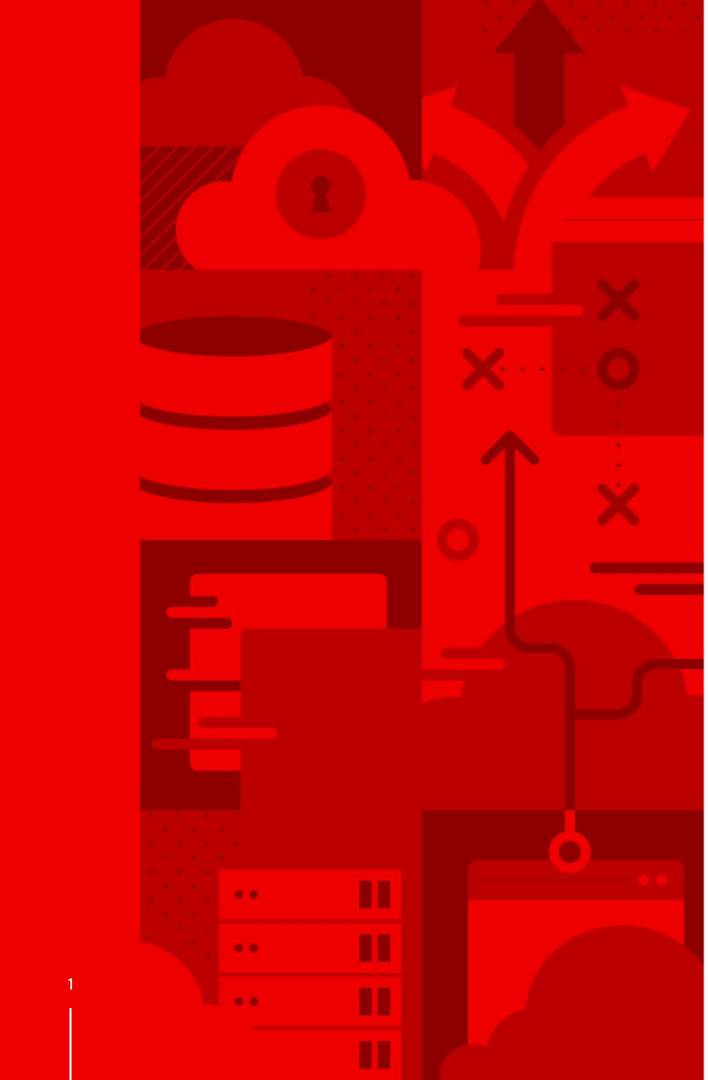


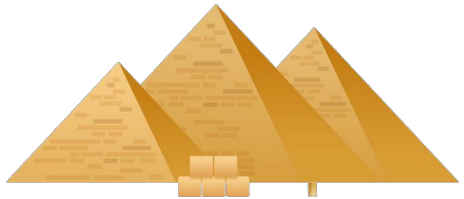
MLMU: Thoth

Reinforcement learning-based dependency resolution

Fridolin Pokorny <fridolin@redhat.com> & Thoth team
https://thoth-station.ninja/
Red Hat AICoE, Project Thoth
2021

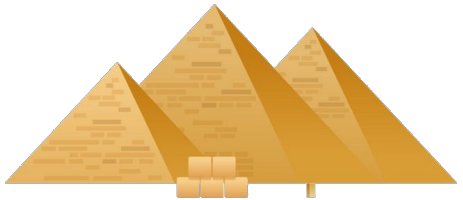
$ whoami && whoarewe

- Member of AICoE - AI Center of Excellence
  - Team in Red Hat's Office of the CTO
- Working on scalable platforms and machine learning applications
- Making Python ecosystem a better universe to code in
  - Improving developer experience when developing and running Python applications
  - Improving data science experience
    - http://opendatahub.io/
  - Service operated by bots to avoid developer's mundane work



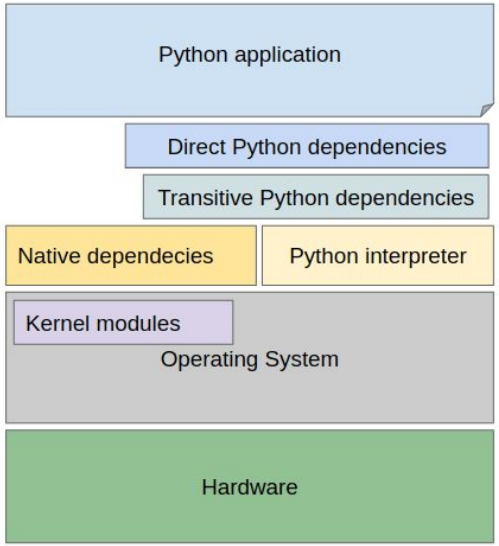
$ whoami && whoarewe

- Project Thoth
  - resurrecting ancient deities
  - https://thoth-station.ninja/
  - Members:
    - Christoph Goern
    - Francesco Murdaca
    - Fridolin Pokorny
    - Kevin Postlethwait
    - Harshad Reddy Nalla
    - Subin Modeel
    - + interns
  - Want to become a member?

