

Analyse der Tags einer Kunst Folksonomie

Tag Analysis of an Art Folksonomy

Martin Weingartner, Max Arends, Josef Froschauer, Doron Goldfarb, Dieter Merkl
Institut für Softwaretechnik und Interaktive Systeme
Technische Universität Wien
Favoritenstraße 9-11/188, A-1040 Wien
Tel.: +43 1 58801 18815, Fax: +43 1 58801 18899
E-mail: vsem@ec.tuwien.ac.at, Internet: <http://vsem.ec.tuwien.ac.at/>

Zusammenfassung

Immer mehr Kulturinstitutionen verwenden Web 2.0 Methoden, um ihre Website für Besucher interessanter und informativer zu gestalten. In der von uns entwickelten Webplattform, *explorARTorium*¹ konnten BenutzerInnen unbeeinflusst von Informationen, nur anhand des Bildes selbst Tags hinzufügen. Die daraus entstandene Folksonomie wurde analysiert, um herauszufinden, welches Vokabular die BenutzerInnen für Ihre Tags verwenden, welches Vorwissen sie haben, welche Bildmotive sie erkennen und von welchen KünstlerInnen sie die Werke bevorzugt mit Tags versehen haben. Diese Erkenntnisse verwenden wir für die Planung künftiger Web 2.0 Anwendungen im kunsthistorischen Umfeld.

Abstract

More and more cultural institutions are using Web 2.0 methods in order to present their site in a more interesting and informative manner. We have developed the Web platform *explorARTorium*, where users are allowed to tag artworks while being unbiased by additional information, using only the image itself. The resulting folksonomy was analyzed to determine the individual user vocabularies used for tagging, which knowledge users have about art history, and which artists' works they preferred to tag. These findings can help to develop future Web 2.0 applications in the domain of art history.

Einleitung

Das Internet bietet Zugang zu einer großen Anzahl von Abbildungen von Kunstwerken. Die *Web Gallery of Art*² (WGA) bietet Zugriff auf mehr als 26.000 Bilder von Kunstwerken, die zwischen 1000 und 1900 n. Chr. geschaffen wurden. Mit *Google Art Project*³ können die BenutzerInnen mehrere Museen virtuell besuchen und dabei mehr als 1.000 Kunstwerke in hoher Auflösung bewundern. Viele Museen und Kulturinstitutionen stellen Bilder aus ihren Bildarchiven bei Projekten wie *Flickr Commons*⁴ zur Verfügung. Ein weiterer Trend der letzten Jahre war, dass zunehmend den virtuellen Besuchern über Web Schnittstellen die Möglichkeit geboten wird, die Bilder mit Informationen anzureichern [4]. Das Hinzufügen von Tags ist dabei besonders beliebt. Das *steve.museum*⁵ Projekt [14] zeigte, dass Kunstmuseen von den Tags profitieren können. Auch im deutschsprachigen Raum bieten vermehrt Museen, wie zum Beispiel das *Städel Museum*⁶, den

¹ <http://www.explorARTorium.info>

² <http://www.wga.hu>

³ <http://googleartproject.com>

⁴ <http://www.flickr.com/commons>

⁵ <http://www.steve.museum>

⁶ <http://www.staedelmuseum.de>

BenutzerInnen diese Möglichkeit an. Da das Hinzufügen von Tags so vergnüglich und lehrreich ist, wurden um die Idee sogar Lernspiele wie *ARTigo*⁷ entwickelt.

Die Sammlung aller Tags wird als *Folksonomie* bezeichnet. Es gibt noch relativ wenige Untersuchungen über deren Anwendung in einem kunsthistorischen Kontext. Kunstmuseen haben vor allem den Fokus auf zwei Aspekte gelegt. Einerseits soll mit Hilfe dieser Folksonomie eine, auch den Nichtkunstexperten zugängliche, alternative Beschreibung der Kunstwerke erzeugt und damit auch eine Verbesserung der computerunterstützten Suchmöglichkeiten erreicht werden. Andererseits sollen die Besucher dazu animiert werden, sich selbst über die Kunstwerke auszudrücken und dadurch ihr Interesse an Kunst zu steigern [9].

