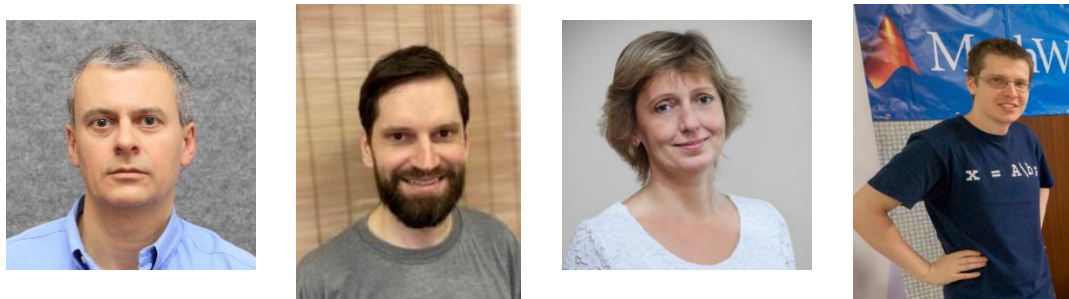
MATLAB Campus-Wide licence v ČR a SR



<https://www.humusoft.cz/univerzity/>



National Czech Programme



Ďakujem za pozornosť