

Jupyter Notebook, JupyterHub and Nbgrader

Nik Klever

University of Applied Sciences Augsburg / Faculty of Computer Science
nik.klever@hs-augsburg.de

Abstract

In modern university teaching, students are encouraged to learn on their own initiative and in a self-regulated way. This applies in particular to e-learning offers and requires an environment that allows to test your own knowledge and skills at any time, i.e. carry out a self-assessment. With the help of digital methods, it is possible to set individualized tasks, automatically test student submissions and provide appropriate feedback. The project "OWL - Offene digitale Werkzeuge in der Lehre" (open digital tools in teaching) originated from this idea and enables

- Supporting students by self-assessment platforms
- E-learning for independent learning
- creation of feedback
- reduced burden on deadlines
- platforms for easy interaction and submission, including bonus tests

In this part, the integration of Nbgrader based on Jupyter Notebook into a JupyterHub server was realized using the authorization system of University of Applied Sciences Augsburg, to build a useful platform for all teachers who are interested in using Jupyter Notebook as a standard tool in teaching.

Keywords: digital teaching tools, project jupyter

1. Overview on Jupyter

1.1 Jupyter Notebook

Jupyter Notebook (Project Jupyter 2020a) is a particularly well-established webserver for creating mathematical-scientific, statistical, IT and programming, but also data and graphics content. Integration of various programming languages such as Python, R, Java, Javascript and many others is possible as well as integration of various markup languages such as HTML, Markdown, Scalable Vector Graphics, LaTeX and others. The base structure to include any of this different content are cells. In a nutshell, Jupyter Notebook is therefore quick and easy to create of all kinds of content on a webpa-

ge. The advantages of Jupyter Notebook are described in more detail in a short manual (Klever 2020a).

1.2 JupyterHub

Like Jupyter Notebook, JupyterHub has been developed as part of the worldwide open source Project Jupyter (2020b) creating single Jupyter Notebook web-servers for every authorized user of the JupyterHub server. The advantage of JupyterHub (Project Jupyter 2020c) is on the one side that a Jupyter Notebook installation on the user's own computer is no longer necessary and on the other side a permanent accessibility of the user's Jupyter Notebook in the cloud. An example of an intensive use of JupyterHub is the Libre Text Project (Libre Text Project 2020) at the University of California, Davis, with an Open Textbooks

Pilot Program and support from the American Department of Education (University of California, Davis 2020).

1.3 Nbgrader

Nbgrader (Project Jupyter 2020d) arose on the basis of Jupyter Notebook to allow the creation of Jupyter Notebooks with exercises in cells whose solutions are hidden in a copy for the students. After solving the exercises in this copy, students can submit it back. After the deadline, the lecturer can collect all submissions and grade the individual solutions either automatically (currently only for Python) or manually (for all programming languages or pure text questions), correct them in a simple manner, provide them with feedback and return them to the students.

2. Teaching and Learning with Jupyter

The idea and motivation for this part of the OWL project is based on the work on Jupyter Notebook for teaching, which was presented to the participants of the first conference "Digital teaching - challenges and opportunities" at the University of Applied Sciences Augsburg (Klever & Rösch 2016). In the meantime, Jupyter Notebook has established itself both as a tool for the exchange of scientific results and publications and, together with Nbgrader and other tools, also has a major influence on teaching. A clear indication of this can be seen in particular in the online book "Teaching and Learning with Jupyter" published in 2019 (Barba et.al. 2019).

2.1. Jupyter Notebook

After Jupyter Notebook had been used successfully by the author for several years, a sufficient number of Jupyter Notebooks was available, especially in programming with Python. These include the lectures network programming, data management 1 and image processing in the course interactive media, computer science 1 and 2 in the course systems engineering, programming 3 and image processing in the course computer science.

