

BILDER DER KUNSTGESCHICHTE IN DER GETTY ULAN UND IN WIKIPEDIA

Doron Goldfarb^a, Dieter Merkl^a

^a Institut für Softwaretechnik und Interaktive Systeme, Technische Universität Wien, Österreich, [vorname.nachname]@ec.tuwien.ac.at

KURZDARSTELLUNG: Diagramme zur Vermittlung von kunsthistorischen Entwicklungen besitzen eine lange Tradition. Digitale Informationen über Kunstsammlungen und deren Kontext ermöglichen die algorithmische Erzeugung solcher Bilder sowie die Analyse der zugrundeliegenden Datensätze mit quantitativen Methoden. Diese Arbeit widmet sich diesem Thema anhand des visuellen und analytischen Vergleichs von kunsthistorischen Personennetzwerken aus der Getty Union List of Artist Names und entsprechenden Informationen aus Wikipedia. Auf diese Weise werden in den Daten vorhandene, kulturelle Präferenzen hervorgehoben und Unterschiede zwischen professionell kuratierten und von der Internet-Community gepflegten Daten sichtbar gemacht.

1. EINFÜHRUNG

Schon seit den Ursprüngen der Kunstgeschichte als wissenschaftliche Disziplin im 19. Jhd. gibt es Bestrebungen, historische Entwicklungen in der bildenden Kunst zu Zwecken des Überblicks und der Vermittlung in visuell abstrahierter Form darzustellen. Ob in der Funktion des Grundrisses des Kunstmuseums als visualisierte Kunstgeschichte, wie von Fischer in 1 am Beispiel der kaiserlichen Gemäldegalerie in Wien beschrieben, oder als veranschaulichendes Diagramm in kunsthistorischer Literatur, wie von Schmidt-Burkhardt in 2 ausführlich untersucht: Viele dieser Versuche spiegeln die Interessen und kulturellen Prägungen der jeweiligen AutorInnen und Institutionen, sowie die der intendierten RezipientInnen wieder. Sie folgen somit oft einer zugrundeliegenden, impliziten oder expliziten Agenda, bestimmte Aspekte der Kunstgeschichte hervorzuheben.

Als Beispiel für diesen individuellen oder institutionellen „Bias“ sei das berühmte, von dem Museumsdirektor und -kurator Alfred H. Barr, Jr. gestaltete Diagramm der Stilentwicklungen der modernen Kunst zur MoMA Ausstellung „Cubism and Abstract Art“ von 1936 genannt. Es diente der Vermittlung der in den USA damals noch weitgehend unbekanntes modernen Kunst an ein breites Publikum. Laut Schmidt-Burkhardt wurde es aufgrund der Auswahl der darin erwähnten Stile durchaus kontrovers diskutiert 2. Das Diagramm spiegelt die persönlichen Forschungsinteressen Barrs

wieder, repräsentiert aufgrund dessen Funktion als Museumsdirektor aber auch die Agenda der Institution.

Während die historischen Vorhaben im Wesentlichen auf dem Fachwissen Einzelner beruhen, lädt die zunehmend freie Verfügbarkeit von Kulturerbe-Informationen in digitaler Form dazu ein, solche Zugänge auf quantitativer Basis zu verfolgen. Gerade in Zeiten der zunehmenden Digitalisierung des Kulturbetriebes können automatisch generierte visuelle Kartierungen von Sammlungsbeständen dazu dienen, virtuellen BesucherInnen Überblick und Orientierung in den wachsenden Beständen des digitalen kulturellen Erbes zu ermöglichen. Wie beispielsweise von Glinka und Dörk in 3 erörtert, sind diese schließlich online nicht nur in Form kuratierter Narrative, sondern auch als „Raw Data“ verfügbar.

Der datenbasierte Zugang verspricht hier potentiell mehr Objektivität, da eine auf Algorithmen basierende Transformation von kunsthistorischen Informationen in visuelle Darstellungen auf den ersten Blick mehr formalen bzw. technischen Gesetzmäßigkeiten folgt. Sie blendet somit den Filter „Mensch“ in dieser Hinsicht aus, bzw. überlässt sie diese Rolle den RezipientInnen selbst. Die zugrundeliegenden Datensammlungen stammen in der Regel aber von einzelnen Institutionen. Es stellt sich daher die Frage, ob anstelle des individuellen Bias nun eine institutionelle Variante davon in den

Vordergrund tritt. In den verfügbaren Daten „versteckt“, tritt sie nur indirekt über die auf der verfügbaren Datenbasis realisierten Anwendungen, beispielsweise in der Ordnung von Suchergebnissen, zutage. Auch hier können Visualisierungen aber hilfreich sein, indem sie dafür genutzt werden, genau solche institutionellen Präferenzen sichtbar zu machen.

Ein im Bereich der Kunstgeschichte relativ neuer Aspekt, der in den letzten Jahren vor dem Hintergrund des zunehmenden Datenaustausches zwischen einzelnen Kulturinstitutionen, der sich etablierenden „Digitalen Kunstgeschichte“, aber auch der generellen Vision des „Semantic Web“ als möglichst allumfassende, parallele und einheitliche Repräsentation von Daten verschiedenster Provenienz zu sehen ist, sind Normdaten. Geboten wird oft mehr als eine bloße Liste von eindeutigen IDs und Varianten der entsprechenden Bezeichner, so werden beispielsweise für Personen oft zusätzliche biographische Fakten mit aufgenommen. Normdaten stellen demnach nicht nur ein wertvolles und notwendiges Mittel für die Datenintegration dar, sondern repräsentieren gleichzeitig auch eine gemeinsame domänenspezifische Wissensbasis, die auf einem Querschnitt der Inhalte der beitragenden Institutionen basiert. Durch ihre überinstitutionelle Natur werden potentiell auch einzelne Sammlungs-Schwerpunkte gegenseitig ausgeglichen, dies allerdings nur hinsichtlich des in den Normdaten repräsentierten Kontextes der jeweiligen Bestände.

