

# Executable Papers in den Computational Humanities

## Von technischen Herausforderungen und erkenntnistheoretischen Mehrwerten

**Walkowski, Niels-Oliver**

niels-oliver.walkowski@uni.lu  
Universität Luxemburg

**Burghardt, Manuel**

burghardt@informatik.uni-leipzig.de  
Universität Leipzig

## Einleitung

Die Gegenüberstellung von Programmcode und narrativem Text bei gleichzeitiger Evaluation ihrer Beziehungen, Ähnlichkeiten sowie Modi der gegenseitigen Bezugnahme ist so alt wie der Computer selbst. Bereits vor 40 Jahren appellierte der Informatiker Donald Knuth, man solle Programme doch besser als "literarische Werke" betrachten (Knuth 1984, 97). Knuths Ansatz bildete die Grundlage vieler gegenwärtiger Softwaredokumentationssysteme wie z.B. Pythons *docutils reStructuredText* und *sphinx*. Der durch Knuth sprichwörtlich gewordene Begriff des *literate programming* weist jedoch eindeutig darauf hin, dass diese Idee sehr viel weitreichender gemeint war als das zum quasi Standard gewordene formatierungs-fähige Dokumentieren eines Programms im Programmcode selbst. Legt der Begriff der Dokumentation immer noch eine Priorisierung nahe – das Eigentliche ist der Programmcode – so eröffnen *Executable Papers* (ExP) unter dem Vorzeichen des wissenschaftlichen Publizierens ein Feld in dem beide Gegenstandsbereiche sehr viel egalitärer und vielfältiger miteinander interagieren als im Bereich der Softwareentwicklung. ExP ist ein Sammelbegriff für einen Diskurs in den (digitalen) Wissenschaften sowie einer Reihe von Aktivitäten, die die gleichzeitige Veröffentlichung von Forschungsartikeln mit dem der Forschung zu Grunde liegendem, ausführbaren Programmcode vorsieht. Diese Aktivitäten sind unter anderem eng verbunden mit *Elseviers Executable Papers Grand Challenge* im Jahr 2011 (Gabriel 2011). Die Genealogie der dahinter liegenden Idee beginnt aber auch hier wesentlich früher. Erste Beispiele lassen sich bereits zum Ende des letzten Jahrtausends entdecken (vgl. Singh et al. 1998; Burg et al. 2000).

Die Realisierung von ExP variiert allerdings sehr stark. Dies betrifft sowohl die Art der Integration von Text und Programmcode als auch den technischen Ansatz, mit dem eine stabile, möglichst kontext-ungebundene Ausführbarkeit des Programmcodes gesichert werden soll. Gerade frühere Ansätze setzten dabei auf Komplettvirtualisierungen von Betriebssystemen, die ein lokal installierbares Image oder einen Remote-Desktop zur Verfügung stellen innerhalb dessen Leser:innen das ExP konsumieren konnten (Brammer et al. 2011; van Gorp and Mazanek 2011). Mittlerweile werden dafür nunmehr eher "leichtgewichtige" Virtualisierungsverfahren wie Docker (Cito, Ferme, und Gall 2016;

Boettiger 2015) genutzt. Ein anderer Ansatz arbeitet mit deklarativen Beschreibungen der für die Ausführbarkeit von ExP aufzulösenden Abhängigkeitsbäume, die wie eine Art Manifest mit dem ExP zusammen veröffentlicht werden (Pebesma et al. 2012; Nüst et al. 2016). Auch im Kontext von PDFs haben *Springer Nature* und der PDF Reader *ReadCube* vor einigen Jahren mit der Integration live ausführbarer, interaktiver Elemente experimentiert.

ExP arrangieren den Programmcode in den Randbereichen des Textes, vergleichbar mit einer Fußnote (Ciepiela et al. 2013; siehe Abbildung 1) oder aber auch auf Tokenebene, eng mit dem Text verwoben (Maciocco et al. 2019). Manchmal bilden Text und Programmcode zwei separate Bestandteile einer übergreifenden Publikation (Agnone 2020; Kray et al. 2019), ein anderes mal wird der vollständige Text zusammen mit den ausführbaren Bestandteilen überhaupt erst generiert (Smith et al. 2013).

