



<https://publications.dainst.org>

iDAI.publications

DIGITALE PUBLIKATIONEN DES  
DEUTSCHEN ARCHÄOLOGISCHEN INSTITUTS

Das ist eine digitale Ausgabe von / This is a digital edition of

Baumeister, Peter

## Das neue maschinenlesbare Zeitschriftenmodell des Deutschen Archäologischen Instituts. Ein Werkstattbericht

aus / from

**Forum for Digital Archaeology and Infrastructure, 1. Faszikel 2022, § 1–67**

DOI: <https://doi.org/10.34780/cf6e-prof>

**Herausgebende Institution / Publisher:**  
Deutsches Archäologisches Institut

**Copyright (Digital Edition) © 2022 Deutsches Archäologisches Institut**  
Deutsches Archäologisches Institut, Zentrale, Podbielskiallee 69–71, 14195 Berlin, Tel: +49 30 187711-0  
Email: [info@dainst.de](mailto:info@dainst.de) | Web: <https://www.dainst.org>

**Nutzungsbedingungen:** Mit dem Herunterladen erkennen Sie die Nutzungsbedingungen (<https://publications.dainst.org/terms-of-use>) von iDAI.publications an. Sofern in dem Dokument nichts anderes ausdrücklich vermerkt ist, gelten folgende Nutzungsbedingungen: Die Nutzung der Inhalte ist ausschließlich privaten Nutzerinnen / Nutzern für den eigenen wissenschaftlichen und sonstigen privaten Gebrauch gestattet. Sämtliche Texte, Bilder und sonstige Inhalte in diesem Dokument unterliegen dem Schutz des Urheberrechts gemäß dem Urheberrechtsgesetz der Bundesrepublik Deutschland. Die Inhalte können von Ihnen nur dann genutzt und vervielfältigt werden, wenn Ihnen dies im Einzelfall durch den Rechteinhaber oder die Schrankenregelungen des Urheberrechts gestattet ist. Jede Art der Nutzung zu gewerblichen Zwecken ist untersagt. Zu den Möglichkeiten einer Lizenzierung von Nutzungsrechten wenden Sie sich bitte direkt an die verantwortlichen Herausgeberinnen/Herausgeber der entsprechenden Publikationsorgane oder an die Online-Redaktion des Deutschen Archäologischen Instituts ([info@dainst.de](mailto:info@dainst.de)). Etwaige davon abweichende Lizenzbedingungen sind im Abbildungsnachweis vermerkt.

**Terms of use:** By downloading you accept the terms of use (<https://publications.dainst.org/terms-of-use>) of iDAI.publications. Unless otherwise stated in the document, the following terms of use are applicable: All materials including texts, articles, images and other content contained in this document are subject to the German copyright. The contents are for personal use only and may only be reproduced or made accessible to third parties if you have gained permission from the copyright owner. Any form of commercial use is expressly prohibited. When seeking the granting of licenses of use or permission to reproduce any kind of material please contact the responsible editors of the publications or contact the Deutsches Archäologisches Institut ([info@dainst.de](mailto:info@dainst.de)). Any deviating terms of use are indicated in the credits.

AN ARTICLE FROM THE



FORUM FOR  
DIGITAL ARCHAEOLOGY AND  
INFRASTRUCTURE

62·20

MITTEILUNGEN DES  
DEUTSCHEN ARCHÄOLOGISCHEN  
INSTITUTS  
RÖMISCHE ABTEILUNG

BAND 127 · 2021



JOURNAL OF  
GLOBAL ARCHAEOLOGY



FORUM FOR  
DIGITAL ARCHAEOLOGY AND  
INFRASTRUCTURE

2022

## ABSTRACT

# The new machine-readable journal model of the German Archaeological Institute

## A workshop report

Peter Baumeister

In all disciplines of archaeology, not only the publication of research data plays a significant role. Interoperability and machine readability of all kinds of content are also becoming increasingly important in order to make the scientific outcome reusable in the best possible way. In the form of a workshop report, this article wants to share the German Archaeological Institute's experiences in developing a new journal model as part of its digitization strategy, which aims not only to publish journal articles in machine-readable form, but also to give research data publications prominent visibility by including them directly into the corresponding research article.

## SCHLAGWORTE/KEYWORDS

Digitale Transformation/Digital Transformation, Maschinenlesbarkeit/Machine-readability, Forschungsdatenmanagement/Research Data Management, Textverarbeitung/Text Processing, XML

# Das neue maschinen- lesbare Zeitschriften- modell des Deutschen Archäologischen Instituts

## Ein Werkstattbericht

### Ausgangspunkt

<sup>1</sup> Auch bei archäologisch arbeitenden Disziplinen spielt nicht nur die Publikation der Forschungsdaten, die im Rahmen eines Projektes entstehen, eine zentrale Rolle, es gewinnen ebenso Interoperabilität und maschinelle Lesbarkeit der zu veröffentlichenden Inhalte zunehmend an Bedeutung, um die Ergebnisse ihrer Forschungsarbeit in bestmöglicher Form rezipierbar zu machen.

<sup>2</sup> Durch die fortschreitende Ausdifferenzierung des Open Access-, Open Data- bzw. Open Science-Paradigmas werden die Anforderungen an die Eigenschaften der zu veröffentlichenden Inhalte zwar zunehmend klarer definiert<sup>1</sup>. Dennoch bedeutet eine praktische Umsetzung für institutionelle Akteure, die Publikationen herausgeben, insbesondere in den Geisteswissenschaften weiterhin eine größere Herausforderung, da Standards, Konzepte zur technischen Umsetzung bzw. diesbezügliche Best-Practice-Beispiele und infrastrukturelle Voraussetzungen oftmals nur partiell oder nur in einer bedingten Übertragbarkeit vorhanden sind<sup>2</sup>. Dies kann aus Gründen resultieren, die sehr disziplinspezifisch sind, und betrifft konkret etwa bestimmte Formate bzw. Konzepte der Datenmodellierung, die für den Einsatz im eigenen Fachbereich noch angepasst oder ausdifferenziert werden müssen. Dies bedeutet gleichzeitig, dass auch

---

<sup>1</sup> Stellvertretend und nur auszugsweise seien hier folgende Positions- und Strategiepapiere angeführt mit Verweis auf die dort weiter genannte Literatur: UNESCO 2021, Bilder – Lin – Neylon 2020, Wilkinson et al. 2016, Bethesda Statement 2003, siehe auch Bosman et al. 2021, Pampel 2021, aktuell für den Publikationsbereich besonders relevant: Wissenschaftsrat 2022, aber auch Pieper et al. 2018, Digitales Publizieren in den Geisteswissenschaften 2021.

<sup>2</sup> Anders stellt es sich dar etwa für die Bereiche rechtliche Rahmenbedingungen/Lizenzierung, Finanzierungs- bzw. Geschäftsmodelle usw., in denen sehr gute Vorlagen bzw. eine aktive unterstützende Community mit Helpdesks, Veranstaltungen usw. existieren, vgl. die in Anm. 1 genannte Literatur und siehe – nur stellvertretend genannt – Initiativen wie <https://open-access.network> (05.04.2022).

die entsprechenden Anwendungen für die konkrete Prozessierung und Bereitstellung der Inhalte für einen belastbaren Regelbetrieb oftmals noch entwickelt oder daraufhin skaliert werden müssen.

3 Nicht zu unterschätzen sind in den Geisteswissenschaften und besonders in den archäologisch arbeitenden Disziplinen aber auch abstraktere Faktoren, d. h.: Etablierte Erwartungen an die Eigenschaften der medialen Zwischen- und Endprodukte und vor allem auch an die Abläufe von Publikationsprozessen<sup>3</sup>. Auch ist die Heterogenität der am Prozess beteiligten Anspruchsgruppen in Bezug auf die notwendigen spezifischen Kenntnisse nach wie vor groß.

4 Der vorliegende Artikel will in Form eines Werkstattberichts zeigen, wie das DAI im Rahmen seiner Digitalisierungsstrategie ein neues Zeitschriftenmodell entwickelt, das dabei den Anspruch hat, Artikelinhalte in maschinenlesbarer Form zu publizieren und darüber hinaus dazugehörigen Forschungsdatenkonvoluten durch die inhaltliche und technische Verschränkung eine prominente Sichtbarkeit zu bieten.

5 Das Modell möchte sich zudem als disziplinspezifischer Beitrag zum Versuch einer Konzeptfindung sehen, an deren Ende ein Spektrum an Best-Practice-Lösungen steht, die andere Akteur:innen für sich übertragen können<sup>4</sup>. Daher werden in dem Werkstattbericht aus Sicht der herausgeberischen bzw. der redaktionellen Praxis die Herausforderungen, aber vor allem die Hintergründe der Entscheidungsfindungen für die gewählten Lösungsansätze nachvollziehbar gemacht und damit ausdrücklich auch zur Diskussion gestellt. Gleichzeitig werden die vom DAI entwickelten Komponenten quelloffen und somit nachnutzbar der Community zur Verfügung stellt, um einen Beitrag zur möglichen Übertragbarkeit zu leisten.

## Ziele

6 Ausgangspunkt war die Neugestaltung des sog. Archäologischen Anzeigers, eines der international traditionsreichsten Publikationsorgane im Bereich der Archäologie und der Altertumswissenschaften<sup>5</sup>. Der Archäologische Anzeiger (AA) wurde erstmals 1889 als Beiblatt zum Jahrbuch des Deutschen Archäologischen Instituts veröffentlicht, seit 1976 erscheint er als eigenständige Zeitschrift in zwei Teilbänden pro Jahr. Von Anfang an war er ein Ort für Aufsätze, aktuelle Berichte und kleinere Mitteilungen einer breiten, internationalen Fachcommunity.

---

3 Erfahrungen, die die Redaktionen des DAI diesbezüglich im Umgang mit den Autor:innen machen, werden durch verschiedene Untersuchungen bestätigt, vgl. Projekt AuROA 2022. Weitere Faktoren nennen Kuhn – Hagenhoff 2019 in ihrer Analyse zur Rolle digitaler Monographien in den Geisteswissenschaften; interessant sind insbesondere die Ausführungen zur Interdependenz zwischen (Text)form und hermeneutischen Prozessen in den Geisteswissenschaften.

4 Wie wichtig disziplinspezifische Beiträge sind, zeigen Böhm et al. 2020, 238: So mag es für manche Fachbereiche der Geisteswissenschaften zutreffen, dass deren Manuskripte »nur geringe Anforderungen an die Automatisierbarkeit [stellen], da sie in der Regel strukturell nur wenig komplex sind«. Dies gilt aber nicht für jene Disziplinen, die bild- und textwissenschaftlich arbeiten, komplexe Abbildungsapparate benötigen oder mit einem breiten Spektrum an Formaten und strukturierten Inhalten umgehen, wie Katalogtexten, Fotos, SFM-Modellen, naturwissenschaftlichen Analyse-, Grabungs-/Vermessungs-, Geo- bzw. Fernerkundungsdaten usw. – alles Punkte, die auf archäologisch und altertumswissenschaftlich arbeitende Disziplinen und deren Nachbarwissenschaften zutreffen.

5 Vgl. das Editorial des ersten neugestalteten Bandes: F. Fless – Ph. von Rummel, Editorial, AA 2019/1, § 1–10, <https://doi.org/10.34780/x04h-kv97>. Zum AA siehe <https://publications.dainst.org/journals/aa>.

Folgende Ziele wurden im Rahmen des Projektes definiert:

1. Sowohl die Artikel selbst als auch die beschreibenden Metadaten sollen in einer maschinenlesbaren Form zur Verfügung gestellt werden, um eine bestmögliche Nachnutzung für die Community – auch durch automatisierte Verfahren – zu gewährleisten.
2. Dies soll in Form einer zitierfähigen Ausgabe in responsiver Webansicht geschehen, bei der die Text- und Bildinhalte durch klare Strukturierung gut erfassbar sind, um alle wichtigen wissenschaftlichen Anforderungen einer zunehmend digital arbeitenden Fachcommunity zu erfüllen.
3. Die Artikel sollen mit möglichst vielen Normdaten anreicherbar sein. Zudem sollen alle relevanten Metadaten beigefügt werden, die notwendig sind, um Inhalt und Kontributoren usw. beschreiben zu können, und zwar unter Verwendung etablierter persistenter Identifikatoren wie DOI, ORCID usw. inkl. der entsprechenden Registrierungsvorgängen bei den zuständigen Agenturen wie DataCite usw.
4. Es soll möglich sein, die Journal-Artikel mit Forschungsdatenpublikationen zu verschränken.
5. Eine hochwertige Druckausgabe soll weiterhin angeboten werden, da diese für die relevanten Zielgruppen – d. h. für Rezipient:innen, aber auch vor allem für Autor:innen – nach wie vor wichtiges Qualitätsmerkmal eines wissenschaftlichen Journals in der Archäologie bzw. auch ein Karriereargument und damit einen Publikationsanreiz darstellt<sup>6</sup>.

