

OWL – Offene digitale Werkzeuge in der Lehre

Teil I - Jupyter Notebook. JupyterHub und Nbgrader

Nik Klever, Fakultät für Informatik der Hochschule Augsburg

Projektbeschreibung

Motivation

In einer modernen Hochschullehre sind Studierende angehalten, eigeninitiativ und selbstreguliert zu lernen. Dies gilt insbesondere auch für e-Learning-Angebote. Dazu ist ein Umfeld erforderlich, das es erlaubt, jederzeit das eigene Wissen und die eigenen Kompetenzen zu testen, also ein Self-Assessment vorzunehmen. Mit Hilfe digitaler Methoden ist es möglich, individualisierte Aufgaben zu stellen, studentische Abgaben automatisch zu testen und entsprechendes Feedback zu geben. Das Projekt „OWL – Offene digitale Werkzeuge in der Lehre“ ist aus dieser Idee heraus entstanden und dient

- der Unterstützung der Studenten durch Self-Assessment Plattformen
- dem E-Learning zum autarken Lernen
- dem generierten Feedback
- der verringerten Belastung bei Abgabeterminen
- den Plattformen für einfache Interaktion und Abgabemöglichkeit inkl. Bonus-Tests

In diesem Teil wurde die Integration von Nbgrader auf der Basis von Jupyter Notebook in einen JupyterHub Server mittels des Authorisierungssystems der Hochschule Augsburg eingebunden, sodass eine Plattform für alle Lehrenden, die an der Nutzung von Jupyter Notebook als Standard-Werkzeug in der Lehre interessiert sind, entstanden und nutzbar ist.

Jupyter Notebook

Jupyter Notebook [1] ist insbesondere für mathematisch-naturwissenschaftliche, statistische, informatik- und programmierlastige aber auch daten- und graphiklastige Inhalte etabliert. Die Integration unterschiedlichster Programmiersprachen wie Python, R, Java, Javascript und vielen weiteren ist ebenso möglich wie die Einbindung von unterschiedlichen Markup-Sprachen wie HTML, Markdown, Scalable Vector Graphics, LaTeX und weiteren. In aller Kürze ist Jupyter Notebook daher eine schnelle und einfache Erstellung dieser unterschiedlichsten Inhalte auf einer Seite im Web. Die Vorteile von Jupyter Notebook sind in einer Kurzanleitung an anderer Stelle eingehender beschrieben und nachlesbar [2].

JupyterHub

JupyterHub ist nun wie Jupyter Notebook ebenfalls im Rahmen des weltweiten Open Source Projekts Jupyter [3] entstanden und als Server dafür zuständig für jeden, der auf diesem Server autorisierten Benutzer, eine Umgebung in einer Webseite auszuliefern, in der eigene Jupyter Notebooks in den Home-Verzeichnissen von den jeweiligen Benutzern erstellt werden können. Der Vorteil von JupyterHub [4] ist daher, dass eine Jupyter Notebook Installation auf dem eigenen

Rechner eines Benutzers nicht mehr notwendig ist. Ein Beispiel für eine intensive Nutzung von JupyterHub ist u.a. das Libre Text Projekt [5] an der University of California mit einem Open Textbooks Pilot Program und der Unterstützung des amerikanischen Department of Education [6].

Nbgrader

Nbgrader [7] ist auf der Basis von Jupyter Notebook entstanden und erlaubt die Erstellung von Jupyter Notebooks mit Aufgabenstellungen deren Lösungen den Studierenden in einer von den Dozenten freigegebenen Kopie verborgen werden, dennoch von ihnen bearbeitet werden können. Nach der entsprechenden Bearbeitung dieser Kopie können die Studierenden ihre Kopie als Abgabe einreichen. Der Dozent kann nach dem Abgabedatum diese Abgaben für die gesamte Klasse einsammeln, entweder automatisch (derzeit nur für Python) oder manuell (für alle Programmiersprachen oder reine Textfragen) in einer einfachen Art und Weise korrigieren, mit einem Feedback versehen und bepunktet den Studierenden wieder zur Verfügung stellen.