Ein Beispiel für eine kunsthistorisch relevante Normdatei ist die Getty Union List of Artist Names (ULAN), die seit 2015 frei als Linked Open Data verfügbar ist. Als Thesaurus organisiert, bietet dieses kontrollierte Vokabular Informationen über ca. 193.823 Personen (Stand Juli 2015) mit Bezug zur Kunstgeschichte, wobei für einen Großteil der Einträge Attribute wie Lebensdaten, Beruf, Nationalität und Geschlecht abrufbar sind. Darüber hinaus ist eine Untermenge von 21.942 Personen über 50.076 eindeutige, als Familie, Lehrer-Schüler und andere Arten typisierte Beziehungen miteinander verbunden und stellt somit ein historisches Personennetzwerk dar.

Besonders die identifizierten Personenbeziehungen bieten sich an, dahingehend untersucht zu werden, ob sich aus einem Blick auf ihre Gesamtheit ein „Bild der Kunstgeschichte“, im Sinne der identifizierten historischen Vorbilder, ergibt. Dies allerdings vor dem Hintergrund des Wandels der Bedeutung von Personenbeziehungen in der Kunstgeschichte, die sich von einer Künstlergeschichte à la Vasari zu einer Vielfalt an Sichtweisen zwischen den Polen des historischen und sozialen Kontextes der KünstlerInnen und ihrer Kunst und eher formalistischen, auf Stilentwicklung fokussierten Sichtweisen entwickelt hat. Die Analyse und Visualisierung der Personenbeziehungen aus der Getty ULAN stellt somit nur eine von vielen verfolgungswerten Möglichkeiten dar, so wurden beispielsweise Geburts- und Sterbeorte von ULAN-Personen von Schich et al. in 4 dazu genutzt, globale Trajektorien der kulturellen Migration über die Jahrhunderte aufzuzeigen.

Ein weiterer Aspekt, der die ULAN gerade in ihrer Funktion als kontrolliertes Vokabular zu einer interessanten Datenbasis macht, ist ihre Vernetzung mit anderen Datenbeständen. Seit mehreren Jahren ist sie in das „Virtual International Authority File“ (VIAF) integriert, eine Initiative, die die Zusammenführung von Einträgen verschiedener internationaler Normdatenbestände zum Ziel hat. Auf diese Weise wird es möglich eine gemeinsame Teilmenge von Personen zu identifizieren, die sowohl in der ULAN als auch in anderen über VIAF verknüpften Datensammlungen vorkommen, und diese auf mögliche Unterschiede zwischen den verschiedenen Beständen hin zu untersuchen.

Besonders vielversprechend ist der Vergleich der ULAN als professionell kuratierte Datenbank mit Quellen, die mehr einen „Crowd-Sourcing“ Hintergrund haben und ebenfalls über assoziative Verknüpfungen zwischen den vorkommenden Personen verfügen, wie es z.B. bei Wikipedia in Form von Personenbiographien und der sie verbindenden Hyperlinks der Fall ist. Von Klein und Kyrios wurde in 5 die Integration von Wikipedia in VIAF beschrieben, die auf diese Weise erzielten Konkordanzen mit der ULAN sind mittlerweile auch über das 2013 etablierte Datenportal Wikidata, siehe Vrandečić und Krötzsch in 6, verfügbar. Dieses Schwesterprojekt zu Wikipedia kann als

Knotenpunkt zwischen der ULAN und den verschiedenen Sprachversionen der freien Enzyklopädie genutzt werden. Derzeit (Stand März 2016) befinden sich dort in 55.122 Personeneinträgen eindeutige Verknüpfungen mit Personen aus der ULAN, wobei für 47.463 dieser Einträge jeweils auf mindestens eine Biographie in einer Wikipedia-Sprachversion verwiesen wird. Die Hyperlinks zwischen diesen Artikeln können wiederum über die Plattform DBpedia, siehe Auer et al. in 7, bezogen werden, deren Überlappung mit der ULAN hinsichtlich der dort verfügbaren Personenattribute bereits von Clough, Ireson und Marlow in 8 untersucht wurde.

Auf diese Weise ist es möglich, die Struktur der ULAN Personenbeziehungen mit jener der Hyperlinks zwischen entsprechenden Personenbiographien in verschiedenen Wikipedia-Sprachversionen zu vergleichen, wobei die Personenattribute der ULAN als „Variablen“ bzw., wie von Lev Manovich in 9 ausgedrückt, „Features“ dienen, anhand derer allfällige Unterschiede auf quantitativer Ebene gezeigt werden können.

2. ULAN-PERSONENATTRIBUTE

Das Attribut „Nationalität“ verfügt in der ULAN über 303 verschiedene Werte und fasst den zugrundeliegenden Begriff recht weit. Personen, auch historische, sind größtenteils gegenwärtigen Staaten („French“, „American“, „Austrian“), aber in geringerem Ausmaß auch vergangenen Entitäten („Holy Roman Imperial“) oder ethnischen Gruppen („Native American“) zugeordnet. In der Tat ist die Inklusion von traditionell aus der Kunstgeschichte eher ausgegrenzten Personengruppen ein erklärtes Ziel der ULAN. Auch die Detailliertheit der Angaben ist unterschiedlich, so gibt es neben historischen Personen mit der Zuweisung „Italian“ z.B. auch solche mit dem Attribut „Venetian“. Grundsätzlich können einer Person mehrere Werte zugewiesen werden, es existiert jedoch immer ein bevorzugter („preferred“) Wert, diese Arbeit zieht nur diese in Betracht. Die Verteilung der bevorzugten Nationalitäten ist höchst ungleich, die größten 26 decken bereits 87,48% aller Personen ab. Die Tatsache, dass knapp unter einem Drittel aller Einträge die Nationalität „American“ oder „British“ besitzt, unterstreicht die (nicht ausschließliche) Herkunft der ULAN aus dem englischsprachigen Raum, gemeinsam mit Franzosen,

Italienern und Deutschen repräsentieren diese Personen 61,31% aller Datensätze.

Abb. 1 zeigt die Anzahlen der Top-26 Nationalitäten auf einer logarithmischen Skala, aufgeteilt zwischen den 171.881 Datensätzen ohne Personenbeziehungen und jenen 21.942, für die solche existieren. 26 deshalb, weil für Darstellungszwecke möglichst vielen Nationalitäten jeweils eigene Farbcodes zugewiesen werden sollen und das „Farbalphabet“ von Green-Armytage aus 10 nahe der größtmöglichen Zahl an visuell unterscheidbaren Farben liegt. Alle anderen Werte werden unter dem Bezeichner „other“ codiert. Die Abbildung dient auch als Legende für die weiteren Grafiken in dieser Arbeit.

