

■ ÜBERLEGUNGEN ZUR NUTZUNG VON PHAIDRA ALS REPOSITORY FÜR DIGITALE ARCHÄOLOGISCHE DATEN

von Dominik Hagmann

Abstract: Im Allgemeinen sind archäologische Daten durch eine Vielzahl an Datentypen und Dateiformaten geprägt, die unterschiedliche Inhalte speichern. Hier spannt sich der Bogen etwa von Texten und tabellarisch erfassten quantitativen Daten über Objektfotos bis hin zu Vektorgrafiken sowie den immer beliebteren 3D-Modellen. Diese ausgeprägte Diversität ist der Anwendung verschiedener Methoden und der Umsetzung vielfältiger Ansätze in der archäologischen Forschung geschuldet. So generieren archäologische Surveys (Oberflächenfundaufsammlungen) andere Daten als Ausgrabungen, luftbildarchäologische Maßnahmen oder geophysikalische Prospektionen. Je nach Projekt kommen zudem noch zahlreiche nachträglich digitalisierte, d.h. ursprünglich analog angefertigte Ressourcen hinzu, die ihrerseits abermals sehr heterogene Informationen abbilden können. Die Herausforderung in der Langzeitarchivierung dieser inhomogenen Datensätze liegt unter anderem im Vorhandensein eines Repositoriums, das die verschiedenen Datentypen effektiv aufnehmen und zudem innerhalb des Archivs sinnvoll in Beziehung zueinander setzen kann. Diese Voraussetzungen sind im System PHAIDRA der Universität Wien gegeben, weswegen es dadurch möglich ist, komplexe archäologische Datensätze dort nachhaltig verfügbar zu machen.

Schlüsselwörter: Digitale Geisteswissenschaften; Digitale Archäologie; Langzeitarchivierung; Open Access; Open Science

CONSIDERATIONS ON THE USE OF PHAIDRA AS A REPOSITORY FOR DIGITAL ARCHAEOLOGICAL DATA

Abstract: Archaeological data are characterized by a variety of data types and file formats that store different kinds of information respectively contents. This spectrum ranges from texts and tables to object photographs, vector graphics and the increasingly popular 3D models for example. This diversity of archaeological data is due to the use of different methods and various approaches in archaeology. Depending on the individual project, numerous digitized and heterogeneous datasets are added as well. Among others, one of the main challenges in long-term data archiving and data dissemination in archaeology is to handle this inhomogeneous information within a repository. In this article it will be shown that the system PHAIDRA at the University of Vienna enables researchers to make complex archaeological datasets available in a sustainable way.

Keywords: *Digital Humanities; Digital Archaeology; Long term archiving; Open Access; Open Science*

DOI: <http://doi.org/10.31263/voebm.v71i1.1974>



Dieses Werk ist lizenziert unter einer
Creative-Commons-Lizenz Namensnennung 4.0 International

Inhalt

1. *Einleitung*
2. *Herausforderung digitaler archäologischer Daten*
3. *Datenarchive in der Archäologie*
4. *Lösungsansatz: Fallbeispiel PHAIDRA und Molino San Vincenzo*
5. *Diskussion und Conclusio*

1. Einleitung

Archäologische Feldforschungen lassen sich grob in zwei Arten gliedern: Zum einen treten sie als mikro- und makro-invasive Eingriffe auf, beispielsweise als stratigrafische Ausgrabungen und Bohrungen. Zum anderen werden sie als non-invasive Untersuchungen durchgeführt, z.B. in Form von intensiven und extensiven Fundaufsammlungen (sog. Surveys) oder als geophysikalische Prospektionen. Weiters sind hier auch verschiedenen Methoden der Fernerkundung, etwa flugzeuggestütztes Laserscanning oder Satellitenfotografie, zu erwähnen.¹ Die Wahl der entsprechenden Methode hängt von der jeweiligen Fragestellung ab: Sollen verhältnismäßig schnell relativ große Flächen, etwa über mehrere km² hinweg, auf Prä- und Absenz archäologischer Funde überprüft werden, bieten sich extensive Surveys oder die Auswertung von Luftbildern an. Soll hingegen ein fest definierter Fundplatz hinsichtlich seiner konkreten Funktion und Chronologie untersucht werden, ist eine Ausgrabung in Erwägung zu ziehen. Wie die Methode umgesetzt wird, ist danach unter anderem von historisch bedingten Forschungstraditionen und nicht zuletzt von den jeweiligen Forscher*innen sowie vorhandenen finanziellen und technischen Mitteln beeinflusst.

Unabhängig von der spezifischen archäologischen Fachausrichtung (z.B. Ur- und Frühgeschichte, Klassische Archäologie, Vorderasiatische Archäologie etc.) wird dabei die materielle Hinterlassenschaft vergangener Kulturen durch die Verwendung unterschiedlicher Verfahren im kleinen und

großen Maßstab in ihren qualitativen und quantitativen Eigenschaften umfassend aufgenommen, verarbeitet und analysiert. Hierbei werden derzeit neben analogen Daten (vor allem Fund- und Befundzeichnungen sowie diverse Formulare zur deskriptiven Dokumentation) vermehrt umfangreiche Bestände an Informationen in digitaler Form erzeugt. Zudem werden ursprünglich analog angefertigte Daten gegebenenfalls ebenso in einem weiteren Schritt digitalisiert und dann gemeinsam mit originär digital erzeugten Daten („*born digital*“) weiterverarbeitet. Aus diesem multiplen Vorgehen im Bereich der Feldforschungen resultieren vielfältige Inhalte, sowohl in Bezug auf ihr archäologisches Informationsgehalt als auch in technischer Hinsicht.

Der nachfolgende Artikel² setzt sich im angesprochenen Umfeld aus diesem Grund mit der Herausforderung der nachhaltigen und effektiven Archivierung archäologischer Daten auseinander und zeigt anhand des Fallbeispiels der archäologischen Ausgrabung „Molino San Vincenzo“ Möglichkeiten der praktischen Umsetzung eines solchen Vorhabens auf.