Teaching and Learning with Jupyter

Die Idee und die Motivation für diesen Teil des OWL Projekts basiert auf Arbeiten über Jupyter Notebook für die Lehre, die bereits im Rahmen des ersten Tags der Lehre „Lehre Digital – Herausforderungen und Möglichkeiten“ Anfang 2016 in einer Poster Session [8] den Teilnehmern vorgestellt wurden. Inzwischen hat sich Jupyter Notebook sowohl als Tool für den Austausch von wissenschaftlichen Ergebnissen und Veröffentlichungen etabliert und hat zusammen mit Nbgrader und weiteren Tools auch einen großen Einfluss auf die Lehre. Ein deutlicher Hinweis hierfür zeigt sich insbesondere in dem 2019 veröffentlichten Online-Buch „Teaching and Learning with Jupyter“ [9].

Zielerreichung

Jupyter Notebook

Nachdem Jupyter Notebook bereits seit einigen Jahren in der Lehre erfolgreich von den Antragstellern genutzt worden ist, konnte auf eine ausreichende Anzahl von Jupyter Notebooks insbesondere in der Programmierung mit Python zurückgegriffen werden. Hierzu zählen die Veranstaltungen Netzwerkprogrammierung, Datenmanagement 1 und Bildverarbeitung im Studiengang Interaktive Medien, Informatik 1 und 2 im Studiengang Systems Engineering, Programmieren 3 und Bildverarbeitung im Studiengang Informatik.

JupyterHub

Für die Erstellung und Nutzung der Jupyter Notebooks in den Veranstaltungen Informatik 1 und 2 des ersten Jahrgangs im Studiengang Systems Engineering wurde bereits vor Beginn des Projekts ein JupyterHub-Server [10] aufgesetzt und eingesetzt, sodass auf diese Erfahrungen aufgebaut werden konnte. Der Studiengang Systems Engineering ist insbesondere infolge seiner Mischung aus digitaler Lehre und Präsenzlehre und seinen modernen Lehrformen wie Flipped Classroom sowie seinen Veranstaltungsorten Nördlingen und Memmingen für gleichzeitig parallel stattfindende Veranstaltungen ideal für diesen Zweck geeignet. Zudem wurde das Konzept dieses Studiengangs

zum einen mit einem Sonderpreis des VDMA-Preises „Bestes Maschinenhaus 2017“ [11] und zum zweiten mit dem Preis für herausragende Lehre des bayerischen Staatsministeriums für Wissenschaft und Kunst [12] bereits zweimal ausgezeichnet. Eine ausführliche Beschreibung des Konzepts für diesen Studiengang findet sich in [13] Für diesen ersten JupyterHub-Server wurde im wesentlichen die Anbindung an das Authorisierungssystem [14] der Hochschule Augsburg [15] zum einen über eine bereits integrierte Erweiterungen des LDAP-Authenticators von JupyterHub [16] sowie zusätzlich einen auf dem LDAP-Authenticator aufbauenden eigenen LDAP2LocalUserAuthenticator [17] durchgeführt. Letzterer erzeugt lokale Benutzer mit entsprechenden Verzeichnissen auf dem lokalen Filesystem, sodass hier eine Trennung der normalen Benutzerdaten von denen über den JupyterHub erfolgt, dennoch die Authentifizierung über das Authentifizierungssystem der Hochschule erfolgt. Die Erfahrungen hierzu als auch die Rückmeldungen der Studierenden waren grundsätzlich positiv, allerdings wurde die Herausgabe und das Einsammeln von Übungsaufgaben-Notebooks noch mit eigenen Skripten durchgeführt und eine individuelle Korrektur der Übungsaufgaben wurde nur auf Anfrage mit den jeweiligen Studierenden besprochen.