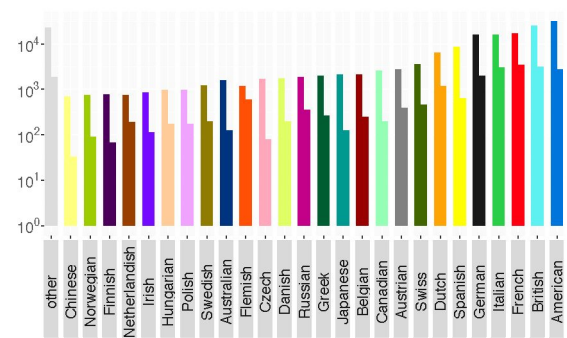


Abb. 1: Top-26 ULAN-Nationalitäten

Das Verhältnis von ca. einer Größenordnung zwischen Personen mit und ohne Beziehungen ist auch bei den einzelnen Nationalitäten zu erkennen, wobei einige davon offensichtlich höhere Anteile an verbundenen Personen aufweisen, als andere. Ein Vergleich der beiden Verteilungen auf Basis eines Chi-Quadrat-Homogenitätstests zeigt, dass beispielsweise trotz ihrer hohen Anzahl signifikant weniger „Americans“ über Personenbeziehungen verfügen, als erwartet. Für die „klassischen Kunstnationen“ („French“, „Italian“, „Dutch“, „Flemish“, „Netherlandish“), aber interessanterweise auch für manch kunsthistorische „Peripherie“ („Native American“, „Brazilian“, in der Abbildung unter „other“) sind wiederum signifikant mehr Beziehungen aufgezeichnet, als erwartet.

Bei der Betrachtung der Verteilung der Geburtsjahre der in der ULAN befindlichen Personen zeigt sich, dass eine kleine Gruppe in der Antike verortet ist, der eine „künstlerlose“ Zeit bis weit über das Jahr 1000 hinaus folgt. Es ist bemerkenswert, dass sich diese Beobachtung bereits 1765 in Priestleys Beschreibung seines berühmten „Chart of Biography“ 11 findet und laut Schich et al. in 4 daher stammt, dass die

KünstlerInnen dieser Periode ihre Autorenschaft nur selten durch Signaturen festhielten. Ab dem Jahr 1200 wächst die Anzahl an Personen bis zum Ende des 19. Jahrhunderts geradezu exponentiell an, um danach sukzessive wieder abzunehmen. Letzteres suggeriert, dass es eine gewisse Zeit dauert, um Teil des kunsthistorischen Kanons zu werden, sofern man die ULAN als eine Abbildung dessen wahrnimmt.

Abb. 2 zeigt die Verteilung der auf halbe Jahrhunderte gerundeten Geburtsjahre für Datensätze ohne (oben) und mit (unten) Personenbeziehungen und es wird bei letzteren sichtbar, wie die dort höheren Anteile der oben erwähnten „klassischen Kunstnationen“ in ihren jeweiligen historischen Zeithorizont fallen. Dies erklärt auch den merklich höheren Anteil an Personen aus der Zeit der Renaissance und des Barock, wo z.B. die Epoche der klassischen italienischen Malerei und das „Dutch Golden Age“ deutlich erkennbar sind. Ein weiterer interessanter Aspekt ist, dass viele der 277 unter „other“ zusammengefassten „Nationalitäten“ eher jüngeren Ursprungs sind, wie deren plötzlicher Anstieg um das Jahr 1800 herum zeigt. Diese Beobachtung deckt sich einerseits mit der von Anderson in 12 für diesen Zeitraum beschriebenen Etablierung des Konzeptes „Nationalität“ als Umfassung von „Imagined Communities“, zeigt andererseits auch die Bestrebungen der jüngeren Zeit, die Kunstgeschichte als globalen Prozess zu betrachten.

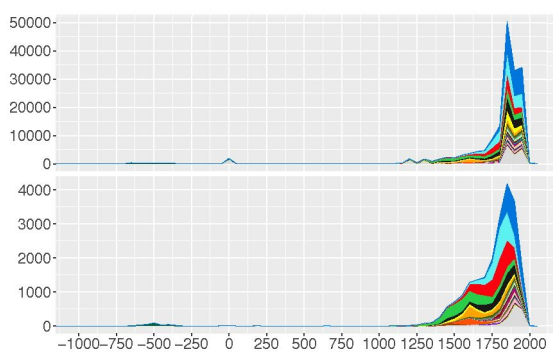


Abb. 2: ULAN Nationalitäten über die Zeit

Die Attribute Rolle und Geschlecht sind noch inhomogener verteilt, als es bereits für Nationalität und Geburtsjahr der Fall ist: 64,11% aller Datensätze haben die Rolle „Artist“, 23,38% „Architect“ und die restlichen 12,5% eine von 498 anderen Rollen. 78% Männer stehen 12% Frauen gegenüber (10% sind unbestimmt), dieses Verhältnis verschärft

sich zusätzlich für verbundene Personen, wo der Männeranteil auf 88,66% steigt und jener der Frauen auf 10,05% sinkt. Über die Zeit betrachtet nimmt der Frauenanteil jedoch besonders im 20. Jahrhundert deutlich zu und liegt in der Gegenwart bei über 30%. Diese Beobachtung reflektiert ebenfalls die zunehmende Sensibilisierung des Faches gegenüber der jahrhundertelangen Fokussierung auf männliche, westliche Künstler.