2. Herausforderung digitaler archäologischer Daten

Archäologische Daten sind in hohem Maße heterogen. Die Diversität der Daten resultiert dabei neben unterschiedlichen Methoden auch aus dem Nichtvorhandensein von einheitlichen Dokumentationsstandards und Terminologien. Diese ist in variierendem Ausmaß etwa durch verschiedene lokale bzw. nationale Gesetzesrichtlinien oder die individuelle Umsetzung der Methoden in der Arbeitspraxis bedingt. Die Anwendung verschiedenster Informations- und Kommunikationstechnologien (IKT) in der Archäologie – ein Zustand, der mit Begriffen wie „Digitale Archäologie“ oder „Archäoinformatik“ beschrieben werden kann und der die Voraussetzung für die Generierung digitaler Daten in der Archäologie ist – erfolgt ebenso in vielförmiger Ausprägung.³ Im Bereich der Datenerhebung kann besonders die digitale Ausgrabungsdokumentation als Beispiel hierfür angeführt werden. Diese erfolgt etwa unter der Verwendung leistungsfähiger digitaler Vermessungsinstrumente (Totalstationen oder terrestrische Laserscanner), aber auch durch die Anwendung elaborierter photogrammetrischer Methoden (*Image Based Modeling*, IBM), die die Generierung von 3D-Oberflächenmodellen ermöglichen.⁴ Diese Modelle können (je nach intendiertem Zweck) sowohl als Punktwolke als auch in der Form von 3D-Geometrien oder weiterverarbeitet als Rasterdatensätze vorliegen. Die Organisation und Analyse solcher Daten erfolgt innerhalb der deutschsprachigen Forschung im-

mer häufiger durch die Verwendung Geographischer Informationssysteme (GIS)⁵, wobei diese Entwicklung bereits in den 1990er-Jahren besonders im anglo-amerikanischen Raum einen nennenswerten Beginn genommen hat.⁶ Neben der Schöpfung verschiedener raumbezogener Daten im Zuge von Feldforschungen gewinnen zudem auch weitere digitale Bereiche der archäologischen Weiterverarbeitung und Verbreitung der Daten zusehends an Bedeutung. Exemplarisch hierfür seien etwa die Anwendung von Ontologien wie CIDOC CRM und die Entwicklung graphenbasierter (Online-) Datenbanken, die Arbeit mit *computer vision* und *machine learning* oder generell das weite Feld von *linked open data* (LOD) genannt.⁷

Insgesamt werden bei archäologischen Untersuchungen also regelmäßig verschiedene Datentypen generiert, die in unterschiedlichen Dateiformaten gespeichert werden und abweichend aufgebaut und benannt sein können. Unter anderem aus diesem Grund existiert eine verwirrende Vielfalt offener und kommerzieller Archäologischer Informationssysteme (AIS), die on- und offline betrieben werden können, ebenso wie ausschließlich „interne“ Lösungen, die ebenfalls teilweise auf offener und teilweise auf kommerzieller Software basieren. Zu nennen sei hier exemplarisch etwa das kommerzielle, webbasierte Service Inari AIS. Im Bereich der freien und offenen Software (FOSS) kann als Beispiel OpenAtlas angeführt werden, ein archäologisches Datenbanksystem, das zum Teil an der Universität Wien entwickelt wurde. Im Bereich der „internen“ Datenverwaltung sind besonders LibreOffice Base, Microsoft Access (diesem kann insgesamt der größte Marktanteil eingeräumt werden) oder FileMaker verbreitet, die aber – als Lösungen für archäologische Datenerfassung und -verwaltung – teilweise auch kommerziell verhandelt werden.⁸ Alle haben aber in unterschiedlicher Art und Weise zum Ziel, die vielfältigen archäologischen Daten nachvollziehbar zu organisieren und (teilweise) auch weiterzuverarbeiten.⁹

Zusätzlich zur erwähnten Vielfalt stammen digitale Daten nicht nur aus jüngster Zeit: Aufgrund der Eigenart der archäologischen Forschung, verschiedenste digitale Tools und Methoden unmittelbar zu adaptieren, ist zudem in Österreichs seit den 2000er-Jahren mit entsprechenden Daten in umfangreicherem Ausmaß zu rechnen, wobei dieser Trend bereits in den 1980er-Jahren in Anfängen vorhanden war und bereits heute mit Problemen verbunden ist. Als Beispiel können etwa elektronisch verarbeitete Texte genannt werden, die auf *de facto* nicht mehr verwendeten Datenträgern (z.B. 8- und 3,5-Zoll Floppy Disks) in veralteten Formaten (z.B. Revisable-Form Text, RFT) gespeichert wurden und nur noch mit erhöhtem Aufwand gelesen werden können. Es ist also davon auszugehen, dass

nicht nur eine ausgeprägte Diversität archäologischer Daten vorhanden, sondern auch ein „*digital dark age*“¹⁰ (im Sinne eines effektiven Verlustes an Wissen aufgrund der hard- und/oder softwaretechnisch bedingten Unlesbarkeit der Daten) bereits angebrochen ist. Es muss an dieser Stelle aber ebenso angemerkt werden, dass vermutlich deutlich mehr vielfältige und aktuelle Daten vorliegen, als Daten auf veralteten Datenträgern in obsoleten Dateiformaten. Darum ist es als besondere Herausforderung anzusehen, archäologische Daten durch effektive Langzeitarchivierung nachhaltig verfügbar zu machen und so ein dunkles digitales Zeitalter nicht anbrechen zu lassen bzw. dessen Auftakt soweit als möglich einzudämmen.

3. Datenarchive in der Archäologie

Um aus diesem Grund effektive sowie nachhaltige Langzeitarchivierung betreiben zu können, ist es zum einen unerlässlich, die Speicherung aller relevanten Daten in für die Archivierung geeigneten Formaten vorzunehmen. Zum anderen muss ein passendes Archiv gewählt werden, das für die langfristige Speicherung der Daten zur Verfügung steht: Im Bereich der Archäologie bieten sich in Österreich besonders institutionelle Repositorien an, aber auch international betriebenen Langzeitarchivierungsplattformen.¹¹ Hierbei gibt es sowohl dezidiert auf archäologische Forschungen fokussierte Projekte¹² als auch fächerübergreifendere Lösungen¹³. Neben der Archivierung dienen Repositorien auch für das digitale Verfügbarmachen von Forschungsdaten und anderen forschungsrelevanten Outputs. Dieses Vorgehen verfügt über zahlreiche Vorteile, so Lizenzmodelle gewählt werden, die eine freie Verbreitung der Daten im Open Access gestatten.¹⁴ Dadurch ist es etwa von Seiten der Rezipienten möglich, die publizierten Forschungsergebnisse anhand solcher Daten zu evaluieren und weiterzuentwickeln. Den Publizierenden ermöglicht ein derartiges Vorgehen, Daten, die etwa aufgrund einer begrenzten Projektlaufzeit nicht mehr weiterbearbeitet werden konnten, so trotzdem zugänglich zu machen. Es ist darum von hoher Bedeutung, Langzeitarchivierung einerseits als bewahrendes Instrument zu betrachten, andererseits jedoch gleichzeitig auch als zusätzliches Publikationsorgan.¹⁵ Im Vergleich zu Printpublikationen, die in der Archäologie immer noch einen sehr hohen Stellenwert genießen, sind bei korrekt langzeitarchivierten und frei verfügbaren Daten zahlreiche Vorteile gegeben, nämlich deutlich umfangreichere Weiterverarbeitungsmöglichkeiten und eine erhöhte Reichweite der Inhalte sowie Resonanz darauf.¹⁶ Auch leiden „*digital born data*“ durch deren Abbildung auf ein (im Vergleich zur