Folksonomie Analyse, speziell mithilfe von Data Mining Algorithmen, ist ein sehr junges Forschungsgebiet. Schwerpunkte liegen bei der Spam Erkennung, sowie der Untersuchung, wie gut man Folksonomies zum Suchen nach bestimmten Inhalten verwenden kann und in der Erzeugung von Vorschlagssystemen [5].

Wir entwickelten eine Webplattform, mittels der BenutzerInnen Tags zu Bildern hinzufügen konnten [1, 3]. In dieser Arbeit wird das von den BenutzerInnen verwendete Vokabular zur Beschreibung von Bildinhalten untersucht. Wir versuchen anhand des Vokabulars zu erkennen, welches Vorwissen die BenutzerIn hat, welche Bildmotive anhand der Tags erkannt wurden und ob Werke von bekannten KünstlerInnen anders mit Tags versehen und damit unter Umständen auch anders wahrgenommen werden. Diese Ergebnisse werden in zukünftigen Weiterentwicklungen am explorARTorium genutzt werden.

Systembeschreibung

Als Datenquelle für unser Tagging Webplattform verwendeten wir die open-source Daten der Web Gallery of Art (WGA). Die WGA umfasst eine Sammlung von 26.000 Bildern von europäischen Kunstwerken, die im Zeitraum von 1000 und 1900 n. Chr. geschaffen worden sind. Zusätzlich zu den Bildern bietet die WGA auch Informationen über den Künstler / die Künstlerin, den Titel des Kunstwerks, das Entstehungsdatum, an welchem Ort sich das Original befindet, die Abmessungen des Kunstwerkes, welcher Schule es zugerechnet werden kann, und welcher Bildgattung es angehört.

Mit Hilfe des *Taggingtools*⁸ wurden zwischen Oktober 2010 und Jänner 2011 80.000 Tags gesammelt. Das Taggingtool hat ein einfaches Interface, und es werden den BenutzerInnen keine zusätzlichen Informationen angeboten, um sie nicht zu beeinflussen. Im Februar 2011 ging das explorARTorium, eine erweiterte Version des Taggingtools, online (siehe Abbildung 1). Dieses Tool bietet zusätzlich die Möglichkeit, kontextabhängig durch den Datenbestand zu navigieren. Die Auswahl der Kontexte zur Darstellung im explorARTorium wird in einer weiteren Arbeit bei dieser Tagung beschrieben [2].