8 Mit den o. g. Merkmalen soll vor allem auch jener Perspektivwechsel ermöglicht werden, die Publikationen selbst als digitale Objekte betrachtbar machen zu können, was in Anlehnung an die FAIR-Prinzipien zunehmend an Bedeutung gewinnt<sup>7</sup>.

## Die Umsetzung

9 Die Konzeption und Umsetzung erfolgte in Form eines gemeinsamen Projektes der Redaktion der Zentrale und des damaligen IT-Referats, heute beides Arbeitseinheiten der Zentralen Wissenschaftlichen Dienste (ZWD)<sup>8</sup>. Aufgrund der breitgefächerten Anforderungen, die überdies wegen des halbjährlichen Erscheinens des AA vor dem Hintergrund eines sehr eng getakteten Zeitplans zu bewältigen waren, war es nicht möglich, die Umsetzung in der Kürze der Zeit komplett inhouse zu bewerkstelligen. Daher wurden zur Ausdetaillierung des Konzeptes und dessen Umsetzung, d. h. auch konkret zur Einrichtung der Produktionsstrecke, externe Dienstleister

6 Vgl. zur Bedeutung von traditionellen Publikationsformen in den Geisteswissenschaften etwa zuletzt das Projekt AuROA 2022, 6; Dreher – Lang – Oberländer 2020, 13; zu den Altertumswissenschaften: Hintze 2020, bes. 197. 205. 206; siehe auch allgemein das DFG-Positionspapier »Elektronisches Publizieren« vom März 2005 (10.02.2022), 9, dessen Gültigkeit durch die eingangs genannten Publikationen weiterhin bestätigt wird.

7 Siehe dazu allgemein Wissenschaftsrat 2022, 43–48.

8 Benedikt Boyxen hatte bzw. hat als verantwortlicher Redakteur des AA maßgeblichen Anteil an der Konzeption, Umsetzung und kontinuierlichen Weiterentwicklung. Wir danken zudem Sebastian Cuy, der das Projekt trotz vielerlei anderer Verpflichtungen stets äußerst engagiert begleitet hat, und vor allem Reinhard Förtsch (beide ehemals IT-Referat) sowie der Präsidentin des DAI, Friederike Fless, für ihre uneingeschränkte Unterstützung. Benedikt Boyxen und Marcel Riedel danke ich für die kritische Durchsicht des Beitrages und ihre Anregungen.





1  
Abb. 1: Darstellung der Viewer-Ausgabe des Archäologischen Anzeigers (Design: Tanja Lemke-Mahdavi)

aus unterschiedlichen Gewerken ins Boot geholt, insbesondere aus den Bereichen Kommunikationsdesign<sup>9</sup>, Realisierung von XML-Prozessen<sup>10</sup> und allgemeine Programmierdienstleistungen für die Anpassungen des Viewers<sup>11</sup>. Dem Arbeitsbereich Wissenschaftliche Informationstechnologien des DAI kommt dabei eine besondere Rolle zu, nicht nur im Hinblick auf entscheidende konzeptionelle Impulse und Fragen der Machbarkeit, sondern auch bei der konkreten technischen Umsetzung und der Begleitung und Steuerung der externen Programmierung.

## Die maschinenlesbare Ausgabe

<sup>10</sup> Für das Ausgangsformat der maschinenlesbaren Ausgabe fiel die Entscheidung aufgrund der Möglichkeiten zur formalen Strukturierung und semantischen Auszeichnung auf das JATS XML-Format, das sich in einigen Bereichen für wissenschaftliche Journals zum Standard zu entwickeln beginnt<sup>12</sup>.

<sup>11</sup> Für die Anzeige fiel die Wahl auf den sog. eLife Lens, eine freie und quelloffene Anwendung (Abb. 1)<sup>13</sup>. Dies hat mehrere Gründe: Da das DAI im Sinne seiner Open Science-Strategie bestrebt ist, seine Informationsinfrastruktur möglichst auf Basis von Open Source-Produkten weiterzuentwickeln, schieden proprietäre Lösungen von vornherein aus. Entscheidend war dann nicht nur, dass sich viele Instanzen des eLife Lens in der Praxis des wissenschaftlichen Publikationsbetriebs bewährt haben, darunter auch in geisteswissenschaftlichen Kontexten<sup>14</sup>, und dass der Viewer viele Anforderungen für wissenschaftliche Journals standardmäßig erfüllt, sondern auch, dass eLife seine Entwicklungen in gewisser Abstimmung mit dem PKP-Konsortium vornahm<sup>15</sup>. Dies ist für das DAI insofern wichtig, da damit zumindest eine gewisse Kompatibilität mit den ebenfalls von im PKP-Umfeld entwickelten OJS- und OMP-Plattformen sichergestellt ist, von denen das DAI im Rahmen seiner iDAI.publications-Systeme betreibt, über die auch seine digitale Zeitschriftenausgaben veröffentlicht werden<sup>16</sup>.

## Verschränken der Journal-Artikel mit Forschungsdaten

<sup>12</sup> Kernbestandteil des Konzeptes ist es, zusätzlich zu den Inhalten, die auch in der Print-Version ausgegeben werden, in die Viewer-Ausgabe der .xml-Version eines Artikels Daten aus anderen Informationsressourcen einzubinden und anzuzeigen, um Normdaten, vor allem aber umfangreichere Forschungsdatenkonvolute mit den auswertenden Beiträgen verschränken und damit letztere prominent publizieren zu können.

<sup>9</sup> Tanja Lemke-Mahdavi, LMK – Büro für Kommunikationsdesign, Berlin.

<sup>10</sup> Fabian Kern, digital publishing competence, München.

<sup>11</sup> Friedrich Stoller, Florian Bergmann, LEAN BAKERY, München.

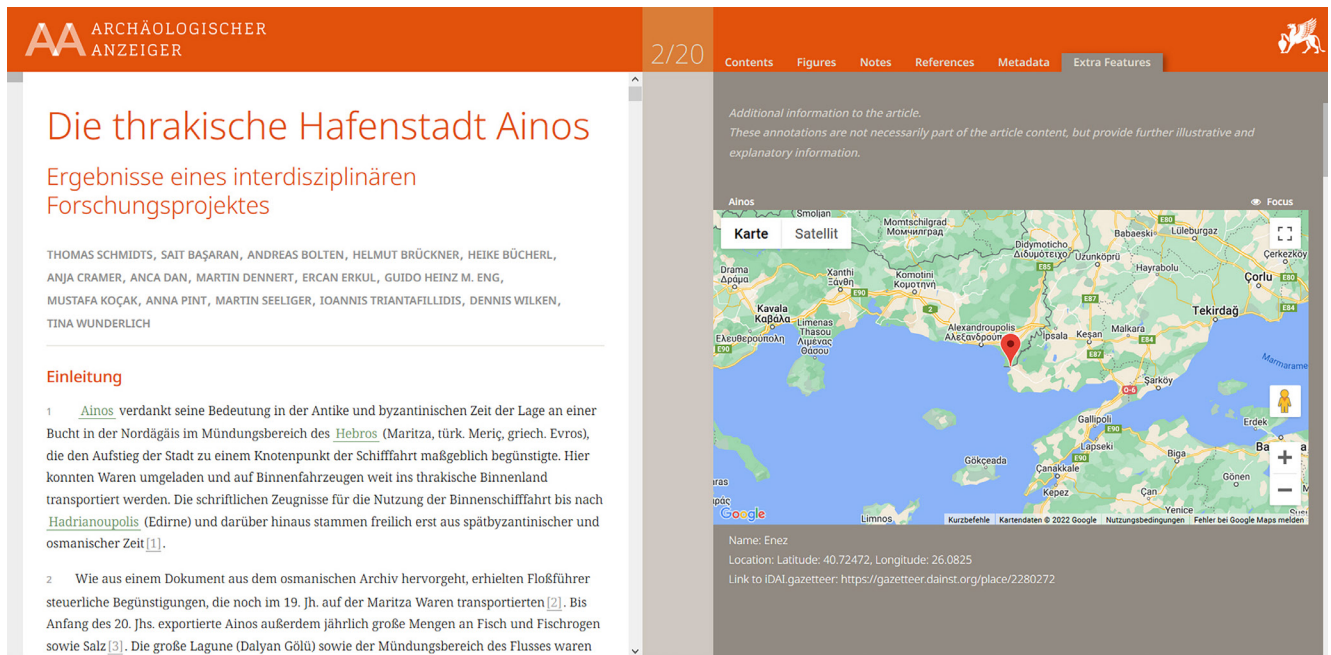
<sup>12</sup> <https://jats.nlm.nih.gov/index.html> (02.04.2022). Eine gute Übersicht zur Genese und zur Literatur bis 2016 findet sich bei Cristiani Sienra 2016, 101–107.

<sup>13</sup> <https://lens.elifesciences.org> (02.04.2022) bzw. <https://GitHub.com/elifesciences/lens>. Siehe das Mission Statement zum Start des Projektes 2013 unter <https://elifesciences.org/inside-elifesciences/0414db99/seeing-through-the-elifesciences-lens-a-new-way-to-view-research>. Das DAI nutzt dabei den Release 2.0.0.

<sup>14</sup> In github sind mehr als 70 Forks verzeichnet. Einen nutzt die UB Heidelberg für ihr Publikationsportal (<https://github.com/UB-Heidelberg/lens>), siehe zu heiUP bzw. heiMPT unten Anm. 33.

<sup>15</sup> Siehe <https://elifesciences.org/for-the-press/ca18fdae/elifesciences-invests-in-texture-to-provide-open-source-content-production-tools-for-publishers> (18.04.2022).

<sup>16</sup> Basierend auf OJS: <https://publications.dainst.org/journals>, basierend auf OMP: <https://publications.dainst.org/books/dai>.



2

13 Das Konzept sieht vor, den Nutzer:innen in einer Vorschau die wichtigsten Informationen zum referenzierten Objekt, Befund, Normdatum usw. inklusive der wichtigsten Kernfelder der beschreibenden Metadaten anzuzeigen. Für alle weiteren Informationen zu den Normdaten bzw. dem Objekt/Befund, dessen Kontextualisierung usw. werden die Nutzer:innen dann in die entsprechende Quell-Systeme weitergeleitet. Dazu wurde bzw. wird der Viewer so modifiziert, dass nach und nach die Einbindung von Inhalten aus allen wichtigen Normdaten-, Objekt- und sonstigen Dokumentationssystemen der iDAI.world, der vernetzten Informationsinfrastruktur des DAI, möglich sein soll.

14 Zu diesem Zweck wurde der Viewer um sog. Labels erweitert, die in Form von Tabs in der rechten Spalte angezeigt werden. Dass dabei die Labels »Extra Features« und »Supplements« unterschieden werden, obgleich beide auf identische Weise Inhalte der iDAI.world anzeigen, liegt am zugrundeliegenden Konzept: Bei den »Extra Features« handelt es sich entweder um referenzierte Normdaten oder um allgemeine Verweise auf illustrierendes Material, das nicht zwingend von dem:der Autor:in selbst erstellt wurde und das daher weder Gegenstand des Begutachtungsprozesses noch der redaktionellen Bearbeitung des Beitrages ist (Abb. 2). Davon unterscheiden sich die »Supplements«, bei denen es sich um Daten-Konvolute handelt, die von dem:der Autor:in entweder erstellt oder für den Beitrag zusammengestellt wurden und die vor allem – wie der Artikel selbst – Gegenstand des fachwissenschaftlichen Peer-review-Prozesses waren (Abb. 3).