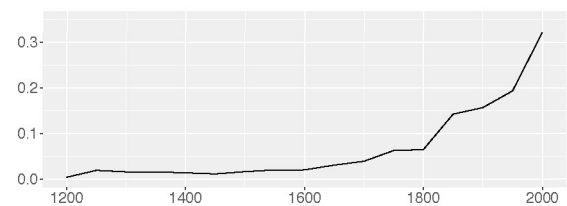


Abb. 3: Anteil der Frauen in der ULAN über die Jahrhunderte

3. ULAN-PERSONENBEZIEHUNGEN

Die 50.076 eindeutigen Personenbeziehungen in der ULAN bestehen zu 45,43% aus Familienbeziehungen, zu 39,41% aus Unterrichtsverhältnissen, zu 8,22% aus professionellen Verbindungen und zu 2,74% aus dem Bereich „Mäzenatentum“, der Rest teilt sich auf andere Typen auf. Alle Verbindungen sind immer bidirektional, zu jeder „Lehrer-von“-Verknüpfung existiert z.B. ein „Schüler-von“-Gegenstück. 10.444 der in der ULAN miteinander vernetzten Personen bilden über 29.956 Beziehungen einen großen zusammenhängenden Graphen („Giant connected component“, im weiteren als Hauptkomponente bezeichnet), während die restlichen 11.498 Personen mit Beziehungen auf 4.082 erheblich kleinere Komponenten verteilt sind. 2.780 davon haben nur zwei, 674 nur drei Mitglieder und der Rest umfasst zwischen vier und 39 Personen. Zu über 65% bestehen diese kleineren Gruppen aus ca. 58% aller Familienbeziehungen, während mehr als 50% der Verbindungen in der Hauptkomponente aus mehr als 80% der Unterrichtsbeziehungen bestehen.

Die großen Unterschiede zwischen den Netzwerkkomponenten werfen die Frage auf, ob die Hauptkomponente gegenüber den kleineren Komponenten und der Gruppe der gänzlich unverbundenen Personen besonders wichtige Persönlichkeiten der Kunstgeschichte enthält. Um dies quantitativ zu untersuchen, muss eine geeignete Referenz identifiziert werden, Quellen dafür finden sich in

kunsthistorischer Literatur: Von Jensen wurde in 13 ein Ranking von 50 KünstlerInnen des 19. Jahrhunderts beschrieben, das auf Bildzitat in 36 europäischen und amerikanischen Kunstgeschichte-Lehrbüchern des späten 20. Jahrhunderts beruht. Im selben Band findet sich in 14 eine von Elkins erstellte Liste der 50 meistzitierten KünstlerInnen auf Basis einer Schlagwortsuche in der „Bibliography of the History of Art“ (BHA), während eine von Dilly in 15 ableitete Reihung 147 Personen mit mindestens fünf Zitaten im Künstlerindex des „Repertoire d'Art et d'Archeologie“ (RAA) aus dem Jahr 1987 enthält. Kombiniert ergeben die drei Zugänge eine Liste von 178 Namen, von denen ausgegangen werden kann, dass es sich um besonders bekannte VertreterInnen der Kunstgeschichte handelt. Ein Blick in die einzelnen Listen bestätigt, dass die für die Allgemeinheit geläufigen Namen darin enthalten sind. Es ist anzumerken, dass das RAA in die BHA eingeflossen ist und letztere von der Getty Foundation verwaltet wird. Die ULAN enthält als kontrolliertes Vokabular der selben Institution naheliegender Weise die Namen aus diesen Bibliographien und ist daher nicht unabhängig von ihnen zu sehen.

Von den 178 Personen aus der kombinierten Liste befinden sich 136 (76,4%) in der Hauptkomponente des ULAN Netzwerkes, während 32 (17,98%) in kleineren Komponenten zu finden sind und die verbleibenden 10 in der ULAN über keine Personenbeziehungen verfügen. Dieses Ergebnis legt nahe, dass Beziehungen in der ULAN eher für kunsthistorisch wichtige Personen aufgezeichnet werden, deren überwiegende Mehrheit sich darüber hinaus in der Hauptkomponente befindet. Daraus kann geschlossen werden, dass es sich bei Letzterer tatsächlich um ein bestimmtes, allerdings unvollständiges Abbild des Prozesses „Kunstgeschichte“ handelt. Ein bemerkenswertes Beispiel für die Unvollständigkeit betrifft gerade berühmte deutsche Künstler wie Albrecht Dürer und Hans Holbein, die sich gemeinsam mit Teilen ihrer Familien, aber auch mit gemeinsamen Schülern in einer gesonderten Komponente finden. Zumindest hinsichtlich der bekannten Kontakte Dürers mit berühmten Vertretern der italienischen Renaissance wirft diese Beobachtung interessante Fragen auf.

In der Tat sind gerade Deutsche, gemeinsam mit Briten und Britinnen, im Verhältnis zu ihrer generellen Präsenz in der ULAN in der Hauptkomponente unterrepräsentiert, der Fokus liegt hier klar auf Personen aus Frankreich (22,71%) und Italien (14,33%), auch die Niederlande und Flandern treten signifikant häufiger auf. Interessanterweise sind auch AmerikanerInnen (13,47%) in der Hauptkomponente mehr präsent als in der Gruppe der Personen, für die Personenbeziehungen aufgezeichnet sind.

Abb. 4 zeigt eine Netzwerkvisualisierung der Hauptkomponente, wobei die Farbe der Knoten analog zu *Abb. 2* die jeweilige Nationalität der dargestellten Personen repräsentiert. Es wird sichtbar, dass das Netzwerk einer groben chronologischen Ordnung folgt und von zusammenhängenden Gruppen von Personen gleicher Nationalität dominiert wird. Diese Beobachtung ist der Tatsache geschuldet, dass 95,55% aller Beziehungen in der ULAN zwischen Personen sind, deren Geburtsdaten nicht weiter als 75 Jahre auseinanderliegen, während 81,71% von ihnen „Landsleute“ verbinden. Die Darstellung suggeriert eine Generationenabfolge von dominierenden Nationalitäten in der Kunstgeschichte, deren Hauptstrang die Reihe Italien-Frankreich-USA zu bilden scheint, wobei anzumerken ist, dass das Attribut „American“ auch vielen ExilantInnen aus Europa zugewiesen ist. Die Entwicklung beginnt links in der italienischen Renaissance (grün) und geht in Zeiten des Barock dann graduell zu französischen Personen (rot) über, die den Hauptteil der Grafik darstellen. Im 19. Jahrhundert intensivieren sich schliesslich die Verbindungen zu Personen aus Amerika (blau), die die rechte Seite der Grafik dominieren. Dieses Bild lässt sich auch in Zahlen beschreiben: Mit 48,95% bezieht sich der größte Teil der Verbindungen zwischen AmerikanerInnen und VertreterInnen anderer Nationalitäten auf Personen aus Frankreich. Dies gilt, allerdings in geringerem Ausmaß, auch für Verbindungen zwischen den Letzteren und Personen aus Italien, die 19,33% aller Verbindungen zwischen ItalienerInnen und anderen Kulturen darstellen.