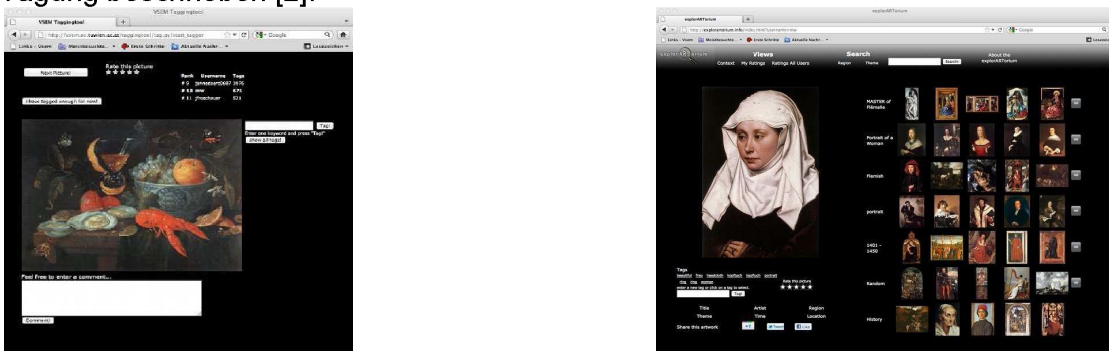


Abbildung 1: Links das Tagging Tool und Rechts das explorARTorium

⁷ <http://www.artigo.org>

⁸ <http://vsem.ec.tuwien.ac.at/taggingtool/>

Die Tags wurden vor der Weiterverarbeitung standardisiert, wie in [7] beschrieben. Es wurden alle Sonderzeichen und Satzzeichen eliminiert und die Worte generell in Kleinbuchstaben transformiert. Deutsche Umlaute wurden in Vokale übersetzt (z.B. ä in ae). Mithilfe der Tags und der Bilder wurde ein binäres Vektorraummodell (Vector Space Model) gebaut, wie es im Information Retrieval üblich ist [10]. In diesem Vektorraum entspricht jeder Tag aus der Folksonomy einer eigenen Dimension. Wir verwendeten ein binäres Model, in dem der Wert 1 besagt, dass der Tag bei einem Bild vorkommt und der Wert 0, dass der Tag nicht vorkommt. Würde der Tag zweimal oder öfter vorkommen oder durch die Normalisierung ein zweites Mal entstehen, ist der Wert für diese Dimension trotzdem nur 1.

Alle Data Mining Verfahren wurden mit *Weka*⁹, einer Data Mining Toolbox [15], ausgeführt. Wir verwendeten die Hauptkomponentenanalyse um die Eigenvektoren zu ermitteln. Die Eigenvektoren bestehen aus Tags, die untereinander eine sehr hohe Korrelation aufweisen. Diese Gruppen von Tags können dann in weiterer Folge als Bildmotive gedeutet werden. Weiters überprüften wir, ob es möglich ist die KünstlerIn eines Bildes anhand der Tags vorherzusagen. Für diese Klassifikationsaufgaben wurde der Naive Bayes Multinomial verwendet [8].

Analyse des BenutzerInnenvokabular

Um einen Überblick zu bekommen, welches Vokabular unsere BenutzerInnen verwendet haben, ordneten wir den Tags unterschiedliche Klassen zu. Die Klassen wurden aus einem Artikel von Sen et. al [11] übernommen, um die Ergebnisse mit anderen Auswertungen von Folksonomies vergleichen zu können. In dieser Terminologie beschreiben *Fakten*-Tags objektive Fakten zu einem Bild, wie zum Beispiel den Ort, Gegenstände oder Personen. *Subjektive*-Tags beschreiben Meinungen und Gefühle zu einem Bild, wie etwa *beängstigend*, *nachdenklich*, *träumerisch*. Persönliche-Tags beschreiben eine persönliche Beziehung zum Bild, zum Beispiel *mein Lieblingsbild*.

<i>Fakten-Tags</i>		<i>Subjektive – Tags</i>	
Tag	Anzahl (#)	Tag	Anzahl (#)
Wolken	1.424	düster	45
Mann	1.343	freundlich	28
Himmel	1.319	Schmerz	27
Frau	1.289	Müde	23
Engel	1.261	ruhig	17

Tabelle 1: Die 5 häufigsten Tags und die Anzahl ihrer Vorkommnis

Das Ergebnis unserer Untersuchung zeigte, dass sich das Tagging Verhalten auf unserer Seite deutlich von Ergebnissen anderer Plattformen wie *flickr*¹⁰, *del.icio.us*¹¹ oder *Cite-U-Like*¹² unterscheidet. Der größte und auffälligste Unterschied ist, dass auf unserer Seite keine persönlichen Tags vorkommen. Der Grund ist vermutlich der, dass es bei uns nicht möglich war, Bilder nach Vorlieben zu ordnen. Wie in Tabelle 1 ersichtlich, sind die subjektiven Tags, wie bei anderen Folksonomies auch, unterrepräsentiert, was gerade in der Kunstwahrnehmung wichtig wäre [13]. Die häufigsten Tags dieser Kategorie waren *düster* (welches bei 45 unterschiedlichen Bildern als Tag verwendet worden ist), *freundlich* (28), *Schmerz* (27), *müde* (23) und *ruhig* (17). Am häufigsten wurden Fakten-Tags vergeben; Solche mit der größten Häufigkeit waren dabei *Wolken* (1.424), *Mann* (1.343), *Himmel* (1.319), *Frau* (1.289) und *Engel* (1.261). Die Fakten Tags

