

MARTIN PAPENBROCK (KARLSRUHE), JOACHIM SCHARLOTH (TOKYO)

## Datengeleitete Analyse kunsthistorischer Daten am Beispiel von Ausstellungskatalogen aus der NS-Zeit: Musteridentifizierung und Visualisierung

### **Abstract**

Im Rahmen einer interdisziplinären Kooperation zwischen der Kunstgeschichte einerseits und der Korpus- und Computerlinguistik andererseits wollen wir die Möglichkeiten der datengeleiteten Analyse und der Visualisierung als Methode der Erkenntnisgewinnung sowie ihr Potenzial für die Analyse kunstwissenschaftlicher Datensammlungen exemplarisch erproben. In einer Pilotstudie untersuchen wir anhand von Ausstellungskatalogen zur Gegenwartskunst aus der NS-Zeit, ob es Künstler gegeben hat, die häufig miteinander ausgestellt wurden, was diese Künstler untereinander verbindet und ob sich nach den Ausstellungsereignissen des Jahres 1937 Veränderungen in den Konfigurationen häufig miteinander ausgestellter Künstler zeigen. Wir versprechen uns Ergebnisse, die nicht nur Rückschlüsse auf die Wirksamkeit der staatlichen Ausstellungspolitik, sondern der zentralistischen Kunstpolitik des Nationalsozialismus insgesamt erlauben.

### **1 Visualisierung kunsthistorischer Daten**

#### **1.1 Visualisierung als Methode der Erkenntnisgewinnung**

<1>

In der Kunstwissenschaft werden digitalisierte Datenbestände, seien es Bild- oder Textdatenbanken, vor allem zu dokumentarischen Zwecken angelegt. Ihr Sinn wird darin gesehen, große Mengen an Informationen verfügbar zu machen, sie abzubilden oder aufzulisten und der rechnergestützten Recherche oder der statistischen Untersuchung zur Verfügung zu stellen. Man kann diesen Gebrauch von Kunstdatenbanken als datenbasiert (data based) bezeichnen. Datenbanken auf der Basis vorgängiger Hypothesen zu befragen, birgt jedoch die Gefahr, in den Daten nur jene Strukturen zu finden, die mit der Theorie kompatibel sind und blind gegenüber Evidenzen zu sein, die quer zu ihr stehen. Das Gegenstück zur datenbasierten Vorgehensweise sind datengeleitete (data driven) Verfahren. Dabei handelt es sich um Verfahren, die induktiv Muster in Datenbeständen aufdecken und deren Algorithmen unabhängig von Hypothesen über mögliche Zusammenhänge in den

Daten entwickelt wurden. Sie haben das Potenzial, in großen und komplexen Datenbeständen Auffälligkeiten und Muster sichtbar zu machen, die beim datenbasierten Verfahren gar nicht in den Blick geraten wären, und so gängige Kategorien und etablierte Theorien in Frage zu stellen. Die Möglichkeiten datengeleiteter Analysen sind von der Kunstgeschichte, anders als von den technischen Disziplinen, der Soziologie oder der Linguistik, bisher noch nicht in Anspruch genommen worden. Anders als bei datenbasierten Verfahren, deren Ergebnisse umso prägnanter sind, je präziser die Hypothesen formuliert wurden, produzieren datengeleitete Analysen als Ergebnisse selbst sehr große Datenmengen, die sich nur schwer einer Gesamtschau unterziehen lassen. Um sie interpretierbar zu machen, werden im Zuge dieser Verfahren Bilder (Visualisierungen) generiert, die es zu interpretieren gilt, zwar keine künstlerischen, sondern technische Bilder, Graphen und Diagramme, aber Bilder, aus denen sich Erkenntnisse ableiten und Zusammenhänge erschließen lassen.

<2>

Mit unserem Projekt *informARTics*,<sup>1</sup> einer interdisziplinären Kooperation zwischen Kunstgeschichte einerseits und Korpus- und Computerlinguistik andererseits, wollen wir die Möglichkeiten der datengeleiteten Analyse und der Visualisierung als Methode der Erkenntnisgewinnung sowie ihr Potenzial für die Analyse kunstwissenschaftlicher Datensammlungen exemplarisch erproben. Beispielgebend sind die technischen Disziplinen, wo vor allem in Mathematik, Informatik, Naturwissenschaften und life-sciences bei der Analyse komplex vernetzter Daten mit Visualisierungen gearbeitet wird, die in der Medizin den Namen »bildgebende Verfahren« tragen. Diese haben nicht nur illustrativen Charakter, sondern sind selbst Mittel der Erkenntnisgewinnung, insofern sie die Daten filtern, gewichten und in eine Form bringen, die vom Menschen erfassbar und interpretierbar ist.

<3>

Dies lässt sich an einem der ältesten bildgebenden Verfahren in der Medizin – dem Röntgenbild – erläutern. Röntgenstrahlen werden, je nach der Dichte des Gewebes, auf das sie gerichtet werden, teilweise gefiltert. Die Menge der Strahlen, die nach dem Durchgang durch ein Gewebe noch messbar sind, lassen also Rückschlüsse auf dessen Beschaffenheit zu. Würde die Verteilung der durchgedrungenen Strahlen auf einer Fläche in Form einer Zahlenmatrix vorliegen, wäre diese Matrix angesichts der unzählbaren Menge an Strahlen nicht erfassbar. Durch die Lenkung dieser Strahlen auf einen geeigneten Film (oder ein digitales Detektorensystem) jedoch kann die Verteilung durchgekommener und gefilterter Strahlen visualisiert werden. Es entsteht ein zweidimensionales Bild, das die Durchdringbarkeit (dunkle Bereiche) bzw. Undurchdringbarkeit (helle Bereiche) des

Gewebes für Röntgenstrahlen repräsentiert und damit Rückschlüsse auf die Beschaffenheit des Gewebes zulässt. Noch erheblich komplexer wird die Analyse im Fall der Computertomographie: sie kombiniert die Daten vieler Röntgenbilder aus unterschiedlichen Perspektiven zu einem dreidimensionalen Bild des Gewebes.

<4>

Visualisierungen sind freilich keine Abbildungen der Wirklichkeit, sondern aufgrund von Relevanzkriterien geordnete Reduktionen von Daten, die auf der Basis gestalterischer Vorgaben visualisiert werden. Beim klassischen Röntgenbild etwa wird ein dreidimensionaler Raum auf eine Fläche reduziert; und in einer Computertomographie werden Repräsentationen von Gewebeteilen mit dem Ziel der Unterscheidbarkeit eingefärbt und nicht ihrer tatsächlichen Färbung entsprechend.

<5>

Visualisierungen sind insbesondere dann Mittel der Erkenntnisgewinnung, wenn andere Formen der Repräsentation wie Listen, Tabellen oder Texte zu umfangreich sind, um als Ganze erfasst und interpretiert werden zu können. Die Entwicklung von Methoden zur Visualisierung ist daher in den genannten Disziplinen integraler Bestandteil des Forschungsprozesses. Auch im Bereich der Sozialwissenschaften fassen Visualisierungen über den Bereich der Illustration statistischer Informationen hinaus mehr und mehr Fuß, etwa im Bereich der Netzwerkanalyse.<sup>2</sup> In den Kultur- und Geisteswissenschaften steckt die Entwicklung von durch Visualisierung geleiteten Methoden dagegen noch in den Kinderschuhen. Und dies, obwohl seit der Begründung der digital humanities und der Computerphilologie große digitale Textarchive sowie kultur- und kunstwissenschaftliche Datenbanken entstanden sind.

## 1.2 Fragestellung

<6>

In einer ersten Fallstudie möchten wir anhand von Ausstellungskatalogen zur Gegenwartskunst aus der NS-Zeit untersuchen, ob sich mittels der Ausstellungsaktivitäten Gruppen von Künstlern identifizieren lassen, deren Werke besonders häufig miteinander ausgestellt wurden. Da es sich um insgesamt 378 Ausstellungskataloge handelt, in denen 10.014 Künstler verzeichnet sind, ist diese Aufgabe mit nicht-maschinellen Methoden nicht zu bewältigen. Zudem interessiert uns die Frage, ob sich in unterschiedlichen Phasen unterschiedliche Gruppen identifizieren lassen. Um den einschneidenden ausstellungspolitischen Ereignissen des Jahres 1937, der Ausstellung *Entartete Kunst* und der erstmals ausgerichteten *Großen Deutschen Kunstausstellung* in München, Rechnung zu

tragen, werden wir den Untersuchungszeitraum in zwei Phasen unterteilen, die wir vergleichend analysieren wollen: die Jahre bis einschließlich 1937 und die Jahre ab 1938.

