

# Zeitschrift für digitale Geisteswissenschaften

---

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

»Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos

---

Autor/in:

Thorsten Wübbena

Kontakt: [twuebbena@dt-forum.org](mailto:twuebbena@dt-forum.org)

Institution: Deutsches Forum für Kunstgeschichte Paris

GND: [123312396](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-23312396) ORCID: [0000-0001-8172-6097](https://orcid.org/0000-0001-8172-6097)

---

Autor/in:

Eric Decker

Kontakt: [decker@asia-europe.uni-heidelberg.de](mailto:decker@asia-europe.uni-heidelberg.de)

Institution: Universität Heidelberg

GND: [1084009560](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1084009560) ORCID: [0000-0003-3035-2413](https://orcid.org/0000-0003-3035-2413)

---

Autor/in:

Matthias Arnold

Kontakt: [arnold@asia-europe.uni-heidelberg.de](mailto:arnold@asia-europe.uni-heidelberg.de)

Institution: Universität Heidelberg

GND: [1049623088](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1049623088) ORCID: [0000-0003-0876-6177](https://orcid.org/0000-0003-0876-6177)

---

DOI des Artikels:

[10.17175/sb001\\_018](https://doi.org/10.17175/sb001_018)


Nachweis im OPAC der Herzog August Bibliothek:

[830205977](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-830205977)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Musikvideo](#) | [Computergestütztes Verfahren](#) | [Annotation](#) |

Zitierweise:

Thorsten Wübbena, Eric Decker, Matthias Arnold: »Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001\\_018](https://doi.org/10.17175/sb001_018).

Thorsten Wübbena, Eric Decker, Matthias Arnold

## »Losing My Religion« – Einsatz der Videoannotationsdatenbank Pan.do/ra in der kunstgeschichtlichen Analyse von Musikvideos

---

### Abstracts

In dem vorliegenden Beitrag werden die Möglichkeiten sowie der daraus resultierende Mehrwert der Anwendung einer webbasierten Datenbank für Videoannotationen (Pan.do/ra) in der kulturwissenschaftlichen Untersuchung von Bewegtbildmaterial beleuchtet. Der erste Teil beschäftigt sich am Beispiel eines Musikvideos konkret mit einer Anzahl von Nutzungsmöglichkeiten, während der anschließende zweite Teil das verwendete System in einen allgemeineren Kontext einbettet.

In this paper, we discuss the possibilities as well as the added value of using a web-based databank for video annotation (Pan.do/ra) in cultural studies research on film. The first part demonstrates a variety of potential uses through the example of a music video, and the second part introduces the software on a more general level.

## 1.

### 1.1 Intro

Im ersten Teil des Beitrags<sup>1</sup> zur Nutzung der Videoannotationsplattform [Pan.do/ra – Open media archive](#)<sup>2</sup> im kulturwissenschaftlichen Kontext wird der Schwerpunkt auf den (mit diesem Werkzeug) zu bearbeitenden Bewegtbildinhalten – in diesem Fall Musikvideos – und den damit verbundenen Forschungsfragen liegen, während Matthias Arnold und Eric Decker von [HRA](#), der *Heidelberg Research Architecture*, im sich daran anschließenden zweiten Teil die dahinter liegenden technischen Aspekte und die globaleren Möglichkeiten darstellen und diskutieren werden. Zusammengefasst soll am Ende deutlich werden, dass durch die Verzahnung der Informationstechnologie mit der Fachwissenschaft ein Mehrwert erreicht werden kann, der durch analoge Möglichkeiten nicht hätte erzielt werden können – sozusagen ein geradezu idealtypisches Modell aus dem Gebiet der Digital Humanities.

### 1.2 Das Musikvideo als Forschungsgegenstand

---

<sup>1</sup> Autor dieses Teils ist Thorsten Wübbena.

<sup>2</sup> Pan.do/ra ist eine Online-Applikation, die es – kurz gesagt – mehreren Nutzern gleichzeitig erlaubt, frei gewählte Abschnitte eines Videos mit unterschiedlichen Arten von Annotationen (Layer oder Spuren) zu versehen. Durch diese Annotationen ist es möglich, sowohl einzelne Sequenzen als auch sämtliche Szenen zu einer Thematik direkt über eine URL zu referenzieren. Jede Spur enthält zeitreferenzierte Annotationen eines bestimmten Typs und neben Text können auch georeferenzierte Ortsinformationen und Bilder hinzugefügt werden. Mit seiner integrierten Rechteverwaltung erlaubt es das System, Filme gezielt bestimmten Nutzergruppen zugänglich zu machen oder Filme samt Annotationen frei zugänglich zu machen.

Zahlreiche Musikvideos weisen über ihren genuinen Zweck als Werbeträger hinaus einen ästhetischen Mehrwert auf, der eine eingehendere bildwissenschaftliche Betrachtung herausfordert und aufzeigt, dass sich in diesem Bereich eine eigene Kunstform entwickelt hat. Die künstlerischen Ausdrucksmöglichkeiten, die hier durch die flexible Verknüpfung von Text, Bild und Musik vorhanden sind, eröffnen Szenarien, in denen Text und Musik nicht nur einen kurzen Schwall von heterogenen Bildern zusammenhalten, vielmehr wird ein komplexer Diskurs auf drei Ebenen gestaltet. Das Musikvideo als visuelles Produkt referenziert dabei auch immer wieder auf »verwandte« Medien (Kino- und Fernsehfilm), aber auch auf Werke der bildenden Kunst, andere Kommunikations- und Unterhaltungsformen und auf sich selbst.

Und obgleich das Thema des Musikvideos inzwischen eine verstärkte Aufmerksamkeit innerhalb von Disziplinen wie der Film-, Musik- und Medienwissenschaft sowie der Kunstgeschichte erfahren hat,<sup>3</sup> fehlt es in der wissenschaftlichen Bearbeitung des Themas immer noch an einer Auswahl adäquater Werkzeuge zur Unterstützung der Analyse des Untersuchungsgegenstandes und der Präsentation der Ergebnisse.<sup>4</sup>

In der Vergangenheit habe ich gemeinsam mit Henry Keazor zu zeigen versucht, dass eine fruchtbare Betrachtung des Musikvideos nur durch den Blick auf die drei oben erwähnten Ebenen sinnvoll ist.<sup>5</sup> Ausgehend von dieser Vorgehensweise war das Ansinnen stets, die internen und externen Bezüge einzelner Clips herauszuarbeiten und zu analysieren, wobei bis dato zu bedauern war, dass nicht in stärkerem Maße im Medium selbst gearbeitet werden konnte, sondern jeweils schon bereits zu Beginn – in der Phase der Beobachtungen – auf die Übertragung in eine Textform zurückgegriffen werden musste. Text, Musik und Bild mussten dabei getrennt verhandelt werden und in der Kategorie »Bild« konnte nur mit Standbildern gearbeitet werden.

### 1.3 Losing My Religion

Am Beispiel des Clips zu R.E.M.'s *Losing My Religion*<sup>6</sup>, der 1991 vom Regisseur Tarsem Singh<sup>7</sup> umgesetzt wurde, soll nun erläutert und zugleich dargestellt – bzw. in diesem Rahmen

---

<sup>3</sup> Vgl. dazu die Bibliografien in Keazor / Wübbena 2010 sowie die Tatsache, dass auf dem Kunsthistorikertag 2013 in Greifswald dem Musikvideo eine eigene Sektion gewidmet wurde (die Vorträge der Sektion erscheinen in Keazor 2014).

<sup>4</sup> Erwähnenswert ist an dieser Stelle einzig der Versuch, der an Universität Basel im Bereich der Medienwissenschaften zwischen April 2008 und Juni 2011 unternommen wurde, ein Analyseinstrumentarium zur Interpretation von Inszenierungsformen in Musikvideos zu entwickeln. Das Ergebnis *traVis* (Musikzentriertes Transkriptionsprogramm für audiovisuelle Medienprodukte) wurde im Februar 2013 an der Universität Basel im Rahmen eines Workshops vorgestellt und getestet. Hierbei erwiesen sich bei allen neuen Möglichkeiten schnell die Grenzen von *traVis*: Wie es die Beibezeichnung schon andeutet, ist das Instrument musikzentriert, d.h. die Parameter von Bild und Text kommen eindeutig zu kurz – in letzter Instanz werden sie der Musik doch wieder untergeordnet, wodurch die eigentlich anvisierte Gegenstandsangemessenheit verfehlt wird. Denn in Musikvideos kommen laufende Bilder, gesungener und vorgetragener Text gleichberechtigt zum Zug. Ein weiterer Nachteil ist zudem, dass das Programm dazu führt, dass dem temporalen Verlauf der Clips (bzw. genauer: der ihnen zugrundeliegenden Musik) zu stark gefolgt wird, so dass der Überblick über die Gesamtstruktur des Musikvideos verloren geht, was bei dem Workshop immer wieder deutlich wurde. Damit ist umso ersichtlicher, welches großes Desiderat auf diesem Gebiet weiterhin existiert.