Nbgrader

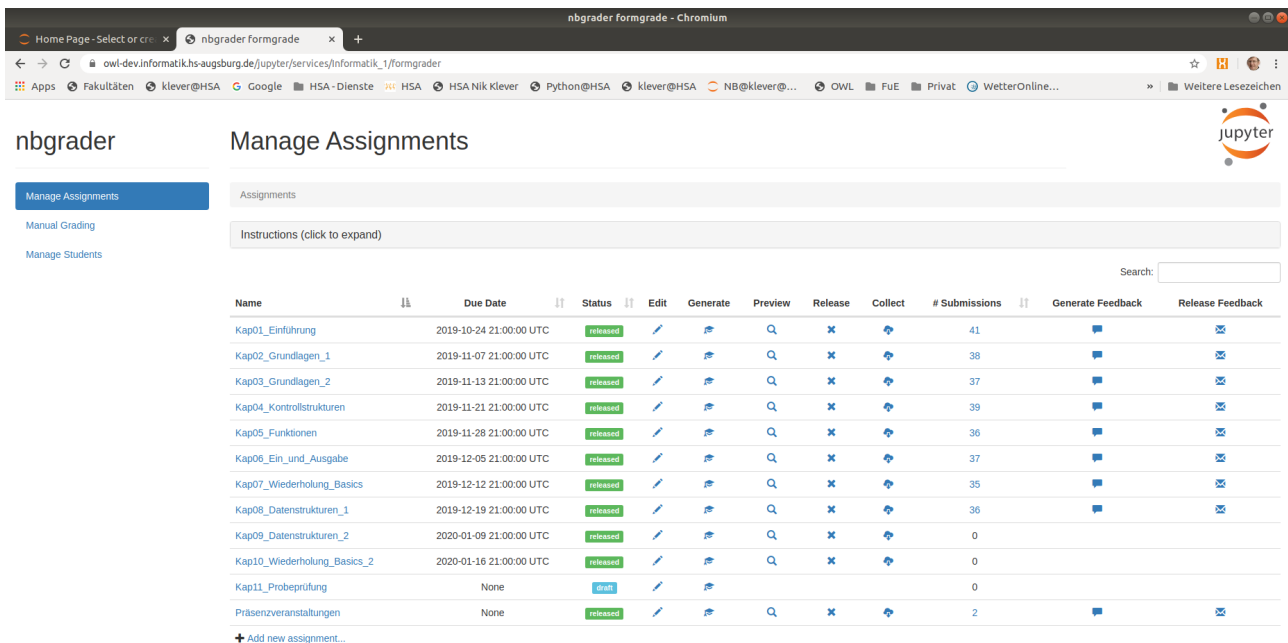
Die erste öffentlich zugängliche Version von Nbgrader war die Version 0.1.0 vom 17.7.2015 [18]. Bereits die Version 0.3.0 [19] vom 25.6.2016 wurde für die ersten Tests im Rahmen dieses Projekts verwendet. In den produktiven Test im Rahmen der Veranstaltungen Informatik 1 und 2 für den dritten Jahrgang des Studiengangs Systems Engineering wurde dann im WS 2018/19 die Version 0.4.0 [20] vom 6.2.2017 eingesetzt und erhielt ebenfalls wieder positive Rückmeldungen, insbesondere infolge der automatisierten und manuell überprüften Feedback Rückmeldungen an die Studierenden. Allerdings mussten auch hierfür wieder eigene Skripte geschrieben werden und die Studierenden konnten noch alle verfügbaren Veranstaltungen gleichzeitig benutzen. Erst mit der Version 0.6.0 [21] war es möglich, verschiedene Veranstaltungen von unterschiedlichen Dozenten und unterschiedlichen Studierenden zu konfigurieren und aufzusetzen. Diese Version wurde im WS 2019/20 für den vierten Jahrgang des Studiengangs Systems Engineering eingesetzt und kann damit erst als Grundlage für eine an der Hochschule Augsburg allgemein nutzbare Plattform dienen. Entsprechend der intensiven Nutzung konnten auch für diese Version bereits entsprechende Patches zu Beseitigung von Bugfixes eingebracht werden [22]. Nachteil dieses Konzepts ist, dass alle Dozenten als JupyterHub-Admin-User registriert sein müssen, was bedeutet, dass alle Dozenten sowohl Benutzer anlegen als auch löschen können – jedoch nur in der Datenbank des JupyterHub Servers. Notwendig ist jedoch hierzu unbedingt, dass in der Konfiguration des JupyterHubs der Parameter

```
JupyterHub.admin_access = False
```