15 In dieser Form angezeigt werden können zur Zeit Inhalte aus:

- iDAI.objects, dem Verzeichnissystem für Objektdaten
- iDAI.gazetteer, dem Normdatensystem für Ortsnamen (Toponyme) und deren geografische Koordinaten
- iDAI.field, dem System zur Dokumentation von archäologischen Ausgrabungen und Feldforschungsprojekten

Abb. 2: Beispiel für einen referenzierten Datendatensatz aus dem Normdatensystem für Ortsnamen und deren geografische Koordinaten iDAI.gazetteer, dessen Vorschau unter »Extra Features« angezeigt wird (Screenshot: Peter Baumeister)



# The original of Max Uhle's map of Pachacamac in Quito

HEIKO PRÜMERS

1 There are many stories about unexpected document finds in archives – this contribution is another example in the list. My search in the Archivo Histórico del Ministerio de Cultura y Patrimonio in Quito was for notes, drawings and photos made by Jacinto Jijón y Caamaño. During his exile in 1925, this Ecuadorian industrialist, politician and archaeologist had carried out excavations in Maranga, Peru, which he later published in a confusing monograph that did not contain a single picture of the actual excavations (Jijón y Caamaño 1949). I had hoped to find exactly such photographs in his estate, which is kept in the Archivo Histórico, because I was interested in reconstructing the exact archaeological context of some textiles he had uncovered in Maranga that were of the few known examples belonging to the Lima Culture. The result of my search was sobering. Only 25 small prints with views of the excavation were found, including repetitive scenes from slightly different angles [1].

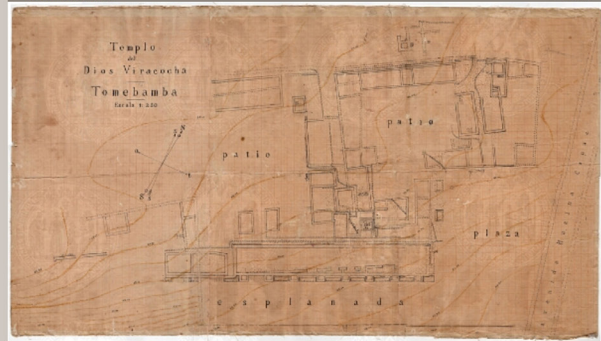
2 As disappointing as the search for the above-mentioned photographs ended, another completely unexpected clue matched it in excitement. I came across it while looking through a bundle with the abbreviation JJC.02525. According to the catalogue entry, it contained: "29 f; paper; miscellaneous; sketches and plans of settlements of several places in Ecuador and Peru; explanatory cross section of the excavations carried out by Jacinto Jijón y Caamaño in the Huaca III in Maranga Peru; in addition an architectural plan of the temple of the god Viracocha in

Supplementary online content of the article. This content is created by the author, peer-reviewed and edited by the editorial office of the DAI.

For further information s. [https://arachne.dainst.org/project/JGA\\_2020\\_Pruemers](https://arachne.dainst.org/project/JGA_2020_Pruemers)

Plan of Tomebamba

Focus



Plan von Tomebamba / Plan of Tomebamba

Cuenca, Tomebamba

Link to IDAI objects: <http://arachne.dainst.org/entity/6915466>

3

Abb. 3: Beispiel für einen Datensatz, der ergänzend zu einem Artikel in einem strukturierten Konvolut in iDAI.objects/Arachne veröffentlicht wurde und dessen Vorschau unter »Supplements« in der Viewer-Version des Artikels angezeigt wird (Screenshot: Peter Baumeister)

16 Ein Sonderfall bietet [iDAI.bibliography](#): Hier wurde beschlossen, zunächst auf die dort existierenden Datensätze zu verlinken, um alle relevanten Funktionen des Fachsystems wie Export in verschiedene Datenformate etc. nutzen zu können (Abb. 4).

17 Der Quellcode zu den oben genannten Anpassungen ist auf GitHub verfügbar<sup>17</sup>.

## Die Druckausgabe

18 Aus dem oben geschilderten Ziel, dass eine hochwertige Druckpublikation weiterhin zur Verfügung gestellt werden muss, leitete sich die Anforderung ab, eine gelayoutete Ausgabe in Form eines PDF-Formats vorzusehen. Der Aufwand, die in archäologischen Publikationen oft komplexe Korrelation von Texten, Bildern, Abbildungstableaus, maßstäblichen Zeichnungen usw. mit alternativen Textsatzsystemen oder darstellenden Auszeichnungssprachen umzusetzen (LaTeX, Print-CSS, XSL-FO o. Ä.), wurde dabei im Vorfeld als nur schwer kalkulierbar eingeschätzt<sup>18</sup>. Da zudem das Risiko als zu hoch erschien, für beide Ausgabemedien gleichzeitig völlig neue Prozesse einzurichten, zumal vor dem Hintergrund der eng getakteten Erscheinungstermine kein Puffer für Fallback-Lösungen existierte, wurde die Entscheidung getroffen, die Druckausgabe in einer Übergangsphase zunächst weiterhin auf herkömmliche Weise in einem Satzprogramm vorzubereiten und damit im Falle des AA für den Satz der PDF-Ausgabe weiterhin InDesign zu verwenden.

## Das Aufbereiten der Manuskriptdaten für Viewer- und die Druck-Ausgabe

19 Bei der Umsetzung wurde zwar darauf geachtet, die Abläufe möglichst im Hinblick auf die Bedürfnisse der technischen Prozesslogik zu entwerfen. Diese Absicht musste allerdings insofern wieder eingeschränkt werden, als andere wichtige Faktoren – allen voran Erwartungen der Autor:innen an den Schreib- und

17 <https://github.com/dainst/dai-lens>. Desiderate bleiben einstweilen die Implementierung für iDAI.geoserver, das iDAI.world-System für Geodaten (<https://geoserver.dainst.org>), sowie eine Erweiterung der Bandbreite anzeigbarer Dateiformate (3D-Modelle, Filme usw.), wobei aber noch unklar ist, ob eine technische Umsetzung möglich ist.

18 Siehe dazu auch Anm. 31, 32 und bes. 33.



Redaktionsprozess – ebenso zu berücksichtigen waren. Ein entscheidender Punkt dabei war, dass nach wie vor viele Autor:innen gewohnt sind, nicht nur ihre Manuskripte in MS Word zu erstellen, sondern auch eine gelayoutete Ausgabe zu Korrekturzwecken zu erhalten, was bedeutet, dass die Korrekturläufe zwischen Redaktion und Autor:in in gelayouteten PDFs zu bewerkstelligen sein müssen, da alle anderen Varianten wenig arbeitsökonomisch wären. Daher wurde eine Reihenfolge gewählt, in der in der .pdf-Korrekturphase keine mehrfachen und fehleranfälligen Konvertierungsprozesse notwendig sind, was nur zu erreichen ist, wenn das Quellformat, aus dem beide Ausgabeformate (.pdf und .xml) generiert werden, die in InDesign vorbereitete Satzdatei ist. Dies soll allerdings ausdrücklich nur eine mittelfristige Lösung darstellen, bis andere Formen des Austauschs zwischen Autor:in, Redaktion und technischer Aufbereitung sich zu etablieren beginnen.

## Tagging-Wortschatz

20 Aufgrund der Komplexität des .xml-Formats, aus dem der Viewer die responsive Web-Ausgabe generiert, wurde darauf geachtet, dass bei der ganzen Prozesskette die formalen Anforderungen dieses End-Formats von Anfang mitgedacht werden, um mögliche strukturelle Inkongruenzen und damit Fehler bei der Konvertierung oder der Darstellung zu vermeiden. Daher wurde in der Satzdatei ein umfangreicher Tagging-Wortschatz in Form von Zeichen- (77), Absatz- (43) und Objektformaten (13) angelegt, bei dem die Kategorien so definiert sind, dass die satztypischen Formatangaben beim Export in die notwendigen .xml-Elemente umgewandelt werden können (Abb. 5).

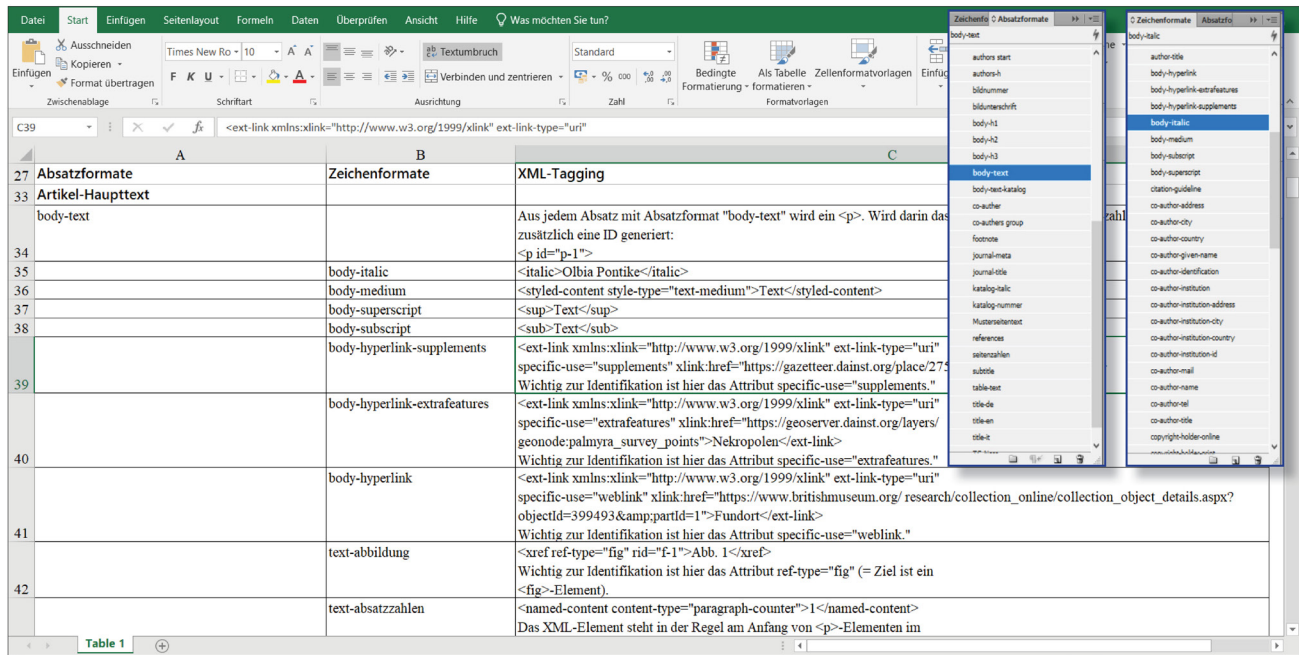
21 So hoch diese Zahlen zunächst klingen, so dienen allerdings nur die wenigsten Formatvorlagen davon zur Auszeichnung der Texte selbst: Der allergrößte Teil dieser Formatvorlagen der Satzdatei – insgesamt 62 – sind für die korrekte metadatentechnische Beschreibung des Beitrages bzw. der Zeitschriftenausgabe notwendig. Das bedeutet, dass es sich in großen Teilen um denselben Aufwand handelt, der auch bei klassischen Print-Publikationen anfällt – bis auf den entscheidenden Unterschied, dass dieser bisher in der Regel erst nach der Veröffentlichung eines Bandes oder Beitrages und zudem zumeist in den Bibliotheken anfiel.

## Die Produktionsstrecke .indd nach .pdf und .xml

22 An die Satzherstellung und damit an die Erstellung der Druck-PDFs werden keine besonderen Anforderungen gestellt, außer dass die Auszeichnung sämtlicher Elemente (Textelemente, aber auch Text- und Bildrahmen usw.) sehr präzise sein muss. Im Falle des Archäologischen Anzeigers wird die korrekte Auszeichnung der Artikel bzw. der einzelnen Elemente mit sämtlichen Formaten vom Lektorat in einem Word-Dokument vorgenommen (siehe dazu unten unter Textverarbeitung). Anschließend wird der Satz der Artikel von einer Prepress-Firma durchgeführt, der eine



Abb. 4: Beispiel für einen verlinkten bibliographischen Datensatz in IDAI.bibliography (Screenshot: Peter Baumeister)



5

Abb. 5: Auszug aus dem Tagging-Wortschatz, eingeblendet die korrespondierenden Absatz- und Zeichenformate der Satzdatei (Screenshot: Peter Baumeister)

entsprechende InDesign-Musterdatei mit den eingerichteten Formatvorlagen sowie die zur Konvertierung nach .xml notwendigen XSLT-Skripte übergeben wurden<sup>19</sup>.