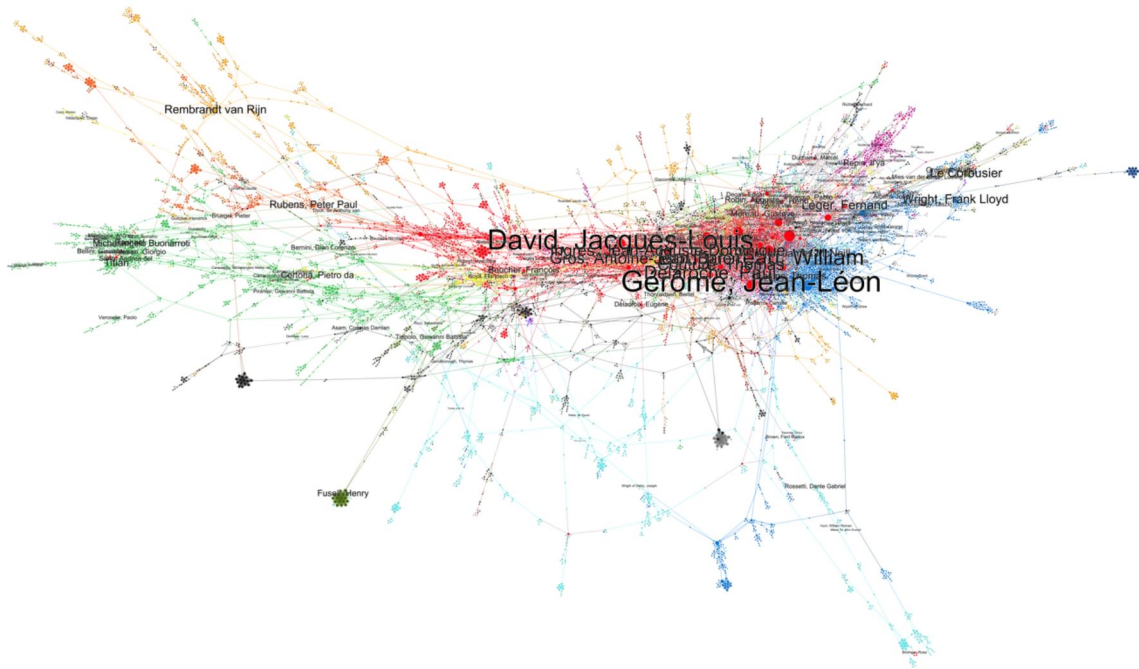


Abb. 4: Zusammenhängender Graph von 29.956 Beziehungen zwischen 10.444 ULAN Einträgen

4. ULAN-BIOGRAPHIEN IN WIKIPEDIA

Bei den 47.463 Wikidata-Personeneinträgen, die auf die ULAN und auf mindestens eine Wikipedia Sprachversion verweisen, wird eine Präferenz der einzelnen Sprachversionen für Personen aus dem sprachnahen kulturellen Kontext erkennbar. Dies kann anhand der Nationalitäten der ULAN Einträge quantifiziert werden. Wie Abb. 5 für die 12 Sprachen mit der höchsten Abdeckung zeigt, deckt die englische Version die meisten Personen aus dem englischsprachigen, die Deutsche jene aus dem deutschsprachigen Raum ab, dies ist auch bei den meisten anderen Versionen der Fall. Die Anzahl an ULAN Biographien nimmt schnell ab, während die englische Version 68,55% aller 47.463 Einträge beinhaltet, ist das bei der an 12. Stelle liegenden japanischen Wikipedia nur noch für 9,4% der Fall.

Obwohl nur 12,5% aller ULAN Datensätze andere Rollen als „Artist“ und „Architect“ haben, nimmt ihr Anteil an den abgedeckten ULAN Biographien mit abnehmender Größe der Sprachversion zu. Dies vor allem zugunsten von Biographien mit mehr politischen Rollen, aber auch von wichtigen Vertretern der Kirche oder berühmten Philosophen aus der Antike. Sind beispielsweise 87,12% der ULAN-Biographien in der englischen Version „Artist“ oder „Architect“, gilt das bei der Japanischen nur noch für 67,51%, bei der an 30. Stelle liegenden weißrussischen Wikipedia nur

noch für 58,28%. Analog sind Biographien der berühmtesten Figuren der Kunstgeschichte und jene von Vertretern der Monarchie, des Klerus und der Antike parallel in vielen Sprachversionen vertreten, während weniger bekannte Persönlichkeiten eher nur in Sprachversionen ihres eigenen kulturellen Kontextes zu finden sind. Von Eom et al. wurden letztere in 16 als Lokale, erstere als Globale historische Persönlichkeiten bezeichnet.

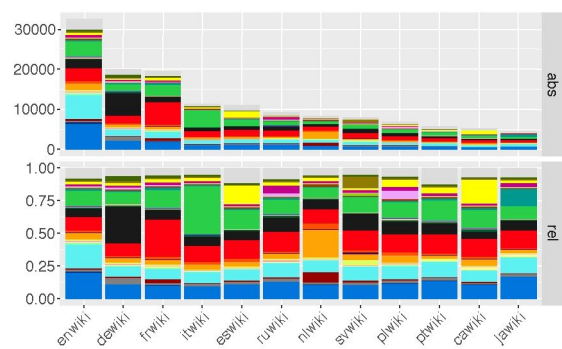


Abb. 5: ULAN-Biographien in Wikipedia

Die Tendenz der Sprachversionen zu Inhalten der eigenen Kultur setzt sich auch bei der Anzahl an Hyperlinks zwischen Personenbiographien sprachlich naher Nationen fort. Liegt das aufgrund der jeweils höher abgedeckten Nationalitäten nahe, gilt es interessanter Weise auch dann, wenn nur eine mehreren Sprachen gemeinsame Untermenge an Personen betrachtet wird. Dies zeigt, dass einer Sprache kulturell nahe Biographien dort auch über mehr einschlägigen Inhalt verfügen.