gesetzt wird. Die Löschung eines Benutzers in der Datenbank lässt sich jedoch relativ einfach wiederherstellen, sodass dieser Nachteil nicht so gravierend ist.

Beispiel für die Nutzung des JupyterHub in Verbindung mit Nbgrader für die Veranstaltung Informatik 1 im Studiengang Systems Engineering

Die Veranstaltungen Informatik 1+2 im Studiengang Systems Engineering sind eine Einführung in die Programmierung mit Python. Hierzu wurden für die Veranstaltung Informatik 1 insgesamt 12 Kapitel – in Nbgrader Assignments genannt – mit jeweils einer unterschiedlichen Anzahl Jupyter Notebooks erstellt (s. Abb. 1).



The screenshot shows the 'Manage Assignments' page in the Nbgrader interface. It features a table with 12 rows, each representing an assignment. The columns include Name, Due Date, Status, Edit, Generate, Preview, Release, Collect, # Submissions, Generate Feedback, and Release Feedback. The assignments are listed as follows:

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_Probeprüfung	None	draft						0		
Präsenzveranstaltungen	None	released						2		

Abb. 1 Manage Assignment Sicht mit allen 12 Kapiteln der Veranstaltung Informatik 1

Zudem wurden für fast alle Notebooks zwischen 1 und 4 Videos als Screenshots mit zu dem jeweiligem Stoff passenden Programmierbeispielen ebenfalls in Jupyter Notebooks entwickelt und einschliesslich der Erklärungen für die jeweiligen Schritte aufgenommen. In Abb. 2 ist das Beispiel eines Notebooks mit jeweils einer **Autograded answer** und einer **Autograder tests** Zelle aus Dozentensicht aufgenommen, während in Abb. 3 das gleiche Notebook aus der Sicht der Studierenden abgebildet ist.