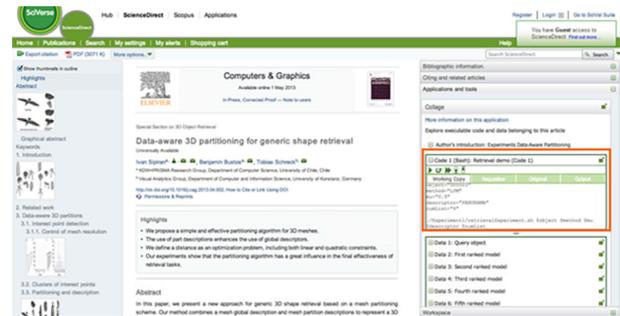


Abb. 1: Einer der Gewinner des Elseviers Executable Papers Grand Challenge mit ausführbaren Codefragmenten im Randbereich des Artikels (Hervorhebung)

Eine der größeren Herausforderungen von ExP betrifft die Möglichkeiten und Mittel ihrer Erstellung. Vereinzelt hat dies, wie im Fall des *Collage Authoring Environment* (Ciepiela et al. 2013) oder *Stencila* (Aufreiter et al. 2018) zur Entwicklung genuiner ExP-"Editoren" geführt. Eine andere Perspektive basiert auf der Weiterentwicklung eines bestehenden Ökosystems aus Werkzeugen, Infrastrukturen und Initiativen. Gemeint sind hier Projekte wie das *Jupyter Notebook*, *Git* und *Docker* u. ä. Gerade das *Jupyter Notebook* ist seit Jahren ein integraler Bestandteil bestimmter Bereiche der Wissenschaftskommunikation. Doch der potenzielle Wert für die Anfertigung vollwertiger Wissenschaftspublikationen wird ebenfalls seit einiger Zeit erkannt (Kluyver et al. 2016; Chandre et al. 2021), und das nicht nur durch die Integration von *Jupyter* in den Publikationsworkflow des O'Reilly-Verlags (Odeh 2015). Mit dem *Jupyter Book Framework* hat das *Executable Book Project* kürzlich einen weiteren interessanten Schritt auf diesem Weg vorgelegt.

## Executable Publications in den Humanities

Die eben zusammengefasste Entwicklung wird zu großen Teilen von der Informatik sowie auch den Lebens- und Naturwissenschaften getragen. Beispiele aus den Geisteswissenschaften, wie etwa die *Executable Music Documents* (De Roure et al. 2014), Melanie Walshs *Introduction to Cultural Analytics & Python* (2021) oder mit Einschränkungen innerhalb von Alan Lius *WhatEveryISays*-Projekt (Liue et al. 2017), sind selten – und das obwohl *Jupyter Notebooks* auch in den Computational Humanities längst zum Grundlagenwerkzeug geworden sind. Schaut man je-

doch auf die etablierte Zeitschriftenlandschaft oder in das akademische Verlagswesen, so sieht die Situation disziplinübergreifend nicht viel anders aus. Zweifellos hat dies mit den großen Herausforderungen zu tun, denen man sich stellen muss, sollen ExP eine nachhaltige und vollwertige Form wissenschaftlichen Publizierens werden. Dabei stehen die technischen Herausforderungen nicht einmal unbedingt im Vordergrund. Publizieren ist mehr als eine Ressource online verfügbar machen, es ist der Übergang von einer klandestineren, informelleren und damit weniger verpflichtenden Wissenschaftskommunikation in eine wesentlich komplexere Kommunikationsökologie mit stärker kodifizierten Normen, Erwartungen, Bewertungsmaßstäben, Rollen und Infrastrukturen – ein Umstand der bei innovativen Publikationsformaten häufig unterschätzt wird (Walkowski 2019). Auf der anderen Seite bieten *Jupyter Notebooks* gerade in den Geisteswissenschaften einen idealen Ausgangspunkt für die Etablierung von ExP, stellen sie doch die bislang weitestgehende Entsprechung mit Knuths Vision des Programmierens als eine literarische Tätigkeit dar. Sie sind daher am stärksten in der Lage, den häufig narrativen Logiken geisteswissenschaftlicher Erkenntnis- und Darstellungsformen zu entsprechen.