<7>

Die Neustrukturierung des staatlichen Kunstbetriebs, die bereits 1933 einsetzte, die Einrichtung von Kontrollinstanzen wie der Reichskulturkammer und der Reichskammer der bildenden Künste, deren Mitgliedschaft für alle ausstellenden Künstler verbindlich war, die Gleichschaltung der Kunstvereine und die Genehmigungspflicht für sämtliche Kunstaussstellungen waren Zentralisierungsmaßnahmen, die eine politische Steuerung der Gegenwartskunst ermöglichen und eine Homogenisierung des künstlerischen Schaffens im Sinne ihrer politischen Verwertbarkeit herbeiführen sollte. Mit der Analyse der Ausstellungskataloge möchten wir prüfen, ob dieses kunstpolitische Ziel erreicht wurde. Wir wollen fragen, ob es in den Ausstellungen typische Konfigurationen von Künstlern gegeben hat, was diese Künstler, die häufig miteinander ausgestellt wurden, untereinander verbindet und ob sich nach den Münchner Großereignissen des Jahres 1937 Veränderungen in den Gruppen häufig miteinander ausgestellter Künstler zeigen. Uns interessiert die Frage, ob die datengeleitete Analyse regionale Kunstszenen abbildet oder eine überregionale, nationale Szene sichtbar macht, wie sie von der nationalsozialistischen Kunstpolitik intendiert war. Darüber hinaus wird zu fragen sein, welche Rolle die prominenten Künstler in den Ausstellungen spielten, die »Staatskünstler«, die mit großen öffentlichen Aufträgen bedacht wurden und wichtige Positionen in den staatlichen Behörden (Kunstkammer) und Institutionen (Akademien, Hochschulen) bekleideten, ob sie in den Ausstellungen eine eigene Gruppe bildeten oder verschiedenen Szenen angehörten. Auch die Frage nach der Präsenz der modernen Künstler, die als »entartet« diffamiert und zum Teil mit Arbeits- und Ausstellungsverbot belegt wurden, wird bei der Analyse eine Rolle spielen.

<8>

Die Forschung geht von einer hohen Wirksamkeit der kunstpolitischen Maßnahmen des Nationalsozialismus aus, insbesondere im Bereich des Ausstellungswesens.<sup>3</sup> Die Propaganda-Schau »Entartete Kunst« gehört schon seit Jahrzehnten zu den bestdokumentierten Kunstaussstellungen des 20. Jahrhunderts,<sup>4</sup> und die Dokumentation der *Großen Deutschen Kunstaussstellungen* ist inzwischen ebenfalls sehr weit fortgeschritten.<sup>5</sup> Die wissenschaftliche Monumentalisierung der beiden ausstellungspolitischen Großprojekte hat dazu geführt, dass sie inzwischen als Prototypen des nationalsozialistischen Ausstellungswesens wahrgenommen werden. Da wir im Bereich der staatlich akzeptierten und geförderten Ausstellungen aber nicht nur die seit 1937 jährlich stattfindenden *Großen Deutschen Kunstaussstellungen*, sondern sämtliche Gruppenausstellungen deutscher Gegenwartskunst, zu denen Kataloge erhalten sind, in den Blick nehmen, wird sich die

Bedeutung der Münchner Jahresausstellungen im Ergebnis zwangsläufig relativieren und das Bild differenzierter werden. Ob und bis zu welchem Grad man tatsächlich von einer normierenden Wirkung der kunstpolitischen Maßnahmen und entsprechenden Effekten der jährlichen *Großen Deutschen Kunstausstellung* sprechen kann, soll die Untersuchung zeigen. Wir versprechen uns Ergebnisse, die nicht nur Rückschlüsse auf die Wirksamkeit der staatlichen Ausstellungspolitik, sondern der zentralistischen Kunstpolitik des Nationalsozialismus insgesamt erlauben.

## 2 Daten und Datenaufbereitung

### 2.1 Kataloge zu Ausstellungen von Gegenwartskunst in der NS-Zeit

<9>

Die Daten wurden im Rahmen des Projekts *Ausstellungsgeschichte des 20. Jahrhunderts* erhoben, das von 1995 bis 1997 im Fach Kunstgeschichte an der Universität Osnabrück durchgeführt wurde. Dieses Projekt war Teil des Projektverbundes *EDV-gestützte Dokumentation in Forschung und Lehre ausgewählter kunstgeschichtlicher Universitätsinstitute*, der vom Bildarchiv Foto Marburg initiiert und geleitet wurde und an dem neben dem Fach Kunstgeschichte der Universität Osnabrück auch die kunstgeschichtlichen Institute der Universitäten Berlin (FU), Bochum, Bonn, Frankfurt, Freiburg, Göttingen und Leipzig mit selbständigen Projekten beteiligt waren. Es ging darum, den Einsatz standardisierter Verfahren in der kunstgeschichtlichen Forschung und Lehre zu fördern und seine Möglichkeiten und Grenzen zu erproben. Die erhobenen Daten sollten in die DISKUS-Datenbank von Foto Marburg einfließen.<sup>6</sup>

<10>

Das Osnabrücker Projekt wurde von Jutta Held, Martin Papenbrock und Gabriele Saure geleitet. Ziele des Projekts waren die bibliographische Erfassung, Auswertung und Kommentierung von Ausstellungskatalogen zur Kunst der 30er und 40er Jahre in Deutschland und die Herausgabe entsprechender kommentierter Bibliographien. In einem Teilprojekt, das von Martin Papenbrock geleitet und von Anette Sohn bearbeitet wurde, wurden Kataloge zu Gruppenausstellungen von Gegenwartskunst in der NS-Zeit erfasst und ausgewertet. Im Rahmen dieses Teilprojekts konnten 378 Ausstellungskataloge ermittelt und aufgenommen werden. Die kommentierte Bibliographie, die die Kataloge auflistet und dokumentiert, ist bis heute das umfangreichste Nachschlagewerk zum Ausstellungswesen im Nationalsozialismus.<sup>7</sup>

## 2.2 Preprocessing

<11>

Die erhobenen Daten zu den Ausstellungskatalogen liegen sowohl in Buchform als auch in Form einer Datenbank vor, die neben den Namen jener Künstler, deren Werke ausgestellt wurden, auch Angaben zum Ort und Termin der Ausstellungen sowie bibliographische Informationen zu den erfassten Ausstellungskatalogen enthält. Dass die Daten bereits in strukturierter Form vorlagen, vereinfachte ihre Analyse erheblich. Dennoch mussten vor dem Beginn der Analyse einige Anpassungen vorgenommen werden.

<12>

Ein Problem ergab sich daraus, dass die Datenbank die Künstlernamen in exakt der Form verzeichnet, in der sie sich im Katalog fand. So ist es möglich, dass sich für den Namen des (mutmaßlich) selben Künstlers mehrere schriftliche Repräsentationen finden. Zudem wurde beim Erstellen der Datenbank dem Namen eines Künstlers dann ein »(A)« hinzugefügt, wenn sich die Abbildung eines seiner Werke im Ausstellungskatalog fand. So finden sich für den Namen des (mutmaßlich) selben Künstlers in der Datenbank die folgenden Schreibvarianten:

<13>

Agricola, R.

Agricola, Rudolf

Agricola, Rudolf (A)

Agricola, Rudolf A.

Agricola, Rudolf A. (A)

Agricola, Rudolf Alexander

<14>

Während sich die Annotation der Abbildung eines Werkes leicht automatisiert entfernen ließ, stellten die Schreibvarianten ein größeres Problem dar. Ideal wäre es gewesen, auf der Basis bio-bibliographischen Wissens die Schreibvarianten auf einen Type zurückzuführen. Da für die meisten der über 10.000 Künstlerinnen und Künstler keine gesicherten biographischen Informationen zur Verfügung stehen, die Rückführung zudem sehr zeitaufwändig gewesen wäre und den Rahmen einer Pilotstudie gesprengt hätte, haben wir uns dafür entschieden, die Künstlernamen auf den Nachnamen und den Anfangsbuchstaben des Vornamens zu kürzen. Dadurch reduzierte sich die Anzahl der in der Datenbank vorkommenden Namen von 14.563 auf 10.014.

<15>

Für die Analysen wurde die Datenbank einerseits in ein XML-Format transformiert, zum anderen wurde eine Datenmatrix (Abb. 1) erstellt. Objekte der Datenmatrix sind die Namen

der Künstler, die Variablen (Merkmale) die Ausstellungen. Die Elemente der Matrix haben die Merkmalsausprägungen 1 (Teilnahme an der Ausstellung) oder 0 (keine Teilnahme an der Ausstellung). Da die Analyse darauf zielte, Künstlergruppen in zwei unterschiedlichen Phasen zu identifizieren, wurden für jede Phase je eine XML-Datei und Datenmatrix erstellt.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
<b>Abeking, H.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Abel, A.</b>	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0
<b>Abel, K.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Abel, W.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Abelen, P.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Abramowski, E.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Abt, K.</b>	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
<b>Achenbach, A.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Achenbach, H.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Achenbach, M.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Achenbach-Junemann, H.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Achepol, E.</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Abb. 1: Ausschnitt Datenmatrix

### 3 Data Mining und linguistische Methoden der Mustererkennung

<16>

Bei der Analyse der Daten benutzen wir Methoden, die teils dem klassischen Data-Mining, teils der Korpus- und Computerlinguistik entlehnt sind. »Data Mining« ist der Oberbegriff für statistische Methoden, die dafür geeignet sind, Wissen aus großen Datenbeständen zu extrahieren. Unter Wissen verstehen mit Data Mining befassende Informatiker »interessante Muster« und bestimmen interessante Muster als solche, »die allgemein gültig sind, nicht trivial, neu, nützlich und verständlich«. <sup>8</sup> Zur Bestimmung interessanter Muster – das zeigt die Definition – sind also maschinelle Verfahren alleine nicht ausreichend; die Ergebnisse bedürfen der Evaluation und Interpretation durch Expertinnen und Experten, mit deren Hilfe die verwendeten Methoden für die gewählte Fragestellung optimiert werden können.