23 Die Generierung des .xml (JATS)-Dokumentes, das vom Viewer angezeigt wird, aus der InDesign-Datei erfordert verschiedene Konvertierungsschritte, die von XSLT-Skripten gesteuert werden; die Skripte sind auf GitHub verfügbar<sup>20</sup>. An zwei Stellen helfen Preflight-Kontrollen, die Prozesse zu überwachen. Der Ablauf gestaltet sich folgendermaßen:

- Export des Dokumentes aus InDesign von .indd nach .xhtml.
- Erster Preflight-Test: Geprüft werden Strukturierung sowie die einzelnen Elemente, anschließend generiert die Produktionsstrecke ein sehr ausführliches Prüfprotokoll in Form eines .html-Reports (Abb. 6)<sup>21</sup>.
- Es folgen drei Schritte, bei denen die Elemente schrittweise nach .xml (JATS) konvertiert werden.
- Zweiter Preflight-Test: Auf Basis des .xml (JATS)-Dokumentes wird eine .html-Preview des Inhalts generiert, um die strukturelle Korrektheit des Dokumentes vor dem Import in den Viewer ein letztes Mal zu prüfen.

24 Diese gesamte Produktionsstrecke kann von einer Satzfirma völlig autonom bearbeitet werden, was bedeutet, dass für diese komplexen Schritte nicht zwingend

19 Mangels Open Source-Alternativen nutzt das DAI für diesen Teil der Produktionsstrecke den Oxygen XML Editor. Zum Problem freier XML-Editoren siehe u. Anm. 34.

20 <https://github.com/dainst/dai-jats-produktionsstrecke>. Entwickelt von Fabian Kern, digital publishing competence, München (DPC).

21 Allgemeine Prüfungen der InDesign-Dokument-Struktur, Prüfung der Vollständigkeit aller Arteikelemente, Prüfung Textrahmen und Formate im Artikel-Haupttext, Prüfungen der Elemente für die JATS-Metadaten, Prüfungen von Abstract, Abstract-Übersetzungen und Keywords, Prüfungen des Title-Textrahmens und der enthaltenen Absatz-Formate, Prüfungen von Abbildungen und Abbildungsnachweis, Prüfungen der Referenzen, Prüfungen von Autor:innen und Autor:innen-Metadaten.

# Prüfbericht für InDesign-Export zur Verarbeitung in JATS-XML

Artikel: Vor den Toren der Stadt

Autoren: JOCHEN FORNASIER, ALLA V. BUJSKICH, ALEKSANDR G. KUZ'MIŠČEV, ARNO PATZELT, MARKUS HELFERT, NATHALIE KRATZSCH

Ausgabe: 1/2019

Konvertiert: 12.11.2019, 14:21:44

#	Prüfung	Ergebnis	Typ
1.	Dokument-Struktur	Allgemeine Prüfungen der InDesign-Dokument-Struktur	Bereich
1.1	Content-Image	Content-Image gefunden: OK	Info
1.2	Abstract	Abstract gefunden: OK	Info
1.3	Abstract-Translation	1 mal Abstract-Translation gefunden: OK	Info
1.4	Title-Container	Title-Container gefunden: OK	Info
1.5	Artikel-Body	Artikel-Body gefunden: OK	Info
1.6	Artikel-Bilder	44 Bildelemente mit dem Objektrahmen-Format 'picture' gefunden: OK	Info
1.7	Referenzen	Referenzen gefunden: OK	Info
1.8	Abbildungsnachweis	Abbildungsnachweis gefunden: OK	Info
1.9	Autoren	Autoren-Angaben gefunden: OK	Info
1.10	Artikel-Metadaten	Artikel-Metadaten gefunden: OK	Info
1.11	Journal-Metadaten	Journal-Metadaten gefunden: OK	Info

6

inhouse eine besondere Kompetenz vorhanden sein muss. Allerdings zeigt sich, dass die angezeigten Fehler oftmals mit unkorrekten Auszeichnungen der Word-Version zusammenhängen, so dass die bzw. der bearbeitende Redakteur:in die Ursachen deutlich schneller ermitteln kann als die Prepress-Firma und daher eine enge Zusammenarbeit empfohlen wird.

Abb. 6: Ausschnitt aus dem Prüfprotokoll des ersten Preflight-Tests der XSLT-Konvertierungsstrecke (Screenshot: Peter Baumeister)

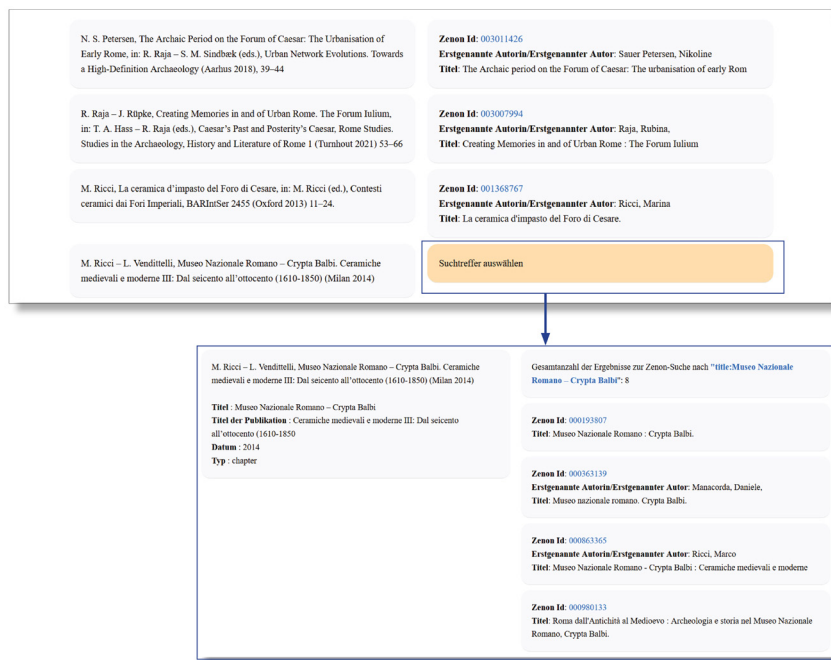
## Maßnahmen zur Aufwandsreduktion

### Semi-automatisiertes Anreichern der Publikationstexte mit Normdaten

25 Die Anreicherung der Publikationstexte mit Referenzen auf Inhalte anderer Informationsressourcen wie Objektdatenbanken oder Normdatenverzeichnisse führt im Vergleich zu bisher etablierten Publikationsworkflows gleich auf zwei Ebenen zu einem Mehraufwand: Zum einen müssen die Einträge der zu referenzierenden Entitäten in den Nachweissystemen bzw. Verzeichnissen gesucht und deren Identifier extrahiert werden. Zum anderen müssen die Verweisstellen im Text mit dem korrekten Format ausgezeichnet und mit dem extrahierten Identifier bzw. der URI/URL hinterlegt werden. Je nach Fallzahl führt dies zu einem nicht unerheblichen Mehraufwand.

26 Ziel ist daher, diese Prozeduren wie das Heraussuchen von IDs oder die händische Auszeichnung von Verweisstellen im Text sukzessive zu semi-automatisieren, um signifikante Zusatzaufwände zu vermeiden, da diese eine Hürde für eine dauerhafte Akzeptanz des Modells und damit dessen Übertragbarkeit darstellen.

27 Ein Beispiel dafür sind Literaturverweise und -verzeichnisse, die von den Autor:innen nach wie vor in den meisten Fällen in Form lediglich typographisch formatierter Text-Dokumente den Artikeln beigelegt werden und die keine Verweise auf Verzeichnissysteme wie iDAI.bibliography o.Ä. enthalten. Auch wenn davon auszugehen ist, dass Autor:innen mittel- und langfristig bibliographische Angaben zunehmend in strukturierter und mit relevanten Identifiern angereicherter Form einreichen werden, so müssen in der momentanen Übergangsphase kurzfristig semi-automatisierte Prozesse helfen, den Aufwand möglichst zu reduzieren.



## Ein erstes Modul: Der iDAI.bibliography-assistant

Zu diesem Zweck hat das DAI in Zusammenarbeit mit Daniel de Oliveira einen webbasierten Bibliographie-Assistenten entwickelt<sup>22</sup>. Die Web-Version wurde über das auf Elixir basierende Web-Development-Framework Phoenix realisiert; als Standard-Parser wird AnyStyle verwendet<sup>23</sup>. Nach dem Kopieren der Literaturliste in das Eingabefeld der Anwendung parst das Werkzeug die Verweise, übergibt die segmentierten Einträge als JSON-Format und gleicht sie über die existierende Zenon-API mit Einträgen aus iDAI.bibliography ab. Danach zeigt das Tool konkordanzartig Treffer an bzw. bietet im Falle ambiguer

7

Abb. 7: Der iDAI.bibliography-assistant zum semi-automatisierten Anreichern bibliographischer Normdaten aus iDAI.bibliography/Zenon (Screenshot: Peter Baumeister)

Matches zu bestätigende Auswahlmöglichkeiten (Abb. 7). Abschließend kann die mit iDAI.bibliography-IDs angereicherte Ausgangsliste exportiert werden.

Im laufenden Probetrieb wird nach Möglichkeiten gesucht, die DAI-seitig zu beeinflussenden limitierenden Faktoren für die Trefferquote (d. h. Fehler beim Segmentieren, aber vor allem unkorrekte bzw. mehrere abweichende Einträge im Nachweissystem) möglichst arbeitsökonomisch zu verringern, wobei sich aber auch zeigt, dass artikelseitige Unzulänglichkeiten wie inkorrekte Literaturangaben ebenfalls eine relevante Anzahl unter den Fehlern ausmachen.

Bei Artikeln, die sich in einem Themenbereich bewegen, der zu den traditionellen Forschungsschwerpunkten des DAI gehört, erreicht die Benutzung des Assistenten bereits jetzt sein Ziel, den Aufwand für die bearbeitende Redaktion maßgeblich zu reduzieren, da die referenzierte Literatur im Verzeichnisbestand der Bibliotheken des DAI und damit in iDAI.bibliography für das Tool auffindbar ist. Dies führt zu der konzeptionellen Frage, wie mit Artikeln umgegangen werden soll, die auf Literatur verweisen, die nicht traditionell zum Erwerbungs-schwerpunkt des DAI gehört, da die Anreicherung von Artikeln mit Normdaten perspektivisch an Bedeutung zunehmen wird. Die Referenzierung auf Inhalte anderer Verzeichnissysteme wäre eine Option, hätte allerdings verschiedene Nachteile und würde u. a. bedeuten, den Verweis auf fehlerhafte Einträge hinnehmen zu müssen, da Korrekturen nicht selbst vorgenommen werden könnten. Es ist daher unbedingt zu präferieren, iDAI.bibliography mit den fehlenden Einträgen im Redaktionsprozess anzureichern. Damit das in einem bibliotheksseitig vertretbaren Aufwand erledigt werden kann, böte sich der Bibliography-Assistent als technisches Instrument an, um Teilschritte zu erledigen, da die entsprechenden Referenzen geparkt im JSON-Format durch den Bibliography-Assistent bereitgestellt werden und so auf einfache Weise nach iDAI.bibliography gepusht werden könnten, wo im Anschluss eine Validierung vonseiten der Bibliothek erfolgen müsste.

<sup>22</sup> Die Entwicklung wurde begleitet von Marcel Riedel. Das Tool wird zur Zeit noch getestet und ist daher noch nicht freigeschaltet; der Quellcode ist aber bereits verfügbar, siehe Anmerkung 23.

<sup>23</sup> <https://github.com/dainst/bibliography-assistant>. Siehe auch <https://anystyle.io/> bzw. <https://github.com/inukshuk/anystyle>. Optional kann auch [Grobid](https://www.grobid.org/) eingebunden werden. Zu Phoenix siehe <https://www.phoenixframework.org>.



## Textverarbeitung und Ansätze für skriptbasierte Bearbeitung

31 Um bereits möglichst früh im Prozess die korrekten Formate zu verwenden, wurde eine .dotx-Formatvorlage für die Bearbeitung in MS Word erstellt, die nicht nur im vorbereitenden Redaktionsprozess genutzt wird, sondern die auch den Autor:innen zur Verfügung gestellt wird. In ihr sind sämtliche Elemente des benötigten Tagging-Wortschatzes in entsprechenden Formatvorlagen abgebildet, die beim Import in die Satzdatei in einfacher Weise den dortigen InDesign-Formatvorlagen zugewiesen werden können.