## Lessons learned aus einem aktuellen ExP-Projekt

Vor diesem Hintergrund hat sich die *Melusina Press Luxemburg* und die Computational Humanities-Gruppe Leipzig im Rahmen des vDHD Bandes “Fabrikation von Erkenntnis” dazu entschlossen, die Realisierbarkeit von ExP als vollwertige Publikationen der Computational Humanities zu erproben. Der Vortrag möchte sowohl die Kontextualisierung und die Potenziale dieses Publikationsformats im zuvor angedeuteten Sinne vorstellen, als auch detailliert die relevanten Fragen und Problemfelder in der Perspektive einer funktionierenden Publikationsökologie evaluieren. Nachfolgend stellen wir einige vorläufige Erkenntnisse zu drei unterschiedlichen Problemfeldern aus dem aktuellen vDHD-Publikationsprojekt vor.

### Zum Verhältnis von epistemischen Mehrwerten und (technischen) Aufwänden

Es hat sich herausgestellt, dass die Aufbereitung eines *Jupyter Notebooks* zu einem ExP trotz der Tatsache, dass ersteres meist bereits während des Forschungsprozesses entsteht, einen Arbeitsaufwand mit sich bringt, der den eines herkömmlichen Artikels deutlich übersteigt. Gleichzeitig bleiben häufig propagierte Mehrwerte, wie die bessere Nachvollziehbarkeit von Forschungsmethoden sowie die Reproduzierbarkeit von Forschungsergebnissen (Lasser 2020), schwer empirisch evaluierbar. Inwieweit ist zum Beispiel die Eleganz der Möglichkeit Code während des Lesens innerhalb des Artikels ausführen zu können ein entscheidender Vorteil gegenüber dem Lesen desselben Codes in einem statischen Export des Notebooks oder dem beiläufigen Verfügbarmachen von Skripten in einem Git-Repository. Solche Fragen lassen sich nur für konkrete Szenarien beantworten. Mit unserer Sektion im vDHD-Band wollten wir einen Beitrag für die bessere Identifizierbarkeit dieser Szenarien liefern. Da geisteswissenschaftliche Forschung sowieso sehr viel seltener einer rein empirischen Evidenzlogik folgt, scheint uns der größte Mehrwert für die Geisteswissenschaften eher in den epistemischen Potenzialen dieses Formats zu liegen. Vor diesem Hintergrund ist es jedoch bemerk-

enswert, dass die meisten Einreichungen das ExP-Format eher in Richtung eines Methodenpapiers bzw. Tutorials interpretiert haben. Diese Beobachtung zeigt, dass ExP, verstanden als eine Form des *literate programming* im weitestmöglichen Sinne, auch auf Autor:innenseite noch Raum für eine kreative Aneignung zulässt

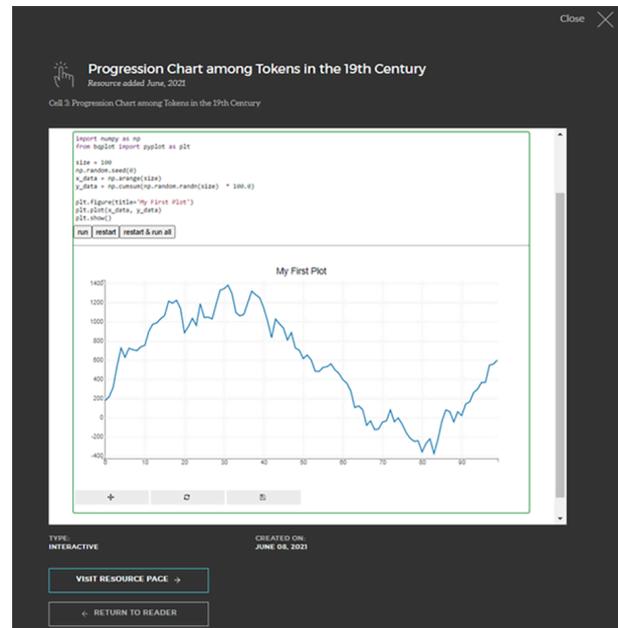


Abb. 2: Artikeloverlay mit ausgeführter und editierbarer Codezelle auf Melusina Press