Die Vielfalt an kulturellen Präferenzen resultiert als Kombination über alle Sprachversionen hinweg in einem Netzwerk von 40.943 Wikipedia Biographien, die über 283.368 Hyperlinks miteinander verbunden sind. Das Netzwerk deckt 12.239 Personen (55,78%) aus der Untermenge derjenigen 21.942 ab, die in der ULAN über Beziehungen zueinander verfügen. Bei der ULAN Hauptkomponente ist die Abdeckung mit 7.097 (67,95%) der 10.444 dort vorkommenden Personen noch höher. Dem gegenüber sind 28.704 (16,7%) der 171.881 in der ULAN nicht verknüpften Personen in mindestens einer Wikipedia zu finden und besitzen dort dennoch Hyperlinks zu anderen ULAN Biographien. Von den 6.520 der mit der ULAN verbundenen Wikipedia Biographien ohne Hyperlinks zu anderen ULAN Personen besitzen wiederum 6.062 (92,98%) auch keine Beziehungen in der ULAN selbst. Diese Beobachtungen legen nahe, dass historisch „wichtigere“ Personen, genauso wie in der ULAN auch in Wikipedia dazu tendieren, mit anderen Personen vernetzt zu sein.

Überlappungen des Wikipedia Netzwerkes mit dem der ULAN können auch anhand der Übereinstimmung der einzelnen Beziehungen quantifiziert werden. 23.730 der 50.076 typisierten Beziehungen aus der ULAN verbinden 11.005 Personen, die mindestens in einer Wikipedia Version vorkommen. 15.747 (66,36%) davon sind auch als Hyperlinks in dem kombinierten Wikipedia Biographie Netzwerk abgebildet. Von den 23.730 sind 7.414 vom Typ „Familie“, 12.452 vom Typ „Unterricht“ und 3.864 anderer Natur. Davon sind jeweils 77,62%, 58,87% und 68,87% in Wikipedia abgedeckt. Familienbeziehungen weisen somit den höchsten Abdeckungsgrad auf, innerhalb dieses Typs sind besonders die Beziehungen „Elternteil von“ (89,13%), „Kind von“ (92,37%), „Bruder oder Schwester von“ (83,80%) sowie „Ehegatte bzw. –gattin von“ (92,98%) in sehr hohem Ausmaß auch in Wikipedia zu finden. Ein wesentlicher Unterschied zwischen Verbindungen in der ULAN und den Hyperlinks in Wikipedia ist, dass ein Hyperlink von einer Biographie zu einer anderen nicht immer entgegnet wird, im kombinierten Wikipedia Netzwerk ist das nur für 92.458 (32,63%) aller 283.368 Hyperlinks der Fall. Vergleicht man die Abdeckung der Personenbeziehungen in Wikipedia auf Basis eines gespiegelten Netzwerkes, in dem jeder einseitige Hyperlink durch ein entsprechendes

Gegenstück ergänzt wird, ergibt sich eine deutlich höhere Abdeckung. Familienbeziehungen sind nun zu 83,41%, Ausbildungsverhältnisse zu 78,83% und die restlichen Typen zu 81,63% abgebildet. In Summe lässt sich auf diese Weise zeigen, dass für 19.154 (80,72%) der 23.730 betrachteten ULAN Beziehungen entsprechende Informationen in Wikipedia präsent sind.

Von der Seite der in Wikipedia vorhandenen Hyperlinks ergibt sich ein entgegengesetztes Bild. Für die soeben betrachtete Untermenge von 11.005 Personen existieren in dem kombinierten Wikipedia Netzwerk 61.639 Hyperlinks, die in der ULAN nicht als Beziehungen aufgezeichnet sind, gespiegelt erhöht sich diese Menge sogar auf 103.418. Im letzteren Fall gibt es demnach knapp fünf mal so viele Personenbeziehungen in Wikipedia. Betrachtet man die gesamten 40.943 in Wikipedia vernetzten ULAN-Biographien, so gibt es dort 267.621 Hyperlinks die keine Entsprechung in der ULAN haben, gespiegelt erhöht sich dies auf die sehr große Anzahl von 455.124. Im Vergleich zur ULAN existieren in Wikipedia für die selben Personen somit knapp 20 mal so viele gegenseitige Beziehungen. Diese Beobachtung legt einerseits nahe, dass Hyperlinks in Wikipedia viel weiter gefasste Arten von Personenbeziehungen beschreiben, die in der ULAN grundsätzlich auf verschiedene Arten direkter Interaktion beschränkt sind. Andererseits suggeriert sie auch, dass der Prozess der Aufzeichnung von Beziehungen in der ULAN wesentlich spezifischeren Vorgaben folgt. So ist beispielsweise Dürers in der ULAN nicht aufgezeichneter Kontakt mit Giovanni Bellini und anderen Italienern seiner Zeit in vielen Wikipedia Sprachversionen dokumentiert. Aspekte wie dieser sind mit hoher Wahrscheinlichkeit einer der deutlichsten Unterschiede zwischen professionell kuratierten und „Crowd-Sourced“ Datensammlungen.

Die Asymmetrie der Hyperlinks zwischen Biographien in Wikipedia wird auch in der zeitlichen Dimension sichtbar. Betrachtet man die Differenz zwischen den Geburtsdaten zweier verknüpfter Personen so wird deutlich, dass 63,43% aller Hyperlinks in Wikipedia in die Vergangenheit zeigen. Ganz im Sinne der zuvor festgestellten weiteren Auffassung von Personenbeziehungen erstrecken sich diese Verweise im Gegensatz zu ULAN-Beziehungen teilweise auch über mehrere

Jahrhunderte, wobei aber auch in Wikipedia eine deutliche Mehrheit der Hyperlinks (85.24%) Personen mit einem Altersunterschied von bis zu 75 Jahren verknüpft. Auffallend ist hierbei auch der Umstand, dass 64,07% aller Hyperlinks mit einer maximalen Distanz von 75 Jahren Personen gleicher Nationalität verbinden, während das bei weiter auseinanderliegenden Verbindungen nur bei 36,17% der Fall ist. Letztere treten vor allem bei Verknüpfungen mit besonderen Persönlichkeiten der Kunst- bzw. Kulturgeschichte auf.