32 Neben der oben skizzierten Anreicherung von Normdaten usw. entsteht bei der redaktionellen und formalen Vorbereitung der Texte insofern noch ein weiterer zusätzlicher Aufwand, da es im Vergleich zu Artikeln, die für eine einfache Druckausgabe vorbereitet werden, mehr Bestandteile gibt, denen Meta-Informationen beigelegt werden müssen. So besteht etwa für Druckausgaben keine Notwendigkeit, einzelne Bestandteile von Abbildungsnachweisen o. Ä. separat auszuzeichnen, für maschinenlesbare Ausgabeformate aber schon. Das führt zu dem, dass die Zahl der auszuführenden Auszeichnungsvorgängen höher ist.

33 Zum anderen müssen die Auszeichnungen selbst auch sorgfältiger erfolgen bzw. es ist mit größerem Korrekturaufwand zu rechnen, da unvollständige oder unkorrekte Auszeichnungen, die bei der Vorbereitung von Printpublikationen nur typographische und damit oftmals tolerierbare Abweichungen im gedruckten Endprodukt hervorrufen, in der formatübergreifenden Prozesskette des Anzeiger-Modells aber entweder zu einem fehlermeldungsbegleiteten Stopp von Konvertierungsvorgängen oder zu Fehlern in der Anzeige führen.

34 So naheliegend die Lösung einer skriptgesteuerten Bearbeitung der Text-Dokumente in der Theorie auch ist, umso schwieriger gestaltet sich ein praktikables Konzept für deren Umsetzung, da eine herausgebende Institution ihre Prozesse nicht allein auf die technische Umsetzbarkeit ausrichten kann, sondern – wie geschildert – auch auf andere Faktoren wie die fachlich etablierte Publikationskultur und damit auch auf die Erwartungen, die die Autor:innen an die Vorbereitung wissenschaftlicher Publikationen und deren Erscheinungsform haben, in angemessener Form Rücksicht nehmen muss. Das bedeutet:

1. In Bezug auf individuelle Gewohnheiten: Einigen Akteur:innen, die zentrale Rollen im Prozess spielen, muss man die Möglichkeit zugestehen, zumindest in einer Übergangsphase mit ihren gewohnten Formaten weiter zu arbeiten (siehe auch oben zu Umbruch-PDFs). Dies gilt vor allem für Autor:innen, die bestimmte Erwartungen an ihre Art der Vorbereitung der Texte stellen, kann aber auch für jene Redakteur:innen zutreffen, die vor allem aufgrund ihrer fachwissenschaftliche Expertise in den Redaktionsprozess involviert sind und denen man in einer Übergangsphase ebenfalls Zeit für eine Umstellung in neue Formen des wissenschaftlichen Publizierens einräumen muss.
2. In Bezug auf fachliche Gewohnheiten der wissenschaftlichen Argumentationsführung: Für die Archäologie und Altertumswissenschaften ist vorauszusetzen, dass Texte zunächst weiterhin mit komplexen Fußnotenapparaten verschränkbar bleiben müssen, was bedeutet, dass man auf Formate von Anwendungen, die solche dynamischen Referenzierungen abbilden können, zurückgreifen muss. Mit anderen Worten: Es ist zu antizipieren, dass man vermutlich noch eine Weile in der Lage sein muss, Inhalte, die im

.docx-Format eingereicht werden, in der ein oder anderen Weise verarbeiten zu können.

3. In Bezug auf fachliche Gewohnheiten bei der Darstellung des Abbildungsapparates: Wie eingangs dargelegt, bleibt – ebenso für eine Übergangszeit mangels alternativer Möglichkeiten – die Anforderung bestehen, wissenschaftliche Beiträge in einem Satzprogramm zu formatieren, um ein hochwertiges Layout im Hinblick auf erwartete Muster der Text-Bild-Kombination erstellen zu können, was bedeutet, dass an irgendeiner Stelle im Prozess Inhalte in einem .indd/.idml-Format (siehe oben) bearbeitet werden müssen.

35 Das führt zu der Frage: In welchem der genannten Formate bzw. im Kontext welcher Anwendung bzw. mit welcher Informationstechnologie kann eine skriptgesteuerte Aufbereitung am sinnvollsten stattfinden?

36 Das .indd/.idml-Format wäre theoretisch sehr gut dafür geeignet: Anwendungen wie InDesign bieten nicht nur eine Umgebung für die Bearbeitung der Dokumente mit JavaScript oder grep, hinzukommt auch, dass es viele spezialisierte Dienstleister gibt, die in der Lage wären, auch sehr spezifische Anforderungen umzusetzen. In Hinblick auf Korrekturläufe usw. wäre ein zusätzlicher Vorteil, dass die automatisierte Formatierung bereits in dem Format erfolgen würde, aus dem beide Ausgabemedien direkt generiert würden.

37 Dagegen sprechen jedoch zwei gewichtige Argumente: Zum einen leitet sich aus den Anforderung an die Zurverfügungstellung wissenschaftlicher Inhalte ab, dass bei der Formatierung von Publikationsinhalten Satzprogramme nur noch eine Übergangslösung darstellen, da bei wissenschaftlichen Publikationen maschinenlesbare Ausgaben in Zukunft im Vordergrund stehen werden. Das bedeutet, dass perspektivisch für eventuell notwendige .pdf-Ausgaben Anwendungen oder Formate genutzt werden, die besser zu diesen Workflows passen müssen (LaTeX, Print-CSS, XSL-FO o. Ä), so dass umfangreichere Investitionen in dieser Übergangsphase gut überlegt sein müssen. Vor allem würde es mangels Open Source-Alternativen zum anderen dazu führen, einen hohen Entwicklungsaufwand in Maßnahmen zur Bearbeitung proprietärer Formate zu investieren, was dem eigenen Anspruch an quelloffene Entwicklungen zuwider laufen würde und was nur in Betracht gezogen werden sollte, wenn pragmatische Alternativen fehlen.

38 Ähnlich verhält es sich mit .docx-Formaten, für die es zweifellos ebenfalls sehr mächtige Skriptmöglichkeiten zur Bearbeitung gibt. Doch auch hier stellt sich in gleicher Weise die Frage, ob es sinnvoll ist, angesichts der oben skizzierten Ziele Ressourcen in Technologien zu investieren, die man nicht dauerhaft nutzen möchte. Dies gilt auch für die notwendige Expertise im Umgang damit: Es wäre ein hoher Aufwand nötig, alle am Prozess Beteiligten für die Arbeit mit spezialisierten und proprietären Formate zu schulen, wissend, dass es sich nur um Übergangslösungen handelt.

*Ein Vorschlag zur Diskussion: .html als Zwischenprodukt?*

39 Eine sehr pragmatische Ansatz könnte sein, die automatisierte Bearbeitung in einem Format stattfinden zu lassen, das deutlich einfacher strukturiert ist als .docx

oder .idml, nämlich .html<sup>24</sup>. Abgesehen von seinen spezifischen Eigenschaften hat das .html-Format vor allem auch den entscheidenden Vorteil, dass viele Akteur:innen ohnehin bereits aus anderen Kontexten mit dessen Umgang vertraut sind. Daher kann das .html-Format als Zwischenprodukt garantieren, dass auch Akteur:innen mit allgemeinen Grundkenntnissen der Syntax in der Lage sind, die skriptgesteuerten Bearbeitungsschritte nachzuvollziehen. Vor allem können Fehler schneller identifiziert und einfacher korrigiert werden, als dies in komplexer strukturierten Formaten der Fall ist, bevor der automatisiert aufbereitete Inhalt zur weiteren Verarbeitung und Finalisierung für die Ausgabemedien in proprietäre Formate zurückkonvertiert wird.

40 Eine einfache Anwendung, die testweise bereits zur Vorstrukturierung der Zeitschriftenbeiträge eingesetzt wird, zeigt, dass dieses Konzept grundsätzlich funktionieren kann<sup>25</sup>. Man könnte demgemäß folgenden Vorschlag für ein Szenario skizzieren:

41 Schritt 1: Im .docx-Format eingereichte Manuskripte werden mit pandoc nach .html konvertiert. Dabei ist der Einsatz von pandoc nicht nur aufgrund seiner Funktion als Konvertierungstool von zentraler Bedeutung: Pandoc kann vielmehr auch insofern eine entscheidende Rolle im Prozess spielen, als es die Dokumente für die von ihm durchgeführte Konvertierung standardisiert<sup>26</sup>, d. h. es gibt Dokumente aus, die mit einer sehr reduzierten und vor allem vereinheitlichten Spektrum an Auszeichnungen versehen sind. Diese Art der Normierung stellt eine wichtige Voraussetzung für nachfolgende erfolgreiche automatische Suchen/Ersetzen-Operationen dar.

42 Schritt 2: Im .html-formatierten Text werden die einzelnen Bestandteile über einfaches Suchen/Ersetzen mit den Elementen ausgezeichnet, die gemäß dem oben geschilderten Tagging-Wortschatz notwendig sind. Dabei können nicht nur Zeichen- und Absatzformate automatisch festgelegt werden oder die notwendige Absatznummerierung durchgeführt werden. Vielmehr können hier mittels einfacher Wertelisten auch die aufwendigen Verlinkungen zu den Normdatensystemen erfolgen: So kann eine Tabelle, die vom iDAI.bibliography-assistant mit Zenon-IDs angereichert wurde (Autor/innen – Titel – Zenon-ID), als Werteliste dienen, auf Basis derer automatisch die Einträge der Literaturliste mit entsprechenden Links hinterlegt, aber vor allem die zahlreichen Autor-Jahr-Zitate in den Fußnoten getaggt werden. Dasselbe gilt auch für die Bildnachweise und Abbildungsunterschriften, deren Bestandteile ebenfalls jeweils separat ausgezeichnet werden müssen. Auch hier können sehr einfache Tabellen, die im .ods-Format vorbereitet und in .csv-Listen konvertiert werden, als Wertelisten dienen.

43 Die Syntax der Auszeichnungen ist so definiert, dass die Tags beim Öffnen von einer aktuellen Microsoft Word-Version als korrekte Formatvorlagen des Anzeiger-Modells interpretiert werden (Abb. 8).

44 Schritt 3: Nach der Überarbeitung und automatischen Auszeichnung kann der Text in Word geöffnet und wie gewohnt redigiert und weiterverarbeitet werden. Da die Bestandteile bereits mit den korrekten Formatvorlagen ausgezeichnet sind (siehe oben),

---

24 In anderen Kontexten wurde bereits darauf hingewiesen, dass .html als Zwischenformat eine sinnvolle Rolle bei Konvertierungsprozessen zwischen .docx und .xml (JATS) spielen kann, siehe beispielsweise Piez 2017.

25 <https://doi.org/10.5218/zenodo.7051409>.

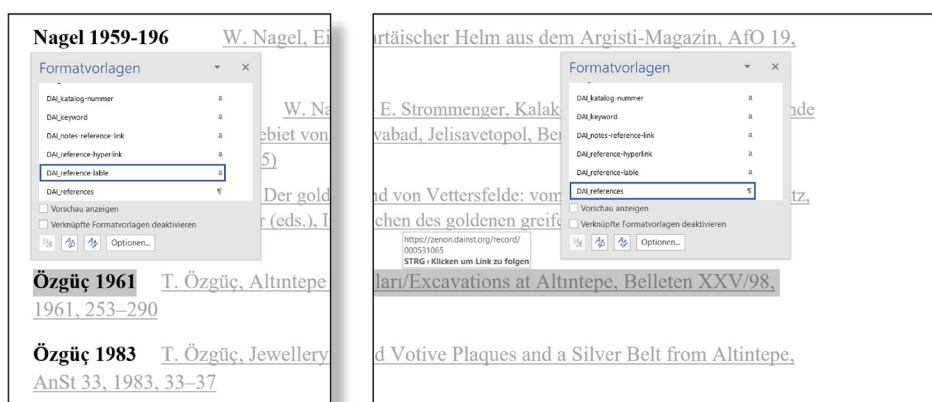
26 Das Problem der Ausgangsdaten aus Word beschreibt Piez 2017 folgendermaßen: »Much as authors labor over documents in MS Word, any ›random Word document‹ turns out, regrettably, as often as not to be lamentably far from ready for ingest into any system much less a publishing system.«

	A	B	C
1	Author/Year	Full citation	Identifier
	Nawroth 2007	M. Nawroth, Der goldenfund von Vettersfelde: vom Schwarzen Meer in die Lausitz, in: W.Menghin – H. Parzinger (eds.), Im zeichen des goldenen greifen. Königsgräber der Skythen, (München 2007) 318–329	
29	Özgüç 1961	T. Özgüç, Altintepe Kazıları/Excavations at Altintepe, Belleten XXV/98, 1961, 253–290	<a href="https://zenon.dainst.org/Record/000531065">https://zenon.dainst.org/Record/000531065</a>
30	Özgüç 1983	T. Özgüç, Jewellery, Gold Votive Plaques and a Silver Belt from Altintepe, AnSt 33, 1983, 33–37	<a href="https://zenon.dainst.org/Record/000486034">https://zenon.dainst.org/Record/000486034</a>
31			