⁹ <http://www.cs.waikato.ac.nz/ml/weka>

¹⁰ <http://www.flickr.com>

¹¹ <http://www.del.icio.us>

¹² <http://www.citeulike.org>

können noch weiter unterteilt werden (siehe Tabelle 2). Die am häufigsten genannten Personen sind *Maria* (795), *Jesus* (704), *Josef* (130) und *Adam* (50). Die häufigsten als Tags verwendeten Ortsnamen sind *Golgatha* (35), *Venedig* (27), *Sixtinische Kapelle* (24), *Dogenpalast* (15) und *Niederlande* (15). Bildthemen, die als Tag verwendet wurden, sind *Anbetung der Könige* (127), *Maria Verkündigung* (80), *Das letzte Abendmahl* (36) und *Pieta* (18). Bei den Bildgattungen sind *Stilleben* (98), *Landschaft* (69), *Mythologie* (28) und *Portrait* (13) am häufigsten genannt. KünstlerInnenamen als Tags waren beispielsweise *Dürer* (80), *El Greco* (66), *Michelangelo* (22) und *Caravaggio* (18). Tags, die Kunstepochen beschreiben, waren *Gotik* (63), *Antike* (28), *Mittelalter* (24), *Renaissance* (23) und *Barock* (16). Anhand dieser Analyse kann festgestellt werden, welche Begriffe den BenutzerInnen präsent sind.

Personen		Orte		Künstler		Bildthemen		Kunstepochen		Bildgattung	
Tag	#	Tag	#	Tag	#	Tag	#	Tag	#	Tag	#
Maria	795	Golgatha	35	Dürer	80	Anbetung der Könige	127	Gotik	63	Stilleben	98
Jesus	704	Venedig	27	El Greco	66	Maria Verkündigung	80	Antike	28	Landscape	69
Josef	130	Sixtinische Kapelle	24	Michelangelo	22	Das letzte Abendmahl	36	Mittelalter	24	Mythologie	28
Adam	50	Dogenpalast	15	Caravaggio	18	Pieta	18	Renaissance	23	Portrait	13
Johannes	48	Niederlande	15	Rembrandt	11	Unbefleckte Empfängnis	14	Barock	16	Religiös	11

Tabelle 2: Die 5 häufigsten Tags der Kategorie Personen, Orte, Künstler, Bildthemen, Kunstepochen, Bildgattungen und die Anzahl ihrer Vorkommnis

Analyse Kunsthistorisches Vorwissen der BenutzerInnen

Jede BenutzerIn nimmt Bilder anders wahr und vergibt dementsprechend unterschiedliche Tags. Mithilfe der Tags kann analysiert werden, wie BenutzerInnen Kunst wahrnehmen und dadurch kann möglicherweise auf ihr kunsthistorisches Vorwissen und ihre Sehgewohnheiten geschlossen werden. Abbildung 2 zeigt, wie zwei unterschiedliche Personen dasselbe Bild¹³ taggen. Person A verwendet ihr kunsthistorisches Fachwissen, wie den Namen des Künstlers (*Michelangelo*), den Ort der Kunstwerks (*Sixtinische Kapelle*) und den Titel des Kunstwerks (*libysche Sibylle*). Anhand dieser Tags ist es offensichtlich, dass sie über kunsthistorisches Vorwissen verfügen muss. Person B nimmt das Bild anders wahr und beschreibt, was sie sieht (z.B. *Aufsteckfrisur*, *Kinder* und *Marmor*). Es scheint also möglich zu sein, mit Hilfe der Tags auf das Vorwissen der BenutzerInnen schließen zu können.