<sup>5</sup> Vgl. hierzu besonders Keazor / Wübbena 2011.

<sup>6</sup> Von dem 1991 veröffentlichten Album *Out of Time*.

<sup>7</sup> Nachdem der 1962 in Indien geborene zukünftige Regisseur ein im Himalaya gelegenes Internat besucht hatte, zog er in die Vereinigten Staaten, um an der Harvard Business School auf einen Master of Business

eher angedeutet – werden, wie stark eine Videoannotationssoftware wie das frei verfügbare Pan.do/ra als webbasiertes Werkzeug in der kunsthistorischen Medienanalyse unterstützend eingesetzt werden kann.<sup>8</sup>

Es ist im Nachhinein schwer zu sagen, inwieweit sich der Erfolg des Songs seiner Präsentation in Clip-Form verdankt, doch allein der Umstand, dass der Clip nicht nur im Erscheinungsjahr Aussichten auf einen Grammy hatte, sondern auch für gleich acht MTV Video Music Awards nominiert war (sechs der Trophäen erhielt das Video dann auch, darunter in den Kategorien »Video of the Year«, »Breakthrough Video« und »Best Direction in a Video«), zeigt, dass man den Anteil der visuellen Interpretation am Erfolg des Stücks keinesfalls unterschätzen darf – zumal das Video den Liedtext in eine Richtung hin interpretiert, die zwar von Texter und Leadsänger Michel Stipe möglicherweise mitgedacht (vgl. auch die Liedzeile »Choosing my confessions«), doch keinesfalls ausschließlich intendiert gewesen ist; und es ist eben diese Deutung, die dem Song den (weite Verbreitung sichernden) Stempel des Skandals und der Anrüchigkeit aufdrückt. Denn die Titelzeile »Losing my religion« ist als Hinweis darauf verstanden worden, dass der Protagonist des Textes seinen Glauben verliert. Tatsächlich aber hat Stipe damit eine vergessene Redewendung aus den Südstaaten der USA aufgegriffen, die soviel bedeutet wie »die Fassung zu verlieren«.<sup>9</sup>

Wie die vielen Hörer nach ihm (und diese dann möglicherweise bereits im Gefolge seiner visuellen Interpretation) konzentriert sich Regisseur Tarsem Singh bei der Konzeption seines Videos auf die mit den Reizworten »religion« und »confessions« angedeutete Sphäre des Glaubens und der Frömmigkeit. Und um diese sowie den im Songtitel angedeuteten Verlust derselben auszugestalten, legt er dem Clip eine Dramaturgie zugrunde, in der es letztendlich um die Durchdringung zweier Ebenen geht: Das ist zum einen die religiöse Sphäre, repräsentiert durch Szenerien, welche dem hinduistischen wie dem christlichen Bilderkosmos entlehnt sind und in letzterem Fall auf Werke der bildenden Kunst des 17. und 20. Jahrhunderts rekurrieren, namentlich auf Schöpfungen des Fotografenduos Pierre et Gilles und des Barockmalers Caravaggio. Ihnen gegenüber gestellt werden nach dem ersten Viertel des Videos und mit Beginn der zweiten Strophe die unter ihnen auf der Erde lebenden Menschen, die in ihrer fahlen Farbigkeit, den starken Hell-Dunkel-Kontrasten und den robusten, eher ärmlich und archaisch anmutenden Typen an Gemälde der Caravaggio-Schule erinnern, die hier gleichwohl nie detailgetreu umgesetzt werden, sondern lediglich einzelne Details liefern (wie z.B. den aus Caravaggios »Berufung des Heiligen Matthäus« bekannten, diagonal die Szene durchziehenden Lichtstrahl.

---

Administration hin zu studieren. Doch schon bald beschloss er, lieber das Fach »Film« zu belegen und schrieb sich daher am renommierten Art Center College of Design in Pasadena ein. Singh entwickelte sodann seinen oft als opulent und dynamisch bezeichneten visuellen Stil, so u.a. in dem Video zu Suzanne Vegas Song *Tired of Sleeping*. Seinen Durchbruch erzielte er jedoch mit dem hier vorgestellten Musikvideo.

<sup>8</sup> Dass an dieser Stelle und im Folgenden kein Verweis auf den Videoclip (oder Teile davon) in der Videoannotationssoftware auf dem Server der HRA zu finden ist, liegt an der momentan vorherrschenden großen rechtlichen Unsicherheit in Bezug auf die Online-Bereitstellung von Bildmaterial (statisch oder bewegt), so dass der Serverbetreiber sich in diesem Fall keinem Risiko aussetzen wollte. Deutlich wird daran, dass trotz der bereitstehenden Optionen von freien Online-Werkzeugen eine Einschränkung der Nutzung für bestimmte Themen der Forschung besteht, da letztlich die Frage der Möglichkeit der Verfügbarmachung des Contents entscheidet.

<sup>9</sup> Vgl. dazu Bruder 2004, S. 85.

## 1.4 Pan.do/ra

Wie sehen die Möglichkeiten des Umgangs mit dem Musikvideo nun in der Annotationssoftware aus (vgl. Abbildung 1)?<sup>10</sup>



Abb. 1: Startseite von Pan.do/ra. Screenshot, erstellt am 17.06.2015, © HRA, CC-Lizenz 0.

Da dieses Musikvideo eine mannigfaltige Anzahl von nachgehenswerten Referenzen beinhaltet, müssen an dieser Stelle einige wenige Beispiele ausreichen. Es soll dabei versucht werden, die Möglichkeiten der Software aus Anwendersicht dicht an den Inhalten zu beleuchten – bei diesem Vorgehen möge man die teils eher deskriptive Art verzeihen.<sup>11</sup>

Der Leadsänger von R.E.M, Michael Stipe, erscheint von Anfang an immer wieder mit riesigen Engelsflügeln, was sich sodann leitmotivisch durch das ganze Video zieht (Abbildungen 2a-d).



Abb. 2a-d: Stills aus dem Clip *Losing My Religion* – Sänger Michael Stipe mit Flügeln und weitere Szenen (beispielhaft) mit dem Motiv der Engelsflügel in der Editor-Ansicht von Pan.do/ra, jeweils mit dem Tag »Engelsflügel« versehen (vgl. im rechten Bereich den Punkt »Keywords«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.

---

<sup>10</sup> Auch an dieser Stelle noch einmal deutlich der Hinweis auf die unbefriedigende Lösung mit Filmstills arbeiten zu müssen. Ein Vorgehen, welches ja gerade durch den Einsatz der webbasierten Videoannotationssoftware überwunden werden sollte, aber hier wegen der rechtlichen Bedenken der Serverbetreiber notwendig ist.

<sup>11</sup> Mehr zur inhaltlichen Auseinandersetzung in Keazor / Wübbena 2011, passim.

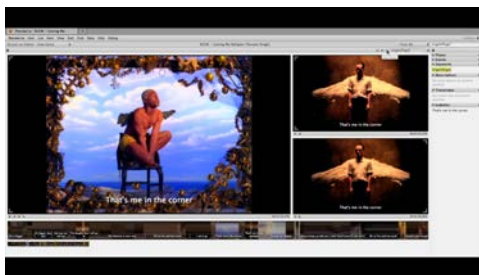


Abb. 2b



Abb. 2c



Abb. 2d

Eine dieser Szenen stellt ein Gemälde des Barockmalers Caravaggio mit dem in seiner Verzückung von einem Engel gehaltenen Heiligen Franziskus nach, der hier im Clip allerdings durch einen zweiten, älteren Himmelsboten ersetzt wird (vgl. Abbildung 3 und Abbildung 4).<sup>12</sup>

---

<sup>12</sup> Open Source-Bilddatenbanksoftware der Firma [Coneda UG](#), einer Ausgründung des Kunstgeschichtlichen Instituts der Goethe-Universität Frankfurt. Der Link an dieser Stelle in Pan.do/ra führt in [die Installation des Frankfurter Kunstgeschichtlichen Instituts](#).