2.2. JupyterHub

A JupyterHub server (Klever 2020b) was set up in 2016 and used for the creation and use of the Jupyter notebooks in the lectures computer science 1 and 2 in the first year of the course systems engineering before the start of the project OWL. The project could therefore be built upon these experiences. The course systems engineering is ideally suited for this pro-

ject in particular due to its mixture of digital teaching and its modern forms of teaching such as flipped classroom and face-to-face teaching in its venues in Nördlingen and Memmingen for simultaneous lectures via videoconference tools. In addition, the concept of this course has already been awarded twice with a special award of the VDMA (Mechanical Engineering Industry Association) award "Bestes Maschinenhaus 2017" (Hochschule Augsburg 2017) and second with the teaching award „Preis für herausragende Lehre“ of the Bavarian Ministry for Science and Art (Jacob et. al. 2018). A detailed description of the concept for this course can be found in Jacob et. al. (2017). For this first JupyterHub server, the connection to the authorization system LDAP (Wikipedia, 2020) of the University of Applied Sciences Augsburg Computing Center (Hochschule Augsburg 2020) was realized using the available integrated extension of the JupyterHub LDAP authenticator (Klever, 2017) as well as a separate LDAP2LocalUserAuthenticator (Klever 2020c) based on the LDAP authenticator. The latter creates local users with corresponding directories on the local file system, so that user data in the IT department are separated from the user data on the JupyterHub, but authentication is still carried out via the university's authentication system. The experiences of using the JupyterHub, as well as the feedback from students was positive, however, submission and collection from student's exercise notebooks was carried out with scripts of the author and an individual correction of the exercises was only discussed with the students on request.

2.3. Nbgrader

One of the features of Nbgrader is submission and collection of Jupyter notebooks. One year after the first publicly available version of Nbgrader (Version 0.1.0 from July 17th, 2015 (Hamrick 2015)) the author used Version 0.3.0 (Hamrick 2016) from June 25th, 2016 for the first tests in the project OWL. Version 0.4.0 (Hamrick 2017) from February 2nd, 2017 was then used in this productive beta test as part of the lectures computer science 1 and 2 for the third year of the course systems engineering. This test received also positive feedback, especially because of the automated, and if necessary manually reviewed, feedback response to the students. However, scripts had to be written for the exchange of this response and all registered students could use all available courses at the same time. With version 0.6.0 (Hamrick 2019) it was possible to configure and set up various courses by different lecturers and different students. This version was used for the fourth year of the course systems engineering and is the basis for a common platform used at the University of Applied Sciences Augsburg. Corresponding to the intensive use, appropriate patches for eliminating bugs

have already been introduced for this version (Kleve, 2019). The disadvantage of this setup is that all lecturers must be registered as JupyterHub admin users, which implies that all lecturers can both create and delete users in the database of the JupyterHub server. As deletion of a user from the database is relatively easy to restore this disadvantage is not that serious. However, it is absolutely necessary that the parameters in the configuration of the JupyterHub

```
JupyterHub.admin_access = False
```

is set to prevent the lecturers from creating and deleting Jupyter Notebooks from other users.

3. Example for the Use of JupyterHub and Nbgrader

The lectures computer science 1 + 2 in the course systems engineering are an introduction to programming with Python. For this, a total of 12 chapters - in Nbgrader named Assignments - were created for the lecture computer science 1, each with a different number of Jupyter notebooks (see Fig. 1).

In addition, for almost all notebooks, between 1 and 4 screencasts were developed with programming examples in Jupyter notebooks, including the explanations for every single programming step. Fig. 2 shows

an example of a notebook with an autograded answer and an autograder tests cell from a lecturer's perspective, while Fig. 3 shows the same notebook from the students' point of view.