1. Auszug aus der Werteliste, bei der die Zenon-IDs aus dem Suchergebnis des iDAI.bibliography-assistant übernommen wurden

```
<p class=DAIreferences><span class=DAIreference-table><span lang=EN-GB style='mso-ansi-language:EN-GB'>Özgüç 1961<span style='mso-tab-count:1'> </span></span></span> <a href="https://zenon.dainst.org/Record/000531065"><span class=DAIreference-hyperlink><span lang=EN-GB style='mso-ansi-language:EN-GB'>T. Özgüç, Altintepe Kazıları/Excavations at Altintepe, Belleten XXV/98, 1961, 253–290</span></span></a><span lang=EN-GB style='mso-ansi-language:EN-GB'> </span></span></p>
<p class=DAIreferences><span class=DAIreference-table><span lang=EN-GB style='mso-ansi-language:EN-GB'>Özgüç 1983<span style='mso-tab-count:1'> </span></span></span> <a href="https://zenon.dainst.org/Record/000486034"><span class=DAIreference-hyperlink><span lang=EN-GB style='mso-ansi-language:EN-GB'>T. Özgüç, Jewellery, Gold Votive Plaques and a Silver Belt from Altintepe, AnSt 33, 1983, 33–37</span></span></a><span lang=EN-GB style='mso-ansi-language:EN-GB'> </span></span></p>
```

2. Auszug aus der auf Basis der Werteliste bearbeiteten .html-Datei mit den entsprechenden Tags und den hinterlegten Zenon-IDs



3. Erkennen der korrekten Formatvorlagen nach Öffnen der .html-Datei in MS Word

8

Abb. 8: Automatisiert ausgezeichnete .html-Version und Übernahme der korrekten Formatvorlagen in MS Word (Screenshot: Peter Baumeister)

können die Texte nach Abschluss der fachwissenschaftlich-redaktionellen Bearbeitung in InDesign importiert werden, das die Formate automatisch mappt.

45 Dieses Szenario wird zur Zeit bei einzelnen Artikeln auf seine grundsätzliche Praxistauglichkeit hin evaluiert. Noch ist unklar, wie sinnvoll es ist, mit .html noch ein weiteres Format einzuführen. Letztendlich könnte man auch die sehr grundsätzliche Frage aufwerfen, welche Notwendigkeiten überhaupt bestehen, an .xml als Zielformat festzuhalten bzw. ob nicht auch einfacher strukturierte Formate wie .html alle notwendigen Anforderungen erfüllen können, zumal diese mit deutlich weniger Aufwand zu verarbeiten sind.

46 Das bedeutet, dass zeitnah auch andere Ansätze sorgfältig geprüft werden müssten, bevor man sich auf eine langfristige Lösung festlegt, die dann idealerweise in einer webfähigen Anwendung aufgesetzt werden müsste.

## Aufbereiten von Forschungsdatenkonvoluten für eine supplementäre Publikation

47 Der Aufwand, der für die Aufbereitung der Forschungsdatenkonvolute für eine supplementäre Publikation notwendig ist, ist am schwierigsten zu beziffern, da dieser von sehr vielen Faktoren abhängig ist und teilweise sehr individuelle Lösungen erfordert. Im Wesentlichen lassen sich zur Zeit zwei Anforderungsszenarien unterscheiden: Autor:innen möchten Daten veröffentlichen,



1. die in Datenmodellen strukturiert sind, die prinzipiell auf einen Austausch und potenzielle Nachnutzung hin konzipiert wurden oder die sich bereits in der iDAI.world befinden;
2. die sehr spezifisch im Hinblick auf die Anforderungen des jeweiligen Forschungsprojektes hin modelliert sind und die erst nachträglich in die iDAI.world-Systeme eingepflegt werden.

48 Im ersten Fall entsteht entweder kein oder nur geringer zusätzlicher Aufwand, da die Daten bei ihrer Aufzeichnung oder Erzeugung entsprechend prä-modelliert<sup>27</sup> wurden bzw. die Strukturierung ein einfaches Mapping erlauben. Im zweiten Fall kann je nach Zustand der Ausgangsdaten ein mitunter sehr großer Aufwand entstehen, da die Ausgangsdaten/Formate der lokalen Datensammlung eines Projektes für eine supplementäre Publikation zunächst re-modelliert werden müssen, bevor sie in das Fachsystem migriert werden können. Die Lösungen, wie diese Aufwände zu reduzieren sind, müssen zwangsläufig ebenfalls individuell sein, was kurz- und mittelfristig ein ungelöstes Problem des Modells darstellt.

49 Für den Fall, dass für eine Forschungsdatenpublikation eine Aufbereitung und ein Import von Daten in die iDAI.world-Systeme notwendig ist, hat das DAI bereits begonnen, Tools zu entwickeln, die den Aufwand erheblich reduzieren. Ein Beispiel dafür ist die ArachneImportAppX (AIAX), die einen semi-automatisierten Import von Datenbeständen in die zentrale Objektdatenbank des DAI iDAI.objects/Arachne ermöglicht, auch wenn je nach Zustand des Quellbestands ein mehr oder weniger großer Zusatzaufwand bestehen bleibt<sup>28</sup>. Das Tool schließt aber nur eine der Lücken, da perspektivisch auch Lösungen für andere iDAI.world-Systeme gefunden werden müssen. Das DAI arbeitet zudem zur Zeit an der Implementierung eines Systems mit klassischem Repositoriencharakter, das auch eine niedrigschwelligere Publikation von Datensupplementen erlauben wird, sofern der hohe Aufwand einer Strukturierung für die anderen Systeme durch das Projekt nicht geleistet werden kann – was dann allerdings auf Kosten der Kontextualisierung mit anderen Datensätzen in der iDAI.world gehen wird.

## Fazit

### Der Blick zurück

50 Die Arbeiten am Konzept haben im Frühjahr 2018 begonnen. Seitdem ist nicht nur die Zahl an Publikationen, die Vorschläge für die Modellierung von Open-Access-Workflows im Publikationsbereich machen<sup>29</sup>, sondern auch die der Initiativen und Vorhaben, die infrastrukturelle Lösungsansätze zu entwickeln beginnen, stark angestiegen<sup>30</sup>. Im Rückblick zeigt sich allerdings, dass der heutige Kenntnisstand bei derselben Zielsetzung vermutlich nur bedingt Einfluss auf die Entscheidungsfindung im Rahmen des Anzeiger-Projektes gehabt hätte, da spezifische Fragen zur Ausdetaillierung

27 Diese Formulierungen verdanke ich Marcel Riedel.

28 Entwickelt von Marcel Riedel. Siehe dazu Boecker et al. 2021, § 20.

29 Allen voran Böhm et al. 2020.

30 Exemplarisch seien folgende Projekte genannt, die erst nach Abschluss des AA-Projektes gestartet wurden: Das vom BMBF geförderte Projekt »Open Source Academic Publishing Suite« (OS-APS), aber auch das Projekt »OPEN-ACCESS-Strukturierte-Kommunikation (OA-STRUKTKOMM)«, das »Werkzeuge und Strategien zur Beschleunigung der Transformation des wissenschaftlichen Publikationssystems hin zu Open Access« inkl. eines »automatisierten XML-first-Satzsystems zur Herstellung von OA-Monografien« entwickelt (<https://doi.org/10.5281/zenodo.5526686>). Zum Projekt <https://libero.pub> siehe unten Anm. 31.

bestimmter Prozessschritte, Standards usw., die oben angesprochen wurden, entweder ausgeklammert, nicht in der notwendigen Detailtiefe betrachtet wurden<sup>31</sup> oder im Falle der aktuelleren Projekte als noch umzusetzende Desiderate formuliert werden<sup>32</sup>. Das gilt letztendlich auch in Bezug auf Projekte, die zu Anfang der Konzeptionsphase des AA bereits weit fortgeschritten waren und die in der Zwischenzeit durch einen erfolgreichen Regelbetrieb den Proof of Concept angetreten haben<sup>33</sup>.

51 Eine rückblickende Erfahrung ist auch, dass selbst bei vielversprechenden Open Source-Anwendungen kaum vorhersehbar ist, ob die Intensität der Community-getriebenen Weiterentwicklung hoch genug bleibt, um für das eigene Projekt mit einer einsatzfähigen Version rechnen zu können – auch wenn sehr engagierte Konsortien und Akteur:innen dahinterstehen<sup>34</sup>.

## Abgleich mit den Zielen

52 Im Hinblick auf die Sichtbarkeit von Forschungsdatenpublikationen erfüllt das Modell insofern seine Erwartungen, als die Zahl der Einreichungen von Manuskripten mit ergänzenden Datenkonvoluten zunimmt und zu manchen Artikeln mittlerweile umfangreichere und komplexere Datenbestände publiziert werden<sup>35</sup>. Damit kann das Modell perspektivisch einen Anreiz zur Veröffentlichung der Projektdaten darstellen. Auch zeigt sich, dass immer mehr Autor:innen die Möglichkeit nutzen, ihre Beiträge mit Normdaten anzureichern. Die Tatsache, dass das Konzept, Forschungsdaten in Journal-Artikel einzubinden und somit Inhalte der iDAI.world sichtbar zu machen, in der Lage war, eine Jury zu überzeugen, den Archäologischen Anzeiger auf der Frankfurter Buchmesse 2020 mit einem der Digital Publishing Awards

- 
- 31 Die Fallstudien zum XML-first-Workflow bei Böhm et al. 2020 (in Fremdleistung: 285–300, bzw. in Fremd- und Eigenleistung: 301–316) belegen, dass auch andere Akteure an denselben Stellen der Prozesskette – wenn auch bei etwas abweichender Ausgangslage – auf sehr ähnliche ungelöste Probleme gestoßen sind, was die Einschätzung bekräftigt, dass vielerorts noch Standards gesucht werden bzw. Best-Practice-Lösungen nicht existent sind oder noch nicht so beschrieben wurden, dass sie einfach übertragen werden können. Dies betrifft letztendlich alle Aspekte, die in diesem Artikel beschrieben und diskutiert werden: Vom Umgang mit Stakeholdern wie Autorinnen und Autoren bzw. deren Arbeitsweise (43) über Fragen zur Anreicherung der XML-Daten (310) bis hin zum Problem der Generieren der Ausgabeformate usw. (312).
- 32 Siehe den Antragstext von OS-APS unter [https://os-aps.de/files/2020\\_08\\_07\\_BMBF\\_OS-APS\\_FAU\\_Antrag\\_kurz.pdf](https://os-aps.de/files/2020_08_07_BMBF_OS-APS_FAU_Antrag_kurz.pdf) (02.04.2022), z. B. zum Import von Word-Dokumenten, bei dem Metadaten teilautomatisiert aus den Manuskripten herausgelesen werden (6), Einrichten der Entwicklungsumgebung zur Entwicklung eigener Satztemplates, auch um die »Lücke des fehlenden Satzes« zu schließen. (1) usw.
- 33 Insbesondere zu nennen ist – nicht nur aufgrund des geisteswissenschaftlichen Bezugs – Heidelberg University Publishing (heiUP) und das Heidelberg Monograph Publishing Tool (heiMPT), siehe <https://heiup.uni-heidelberg.de>, <https://github.com/witthanage/heiMPT> (18.04.2022). Siehe dazu, inkl. sehr hilfreicher Erfahrungsberichte, Hacker – Corrao 2017; Effinger et al. 2018. Zur Schwierigkeit einer PDF-Ausgabe ohne InDesign besonders Effinger et al. 2018, 400 f.
- 34 Exemplarisch genannt sei der Texture Editor, ein web-basierter Open Source-XML-Texteditor für wissenschaftliche Dokumente, der als eine musterhafte Lösung für viele Anforderungen konzipiert war (siehe Garnett et al. 2017), deren Hauptentwickler die Arbeit daran allerdings eingestellt haben. Siehe dazu das Fazit eines der Entwickler, der die Herausforderungen beschreibt: <https://letsken.com/michael/online-plattform-entwickeln#hcc54f> (02.04.2022): »For about a decade my big mission was helping scientists publish their research more effectively. We failed – the process was too complex and bureaucratic.« (<https://letsken.com/about> [02.04.2022]). Verfolgt wurden ebenfalls die Entwicklungen im Rahmen der »Libero publishing suite«, da diese maßgeblich von eLife, also jener Institution, die auch den eLife-Lens-Viewer entwickelt hat, betrieben wurde, siehe <https://libero.pub> (02.04.2022) bzw. <https://github.com/libero> (»Open-source platform of services and applications to help publishers do more with everything they publish«). eLife scheint sich aus der Entwicklung zurückzuziehen, weswegen abzuwarten bleibt, ob und wie das Projekt weiterentwickelt wird, vgl. die entsprechende Meldung unter <https://elifesciences.org/inside-elife/daf1b699/elife-latest-announcing-a-new-technology-direction> (02.04.2022).
- 35 Exemplarisch dafür steht ein Beitrag zur Stratigraphie und Fundvorlage einer Sonde in Henchir Bourgou (Djerba, Tunesien), bei dem die Grabungsdokumentation in iDAI.field als digitales Supplement in den Artikel eingebunden wird, siehe <https://doi.org/10.34780/f05j-59fd>.
-

auszuzeichnen (Abb. 9)<sup>36</sup>, macht in erster Linie Hoffnung, dass Lösungen aus der Archäologie und der Altertumswissenschaften den Diskurs auch außerhalb der eigenen Disziplinen bereichern und damit die Suche nach Best-Practice-Lösungen unterstützen können.