bloßen Datei) in der Regel ökologisch und ökonomisch aufwändiges analoges Medium (z.B. aufgrund der Satz- und Druckkosten oder nicht nachhaltiger Waldbewirtschaftung hinsichtlich der Papierproduktion) nicht an den damit verbundenen Informationseinbußen oder Qualitätsverlusten, etwa aufgrund von Einschränkungen des gewählten Druckverfahrens (z.B. Graustufen).¹⁷

4. Lösungsansatz: Fallbeispiel PHAIDRA und Molino San Vincenzo

Eines der breit angelegten Systeme, das über archäologische Daten hinaus als Archiv allen Fächern und ihren Daten zur Verfügung steht, ist das Repositorium „PHAIDRA“ der Universität Wien.¹⁸ Als universitäres Repositorium bietet es die Möglichkeit, digitale Daten wissenschaftlicher Forschungen langfristig in einer kontrollierten Umgebung unter der Verwendung nachhaltiger Standards zu speichern.¹⁹ Aufgrund dieser breiten Auslegung ist PHAIDRA bestens für archäologische Daten geeignet.

PHAIDRA²⁰ basiert auf dem Fedora Commons Repository²¹ und wird als Langzeitarchivierungsservice an verschiedenen Universitäten in mehreren Ländern Europas verwendet²². Das Repositorium, das im Registry of Research Data Repositories gelistet ist, entspricht dem Open Archival Information System (OAIS) Referenzmodell.²³ Alle Datensätze verfügen über einen permanenten Link und strukturierte Metadaten. Diese werden anhand verschiedener Schemata (modifiziertes Metadata Object Description Schema, Dublin Core und Learning Object Metadata) beschrieben. Die Datensätze werden zudem für Suchmaschinen indiziert. Das System bietet die Möglichkeit, die darin archivierten Ressourcen untereinander zu verlinken und weiters in Sammlungen (so genannte Collections) zusammenzufassen, was umfangreiche Organisationsstrukturen ermöglicht. Auch ist eine Versionierung der Datensätze verfügbar.²⁴ Zur Beschreibung der Objekte kann auf kontrolliertes, standardisiertes Vokabular aus verschiedenen Thesauri zugegriffen werden.²⁵

Als Beispiel für die Archivierung archäologischer Daten im Repositorium PHAIDRA sei der Fundplatz Molino San Vincenzo in der italienischen Toskana angeführt, auf dem durch die Universität Wien seit 2012 jährlich verschiedene archäologische Untersuchungen stattfinden, darunter Ausgrabungen, geophysikalische Prospektionen und Oberflächenfundaufsammlungen. Die archäologische *site* ist durch Strukturen charakterisiert, die schwer durch rezente Aktivitäten im Zuge der intensiven Landwirtschaft in Mitleidenschaft gezogen sind. Anhand der angetroffenen materiellen

Kultur (neben diversen Funden u. a. bauliche Befunde in Form von Mauerfundamenten) kann von mehrphasigen Siedlungsaktivitäten gesprochen werden, die sich wohl etwa vom 6. Jh. v. Chr. bis in das 6. Jh. n. Chr. erstrecken; der Schwerpunkt dieser Aktivitäten kann für die römische Epoche angenommen werden. Im Zuge der archäologischen Untersuchungen auf dem Fundplatz wurden verschiedene Forschungs- und Dokumentationsmethoden angewandt, die unterschiedliche Datensätze schufen. Bei diesen handelt es sich etwa um digitale Dokumentationsfotos oder um mittels *Computer-Aided Design* (CAD) generierte 3D-Geometrien, die als geschlossene Polygone die Grenze der jeweiligen Stratifikationseinheiten (i.e. Oberflächen/Interfaces [IF] und Volumina/stratigrafische Einheiten [SE], „Schichten“) beschreiben. Die Topografie der Stratifikationseinheiten wird durch hochauflösende digitale Oberflächenmodelle detailliert beschrieben, die durch IBM generiert werden. Analog angefertigte und später digitalisierte, textbasierte, deskriptive Daten zur Befundbeschreibung ergänzen diese Informationen ebenso wie Videodaten zur Befundaufnahme und der Dokumentation des Arbeitsprozesses. Alle Informationen werden in einem GIS verarbeitet und analysiert sowie auf Servern der Universität Wien gespeichert.²⁶

Hinsichtlich des archäologischen Datenbestandes von ca. 1,6 TB²⁷ wird angestrebt, eingedenk des Leitspruchs „so viel wie nötig, so wenig wie möglich“ alle relevanten archäologischen Daten zu archivieren. In diesem Kontext muss nun der Frage nachgegangen werden, welche Daten als relevant angesehen werden können und ob es tatsächlich notwendig ist – besonders unter der Berücksichtigung ökonomischer Faktoren –, auch die Rohdaten selbst zu archivieren oder ob es genügt, sich lediglich auf die entsprechenden Endprodukte zu fokussieren? Direkt (und im wahrsten Sinne des Wortes) *im Feld* erhobene Daten als tatsächliche Rohdaten dienen gegebenenfalls nur als Grundlage für weiterführende Verarbeitungen, die dann erst in einem nachfolgenden Schritt das gewünschte Produkt generieren – ein Umstand, der etwa im Bereich des IBM von Bedeutung ist.²⁸ So wurden etwa für ein 3D-Modell des gesamten Schnitts S2/2014 der Ausgrabung Molino San Vincenzo 762 Fotos mit einer Dateigröße von insgesamt 26 GB (unkomprimiertes Nikon Electronic Format, NEF; ein RAW-Dateiformat) beziehungsweise von 2,2 GB als Dateien im Format JPEG angefertigt, wobei nicht alle Fotos von gleicher Qualität waren und verwendet wurden (Abbildung 1). Die JPEG-Dateien wurden für die Generierung der Modelle verwendet. Die Fotos verfügen nur indirekt über ein eindeutiges Motiv und dienen vielmehr erst in der Masse der Erfassung eines bestimmten Befundes. Jedes einzelne Foto