While working on the notebook, students can validate the code at any time, which lights up green when executed correctly, otherwise red. In order to enable the students to carry out a simple self-examination, the programming tasks should be divided into minimal steps to clarify when and where an error occurs in a non-validating programming step. This is not always easy and sometimes impossible. After submitting the finished Jupyter notebook, feedback is provided via the autograding process including the points achieved, whereby either no points or the entire number of points can be awarded (according to the validation of the notebook) (see Fig. 4)

Manual correcting is essential for an accurate grading. However, this is complex and cannot be performed for all 30-50 notebooks in a lecture during the everyday life of a lecturer. For this reason, only one Jupyter notebook for each assignment was created as an exercise notebook and therefore manually corrected and used for a bonus rating. Both the feedback on the individual exercise notebooks and the bonus rating was very well received and acknowledged by the students.

Name	Due Date	Status	Edit	Generate	Preview	Release	Collect	# Submissions	Generate Feedback	Release Feedback
Kap01_Einführung	2019-10-24 21:00:00 UTC	released						41		
Kap02_Grundlagen_1	2019-11-07 21:00:00 UTC	released						38		
Kap03_Grundlagen_2	2019-11-13 21:00:00 UTC	released						37		
Kap04_Kontrollstrukturen	2019-11-21 21:00:00 UTC	released						39		
Kap05_Funktionen	2019-11-28 21:00:00 UTC	released						36		
Kap06_Ein_und_Ausgabe	2019-12-05 21:00:00 UTC	released						37		
Kap07_Wiederholung_Basics	2019-12-12 21:00:00 UTC	released						35		
Kap08_Datenstrukturen_1	2019-12-19 21:00:00 UTC	released						36		
Kap09_Datenstrukturen_2	2020-01-09 21:00:00 UTC	released						0		
Kap10_Wiederholung_Basics_2	2020-01-16 21:00:00 UTC	released						0		
Kap11_Probepf Prüfung	None	draft						0		
Präsenzveranstaltungen	None	released						2		

Figure 1. Manage assignment view with all 12 chapters of the lecture computer science 1.

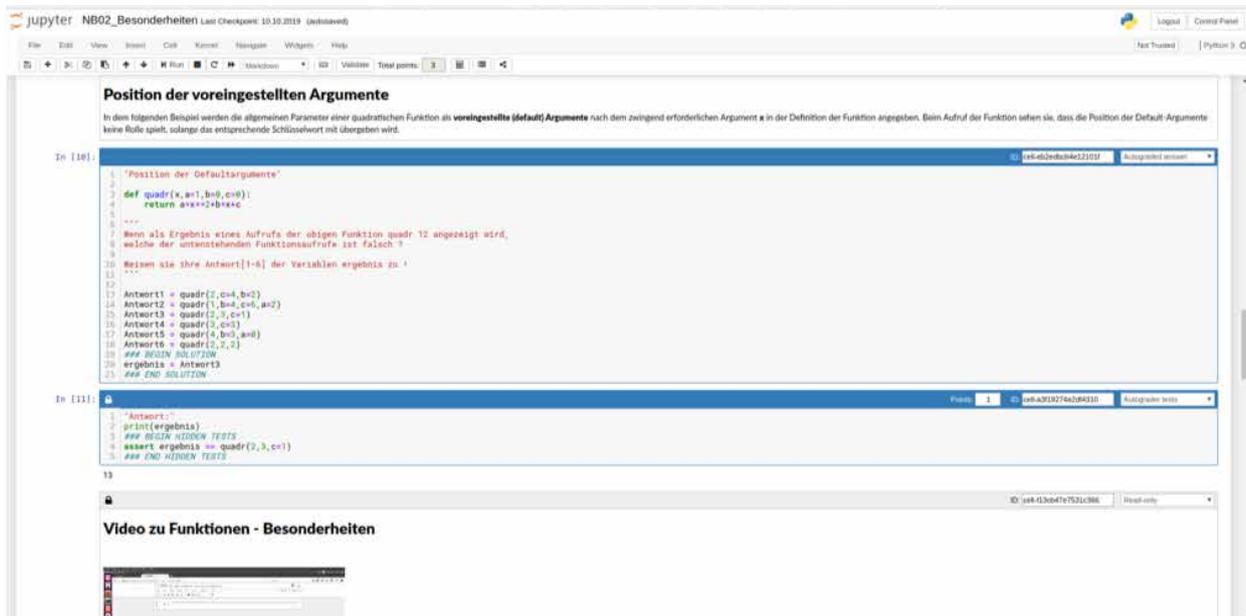


Figure 2. Section from a Jupyter notebook with an Autograded answer and an Autograder tests cell (from the lecturer's perspective).