<17>

Typische Analysemethoden des Data Mining sind Korrelation, Regression, Prognose, Klassifikation und Clusteranalyse. Diese Methoden können auf ganz unterschiedliche Datentypen angewendet werden: auf industrielle Prozessdaten mit dem Ziel der Optimierung von Abläufen, auf Kundendaten zur Verbesserung des Marketings, <sup>9</sup> auf Bilddaten etwa zur Identifizierung von Layouts, <sup>10</sup> auf Textdaten zur Filterung relevanter Dokumente, zur Erkennung von Spam oder zur Erstellung von Autorenprofilen, <sup>11</sup> auf biomedizinische Daten etwa im Bereich der Erbgutanalyse <sup>12</sup> und auf forensische Daten etwa zur Identifizierung von

Netzwerken mit abweichendem Verhalten.<sup>13</sup> Viele der Methoden verwenden Visualisierungsverfahren, um die entdeckten Muster lesbar zu machen.

<18>

Auch in der Korpus- und Computerlinguistik werden Verfahren der Mustererkennung dafür eingesetzt, um sprachliche Regularitäten zu identifizieren, semantische Einheiten oder Relationen zu finden und Sprachwandelprozesse zu analysieren oder gar zu prognostizieren. Ein Korpus ist eine »[e]ndliche Menge von konkreten sprachlichen Äußerungen, die als empirische Grundlage für sprachwissenschaftliche Untersuchungen dienen.«<sup>14</sup> Forschungspraktisch sind Korpora große Mengen digitaler Texte.<sup>15</sup> Häufig eingesetzte Verfahren zur Auffindung von Mustern in Korpora sind die Analyse von n-Grammen (Folge von n sprachlichen Einheiten) und von Kookkurrenzen (gemeinsames Auftreten von Wörtern in einem bestimmten Textfenster).<sup>16</sup>

<19>

Da unsere Fragestellung darauf zielt, Künstler auf der Basis ihrer (gemeinsamen) Beteiligung an Ausstellungen zu gruppieren, haben wir uns dafür entschieden, je eine Methode zur Identifizierung von Gruppen ähnlicher Objekte des klassischen Data Mining und der Korpuslinguistik auf die Daten anzuwenden: zum einen die hierarchische Clusteranalyse und zum anderen die linguistische Kollokationsanalyse. Da die Clusteranalyse zu den in der empirischen Sozialforschung etablierten Verfahren gehört, wird unser Vorgehen hier nur skizziert. Etwas ausführlicher werden wir die Analyseschritte bei der Kollokationsanalyse und ihrer Adaption für unsere Daten und Fragestellung darstellen.

### **3.1 Kollokationsanalyse: Kollokationsnetze als gerichtete Graphen**

<20>

Der Begriff der Kollokation ist in der Linguistik durchaus umstritten.<sup>17</sup> Einig sind sich jedoch die meisten Definitionen darin, dass es sich bei Kollokationen um rekurrente Kookkurrenzen von mindestens zwei lexikalischen Einheiten in einer begrenzten Menge sprachlicher Daten handelt. Typische Beispiele für Kollokationen sind »Weg« und »einschlagen« oder »Geschichte« und »erzählen«, die in sprachlichen Daten jeweils überzufällig häufig miteinander auftreten. Je nach Erkenntnisinteresse und Anwendungsgebiet der Kollokationsanalysen werden weitere Bestimmungen als konstitutive Bestandteile des Kollokationsbegriffs formuliert: etwa im Hinblick auf die Art der Relation zwischen den Wörtern oder im Hinblick auf die Menge der Sprachdaten, in denen die lexikalischen Einheiten gemeinsam vorkommen.

<21>

Die Grundlage für die Berechnung von Kollokationen ist die Bestimmung von Kollokatoren. Kollokatoren sind Wörter, die gemeinsam mit einem bestimmten Wort (im Folgenden als »Basis« bezeichnet) auftreten. Je nach linguistischem Erkenntnisinteresse legt man ein Fenster von n Wörtern links und m Wörtern rechts der Basis fest und wertet dieses Fenster bei jedem Auftreten der Basis im Korpus aus. So erhält man eine Liste von Wörtern, die gemeinsam mit der gesuchten Basis auftreten, sowie ihrer Frequenzen. Die Liste der Kollokatoren alleine ist jedoch noch nicht aussagekräftig. Um zu überprüfen, ob eine Basis und ein Kollokator überzufällig häufig miteinander auftreten, muss untersucht werden, ob der Kollokator in anderen Kontexten nicht ähnlich häufig auftritt wie im Kontext der Basis. Hierfür teilt man das Korpus in zwei Teilkorpora. Das erste Teilkorpus besteht aus allen Basen und den dazugehörigen Textfenstern, also von den n Wörtern links und m Wörtern rechts der Basis. Das zweite Teilkorpus besteht aus allen restlichen Wörtern. Nun zählt man die Frequenzen des Kollokatorwortes im zweiten Korpus und berechnet, wie häufig es in beiden Korpora relativ zur Wortzahl des jeweiligen Korpus vorkommt. Zur Beantwortung der Frage, ob das gemeinsame Auftreten von Kollokator und Basis als zufällig oder überzufällig gelten kann, setzt man die relativen Frequenzen miteinander in Beziehung: als statistische Verfahren werden unterschiedliche Signifikanztests verwendet.<sup>18</sup> Je nach Wahl der Parameter (Korpus, Fenstergröße, Grenzwerte, Signifikanztest) erhält man unterschiedliche Ergebnisse.

<22>

Das beschriebene korpuslinguistische Verfahren lässt sich – wenn auch mit einigen Anpassungen – für die Bearbeitung unserer Fragestellung fruchtbar machen. Da es in unserem Projekt darum geht, Künstler in Gruppen einzuteilen, die häufig gemeinsam in Ausstellungen vertreten waren, ist das Konstrukt der Kollokation, verstanden als ein überzufälliges gemeinsames Auftreten zweier Phänomene, ein geeigneter Anknüpfungspunkt. Jedoch unterscheiden sich sprachliche Daten von den Daten in unserem Projekt in einem wichtigen Punkt: (1) Während sprachliche Daten linear in Texten repräsentiert sind und vor diesem Hintergrund die Wahl einer Wortspanne links und rechts der Basis plausibel wird, bilden die aus einem Ausstellungskatalog extrahierten Künstlernamen für diese Ausstellung eine ungeordnete Menge. (2) In dieser Menge kann jeder Name genau einmal vertreten sein, ganz gleich wie groß die Ausstellung war, während es bei sprachlichen Daten möglich ist, dass ein Kollokator mehrmals im gleichen Fenster auftritt und die Wahrscheinlichkeit dafür wächst, je größer das Fenster gewählt wird.

<23>

Ausgehend von diesen Unterschieden in der Datenstruktur waren bei der Anwendung eines kollokationsanalytischen Verfahrens auf unsere Fragestellung folgende Anpassungen nötig: (1) die Größe des Fensters zur Bestimmung von Kollokatoren liegt nicht fest, sondern entspricht jeweils der Menge der in der jeweiligen Ausstellung vertretenen Künstler. (2) Als Basis für die Berechnung der Typizität des gemeinsamen Auftretens von Künstlern wurde nicht die Anzahl der Token in den beiden Teilkorpora gewählt, sondern die Anzahl der Ausstellungen mit bzw. ohne gemeinsames Auftreten. Als statistisches Verfahren haben wir den Vierfelder-Chi<sup>2</sup>-Test benutzt.

<24>

Bei der Analyse sind wir bei beiden phasenspezifischen Teilkorpora wie folgt vorgegangen: Zunächst wurde für jeden Künstler, der im Korpus mindestens sieben Mal vorkam, ein Kollokationsprofil erstellt. Nur Kollokatoren, die mindestens drei Mal auftraten, wurden auf ihre Signifikanz geprüft. Da dieses Verfahren nur Informationen über die Typizität des gemeinsamen Ausstellens von je zwei Künstlern liefert, uns aber an der Gruppierung der Künstler gelegen war, haben wir in einem zweiten Schritt Kollokationsnetze gebildet. Hierfür wurden aus allen künstlerspezifischen Kollokationslisten die n signifikantesten Verbindungen herausgefiltert.

<25>

Während sich Informationen über die Kollokationen zu einer einzigen Basis leicht als Liste oder Tabelle überblicken lassen, ist die Analyse von Kollokationsnetzen, also Informationen über die Kollokationen mehrerer Basen, die selbst Kollokationen anderer Basen sein können, bei einer nennenswerten Menge an Verbindungen ohne Visualisierung nicht mehr möglich. Basis-Kollokationen-Verbindungen und Kollokationsnetze lassen sich aber beispielsweise als gerichtete Graphen visualisieren, eine Darstellungsform, der wir uns in unserem Projekt ebenfalls bedienen. Graphen sind – einfach gesprochen – eine Menge von Punkten (»Knoten«/»vertices«/»nodes«), zwischen denen Linien (»Kanten«/»edges«) verlaufen. In unserem Projekt repräsentieren die Punkte Namen von Künstlern, die Kanten signifikante Verbindungen zwischen der Ausstellungstätigkeit zweier Künstler. Die Richtung des Graphen zeigt an, ob Künstler A für Künstler B, Künstler B für Künstler A oder beide füreinander signifikante Kollokationen sind.