gibt nur einen Ausschnitt der Ausgrabung wieder, erst als anhand qualitativer Kriterien bestimmte Selektion können die Fotos unter der Verwendung verschiedener Algorithmen im Zuge mehrerer Arbeitsschritte zur Generierung eines 3D-Modells verwendet werden. Das 3D-Modell, dessen Ziel die Anfertigung der Fotos war, verfügt bei durchschnittlicher Qualität hinsichtlich der Geometrien und ohne weitere Nachbearbeitung über eine Dateigröße von 55,8 MB beziehungsweise von 15,3 MB für die Textur.²⁹ In weitere Folge kann dieses Modell dazu verwendet werden, Orthofotos und hochauflösende digitale Oberflächenmodelle der jeweiligen Stratifikationseinheiten anzufertigen. Eine zusätzliche Archivierung der Rohdaten würde einen deutlich erhöhten Speicherplatzbedarf bedeuten, was auch einer erhöhten ökonomischen Belastung für die Langzeitarchivierung gleichkommt. Wiederum liefern die Rohdaten umfangreiche Weiterverarbeitungsmöglichkeiten, wobei es noch nicht absehbar ist, inwiefern das NEF-Format überhaupt langzeitarchivierungsfähig ist. Auch ist unklar, ob und wie IBM mittel- und langfristig in der Archäologie Verwendung finden wird und die Rohdaten noch verwendet werden können oder generell noch Interesse daran besteht, derartige Methoden anzuwenden. Diesen Überlegungen folgend ist derzeit keine dezidierte Langzeitarchivierung solcher Rohdaten vorgesehen. Derartige Fragen machen zudem die Notwendigkeit der Formulierung einer eindeutigen Strategie hinsichtlich des Forschungsdaten-*lifecycles* im Rahmen eines Datenmanagementplans offensichtlich.³⁰

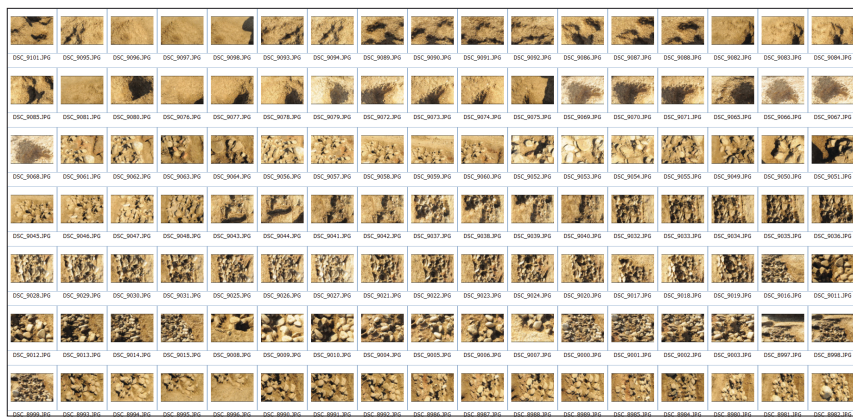


Abb. 1: Selektion der für die 3D-Modellgenerierung anhand von IBM angefertigten Fotos für ein texturiertes Oberflächenmodell des Zustandes von Schnitt 4/2015 (© Dominik Haggmann 2018)

Das Archivieren der Daten im Repository PHAIDRA folgt darum einer Archivierungs- und Nachnutzungsstrategie: Entsprechend der technischen Möglichkeiten innerhalb PHAIDRAs wird die Organisation der Daten in unterschiedliche Collections durchgeführt, die zueinander in Beziehung stehen und nach methodischen Gesichtspunkten angelegt werden. Die Grundidee ist hierbei, etwa im Bereich der Ausgrabung die während des Arbeitsprozess erfassten Daten über die Stratifikationseinheiten, die die kleinste Dokumentationseinheiten darstellen, soweit wie möglich in entsprechende Collections zusammenzufassen und miteinander zu verbinden. Solche Collections lassen sich nach Datenarten (Fotos, Geometrien etc.) ordnen, wobei hier wieder die Möglichkeit besteht, diese in Untersammlungen aufzugliedern. Alle zu publizierenden und publizierten Datensätze werden soweit möglich hinsichtlich eines freien Open Access mit einer Creative Commons Attribution 4.0 International (CC BY 4.0) Lizenz versehen. Die Speicherung der Daten in Archivformaten wird gemäß archäologischer Best Practice vorgenommen.³¹ Für Molino San Vincenzo sind alle Daten auf PHAIDRA in einer (im Aufbau befindlichen) Collection o:528478 zusammengefasst (<http://phaidra.univie.ac.at/o:528478>), die sich in weitere Untersammlungen aufgliedert: Collection o:528478 ist so auf der obersten Ebene die Sammlung aller fundplatzbezogenen Daten; dies gestattet die Publikation aller langzeitarchivierten Datensätze mittels nur eines URI, der dafür auch als DOI (<https://dx.doi.org/10.25365/phaidra.45>) zur Verfügung steht. Auf der Ebene darunter sind im Bereich der Ausgrabung die einzelnen, archäologischen Schnitt-Sammlungen angesiedelt, die nach den jährlichen Kampagnen getrennt sind. Hierfür kann als Beispiel etwa Collection o:775596 für Schnitt 2/2015 (<http://phaidra.univie.ac.at/o:775596>) genannt werden. Abermals eine Ebene darunter sind in der entsprechenden Schnittsammlung die Daten zu den jeweiligen Stratifikationseinheiten subsummiert: Großteils sind die Daten dabei wiederum in Untersammlungen angeordnet (etwa die Fotos einer Ascheschicht in der Brennkammer eines Ofens [US 12000/S2/2015]: <http://phaidra.univie.ac.at/o:512223>). Alleinstehende digitale Objekte – so z.B. o:775606, bei dem es sich um das digitale Oberflächenmodell der stratigrafischen Einheit US 13600/S2/2015 (Ofensohle) handelt (<https://phaidra.univie.ac.at/o:775606>) – werden auf gleicher Inhaltsebene keiner eigenen Gruppe zugeordnet, sondern direkt mit den korrespondierenden Sammlungen (hier die Fotosammlung zu US 13600/S2/2015: <https://phaidra.univie.ac.at/o:512226>) in Beziehung gesetzt, indem Verbindungen zwischen dem Einzelobjekt und der jeweiligen Collection hergestellt werden. Innerhalb jeder Collection werden zudem im Bedarfsfall ebensolche Verbindungen zwischen den jeweiligen Einzelobjekten angelegt (Abbildung 2).

