

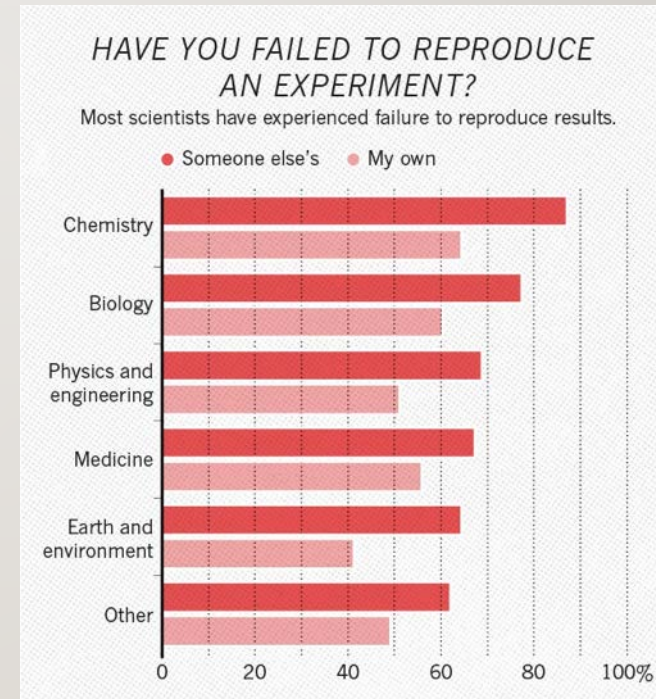
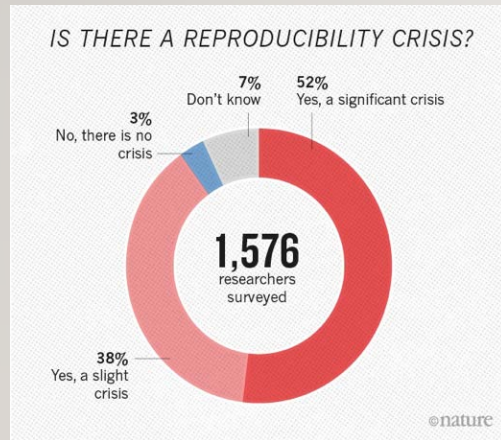
JUPYTER IN GALAXY ZA ANALIZO PODATKOV V ODPRTI ZNANOSTI