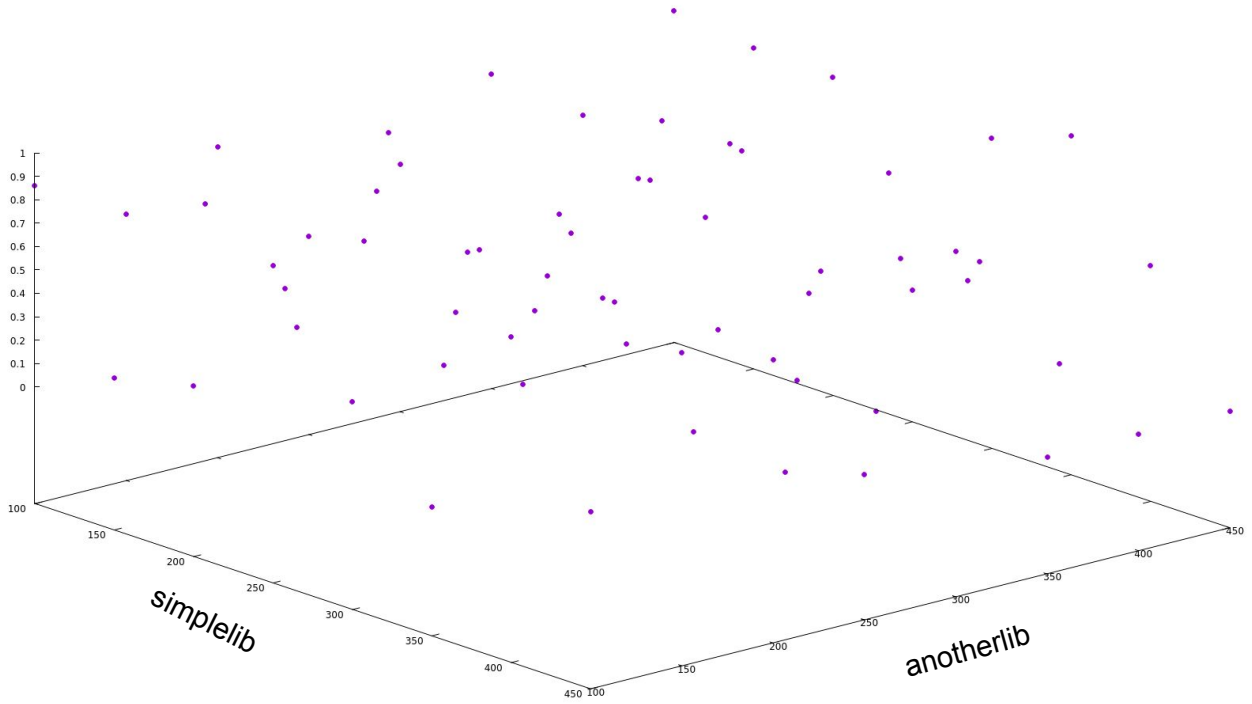


Thoth

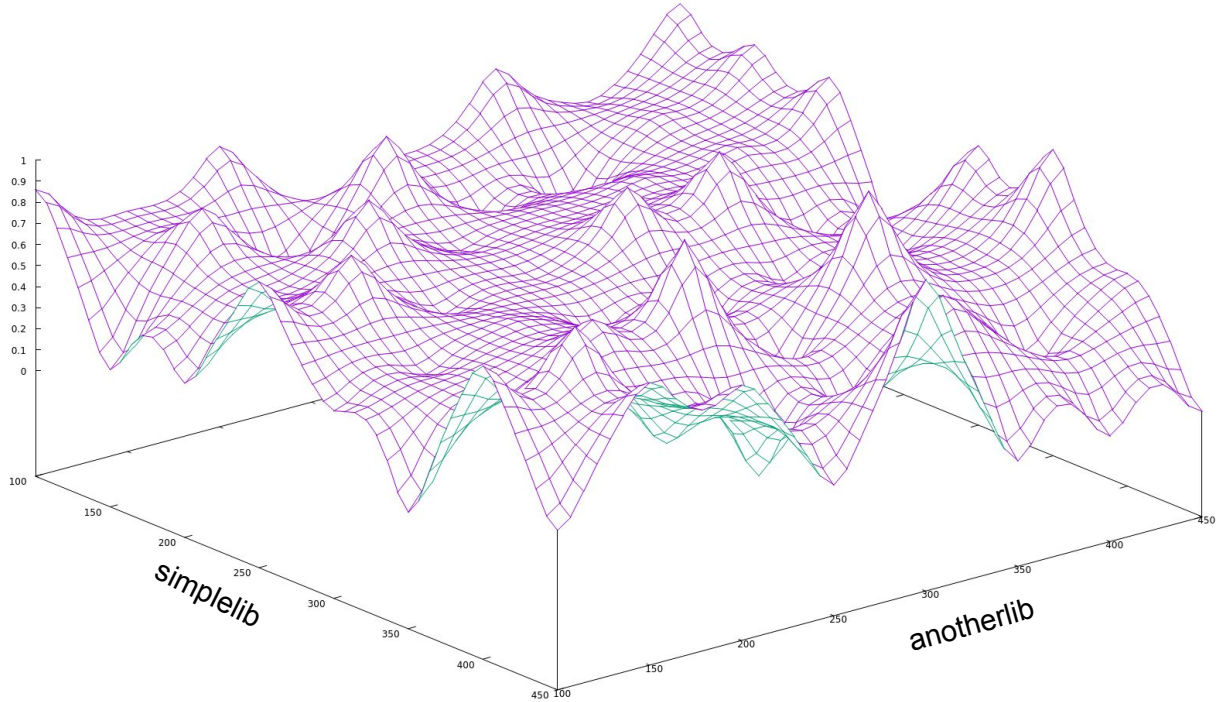
- Why Thoth?
  - Large stack of dependencies to run an application
  - Each layer in the stack introduces complexity
  - Any component on any layer may cause issues
  - An issue on any layer creates an observable aspect
    - Performance
    - Security
    - ...
  - How to find the right stack for user needs?