Abbildung 2: Unterschiedliches Tagging Verhalten:
 Person A (Michelangelo, Sixtinische Kapelle, libysche Sibylle)
 Person B (Aufsteckfrisur, Kinder, Marmor)

¹³ http://upload.wikimedia.org/wikipedia/commons/thumb/5/5c/Michelangelo_the_libyan.jpg

Abigail Housen¹⁴ hat ein 5 stufiges Model entwickelt, das die ästhetische Entwicklung von BildbetrachterInnen darstellt [6]. In der ersten Stufe ist das Sammeln von Fakten das Hauptanliegen der BetrachterInnen. Es werden einfach Erfahrungen beobachtet, wie zum Beispiel welche Linien, Kreise, Menschen und Tiere auf dem Bild zu sehen sind. In der zweiten Stufe werden die vorher genannten Beobachtungen mit dem eigenen Wissen kombiniert, und die Ergebnisse mit der Realität verglichen und Folgerungen daraus konstruiert. In der Stufe 3 rücken Informationen, welche KünstlerIn, welche Stillepoche einem Bild zugeordnet werden kann, in den Vordergrund. In der Stufe 4 wird spontan und interaktiv mit dem Kunstwerk umgegangen, neue Interpretationen werden gewonnen. Auf Stufe 5 haben die BetrachterInnen bereits sehr viel Erfahrung mit dem Betrachten von und der Reflektion über Kunstwerke. Für diese BetrachterInnen ist das Bild wie ein alter, lange bekannter Freund.

Die Tags können nun dazu verwendet werden, die einzelnen BenutzerInnen den unterschiedlichen Entwicklungsstufen zuzuordnen [12]. Dabei zeigt sich, dass von den 123 BenutzerInnen 98 BenutzerInnen der Stufe 1 zuzuordnen sind, die übrigen 25 der Stufe 3. Beschreibungselemente, die der Stufe 2 zugerechnet werden, wurden sehr selten verwendet. Dies liegt natürlich an der Natur des Tagging, welches mehr das Sammeln von Informationen unterstützt, während die Reflektion über das Bild selten in Tags zum Ausdruck gebracht wird. Beschreibungselemente der Stufen 4 und 5 sind überhaupt nicht vorhanden; sie sind vermutlich nur in qualitativen Interviews mit den einzelnen BenutzerInnen feststellbar.

Bildmotive

Um zu überprüfen, ob möglicherweise interessante Muster in den Daten vorhanden sind, wurde die Teilmenge aller Tags zu italienischen Bildern analysiert. Dieser Teildatensatz umfasst 2.702 Bilder, bei denen 2.188 unterschiedliche Tags vergeben wurden. Es wurde eine Hauptkomponentenanalyse durchgeführt. Wie in Tabelle 3 ersichtlich, können die Ergebnisse sehr gut interpretiert werden. Die Tags *Wolken*, *Himmel*, *Bäume*, *Hügel* und *Berge* bilden eine Hauptkomponente, die offensichtlich landschaftliche Motive beschreibt. Die Hauptkomponente aus den Tags *Gondeln*, *Venedig*, *Gondolieri* und *Kanal* beschreiben allesamt einen bestimmten geographischen Ort. Die Tags *Abel*, *Kain*, *Brudermord* und *Unbarmherzig* ergeben eine weitere Hauptkomponente, die ein alttestamentarisches Motiv ergibt. Mit dieser Methode ist es möglich, Gruppen von Tags zu erkennen, die in kollektiver Wahrnehmung zusammen auftreten.

Eigenwert	Hauptkomponente
8.97441	0.197 wolken +0.196 himmel +0.157 baeume +0.14 huegel ...
7.23945	0.204 gondeln +0.186 venedig +0.171 gondolieri +0.147 kanal ...
6.13203	0.165 weintrauben +0.142 feigen +0.14 birnen +0.131 kuerbis ...
5.68011	0.14 lendenschurz +0.115 naegel +0.114 wundmale +0.114 inri ...
5.36131	0.222 abel +0.222 kain +0.203 pruegel +0.202 brudermord ...

Tabelle 3: Die 5 Hauptkomponenten mit den größten Eigenwerten

Diese Ergebnisse können dazu verwendet werden, gleiche Bildmotive von unterschiedlichen KünstlerInnen zu ermitteln. 123 Bilder wurden mit den Tags *Wolken*, *Himmel*, *Bäume* und *Hügel* versehen. Mittels der Hauptkomponentenanalyse kann man solche Gruppen von Bildern automatisch erkennen und diese Information auch für Navigationsmöglichkeiten oder Bildervergleiche nutzen.