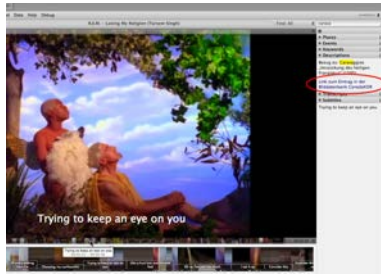


Abb. 3: Link aus Pan.do/ra zum Bilddatenbanksystem ConedaKOR. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.

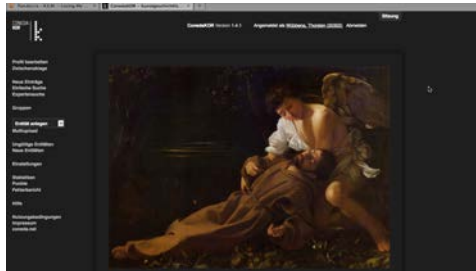


Abb. 4: Entitätsansicht in ConedaKOR: Michelangelo Merisi da Caravaggio, *Verzückung des heiligen Franziskus*, um 1595, Hartford, Wadsworth Atheneum Museum of Art. Quelle: KGI (Frankfurt/Main), [Web Gallery of Art](#), erstellt am 12.10.2014.

In dem Moment, wo der alte Engel versehentlich zur Erde herabstürzt und der andere Engel nach ihm zu greifen versucht, zitiert Singh ein weiteres Gemälde Caravaggios (*Die sieben Werke der Barmherzigkeit*), in dem ebenfalls prägnant das Motiv eines zur Erde hinabgereckten Engels-Armes erscheint (Abbildung 5).



Abb. 5: Interner Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Die sieben Werke der Barmherzigkeit*, 1607, Neapel, Chiesa del Pio Monte della Misericordia. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.

Wenn der ältere, gefallene Engel auf der Erde sodann auf neugierige, argwöhnische und grausame Weise von den Menschen untersucht wird, spielt Singh wiederum auf ein Bild Caravaggios (*Der ungläubige Thomas*) an (Abbildung 6).



Abb. 6: Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Der ungläubige Thomas*, 1603, Potsdam, Bildergalerie Schloss Sanssouci. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.

Die solcherart nachgestellten oder zitierten Gemälde werden dabei zuweilen in einer poppig-bunten und kitschig-schriellen Ästhetik präsentiert, die sich an den Werken eines französischen Künstlerpaars, des Fotografen Pierre (Commy) und des Malers Gilles (Blanchard)<sup>13</sup> orientiert, deren *Heiliger Sebastian* dann auch in einer Einstellung des Clips zitiert wird (Abbildung 7).



Abb. 7: Einbettung: Pierre & Gilles, *St. Sebastian*, 1987. Quelle: HRA, KGI (Frankfurt/Main), [Art Forum](#), Warner Bros., erstellt am 12.10.2014.

Abgerundet wird das Personal des Videos durch Männer, die zum einen für die Beleuchtung zuständig zu sein scheinen, zugleich jedoch – als eine eiserne Flügel-Skulptur herstellende Schmiede – als Vertreter der (russisch-kommunistischen) Arbeiterklasse ausgewiesen werden, die dem Jenseitsbezug der Religion das erdverhaftete Vertrauen in die technischen Fähigkeiten des Menschen gegenüberstellen (Abbildung 8a und Abbildung 8b).

---

<sup>13</sup> Vgl. zum Werk Marcadé / Cameron 1997.





Abb. 8a-b: Interner Bezug durch Tagging, verdeutlicht durch eine farbige Kennzeichnung auf der unteren Timeline (in diesem Fall »Arbeiter«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.



Abb. 8b

Der Bezug zu Russland wird ebenfalls in Form des Kostüms und der Choreographie von Stipe hergestellt: Sie sind beide an Aufnahmen vom Auftritt des berühmten Tänzers Vaslav Nijinsky in dem von Claude Debussy geschriebenen Ballett *Jeux* aus dem Jahre 1913 angelehnt (Abbildung 9a und Abbildung 9b).



Abb. 9a-b: Einbettung: Abbildungen des Tanzes von V. Nijinsky zu *Jeux*. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.



Abb. 9b

Allein die hier aufgezeigten Bezüge (die z.B. noch um den Verweis auf die 1968 erschienene Kurzgeschichte *Ein sehr alter Mann mit riesengroßen Flügeln* des kolumbianischen Schriftstellers Gabriel García Márquez zu ergänzen wäre, aus der das Motiv des zur Erde gestürzten und dort von Menschen misshandelten Engels stammt) machen deutlich, welche Vielfalt an Verweisen und Zitaten in einen Clip wie demjenigen Singhs eingewoben sein kann.

Im Jahr 2000 hat Tarsem Singh diese im Musikvideo entwickelte Bilderwelt dann auch in seinem Kinodebüt *The Cell* unterzubringen versucht, und so begegnet man dort den bereits bekannten Motiven und Vorbildern, die zuweilen sogar zu identisch anmutenden Szenarien zusammengeführt werden – fast scheint es, als seien einzelne Szenen in der Kulisse des R.E.M.-Videos gedreht worden (vgl. Abbildung 10a und Abbildung 10b).



Abb. 10a-b: Interner Link zur korrespondierenden Filmszene aus *The Cell*, 2000 (Director: Tarsem Singh).  
Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., New Line Cinema (theatrical), Warner Home Video (DVD),  
erstellt am 12.10.2014.



Abb. 10b

## 1.5 Vorteile

Wo liegen aber nun die Vorteile in der Bearbeitung des Materials mit einer – und im Speziellen – mit dieser Videoannotationssoftware?

Für webbasierte Systeme gilt allgemein – nach wie vor erwähnenswert – die zeit- und ortsunabhängige Verfügbarkeit der Inhalte. Für die kollaborative Bearbeitung bei verteilten Forscher- und Studierendengruppen ist dies ein basaler, aber wichtiger Faktor.

Wenn ein Clip mit der Annotationssoftware bearbeitet ist (was aber natürlich nie als »abgeschlossen« betrachtet werden kann) und er für eine spätere Weiterbearbeitung unter einem neuen Aspekt angesehen wird, dann sind die Referenzen an der entsprechenden Stelle im Clip zu finden, ohne dass die entscheidenden Passagen wieder aus einem Text zusammengesucht werden müssen.

Dadurch ist auch für Nichtkenner ein exploratives Sehen des Clips möglich, da sie an den entscheidenden Punkten auf die Bezüge hingewiesen werden, nach denen sie nicht recherchieren konnten, weil sie schlicht nichts davon wussten.

Daneben ist sicherlich aber das Arbeiten im Clip selbst der wichtigste Punkt wissenschaftlicher Auseinandersetzung. Die Bezüge zu bewegten Einheiten, zu Filmausschnitten und nicht nur zu Stills setzen zu können, entspricht dem Medium und bietet die Möglichkeit, sozusagen »in Film zu denken« und zu arbeiten.

## 1.6 Things-To-Do

Bei der Beschäftigung mit Pan.do/ra konnten eine Reihe von Punkten identifiziert werden, um die das System noch erweitert werden muss, um Musikvideos in ihrem medienübergreifenden Kontext befriedigend analysieren zu können.

Als Beispiele seien hier nur die folgenden genannt:

1. Die direkte Gegenüberstellung mehrerer unterschiedlicher Videos.
2. Die Anpassbarkeit der Benutzeroberfläche an unterschiedliche Analysesituationen.
3. Die Möglichkeit der Annotation reiner Audiodaten ohne Bild.

Wie eingangs gezeigt, stellt sich dieser Forschungsbereich durch seine enge Verbindung von Bild-, Text- und Musik/Audio-Schichten bisher im Hinblick auf Informationsverwaltung und Visualisierung des komplexen Bezugssystems als Problemfall dar, und es erweist sich als Glücksfall, dass im Rahmen des Heidelberger DFG-Projekts **Zur ästhetischen Umsetzung von Musikvideos im Kontext von Handhelds** (Laufzeit: 2011–13) ein möglicher Weg aufgezeigt werden konnte, gemeinsam mit den Kollegen vom Heidelberg Research Architecture an ähnlichen Fragestellungen zu arbeiten.