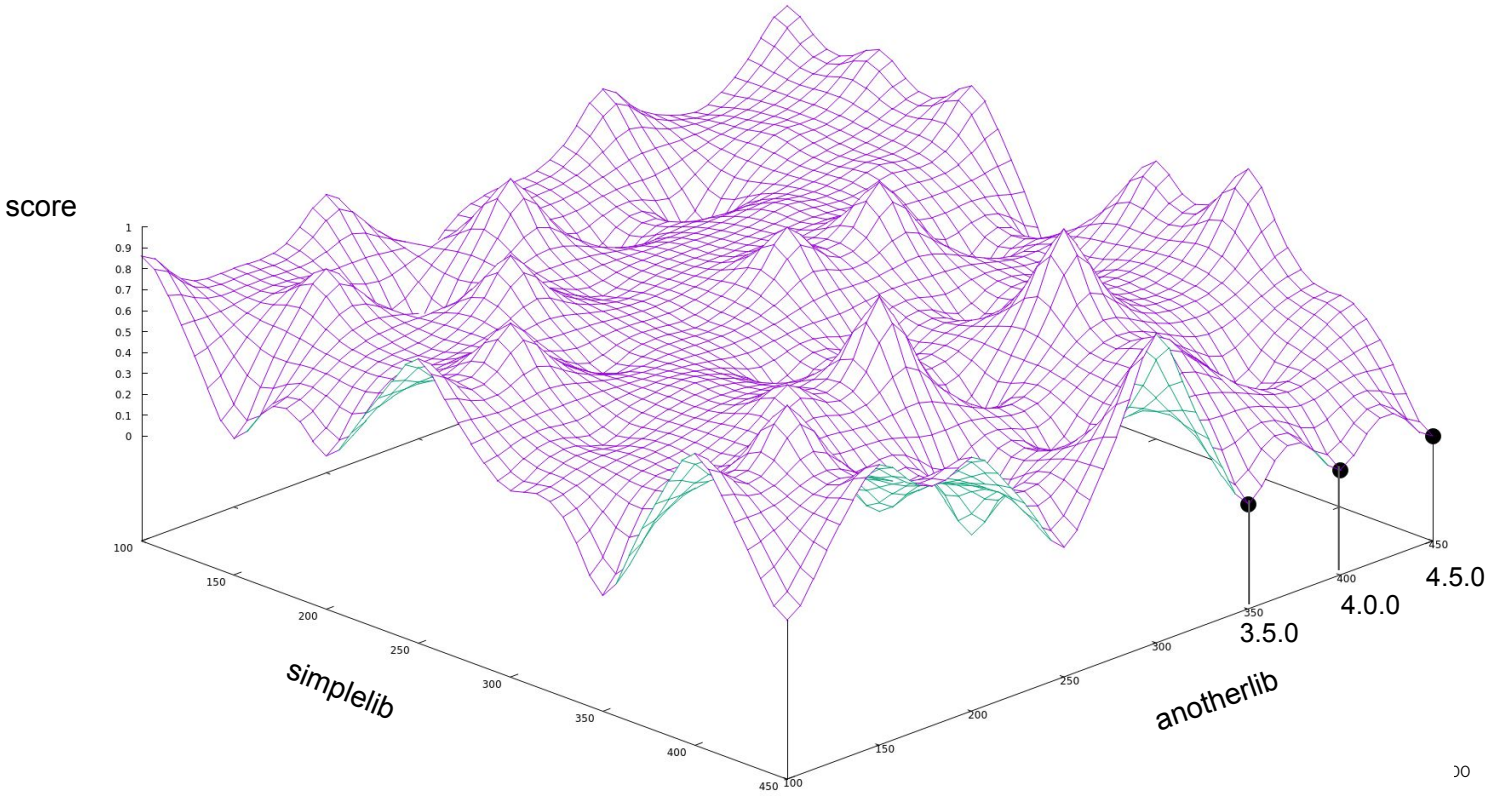
Thoth: score of the stack



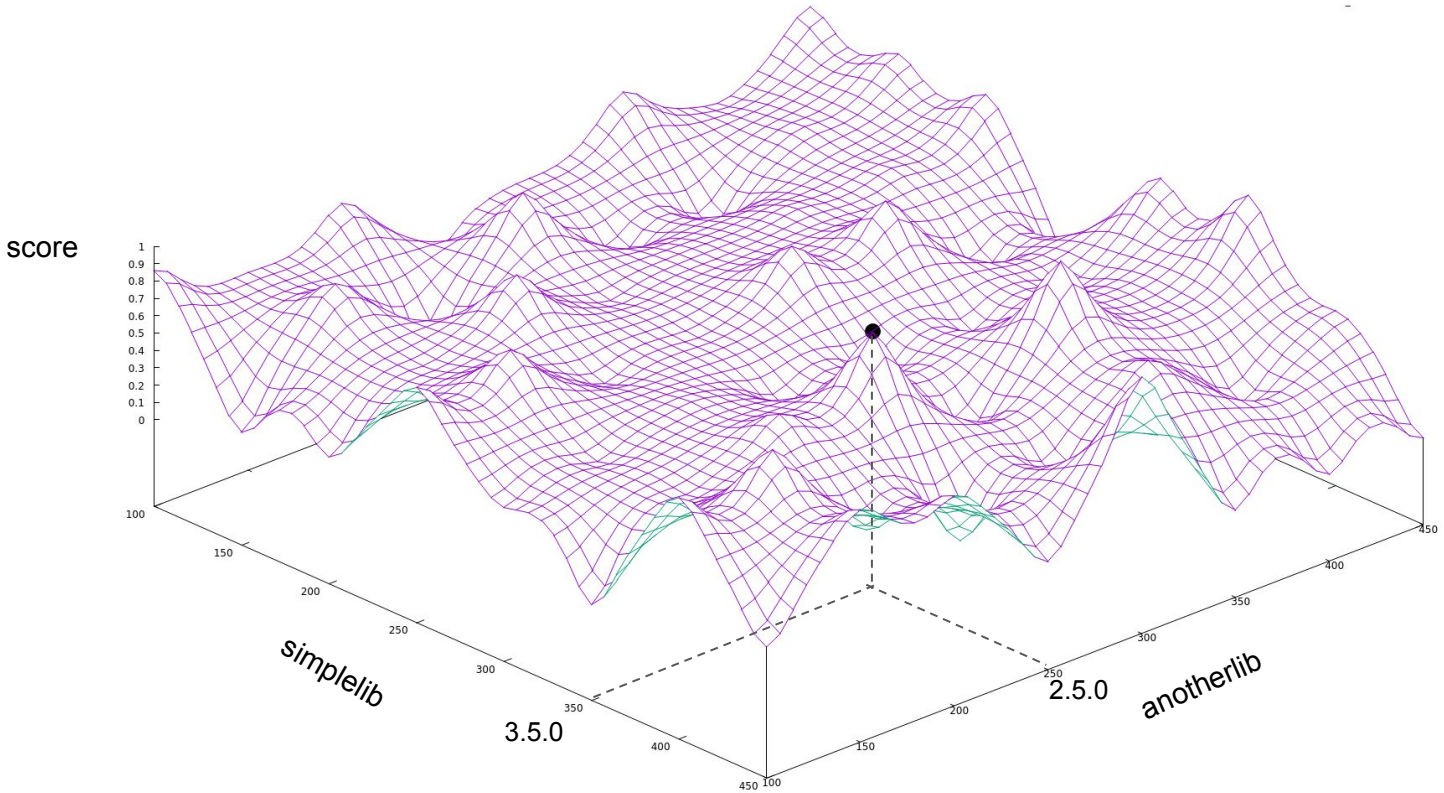
Thoth: score of the stack



Thoth: score of the stack



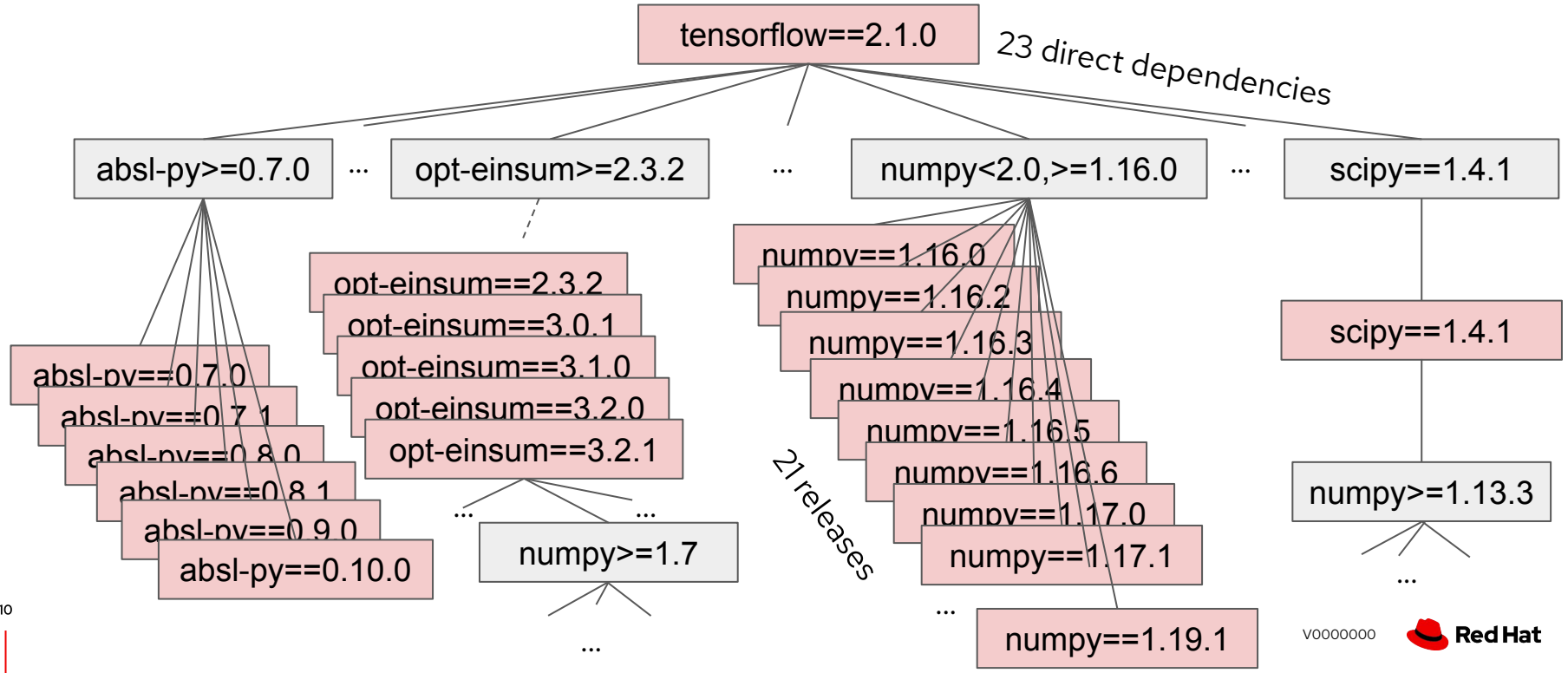
Thoth: score of the stack



Thoth

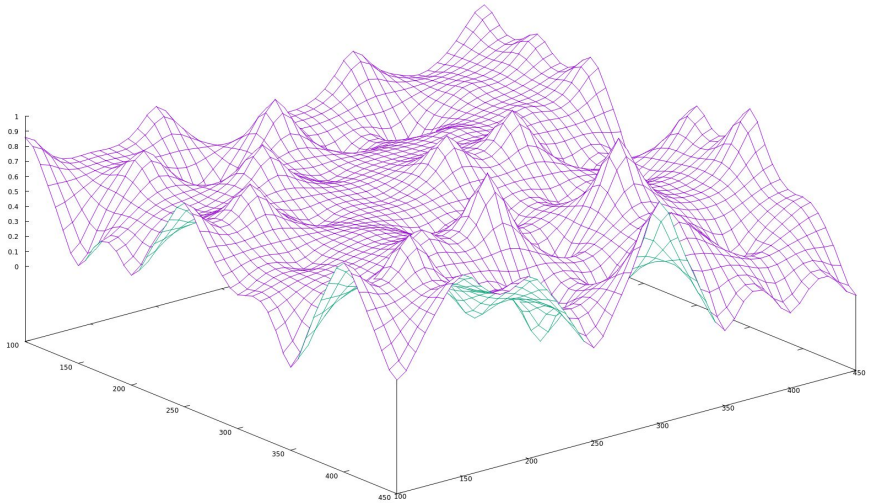
- Not just 2 dependencies but the whole dependency graph
  - Consider direct dependencies
  - Consider transitive dependencies
  - Consider runtime environment and other aspects of the software
- The state space is often too large to explore in real time
  - Number of combinations for a TensorFlow==2.1.0 stack: ~$3x10^{13}$
  - Red Hat Developer: AI software stack inspection with Thoth and TensorFlow
- Can we split this into smaller sub-problems and find a solution to the optimization problem?

Thoth: TensorFlow dependency graph



Thoth: random state space sampling