### Zu technischen Anforderungen und nachhaltiger Infrastruktur

Technische und infrastrukturelle Herausforderungen bestehen am augenscheinlichsten zunächst erst einmal in der situativen Bereitstellung jeder Zeit abrufbarer Rechnerressourcen, die für das Ausführen des Codes notwendig sind, und bei der verlässlichen Auflösung von Abhängigkeiten. Hier hat sich mit dem Projekt *myBinder* ein Verfahren etabliert, das unter Einhaltung bestimmter Vorgaben aus einem Git-Repository mit einem *Jupyter Notebook* ein lauffähiges *Docker Image* generiert und dieses mittels Link über virtuelle Server kooperierender Institutionen im Browser verfügbar macht. Zwar erfreut sich dieses Verfahren aufgrund seiner Einfachheit großer Beliebtheit, allerdings zeigten die Einreichungen zu unserer vDHD-Sektion, dass die auf diese Weise zur Verfügung stehenden Ressourcen den Bedarfen, z. B. im Bereich des statistischen Lernens, schnell nicht mehr gerecht werden. Gleichzeitig steht der Nutzung institutioneller Kapazitäten aus dem High Performance Computing-Bereich (HPC) wie sich in unserem Falle erwiesen hat – nicht selten eine andere Dienstphilosophie entgegen. HPCs arbeiten mit definierbaren Aufgabenbeschreibungen, Ressourcenallokation und geplanten Abarbeitungen und sehen daher keine Benutzerinteraktion zum Zeitpunkt der Abarbeitung vor. Aus der Perspektive der Nachhaltigkeit gibt es nicht zuletzt auch immer wieder generelle Bedenken gegen die Verwendung von *Docker* und anderen Containerisierungsverfahren (Nüst et al. 2016). Festzuhalten bleibt auf jeden Fall, dass die Perspektive der Langzeitverfügbarkeit von ExP andere Lösungen erfordert, als die zur Zeit für die unmittelbare Verfügbarmachung gängigen Mittel. Weitere Probleme dieses Verfahrens sind

unzufriedenstellende Mittel der Absicherung copyright-geschützten Materials vor Zugriffen aus dem ExP, der Einbettung ausführbarer Bestandteile aus einem Notebook, geeignete Leseumgebungen und Webseiten von Verlagen (siehe auch Abbildung 2) sowie damit in Beziehung stehend, die Absicherung der Verlagsinfrastruktur vor Attacken, etwa durch *Code Injections*. Der Vortrag wird einen Überblick über die Melusina Press-Infrastruktur für die Realisierung von ExP im Rahmen des vDHd-Bandes geben und Stärken und Schwächen im Kontext der zuvor genannten Problemfelder diskutieren. Die Infrastruktur besteht im Wesentlichen aus der Nutzung der GESIS *myBinder*-Instanz, einem virtuellen Server zur technischen Isolierung der ausführbaren Publikationsabschnitte, einem virtuellen Server für die Präsentationsschicht, einer GitLab-Instanz der Universität Luxemburg und der *thebe*-Bibliothek (Abbildung 3).