<26>

Es gibt freilich unzählige Möglichkeiten einen Graphen mit mehreren Knoten und Kanten zu zeichnen.<sup>19</sup> Insbesondere für komplexe Netze haben Informatiker und Mathematiker verschiedene Algorithmen entwickelt, die automatisiert das Layout eines Graphen erzeugen. Da über die Struktur der Graphen in unserem Projekt nichts vorgängig bekannt ist, wir

beispielsweise keine Hierarchien oder Verwandtschaften voraussetzen können, bietet sich die Verwendung einer Layout-Methode an, die prinzipiell für alle Arten von Graphen geeignet ist. Die Wahl fiel auf die forced-based-Methode (Kräfte-Diagramm), die einen Graphen als ein physikalisches System modelliert, das von Kräften (forces) zwischen den Knoten gebildet wird.<sup>20</sup> Weil wir teilweise Graphen mit mehr als 10.000 Kanten setzen mussten, haben wir auf einen Force-directed-Algorithmus zurückgegriffen, der sehr große Graphen in hoher Geschwindigkeit layoutet und über einen speziellen Algorithmus zur Verhinderung des Überschreibens von Knoten verfügt: den Scalable Force Directed Placement-Algorithmus.<sup>21</sup>

<27>

Die Implementierung von Algorithmen zur Sicherung der Lesbarkeit der Graphen hat zur Folge, dass nicht alle Aspekte eines Graphen als Visualisierung relevanter Informationen gelesen werden können. So sind beispielsweise die Kantenlängen und die Winkel zwischen zwei Kanten nicht bedeutungstragend. Der Algorithmus sorgt jedoch dafür, dass zwei oder mehr Knoten, die mit einer gleichen oder sehr ähnlichen Menge anderer Knoten verbunden sind, in relativer Nähe zueinander dargestellt werden. Verdichtungsgebiete im Graphen lassen also auf eine Gruppe von Knoten mit ähnlichen Vernetzungsmerkmalen schließen. Für die Auswertung der Graphen vor dem Hintergrund unserer Fragestellung ist dies insofern relevant, als Verdichtungsgebiete als Indikatoren für Gruppierungen gelesen werden können, denn sie verweisen auf eine Menge von Künstlern, die mit einer jeweils ähnlichen Menge an Künstlern signifikant häufig gemeinsam ausgestellt haben.

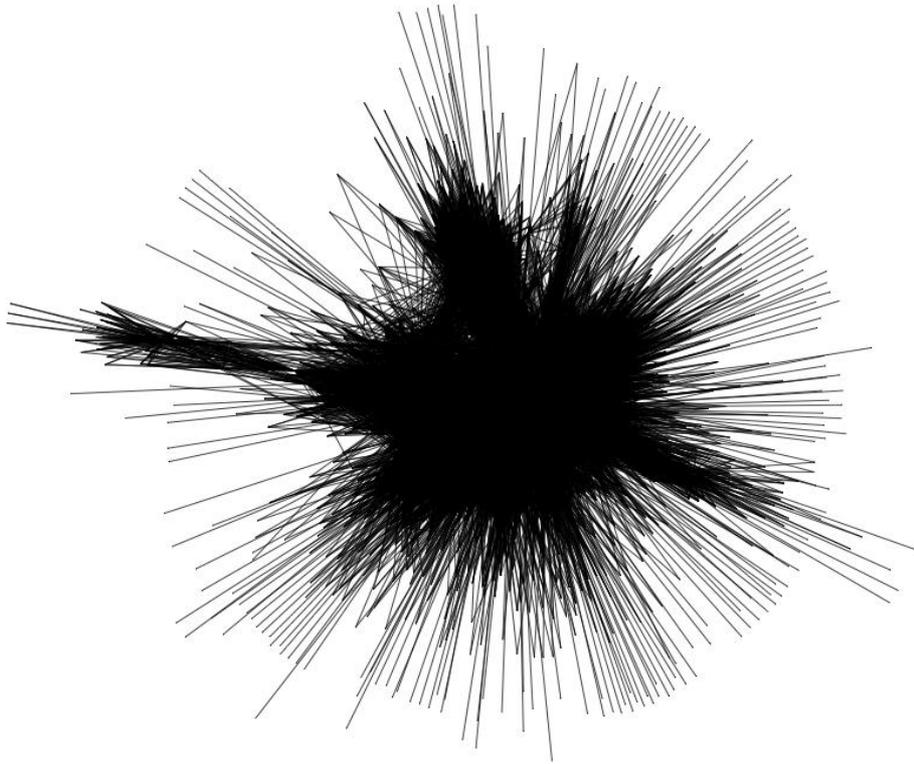


Abb. 2: Graph mit allen signifikanten Verbindungen für den gesamten Zeitraum

<28>

Die Kollokationsanalyse wurde mittels eines selbst geschriebenen Perl-Scripts durchgeführt, ebenso wie die Überführung der Einzelergebnisse in ein Kollokationsnetz. Das Setzen des Graphen erfolgte mit Hilfe des GraphViz-Software-Pakets, in das der sfdp-Algorithmus implementiert ist.<sup>22</sup> Ein Graph (Abb. 2), der sämtliche signifikante Verbindungen auf einem Signifikanzniveau von  $p < 0.05$  in der Phase 2 enthält, ist jedoch trotz sfdp-Algorithmus unübersichtlich und kaum lesbar, da die Verdichtungsgebiete zu stark in einander übergehen. Er hat 2.246 Knoten (Künstlernamen) und 15.749 Kanten (signifikante Verbindungen). Um die Darstellung luzider zu machen, haben wir uns dafür entschieden, die Anzahl der Kanten zu begrenzen und zwar nach dem Kriterium der Höhe des Chi2-Wertes. Wir haben die Anzahl der Kanten Schritt für Schritt reduziert (Abb. 3), indem wir nur noch jene Verbindungen mit den 7000, 6000, 5000 etc. höchsten Chi2-Werten berücksichtigt haben. Durch die Filterung nach dem Kriterium des Chi2-Wertes ergeben sich Graphen, in denen die Verdichtungsgebiete unterschiedlich stark von einander abgegrenzt auftreten. Je kleiner man die Anzahl der Kanten wählt, desto mehr werden die Verdichtungsgebiete auf ihre Kerne begrenzt oder verschwinden. Auf diese Weise ist es möglich, unterschiedliche Festigkeitsstufen sowie zentrale und periphere Regionen von Verdichtungsgebieten zu

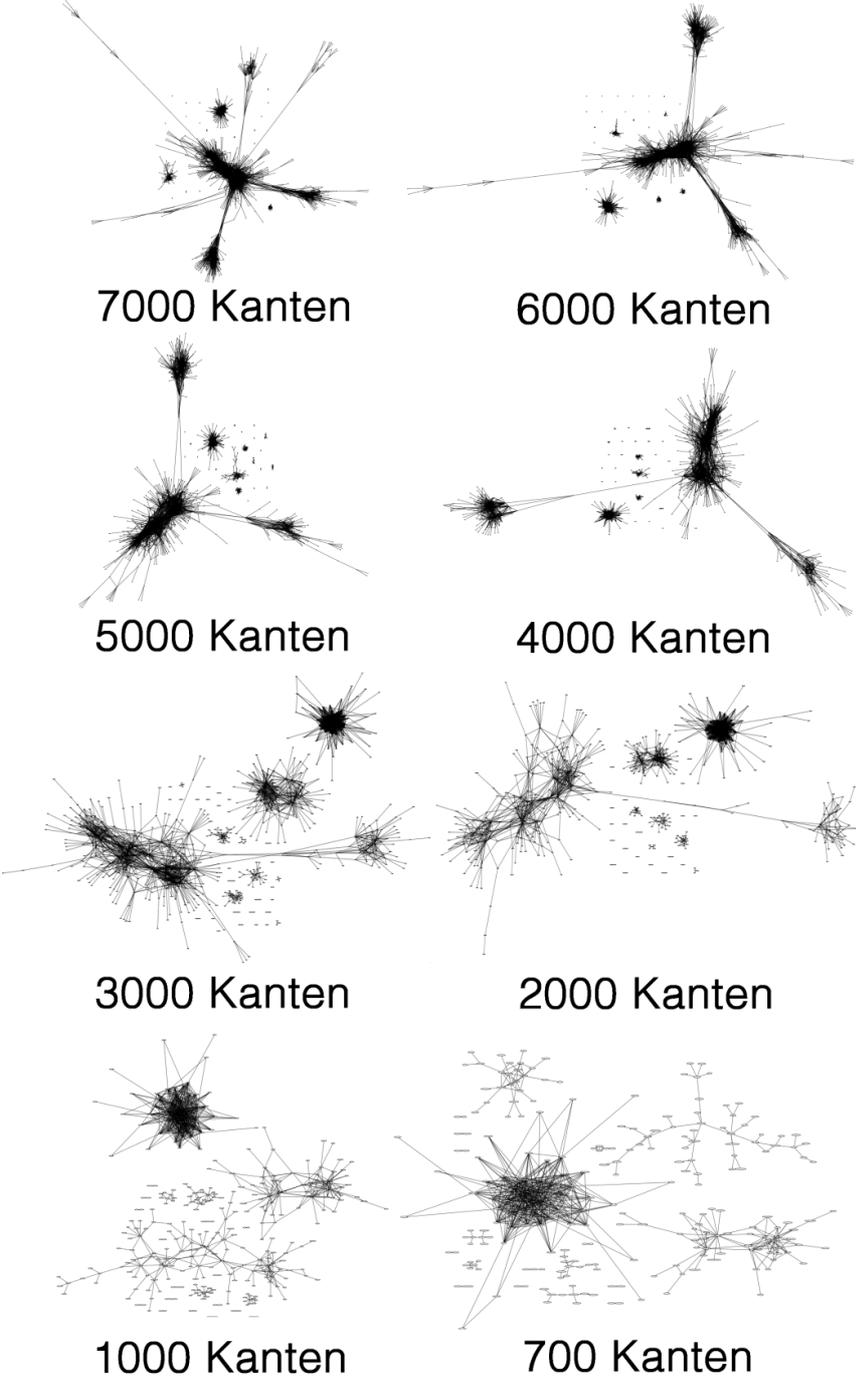


Abb. 3 Schrittweise Reduzierung der Anzahl der Kanten

identifizieren. Die Analyse von als Graphen visualisierten Kollokationsnetzen ermöglicht eine grobe Gruppierung der Daten. Differenziertere Ergebnisse sind mittels Clusteranalysen möglich.