- Randomly resolve software stacks based on the dependency graph
- Check “how good” they are based on knowledge about the software



Thoth: reinforcement learning

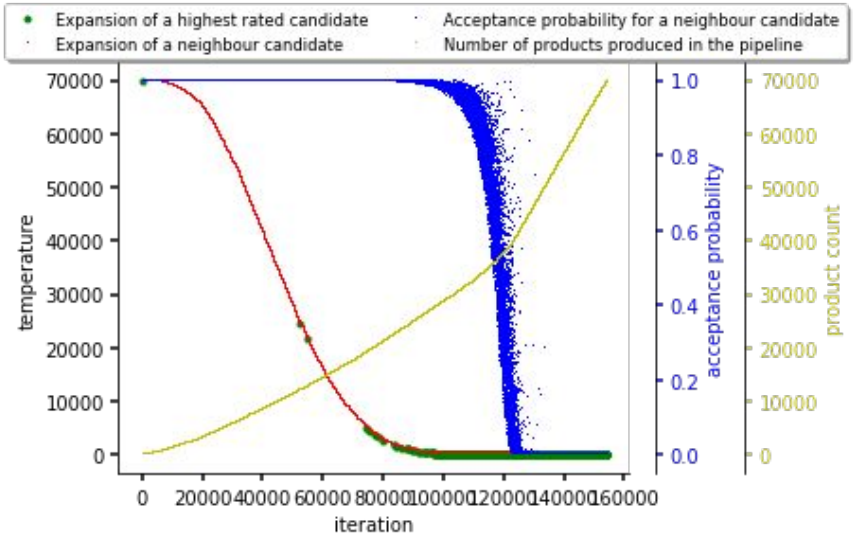
- Create a mechanism that can learn how to resolve software stacks
  - Using knowledge about the software to score software stacks
  - Learning how to navigate through the dependency graph to obtain best possible set of packages
- Gradient-free methods
  - Temporal Difference learning (TD-learning) and Monte Carlo Tree Search (MCTS)
  - Learning how to navigate through the state space and predicting score of trajectories in the dependency graph
  - Using “scoring pipeline” to score actions taken in the dependency graph and propagating a reward signal
  - Adaptive simulated annealing helps to balance exploration and exploitation