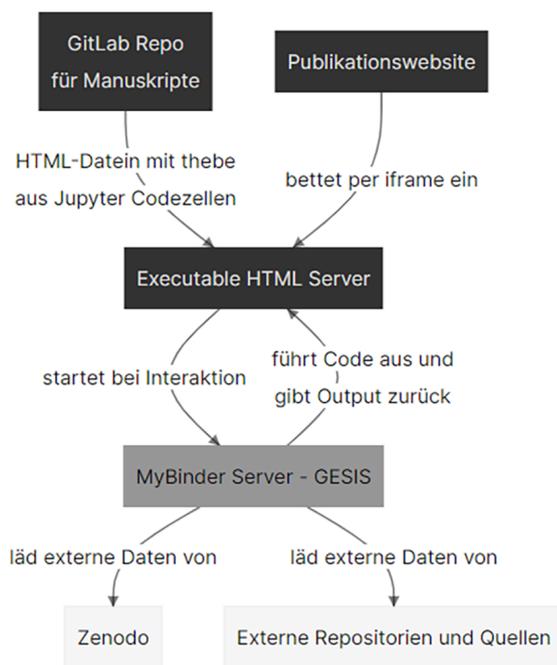


Abb. 3: Kerninfrastruktur von ExP auf Melusina

## Zu Publikationsworkflows, Rechten, Versionierung und Reviewkriterien

Während es für viele technische Fragestellungen von ExP immerhin schon erste Lösungsvorschläge gibt und die Wahrnehmung potenzieller Mehrwerte der Entwicklung eher vorausgeht, wurde das Problemfeld der Publikationsworkflows bisher am wenigsten bearbeitet. Ob diesem Umstand mangelndes Bewusstsein gegenüber der Notwendigkeit ExP anders "bewerten" und "behandeln" zu müssen als traditionelle Forschungsartikel zugrunde liegt, oder ob sich Schwierigkeiten und Ausmaß dieser Perspektiven erst in konkreten Publikationssituationen von ExP so richtig identifizieren lassen, bleibt offen. Nichtsdestotrotz hat sich während der Konzeptualisierung des vDHd-Bandes und bei der Gestaltung von Autor:innenhandreichung zur Anfertigung der ExP schnell gezeigt, dass nicht nur mehrere Dimensionen, wie etwa Code, Interaktivität, multimodale Komposition, Organisation aller für die Lauffähigkeit notwendiger Ressourcen usw., zu bewer-

ten sind, sondern das in dem Moment wo man diese Mittel ermöglicht ebenfalls Kriterien für Einschränkung ihres Gebrauchs entwickelt werden müssen. Wie Lasser (2020) für das Element der Reproduktionsfähigkeit beschriebener Ergebnisse zeigt, ist mehr nicht immer besser – insbesondere für die Frage der "Lesbarkeit" von ExP. Daran schließt sich sogleich die Fragestellung an, was Lesbarkeit im Kontext von ExP überhaupt bedeuten soll: was ist eine lesbare multimodale Komposition, wie ist Interaktivität unter dem Gesichtspunkt der Lesbarkeit zu fassen, sind die Ansprüche an lesbaren Code dieselben wie in der traditionellen Softwareentwicklung? Wahrscheinlich sind diese und andere Fragen nur im Kontext von spezifischen Disziplinen und konkreten Forschungsbereichen beantwortbar. Da es bisher aber kaum ExP gibt, konnte ein solcher Entwicklungsprozess auch noch nicht wirklich stattfinden. Folgen hat dies sowohl für die Bewertbarkeit der Qualität von ExP, d. h. der Durchführbarkeit transparenter Reviewverfahren, als auch für die Infrastruktur und die Workflows der beteiligten Verlage, die die Gestalt von ExP in der Praxis antizipieren können müssen. Es ergibt sich aus der zuvor beschriebenen Situation von selbst, dass ein Standardisierungsprozess und die Entwicklung von Best Practices in Teilbereichen von ExP, wo dies sinnvoll und möglich erscheint, noch nicht eingesetzt hat. Es darf nicht unterschätzt werden, dass es sich hierbei eben nicht nur um eine Bedingung der Möglichkeit ihrer Realisierung handelt, sondern ebenso um die Grundlage dafür ExP zu einer vertrauensvollen Ressource, zu einer Währung mit Wert innerhalb der Wissenschaftskommunikation, werden zu lassen. Angrenzende Aktivitäten (Leipzig 2019; Sackmann 2020) finden bisher eher in informellen Kontexten statt, inkludieren meist nur ein Bruchteil der Akteursgruppen, die an solch einem Prozess beteiligt sein müssten, und versuchen zunächst eher das Feld zu kartographieren als es zu organisieren. Hervorgehoben werden kann "A guide to reproducible code in ecology and evolution" (Cooper & Hsing 2017), doch wie der Name anzeigt bewegen sich diese Aktivitäten stärker im *Reproducibility*-Bereich, der zwar viele Schnittmengen mit ExP aufweist, aber eben nicht mit ihnen deckungsgleich ist.