## 2.

### 2.1 Digital Humanities am Heidelberger Exzellenzcluster

Die Heidelberg Research Architecture, kurz HRA, ist die Digital Humanities Unit am Exzellenzcluster **Asien und Europa im Globalen Kontext** der Universität Heidelberg. Zum Kerngeschäft der HRA gehört die Entwicklung des modularen Metadatenframeworks **Tamboti**. Es ermöglicht das gemeinsame Sammeln, Organisieren und Annotieren von digitalisiertem Forschungsmaterial (Text, Bild, Video, Audio). Das System erlaubt das Teilen der Daten und Metadaten innerhalb von Gruppen, von Forschungsgruppen bis hin zu einer weltweiten Veröffentlichung. Dabei setzt die HRA auf eine Kombination hochdeskriptiver internationaler Metadatenstandards (derzeit MODS, TEI und VRA Core) mit einem hohen Anteil kontrollierter multilingualer Vokabulare (wie die Getty Thesauri oder das Virtual International Authority File) und einer modularen Systemarchitektur, die das nachhaltige Erfassen und Annotieren multilingualer Metadaten nach wissenschaftlichen Standards erlaubt.

Tamboti ist für eine Kooperation auf Universitätsebene und darüber hinaus ausgelegt, wobei Kooperationen im Sinne einer gemeinsam nutzbaren Infrastruktur verstanden wird. Mit Tamboti entwickelt die HRA eine auf dem XML-Datenbanksystem **eXist-db** basierende Open Source-Lösung. Tamboti ist als Application für eXist-db weltweit frei nutz- und erweiterbar. Die HRA arbeitet derzeit an zusätzlichen Applikationen (Bildmodul Ziziphus und das Wiki-Modul), die Tamboti um weitere Funktionsbereiche modular erweitern.

Eine größere Nutzergruppe im Exzellenzcluster setzt in der Forschung auf die Arbeit an und mit visuellen Medien, daher liegt hier ein Schwerpunkt der Entwicklungen der HRA. Dabei wurde bereits während der ersten Förderphase des Exzellenzclusters auf die Kombination vorhandener Softwarelösungen mit Eigenentwicklungen gesetzt.

### 2.2 Videoannotationsdatenbank Pan.do/ra

Im Bereich der kollaborativen Videoannotation hat die HRA bereits in der ersten Förderphase des Exzellenzclusters (2007–2012) in mehreren Projekten Erfahrungen sammeln können. Nach einer Bedarfsanalyse und Tests unterschiedlicher frei zugänglicher Softwarelösungen zur Videoannotation fiel die Wahl auf **pad.ma**, die von einem freien Entwicklerteam<sup>14</sup> als Open Source-Plattform entwickelt wird und jetzt in der überarbeiteten

---

<sup>14</sup> Die Pan.do/ra-Software wurde ursprünglich im Rahmen des Pad.ma (Public Access Digital Media Archive) Projekts federführend von Jan Gerber, Sebastian Lütgert und Sanjay Bhangar entwickelt. Pad.ma wurde von einer Gruppe von Organisationen aus Indien und Deutschland ins Leben gerufen: CAMP (Mumbai), Alternative Law Forum (Bangalore) und 0x2620 (Berlin). Des Weiteren waren Majlis und Point of View aus Mumbai mit beteiligt. Ausgestattet mit einer Anschubfinanzierung des niederländischen Humanistisch Instituut voor Ontwikkelingssamenwerking (HIVOS) konnte eine erste Version der Videoannotationsdatenbank entwickelt werden. Der weitere Ausbau wird seit 2010 durch die Unterstützung der Bohn Foundation und Foundation for Arts Initiatives ermöglicht. Vgl. [About Pad.ma](#), [About 0xDB](#) und die [faqs](#) der Pan.do/ra Webseite.

Pan.do/ra-Version in Heidelberg installiert ist.<sup>15</sup> Eine Reihe anderer Projekte setzen auch Pan.do/ra ein, darunter 0xDB, Pad.ma und Indiancine.ma.<sup>16</sup>

Ausschlaggebend für die Wahl von pad.ma war, dass sie es damals als einzige uns bekannte Software erlaubte, beliebige Sequenzen innerhalb eines Videos mit mehreren Nutzern gleichzeitig zu annotieren. Zudem verfügt sie über eine ausreichende Nutzer- und Rechteverwaltung und kann mit ihrer einfachen Handhabung mehrerer unterschiedlicher Annotationsspuren überzeugen. Der Funktionsumfang ist in der Zwischenzeit, auch vorangetrieben durch gezielte Investitionen der HRA, deutlich gewachsen und in die ebenfalls als Open Source-Software veröffentlichte Nachfolgeversion Pan.do/ra eingeflossen. So wurden beispielsweise eine LDAP-Authentifizierung integriert sowie mehrsprachige Audio- und Annotationsspuren implementiert. Zeitbasierte Annotationen können im .srt und JSON Format im- und exportiert werden. Das erlaubt es zum Beispiel, Audiotranskriptionen in Tools wie [Speedtrans](#)<sup>17</sup> anzufertigen und dann in Pan.do/ra zu importieren.

Für Annotationen stehen in Pan.do/ra unterschiedliche Spuren oder Layer zur Verfügung. So können beispielsweise Annotationen vom Typ ›Subtitle‹ direkt in den Film eingeblendet oder Annotationen vom Typ ›Location‹ auf einem Karteninterface dargestellt werden. Es wird festgehalten, welcher Nutzer oder welche Nutzerin wann welche Annotationen hinzugefügt hat und es ist möglich, die Anzeige von Annotationen nach Autor bzw. Autorin zu filtern – eine Funktion, die sich nicht nur während des Einsatzes im Unterricht bewährt hat.

In den Annotationen ist es grundsätzlich möglich Markup zu verwenden. Die bereits unterstützten HTML5-Tags erlauben Textformatierungen sowie das Einfügen von Hyperlinks und eröffnen auch die interessante Option, Bilder direkt in Annotationen einzubinden (vgl. die Beispiele im ersten Teil). Da sämtliche Annotationen mit URLs versehen sind, lassen sich bequem Szenen in einem oder mehreren Filmen verlinken. Ebenso können Verweise auf einzelne Annotationen von externen Websites gesetzt werden. Dank der Möglichkeit, auch Suchen über eine URL auszulösen, können Anwender beispielsweise sämtliche Stellen im Film aufrufen, die mit demselben Keyword annotiert wurden.<sup>18</sup>

Mit dem im Frontend für jede Annotation bereitgestellten *embed code* können darüber hinaus ohne großen Aufwand Filme oder ›Filmschnipsel‹ in andere Webseiten eingebunden werden. Am Exzellenzcluster wird diese Funktion beispielsweise im Open Access e-Journal<sup>19</sup> bereits eingesetzt. Ein derzeit noch eher ungewohnt anmutender Einsatzbereich ist das Einbetten von Videos in Annotationen. So erlaubt der HyperImage Nachfolger Yenda die Nutzung von HTML5-Tags an verschiedenen Stellen. Damit können in Bildannotationen nicht nur Texte und zusätzliche Bilder, sondern auch Videos aus Pan.do/ra eingebettet werden.

---

<sup>15</sup> [HRA Videoannotationsdatenbank](#).

<sup>16</sup> Vgl. [Demo auf Pan.do/ra](#).

<sup>17</sup> Eine Firefox Erweiterung, die für den Einsatz mit pad.ma entwickelt wurde.

<sup>18</sup> [Beispiel](#).

<sup>19</sup> Vgl. [Transcultural Studies](#).

Die Nutzung von *embed codes* ist nur eine Möglichkeit aus anderen Webseiten heraus auf Daten in Pan.do/ra zuzugreifen. Per API stellt Pan.do/ra eine ganze Reihe von Funktionen für die Anbindung an andere Webapplikationen zur Verfügung.

## 2.3 Einsatz am Exzellenzcluster Asien und Europa – Global Politics on Screen

Im akademischen Bereich wird die Software am Exzellenzcluster Asien und Europa seit 2010 in unterschiedlichen Projekten und Fachbereichen in Forschung und Lehre eingesetzt. Zusätzlich zu dem in diesem Artikel ausführlich beschriebenen Projekt sei noch das öffentlich zugängliche Global Politics on Screen-Projekt von Madeleine Herren-Oesch et al. erwähnt.<sup>20</sup> In diesem Projekt wurde von einer Gruppe von Studierenden und Wissenschaftlern des Historischen Seminars und des Exzellenzclusters Filmmaterial zu einer Mission des Völkerbunds untersucht. Erst während des Projekts wurde immer klarer, dass es sich dabei um eine frühe Version eines japanischen Propagandafilms handelt. Der von der South Manchurian Railway Company (SMR) zu Beginn der 1930er produzierte Film begleitet die sogenannte Lord Lytton Mission des Völkerbundes bei ihrer Untersuchung des Einmarsches der japanischen Streitkräfte in die Mandschurei. Das dabei entstandene Filmmaterial war für die Vorführung in Botschaftskreisen in Europa gedacht. Der Stummfilm wurde für die Vorführung mit Texttafeln in unterschiedlichen Landessprachen aufbereitet.