### **3.2 Clusteranalyse: Klassenbildung als Dendrogramm**

<29>

Die Clusteranalyse gehört zu den strukturentdeckenden Verfahren der multivariaten Statistik. Sie dient dazu, die Objekte eines unklassifizierten Merkmalsdatensatzes in Gruppen einzuteilen, die in sich möglichst homogen sind und sich von den anderen Gruppen möglichst stark unterscheiden.<sup>23</sup> Sie ist daher eine geeignete Methode für unsere Fragestellung. Wir haben auf der Basis der Datenmatrix eine hierarchische Clusteranalyse (averagelinkage) mit dem Statistikprogramm R durchgeführt.<sup>24</sup> Die Ergebnisse von Clusteranalysen können auf unterschiedliche Weise visualisiert werden. Die am häufigsten verwendete Methode der Visualisierung der Ergebnisse hierarchischer Clusteranalysen ist das Dendrogramm. Dendrogramme sind Baumgraphen, die die Zerlegung einer Datenmenge in immer kleinere Teilmengen darstellen. Die äußersten Enden des Graphen (»Blätter«) bestehen aus Clustern, die je ein einzelnes Objekt der Datenmenge umfassen; die »Wurzel« aus einem Cluster, das sämtliche Objekte der Datenmenge umfasst. Ein Knoten eines Dendrogramms repräsentiert die Vereinigung aller seiner Kindknoten. Die Länge der Kanten zwischen zwei Knoten bildet die Distanz zwischen den beiden Mengen ab.<sup>25</sup> Zwar können mit R auch Graphen gezeichnet werden; jedoch sind keine Algorithmen zur Verhinderung von Überschneidungen von Knoten und Kanten implementiert, was dazu führt, dass große Graphen unlesbar werden. Um das Dendrogramm mit seinen 10.014 Blättern dennoch lesbar zu machen, haben wir auf eine Graphen-Software zurückgegriffen, die für die Visualisierung biomedizinischer Daten, insbesondere der Genomanalyse entwickelt wurde: die Visualisierungssoftware Dendroscope.<sup>26</sup> Die mittels Dendroscope gesetzten Dendrogramme können als Hilfsmittel für die Identifizierung von Gruppen benutzt werden, wobei das Mindestabstandsmaß je nach Erkenntnisinteresse variiert werden kann. So können, je nachdem, wie fein die Analysen ausfallen sollen, alle der im Dendrogrammausschnitt in Abbildung 4 enthaltenen Künstlernamen als Angehörige einer einzigen Gruppe oder von fünf unterschiedlichen Gruppen aufgefasst werden.

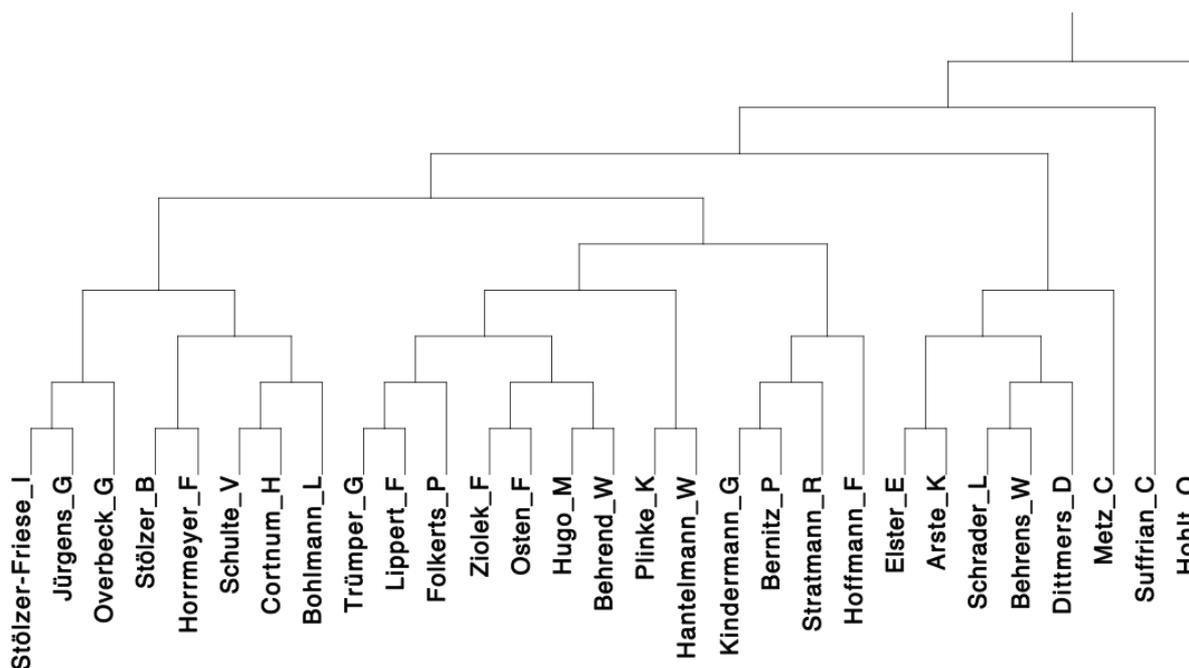


Abb. 4: Ausschnitt aus dem Dendrogramm aus Phase 2

## 4 Datenanalyse

### 4.1 Interpretationsschritte

<30>

Für unser spezifisches Datenmaterial – die Kataloge von Gruppenausstellungen zur Gegenwartskunst in der NS-Zeit – ermöglichen es die Kollokations- und die Clusteranalyse, Gruppen von Künstlern zu identifizieren, deren Werke häufig und typischerweise miteinander ausgestellt wurden. Aus der Häufigkeit und Typizität gemeinsamer Ausstellungspräsenz – das ist die Prämisse unserer Interpretation – lassen sich Rückschlüsse auf die Wirksamkeit der nationalsozialistischen Kunst- und Ausstellungspolitik ziehen. Wenn wir z.B. sehen, dass die mit Staatsaufträgen bedachten Künstler und die als »entartet« diffamierten modernen Künstler jeweils eigene Gruppen bilden, wenn wir lokale, regionale oder überregionale Gruppenbildungen erkennen und bei einer zeitlich differenzierten Betrachtung Veränderungen feststellen können, so kann dies im Hinblick auf normierende und segregierende Tendenzen und Effekte der staatlichen Kunst- und Ausstellungspolitik durchaus signifikant sein. Die kunstpolitische Signifikanz der Gruppenbildungen und ihrer Veränderungen vorausgesetzt, geht es bei der Analyse vor allem darum, die Gruppen zu bestimmen und zu qualifizieren.

<31>

Die Kollokationsanalyse und die Clusteranalyse bieten in ihren visualisierten Formen, den Kollokationsnetzen und den Dendrogrammen, unterschiedliche Ansatzpunkte für eine Interpretation. Die Kollokationsnetze mit ihren Verdichtungen, aber auch mit ihrer auffälligen zentripetalen Gestalt, die auf ein signifikantes Verhältnis von Kern und Randbereich schließen lässt, sind – bei einer Datenmenge von mehreren Tausend Künstlernamen – übersichtlicher als die Dendrogramme und daher vor allem geeignet, mit einem ersten Blick grundlegende Muster und Gruppenbildungen zu erkennen. Die Dendrogramme dagegen definieren die Gruppen präziser und eignen sich deshalb insbesondere zur Überprüfung und Verifizierung der Interpretation der Kollokationsnetze. Es wurden also in einem ersten Schritt die Kollokationsnetze nach visuellen Auffälligkeiten überprüft, das heißt die Verdichtungen in den Blick genommen, das Verhältnis der Verdichtungen zueinander sowie das Verhältnis von Zentrum und Peripherie analysiert und in einem zweiten Schritt dann die Befunde anhand der Dendrogramme überprüft.

<32>

Um die auffälligen Verdichtungen genauer in den Blick zu nehmen, wurden die entsprechenden Stellen in den Kollokationsnetzen so weit vergrößert, bis die Knoten, das heißt die einzelnen Künstlernamen samt ihrer unmittelbaren Umgebung lesbar wurden. Dann ging es darum zu ermitteln, was die Künstler der verdichteten Bereiche über ihre gemeinsame Ausstellungstätigkeit bzw. ihre gemeinsamen Kollokatoren hinaus miteinander verbindet. Eine Schwierigkeit bestand darin, dass über die meisten der mehr als 10.000 Künstler, die in der NS-Zeit ausgestellt haben, kaum etwas bekannt ist. Nur von einem Bruchteil dieser Künstler kennen wir grundlegende biographische Daten wie Geburtsdatum und Geburtsort, Daten zur künstlerischen Ausbildung und zum künstlerischen Werdegang, zum Tätigkeitsort und zu Gruppen-, Verbands- und Institutionszugehörigkeiten.<sup>27</sup> Gruppendifinitionen konnten deshalb nur über die bekannten Künstler erfolgen. Sie waren die Ankerpunkte bei der Interpretation der Kollokationsnetze. Es ging also darum, die Verdichtungen in den Netzen zunächst näher zu bestimmen und voneinander abzugrenzen, wobei die Graphen mit reduzierter Kantenzahl hilfreich waren, und die identifizierten Gruppen dann durch das Auffinden bekannter, das heißt in der Regel oft ausgestellter Künstler näher zu qualifizieren. Bei der Auswahl dieser Künstler half die quantitative Analyse, die bei der ursprünglichen Erhebung der Daten vorgenommen worden war.<sup>28</sup> Untersucht wurden nicht nur die Verdichtungen nach bekannten Künstlernamen, sondern umgekehrt auch die Liste der bekannten, oft ausgestellten Künstler nach ihrer Position im Kollokationsnetz. Stichprobenartig wurden auch die Positionen der als »entartet« diffamierten Künstler im Netz überprüft.