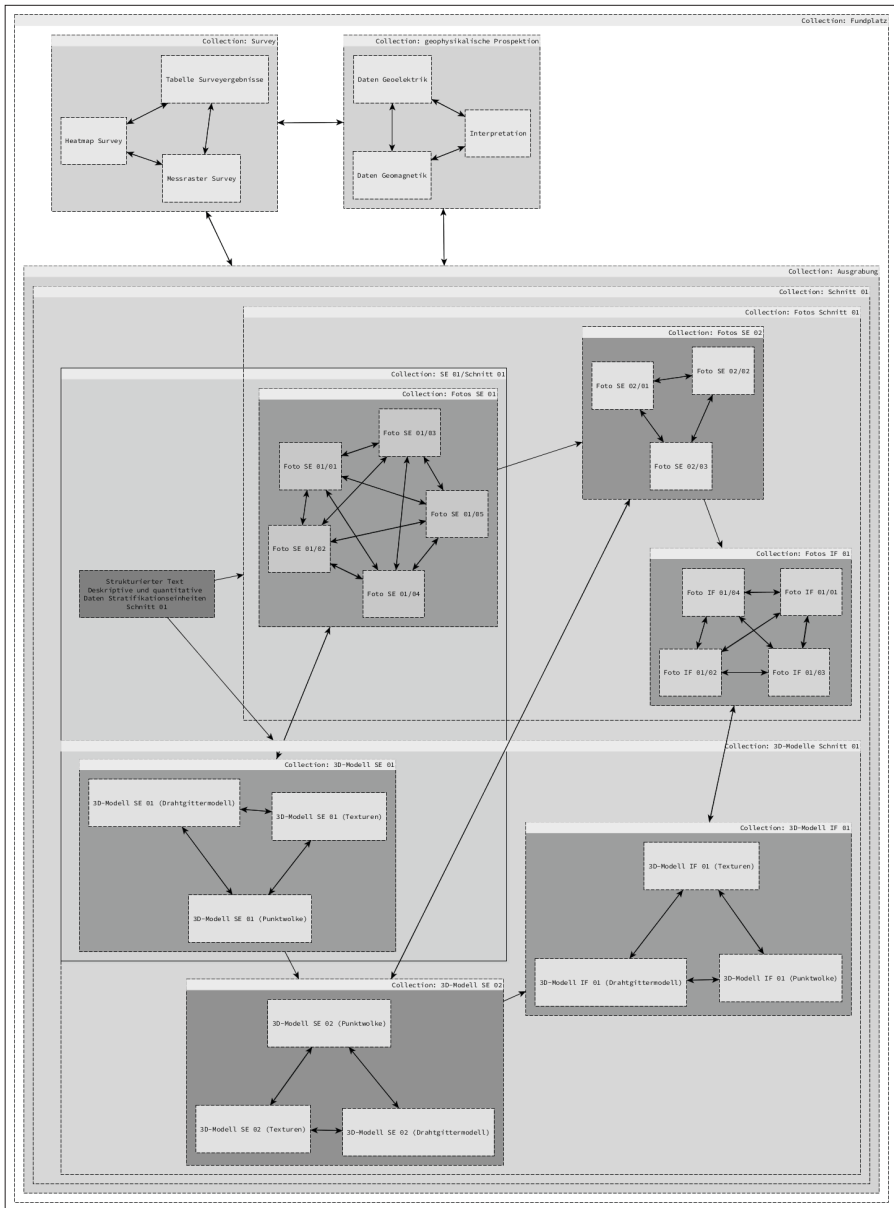


Abb. 2: Exemplarisches netzwerkartiges Modell der archäologischen Daten des Fundplatzes Molino San Vincenzo innerhalb des Systems PHAIDRA (© Dominik Hagmann 2018)

Alle Objekte verfügen über umfangreiche Metadateninformationen: So ist innerhalb PHAIDRAS z. B. für das Objekt o:512225 (<https://phaidra.univie.ac.at/o:512225>), welches zur Fotodokumentation der Ofensohle (US 13600/S2/2015) gehört, neben qualitativen Angaben zur Klassifikation aus kontrollierten Vokabularen sofort ersichtlich, dass es sich etwa um ein TIFF mit einer Dateigröße von 68,90 MB handelt, das Objekt in seinem Lebenszyklus in Version 1 vorliegt und mit einer CC BY 4.0 International Lizenz versehen ist. Neben zahlreichen weiteren Metainformationen sind Informationen zum Uploaddatum der Datei (10.02.2017 um 16:43:58 Uhr) ebenso vorhanden wie sämtliche verfügbaren EXIF Daten, wodurch problemlos etwa das Aufnahmedatum (22.08.2015, 13:41:16 Uhr) ermittelt werden kann.