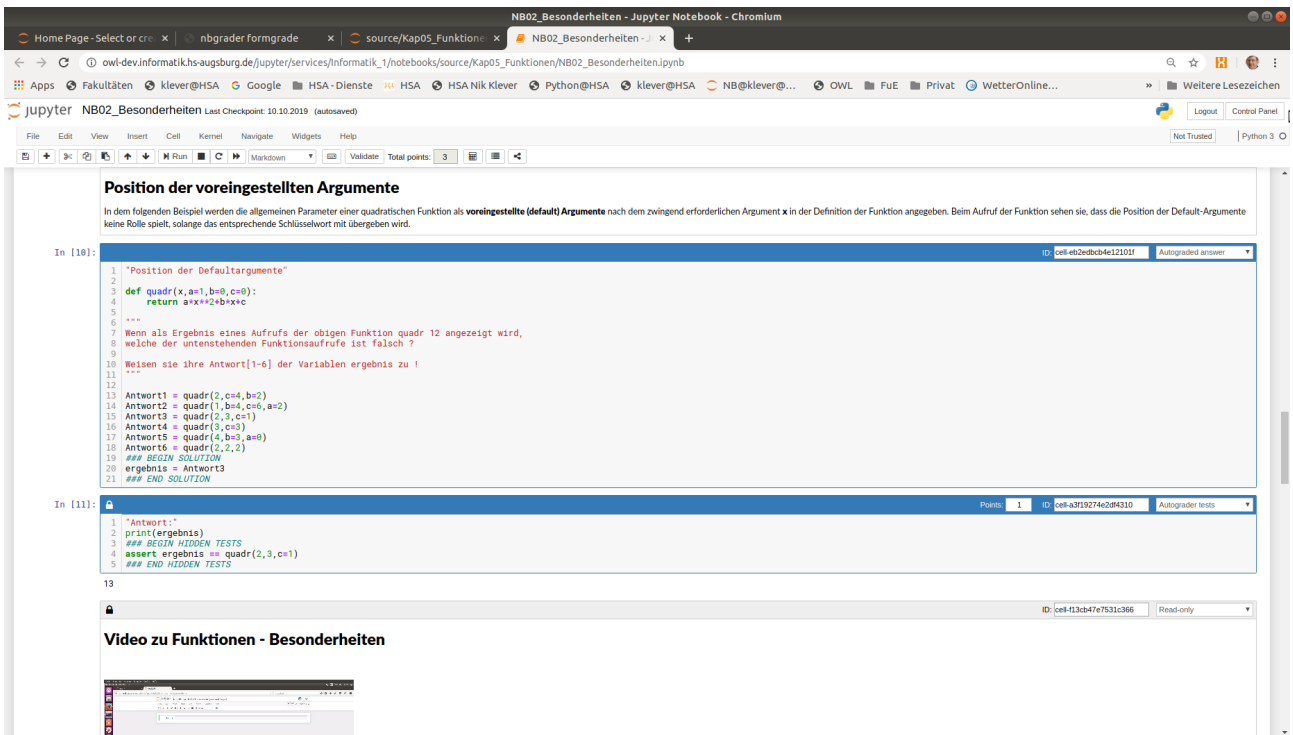


Abb. 2 Ausschnitt aus einem Jupyter Notebook mit einer Autograded answer und einer Autograder tests Zelle (aus Dozentensicht)

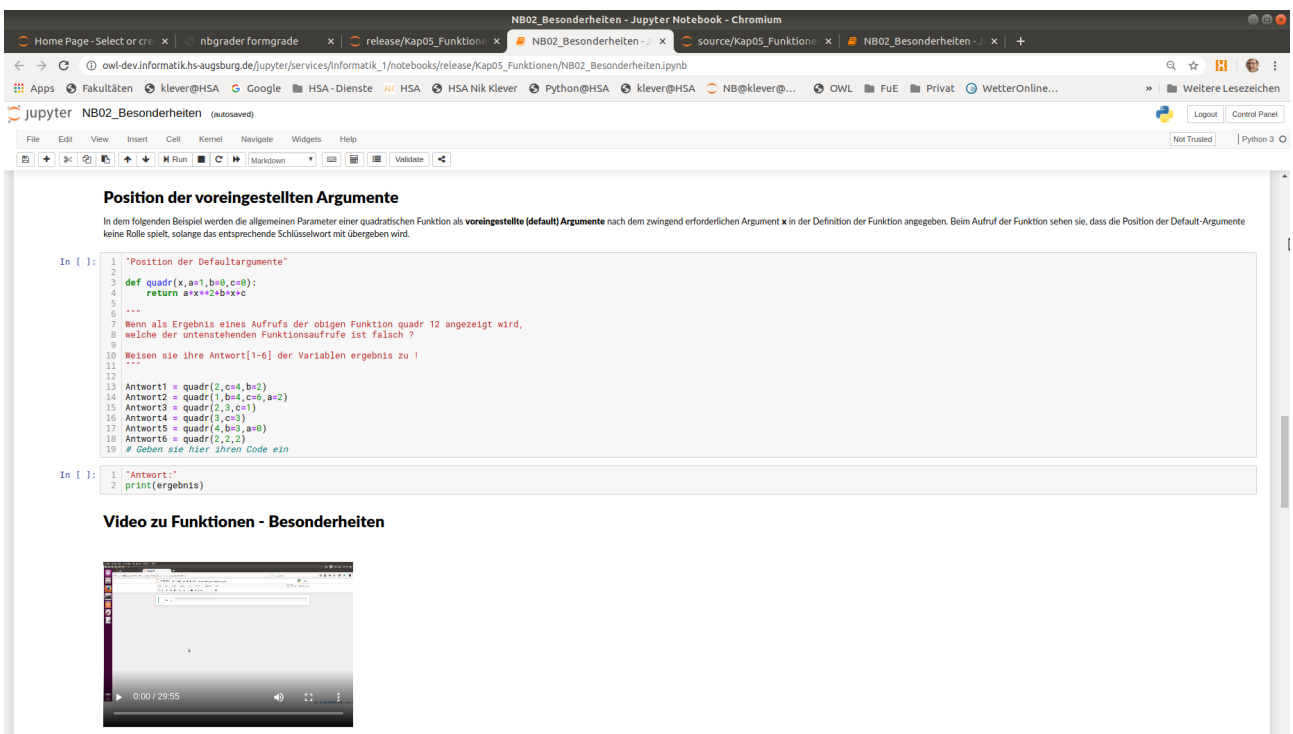


Abb. 3 Dem in Abb. 2 entsprechenden Ausschnitt aus dem Jupyter Notebook diesmal aus Studierendensicht

Die Studierenden können in ihrer Bearbeitung des Notebooks jederzeit eine Validierung des Notebooks durchführen, die bei korrekter Bearbeitung grün aufleuchtet, ansonsten rot. Um den Studierenden daher eine leichte Selbstüberprüfung zu ermöglichen, sollten die Aufgaben in minimalste Schritte aufgeteilt werden, sodass immer klar wird, wann und wo ein Fehler bei einer nicht validierenden Aufgabe auftritt. Dies ist nicht immer einfach und manchmal auch unmöglich. Wenn der/die Studierende seine/ihre Bearbeitung abgegeben hat, dann erhalten sie über

Autograding beispielhaft ein Feedback als HTML-Seite einschliesslich der erreichten Punkte, wobei über das Autograding Verfahren entweder keine Punkte oder nur volle Punktzahl vergeben wird (entsprechend der Validierung des Notebooks) (s. Abb. 4)

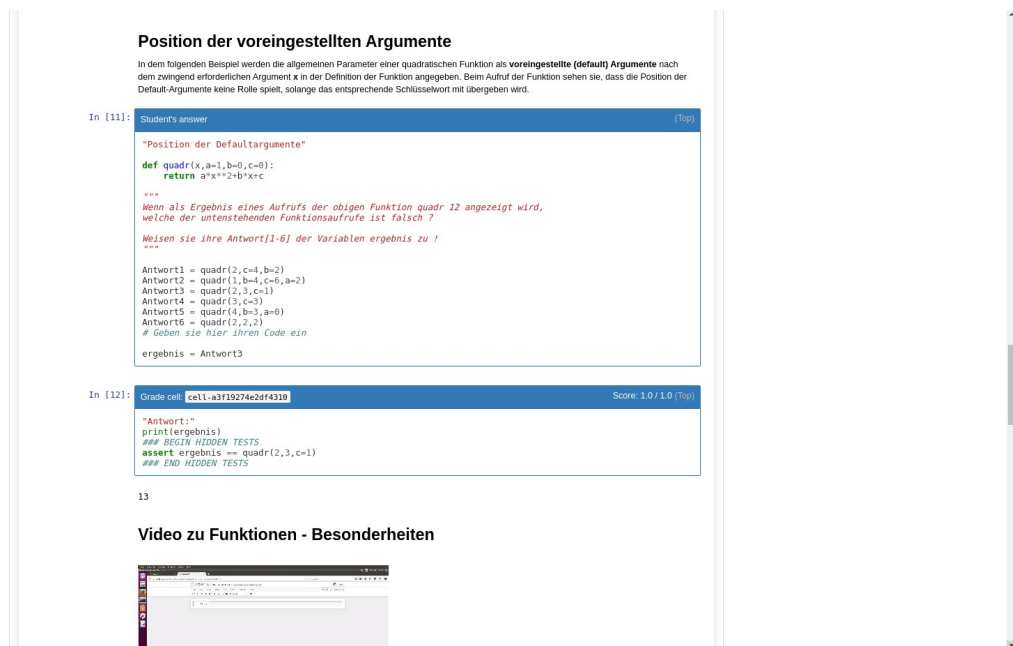


Abb. 4 Wieder dem in Abb. 2 sowie Abb. 3 entsprechenden Ausschnitt aus dem Jupyter Notebook diesmal nach dem Autograding als Feedback an die Studierenden