<33>

Neben den Verdichtungen galt es, das Verhältnis von Zentrum und Peripherie im Blick zu behalten. Die Nähe eines Künstlers oder eine Gruppe zum Zentrum eines Kollokationsnetzes ist Ausdruck häufiger Verbindungen zu anderen Künstlern und Künstlergruppen. Ein Künstler, der zentrumsnah, aber außerhalb der verdichteten Bereiche positioniert ist, hat häufig und breit ausgestellt. Eine Verdichtung im Zentrum bedeutet, dass diese Gruppe häufiger ausgestellt hat als die Gruppen an der Peripherie. Es ist also durchaus möglich, von der zentralen oder peripheren Position der Künstler und Künstlergruppen in den Kollokationsnetzen auf eine im übertragenen Sinne zentrale oder periphere Rolle im Ausstellungsbetrieb rückzuschließen.

<34>

Die Interpretation der Kollokationsnetze und der Dendrogramme wurde jeweils für den gesamten Zeitraum (1933–1945) sowie für die Zeiträume 1933–1937 (Phase 1) und 1938–1945 (Phase 2) vorgenommen. Auf diese Weise konnte einmal das Ausstellungswesen zur Gegenwartskunst in der NS-Zeit insgesamt in den Blick genommen und zum anderen mögliche Veränderungen, die sich als Folge der nationalsozialistischen Kunst- und Ausstellungspolitik, insbesondere der beiden großen Ausstellungsereignisse von 1937, einstellen, erfasst und interpretiert werden.

## **4.2 Ergebnisse**

<35>

Bei der Analyse des Kollokationsnetzes für den gesamten Zeitraum von 1933 bis 1945 (vgl. Abb. 2) zeigte sich, dass die erkennbaren, großen Verdichtungen, die sich im Zentrum überlagern, in der Regel regionale Kunstszenen abbilden. Das heißt, die Verdichtungen enthalten jeweils die Namen von Künstlern, die einer gemeinsamen Region zuzuordnen sind. Identifizierbar waren z.B. die Städte bzw. Regionen Berlin, das Rheinland, München oder Hannover. Offenbar wurden die Künstler typischerweise zusammen mit anderen Künstlern aus ihrer Stadt oder ihrer Region ausgestellt. Dieser Befund überrascht nicht, ist aber ein Beleg für eine tendenziell regionale Orientierung des Ausstellungswesens im Nationalsozialismus, wenn man den gesamten Zeitraum von 1933 bis 1945 zugrunde legt.

<36>

Das Bild ändert sich, wenn man die Zeiträume 1933 bis 1937 (Phase 1) und 1938 bis 1945 (Phase 2) einzeln betrachtet und insbesondere die kantenreduzierten Graphen (Abb. 5) in den Blick nimmt. In der ersten Phase ordnen sich die Verdichtungen an der Peripherie an und weisen kaum Verbindungen untereinander auf, das heißt die regionale Orientierung des Ausstellungswesens war offenbar in den ersten Jahren des Nationalsozialismus

ausgeprägter, das Ausstellungswesen eher dezentral als zentralistisch strukturiert. Das Kollokationsnetz zur Phase 2 zeigt dagegen eine starke Verdichtung, von der Verbindungen zu kleineren Verdichtungen an den Rändern bestehen. Hier zeigen sich die Effekte der auf Vereinheitlichung ausgerichteten nationalsozialistischen Kunst- und Ausstellungspolitik sehr deutlich. Die starke Verdichtung im Zentrum des Graphen kann als ein Indikator für eine Homogenisierung des Ausstellungswesens, für eine Auflösung der regionalen Strukturen und damit für eine tendenziell vereinheitlichende, d.h. nationale Entwicklung gelesen werden.

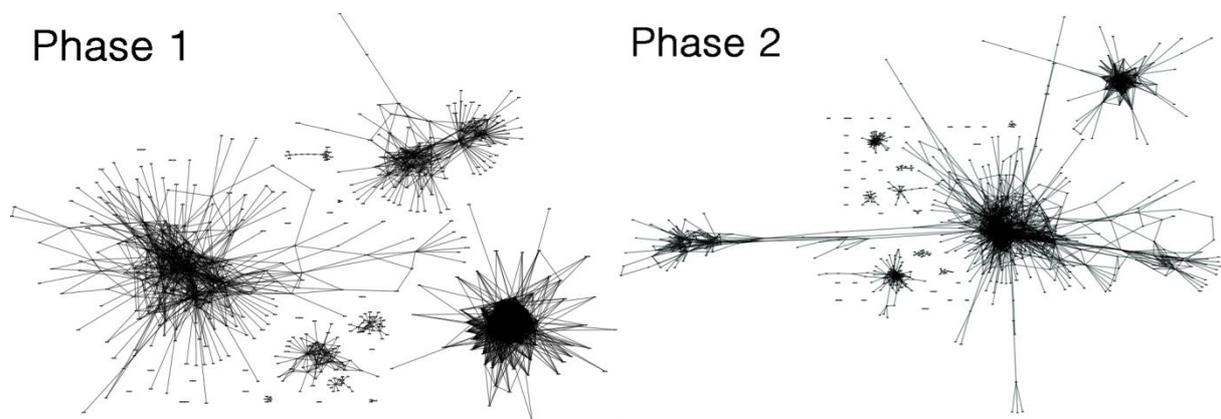


Abb 5.: Phasenspezifische Kollokationsnetze mit je 3000 Kanten

<37>

Neben den unterschiedlichen Graden der Verdichtung ist die Position der Verdichtungen, insbesondere ihre Lage zum Zentrum des jeweiligen Netzes, ein wichtiges Element des visuellen Befundes. So ist es auffällig, dass im nicht-reduzierten Kollokationsnetz der ersten Phase (o. Abb.) die Verdichtungen, die die künstlerischen Szenen Berlins und des Rheinlandes abbilden, näher zum Zentrum stehen als die Szenen der Dresdner, der Hannoveraner oder auch der Münchner Künstler. Dies kann als ein Hinweis darauf interpretiert werden, dass die Bemühungen der nationalsozialistischen Kunstpolitik, München als neues künstlerisches Zentrum zu etablieren und Berlin, das künstlerische Zentrum der Weimarer Zeit, in dieser Funktion abzulösen, bis 1937 noch nicht gegriffen hatten. Interpretiert man das Kollokationsnetz der zweiten Phase richtig, verloren die Kunstzentren nach 1937 ihre regionale Prägung. Dies gilt nicht zuletzt für München.

<38>

Bei den Interpretationen der Kollokationsnetze, die ja nicht mehr als das gemeinsame Auftreten von Künstlern in Ausstellungen auf der Basis der erhaltenen Kataloge abbilden, ist

immer zu berücksichtigen, dass die Ausstellungskataloge, die die Datenbasis bilden, nur einen Teil der tatsächlich stattgefundenen Ausstellungen repräsentieren. Hochrechnungen der Ausstellungsankündigungen aus dem Mitteilungsblatt der Reichskammer der bildenden Künste zeigen, dass nur zu jeder zehnten Ausstellung auch ein Katalog erschienen ist.<sup>29</sup> Die erfassten Ausstellungskataloge repräsentieren also nur ein Zehntel aller Gruppenausstellungen zur Gegenwartskunst, die in den Jahren 1933 bis 1945 in Deutschland stattgefunden haben, allerdings wohl das hinsichtlich ihrer orientierenden Funktion bedeutendste und öffentlichkeitswirksamste Zehntel. Stellt man die Ausstellungen ohne Katalog mit in Rechnung, so war das Ausstellungswesen insgesamt und über den gesamten Zeitraum von 1933 bis 1945 betrachtet vermutlich wesentlich provinzieller und auch dezentraler strukturiert, als es die auf der Basis der Kataloge erstellten Kollokationsnetze den Anschein hat.

<39>

Bis 1937 konnten viele der öffentlich diffamierten und ihrer Ämter enthobenen modernen Künstler, wenn sie nicht jüdischer Herkunft oder politisch auffällig waren, trotz Arbeitsverbot weiter ausstellen, wenn auch in weitaus geringerem Umfang als zuvor. Im Kollokationsnetz der ersten Phase besetzen diese Künstler – meist prominente Mitglieder des früheren Deutschen Künstlerbundes wie Ernst Barlach, Käthe Kollwitz oder Max Beckmann – einen Bereich in der Nähe des Zentrums, der nicht den Grad der Verdichtung einer regionalen Szene erreicht, aber dennoch auf eine eigene Gruppe schließen lässt. Das Dendrogramm bestätigt die Clusterbildung (vgl. Abb. 6). Die prominenten Staatskünstler bilden in dieser Phase noch kein eigenes Cluster aus, sondern sind noch Teil regionaler Gruppen.