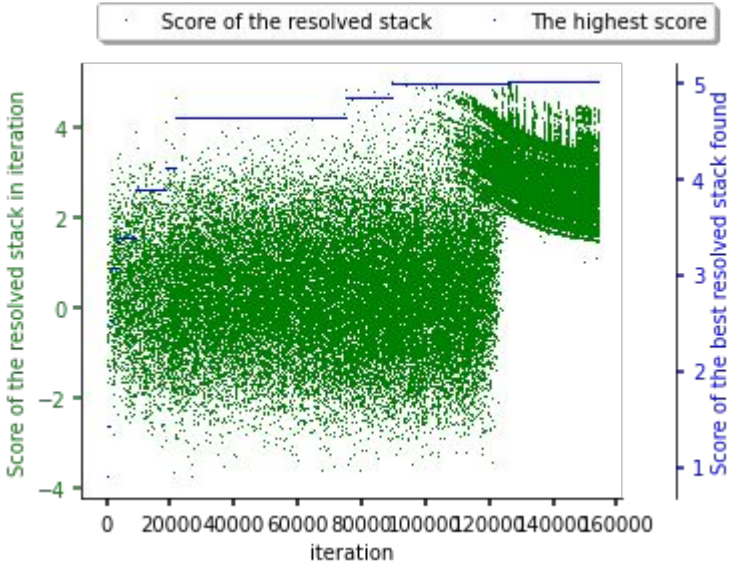
Thoth: example scenario

- Recommending a flask application
  - 7 packages considered in total
  - 7,861,340 possible valid software stack resolutions
  - Reinforcement learning based resolver run to resolve and score 70,000 software stacks (~1%)
  - Using randomly assigned scores for packages

Thoth: an example scenario

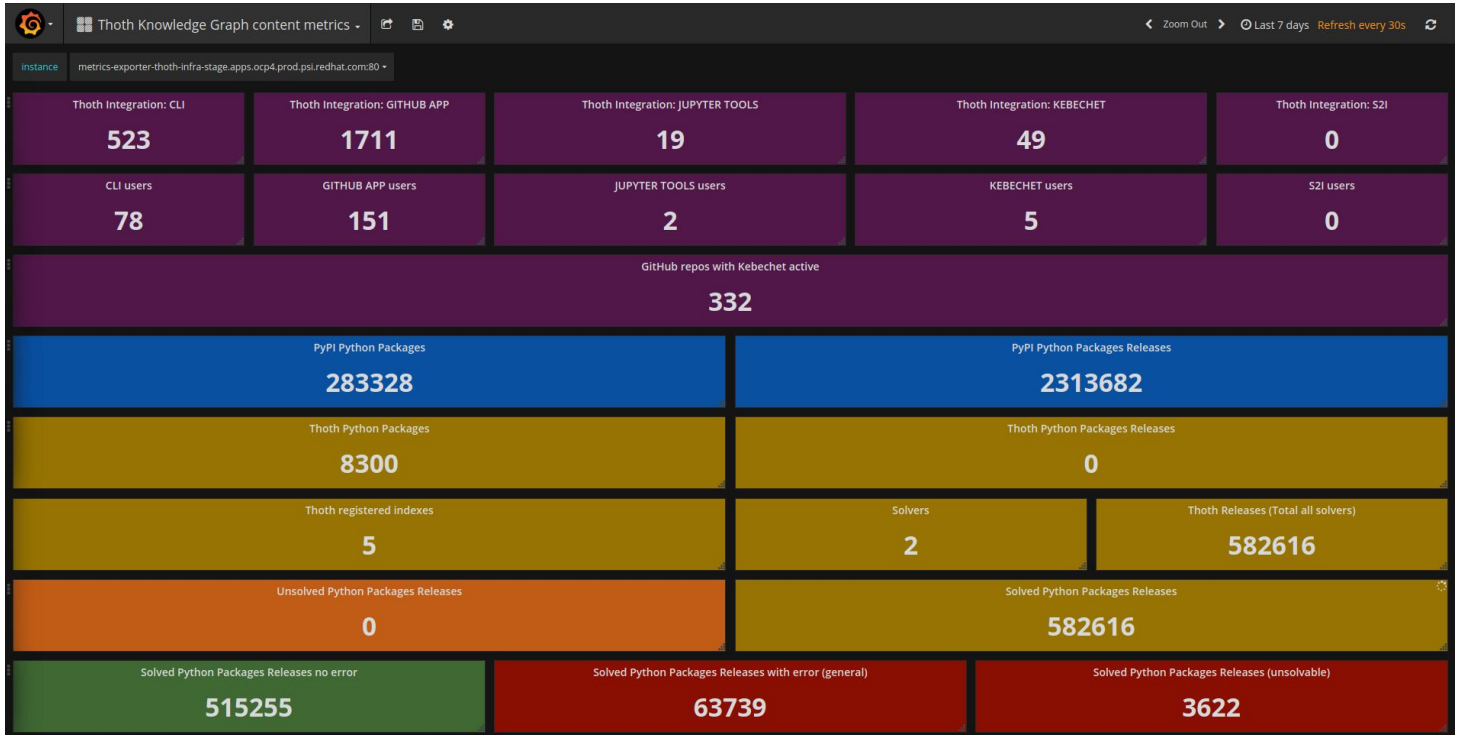
- Finding a software stack with a score of 4.99, the best possible candidate has 5.04