53 Nach dem Start des neuen Modells im Jahre 2019 haben mittlerweile fünf weitere wissenschaftliche Zeitschriften des DAI dieses Modell – mit kleineren Abweichungen – übernommen (Abb. 10)<sup>37</sup>. Das zeigt, dass ein Workflow, der zu einem maschinenlesbaren Endprodukt führt, grundsätzlich auch von Redaktionen übernommen werden kann, die ihre Publikationen bisher auf herkömmliche Weise bearbeitet bzw. produziert haben.

54 Ein wichtiges Fazit ist allerdings auch, dass die Prozesse und unterstützenden Werkzeuge noch weiter verbessert werden müssen. Im Hinblick auf die Aufwände und eine mögliche Übertragbarkeit ist daher Folgendes zu konstatieren:

### Maschinenlesbarkeit

55 Für viele der zentralen Komponenten, die nicht nur für die Publikation, sondern auch für die Vorbereitung maschinenlesbarer Inhalte notwendig sind, gibt es – wie gezeigt – quelloffene oder frei nutzbare Anwendungen, die – je nach Anspruch an die Gestaltung und Funktionalität – mit mehr oder weniger Aufwand auf die eigenen Bedürfnisse hin konfiguriert werden können. Allerdings klaffen für die Geisteswissenschaften im allgemeinen sowie für die Archäologie und Altertumswissenschaften im besonderen einige Lücken, da bestehende Lösungen die fachspezifischen Anforderungen nicht vollständig erfüllen bzw. in Teilen auch den Erwartungen an die Abläufe und Endprodukte, die durch die geisteswissenschaftliche Publikationskultur geprägt sind, nicht gerecht werden.

56 Die Zurverfügungstellung einer maschinenlesbaren Ausgabe führt daher im Vergleich zur Vorbereitung und Veröffentlichung herkömmlicher Publikationen zu Mehraufwänden, wobei diese in zwei Arten zu unterscheiden sind:



9



Print- und Online-Ausgaben



Nur Online-Ausgaben

Abb. 9: Der Digital Publishing Award in der Kategorie Prozess/Technologie für das Konzept der in Journal-Artikel eingebundenen Forschungsdaten (Screenshot: Peter Baumeister)

Abb. 10: Zeitschriften des DAI, die das neue Modell bereits übernommen haben (Screenshot: Peter Baumeister)

10

36 Auszeichnung in der Kategorie Prozess/Technologie. In der Jury saßen Vertreter:innen von Google, den Ullstein Buchverlagen, der Frankfurter Buchmesse, smart digits, Kohl Consulting, Berlin Consulting und dem Springer Gabler Verlag. Siehe <https://www.buchreport.de/news/das-sind-die-sieger-des-digital-publishing-award-2020> (02.04.2020), <https://www.dainst.org/presse/pressemitteilung/-/article-display/L11mBpIClzu5/4662694> (02.04.2022). Allgemein zum Award siehe <https://www.dpr-award.com> (09.11.2022).

37 Das [Journal of Global Archaeology](#), die [e-Forschungsberichte](#), das [Forum for Digital Archaeology and Infrastructure](#), die [Madrider Mitteilungen](#) sowie die [Mitteilungen des Deutschen Archäologischen Instituts, Römische Abteilung](#).

1. Mehraufwände, die dadurch entstehen, dass die Ausgabeformate eine höhere Informationsdichte erreichen müssen (Anreicherung mit Normdaten, Verschränken mit bestehenden Forschungsdatenpublikationen usw.) und
2. Mehraufwände, die daraus resultieren, dass die Arbeitsweise der am Prozess Beteiligten noch auf den Umgang mit anderen Formaten ausgerichtet ist.

57 Dass sich erstgenannte Aufwände perspektivisch zuverlässig reduzieren lassen, zeigen Anwendungen wie der iDAI.bibliography-assistant, die auch für andere Aufgaben Problemlösungen werden können. Ziel muss daher sein, diesen Werkzeugkasten sukzessive zu erweitern.

58 Die Umstellung der Arbeitsweise wird hingegen ein längerer Prozess sein, da eine arbeitsökonomische Verarbeitung der Dokumente – insbesondere im Hinblick auf die zahlreichen notwendigen und rekursiven Übergabevorgänge (Autor:in <> Redakteur:in <> technische Aufbereitung usw.) – nur möglich ist, wenn die Zahl der notwendigen Konvertierungs- und Formatierungsschritte deutlich reduziert wird. Das setzt allerdings voraus, dass alle am Prozess Beteiligten auf ähnlichem Kenntnisstand und mit einheitlichen Standards operieren. Anwendungen, die in anderen Disziplinen im Publikationsbereich längst etabliert sind und für sehr reibungslose Abläufe sorgen, stellen nur bedingt eine Lösungsmöglichkeit dar, da sie bestimmte fachliche Anforderungen und individuelle Erwartungen im Bereich der Geisteswissenschaft bzw. der Archäologie und Altertumswissenschaften nur mit Abstrichen erfüllen.

59 Anzuregen ist auch, die Textverarbeitung als Problem an sich mehr in den Fokus zu rücken, da der informationstechnologisch bezogene Umgang mit Texten in den Geisteswissenschaften zwar durchaus ein sehr zentrales Thema darstellt, allerdings oft andere Schwerpunkte setzt, wodurch gewisse Lücke bleiben<sup>38</sup>.

60 Zudem ist zu konstatieren, dass bestimmte Aufwände an einer anderen Stelle der Prozesskette (relativ am Anfang) und zu einem anderen Zeitpunkt (relativ früh) auftreten, als dies bei der Vorbereitung herkömmlicher Publikationen der Fall war. Durch diese Vorverlagerung von Aufgaben und Aufwänden müssen mittel- wie langfristig die Abläufe in allen beteiligten Arbeitseinheiten sowie deren Rollen neu gedacht werden. Das bedeutet vor allem, dass Kapazitäten und Ressourcen anders alloziert werden müssen.

### Verschränkung mit Forschungsdatenpublikationen

61 Die Komponenten, die den entscheidenden Mehrwert des neuen Zeitschriftenmodells (jenseits der Maschinenlesbarkeit) bedingen, liegen genau genommen außerhalb des hier geschilderten Konzeptes, nämlich in der vernetzten Informationsinfrastruktur des DAI (iDAI.world). Die in den entsprechenden Fachsystemen und -diensten vorhandenen oder zu veröffentlichenden Datenbestände erfahren durch das neue Zeitschriftenmodell aufgrund ihrer inhaltlichen und technischen Verschränkung mit den klassischen Zeitschriften-Artikeln lediglich eine prominente Sichtbarkeit. Insofern ist das Modell zwar in Teilen übertragbar,

---

38 Schwerpunkte z. B. Textmining/Textanalyse, Anwendungen im editionswissenschaftlichen Kontext usw. Einen repräsentativen Eindruck kann der Blick in die Tagungsakten der CAA-Konferenzen der letzten Jahre (<https://caa-international.org> [02.04.2022]) oder auf die Seiten relevanter Informationsaggregatoren (<https://archdigi.hypotheses.org> [02.04.2022]) geben. Beispiel für den editionswissenschaftlichen Kontext: Funk 2018. Vgl. allgemeiner und hier lediglich zur Illustration herausgegriffen die Inhaltsverzeichnisse der Zeitschrift für digitale Geisteswissenschaften (<https://zfdg.de/alle-artikel> [02.04.2022]).



als auch Forschungsdatenbestände eines beliebigen Repositoriums referenziert und angezeigt werden können, sofern dieses die in der Wissenschaft ohnehin notwendigen infrastrukturellen Voraussetzungen (Langzeitverfügbarkeit, nachhaltige Referenzierbarkeit, Zugreifbarkeit über Schnittstellen usw.) erfüllt. Allerdings entfallen aber dann auch jene entscheidende Mehrwerte, wie sie die vernetzte Informationsinfrastruktur des DAI bietet.

62 Der teilweise recht hohe Aufwand, der in manchen Fällen für die Aufbereitung und den Import der zu publizierenden Forschungsdaten in spezifischen Fachsystemen mit semantik-orientierten Datenmodellen anfallen kann, ist zweifellos ein Problem. Allerdings kann dies mittelfristig überwunden werden, sofern weitere Tools zur semi-automatisierten Vorbereitung entwickelt werden und darüber hinaus eine Institution auch bereit ist, alle zur Verfügung stehenden Optimierungspotenziale im Bereich der bisherigen redaktionellen Bearbeitung und der Publikationsvorbereitung zu nutzen, um die auf diese Weise freiwerdenden Kapazitäten und Ressourcen in den Bereich der Forschungsdatenaufbereitung umschichten zu können. Beispiele seien hierfür eine streng prozesslogische Segmentierung der Bearbeitungsschritte, die konsequente Anwendung standardisierter Vorgaben und Prozeduren und damit der Verzicht auf alle Arten von Sonderlösungen, auch im Hinblick auf die medialen Eigenschaften (sofern diese nicht zwingend wissenschaftlich notwendig sind), bis hin zum konsequenten Nutzen von Open Source-Anwendungen in möglichst allen Stadien der Vorbereitung.

63 Auch wenn noch nicht abzusehen ist, ob diese Maßnahmen in der Summe ausreichend sein werden, um die Zusatzaufwände so zu reduzieren, dass in der Gesamtbetrachtung keine Aufwüchse erforderlich sein werden, so werden sie doch dafür sorgen können, die Prozesse beherrschbar zu machen, zumal davon auszugehen ist, dass Anforderungen zum Umgang mit Forschungsdaten, wie sie Forschungsförderer oder der Wissenschaftsrat formulieren, mittelfristig dazu führen, dass die zu publizierenden Forschungsdaten ohnehin in einem Zustand vorgehalten werden, der aufwendige Re-Modellierungen überflüssig macht.

64 Im Hinblick auf eine mögliche Übertragbarkeit der Mechanismen in Bezug auf die Forschungsdatenpublikationen stößt das vom DAI entwickelte Modell insofern an seine Grenzen, als die Werkzeuge, die es zur semi-automatisierten Aufbereitung bzw. zum Import entwickelt, sehr spezifisch auf die Datensysteme der iDAI.world zugeschnitten sind und daher für andere Institutionen nur sehr bedingt nachnutzbar sein werden.

## Ausblick

65 Zusammenfassend ist festzuhalten, dass ein dauerhaft tragfähiger Betrieb nur dann zu gewährleisten ist, wenn eine Institution bereit ist, die oben geschilderten Maßnahmen zur Aufwandsreduzierung sowie zur Umschichtung von Kapazitäten konsequent umzusetzen.