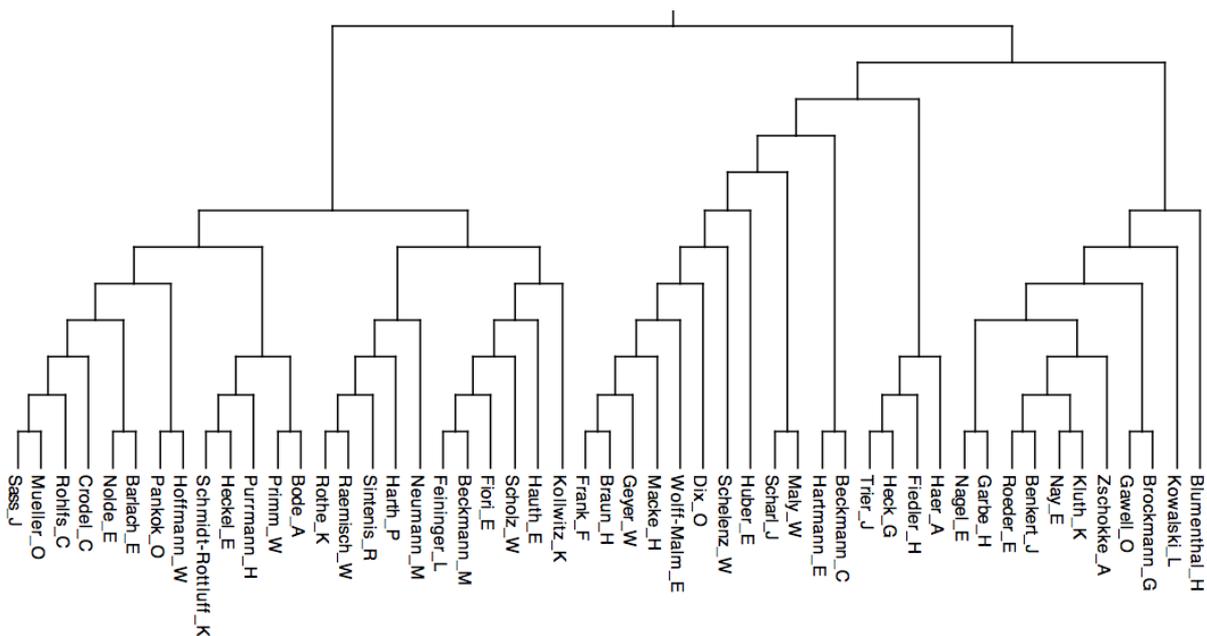


Abb. 6: Ausschnitt Dendrogramm Phase 1 mit Cluster »entartete Künstler«

<40>

Dies ändert sich nach 1937. Insbesondere die prominenten, mit Staatsaufträgen bedachten Bildhauer wie Arno Breker, Georg Kolbe, Richard Scheibe, Fritz Klimsch, Anton Grauel und Richard Knecht bilden nun in den Ausstellungen eine eigene Gruppe, wie insbesondere der Blick auf das Dendrogramm der zweiten Phase zeigt, in dem diese Künstler ein eigenes Cluster bilden (vgl. Abb. 7). Dies unterstreicht die programmatische Rolle, die insbesondere den Monumental-Bildhauern seitens der staatlichen Kunstpolitik zugewiesen wurde. Sie bestand darin, stilbildend auf die künstlerische Entwicklung einzuwirken und eine nationale künstlerische Identität herzustellen.

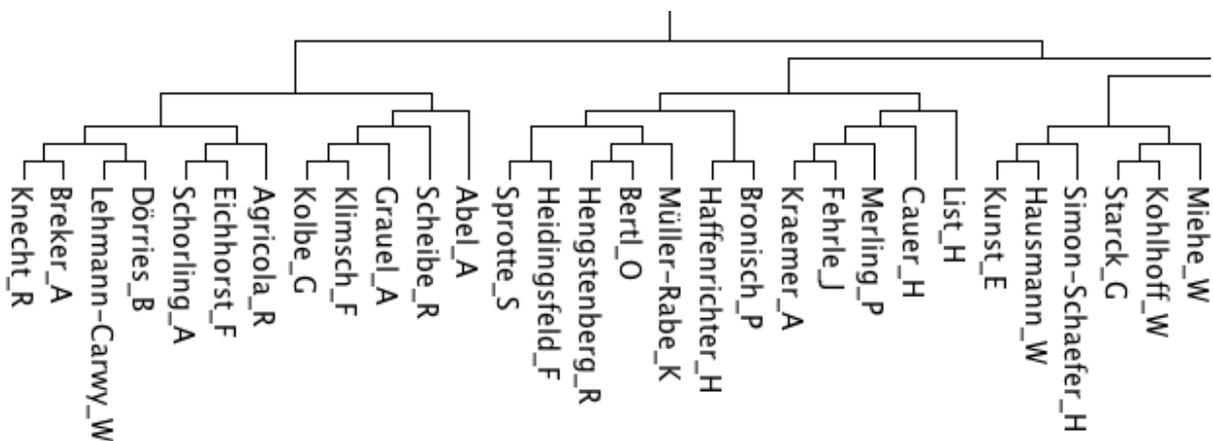


Abb. 7: Ausschnitt Dendrogramm Phase 2 mit Cluster Breker/Kolbe/Scheibe

## 5 Fazit

<41>

Die Ergebnisse der Kollokations- und Clusteranalyse bestätigen, differenzieren und ergänzen die bisherigen Annahmen zur Ausstellungspolitik im Nationalsozialismus. Die Kollokationsnetze und Dendrogramme bilden die Auswirkungen der kunstpolitischen Maßnahmen, insbesondere des Jahres 1937, deutlich ab. Sie dokumentieren den Wandel von einem dezentral zu einem stärker zentral orientierten Ausstellungswesen und geben Hinweise auf die Bedeutung und die Funktion einzelner Künstlergruppen in diesem Prozess.

<42>

Um die Validität und Prägnanz der Analyse noch zu erhöhen, wäre es notwendig, eine biographische Datenbank zu den Künstlern im Nationalsozialismus zu erarbeiten, die eine genauere Bestimmung der in den Kollokationsnetzen und Dendrogrammen sichtbaren Künstlergruppen ermöglicht. Eine solche Datenbank, die detaillierte Informationen zur Biographie und zum künstlerischen Werdegang, aber auch zu Verbindungen und

Mitgliedschaften in politischen und institutionellen Einrichtungen enthielte, erlaubte eine differenzierte Identifizierung der abgebildeten Gruppen, als sie bisher möglich ist, eine Identifizierung der Künstlergruppen, die weit über die Bestimmung ihrer regionalen Tätigkeitsfelder und ihrer Akzeptanz innerhalb des Kunstbetriebs hinausginge. Es wäre z.B. möglich, nach dem Einfluss von Akademien und Schulen, der Partei(en), der Kunstvereine oder der Kirchen auf die Gruppenbildung der Künstler zu fragen.

<43>

Denkbar ist darüber hinaus, die Datenbasis zu erweitern, und zwar einmal um die Kunstaussstellungen, die seit Kriegsbeginn in den von Deutschland besetzten Gebieten durchgeführt wurden,<sup>30</sup> und zum anderen durch die Ausstellungen moderner Kunst, d.h. die Ausstellungen »Entarteter Kunst« in Deutschland einerseits<sup>31</sup> und die Kunstaussstellungen geflüchteter Künstler in den Exilländern andererseits.<sup>32</sup> Mit diesen Erweiterungen wären alle Ausstellungen deutscher Gegenwartskunst in den Jahren 1933 bis 1945 erfasst, ein wichtiges Kapitel der Ausstellungsgeschichte des 20. Jahrhunderts in größtmöglichem Umfang und dabei kunst- und ausstellungspolitisch differenziert dokumentiert, und es wäre möglich, die Unterschiede, aber auch die Verbindungen zwischen den verschiedenen Feldern des Ausstellungswesens aufzuzeigen.

<44>

Die datengeleitete Analyse, wie wir sie hier vorgestellt und erprobt haben, bietet für die Kunstwissenschaft eine Vielzahl neuer Erkenntnismöglichkeiten, die weit über die Auswertung von Katalogen, weit über das Feld der Ausstellungsgeschichte hinausreichen. Wenn man bedenkt, dass nicht nur bio-bibliographische, sondern auch kunstwerkbezogene Daten (Titel, Datierung, Ort, Technik, Stil, Ikonographie usw.) einer Kollokations- und Clusteranalyse unterzogen werden können, lässt sich erahnen, welche Perspektiven diese Analyseverfahren für die Kunstgeschichte eröffnen kann.

### **Bildnachweis:**

Copyright für alle Abbildungen bei den Autoren

### **Zu den Autoren:**

PD Dr. Martin Papenbrock, Studium der Kunstgeschichte, Literatur- und Editionswissenschaft an der Universität Osnabrück, 1991 Promotion, 1999 Habilitation, seit 1998 am Institut für Kunstgeschichte der Universität Karlsruhe (TH), derzeit Lehrstuhlvertretung im Fachgebiet Kunstgeschichte am Karlsruher Institut für Technologie (KIT), Forschungsschwerpunkte: Niederländische Malerei der frühen Neuzeit, Kunst und Politik im 20. Jahrhundert, Theorie- und Fachgeschichte der neueren Kunstwissenschaft.