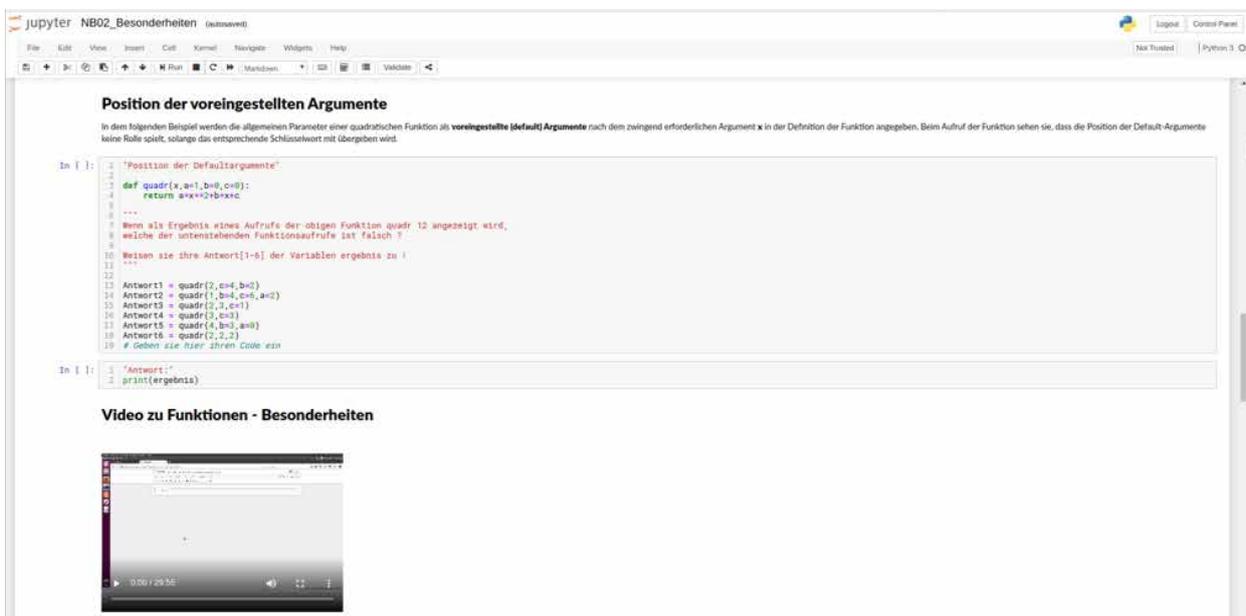


Figure 3. The section from the Jupyter notebook corresponding to Fig. 2 from a student's perspective.



Figure 4. The same section from the Jupyter notebook corresponding to Fig. 2 and Fig. 3, after the autograding as feedback to the students.

4. Sustainability

The positive experiences using Nbgrader in the lectures computer science 1 + 2 in the course systems engineering encouraged the author to open and release the JupyterHub server for other courses and lecturers. Therefore, first discussions about the administration of this server by the IT department of the University of Applied Sciences Augsburg were held. Setting up new courses is more complex in the environment at the University of Applied Sciences Augsburg with its own authorization service than in the multiple courses demos provided by Nbgrader (Project Jupyter 2020e). To set up new courses in a suitable manner without major effort, a script, which has already been published in Github (Klever 2020d) was written as a significant relief for administrators. To install a new course is therefore reduced to call a script with two parameters, the name of the course and the user-account of the corresponding lecturer:

```
$ python createMultipleCourse.py -a Name_of_Course
Useraccount_of_Teacher
```