DAVID BAJŠ

UNIVERZA V MARIBORU, FAKULTETA ZA ELEKTROTEHNIKO, RAČUNALNIŠTVO IN INFORMATIKO

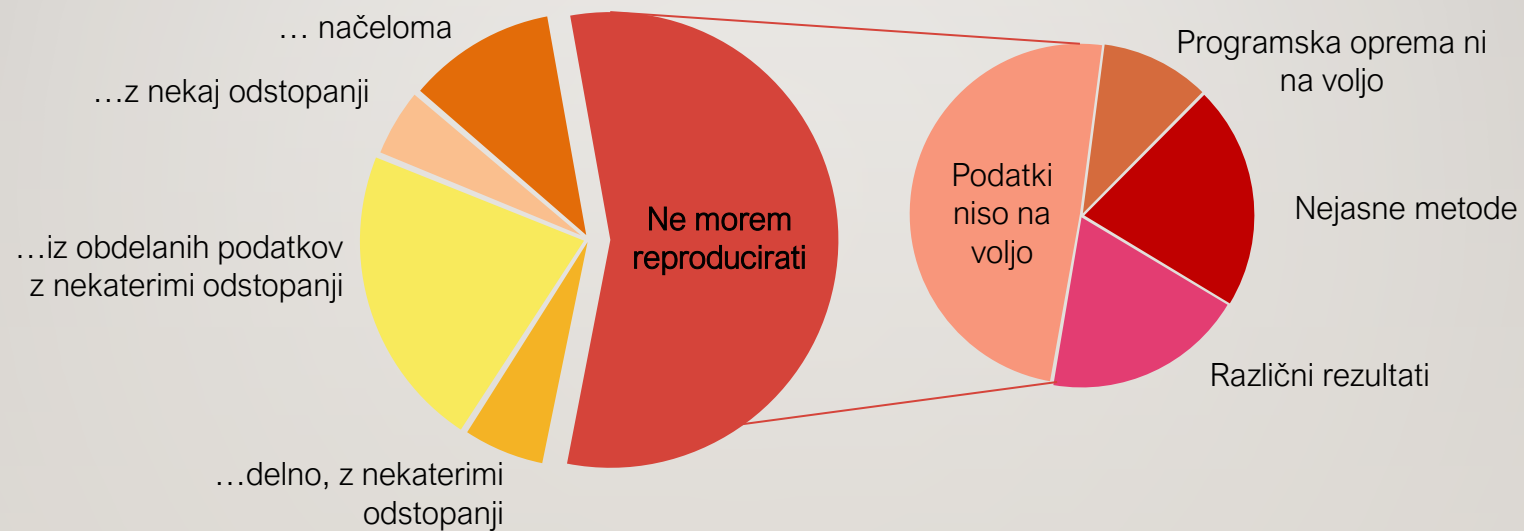
DAVID.BAJŠ1@UM.SI

KRIZA PONOVLJIVOSTI



KRIZA PONOVLJIVOSTI

- Lahko ponovimo...?



PROJEKT JUPYTER

- Brezplačna programska oprema
- Odprti standardi
- Spletne storitve za interaktivno računalništvo
- V več desetih programskih jezikih
- Od 2014, iz projekta IPython
- Jupyter Notebook, JupyterLab in JupyterHub



JUPYTER NOTEBOOK

- Ustvarjanje in skupna raba dokumentov
- Koda, besedilo, enačbe, vizualizacija
- Preprosta izkušnja

The image displays two overlapping Jupyter Notebook windows. The background window shows the 'Welcome to the Jupyter Notebook Server' page, which includes a warning message: 'WARNING: Don't rely on this server. Your server is hosted through a third party.' Below this, there are instructions on how to run Python code in a notebook cell.

The foreground window is titled 'Exploring the Lorenz System' and contains the following content:

In this Notebook we explore the [Lorenz system](#) of differential equations:

$$\begin{aligned}\dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy\end{aligned}$$

This is one of the classic systems in non-linear differential equations. It exhibits a range of complex behaviors as the parameters (σ, β, ρ) are varied, including what are known as *chaotic solutions*. The system was originally developed as a simplified mathematical model for atmospheric convection in 1963.

In [7]:

```
Interact(Lorenz, N=fixed(10), angle=(0.,360.),
         sigma=(0.0,50.0), beta=(0.,5), rho=(0.0,50.0))
```

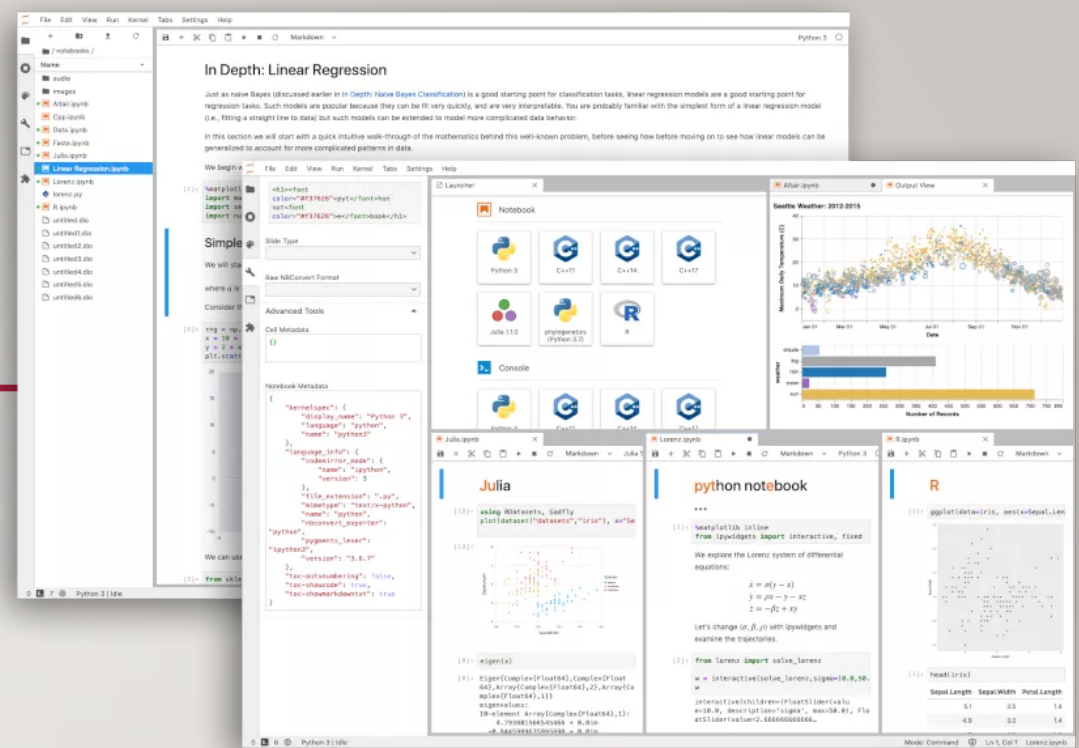
The interactive interface shows sliders for the following parameters:

- angle: 308.2
- max_time: 12
- σ : 10
- β : 2.6
- ρ : 28

Below the sliders is a 3D plot of the Lorenz attractor, showing its characteristic butterfly shape with multiple colored trajectories.