Anschrift: Karlsruher Institut für Technologie (KIT), Institut für Kunst- und Baugeschichte, Englerstraße 7, D-76131 Karlsruhe, [martin.papenbrock@kit.edu](mailto:martin.papenbrock@kit.edu)

Prof. Dr. Joachim Scharloth, Studium der Germanistik, Politikwissenschaft und Philosophie an den Universitäten Mainz und Heidelberg, Promotion an der Universität Heidelberg (2002), Habilitation an der Universität Zürich (2008). Seit 2010 Professur für Germanistische Linguistik an der Dokkyo-Universität in Tokyo. Forschungsschwerpunkte: Korpuslinguistik, Soziolinguistik, Linguistische Pragmatik, Sprachgeschichte.

Anschrift: Dokkyo University, Department of German, Faculty of Foreign Languages, Gakuen-cho 1-1, Soka-shi, Saitama-ken, 340-0042, Japan, [scharloth1@dokkyo.ac.jp](mailto:scharloth1@dokkyo.ac.jp)

---

1 <http://www.informartics.com/>

2 Vgl. Ulrik Brandes, Linton C. Freeman, Dorothea Wagner, Dorothea: Social Networks, in: Roberto Tamassia (Hg.): Handbook of Graph Drawing and Visualization, Boca Raton, London (im Druck).

3 Vgl. Hildegard Brenner: Die Kunstpolitik des Nationalsozialismus, Reinbek bei Hamburg 1963, S. 112–114; Reinhard Merker: Die bildenden Künste im Nationalsozialismus. Kulturideologie, Kulturpolitik, Kulturproduktion, Köln 1983, S. 163–168; Mario-Andreas von Lüttichau: »Deutsche Kunst« und »Entartete Kunst«: Die Münchner Ausstellungen 1937, in: Klaus-Peter Schuster (Hg.): Nationalsozialismus und »Entartete Kunst«. Die »Kunststadt« München 1937, München 1987, S. 83–118; Peter Adam: Kunst im Dritten Reich, Hamburg 1992, S. 92–119; Ines Schlenker: Die »Große Deutsche Kunstausstellung« und ihre Auswirkungen auf den nationalsozialistischen Kunstbetrieb, in: Hans-Jörg Czech, Nikola Doll (Hg.): Kunst und Propaganda. Im Streit der Nationen 1930–1945, Dresden 2007, S. 258-267.

4 Vgl. Mario-Andreas von Lüttichau: Rekonstruktion der Ausstellung »Entartete Kunst«, in: Schuster 1987 (wie Anm. 3), S. 120–181; Stephanie Barron (Hg.): »Entartete Kunst«. Das Schicksal der Avantgarde im Nazi-Deutschland, Berlin 1992; Christoph Zuschlag: »Entartete Kunst«. Ausstellungsstrategien im Nazi-Deutschland, Worms 1995. Hinzuweisen ist vor allem auch auf die Arbeit der *Forschungsstelle »Entartete Kunst«*, die am Kunsthistorischen Institut der Freien Universität Berlin und am Kunsthistorischen Seminar der Universität Hamburg betrieben wird. Vgl. das Beschlagnahmeinventar »Entartete Kunst« auf [http://www.geschkult.fu-berlin.de/e/db\\_entart\\_kunst/datenbank/index.html](http://www.geschkult.fu-berlin.de/e/db_entart_kunst/datenbank/index.html).

5 Vgl. Ines Schlenker: Hitler's Salon. The »Große Deutsche Kunstausstellung« at the Haus der Deutschen Kunst in Munich 1937–1944, Oxford et al. 2007; Schlenker 2007 (wie Anm. 3); Robert Thoms: Große Deutsche Kunstausstellung München 1937–1944. Verzeichnis der Künstler in zwei Bänden. Bd. 1: Maler und Graphiker, Berlin 2010, Bd. 2: Bildhauer (im Druck) sowie das Projekt *GDK-online* des Zentralinstituts für Kunstgeschichte in München und der Kunstbibliothek der Staatlichen

---

Museen zu Berlin/Preußischer Kulturbesitz auf  
<http://www.arthistoricum.net/ressourcen/monographien/gdk/>.

6 Vgl. Thomas Brandt: Sieben Jahre DISKUS (Digitales Informationssystem für Kunst- und Sozialgeschichte). Ein Arbeitsbericht ([http://www.edvtage.de/vortrag.php?kapitel=1999\\_05](http://www.edvtage.de/vortrag.php?kapitel=1999_05)).

7 Vgl. Martin Papenbrock, Gabriele Saure (Hg.): Kunst des frühen 20. Jahrhunderts in deutschen Ausstellungen. 2 Bde. Teil 1: Ausstellungen deutscher Gegenwartskunst in der NS-Zeit. Eine kommentierte Bibliographie, Weimar 2000 (= Schriften der Guernica-Gesellschaft 10).

8 Vgl. Thomas A. Runkler: Data Mining. Methoden und Algorithmen intelligenter Datenanalyse, Wiesbaden 2010, S. 2.

9 Vgl. Runkler 2010 (wie Anm. 8), S. 1 f.

10 Vgl. Donato Malerba, Margherita Berardi, Michelangelo Ceci: Discovering Spatio-Textual Association Rules in Document Images, in: Pascal Poncelet, Maguelonne Teisseire, Florent Maguelonne (Hg.): Data Mining Patterns: New Methods and Applications, Hershey PA, London 2008, S. 176–197.

11 Vgl. Dominique Estival: Authorattribution with email messages, in: Journal of Science, Vietnam National University, 2008, S. 1–9.

12 Vgl. Simona Ester Rombo, Luigi Palopoli: Pattern Discovery in Biosequences: From Simple to Complex Patterns, in: Poncelet, Teisseire, Maseglia 2008 (wie Anm. 10), S. 85–105.

13 Vgl. Hsinchun Chen, Wingyan Chung, Yi Qin, Michael Chau, Michael, Jennifer Jie Xu, Gang Wang, Rong Zheng, Homa Atabakhsh: Crime Data Mining: An Overview and Case Studies, in: Proceedings of the 2003 annual national conference on digital government research. Digital Government Society of North America, 2003, S. 1–5.

14 Vgl. Hadumod Bußmann: Lexikon der Sprachwissenschaft, 4. durchgesehene Auflage, Stuttgart 2008.

15 Vgl. Lothar Lemnitzer, Heike Zinsmeister: Einführung in die Korpuslinguistik. Tübingen 2006.

16 (Hg.): Corpus Linguistics: An International Handbook. Vol. 2. 2009 (=Handbooks of linguistics and communicationscience; 29.2), S. 1212–1248.

17 Vgl. Evert 2009 (wie Anm. 16), S. 1212 f.

18 Für eine weiterführende Darstellung in Frage kommender statistischen Verfahren vgl. Evert 2009 (wie Anm. 16), S. 1224–1236.

19 Eine Übersicht bieten neben den Standardwerken: Michael Kaufmann, Dorothea Wagner (Hg.): Drawing Graphs. Methods and Models, Berlin, Heidelberg, New York et al. 2001; Tamassia (im Druck) (wie Anm. 2).

20 Vgl. Rudolf Fleischer, Colin Hirsch: Graph Drawing and Its Applications, in: Kaufmann/Wagner 2001 (wie Anm. 19), S. 1–22, hier S. 20.

21 Vgl. Yi Fan Hu: Efficient and high quality force-directed graph drawing, in: The Mathematica Journal, 10/2005, S. 37–71.

22 <http://www.graphviz.org>; J. Ellson, E. R. Gansner, E. Koutsofios, S. C. North, G. Woodhull: Graphviz and Dynagraph – Static and Dynamic Graph Drawing Tools, in: M. Junger, P. Mutzel (Hg.): Graph Drawing Software. 2003, S. 127–148.

23 Vgl. Runkler 2010 (wie Anm. 8), S. 105.

24 R statisticspackage: [www.r-project.org/](http://www.r-project.org/); Ross Ihaka, Robert Gentleman: R: A Language for Data Analysis and Graphics, in: Journal of Computational and Graphical Statistics, 5,3/1996, S. 299–314.

25 Zu den Layout-Prinzipien von hierarchischen Graphen vgl. Ralf Brockenauer, Sabine Cornelsen: Drawing Clusters and Hierarchies, in: Kaufmann/Wagner 2001 (wie Anm. 19), S. 193–227, hier S. 212.

26 <http://www-ab.informatik.uni-tuebingen.de/data/software/dendroscope/>; D. H. Huson, D. C. Richter, C. Rausch, T. Dezulian, M. Franz, R. Rupp: Dendroscope – An interactive viewer for large phylogenetic trees, in: BMC Bioinformatics, 8/2007, S. 460.

27 Ein kritisches Lexikon der Künstler im Nationalsozialismus gibt es bis heute nicht. Die meisten Informationen enthält Mortimer G. Davidson: Kunst in Deutschland 1933–1945. Eine wissenschaftliche Enzyklopädie der Kunst im Dritten Reich. 4 Bde., Tübingen 1988–1995.

28 Vgl. Papenbrock/Saure 2000 (wie Anm. 7), S. 43.

29 Vgl. ebd., S. 368–378.

30 Vgl. dazu Otto Thomae: Die Propaganda-Maschinerie. Bildende Kunst und Öffentlichkeitsarbeit im Dritten Reich, Berlin 1978, S. 37–76.

31 Vgl. Zuschlag 1995 (wie Anm. 4).

32 Vgl. Cordula Frowein (1986): Ausstellungsaktivitäten der Exilkünstler, in: Kunst im Exil in Großbritannien 1933–1945, Berlin 1986, S. 35–48; Stationen der Moderne. Die bedeutenden Kunstausstellungen des 20. Jahrhunderts in Deutschland, Berlin 1989.