Red Hat

Grafana dashboards

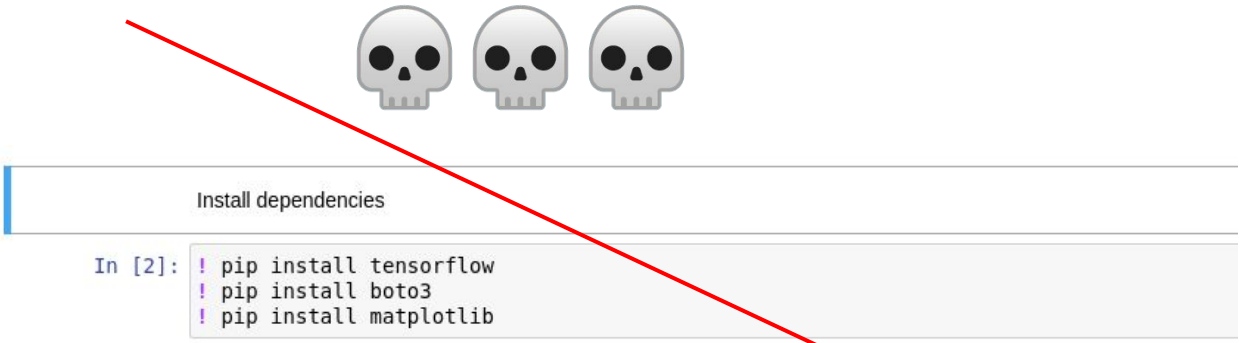


Jupyter Notebooks

Install dependencies

In [2]:
```
! pip install tensorflow
! pip install boto3
! pip install matplotlib
```

Jupyter Notebooks



Follow developers.redhat.com for more info (blog soon published):

- https://developers.redhat.com/

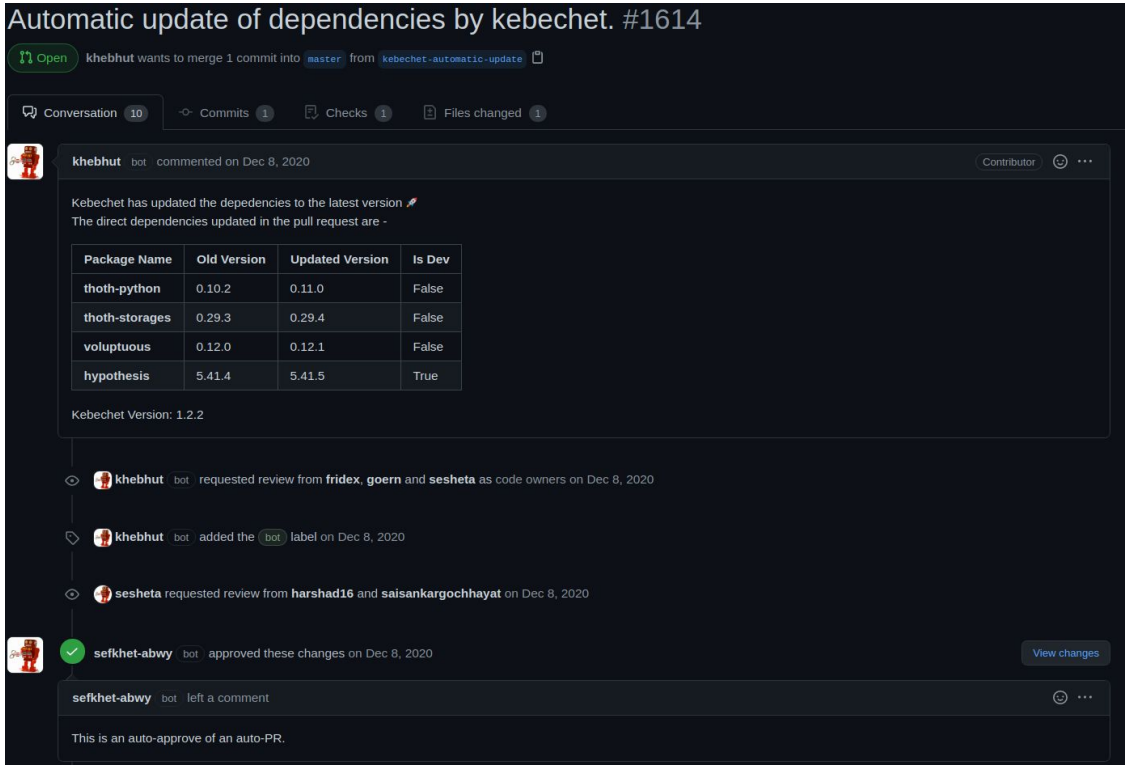
And our Twitter for updates:

- https://twitter.com/ThothStation

Jupyter Notebooks

- Jupyter Notebooks integration
  - https://github.com/thoth-station/jupyterlab-requirements

Kebechet



Automatic update of dependencies by kebechet. #1614

Open khebhut wants to merge 1 commit into master from kebechet-automatic-update

Conversation 10 Commits 1 Checks 1 Files changed 1

khebhut bot commented on Dec 8, 2020 Contributor

Kebechet has updated the depedencies to the latest version 🚀
The direct dependencies updated in the pull request are -

Package Name	Old Version	Updated Version	Is Dev
thoth-python	0.10.2	0.11.0	False
thoth-storages	0.29.3	0.29.4	False
voluptuous	0.12.0	0.12.1	False
hypothesis	5.41.4	5.41.5	True

Kebechet Version: 1.2.2

khebhut bot requested review from fridex, goern and sesheta as code owners on Dec 8, 2020