Die im Rahmen des Heidelberger Projekts vorgenommene Untersuchung des Films konzentriert sich auf eine vermutlich unveröffentlichte Version mit englischen Texttafeln, die eine frühe Fassung des Filmmaterials mit dem Charakter eines Rohschnitts zeigt. Bei der ersten Durchsicht des Filmmaterials waren noch vielen Fragen offen. Es war lediglich klar, dass es sich um ein filmisches Dokument zur Lord Lytton Mission handelte. Sowohl Auftraggeber als auch das anvisierte Zielpublikum waren zunächst nicht bekannt. Mit Hilfe von Pan.do/ra konnte die HRA dem Projekt ein Werkzeug an die Hand geben, diesen und anderen offenen Fragen nachzugehen.

In der ersten Arbeitsphase wurde Pan.do/ra genutzt, um subjektive Auffälligkeiten und Besonderheiten im Film zu finden, zu markieren und zu kommentieren. In den nächsten Phasen hat sich dann das Team Schritt für Schritt dem Film immer stärker und systematisch angenähert. So wurden sämtliche englische Texttafeln transkribiert und dann nach Querverweisen zu Dokumenten in Archiven gesucht. Im [Japan Center for Asian Historical Records](#) wurden mehrere Dokumente mit unterschiedlichen Sprachversionen und in unterschiedlicher Reihenfolge der Texttafeln gefunden.

Da es sich um einen Stummfilm handelt, wurde sämtlichem geschriebenen Text im Film schnell dieselbe Aufmerksamkeit zu Teil wie den Texttafeln. Im Film gezeigte Schriftzüge auf Bannern, Hinweisschildern oder ähnlichem wurden transkribiert und wo nötig ins Englische übersetzt. Dabei hat es sich als sehr vorteilhaft herausgestellt, dass Übersetzungen sofort bei

---

<sup>20</sup> [Global Politics on Screen](#): A Japanese Film on the Lytton Commission in 1932.

Eingabe in die Datenbank für alle Projektmitglieder zur Verfügung standen. Die Banner, die an vielen Stellen im Film den Weg der Kommission begleiten, wirken im Film oft wie Untertitel, die sich je nach verwendeter Sprache an ein unterschiedliches Publikum wenden. Oft stellt sich die Frage, ob die Kommission, das Publikum vor Ort oder das Publikum vor der Leinwand angesprochen werden soll.

Durch das Transkribieren wurde den Projektteilnehmern, die sonst wenig mit Stummfilmen in Berührung gekommen waren, der Stellenwert von geschriebenem Text im Stummfilm deutlich vor Augen geführt. Hinweise, die man ohne die gezielte Verschriftlichung mit den heutigen Sehgewohnheiten vielleicht einfach übersehen hätte, wurden so relevant. Als Beispiel sei hier eine **Szene in einem Hafen** (vermutlich Dairen) angeführt. In einem Kameraschwenk wird gezeigt, wie ein Schiff entladen wird. Dabei ist beiläufig für ca. 3 Sekunden ein Schild, das den Hersteller des Krans identifiziert, zu sehen. Zeit genug, um für einen zeitgenössischen Betrachter aus europäischen Diplomatenkreisen zu bemerken, dass es sich um einen Kran der Demag AG aus Duisburg handelt. Der Rückschluss, dass Japan mit der Besetzung der Mandschurei auch für Sicherheit und Stabilität von Großkunden deutscher Unternehmen sorgt, bleibt den informierten Betrachtern überlassen.

Im weiteren Verlauf des Projekts wurden noch zusätzliche Annotationsebenen verwendet. Mit den Annotationen vom Typ »Location« konnten die gezeigten Schauplätze im Film nachgezeichnet werden und mit **Karten aus dem Kommissionsreport** verglichen werden (vgl. Abbildung 11).



Abb. 11: Gegenüberstellung eines Kartenausschnitts aus dem Kommissionsbericht und eines Screenshots der Location Annotationsspur in Pan.do/ra. Quelle: Route Maps: Itineraries of the Commission in the Far East (League of Nations (ed.): Appeal by the Chinese Government. Report of the Commission of Enquiry, Appendix, Maps No. 13 and 14, Geneva: League of Nations, 1932).

Zur thematischen Aufarbeitung wurden häufig wiederkehrende Argumentationsmuster und Themen des Films mit Keywords versehen. Eines der übergeordneten Ziele des Filmes ist es beispielsweise, den Einmarsch japanischer Truppen in die Mandschurei zu legitimieren. Ein Legitimationsgrund, der immer wieder aufgeführt wird, ist die Zerstörung wichtiger Eisenbahnbrücken durch Anschläge sogenannter »Soldier-Bandits«. Heute geht man davon aus, dass die Anschläge inszeniert waren, um einen Einmarsch vor der Weltöffentlichkeit zu rechtfertigen.

Ebenfalls mit Keywords versehen wurden die Auftritte bestimmter Kommissionsmitglieder im Film. Dabei wurden lediglich die Auftritte von Lord Lytton und des deutschen Abgesandte

Heinrich Schnee über die gesamte Länge des Films transkribiert. Das manuelle Suchen und Markieren hat sich jedoch als sehr zeitaufwändig und nur bedingt aussagekräftig herausgestellt. Um tatsächlich sinnvolle und quantitativ belegbare Aussagen darüber machen zu können, wer wann wie oft und mit wem zusammen gezeigt wird, wäre eine komplette Erfassung aller Gruppenmitglieder nötig. Der dazu nötige Aufwand stand im Rahmen des Projektbudgets aber in keinem vernünftigen Verhältnis zum möglichen Erkenntnisgewinn. Hier wäre eine (semi-)automatisierte Lösung wünschenswert.

## 2.4 Recherchen – Verfügbare Videoannotationssysteme

Im November 2013 und April 2014 wurde eine größere, ergebnisoffene Recherche der aktuell verfügbaren Anwendungen im Bereich Videodatenbanken und -annotation durchgeführt. Ziel war es, einen genaueren Überblick über den Stand der Entwicklungen auf diesem Gebiet zu erhalten. Darüber hinaus sollte auch geprüft werden, ob neuere oder eventuell für den Einsatz am Exzellenzcluster besser geeignete Lösungen entwickelt wurden, um gegebenenfalls darauf umzusteigen. Die Recherche gestaltete sich schwieriger als erwartet. Die Informationen waren weit verstreut und nur wenige Zusammenstellungen von Tools für Videoanalyse oder Videoannotation bis dato veröffentlicht (vgl. Liste von Video-Annotationssystemen).

Vielversprechend schienen die Einträge in der »Digital Research Tools« (DIRT) Datenbank des Project Bamboo, doch waren sie leider nur bedingt hilfreich. Zum einen waren viele Einträge veraltet und die Liste sehr unvollständig. Auch war es zwar möglich, die Einträge nach Tags zu browsen, doch schon eine kombinierte Suche beispielsweise nach »video« und »annotation« war nicht möglich. Die 24 Einträge beim Tag »video« waren sehr divers und enthielten selbst Bilddatenbanken (Flickr) oder Screen recording Tools (Screenr). Nur 6 Einträge in »video« waren auch mit »annotation« getagged (Advene, Annotator's Workbench, ELAN, Project Pad, Vertov, VideoANT). Darüber hinaus waren die Einträge zwar annotiert, die Inhalte aber meist den Selbstdarstellungen der Webseiten entnommen.<sup>21</sup>

Um die Verlässlichkeit der Informationen und deren Potential als Entscheidungshilfe auf der Suche nach einer geeigneten Software war es auf der »Video«-Seite der »Annotations at Harvard« besser bestellt. Die dort veröffentlichte Liste war zwar mit 18 Einträgen auch überschaubar und die Inhalte stammen aus dem Jahr 2010, aber die Zusammenstellung war inhaltlich deutlich fokussierter und die Anmerkungen oft nicht einfach nur wörtlich von der Website kopiert.<sup>22</sup>

Interessant war auch festzustellen, dass das Thema Annotation auch bei Online Guides ausgelassen wurde, so beispielsweise zu beobachten beim ansonsten vorbildlich zusammengestellten »Activists Guide to Archiving Video«.<sup>23</sup>

---

<sup>21</sup> Im Internet Archive kann noch ein [Snapshot vom 2. Oktober 2013](#) eingesehen werden.