A further step is necessary to register all students of this course afterwards. If a list of the students enrolled in this course is available electronically in a file with their first name, last name, email address and user account, these students can be registered in one step into Nbgrader using the same script with the following call:

```
$ python createMultipleCourse.py Name_of_Course
Useraccount_of_Teachers Students_list.csv
```

As an alternative to the disadvantage of the JupyterHub administration rights for all lecturers, it would also be conceivable to assign a JupyterHub server for each lecturer in separate LXC containers instead of one JupyterHub server with multiple lecturers in one LXC container (linuxcontainers.org 2020). In this approach the corresponding lecturer could get also extended administration rights with the parameter

```
JupyterHub.admin_access = True
```

to be able to access the Jupyter notebooks of the students in his lecture and to be able to respond to questions or problems better and easier.

In summary, it can be said that the approach of this project is successful and can be used sustainably with the planned reorganisation of the JupyterHub server to be operated by the computing center. Further and additional research could also be done with this JupyterHub environment, e.g. with regard to automatic hint generation with support for faulty or partial solutions by evaluating solutions from previous years in order to form the basis for a data-driven algorithm (Rivers 2017).

References

- Barba, L., Barker, L., Blank, D., Brown, J., Downey, A., George, T., Heagy, L., Mandli, K., Moore, J., Lippert, D., Niemeyer, K., Watkins, R., West, R., Wickes, E., Willing, C., Zingale, M. 2019. Teaching and Learning with Jupyter. [Cited 7 Jan 2020]. Available at: <https://github.com/jupyter4edu/jupyter-edu-book>
- Hamrick, J. 2015. Version 0.1.0 of nbgrader released! In: Google Groups Forum Teaching with Jupyter Notebooks. [Cited 7 Jan 2020]. Available at: <https://groups.google.com/forum/#!topic/jupyter-education/e4pmPsCR9Jk>
- Hamrick, J. 2016. nbgrader v0.3.0 released. In: Google Groups Forum Teaching with Jupyter Notebooks. [Cited 7 Jan 2020]. Available at: https://groups.google.com/forum/#!msg/jupyter-education/gS0oqY_rK8s/IEFGigjNAAAJ
- Hamrick, J. 2017. nbgrader v0.4.0. In: Github Project Jupyter Repository nbgrader. [Cited 7 Jan 2020]. Available at: <https://github.com/jupyter/nbgrader/releases/tag/v0.4.0>
- Hamrick, J. 2019. nbgrader v0.6.0. In: Github Project Jupyter Repository nbgrader. [Cited 7 Jan 2020]. Available at: <https://github.com/jupyter/nbgrader/releases/tag/v0.6.0>
- Hochschule Augsburg. 2017. Projekt Digital und Regional – Ausgezeichnetes Lehrkonzept: Studiengang Systems Engineering erhält VDMA-Sonderpreis „Bestes Maschinenhaus 2017“ [Cited 7 Jan 2020]. Available at: <http://www.digital-und-regional.de/aktuelles.php?p=121>
- Hochschule Augsburg. 2020. LDAP Anbindung des Rechenzentrums der Hochschule. [Cited 7 Jan 2020]. Available at: <https://www.hs-augsburg.de/Rechenzentrum/RZ-Account.html>
- Jacob, D., Klever, N., Thalhofer, U. 2017. Studiengang Systems-Engineering im Projekt Digital und Regional. In: Meissner, B., Walter, C., Zinger, B. (eds.) Tagungsband zum 3. Symposium zur Hochschullehre in den MINT-Fächern. DiNa 09/2017. ISSN 1612-45377. Nürnberg: Technische Hochschule Nürnberg Georg Simon Ohm - Ingolstadt: DiZ - Zentrum für Hochschuldidaktik Ingolstadt. 123-128. [Cited 7 Jan 2020]. Available at: https://www.diz-bayern.de/images/cwattachments/381_c76aeb3b834947e657aef252c88f6e62.pdf