khebhut bot added the bot label on Dec 8, 2020

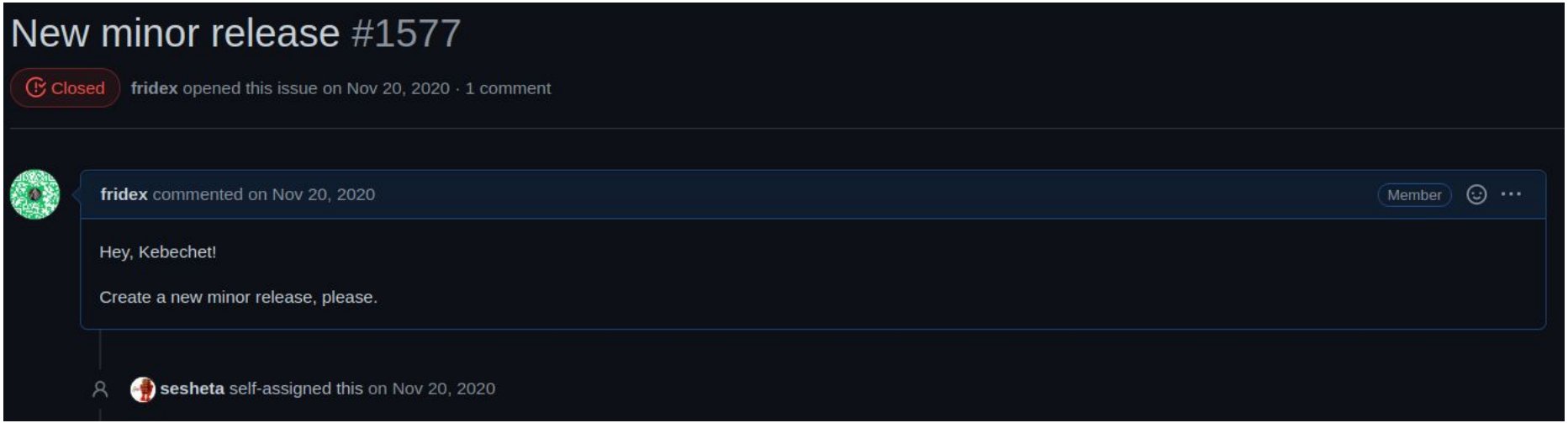
sesheta requested review from harshad16 and saisankargochhayat on Dec 8, 2020

sefkhet-abwy bot approved these changes on Dec 8, 2020 View changes

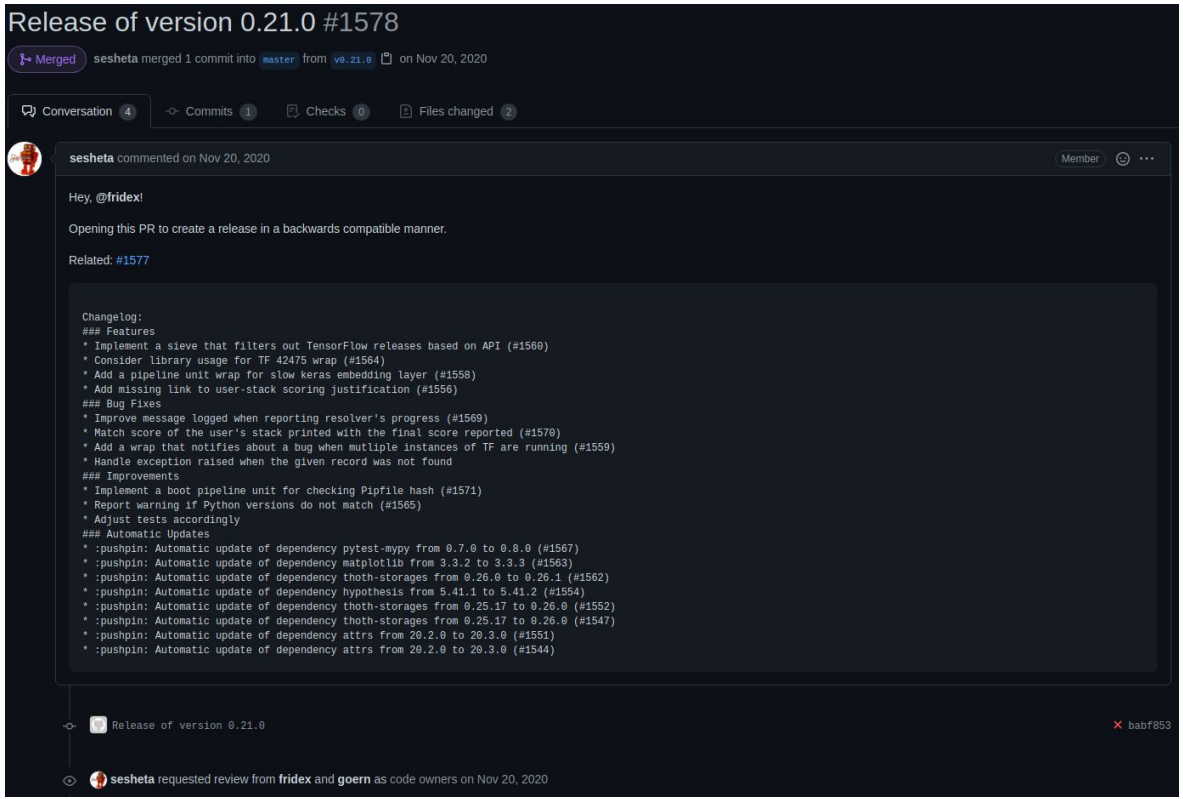
sefkhet-abwy bot left a comment

This is an auto-approve of an auto-PR.

Kebechet


New minor release #1577
Closed
fridex opened this issue on Nov 20, 2020 · 1 comment
fridex commented on Nov 20, 2020
Member
Hey, Kebechet!
Create a new minor release, please.
sesheta self-assigned this on Nov 20, 2020

Kebechet



Release of version 0.21.0 #1578

Merged sesheta merged 1 commit into master from v0.21.0 on Nov 20, 2020

Conversation 4 Commits 1 Checks 0 Files changed 2

sesheta commented on Nov 20, 2020 Member

Hey, @fridex!

Opening this PR to create a release in a backwards compatible manner.