Zur Veranschaulichung des Archivierungskonzeptes sei konkret auf die Collection o:512220 („Fotos: US 11300/S2/2015/01–06“) für alle Dokumentationsfotos des Verfüllungsmaterials der SE „US 11300/S2/2015“ verwiesen (<https://phaidra.univie.ac.at/o:512220>). Hierin sind sechs unterschiedliche Aufnahmen enthalten, die den etappenweisen Abtrag der Verfüllung eines Ofens mit Abbruchmaterial aus verschiedenen Perspektiven abbilden. Kongruent sind weitere Collections der Fotodokumentation der einzelnen Stratifikationseinheiten angelegt, ebenso wie einzelne digitale Objekte, die als Einzelfoto eine separate Stratifikationseinheit dokumentieren (z.B. im Fall von SE „US 9900/S2/2015“: <https://phaidra.univie.ac.at/o:512211>). In einer nächsten Hierarchieebene können die verschiedenen Foto-Collections im Rahmen einer weiteren Sammlung („Fotos S2/2015“) zusammengefasst werden, die für alle Fotos steht, die im jeweiligen Schnitt vorliegen (<https://phaidra.univie.ac.at/o:775598>). Alle Objekte, unabhängig davon, ob inner- oder außerhalb der Sammlungen, beziehungsweise die verschiedenen Collections selbst, können so gezielt mit einander verknüpft werden, um zusammenhängende Datensätze aufzuzeigen. So lassen sich etwa spezialisierte Informationen innerhalb einer Metadatenfilei zum verwendeten Koordinatensystem (<http://phaidra.univie.ac.at/o:528464>) als Objekt mit der gesamten Molino San Vincenzo-Collection verknüpfen. Auch die Geometriedaten in Form von 3D-Polygonen, die zwar mehrere Stratifikationseinheiten des Fundplatzes beschreiben, aber hier nur in einer einzigen SQLite-Datenbankdatei vorliegen und die Attribute zur qualitativen und quantitativen Beschreibung der einzelnen Befunde beinhalten (<http://phaidra.univie.ac.at/o:528465>), können so in den Archivdatenbestand eingegliedert werden.

Innerhalb der alles umfassenden Fundplatz-Collection können insgesamt also beinahe beliebig viele Sammlungen und Objekte angelegt und

laufend ergänzt werden, die z. B. die quantitativen, qualitativen und räumlichen Daten der Untersuchungsfläche der Oberflächenfundaufsammlung oder die Daten der geophysikalischen Prospektion beinhalten.³²

Diese Art der Langzeitarchivierung kann – als Alternative zum Beispiel zu einer strikten Ordnerstruktur – innerhalb von PHAIDRA ein netzwerkartiges Modell der archäologischen Ausgrabungsdaten auf diesem Fundplatz darstellen, womit auf einer Meta-Ebene auch eine Art der Nachvollziehbarkeit gegeben ist, die es gleichfalls für Dritte zulässt, die archivierten Daten und ihre Relationen zueinander in ihrer Gesamtheit einfacher erfassen zu können.

5. Diskussion und Conclusio

Für die Vielfalt archäologischer Daten ist ein Archivierungssystem nötig, das die Diversität der archäologischen Informationen detailliert beschreiben kann. Das angeführte Fallbeispiel zeigt, dass mit PHAIDRA unter Anwendung einer konkreten Archivierungsstrategie genau ein solches System zur Verfügung. Hier ist es möglich, Daten unterschiedlicher Ausprägung in das Repositorium aufzunehmen, ein Umstand, der sich sehr zum Vorteil der Archivierung digitaler archäologischer Daten auswirkt. Zudem eignet sich PHAIDRA besonders, die aufgezeigte Diversität archäologischer Daten nachvollziehbar zu mappen. Dies setzt jedoch zuzüglich zur Generierung von archivfähigen (Meta-)Daten zeitintensive Überlegungen und Vorarbeiten voraus, die die mitunter komplexen Relationen der archäologischen Daten zueinander definieren. Erhöhter Arbeitsaufwand ist auch durch das Anlegen der verschiedenen Collections gegeben, obgleich diese Arbeit durch die Verfügbarkeit von frei definierbaren Vorlagen erleichtert wird.

Nichtsdestoweniger können in PHAIDRA die aus verschiedenen Untersuchungsmaßstäben stammenden Daten hierarchisch, kategorisch und logisch organisiert werden, indem die einzelnen Datenobjekte (etwa 3D-Geometrien mit Attributtabelle, Dokumentationsfotos und 3D-Oberflächenmodelle) in skalierbaren Sammlungen und Untersammlungen (Collections) zusammengefasst und zusätzlich miteinander verknüpft werden. Angaben zur Urheberschaft des jeweiligen digitalen Objektes, der verwendeten Lizenzierung und detaillierte Beschreibungen hinsichtlich der forschungsgeschichtlichen Rahmenbedingungen können den Datenobjekten ebenso hinzugefügt werden wie Angaben zur Versionierung, den technischen Spezifikationen und einer Klassifikation bzw. näheren Charakteri-

sierung der einzelnen Ressourcen anhand kontrollierter Vokabulare. Diese und weitere Metadaten werden strukturiert anhand verschiedener Schemata erfasst und können bei jedem Objekt separat abgefragt werden. Alle Objekte und Sammlungen verfügen über eindeutige *persistent identifiers* (Permalink, Handle und [selektiv] DOI) und können so zitiert werden.³³ All diese Vorteile wiegen einen etwaigen Mehraufwand in der Anlage des Archivdatenbestandes deutlich auf. Durch Systeme wie PHAIDRA wird es ermöglicht, die inhomogene Struktur archäologischer Datensätze effektiv zu speichern, die durchgeführten Forschungen so zu dokumentierten und für eine nachhaltige wissenschaftliche Nutzung anhaltend zur Verfügung zu stellen.

Dominik Hagmann, BA MA

ORCID: <http://orcid.org/0000-0002-4481-6234>

Universität Wien, Institut für Klassische Archäologie

E-Mail: dominik.hagmann@univie.ac.at

- 1 vgl. M. Doneus, Die hinterlassene Landschaft. Prospektion und Interpretation in der Landschaftsarchäologie, Mitteilungen der Prähistorischen Kommission der Österreichischen Akademie der Wissenschaften 78 (Wien 2013). <http://hw.oeaw.ac.at/7197-3?frames=yes> (16.06.2018); E. C. Harris, Principles of Archaeological Stratigraphy² (London 1989); G. Schörner – D. Hagmann, Intensiver archäologischer Survey im nördlichen Etrurien, Forum Archaeologiae 79/IX, 2015. <https://hdl.handle.net/11353/10.407473> (16.06.2018); R. Vogt, Archäologische Prospektion. Bohrungen und Sondagen, Grabungstechnikerhandbuch, Aktualisierung 2011. http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Grabungstechniker/Grabungstechnikerhandbuch/26_2_Bohrungen_und_Sondagen.pdf (16.06.2018).
- 2 Ich danke an dieser Stelle den Herausgeber*innen sowie dem Redaktionsteam für die Aufnahme des Artikels in die aktuelle Ausgabe, Susanne Blumesberger für ihre Unterstützung sowie den anonymen Gutachter*innen für die ausführlichen und überaus hilfreichen Hinweise.
- 3 vgl. J. Huggett – S. Ross, Archaeological Infomatics – Beyond Technology. Themed Issue Edited by Jeremy Huggett and Seamus Ross. Introduction, Internet Archaeology 15, 2004. <http://dx.doi.org/10.11141/ia.15.13> (07.03.2018); E. B. W. Zubrow, Digital Archaeology. A Historical Context. In: T. L. Evans – P. T. Daly (Hrsg.), Digital Archaeology. Bridging Method and Theory (London, New York 2006).