- Jacob, D., Hiemer, A., Klever, N., Schurk, H.-E., Thalhofer, U. 2018. Systems Engineering – Genial Digital und Regional studieren. In: Waldherr, F., Walter, C., Tagungsband Forum der Lehre 2018. DiNa 11/2018. Ingolstadt: Technische Hochschule Ingolstadt. 42-47. [Cited 7 Jan 2020]. Available at: https://www.diz-bayern.de/images/cwattachments/444_dc48ac19616df67379786f19ec1e5a1f.pdf
- Klever, N., Rösch, P. 2016. Jupyter Notebook. In: Hochschule Augsburg. Lehre Digital – Herausforderungen und Möglichkeiten. 13.-14.1.2016. Augsburg, 2016. [Cited 7 Jan 2020]. Available at: <https://klever.hs-augsburg.de/nb/OWL/Jupyter%20Notebook.pdf>
- Klever, N. 2017. Extending ldapauthenticator to allow arbitrary LDAP search-filters #24. [Cited 7 Jan 2020]. Available at: <https://github.com/jupyterhub/ldapauthenticator/pull/24>
- Klever, N. 2019. timestamp_format raises an exception #1249. [Cited 7 Jan 2020]. Available at: <https://github.com/jupyter/nbgrader/pull/1249>
- Klever, N., 2020a. Kurzanleitung Jupyter Notebook, [Cited 7 Jan 2020]. Available at: <https://klever.hs-augsburg.de/nb/OWL/Hinweise%20und%20Anmerkungen%20zu%20Jupyter%20Notebooks.ipynb>
- Klever, N. 2020b. JupyterHub Server. [Cited 7 Jan 2020]. Available at: <http://python.hs-augsburg.de>
- Klever, N. 2020c. ldap2localuserauthenticator. [Cited 7 Jan 2020]. Available at: <https://github.com/nklever/ldap2localuserauthenticator>
- Klever, N. 2020d. createMultipleCourse.py. [Cited 7 Jan 2020]. Available at: <https://github.com/nklever/nbgrader/blob/master/tools/createMultipleCourse.py>
- Libre Text Project. 2020. Open Textbook Pilot Program. [Cited 7 Jan 2020]. Available at: <https://jupyter.libretexts.org>
- linuxcontainers.org. 2020. Infrastructure for container projects - LXC, LXD und LXCFS. [Cited 7 Jan 2020]. Available at: <https://linuxcontainers.org/>
- Project Jupyter. 2020a. Jupyter Notebook Documentation. [Cited 7 Jan 2020]. Available at: <https://jupyter-notebook.readthedocs.io/en/latest/>

Project Jupyter. 2020b. Project Jupyter. [Cited 7 Jan 2020]. Available at: <https://jupyter.org>

Project Jupyter. 2020c. JupyterHub Documentation. [Cited 7 Jan 2020]. Available at: <https://jupyterhub.readthedocs.io/en/stable/>

Project Jupyter. 2020d. Nbgrader Documentation. [Cited 7 Jan 2020]. Available at: <https://nbgrader.readthedocs.io/en/stable/>

Project Jupyter. 2020e. Example Use Case: Multiple Classes. In: Nbgrader Documentation. [Cited 7 Jan 2020]. Available at: https://nbgrader.readthedocs.io/en/stable/configuration/jupyterhub_config.html#example-use-case-multiple-classes

Rivers, K. 2017. Automated Data-Driven Hint Generation for Learning Programming. PhD Thesis. Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University. Pittsburgh. [Cited 7 Jan 2020]. Available at: <http://krivers.net/files/thesis.pdf>

University of California, Davis. 2020. Libre Text Project. [Cited 7 Jan 2020]. Available at: <https://libretexts.org/official-PressRelease.pdf>

Wikipedia. 2020. LDAP - Lightweight Directory Access Protocol. [Cited 7 Jan 2020]. Available at: https://de.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol