



Disseminationsstrategien für digitale Publikationen (R 1.2.1)

Version 4. März 2019

Cluster 1

Verantwortlicher Partner MPIWG

DARIAH-DE Überführung der digitalen Forschungsinfrastrukturen für die e-Humanities in die Operational Phase (Betriebsphase)

Dieses Forschungs- und Entwicklungsprojekt wird / wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF), Förderkennzeichen 01UG1610A bis J, gefördert und vom Projektträger im Deutschen Zentrum für Luft- und Raumfahrt (PT-DLR) betreut.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Projekt: DARIAH-DE: Überführung der digitalen Forschungsinfrastrukturen für die e-Humanities in die Operational Phase (Betriebsphase)

BMBF Förderkennzeichen: 01UG1610A bis J

Laufzeit: März 2016 bis Februar 2019

Dokumentstatus final

Verfügbarkeit öffentlich

AutorInnen

Malte Vogl, MPIWG
Hanna-Lena Meiners, SUB
Klaus Thoden, MPIWG
Jan Rohden, SUB
Michael Haft, ADWLM
Oliver Schmid, TUD

Revisionsverlauf:

Datum	Autor	Kommentare
01.02.2019	Malte Vogl	Vorlage
11.02.2019	Malte Vogl	Notebooks
20.02.2019	Jan Rohden	Zusammenfassung
21.02.2019	Hanna-Lena Meiners	Einleitung
22.02.2019	Klaus Thoden	Working Papers
26.02.2019	Jan Rohden, Malte Vogl	Korrekturen
27.02.2019	Malte Vogl	Versand an Konsortium
04.03.2019	Malte Vogl	Endredaktion und Veröffentlichung



Dieses Werk ist unter einer Creative Commons Lizenz vom Typ Namensnennung 3.0 Deutschland zugänglich. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie <http://creativecommons.org/licenses/by/3.0/de/> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.

Inhaltsverzeichnis

1. Einführung	4
1.1. Zielsetzung des AP 1.2	4
1.2. Entwicklung des AP 1.2 in der Förderphase	4
2. Use-Case Working Papers	5
2.1. Einführung	5
2.2. Produktion der Working Papers	5
2.3. Angebot zusätzlicher Formate	7
2.4. Zusammenfassung Working Papers	7
3. Use-Case Jupyter Notebooks	7
3.1. Einführung	7
3.2. Strategie	9
3.3. Editieren	9
3.4. Publizieren	10
3.4.1. Publikation	10
3.4.2. Referee-Workflow	11
3.5. Darstellung	11
3.6. Zusammenfassung Jupyter Notebooks	13
4. Ausblick	13

1. Einführung

1.1. Zielsetzung des AP 1.2

AP 1.2. befasste sich mit neuen Publikationsstrategien für die Geisteswissenschaften. Deren Notwendigkeit ergibt sich nicht zuletzt aus der zunehmenden Digitalisierung der geistes- und kulturwissenschaftlichen Forschung. Daraus resultiert der Bedarf nach einer umfassenden Publikationsstrategie, welche eine offene bzw. hybride Publikation ermöglicht, die Einbindung digitaler Forschungsdaten und interaktiver Ressourcen sowie deren Nachnutzung erlaubt und einer gezielten Dissemination all dieser Forschungsergebnisse zuträglich ist. Dafür konnte auf diverse Vorarbeiten zurückgegriffen werden. Hierzu zählen unter anderem die Erfahrungen mit den DARIAH-DE Working Papers¹, die ein bereits etabliertes Medium darstellen, um einem größeren Publikum Ergebnisse der computergestützten geisteswissenschaftlichen Forschung in digitaler Form unter einer freien Lizenz verfügbar zu machen, und die TextGrid Nutzertreffen, die sich als wertvolles Instrument für den Austausch mit der Nutzercommunity erwiesen haben.

1.2. Entwicklung des AP 1.2 in der Förderphase

Das Arbeitspaket 1.2 hatte in der dritten Förderphase von DARIAH-DE das Ziel, Anforderungen und Spezifikationen zu identifizieren, um während des geisteswissenschaftlichen Publikationsprozesses Forschungsdaten mit einbeziehen zu können. Darauf basierend wurden bereits bestehende Publikations- und Disseminationsstrategien evaluiert bzw. gesichtet, sowie ein Use-Case entwickelt, um die Publikationen der "DARIAH-DE Working Paper" zu vereinfachen und einem einheitlichen Workflow zu unterwerfen. Der neue Workflow wurde unter anderem als Code in GitHub veröffentlicht². Dieser Vorgang wird in Kapitel 2 näher beschrieben. In einem Workshop wurden die gesammelten Erkenntnisse und Erfahrungen zum Thema Erweitertes Publizieren an eine interessierte Fachcommunity weitergegeben sowie Empfehlungen ausgesprochen. Leitfragen und Themen waren hierbei:

- die Einbindung von Forschungsdaten in die Publikation,
- die stabile Referenzierung von Publikationen und Publikationsteilen und
- die frühzeitige Dissemination von Forschungsergebnissen.

Die Ergebnisse des Workshops sind zusammengefasst in zwei Blogbeiträgen³ sowie

¹<https://de.dariah.eu/working-papers>

²<https://github.com/DARIAH-DE/dwp-template>

³<https://dhd-blog.org/?p=8696> und <https://dhd-blog.org/?p=10799>

einer kurzen Zusammenfassung auf dem DARIAH-DE Portal⁴. Aus dem Workshop und ausgehend von den weiteren Erfahrungen aus den vorherigen Förderphasen wurde neben der Strategie des Erweiterten Publizierens via DARIAH-DE Working Paper eine weitere Publikationsstrategie mittels Jupyter Notebooks entwickelt. Diese wird in Kapitel 3 detailliert beschrieben, bringt dieses Format doch einige Besonderheiten mit sich. Somit wurden in der letzten Förderphase von DARIAH-DE schwerpunktmäßig zwei Use-Cases erforscht und umgesetzt, die beide bereits in der Praxis erprobt sind und z.T. ausgiebig genutzt werden.

2. Use-Case Working Papers

2.1. Einführung

Die DARIAH-DE Working Papers sind eine seit 2013 existierende Publikationsform, die über die Entwicklungen im Kontext von DARIAH-DE und auch darüber hinaus berichtet. Die Beiträge werden im Open Access Format mit einer CC-BY-Lizenz⁵ im DARIAH-DE-Portal⁶ veröffentlicht und auch gleichzeitig im Repositorium der Universität Göttingen archiviert sowie mit URN⁷ versehen. Die Beiträge werden vier verschiedenen Reihen zugeordnet, "Berichte und Empfehlungen", "Tagungen und Konferenzbeiträge", "Dokumentationen und Lehrmaterialien" sowie "DARIAH plus", wobei der Großteil der Beiträge in den Reihen "Berichte und Empfehlungen" und "DARIAH plus" zu finden ist. Bislang sind 28 Working Papers erschienen.

2.2. Produktion der Working Papers

Der ursprüngliche Workflow sah vor, dass AutorInnen die Beiträge in gängigen Office-Dokumenten auf Grundlage von Vorlagen einreichen konnten, woraus dann in der Redaktion eine PDF-Datei generiert wurde. Dieser Prozess war schwierig und führte zu einer langen Bearbeitungszeit der einzelnen Working Papers.

Im Laufe des Jahres 2016 wurde mit einer Neuentwicklung begonnen, um den bisherigen Workflow abzulösen. Grundlage des Formats ist *Markdown*⁸, eine Auszeichnungssprache, die auf einfachem Text mit minimalen Codes für typographische Phänomene wie Kursiv oder Textgliederung basiert. Der Vorteil für die AutorInnen ist, dass

⁴<https://de.dariah.eu/erweitertes-publizieren-workshop>

⁵<https://creativecommons.org/licenses/by/3.0/de/>

⁶<https://de.dariah.eu/working-papers>

⁷Uniform Resource Name

⁸<http://pandoc.org/MANUAL.html#pandocs-markdown>

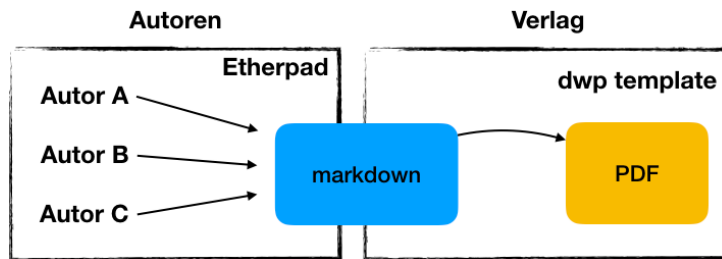


Abbildung 1: Workflow für die kollaborative Erstellung eines Working Papers.

damit beim Schreiben der reine Inhalt eines Textes in den Vordergrund gestellt wird und die Formatierung und Seitenaufteilung weniger wichtig wird. Außerdem stehen für alle Rechnerplattformen Programme zur Verfügung, die eine Ansicht bereitstellen, die an konventionelle Textverarbeitungsprogramme angelehnt ist. Über Dienste wie Etherpad, wie er zum Beispiel von der GDWG angeboten wird, ist es gleichzeitig möglich Markdown-Text direkt im Browser zu schreiben. Dadurch wird auch kollaboratives Arbeiten ermöglicht.

Wenngleich sich die Form der Texteingabe geändert hat, wird am Ende eine PDF-Datei erzeugt. Dabei wird das Programm *pandoc*⁹ verwendet, das im Hintergrund wiederum \LaTeX zum Setzen des Textes aufruft. Über die in \LaTeX integrierten Programme zur Literaturverwaltung ist es auch möglich, eine Bibliographiedatenbank zu verwenden, so dass Literaturhinweise nicht manuell formatiert werden müssen, sondern automatisch in dem Zitationsstil der Working Papers ausgegeben werden.

Diese gesamte Toolchain wurde dann in dem Projekt *dwp-template*¹⁰ zusammengefasst, das das Erstellen des PDF aus dem Markdown-Dokument und der Bibliographiedatenbank in einem Kommando bündelt. Gleichzeitig ist es auch als Dokumentation der Richtlinien für AutorInnen gedacht, um die Befehle in Markdown richtig einzusetzen. Gängige Office-Anwendungen funktionieren nach der WYSIWYG¹¹-Methode, so dass die BenutzerInnen immer kontrollieren können, wie das Endergebnis aussieht. Beim Schreiben in Markdown allerdings fehlt die Information, wie der Text letztendlich als fertiges Produkt, also beispielsweise als gesetztes PDF, aussieht. Bei der Entwicklung des *dwp-template* wurde daher darauf geachtet, dass auch die gesamte Redaktionsumgebung lokal installiert werden kann, so dass die BenutzerInnen selbst ein PDF aus der Markdown-Vorlage erzeugen können. Damit ist eine ständige Kontrolle über das spätere Erscheinungsbild möglich.

⁹<http://pandoc.org/>

¹⁰<https://github.com/DARIAH-DE/dwp-template>

¹¹“What you see ist what you get.”

2.3. Angebot zusätzlicher Formate

Dadurch, dass die Dokumentenquelle nun in einem sehr wandelbaren Format vorliegt, wurden technische Hürden gemindert, die Working Papers auch in anderen Formaten als nur PDF anzubieten. Der *dwp-template*-Workflow folgt mit der Verwendung von Markdown und Pandoc einem Single-Source-Publishing Modell, also einem Paradigma, ohne viel Aufwand aus einem zentralen Dokumentenformat verschiedene Ausgabeformate zu erzeugen. Denkbar wären hier einerseits Formate wie epub, ein Format, das gut auf Ebookreadern oder Tablets funktioniert und gegenüber PDF den Vorteil besitzt, dass der Text an die Bildschirmgröße angepasst werden kann und sich leichter umbrechen lässt. Andererseits können aber auch Formate wie HTML oder XML erstellt werden, wodurch die Working Papers auch direkt im Browser ansehbar sind und dadurch bessere Möglichkeiten zur Vernetzung bilden.

2.4. Zusammenfassung Working Papers

Durch die Umstellung des Workflows auf ein Single-Source-Publishing Modell konnte die redaktionelle Bearbeitungszeit deutlich gesenkt werden. In Verbindung mit der Möglichkeit mittels Etherpad kollaborativ an Markdown-Texten zu arbeiten, konnte insgesamt die Zeitspanne zwischen Konzipierung eines Working Papers und seiner Veröffentlichung verkürzt werden. Zusammen mit der Verfügbarkeit als Open-Access Dokumente ist dies ein wichtiger Schritt, um Forschungsergebnisse zeitnah einer breiten Öffentlichkeit zugänglich zu machen.

3. Use-Case Jupyter Notebooks

3.1. Einführung

Digitale Methoden entwickeln sich zunehmend zur alltäglichen Praxis in den Geisteswissenschaften. Von der Erstellung einer Sammlung, über Bildanalysen, Text-Annotierung, bis hin zu Netzwerk-Analysen für die Ideengeschichte, keiner dieser Bereiche wird heutzutage noch ohne Computer angegangen. Durch diese wachsende Bedeutung treten Themen, die ursprünglich in der Software-Entwicklung angesiedelt waren, in den Fokus der Geisteswissenschaften. So stellt sich zum Beispiel die Frage der Nachhaltigkeit einer Software-Entwicklung für ein geisteswissenschaftliches Forschungsvorhaben oder welche Best-Practice-Maßstäbe dabei verfolgt werden sollten. Eine relativ neue Entwicklung im Bereich der Software-Entwicklung ist das sogenannte „literate programming“, wobei ein Programm eher narrativ mit starkem Fokus auf die Lesbarkeit

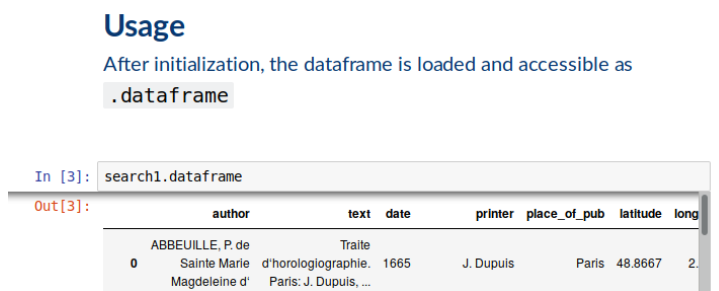


Abbildung 2: Ein Jupyter Notebook verbindet Dokumentation und Programm-Code.

des Codes geschrieben wird. Dieses Paradigma ist besonders für wissenschaftliche Fragestellungen in der Lehre interessant.

Ein Vorreiter in diesem Gebiet sind Jupyter Notebooks¹², oder kurz Notebooks, in denen neben Quellcode auch Markdown- oder \LaTeX -Texte erlaubt sind. Die grundlegende Idee einer Notebook-Veröffentlichung ist die enge Verzahnung von Code, der meistens in Python, Julia oder R geschrieben wurde, mit den zugehörigen Erklärungen und Motivierungen. In einer interaktiven Ansicht können die LeserInnen des Textes dann zum einen den Code ausführen, zum anderen diesen aber auch anpassen um beispielsweise statistische Behauptungen durch eine andere Parameterwahl zu überprüfen oder eine Berechnung mit anderen Startwerten noch einmal auszuführen. Ebenso können Wissenschaftler die Erstellung eines Referenzdatensatzes mit allen Schritten der Daten-Akquise und Selektion sowie Bereinigung dokumentieren. Kombiniert mit öffentlich verfügbaren Forschungsdaten können Jupyter Notebooks einen wichtigen Beitrag zu reproduzierbarer Forschung leisten.

Durch eine einfache Gliederungsstruktur mit Kapiteln oder Absätzen kann um einen Programm-Code eine vollwertige Veröffentlichung samt Bibliographie geschrieben werden. Bereits 2014 wurde dies durch einen Nature Artikel¹³ beispielhaft gezeigt. Inzwischen wird die Umgebung zum Beispiel auch von Verlagen wie O'Reilly¹⁴ unterstützt. Eine kuratierte Sammlung von Notebooks zu wissenschaftlichen Fragestellungen findet man zum Beispiel auf Github¹⁵.

Für eine größere Akzeptanz dieser Veröffentlichungsart sind zwei wichtige Aufgaben von Forschungsinfrastrukturen zu lösen. Zum einen muss für die Unterstützung von Jupyter Notebook-Publikationen geworben werden, bspw. sollten in den Einrichtungen eigene Darstellungsplattformen für Wissenschaftler bereitgestellt werden. Als Vorreiter kann hier zum Beispiel die Edition Topoi genannt werden, die Notebooks be-

¹²<https://jupyter.org/>

¹³<https://www.nature.com/news/interactive-notebooks-sharing-the-code-1.16261>

¹⁴<https://www.oreilly.com/topics/jupyter>

¹⁵<https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>

reits zitier-fähig mit den zugehörigen Daten publiziert¹⁶. Für das Erstellen solcher Erweiterter Publikationen im Geiste des Open-Access¹⁷, sollten Forschungseinrichtungen für Mitglieder eine Editions-umgebung für Notebooks anbieten, die das Teilen von Programmbeispielen und/oder Datensätzen erlaubt. Eindrucksvolle Umsetzungen hierfür findet man zum Beispiel am Cern¹⁸ oder bei der GWDG¹⁹.

3.2. Strategie

Ein Ansatz um den Einstieg in das Erweiterte Publizieren mit Jupyter Notebooks für Forschungseinrichtungen zu erleichtern, ist die Bereitstellung von „Publikations-Bausätzen“ für die Darstellung von Notebooks, zum Beispiel basierend auf Container-Technologie. Auf ähnliche Weise soll ein „Editier-Bausatz“ für WissenschaftlerInnen angeboten werden, der einen möglichst barrierefreien Einstieg in das „literate programming“ ermöglicht.

3.3. Editieren

Ein im Rahmen einer Max Planck Research Group²⁰ am MPIWG²¹ entwickelter Prototyp einer Editions-umgebung²² wird im Folgenden beschrieben. Die Idee ist eine einfache Bereitstellung einer Editions-umgebung für alle DARIAH-DE NutzerInnen, wobei die Authentifizierung über die Dariah-AAI²³ erfolgen sollte. In dieser Editions-umgebung steht damit automatisch eine Sammlung von Beispiel-Notebooks zur Verfügung, die als Grundlage für eigene Forschungsarbeiten dienen können. Dabei werden typische Themen der DH abgedeckt, wie etwa das Laden von Webressourcen, die Analyse für Netzwerke historischer Persönlichkeiten oder das Visualisieren von Orten oder geographischen Verteilungen, z.B. Fundort-Datenbanken, auf Landkarten. Ein besonderer Fokus liegt hierbei auf der Integration von weiteren DARIAH-DE Diensten wie dem Repository, dem Geobrowser oder dem Text-Analyse-Tool DARIAH-DE Topic.

Die Arbeits-umgebung ist als Plattform für das kollaborative Arbeiten an Jupyter Notebooks gedacht. Dabei werden zwei Ziele parallel verfolgt: Hauptsächlich soll die Erstellung von vollwertigen *Erweiterten Publikationen* ermöglicht werden. Gleichzeitig soll den

¹⁶<https://dx.doi.org/10.17171/2-2-281-1>

¹⁷<http://www.topoi.org/news/open-access-forschungsdaten-uneingeschrankt-und-zitierfahig-publizieren/>

¹⁸<https://swan.web.cern.ch/content/basic-examples>

¹⁹<https://www.gwdg.de/application-services/jupyter>

²⁰<https://www.mpiwg-berlin.mpg.de/computational-history>

²¹<https://www.mpiwg-berlin.mpg.de/>

²²<https://c105-188.cloud.gwdg.de:442>

²³<https://wiki.de.dariah.eu/display/publicde/DARIAH+AAI+Documentation>

Forschenden die Möglichkeit gegeben werden, während des Prozesses der Erstellung entstehende Programme oder Programm Teile mit anderen Wissenschaftlerinnen zu teilen. Dies könnten zum Beispiel Programme für Datenmanagement und -bereinigung, Bild- oder Text-Erkennung oder Visualisierungen sein.

Die Editions Umgebung für Jupyter Notebooks baut auf der JupyterHub Plattform²⁴ auf, die Arbeitsumgebungen für angemeldete Nutzer bereitstellt. Das kollaborative Arbeiten wird durch ein NextCloud-Daten-Backend²⁵ ermöglicht. Die Authentifizierung der Benutzer erfolgt über ein Relay, basierend auf CoreOS Dex²⁶, um verschiedene Identitäts-Provider einstellen zu können. Die technische Beschreibung der Komponenten findet sich in den entsprechenden Repositorien (auf Englisch).

3.4. Publizieren

Um Jupyter Notebooks als vollwertige Publikationen anbieten zu können, müssen grundlegende Voraussetzungen wie langfristige Verfügbarkeit, Zitierfähigkeit oder etablierte Peer-Review-Prozesse erfüllt sein. Insbesondere stellt die langfristige Lauffähigkeit von Programmcode ein zu lösendes Problem da. Für Erweiterte Publikationen mit Jupyter Notebooks bietet sich hier ebenso wie bei der Editions Umgebung eine Mischung aus Container-Lösung und virtualisierten Laufzeitumgebungen an. Jede einzelne Publikation hat eine eigene Laufzeitumgebung, im Jupyter-Kontext *Kernel* genannt. Dieser Kernel enthält alle für das Ausführen nötigen Python-Pakete und kann auch wiederverwendet werden, z.B. bei einer Reihe von inhaltlich ähnlichen Notebook-Veröffentlichungen.

3.4.1. Publikation

Jede Publikation liegt als Ordner mit festgelegter Struktur in einem Repository, z.B. dem DARIAH-DE Repository oder auf Github. Der Name des Ordners sollte Aufschluss über das Thema der Publikation geben, z.B. *networks / networks_in_dh*, und wird in ein zentrales Verzeichnis aller Publikationen aufgenommen. Der Publikations-Ordner enthält auf erster Ebene drei Unter-Ordner: *config*, *data*, und *packages*.

Im Ordner *packages* liegt die *requirements.txt* Datei, in der die benötigten Python Pakete mit Angabe der jeweiligen Versionsnummer verzeichnet sind. Zusätzlich können in einer *packages.txt* Datei Linux Pakete angegeben werden, die für die Auswertung des Notebooks benötigt werden.

Das eigentliche Publikations-Notebook befindet sich in dem Ordner *data* und wird immer mit dem Namen *index.ipynb* versehen. Hierbei kann es sich um die Eingangsseite

²⁴<https://jupyterhub.readthedocs.io/en/stable>

²⁵<https://nextcloud.com/de/>

²⁶<https://github.com/dexidp/dex>

zu einem ganzen Buch handeln oder auch um eine einseitige reportartige Anleitung für die Verwendung eines Datensatzes.

Die im Notebook bearbeiteten Daten sollten wenn möglich über Webquellen oder bevorzugt über eine DOI²⁷ geladen werden. Sollte dies nicht möglich sein, können die benötigten Daten auch direkt in dem *data* Ordner abgelegt werden. Diese werden dann mit in den Publikations-Container kopiert.

Um eine Publikation zu erstellen müssen die Autoren ihre Abhängigkeiten von Pythonpaketen feststellen und in eine Textdatei eintragen. Diese wird dann zusammen mit dem eigentlichen Notebook an den Verlag geschickt. Dort werden die Dateien in die Ordnerstruktur eingepflegt und außerdem wird einmalig das Ausführen des Notebooks getestet. Dann kann der Review-Prozess beginnen.

3.4.2. Referee-Workflow

Soll eine Erweiterte Publikation bei einem Verlag zur Publikation eingereicht werden, erfolgt der gewohnte redaktionelle Workflow. Ähnlich wie die Forschungseinrichtung, hat auch jeder Verlag die Möglichkeit, eine Präsentationsumgebung für Jupyter Notebooks aufzusetzen. Ein eingereichtes Notebook wird in diese Umgebung geladen und für die ausgewählten Referees in Kopien freigegeben. Referees erhalten persönliche Zugangsdaten, die den Einblick in die eingereichte Publikation erlauben. Hierdurch soll sichergestellt werden, dass das zu begutachtende Notebook mit dem später veröffentlichten Notebook inhaltlich und in der Präsentation übereinstimmt. So kann jeder Referee getrennt das Notebook interaktiv anschauen und kommentieren. Die Reaktionen der Referees werden anonym an die Autoren weitergeleitet, die wiederum das Notebook anpassen können. Ist der Review-Prozess erfolgreich abgeschlossen, kann der Verlag das Notebook veröffentlichen.

3.5. Darstellung

Für die Darstellung von Jupyter Notebooks bietet sich eine Trennung von Daten und Ansicht an. Die eigentlichen Daten eines Notebooks, also die Ordner mit Notebook-Datei und Dokumentation, sollten zitierbar als Datensatz veröffentlicht werden. Dies kann z.B über externe Partner wie das DARIAH-DE Repository geschehen. Bei der Veröffentlichung ist darauf zu achten, dass die Dokumentation des Notebooks alle Schritte enthält um dieses auf einem beliebigen Rechner lauffähig zu machen. In der Ansicht eines Notebooks wird dann die DOI für diesen Notebook-Datensatz für das Zitieren angegeben.

²⁷Digital object identifier

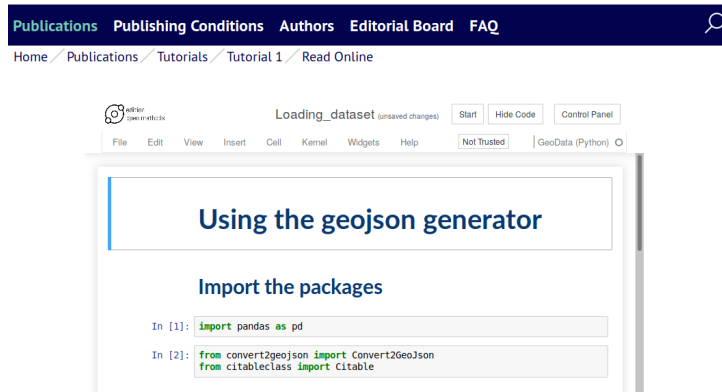


Abbildung 3: Einbindung eines Jupyter Notebooks in eine Publikationsplattform.

Ebenso sollten Daten, die in einem Publikations-Notebook genutzt werden, getrennt von den Notebookdaten ebenfalls in einem Repository veröffentlicht sein. Dies erhöht die Nachnutzbarkeit von Forschungsdaten und erlaubt das getrennte Zitieren von Daten und Auswertung. Ein Notebookzitat könnte also z.B. folgendermaßen Aussehen

Vogl, M.: „Analyse historischer semantischer und soziologischer Netzwerke“, in: Edition Open Methods 2019, 1, DOI:1234.4567.5678, Ansicht unter: <https://c105-187.cloud.gwdg.de/convert2geojson.html> (abgerufen am 24.02.2019).