66 Wichtiger sind aber die Veränderungen, die die Community herbeiführen muss: Zum einen müssen zeitnah noch jene Standards vereinbart und jene spezifischen Werkzeuge entwickelt werden, die für eine effiziente semi-automatisierte Vorbereitung und Publikation der Dokumente in der Prozesskette noch fehlen. Zum anderen müssen vor allem die archäologisch arbeitenden Disziplinen beginnen, bei der Vorbereitung ihrer Ergebnispublikationen einen formatkritischeren Blick auf die Inhalte zu werfen, damit vorrangig die Eigenschaften der zu veröffentlichenden Inhalte die Wahl der Medien und Anwendungen für die Publikation bestimmen. Nur unter

dieser Voraussetzung wird es mittelfristig möglich sein, Komponenten, die andernorts bereits erfolgreich im Einsatz sind bzw. die sich im Stadium des Entstehens befinden, in die eigene Produktionsstrecke zu integrieren.

<sup>67</sup> Die vom DAI entwickelten und zur Nachnutzung zur Verfügung gestellten Komponenten können zwar bis zu einem gewissen Grad einen Beitrag dazu leisten, dass auch andere Institutionen vergleichbare Modelle einführen. Entscheidend wird jedoch sein, dass sich die Fachcommunity in Zukunft noch intensiver über Standards im Hinblick auf Digitalisierungsprozesse im Publikationsbereich austauscht, damit das Repertoire an Best Practice-Lösungen, das auf die Bedarfe anderer Wissenschaftsbereiche zugeschnitten ist, für die spezifischen Anforderungen der archäologisch arbeitenden Disziplinen und deren Nachbarwissenschaften ergänzend erweitert werden kann.

## Referenzen

- Bethesda Statement 2003** Bethesda Statement on Open Access Publishing (2003), <http://legacy.earlham.edu/~peters/fos/bethesda.htm> (01.04.2022)
- Bilder – Lin – Neylon 2020** G. Bilder – J. Lin – C. Neylon, The Principles of Open Scholarly Infrastructure (2022), <https://doi.org/10.24343/C34W2H>
- Boecker et al. 2021** V. Boecker – T. Burmeister – S. Moshfegh Nia – J. Watson, iDAI.werkstatt@dainst.de. Reflexionen aus der ersten Arbeitswerkstatt zur Rekontextualisierung und Aufbereitung archäologischer Forschung mit iDAI.world-Systemen 2021, FdAI 2021/1, § 1–37, <https://doi.org/10.34780/h9z4-61v1>
- Böhm et al. 2020** D. Böhm – A. Grossmann – M. Reiche – A. Schrader, Open-Access-Publikationsworkflow für akademische Bücher. Ein Handbuch für Hochschulen und Universitäten (Leipzig 2020), <http://doi.org/10.33968/9783966270175-00>
- Bosman et al. 2021** J. Bosman – H. de Jonge – B. Kramer – J. Sondervan, Advancing Open Access in the Netherlands After 2020: From Quantity to Quality, Insights 34 (1) 16 (2021), <https://doi.org/10.1629/uksg.545>
- Breuer – Trilcke 2021** C. Breuer – P. Trilcke, Die Ausweitung der Wissenschaftspraxis des Publizierens unter den Bedingungen des digitalen Wandels. Herausgegeben von der Arbeitsgruppe »Wissenschaftspraxis« im Rahmen der Schwerpunktinitiative »Digitale Information« der Allianz der deutschen Wissenschaftsorganisationen (2021), <https://doi.org/10.48440/allianz0a.041>
- Cristiani Sienra 2016** A. Cristiani Sienra, Journal Article Tag Suite (JATS). Situación actual y análisis de implantación del estándar promovido por NISO para revistas científicas basado en XML (Madrid 2016), <http://hdl.handle.net/10016/24418> (09.04.2022)
- Chronopoulos – Maier – Novokhatko 2020** St. Chronopoulos – F. K. Maier – A. Novokhatko (Hrsg.), Digitale Altertumswissenschaften. Thesen und Debatten zu Methoden und Anwendungen (Heidelberg 2020), <https://doi.org/10.11588/propylaeum.563>
- Digitales Publizieren in den Geisteswissenschaften 2021** Digitales Publizieren in den Geisteswissenschaften. Begriffe, Standards, Empfehlungen, in: AG Digitales Publizieren (Hrsg.), Zeitschrift für digitale Geisteswissenschaften / Working Papers 1 (Wolfenbüttel 2021), [https://doi.org/10.17175/wp\\_2021\\_001](https://doi.org/10.17175/wp_2021_001)
- Dreher – Lang – Oberländer 2020** L. Dreher – I. Lang – A. Oberländer, Bereitschaft zu Open Access in den Geisteswissenschaften: Ergebnisse einer Umfrage unter geisteswissenschaftlichen Herausgeber\*innen im Rahmen des Projekts OLH-DE, o-bib. Das offene Bibliotheksjournal 7, 2020, 1–20, <https://dx.doi.org/10.5282/o-bib/5512>
- Effinger – Krabbes – Withanage 2018** M. Effinger – F. Krabbes – D. Withanage, Crossmediales Publizieren bei Heidelberg University Publishing (heiUP), in: BIT online 21, 5 (Wiesbaden 2018) 393–403, <https://www.b-i-t-online.de/heft/2018-05-fachbeitrag-effinger.pdf> (02.04.2022)
- Funk 2018** St. E. Funk, Elektronisches Publizieren von Digitalen Forschungsdaten am Beispiel des TextGrid Repositorys – Umsetzung von Digitalen Publikationsworkflows für die eHumanities (2018), <http://dx.doi.org/10.20375/0000-000B-D269-2>
- Garnett et al. 2017** A. Garnett – M. Aufreiter – O. Buchtala – J. P. Alperin, Introducing Texture: An Open Source WYSIWYG Javascript Editor for JATS, Journal Article Tag Suite Conference (JATS-Con) Proceedings 2017 [Internet]. Bethesda (MD): National Center for Biotechnology Information (US) (2017), <https://www.ncbi.nlm.nih.gov/books/NBK425544/>
- Hacker 2017** A. Hacker, Software für den Publikationsworkflow und den Peer-Review Prozess, in: K. Söllner – B. Mittermaier (Hrsg.), Praxishandbuch Open Access (Berlin 2017), 283–290, <https://doi.org/10.1515/9783110494068-033>
- Hacker – Corrao 2017** A. Hacker – E. Corrao, Laying Tracks as the Train Approaches: Innovative Open Access Book Publishing at Heidelberg University from the Editors' Point of View, Journal of Scholarly Publishing 48, 2017, <https://doi.org/10.3138/jsp.48.2.76>
- Hintze 2020** M. Hintze, Die digitale Online-Publikation in den Geisteswissenschaften. Ein ungenutztes Potential?, in: Chronopoulos – Maier – Novokhatko 2020, 191–206
- Kuhn – Hagenhoff 2019** A. Kuhn – S. Hagenhoff, Nicht geeignet oder nur unzureichend gestaltet? Digitale Monographien in den Geisteswissenschaften, Zeitschrift für digitale Geisteswissenschaften (Wolfenbüttel 2019), [https://doi.org/10.17175/2019\\_002](https://doi.org/10.17175/2019_002)
- Pampel 2021** H. Pampel, Strategische und operative Handlungsoptionen für wissenschaftliche Einrichtungen zur Gestaltung der Open-Access-Transformation (Berlin 2021), <https://doi.org/10.18452/22946>
- Piez 2017** W. Piez, HTML First? Testing an alternative approach to producing JATS from arbitrary (unconstrained or »wild«) .docx (WordML) format, in: Journal Article Tag Suite Conference (JATS-Con) Proceedings 2017 [Internet]. Bethesda (MD): National Center for Biotechnology Information (US) (2017), <https://www.ncbi.nlm.nih.gov/books/NBK425546/>
- Pieper et al. 2018** D. Pieper – S. Fund – K. Werner – A. Jobmann, Qualitätsstandards für den Einstieg in die Open-Access-Stellung von Büchern<sup>2</sup> (Bielefeld 2018), <http://nbn-resolving.de/urn:nbn:de:0070-pub-29321893>

**Projekt AuROA 2022** Publizieren und Open Access in den Geisteswissenschaften. Erkenntnisse aus dem Projekt AuROA zu den Stakeholdern im Publikationsprozess (Essen 2022), <https://projekt-auroa.de/wp-content/uploads/2022/03/AuROA-Publizieren-und-Open-Access-in-den-Geisteswissenschaften.pdf> (05.04.2022)

**UNESCO 2021** First draft of the UNESCO Recommendation on Open Science, <https://unesdoc.unesco.org/ark:/48223/pf0000374837> (02.04.2022)

**Wilkinson et al. 2016** M. Wilkinson – M. Dumontier – I. Aalbersberg et al., The FAIR Guiding Principles for scientific data management and stewardship, *Sci Data* 3, 160018 (2016), <https://doi.org/10.1038/sdata.2016.18>

**Wissenschaftsrat 2022** Wissenschaftsrat (Hrsg.), Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access (Köln 2022), <https://doi.org/10.57674/fyrc-vb61>



---

## ZUSAMMENFASSUNG

**Das neue maschinenlesbare  
Zeitschriftenmodell des Deutschen  
Archäologischen Instituts**  
Ein Werkstattbericht  
Peter Baumeister

Auch bei archäologisch arbeitenden Disziplinen spielt nicht nur die Publikation der Forschungsdaten, die im Rahmen eines Projektes entstehen, eine zentrale Rolle, es gewinnen ebenso Interoperabilität und maschinelle Lesbarkeit der zu veröffentlichen Inhalte zunehmend an Bedeutung, um die Ergebnisse ihrer Forschungsarbeit in bestmöglicher Form rezipier- und nachnutzbar zu machen. Der vorliegende Artikel will in Form eines Werkstattberichts zeigen, wie das Deutsche Archäologische Institut im Rahmen seiner Digitalisierungsstrategie ein neues Zeitschriftenmodell entwickelt, das dabei den Anspruch hat, Artikelinhalte in maschinenlesbarer Form zu publizieren und darüber hinaus dazugehörigen Forschungsdatenkonvoluten durch die inhaltliche und technische Verschränkung eine prominente Sichtbarkeit zu bieten.

## SCHLAGWÖRTER

Digitale Transformation, Maschinenlesbarkeit, Forschungsdatenmanagement, Textverarbeitung, XML

---

## AUTHOR

Peter Baumeister  
Deutsches Archäologisches Institut, Zentrale  
Wissenschaftliche Dienste  
Podbielskiallee 69–71  
14195 Berlin  
Deutschland  
peter.baumeister@dainst.de  
ORCID-ID: <https://orcid.org/0000-0001-5430-1456>  
ROR: <https://ror.org/041qv0h25>

---

## METADATA

Titel/*Title*: Das neue maschinenlesbare  
Zeitschriftenmodell des Deutschen  
Archäologischen Instituts. Ein Werkstattbericht  
Band/*Issue*: FdAI 2022/1

Cover Illustration: Screenshot: Peter Baumeister

Bitte zitieren Sie diesen Beitrag folgenderweise/  
*Please cite the article as follows*:  
P. Baumeister, Das neue maschinenlesbare

Zeitschriftenmodell des Deutschen  
Archäologischen Instituts. Ein Werkstattbericht,  
FdAI 2022/1, § 1–67, <https://doi.org/10.34780/cf6e-prcf>

Copyright: CC-BY-NC-ND 4.0

Online veröffentlicht am/*Online published on*:  
19.12.2022  
DOI: <https://doi.org/10.34780/cf6e-prcf>

Schlagworte/*keywords*: Digitale  
Transformation/*digital transformation*,  
Maschinenlesbarkeit/*machine readability*, For-  
schungsdatenmanagement/*research data*  
*management*, Textverarbeitung/*text processing*,  
XML  
Bibliographic reference: <https://zenon.dainst.org/Record/003024705>

---

## JOURNAL METADATA

Forum for Digital Archaeology and Infrastructure  
published since 2021

Publisher/Editors  
Benjamin Ducke, Friederike Fless, Fabian Riebschläger, Henriette Senst  
Deutsches Archäologisches Institut  
Podbielskiallee 69–71  
14195 Berlin  
Deutschland  
<http://www.dainst.org>

Editing and Typesetting  
Publishing editor: Deutsches Archäologisches  
Institut, Zentrale – Stabsstelle Kommunikation,  
Redaktion  
Editing: Antonie Brenne, Janina Rücker M.A. ([fdai-journal@dainst.de](mailto:fdai-journal@dainst.de))

Corporate Design: LMK Büro für Kommuni-  
kationsdesign, Berlin  
Webdesign: LMK Büro für Kommunikationsdesign,  
Berlin ([lm-kommunikation.de](http://lm-kommunikation.de))  
Programming Viewer: LEAN BAKERY, München  
([leanbakery.com](http://leanbakery.com))