Für eine korrekte Bepunktung ist daher eine manuelle Nachkorrektur unbedingt vonnöten. Diese ist jedoch aufwändig und kann für sämtliche 30-50 Notebooks einer Veranstaltung im normalen Dozentenalltag nicht geleistet werden. Deshalb wurden für die Veranstaltung Informatik 1 nur die Übungsaufgaben, d.h. pro Assignment ein Notebook manuell nachkorrigiert und für eine Bonusbepunktung verwendet. Sowohl das Feedback zu den einzelnen Notebooks als auch die Bonusbepunktung kam bei den Studierenden sehr gut an und wurde einhellig begrüßt.

Nachhaltigkeit / Verwertung der Ergebnisse

Die positiven Erfahrungen in der Nutzung von Nbgrader durch den Autor in den Veranstaltungen Informatik 1+2 im Studiengang Systems Engineering hat den Autor dazu bewegen den JupyterHub-Server entsprechend für weitere Veranstaltungen und Dozenten zu öffnen und freizugeben. Daher wurden erste Gespräche zur Übernahme der Administration dieses Servers durch das Rechenzentrum der Hochschule Augsburg geführt. Die Einrichtung neuer Veranstaltungen ist insbesondere für die in der Hochschule Augsburg notwendige Umgebung mit einem eigenen Authorisierung-Server komplexer als in den von Nbgrader zur Verfügung gestellten Demos für sogenannte multiple Kurse [23]. Damit weitere Veranstaltungen dennoch in geeigneter Weise ohne größeren Aufwand eingerichtet werden können, wurde dafür ein Skript geschrieben, welches ebenfalls bereits in Github veröffentlicht wurde [24] und für Administratoren eine wesentliche Erleichterung bedeutet. Die Einrichtung einer neuen Veranstaltung läuft damit auf den Aufruf des Skripts mit den beiden Parametern des Namens der Veranstaltung und dem Rechenzentrums-Benutzernamen des entsprechenden Dozenten hinaus:

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

Ein weiterer Schritt ist anschliessend noch notwendig, um alle Studierenden in diesen Kurs einzutragen. Wenn eine Liste der in diesem Kurs eingeschriebenen Studierenden elektronisch mit Vorname, Nachname, E-Mail-Adresse und Benutzeraccount in einer Datei vorhanden ist, so können diese Studierenden ebenfalls über dieses Skript in Nbgrader mit dem folgenden Aufruf in einem Schritt eingetragen werden:

```
$ python createMultipleCourse.py Name_of_Course Useraccount_of_Teacher -s Students_list.csv
```

Infolge des Nachteils der JupyterHub-Administrations-Rechte für alle Dozenten wäre als Alternative auch denkbar, statt eines einzigen JupyterHub-Servers, der im Rahmen dieses Projekts in einem LXC-Container [25] erstellt wurde, für jeden Dozenten einen eigenen LXC-Container zu verwenden, sodass der entsprechende Dozent dann auch erweiterte Administrationsrechte über den Parameter

```
JupyterHub.admin_access = True
```

erhalten kann, um z.B. auf die Jupyter Notebooks der Studierenden in seinen Veranstaltungen zugreifen zu können und damit auf Fragen oder Probleme besser und leichter eingehen zu können.