<sup>22</sup> [Annotations at Harvard – Video Annotations](#). Harvard Annotation Project, 2010.

<sup>23</sup> [Activists Guide to Archiving Video](#), Version 1.0 (Juli 2013).



Auflistungen in den Buchpublikationen waren hingegen inhaltlich sehr eng gefasst und beschränkten sich auf wenige Tools aus den Bereichen qualitativer Datenanalyse (QDA) bzw. Broadcast Video. Auch lag der Zeitpunkt der Zusammenstellung meist mehrere Jahre vor dem Publikationsdatum. Die folgende Liste zeigt eine Auswahl veröffentlichter Zusammenstellungen von Video-Annotationssystemen:

- Jost et al. 2013: Christofer Jost et al.: *Computergestützte Analyse von audiovisuellen Medienprodukten* (= Qualitative Sozialforschung 22). Springer Fachmedien: Wiesbaden 2013. [\[online\]](#)
- Linked TV: Evlampios Apostolidis et al.: *State of The Art and Requirements Analysis for Hypervideo*. LinkedTV - Television Linked To The Web. 30. September 2012. [\[online\]](#)
- Bamboo: Bamboo DiRT. Digital Research Tools. [2013]. [\[online\]](#)
- Dasiopoulou et al. 2011: Stamatia Dasiopoulou et al.: »A Survey of Semantic Image and Video Annotation Tools.« In: *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*. Edited by Georgios Paliouras / Constantine D. Spyropoulos / George Tsatsaronis (= Lecture Notes in Computer Science 6050.) Springer: Berlin and Heidelberg 2011, S. 196–239. [\[online\]](#)
- Harvard: Annotations at Harvard – Video. Video Annotation Tools. [2010]. [\[online\]](#)
- UCLA-KB: UCLA Knowledgebase – Video annotation tools. [2009]. [\[online\]](#)

Neben der Verstreutheit der Informationen wurde während der Recherchen sehr schnell deutlich, dass Analyse und Annotation von Videos in sehr vielen Bereichen und mit sehr unterschiedlichen Zielsetzungen eingesetzt werden. So gibt es Anwendungen<sup>24</sup>, die besonders für linguistische Analysen konzipiert sind, beispielsweise das vom Max Planck Institute for Psycholinguistics in Nijmegen entwickelte ELAN. Dann gibt es Anwendungen, die sich auf die in großen Fernsehanstalten anfallenden bzw. dort archivierten Materialien spezialisieren, wie beispielsweise die Entwicklungen innerhalb des FP7 Projektes LinkedTV. Wieder andere Systeme konzentrieren sich auf den Bereich E-Learning oder sind als Streaming-Plattformen oder Archivportale konzipiert. Darüber hinaus gibt es Lösungen die besonders die Musikanalyse von Filmen unterstützen, wie das bereits erwähnte, in Basel entwickelte trAVis<sup>25</sup>, oder sich auf Performances spezialisieren, wie der innerhalb des European Collected Library of Performing Arts (ECLAP) entwickelte MyStoryPlayer. Noch andere Lösungen zielen auf die Erstellung interaktiver Narrationen ab, sind Teil von Softwarepaketen zur qualitativen Datenanalyse, werden in der Lehre eingesetzt oder dienen der Erstellung von Untertiteln. Die folgende Auflistung zeigt Beispiele für Videoanalyse- und -annotationslösungen in verschiedenen Anwendungsbereichen.<sup>26</sup>

- Linguistische Analyse: ELAN, ANVIL, EXMARaLDA
- Medienanstalten / Archive: LinkedTV, MediaGlobe
- E-Learning / Vortragsportale: Catool, AAV/OpenCast Matterhorn, Alexander Street
- Streaming Plattformen / Archivportale: Kaltura, AV Portal

---

<sup>24</sup> Die Links zu den hier aufgeführten Projekten und Softwarelösungen finden sich zusammen mit weiterführenden Informationen in einem vom Autor zusammengestellten frei verfügbaren Google Docs-Dokument [Overview: Video Annotation Tools](#).

<sup>25</sup>Vgl. Abschnitt 1.2, Fußnote 4.

<sup>26</sup> Vgl. [Overview: Video Annotation Tools](#).

- Musik / Performance: trAVis, MyStoryPlayer
- Video: Mediathread, Pan.do/ra
- Interactive storytelling: Klynt, HyperImage
- Subtitling: Speedtrans, DotSub
- Service based analysis: Semex, Videana

Alle Tools bieten im weiteren Sinne Annotationen an. Viele fanden dabei auch elegante - aber eben doch bereichsspezifische Lösungen. Keins deckt alle Bereiche ab.

Es schien daher sinnvoll, die Ergebnisse der eigenen Recherchen zu systematisieren und einer breiteren Community zur Einsicht und Ergänzung zur Verfügung zu stellen. Darüber hinaus sollten in der Zusammenstellung auch inhaltliche Kriterien erfasst werden, die es Lesern erleichtern, auf der Suche nach einer für sie geeigneten Software eigene Entscheidungen zu treffen. Dazu wurde auch erfasst, ob ein Tool noch aktiv entwickelt wird, ob der Sourcecode veröffentlicht wurde und ob es frei benutzbar ist. Weiterhin war für uns relevant, ob es sich um ein online Tool handelt, es kollaborativ nutzbar ist und ob es potentiell in andere Datenbanken integriert werden kann. Es wurden auch (sehr grobe) Kategorien vergeben, um die jeweilige »Zielgruppe« festzuhalten, beispielsweise »filmstudies«, »linguistic« oder »video platform«. Schließlich wurde auch erfasst, welche Personen oder Institutionen die Entwickler sind und ein kurzer Beschreibungstext aus den online Informationen entnommen. In den etwa fünf Monaten unterschiedlich intensiver Recherchen kamen so knapp 60 Einträge zusammen. Das Dokument ist auf Google Docs veröffentlicht und kann von jedem Nutzer eingesehen, ergänzt und bearbeitet werden.<sup>27</sup>

## 2.5 Anforderungen an ein zukünftiges Videoannotationssystem

Bei den Recherchen kamen auch Ressourcen zum Vorschein, in denen Anforderungen für Videoannotationssysteme formuliert werden. Hier sind insbesondere die im Rahmen des Harvard Annotation Projektes zusammengestellten »Requirements of a Video Annotation System«<sup>28</sup> zu erwähnen (Abbildung 11), die auch den Interessen und Zielen am Heidelberger Exzellenzcluster Asien und Europa sehr nahe kommen, aber in der Praxis in dieser Art noch in keiner Software verfügbar sind.

---

<sup>27</sup> [Link zum Dokument.](#)

<sup>28</sup> [Requirements of a Video Annotation System](#), Harvard Annotation Project, 2010.



Abb. 12: **Requirements of a Video Annotation System**. Quelle: Harvard Annotation Project, 2010. Mit freundlicher Genehmigung.

Ein solches Annotationssystem sollte lokal oder online verfügbar sein und auf offene Datenstandards setzen. Dabei sollte das Interface erlauben, Videos insgesamt oder in Teilen zu annotieren und Bilder oder andere Videos neben der Visualisierung der Soundspur als Overlays möglich sowie gesprochener Text automatisch als ›baseline annotation‹ verfügbar sein. Außerdem sollten Marker im Bild gesetzt werden können und ihr Aussehen anpassbar sein, wobei einfache geometrische Flächen, Linien und Polygone ebenso erwünscht wären wie kurze, in Farbe, Füllung, Strichstärke usw. anpassbare und möglichst mit der Zoomstufe mitskalierende Text-Tags.<sup>29</sup>

Viele dieser Anforderungen stimmen mit den Wünschen der Nutzer am Exzellenzcluster Asien und Europa überein. Während die Videoannotationen durch die Funktionalitäten in Pan.do/ra gut abgedeckt sind, ist besonders spannend, wie nah viele der weiterführenden Punkte den bereits im Moment verfügbaren Features der ebenfalls im Cluster eingesetzten virtuellen Forschungs- und Publikationsumgebung HyperImage und dessen durch semantische Annotationen erweiterten Nachfolgers Yenda kommen.<sup>30</sup> Zwar sind diese Tools im Moment noch auf Still Images beschränkt, aber die Einbeziehung von Bewegtbildern in die vielfältigen Annotationsmöglichkeiten ist möglich und könnte gemeinsam mit interessierten Kooperationspartnern entwickelt werden.