Durch die Diskussion von Entscheidungen und Vorgaben aus den Autor:innenhandreichung sowie der Illustration des Publikationsworkflows von ExP bei Melusina Press soll der Vortrag einen Beitrag zu der Entwicklung künftiger Best Practices und Standards leisten.

## Bibliographie

- Agnone, Anthony** (2020): "Papers with Code + ArXiv = Reproducible, Organized" in: *Towards Data Science*. <https://towardsdatascience.com/papers-with-code-arxiv-reproducible-organized-research-f5404eb6a22e> [letzter Zugriff 9. Juli 2021]
- Aufreiter, Michael / Pawlik, Aleksandra / Bentley, Nokome** (2018): "Stencila – an Office Suite for Reproducible Research" in: *ELife*, July 2, 2018 [https://elifesciences.org/labs/c496b8bb/stencila-an-office-suite-for-reproducible-research?utm\\_source=labworm&utm\\_medium=feed&utm\\_campaign=stencila](https://elifesciences.org/labs/c496b8bb/stencila-an-office-suite-for-reproducible-research?utm_source=labworm&utm_medium=feed&utm_campaign=stencila).
- Brammer, Grant / Crosby, Ralph / Matthews, Suzanne / Williams, Tiffani** (2011): „Paper Mâché: Creating Dynamic Reproducible Science" in: *Procedia Computer Science, Proceedings of the International Conference on Computational Science* 4: 658–67 10.1016/j.procs.2011.04.069.
- Burg, Jennifer / Wong, Yue-Ling / Yip, Ching-Wan / Boyle, Anne** (2000): „The state of the art in interactive multimedia jour-

nals for academia" in: *Proceedings of EdMedia: World Conference on Educational Media and Technology 2000 2*: 37–42. Montréal: Association for the Advancement of Computing in Education <http://www.editlib.org/p/16036>.

**Ciepiela, Eryk / Harezlak, Daniel / Kasztelnik, Marek / Meizner, Jan / Dyk, Grzegorz / Nowakowski, Piotr / Bubak, Marian** (2013): "The Collage Authoring Environment: From Proof-of-Concept Prototype to Pilot Service" in: *Procedia Computer Science* 18: 769–78 [10.1016/j.procs.2013.05.241](https://doi.org/10.1016/j.procs.2013.05.241).

**Chandre, Cristel / Dubois, Jonathan** (2021): "Notebook Articles: Towards a Transformative Publishing Experience in Non-linear Science" in: *Communications in Nonlinear Science and Numerical Simulation* 97: 105753 [10.1016/j.cnsns.2021.105753](https://doi.org/10.1016/j.cnsns.2021.105753).

**Cooper, Natalie / Hsing, Pen-Yuan** (2017): *A guide to reproducible code in ecology and evolution*. London: British Ecological Society.

**Gabriel, Ann / Capone, Rebecca** (2011): "Executable Paper Grand Challenge Workshop" in: *Procedia Computer Science* 4: 577–78 [10.1016/j.procs.2011.04.060](https://doi.org/10.1016/j.procs.2011.04.060).

**Gorp, Pieter van / Mazanek, Steffen** (2011): „SHARE: a web portal for creating and sharing executable research papers" in: *Procedia Computer Science, Proceedings of the International Conference on Computational Science* 4: 589–97 [10.1016/j.procs.2011.04.062](https://doi.org/10.1016/j.procs.2011.04.062).

**Thomas, Kluyver / Benjamin, Ragan-Kelley / Fernando, Pérez / Brian, Granger / Matthias, Bussonnier / Jonathan, Frederic / Kyle, Kelley** (2016): "Jupyter Notebooks – a publishing format for reproducible computational workflows" in: Loizides, Fernando / Schmidt, Birgit (eds.): *Positioning and Power in Academic Publishing: Players, Agents and Agendas* 87–90. Amsterdam: IOS Press [10.3233/978-1-61499-649-1-87](https://doi.org/10.3233/978-1-61499-649-1-87).