Eingeschränkt auf Verbindungen zwischen Zeitgenossen haben ULAN und Wikipedia relativ große strukturelle Ähnlichkeit bei intra- und internationalen Beziehungen. Werden nur Wikipedia-Hyperlinks mit maximaler Distanz von 75 Jahren betrachtet, ergibt sich trotz der nur zum Teil übereinstimmenden Biographien und der wesentlich größeren Anzahl an Personenbeziehungen ein ähnliches, wenn auch wesentlich dichteres und dadurch weniger detailliertes Bild, als in *Abb. 4*. für die ULAN-Hauptkomponente zu sehen ist. *Abb. 6* zeigt das für eine Auswahl von Wikipedia-Sprachen sowie für das kombinierte Netzwerk aller Versionen. Während besonders die englische Wikipedia dem ULAN Netzwerk dahingehend ähnelt, dass ein zeitliche Abfolge von Italien, Frankreich und Amerika zentral erscheint, unterscheidet sich ihr Bild gleichzeitig aber auch z.B. durch die wesentlich dichtere Gruppe von Personen aus Großbritannien von der professionellen Sicht. Bei der deutschen Sprachversion ist wiederum der deutschsprachige Beitrag zur Kunstgeschichte deutlich sichtbar, viele Biographien über deutsche KünstlerInnen finden sich nur in dieser Sprachversion. Die Bilder der anderen Sprachen sind aufgrund der schnell sinkenden Anzahl an abgedeckten ULAN Biographien auffallend dünner besetzt, zeigen jedoch ebenfalls deutliche Präferenzen für Personen und Beziehungen aus dem jeweiligen kulturellen Kontext.

Trotz aller Unterschiede ist die Grundstruktur der verknüpften ULAN-Biographien über alle Sprachversionen hinweg ähnlich, nicht zuletzt aufgrund der vielen berühmten Persönlichkeiten, die sich in fast allen Versionen finden und wichtige Knotenpunkte darstellen. Als Beispiel ist Pablo Picasso genannt, der in vielen Fällen die Person mit den meisten eingehenden Hyperlinks darstellt. Gerade die unmittelbare Umgebung dieses Künstlers stellt ein im

Vergleich zu anderen Teilen des Netzwerkes sehr internationales Umfeld dar. Die kombinierte Darstellung der in den betrachteten Wikipedia-Sprachen jeweils vorkommenden ULAN-Biographien und ihrer Verknüpfungen vereint unterschiedliche kulturelle Foci in ein Gesamtbild, dass die vertretenen Nationalitäten auf deutlich ausgewogenere Weise zueinander in Bezug setzt. Dennoch bleibt der eurozentristische Aspekt der in der ULAN abgebildeten Kunstgeschichte weiterhin stark sichtbar.

5. SCHLUSS

Die aus der ULAN und aus Wikipedia erzeugten Bilder haben eine visuelle Grundform gemeinsam, die bereits 1933 von Alfred H. Barr für ein weiteres Diagramm zur Darstellung der idealen Sammlung des Museum of Modern Art in New York gewählt wurde. Ausgehend von seinem „Report on the Permanent Collection“ 17 visualisiert die Form eines sich stets in Richtung Zukunft bewegenden Torpedos die sich ändernde Zusammensetzung der Sammlung zu unterschiedlichen Zeitpunkten (vgl. Schmidt-Burkhardt in 2). Sammlungsschwerpunkte werden durch Abschnitte innerhalb des sich zur Vergangenheit hin verjüngenden Torpedos dargestellt, die regional-/nationalen „Schulen“ zugeordnet sind und durch die Bewegung des Torpedos über die Zeit darin auftauchen, wachsen und wieder verschwinden.

Obwohl aus ihrem spezifischen Kontext heraus nicht direkt vergleichbar, ist die Ähnlichkeit der sich aus den Netzwerkvisualisierungen ergebenden Bilder zu dieser Metapher aus der prädigitalen Zeit erstaunlich und regt zu Überlegungen an. Während Barrs Torpedo den Ausschnitt der sich ändernden Sammlung des MoMA aus einer potentiell unendlichen Menge an Kunstwerken repräsentieren soll, zeigen die aus der ULAN und Wikipedia abgeleiteten Formen gegenwärtige Ausschnitte der Kunstgeschichte, die sich aus den digital verfügbaren Kontextinformationen vieler verschiedener Sammlungen und anderer Quellen ergeben. Im Unterschied zu Barrs Intention einer Selektion steht hier die Frage nach der Vollständigkeit im Vordergrund, ob und welche Informationen sich möglicherweise nicht in der betrachteten Datenquelle befinden, stattdessen anderswo vorhanden sind, gar nie existiert haben, nicht aufgezeichnet wurden, oder einfach in Vergessenheit geraten sind.