Die Ansicht einer Notebook-Publikation erfolgt über die Einbettung einer speziellen Variante der Editions Umgebung in eine gewöhnliche Webseite. Diese Variante der Editions Umgebung enthält eine Anmeldeumgebung, die jedem Besucher der Webseite einen zufälligen Nutzernamen gibt und ihn/sie ohne Passwort auf einen extra gestarteten Docker-Container umleitet. In diesem Container sind direkt alle Veröffentlichungen des Verlags eingebaut. Über die verschiedenen Jupyter-Kernel können die einzelnen Paket-Abhängigkeiten aufgelöst werden. Im *config* Ordner können neben technischen Konfigurations-Dateien für das Jupyter Notebook auch Designanpassungen liegen. So kann zum Beispiel für verschiedene Editionen eines Verlags ein angepasstes Design für das Jupyter Notebook über eine *custom.css* Datei geladen werden oder es können über eine *custom.js* Datei Anpassungen an der Bedienungs Oberfläche, wie zusätzliche Knöpfe o.ä., vorgenommen werden.

Durch die Umleitung öffnet sich für die Leser der Veröffentlichung direkt das gewünschte Notebook. Während vor dem Laden des ersten Notebooks eine kurze Zeit vergehen kann (ca. 10 Sekunden), erfolgt das Öffnen von weiteren Publikationen dann

sehr schnell, da wie oben beschrieben alle Publikationen in dem Container bereits geladen sind. Leser können das Notebook ausführen und auch verändern, um z.B. Startparameter einer Simulation zu verändern. Jede Änderung wird allerdings nach einer gewissen Wartezeit verworfen und bei einem späteren Wiederaufruf der Publikation ist der originale Zustand automatisch wieder hergestellt. So kann zum einen eine interaktive Ansicht für ein kritisches Lesen und zum anderen eine reproduzierbare Darstellung einer Notebook-Publikation sichergestellt werden.

Ein Beispiel für die Umsetzung einer Präsentationsumgebung findet sich auf <https://c105-187.cloud.gwdg.de/>. Es ist geplant, sie nach der Testphase in die Infrastruktur der *Edition Open Access*²⁸ zu integrieren.

3.6. Zusammenfassung Jupyter Notebooks

Jupyter Notebooks bieten eine neuartige Möglichkeit des Erweiterten Publizierens. In Verbindung mit der Veröffentlichung als Open Access können sie eine Grundlage für die Zukunft der Open Science bieten. Durch die narrative Verknüpfung von wissenschaftlichem Text, Dokumentation und ausführbarem Programmcode ist es Lesern möglich wissenschaftliche Argumentation kritisch zu verfolgen und zu analysieren. Die Verfügbarkeit von Programmcode und Datensätzen als zitierbare Open Access Dokumente fördert die gute wissenschaftliche Praxis ebenso wie die Nachnutzbarkeit.

DARIAH-DE wird die nötigen Bausätze bzw. Anleitungen für die Erstellung solcher Bausätze auf seinen Github-Seiten in Kürze veröffentlichen.

4. Ausblick

Die im Rahmen von AP 1.2. durchgeführten Arbeiten waren in zweifacher Hinsicht erfolgreich: Einerseits konnten am Beispiel von zwei Use Cases einschlägige Publikationsstrategien entwickelt werden, die den vielfältigen Anforderungen und Möglichkeiten der digitalen geistes- und kulturwissenschaftlichen Forschung Rechnung tragen und eine gezielte Dissemination der Resultate erlauben. Andererseits fördern die erarbeiteten Publikationsstrategien die Transparenz und Nachnutzbarkeit der im Zuge des jeweiligen Forschungsprozesses entstandenen digitalen Daten und Ressourcen, unter anderem Forschungsdaten und interaktive Objekte. Auf diese Weise leistet AP 1.2. nicht nur einen Beitrag zur wissenschaftlichen Nachhaltigkeit, sondern auch zur Vision von DARIAH-DE: der Schaffung eines kooperativen Großraums zur digitalen Forschung in den europäischen Geistes- und Kulturwissenschaften.

²⁸<http://www.edition-open-access.de>