**Knuth, Donald E.** (1984): "Literate Programming" in: *The Computer Journal* 27,2: 97–111 [10.1093/comjnl/27.2.97](https://doi.org/10.1093/comjnl/27.2.97).

**Kray, Christian, / Pebesma, Edzer / Konkol, Markus / Nüst, Daniel** (2019): "Reproducible Research in Geoinformatics: Concepts, Challenges and Benefits" in: Timpf, Sabine / Schlieder, Christoph / Kattenbeck, Markus / Ludwig, Bernd / Stewart, Kathleen (eds.): *COSIT2019* 142: 8:1–8:13 [10.4230/lipics.cosit.2019.8](https://doi.org/10.4230/lipics.cosit.2019.8).

**Lasser, Jana** (2020): „Creating an Executable Paper Is a Journey through Open Science" in: *Communications Physics* 3,1: 1–5 [10.1038/s42005-020-00403-4](https://doi.org/10.1038/s42005-020-00403-4).

**Leipzig, Jeremy** (2019): *Awesome Reproducible Research* [10.5281/ZENODO.3564746](https://doi.org/10.5281/ZENODO.3564746).

**Liu, Alan / Kleinman, Scott / Douglass, Jeremy / Thomas, Lindsay / Champagne, Ashley / Russell, Jamal** (2017): „Open, Shareable, Reproducible Workflows for the Digital Humanities: The Case of the 4Humanities.Org ‚WhatEvery1Says‘ Project" in: *Digital Humanities 2017*. Conference Abstracts. <https://alanyliu.org/research/talks/2017-dh2017/we1sdh2017-panel-abstract.pdf> [letzter Zugriff 7. Juli 2021].

**Maciocci, Giuliano / Aufreiter, Michael / Bentley, Nokome** (2019): "Introducing ELife's First Computationally Reproducible Article" in: *eLife* [https://elifesciences.org/labs/ad58f08d/introducing-elife-s-first-computationally-reproducible-article](https://doi.org/https://elifesciences.org/labs/ad58f08d/introducing-elife-s-first-computationally-reproducible-article) [letzter Zugriff 7. Juli 2021].

**Nüst, Daniel / Konkol, Markus / Pebesma, Edzer / Kray, Christian / Klötgen, Stephanie / Schutzzeichel, Marc / Lorenz, Jörg / Przibytzin, Holger / Kussmann, Dirk** (2016): „Opening Reproducible Research" in: *EGU General Assembly Conference Abstracts* <http://adsabs.harvard.edu/abs/2016EGUGA..18.7396N> [letzter Zugriff 7. Juli 2021].

**Odehahn, Andrew** (2015): „Embracing Jupyter Notebooks at O'Reilly" in: O'Reilly Media Blog <https://beta.oreilly.com/ideas/jupyter-at-oreilly> [letzter Zugriff 7. Juli 2021].

**Pebesma, Edzer / Nüst, Daniel / Bivand, Roger** (2012): „The R Software Environment in Reproducible Geoscientific Research" in: *Eos, Transactions American Geophysical Union* 93,16: 163–163 [10.1029/2012EO160003](https://doi.org/10.1029/2012EO160003).

**Roure, David De** (2014): "Executable Music Documents" in: *Proceedings of the 1st International Workshop on Digital Libraries for Musicology - DLfM '14*, 2014, 1–3 [10.1145/2660168.2660183](https://doi.org/10.1145/2660168.2660183).

**Singh, Ripudaman / Chudoba, Rostislav / Gopal, K. / Koenke, Carsten** (1998): „IMMJ: interactive multi-media journals in science and technology" in: *Ejournal* 8, 2.

**Smith, Vincent / Georgiev, Teodor / Stoev, Pavel / Biserkov, Jordan / Miller, Jeremy / Livermore, Laurence / Baker, Edward** (2013): „Beyond dead trees: integrating the scientific process in the Biodiversity Data Journal" in: *Biodiversity data journal* 1: e995 [10.3897/BDJ.1.e995](https://doi.org/10.3897/BDJ.1.e995).

**Walsh, Melanie** (2021): *Introduction to Cultural Analytics & Python*. Ithaca, NY [10.5281/zenodo.4411250](https://doi.org/10.5281/zenodo.4411250).