Zusammenfassend kann ausgesagt werden, dass der Ansatz dieses Projekts vollumfänglich erfüllt worden und mit der geplanten Umstellung auf Rechenzentrumsbetrieb auch nachhaltig weiter verwendet werden kann. Weiterer und zusätzlicher Forschung sollte damit ein vernünftiges Umfeld bereitet worden sein z.B. im Hinblick auf eine automatische Hinweisgenerierung mit einer Unterstützung für fehlerhafte bzw. Teillösungen über die Auswertung von Lösungen aus den Vorjahren um damit die Grundlage für einen Data-Driven-Algorithmus zu bilden [26].

Quellen

1. Project Jupyter, „Jupyter Notebook“, <https://jupyter-notebook.readthedocs.io/en/latest/>
2. Nik Klever, „Kurzanleitung Jupyter Notebook“, <https://klever.hs-augsburg.de/nb/OWL/Hinweise%20und%20Anmerkungen%20zu%20Jupyter%20Notebooks.ipynb>
3. Project Jupyter, „Project Jupyter“, <https://jupyter.org>
4. Project Jupyter, „JupyterHub“, <https://jupyterhub.readthedocs.io/en/stable/>
5. University of California, „Libre Text Project“, <https://libretexts.org/>
6. Libre Text Project, „Open Textbook Pilot Program“, <https://jupyter.libretexts.org>
7. Project Jupyter, „Nbgrader“, <https://nbgrader.readthedocs.io/en/stable/>
8. Nik Klever, Peter Rösch, Jupyter Notebook, Poster, Konferenz „Lehre Digital – Herausforderungen und Möglichkeiten“, Augsburg, 2016
9. Lorena A. Barba et. al, „Teaching and Learning with Jupyter“, <https://github.com/jupyter4edu/jupyter-edu-book>, 6.12.2019
10. JupyterHub Server, <http://python.hs-augsburg.de>
11. Sonderpreis Bestes Maschinenhaus 2017, <http://www.digital-und-regional.de/aktuelles.php?p=121>
12. Preis für herausragende Lehre 2018, <https://diz-bayern.de/tagungen/forum-der-lehre/preistraegerinnen-und-preistraeger>

13. Dirk Jacobs et. al, Studiengang Systems Engineering im Projekt Digital und Regional, pg. 123-128, in Tagungsband 3. MINT Symposium 2017, https://www.diz-bayern.de/images/cwattachments/381_c76aeb3b834947e657aef252c88f6e62.pdf
14. LDAP - Lightweight Directory Access Protocol, https://de.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol
15. LDAP Anbindung des Rechenzentrums der Hochschule Augsburg, <https://www.hs-augsburg.de/Rechenzentrum/RZ-Account.html>
16. Patch für den JupyterHub LDAP Authenticator <https://github.com/jupyterhub/ldapauthenticator/pull/24>
17. LDAP2LocalUserAuthenticator <https://github.com/nklever/ldap2localuserauthenticator>
18. Nbgrader Version 0.1.0 <https://groups.google.com/forum/#!topic/jupyter-education/e4pmPsCR9Jk>
19. Nbgrader Version 0.3.0 https://groups.google.com/forum/#!msg/jupyter-education/gS0oqY_rK8s/IEFGigjNAAAJ
20. Nbgrader Version 0.4.0 <https://github.com/jupyter/nbgrader/releases/tag/v0.4.0>
21. Nbgrader Version 0.6.0 <https://github.com/jupyter/nbgrader/releases/tag/v0.6.0>
22. Patch für Nbgrader <https://github.com/jupyter/nbgrader/pull/1249>
23. Nbgrader Demos Multiple Classes https://github.com/jupyter/nbgrader/tree/master/demos/demo_multiple_classes
24. Tool createMultipleCourse.py <https://github.com/nklever/nbgrader/blob/master/tools/createMultipleCourse.py>
25. Linuxcontainer LXC, LXD und LXCFS <https://linuxcontainers.org/>
26. Kelly Rivers, „Automated Data-Driven Hint Generation for Learning Programming“, PhD Thesis