JUPYTERLAB

- Naslednja generacija Notebook vmesnika.
- Najnovejše spletno interaktivno razvojno okolje.
- Za Notebooke, kodo in podatke.
- Za področja podatkovne znanosti, znanstvenega računalništva, računalniškega novinarstva in strojnega učenja.



JUPYTERLAB - NAMESTITEV

- V brskalniku
 - JupyterLite
 - CoCalc
 - Colab notebooks
- Lokalno
 - pip install jupyterlab
 - conda, mamba, pip, pipenv ali docker
 - Anaconda

Jupyter Notebook: The Classic Notebook Interface

The Jupyter Notebook is the original web application for creating and sharing computational documents. It offers a simple, streamlined, document-centric experience.

[Try it in your browser](#)

[Install the Notebook](#)

JUPYTERLAB - PRIMER

- Izgled JupyterLab
- Print()
- Funkcija
- Markdown
- Import
- Branje datotek
- DataFrame
- Statistika
- Vizualizacija



JUPYTERHUB

- Večuporabniško vozlišče za interaktivne računalniške seje
- Za podjetja, učilnice in raziskovalne laboratorije
- Dostop preko brskalnika
- Potrebno naprednejše znanje
- Plačljiva oblačna okolja ali dobro poznavanje serverjev
- Google Cloud, Microsoft Azure, Amazon EC2, IBM Cloud



DELJENJE NOTEBOOK-OV

- Github in nato nbviewer (JAVNO)
- Github in nato nekatero izmed drugih orodij (binder, ...) <https://mybinder.org/>
- Gist/GitHub (Javno, z opcijami za privatno)
- Github in nato gostite svoj nbviewer (Za SKUPINO). To je napredno, za svojo skupino morate imeti strežnik in potem lahko dostopajo do datoteke

PLATFORMA GALAXY

- Odprtokodna platforma
- FAIR analiza podatkov
- Grafični spletni vmesnik
- Orodja iz različnih domen
- Interaktivna delovna okolja (Jupyter,...)
- Deljenje
- Objava
- Vizualizacija
- Ponovljivost

GALAXY STREŽNIKI

- Strežnikov je veliko in skoraj vsi so odprti ter brezplačni za uporabo, za vsakogar (125+)
<https://galaxyproject.org/use/>
- Glavni strežniki: usegalaxy.org (US), usegalaxy.eu (EU) in usegalaxy.org.au (Avstralija)
- Na primer, usegalaxy.eu je največji primerek Galaxy v Evropi in eden največjih na svetu, ki ima celo bolj specifične poddomene. <https://galaxyproject.eu/about.html>

GALAXY PRIVATNI STREŽNIKI

- Galaxy Flavor
- Uvoz orodij in knjižnic podatkov
- Strežnik po lastnem okusu (orodja, ki jih potrebujemo)
- Ni priporočljivo iz več razlogov

GALAXY ORODJA

- Enostaven način za analizo, upravljanje, obdelavo podatkov
- Nameščena orodja na strežniku
- Uvoz orodij iz Galaxy Tool Shed na svoj strežnik
- Napišite svoja orodja (za vaš strežnik, Galaxy Tool Shed, useGalaxy.*)
- useGalaxy.eu server ima nameščena sledeča orodja: (2500+) <https://galaxyproject.eu/tools>

DELJENJE PODATKOV

- Posamezni uporabniki (uporabniški račun Galaxy)
- Preko spletne povezave (ni treba biti registriran ali prijavljeni)
- Objavi (seznam objavljenih artiklov in prek povezave)

GALAXY SKUPNOST

- Odprti dogodki, konference
- Seminarji, usposabljanja
- Usposabljanja za orodja
- Za določeno področje raziskav potrebujete posebno usposabljanje za uporabo določenega orodja.
Enako velja za strežnik po meri

PRAKTIČNI PRIMER UPORABE PLATFORME GALAXY

- Uvoz podatkov
- Obdelava podatkov znotraj Galaxy platforme in JupyterLab
- Uporaba orodij in okolij
- Vizualizacija podatkov
- Shranjevanje podatkov
- Deljenje zgodovine oz. procesa obdelave podatkov

**HVALA ZA
POZORNOST**