Die Wahrnehmung von bekannten KünstlerInnen

Eine weitere Fragestellung war, ob die Bilder unterschiedlicher KünstlerInnen unterschiedlich wahrgenommen und deshalb auch mit unterschiedlichen Tags versehen werden. Um diese Frage

¹⁴ <http://www.vtshome.org>

zu beantworten, wurde wieder der italienische Teildatensatz verwendet. Es wurden verschiedene Klassifikationsverfahren angewandt und mit Hilfe einer 10-fachen Kreuzvalidierung [15], einem Standardverfahren zur Messung der Güte von Klassifikationsverfahren, verglichen. Der erfolgreichste Algorithmus konnte 11.2% der Bilder den richtigen KünstlerInnen zuordnen. Das ist signifikant besser als der Intelligente Rater mit einer Erkennungsrate von 3.6 %, der lediglich KünstlerInnen mit der größten Anzahl der Bilder vorhersagt. In Tabelle 4 sind die 5 Künstler mit dem größten F-Measure Wert abgebildet. Im Information Retrieval existieren verschiedene Gütemasse um die Klassifikationsalgorithmen zu beurteilen. Die wichtigsten sind die Genauigkeit und die Trefferquote. Die Genauigkeit gibt den Anteil an, wie oft die vorhergesagten KünstlerIn auch tatsächlich diese KünstlerIn war, während die Trefferquote den Anteil angibt, wie oft die KünstlerIn richtig vorhergesagt wurde im Verhältnis zu ihrem Vorkommen. Um die Güte der Algorithmen mit einer Kennzahl vergleichen zu können werden oft kombinierten Maßzahlen verwendet. Der F-Measure ist das gewichtete harmonische Mittel aus Genauigkeit und Trefferquote.

Für unterschiedliche Bildgattungen werden unterschiedliche Tags verwendet, beispielsweise kommt der Tag *Venedig* ausschließlich in Landschaftsbilder vor, während der Tag *Jesus* in religiösen Bildern vorkommt. Bei 62 % des italienischen Datensatzes handelt es sich um religiöse Bilder. Deshalb sind Künstler, die andere Gattungen malten wie Landschaftsbilder (*Canaletto*) oder Genre (*Longhi*) auch weit vorne zu finden, da ihre Werke leichter erkannt werden können. Die Detailanalyse zeigt, dass nur Bilder KünstlerInnen zugeordnet werden konnten, deren Bilder in unserem Kulturkreis über einen sehr großen Bekanntheitsgrad verfügen wie zum Beispiel Bilder von *Da Vinci*, *Michelangelo*, *Giotto*, *Tizian*, *Veronese* und *Caravaggio*. Das lässt den Schluss zu, dass bekannte Bilder, auch wenn sie nur in unserem Unterbewusstsein präsent sind, anders getagged werden, als uns unbekannte Werke.

Künstler	F-Measure
LEONARDO da Vinci	0.458
CANALETTO	0.422
MICHELANGELO Buonarroti	0.197
Pietro LONGHI	0.167
GIOTTO di Bondone	0.162

Tabelle 4: Die 5 anhand der Tags am besten erkennbaren italienischen Künstler

Zusammenfassung und Ausblick

In dieser Arbeit analysierten wir die Folksonomie, die die BenutzerInnen unserer unterschiedlichen Webplattformen erstellt haben. Bei der Analyse der Tags zeigte sich, dass am häufigsten Fakten-Tags verwendet werden, wie zum Beispiel *Wolken*, *Mann* und *Himmel*. Subjektive-Tags, die Gefühle ausdrücken, werden demgegenüber kaum verwendet, und Persönliche-Tags kommen gar nicht vor. Es wurde anhand der Tags der einzelner BenutzerIn analysiert. Dabei ergab sich, dass der größte Teil der BenutzerInnen keine kunsthistorischen Fachtermini für das Taggen verwendet. Wir zeigten weiter, dass es mithilfe einer Hauptkomponentenanalyse möglich ist, wiederkehrende Bildmotive zu ermitteln. Mit Hilfe eines Klassifikationsalgorithmus konnten wir herausfinden, dass Bilder von bekannten KünstlerInnen wie zum Beispiel *Leonardo Da Vinci* anders beschrieben werden als Bilder von weniger bekannten KünstlerInnen.