Ein Bericht über den Weg der Technischen Informationsbibliothek (TIB) von einer Bedarfsanalyse hin zum im Frühjahr 2014 gemeinsam mit dem Hasso-Plattner-Institut für Softwaresystemtechnik an der Universität Potsdam (HPI) eröffneten AV-Portal, einer webbasierten Plattform für audiovisuelle Medien, wurde von Margret Plank auf der LIBER Konferenz 2012 vorgestellt und auf dem LIBER-Blog veröffentlicht.<sup>31</sup>

<sup>29</sup> Vgl. **Requirements of a Video Annotation System**, Harvard Annotation Project, 2010.

<sup>30</sup> Offizielle Website des Open-Source-Projekts **HyperImage – Eine virtuelle Forschungs- und Publikationsumgebung für den Bilddiskurs**, Yenda – **Picture Knowledge**. Open-Source semantische virtuelle Forschungsumgebung der bitGilde IT Solutions UG.

<sup>31</sup> Plank 2012 (Der Blogbeitrag war leider im Oktober 2014 online nicht mehr erreichbar.) Vgl. auch Plank / Weichert 2014.

Das AV-Portal ist allerdings nicht für den Einsatz im Forschungsalltag von Geisteswissenschaftlern konzipiert, sondern »optimiert den Zugang zu und die Nutzung von wissenschaftlichen Filmen aus Technik und Naturwissenschaften, wie z.B. Computeranimationen, Vorlesungs- und Konferenzaufzeichnungen.«<sup>32</sup> Die in diesen Service integrierten Features wie automatische Sprach- und Texterkennung sowie Anreicherung mit semantischen Informationen sind jedoch auch für die Nutzung über die Naturwissenschaften hinaus spannend und überaus vielversprechend.

## 2.6 Einbindung von Videoannotationen in Tamboti

Um die in der Datenbank enthaltenen Annotationen der Wissenschaftler über das zentrale Suchinterface in Tamboti verfügbar zu machen wurde als »Proof-of-concept« bereits eine Anbindung der Metadaten implementiert. Dabei werden die Metadaten aus der entsprechenden .srt-Spur als <relatedItem> in den MODS Eintrag des Videos integriert. Der Inhalt der Annotation wird im Subelement <note> und die Zeitspanne in <part> ausgegeben. Dies ist eine Lösung, die für Objektmetadaten inklusive Untertiteln noch sinnvoll erscheint, bei gleichzeitigem mapping von Nutzerannotationen nach MODS aber zu einer Überfrachtung des Eintrags führt. Hier ist in Zukunft eine Trennung von Objektmetadaten (MODS) und Annotation (OADM)<sup>33</sup> angestrebt.

```
<relatedItem type="constituent">
  <titleInfo>
    <title supplied="yes">constituent</title>
  </titleInfo>
  <subject>
    <topic>[TOPIC1, TOPIC2, ...]</topic>
  </subject>
  <place>
    <placeTerm>[NAME-OF-LOCATION]</placeTerm>
  </place>
  <note type="transcription">[CONTENT-OF-TRANSCRIPTION]</note>
  <part>
    <physicalDescription>
      <extent>
        <start>[START-TIME]</start>
        <end>[END-TIME]</end>
        <total>[TOTAL-TIME]</total>
      </extent>
    </physicalDescription>
  </part>
</relatedItem>
```

Abb. 13: Schematische Darstellung der Notation von Annotationen in MODS. © Arnold, 2015, CC-Lizenz 0.

---

<sup>32</sup> Erläuterungen zum AV-Portal der TIB Hannover.

<sup>33</sup> Open Annotation Data Model: Das Datenmodell wurde im Rahmen der Open Annotation Collaboration entwickelt. Zum Einsatz von OADM bei der Annotation von Multimedia Objekten und für einen Vergleich mit anderen Annotationsmodellen vgl. Haslhofer et al. 2014, passim.

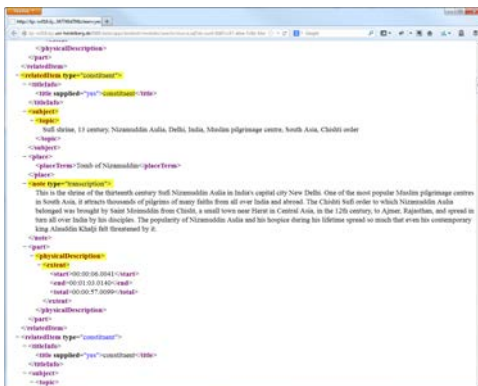


Abb. 14: Beispiel einer Annotation in MODS. Screenshot, erstellt am 30.10.2014, © HRA, CC-Lizenz 0.

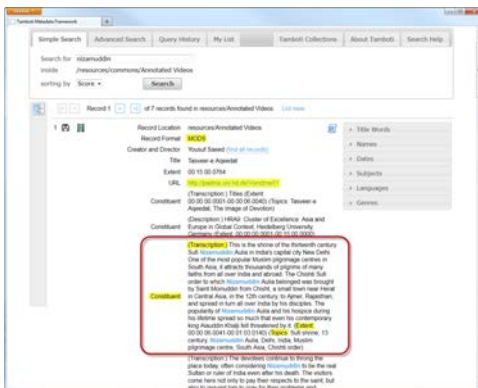


Abb. 15: Anzeige von Video Annotationen in einem Tamboti Datensatz. Screenshot, erstellt am 30.10.2014, © HRA, CC-Lizenz 0.

Dank dieser Erweiterung des Metadatenframeworks ist es möglich, in Tamboti parallel in Metadaten und Annotationen von Videos, Bildern, Texten und bibliographischen Datensätzen zu suchen.



Abb. 16: Eine Tamboti-Suche nach »Krishna« ergibt Treffer in verschiedenen Medien und Datenformaten: Bildmetadaten (VRA Core 4), bibliographische Daten (MODS), Videoannotationen (related items in MODS), Textannotationen (TEI). Screenshot, erstellt am 30.10.2014, © Arnold, [CC-Lizenz 0](#).

## 2.7 Ausblick

Es gibt noch eine ganze Reihe von Features, die für unterschiedliche Anwendungsszenarien in Forschung und Lehre wünschenswert erscheinen.<sup>34</sup> Diese sind natürlich nicht auf den Einsatz in der (digitalen) Kunstgeschichte oder die Geschichtswissenschaft beschränkt, sondern auf viele Bereiche der geisteswissenschaftlichen Forschung anwendbar; ob das nun das (semi-)automatische Erkennen und Verfolgen von Gesichtern, Gegenständen bzw. Mustern im Film oder Standardisierungen bei den Metadaten und Annotationen betrifft.<sup>35</sup> Das Geschäftsmodell, das hinter Pan.do/ra steht, macht es sehr wahrscheinlich, dass diese und andere Features implementiert werden können, da die Kosten nicht von einer Institution oder einem Projekt alleine getragen werden müssen.

Was die Weiterentwicklung im Rahmen des Exzellenzclusters angeht, so stehen die Einbindung kontrollierter Vokabulare für Schlagworte und Personen sowie die dynamische Anbindung an das Metadatenframework Tamboti auf der Agenda. Ebenso soll beim Datenaustausch zwischen Videoobjekt-Metadaten (z.B. als MODS Repräsentation) und Nutzerannotationen (im Open Data Annotation Model) unterschieden werden, auch um eine Überfrachtung der Objektmetadaten zu vermeiden.

## 2.8 Nachtrag

Die Vorstellung des Projekts und der Austausch mit den Kollegen auf der DHd 2014 in Passau erwies sich als überaus fruchtbar. Fabian Cremer beschrieb das Projekt in seinem DHd-Blogbeitrag als »idealtypisch« und »nicht nur leichtgängiger sondern auch nachhaltiger«.<sup>36</sup> Anfang Juni 2014 organisierte die Heidelberg Research Architecture einen Round Table *Annotation von digitalen Medien* im Heidelberger Karl Jaspers Centre mit Teilnehmern verschiedener Fachdisziplinen aus Berlin, Darmstadt, Essen, Heidelberg und Rom.<sup>37</sup> Ein Folgeworkshop fand im Frühjahr 2015 unter Federführung des DARIAH-DE Cluster 6 *Research Annotations* statt.<sup>38</sup> Darüber hinaus wurde die im Artikel vorgestellte Zusammenstellung von Videoannotationstools in die im Rahmen des DARIAH-DE Cluster 6 erarbeitete Liste von Annotationstools eingearbeitet. Diese soll auch an das [DIRT-Verzeichnis](#) weitergegeben und damit einem noch breiteren Nutzerkreis verfügbar gemacht werden.