- 4 vgl. M. Forte – S. Campana (Hrsg.), *Digital Methods and Remote Sensing in Archaeology* (Cham 2016). <http://dx.doi.org/10.1007/978-3-319-40658-9> (06.03.2018); P. Jensen, *Where are We? Reviewing the Integration of Complex Spatial Data in Current Field Archaeology*, *Internet Archaeology* 44, 2017. <http://dx.doi.org/10.11141/ia.44.8>.
- 5 s. dazu kritisch M. Brouwer Burg, *It Must Be Right, GIS Told Me So! Questioning the Infallibility of GIS as a Methodological Tool*, *Journal of Archaeological Science* 84, 2017, 115–120. <http://dx.doi.org/10.1016/j.jas.2017.05.010> (06.03.2018).
- 6 P. Allison, *Dealing with Legacy Data. An Introduction*, *Internet Archaeology* 24, 2008. <http://dx.doi.org/10.11141/ia.24.8> (12.03.2018).
- 7 vgl. etwa ArchAIDE. <http://www.archaide.eu/> (07.03.2018); E. Aspöck – H. Di Giuseppe – E. Fentress – G. Geser – A. Masur – J. Ogden – T. Oikarinen – D. Tsiafakis, *Survey of Excavation Data Management*. DA-RIAH-EU Community Engagement Working Group blog, 21.04.2015. <https://dariahre.hypotheses.org/327> (07.03.2018); B. R. Barricelli – S. Valtolina – M. Marzullo, *ArchMatrix. A Visual Interactive System for Graph-Based Knowledge Exploration in Archaeology*. In: G. Tortora – S. Levaldi – M. Tucci (Hrsg.), *Proceedings of the International Working Conference on Advanced Visual Interfaces. AVI'12 International Working Conference on Advanced Visual Interfaces Capri Island (Naples), Italy. May 22–25, 2012 (New York 2012)*. <http://dx.doi.org/10.1145/2254556.2254684> (07.03.2018); CIDOC CRM. <http://www.cidoc-crm.org/> (08.03.2018); M. L. Gualandi – R. Scopigno – L. Wolf – J. Richards – J. B. i. Garrigos – M. Heinzelmann – M. A. Hervas – L. Vila – M. Zallocco, *ArchAIDE. Archaeological Automatic Interpretation and Documentation of cEramics*. In: C. E. Catalano – L. D. Luca (Hrsg.), *Eurographics Workshop on Graphics and Cultural Heritage (2016)*. <http://dx.doi.org/10.2312/gch.20161408> (07.03.2018); K. May – C. Binding – D. Tudhope, *Barriers and Opportunities for Linked Open Data Use in Archaeology and Cultural Heritage*, in: W. Schön – F. Siegmund (Hrsg.), *Forum Open Access und Open Data, Sammlungsstrategien auf dem Prüfstand, Denkmalschutz in NRW, Schatzregal, Deutscher Studienpreis für Archäologie, Archäologische Informationen* 38 (Bonn 2015), 173–184. <http://dx.doi.org/10.11588/ai.2015.1.26162> (07.03.2018).
- 8 FileMaker. <http://www.filemaker.com/de/> (12.03.2018); Inari AIS. <https://www.inari-software.com/de/archaeologie/> (08.03.2018); LibreOffice Base. <https://de.libreoffice.org/discover/base/> (08.03.2018); Microsoft Access. <https://products.office.com/de-at/access> (08.03.2018);

- OpenAtlas. <http://www.openatlas.eu/website/> (08.03.2018).
- 9 eine umfangreiche Aufstellung findet sich bei H. Wright – S. Cuy – L. Fentress – K. Fernie – P. Gerth – C. Gonzalez-Perez – H. Hollander – A. Kreiter – F. Mambrini – A. Masur – S. Matskevich – W. Schmiedle – S. Thänert – J. Urzúa, Data Management Systems. Online-Dokument, 08.05.2015. <http://www.ariadne-infrastructure.eu/Community/Special-Interest-Groups/Excavation-and-Monument-Data> (12.03.2018).
- 10 S. Jeffrey, A New Digital Dark Age? Collaborative Web Tools, Social Media and Long-term Preservation, *World Archaeology* 44,4, 2012, 553–570. <http://dx.doi.org/10.1080/00438243.2012.737579> (07.03.2018).
- 11 z.B. GAMS. <http://gams.uni-graz.at/> (08.03.2018); Humanities Commons CORE Repository. <https://hcommons.org/> (08.03.2018); tDAR. <https://www.tdar.org/> (08.03.2018); Zenodo. <https://zenodo.org/> (08.03.2018).
- 12 z. B. Archaeology Data Service. <http://archaeologydataservice.ac.uk/> (08.03.2018); IANUS Forschungsdatenzentrum Archäologie und Altertumswissenschaften. <https://www.ianus-fdz.de/> (08.03.2018).
- 13 ARCHE. <https://arche.acdh.oeaw.ac.at/browser/> (08.03.2018).
- 14 s. hierzu allgemein etwa die Vienna Principles. <http://viennaprinciples.org/> bzw. P. Kraker et al., The Vienna Principles: A Vision for Scholarly Communication in the 21st Century. Zenodo 2016. <http://doi.org/10.5281/zenodo.55597> (08.03.2018).
- 15 vgl. hierzu ergänzend E. C. Kansa – S. W. Kansa, Open Archaeology. We All Know That a 14 Is a Sheep. Data Publication and Professionalism in Archaeological Communication, *Journal of Eastern Mediterranean Archaeology & Heritage Studies* 1,1, 2013, 88. <http://dx.doi.org/10.5325/jeasmedarcherstu.1.1.0088> (07.03.2018).
- 16 vgl. etwa B.-C. Björk – D. Solomon, Open Access versus Subscription Journals. A Comparison of Scientific Impact, *BMC Medicine* 10,73, 2012. <http://dx.doi.org/10.1186/1741-7015-10-73> (08.03.2018). S. konkret im Hinblick auf die Ausgrabung Molino San Vincenzo hinsichtlich der Datendisseminationsaktivitäten D. Hagmann, Reflections on the Use of Social Networking Sites as an Interactive Tool for Data Dissemination in Digital Archaeology – Online First, *Interdisciplinaria Archaeologica. Natural Sciences in Archaeology* 9/1, 2018. <http://dx.doi.org/10.17613/M6V56W> (06.06.2018).
- 17 G. Beale – P. Reilly, Digital Practice as Meaning Making in Archaeology, *Internet Archaeology* 44, 2017. <http://dx.doi.org/10.11141/ia.44.13> (08.03.2018); s. als dezidiertes Beispiel: D. Hagmann – V.