Durch Anwendung von Data Mining auf eine Folksonomie lassen sich Rückschlüsse über die Wahrnehmung von Kunstwerken ziehen. Die Erkenntnisse dieser Analyse bieten einen nützlichen Input, um Kunstwerke im Web zu präsentieren und BenutzerInnen einen personalisierten und einfacheren Zugang zur Kunst zu ermöglichen.

Danksagung

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

Referenzen

- [1] Arends, M., Froschauer, J., Goldfarb, D., Merkl, D. und Weingartner, M. (2010). Interaktion mit musealen Inhalten in Web3D, In *EVA 2010 Berlin: Elektronische Medien & Kunst, Kultur, Historie, die 17. Berliner Veranstaltung der Internationalen EVA-Serie Electronic Imaging & the Visual Arts, Konferenzband*.
- [2] Arends, M., Froschauer, J., Goldfarb, D., Merkl, D. und Weingartner, M. (2011). Vermittlung kunstgeschichtlicher Inhalte durch die Kontextualisierung von Kunstwerken im explorARTorium, In *EVA 2011 Berlin: Elektronische Medien & Kunst, Kultur, Historie, die 18. Berliner Veranstaltung der Internationalen EVA-Serie Electronic Imaging & the Visual Arts, Konferenzband*.
- [3] Arends, M., Froschauer, J., Goldfarb, D., und Merkl, D. (2011). Analysing user generated content related to art history. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, Graz, Austria.
- [4] Arends, M., Goldfarb, D., Merkl, D., und Weingartner, M. (2011). Museums on the web: Interaction with visitors. In F. L. Georgios Styliaras, Dimitrios Koukopoulos, editor, *Handbook of Research on Technologies and Cultural Heritage: Applications and Environments*, pages 144–165. IGI-Global.
- [5] Hotho, A. (2010). Data mining on folksonomies. In G. Armano, M. de Gemmis, G. Semeraro, and E. Vargiu, editors, *Intelligent Information Access*, volume 301 of *Studies in Computational Intelligence*, pages 57–82. Springer.
- [6] Housen, A. (2007). Art Viewing and Aesthetic Development: Designing for the Viewer. In *From Periphery to Center: Art Museum Education in the 21st Century*, P. Villeneuve (Ed). VA: NAEA.
- [7] Hsu, M.-H., und Chen, H.-H. (2008). Tag normalization and prediction for effective social media retrieval. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*, pages 770–774.
- [8] McCallum, A, und Nigam, K. (1998). A comparison of event models for naive bayes text classification. In *AAAI-98 Workshop on 'Learning for Text Categorization'*.
- [9] Oomen, J., und Aroyo, L. (2011). Crowdsourcing in the cultural heritage domain: Opportunities and challenges. In *Proceedings of the 5th International Conference on Communities Technologies*.
- [10] Salton, G., Wong, A. und Yang C., S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18:613–620.
- [11] Sen, S., Lam S., K., Rashid, A., M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F., M., Riedl, J. (2006). Tagging, communities, vocabulary, evolution. In *Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work*.
- [12] Smith, M. K. (2006) Viewer tagging in art museums: Comparisons to concepts and vocabularies of art museum visitors. In *Proceedings of the 17th SIG Classification Research Workshop*.
- [13] Schuster. M. (2011). *Wodurch Bilder wirken. Psychologie der Kunst*. DuMont Buchverlag, Köln.
- [14] Trant. J. (2009). Tagging, folksonomy and art museums: Results of steve.museum' s research. In *Proceedings of Museums and the Web Conference 2009*. Archives Museum Informatics.
- [15] Witten, I. H. und Frank, E. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, 3 Edition.