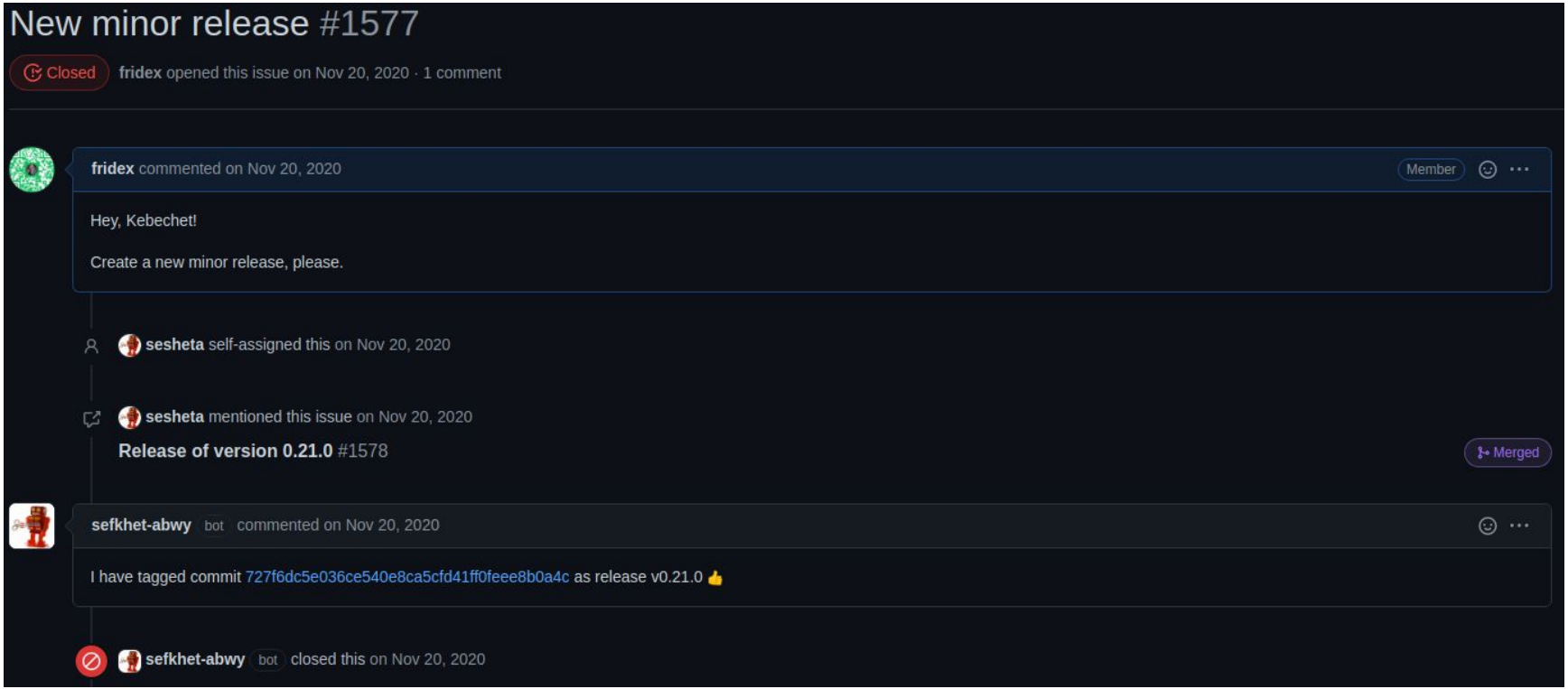
Related: #1577

```
Changelog:
### Features
* Implement a sieve that filters out TensorFlow releases based on API (#1560)
* Consider library usage for TF 42475 wrap (#1564)
* Add a pipeline unit wrap for slow keras embedding layer (#1558)
* Add missing link to user-stack scoring justification (#1556)
### Bug Fixes
* Improve message logged when reporting resolver's progress (#1569)
* Match score of the user's stack printed with the final score reported (#1570)
* Add a wrap that notifies about a bug when mutliple instances of TF are running (#1559)
* Handle exception raised when the given record was not found
### Improvements
* Implement a boot pipeline unit for checking Pipfile hash (#1571)
* Report warning if Python versions do not match (#1565)
* Adjust tests accordingly
### Automatic Updates
* :pushpin: Automatic update of dependency pytest-mypy from 0.7.0 to 0.8.0 (#1567)
* :pushpin: Automatic update of dependency matplotlib from 3.3.2 to 3.3.3 (#1563)
* :pushpin: Automatic update of dependency thoth-storages from 0.26.0 to 0.26.1 (#1562)
* :pushpin: Automatic update of dependency hypothesis from 5.41.1 to 5.41.2 (#1554)
* :pushpin: Automatic update of dependency thoth-storages from 0.25.17 to 0.26.0 (#1552)
* :pushpin: Automatic update of dependency thoth-storages from 0.25.17 to 0.26.0 (#1547)
* :pushpin: Automatic update of dependency attrs from 20.2.0 to 20.3.0 (#1551)
* :pushpin: Automatic update of dependency attrs from 20.2.0 to 20.3.0 (#1544)
```

Release of version 0.21.0 babf853

sesheta requested review from fridex and goern as code owners on Nov 20, 2020

Kebechet

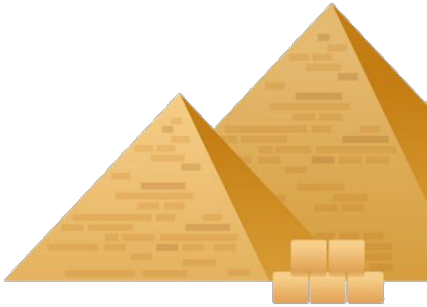

New minor release #1577
Closed
fridex opened this issue on Nov 20, 2020 · 1 comment
fridex commented on Nov 20, 2020
Member
Hey, Kebechet!
Create a new minor release, please.
sesheta self-assigned this on Nov 20, 2020
sesheta mentioned this issue on Nov 20, 2020
Release of version 0.21.0 #1578
Merged
sefkhet-abwy bot commented on Nov 20, 2020
I have tagged commit 727f6dc5e036ce540e8ca5cfd41ff0feee8b0a4c as release v0.21.0
sefkhet-abwy bot closed this on Nov 20, 2020

Call to action

- Use our data
  - Kaggle datasets at https://www.kaggle.com/thothstation/datasets
  - By conducting experiments on our datasets
  - By asking our services for advise
- Use our Services
  - GitHub App - https://github.com/apps/qeb-hwt
  - Thamos
  - OpenShift Pipeline Tasks
- Create Open Source software
  - Contribute a tiny feature (or a large)
  - Project Planning is done openly

Project Thoth

- AICoE, Office of the CTO
- Homepage
  - http://thoth-station.ninja/
- GitHub organization
  - http://github.com/thoth-station/
- Twitter account with updates
  - https://twitter.com/thothstation
- YouTube channel
  - https://www.youtube.com/channel/UCIUIDuq_hQ6vlzmqM59B2Lw
  - http://bit.ly/thoth-on-youtube



Thanks for your attention!

https://github.com/thoth-station

https://twitter.com/thothstation

https://www.youtube.com/channel/UCIUIDuq_hQ6vlzmqM59B2Lw

References

Website https://thoth-station.ninja/

Twitter https://twitter.com/thothstation

GitHub https://github.com/thoth-station


linkedin.com/company/red-hat


youtube.com/user/RedHatVideos


facebook.com/redhatinc


twitter.com/RedHat

Red Hat