- Schreck, Neue Forschungen zum ländlichen Fundplatz Molino San Vincenzo (Toskana, Italien). In: G. Schörner – K. Meinecke (Hrsg.), Akten des 16. Österreichischen Archäologentages am Institut für Klassische Archäologie der Universität Wien vom 25. bis 27. Februar 2016, Wiener Forschungen zur Archäologie 17 (Wien 2018), 171–174 (163–174).
- 18 PHAIDRA (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets). <https://phaidra.univie.ac.at/> (08.03.2018).
- 19 S. Blumesberger, Repositorien: mehr als Archivierungsorte. Am Beispiel von PHAIDRA im Bereich der Kinder- und Jugendliteraturforschung, *biblos. Beiträge zu Buch, Bibliothek und Schrift* 2013/62, 141–157. Handle: <https://hdl.handle.net/11353/10.296060> (08.03.2018).
- 20 Einen grundlegenden Überblick findet sich bei P. Budroni, PHAIDRA. Ein Projekt zur Langzeitarchivierung von digitalen Beständen an der Universität Wien. <http://phaidra.univie.ac.at/o:153> (06.06.2018).
- 21 PHAIDRA auf Github. <https://github.com/phaidra> (06.06.2018).
- 22 PHAIDRA. <http://phaidra.org/> (06.06.2018).
- 23 Registry of Research Data Repositories. <http://www.re3data.org/> (06.06.2018); nestor-Arbeitsgruppe OAIS-Übersetzung / Terminologie – G. Büchler – N. Bütikofer – S. Dobratz – K. Ernst – C. Keitel – J. Ludwig – S. Schrimpf – K. Schwarz – S. Gregorio, Referenzmodell für ein Offenes Archiv-Informationssystem. Deutsche Übersetzung 2.0, nestor-materialien 16, 2013. <http://d-nb.info/104761314X/34> (06.06.2018).
- 24 I. Subirat – I. Solodovnik – P. Budroni – R. Ganguly – R. Hudak, COAR Case Study Controlled Vocabularies and PHAIDRA International. In: F. Loizides – B. Schmidt (Hrsg.), *Positioning and Power in Academic Publishing. Players, Agents and Agendas: Proceedings of the 20th International Conference on Electronic Publishing* (Amsterdam 2016), 117. <https://elpub.architecturez.net/doc/oai-elpub-id-117-elpub2016> (16.06.2018).
- 25 P. Harpring, *Introduction to Controlled Vocabularies. Terminology for Art, Architecture, and Other Cultural Works* (Los Angeles 2010). http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/index.html (16.06.2018), 12–25; M. Lang – G. Carver – S. Printz, Standardised Vocabulary in Archaeological Databases. In: G. P. Earl – T. Sly – A. Chrysanthi – P. Murrieta-Flores – C. Papadopoulos – I. Romanowska – D. Wheatley (Hrsg.), *Archaeology in the Digital Era II. e-Papers from the 40th Conference on Computer Applications and Quantitative Methods in Archaeology*. Southampton, 26–30 March 2012 (Amsterdam 2013), 468f. <http://arno.uva.nl/cgi/arno/show.cgi?fid=545855> (16.06.2018).
- 26 vgl. etwa D. Hagmann – V. Schreck, *L'insediamento rurale romano di*

- Molino San Vincenzo a Montespertoli, *Milliarium* 12, 2018, 42–59.
- 27 Stand: 06.06.2018, inklusive aller Rohdaten, Arbeitskopien etc.
- 28 vgl. etwa M. Doneus – G. Verhoeven – M. Fera – C. Briese – M. Kucera – W. Neubauer, *From Deposit to Point Cloud. A Study of Low-Cost Computer Vision Approaches for the Straightforward Documentation of Archaeological Excavations*, *Geoinformatics FCE CTU* 6, 2011, 81–88. <http://dx.doi.org/10.14311/gi.6.11> (07.03.2018).
- 29 Dominik Hagmann – Johannes Reiter, *Molino San Vincenzo: 3d model of trench 2/2014 (3d model)*, 2016. <http://dx.doi.org/10.13140/RG.2.2.33108.50565> (08.03.2018); Dominik Hagmann – Johannes Reiter, *Molino San Vincenzo: 3d model of trench 2/2014 (texture)*, 2016. <http://dx.doi.org/10.13140/RG.2.2.16331.28967> (08.03.2018).
- 30 P. Budroni – S. Blumesberger – T. Miksa – B. Sánchez Solís, *Datenmanagementplan für Projekte an der Universität Wien. Empfohlenes Repositorium: Phaidra*, 2015. <http://phaidra.univie.ac.at/o:407954> (06.06.2018).
- 31 z.B. *Archaeology Data Service*. <http://archaeologydataservice.ac.uk/> (06.05.2018); *IANUS Forschungsdatenzentrum*. <http://www.ianus-fdz.de/> (06.06.2018).
- 32 Die Möglichkeit, Collections zu erstellen beziehungsweise einzelne Sammlungen und Objekte miteinander zu verknüpfen ist innerhalb des Repositoriums PHAIDRA innerhalb der technischen Möglichkeiten nicht weiter eingeschränkt und erlaubt darum beinahe zahllose weitere Verbindungen mit unterschiedlichen Daten.
- 33 S. Blumesberger, *Digitale Objekte sichern, beschreiben, archivieren und rasch verbreiten. Wie das digitale Langzeitarchivierungssystem PHAIDRA an der Universität Wien eingesetzt werden kann*. In: P. Mismelius – W. Stützl – T. Hug – P. Grell – R. Kammerl (Hrsg.), *Freie Bildungsmedien und Digitale Archive* (Innsbruck 2014). <http://books.openedition.org/iup/980> (12.03.2018).