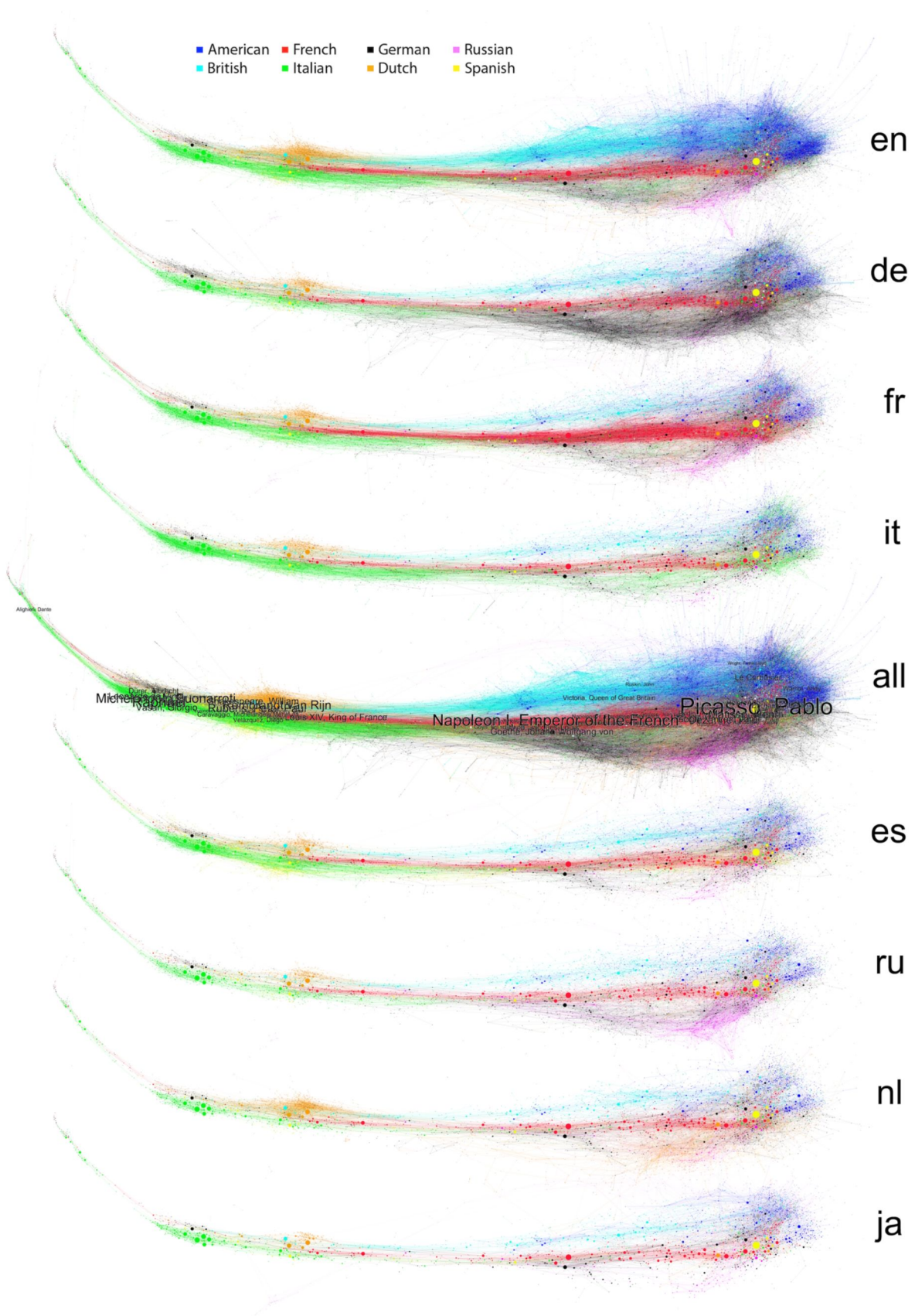


Abb. 6: Hyperlinks zwischen ULAN-Personenbiographien in verschiendensprachigen Wikipedien.

6. DANKSAGUNG

Teile dieser Arbeit wurden aus Mitteln des FWF (Fonds zur Förderung der wissenschaftlichen Forschung) gefördert, Projektnummer L602 „The Virtual 3D Social Experience Museum“.

7. LITERATURHINWEIS

1. Fischer, Nora: Kunst nach Ordnung, Auswahl und System. In: Swoboda, Gudrun, *Die kaiserliche Gemäldegalerie in Wien und die Anfänge des öffentlichen Kunstmuseums, Band 1*, Böhlau, Wien, 2013, S. 22-90.
2. Schmidt-Burkhardt, Astrit: *Stammbäume der Kunst*, Akademie Verlag, Berlin, 2005.
3. Glinka, Katrin und Dörk, Marian: Museum im Display. Visualisierung kultureller Sammlungen (Vikus). *Elektronische Medien & Kunst, Kultur und Historie: 22. Berliner Veranstaltung der internationalen EVA-Serie: Electronic Media and Visual Arts*, Berlin, 4. - 6. November 2015, arthistoricum.net, Heidelberg, S. 18-23.
4. Schich, Maximilian, et al.: A Network Framework of Cultural History. *Science*, 345, 6196, S. 558–562, 08/2014.
5. Klein, Maximilian und Kyrios, Alex: VIAFbot and the Integration of Library Data on Wikipedia. *Code4Lib Journal*, 0, 22, S. 85–107, 10/2013.
6. Vrandečić, Denny und Krötzsch, Markus: Wikidata: A Free Collaborative Knowledgebase. *Communications of the ACM*, 57, 10, S. 78-85, 09/2014.
7. Auer, Sören, et al.: DBpedia: A Nucleus for a Web of Open Data. *6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007*, Busan, 11. - 15. November 2007, Springer, Berlin/Heidelberg, S. 722-735.
8. Clough, Paul D., Ireson, Neil und Marlow, Jennifer: Extending Domain-Specific Resources to Enable Semantic Access to Cultural Heritage Data. *Journal of Digital Information*, 10, 6, 12/2009.
9. Manovich, Lev: Data Science and Digital Art History. *International Journal for Digital Art History*, 0, 1, S. 13-35, 06/2015.
10. Green-Armytage, Paul: A Colour Alphabet and the Limits of Colour Coding. *JAIC - Journal of the International Colour Association*, 5, 0, 08/2010.
11. Priestley, Joseph: *A Description of a Chart of Biography; with a catalogue of all the names inserted in it, and the dates annexed to them ... Second edition*, J. Bowles, London, 1765.
12. Anderson, Benedict: Imagined Communities. In: Spencer, Philip und Wollmann, Howard, *Nations and Nationalism: A Reader*, Rutgers University Press, New Brunswick, 2005, S. 48-60.
13. Jensen, Robert: Measuring Canons: Reflections on Innovation and the Nineteenth-century Canon of European Art. In: Brzyski, Anna, *Partisan Canons*, Duke University Press, Durham, 2007, S. 27-54.
14. Elkins, James: Canon and Globalization in Art History. In: Brzyski, Anna, *Partisan Canons*, Duke University Press, Durham, 2007, S. 55-77.
15. Dilly, Heinrich: Der kunsthistorische Nachthimmel. In: Berndt, Andreas, Kaiser, Peter, Rosenberg, Angela und Trinker, Diana, *Frankfurter Schule und Kunstgeschichte*, Reimer, Berlin, 1992, S. 69-84.
16. Eom, Young-Ho et al.: Interactions of Cultures and Top People of Wikipedia from Ranking of 24 Language Editions. *PLoS ONE*, 10, 3, 2015, Public Library of Science, San Francisco.
17. Barr, Alfred H.: Report on the Permanent Collection, In: *The Museum of Modern Art Archives, Alfred H. Barr, Jr. Papers, 9a.7A*, Museum of Modern Art, New York, 1933.