---

<sup>34</sup> Für einen Einblick in den Bereich Lehre vgl. Lemon et al. 2013, passim.

<sup>35</sup> Alles Themen, die derzeit Gegenstand von Forschungsprojekten oder in Websites wie dem AV Portal bereits verfügbar sind. Vgl. Nixon / Troncy 2014, S. 11–15 und 18ff.

<sup>36</sup> Cremer 2014.

<sup>37</sup> Borek / Reiche 2014.

<sup>38</sup> Programm [\[online\]](#)

## Bibliographische Angaben

- Activists Guide to Archiving Video, Version 1.0 (July 2013). [\[online\]](#) [\[Nachweis im GBV\]](#)
- Annotations at Harvard – Video Annotations. Harvard Annotation Project 2010. [\[online\]](#)
- Matthias Arnold: Overview: Video Annotation Tools. [\[online\]](#)
- Luise Borek / Ruth Reiche: Round Table »Annotation von digitalen Medien«. In: DHd-Blog vom 29. Juli 2014. [\[online\]](#)
- Frank Bruder: Pop-Splits – Die besten Songs aller Zeiten und ihre Geschichte. Berlin 2004. [\[Nachweis im GBV\]](#)
- Fabian Cremer: Erfolg und Emergenz in den Digital Humanities – ein Tagungseindrucksausschnitt. In: DHd-Blog vom 29. März 2014. [\[online\]](#)
- Gabriel García Márquez: Ein sehr alter Mann mit riesengroßen Flügeln (Un señor muy viejo con unas alas enormes). In: Gabriel García Márquez: Laubsturm. Frankfurt/Main 2004. [\[Nachweis im GBV\]](#)
- Global Politics on Screen: A Japanese Film on the Lytton Commission in 1932. [\[online\]](#)
- Bernhard Haslhofer / Robert Sanderson / Rainer Simon / Herbert van de Sompel: Open Annotations on Multimedia Web Resources. In: Multimedia Tools and Applications 70 (2014) H. 2, S. 847–867. doi:10.1007/s11042-012-1098-9. [\[Nachweis im GBV\]](#)
- Henry Keazor / Thorsten Wübbena: Video thrills the Radio Star: Musikvideos: Geschichte, Themen, Analysen. 3. Auflage. Bielefeld 2011 (zuerst 2005). [\[Nachweis im GBV\]](#)
- Kunstgeschichte, Musikvideo und Bildwissenschaften. Eine Einführung: Come ride with me through the veins of history. Hg. von Henry Keazor. Berlin 2014 [\[online\]](#).
- Narelle Lemon / Meg Colasante / Karen Corneille / Kathy Douglas: Video Annotation for Collaborative Connections to Learning: Case Studies from an Australian Higher Education Context. In: Increasing Student Engagement and Retention Using Multimedia Technologies: Video Annotation, Multimedia Applications, Videoconferencing and Transmedia Storytelling. Hg. von Laura A. Wankel / Patrick Blessinger. 2013, S. 181–214 (= Cutting-Edge Technologies in Higher Education, 6 Part F.). [\[Nachweis im GBV\]](#)
- Bernard Marcadé / Dan Cameron: Pierre et Gilles – The Complete Works, 1976–1996. Köln 1997. [\[Nachweis im GBV\]](#)
- Lyndon Nixon / Raphael Troncy: Survey of Semantic Media Annotation Tools – towards New Media Applications with Linked Media. Paper presented at the Linked Media Workshop, European Semantic Web Conference am 25. Mai 2014. [\[online\]](#)
- Margret Plank: A Portal for Scientific Audiovisual Media: Analysing User Needs. In: LIBER, 3. Oktober 2012. [\[online\]](#)
- Margret Plank / Steffen Weichert: Don't Forget the Users – Developing a Portal for Audiovisual Media with a User-Centred Approach. In: IFLA WLIC 2014. Libraries, Citizens, Societies: Confluence for Knowledge (Session 139 - Audiovisual and Multimedia with Information Technology). Lyon. 16.–22. August 2014. [\[online\]](#)
- Rewind – Play – Fast Forward: The Past, Present and Future of the Music Video. Hg. von Henry Keazor / Thorsten Wübbena. Bielefeld 2010. [\[online\]](#)

## Abbildungslegenden und -nachweise

- Abb. 1: Startseite von Pan.do/ra. Screenshot, erstellt am 17.06.2015, © HRA, [CC-Lizenz 0](#).
- Abb. 2: (a-d) Stills aus dem Clip *Losing My Religion* – Sänger Michael Stipe mit Flügeln und weitere Szenen (beispielhaft) mit dem Motiv der Engelsflügel in der Editor-Ansicht von Pan.do/ra, jeweils mit dem Tag »Engelsflügel« versehen (vgl. im rechten Bereich den Punkt »Keywords«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.
- Abb. 3: Link aus Pan.do/ra zum Bilddatenbanksystem ConedaKOR. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.
- Abb. 4: Entitätsansicht in ConedaKOR: Michelangelo Merisi da Caravaggio, *Verzückung des heiligen Franziskus*, um 1595, Hartford, Wadsworth Atheneum Museum of Art. Quelle: KGI (Frankfurt/Main), [Web Gallery of Art](#), erstellt am 12.10.2014.
- Abb. 5: Interner Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Die sieben Werke der Barmherzigkeit*, 1607, Neapel, Chiesa del Pio Monte della Misericordia. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.
- Abb. 6: Link bzw. Einbettung: Michelangelo Merisi da Caravaggio, *Der ungläubige Thomas*, 1603, Potsdam, Bildergalerie Schloss Sanssouci. Quelle: HRA, KGI (Frankfurt/Main), [Web Gallery of Art](#), Warner Bros., erstellt am 12.10.2014.
- Abb. 7: Einbettung: Pierre & Gilles, *St. Sebastian*, 1987. Quelle: HRA, KGI (Frankfurt/Main), [Art Forum](#), Warner Bros., erstellt am 12.10.2014.

Abb. 8: Interner Bezug durch Tagging, verdeutlicht durch eine farbige Kennzeichnung auf der unteren Timeline (in diesem Fall »Arbeiter«). Quelle: HRA, Warner Bros., erstellt am 12.10.2014.

Abb. 9: Einbettung: Abbildungen des Tanzes von V. Nijinsky zu *Jeux*. Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., erstellt am 12.10.2014.

Abb. 10: Interner Link zur korrespondierenden Filmszene aus *The Cell*, 2000 (Director: Tarsem Singh). Quelle: HRA, KGI (Frankfurt/Main), Warner Bros., New Line Cinema (theatrical), [Warner Home Video](#) (DVD), erstellt am 12.10.2014.

Abb. 11: Gegenüberstellung eines Kartenausschnitts aus dem Kommissionsbericht und eines Screenshots der Location Annotationsspur in Pan.do/ra. Quelle: Route Maps: Itineraries of the Commission in the Far East (League of Nations (ed.): *Appeal by the Chinese Government. Report of the Commission of Enquiry, Appendix, Maps No. 13 and 14*, Geneva: League of Nations, 1932).

Abb. 12: [Requirements of a Video Annotation System](#). Quelle: Harvard Annotation Project, 2010. Mit freundlicher Genehmigung.

Abb. 13: Schematische Darstellung der Notation von Annotationen in MODS. © Arnold, 2015, [CC-Lizenz 0](#).

Abb. 14: Beispiel einer Annotation in MODS. Screenshot, erstellt am 30.10.2014, © HRA, [CC-Lizenz 0](#).

Abb. 15: Anzeige von Video Annotationen in einem Tamboti Datensatz. Screenshot, erstellt am 30.10.2014, © HRA, [CC-Lizenz 0](#).

Abb. 16: Eine Tamboti-Suche nach »Krishna« ergibt Treffer in verschiedenen Medien und Datenformaten: Bildmetadaten (VRA Core 4), bibliographische Daten (MODS), Videoannotationen (related items in MODS), Textannotationen (TEI). Screenshot, erstellt am 30.10.2014, © Arnold, [CC-Lizenz 0](#).