

Visuelle Suche für Bild- und Produktdatenbanken

Visual Search for Image and Product Databases

Daniel Burckhardt
Zentrum für Zeithistorische Forschung Potsdam
Am Neuen Markt 1, D-14467 Potsdam
Tel.: 0331/28991-12
E-Mail: burckhardtd@geschichte.hu-berlin.de, Internet: www.zzf-pdm.de

Pepe Jürgens
Weltformat Design
Winsstraße 6, D-10435 Berlin
E-Mail: juergens@weltformat.de, Internet: www.weltformat.de

Zusammenfassung:

Suchen ist notwendig. Suchen soll Spaß machen. Sammlungs- und Produktdatenbanken umfassen häufig umfassende Bildbestände, die weniger nach Volltexten und Metadaten als aufgrund visueller Eigenschaften gesichtet werden. Die gebräuchlichste Präsentation in Form langer Textlisten unterstützt dies nur unzureichend und wirkt deshalb auf viele Nutzer eher abstoßend als einladend. Content-based image retrieval als Suchmöglichkeit, die direkt auf den Bildinhalten und nicht deren Beschreibung operiert, steht deshalb im Zentrum einer Reihe von Forschungsprojekten. Hinter der visuellen Suche steht dagegen eine einfache Beobachtung aus dem vor-digitalen Zeitalter: Umfangreiche Bildbestände – die Diapositive auf dem Leuchttisch – lassen sich im Gegensatz zu größeren Textmengen sehr schnell sichten und nach einer Vielzahl von Kriterien gruppieren. Die visuelle Suche transformiert diesen Vorgang ins Web und ermöglicht so eine intuitive und flüssige Navigation, die auf die Wünsche und Bedürfnisse der Anwender eingeht und die Ergebnisse klar, sinnvoll und ästhetisch präsentiert.

Abstract:

Searching is a necessity. Search should be fun. Collection and product databases often contain large amounts of pictures. When screening such an image pool, visual characteristics are often more important than text or metadata. But the most common presentation of the results in the form of long lists of text seems to be rather repulsive. Content-based image retrieval, which operates directly on the image data instead of its description and classification, is therefore at the center of numerous research projects. Visual search is a low-tech alternative based on a simple observation from the pre-digital era: extensive image libraries - the slides on a light table - allow for easy scanning and ordering of large result sets. Visual search aims to bring this approach to the web. By combining an intuitive and fluid navigation with a clear, meaningful and aesthetically pleasing presentation, these applications attract the users' attention to a larger set of results.

Nicht erst seit Google ist die Recherche eine der zentralen Tätigkeiten im Internet. Vier der zehn meistbesuchten Websites sind auf die Web-Suche fokussiert.¹ Zwei Dinge hoben Google von Vorgängern wie Altavista oder Hotbot ab: eine radikal verschlankte Startseite, bei der das Suchfeld im Zentrum der Seite steht, sowie eine deutlich verbesserte Sortierung durch Einbezug des PageRanks, der jeden Link auf eine Seite als mehr oder weniger stark gewichtete Empfehlung zählt. Der rasante Aufstieg von Google hat die Beschäftigung mit Suchen fast ausschließlich auf

¹ <http://www.alexa.com/topsites>

dieses Ranking der Trefferlisten verengt. Morville und Callender weisen in ihrem Buch „Search Design Patterns“ darauf hin, dass „Best First“ aber nur eines von vielen Merkmalen ist, die bei der erfolgreichen Konzeption von Suchen von Bedeutung sind.² In der Folge konzentrieren wir uns auf die visuelle Suche als Alternative zur herkömmlichen Präsentation der Suchresultate als reine Textliste. Diese Präsentation eignet sich primär für Datenbestände, bei denen ein visueller Platzhalter – das Vorschaubild – einen Treffer aussagekräftig repräsentieren kann. Dies trifft auf viele Bild- und Medienarchive ebenso zu wie auf Datenbanken und Online-Shops, bei denen nicht nur Metadaten, technische Daten oder Preis entscheidend sind, sondern das Äußere die Wahl eines bestimmten Produktes wesentlich mitbestimmt.

Die Leitmetapher für die visuelle Suche ist das Leuchtpult für Diapositive. Der Anwender sortiert, gruppiert, ordnet, er ist aktiv, experimentiert, sucht sich Feedback beim gemeinsamen Arbeiten mit Kollegen. Das ist ein schneller kreativer Prozess: „Ich sehe, was ich finden kann und wie ich es finden kann.“ Die Ergebnisse werden dabei über einen eindeutigen visuellen Platzhalter repräsentiert und über die Anordnung in der Ebene graphisch visualisiert. Nach der Übertragung des Diapositivs in eine Multimedia-Anwendung kann das Vorschaubild um eine Reihe von Zusatzinformationen ergänzt werden. Diese erscheinen in reduzierter Form beim Darüberfahren mit der Maus. Nach einem Mausclick oder einem längerem Verharren auf dem Bild werden Detailinformationen nachgeladen und angezeigt. Dabei sollte die flüssige, schnelle und einfache Bedienung im Vordergrund stehen und durch Animationen sinnvoll unterstützt werden.

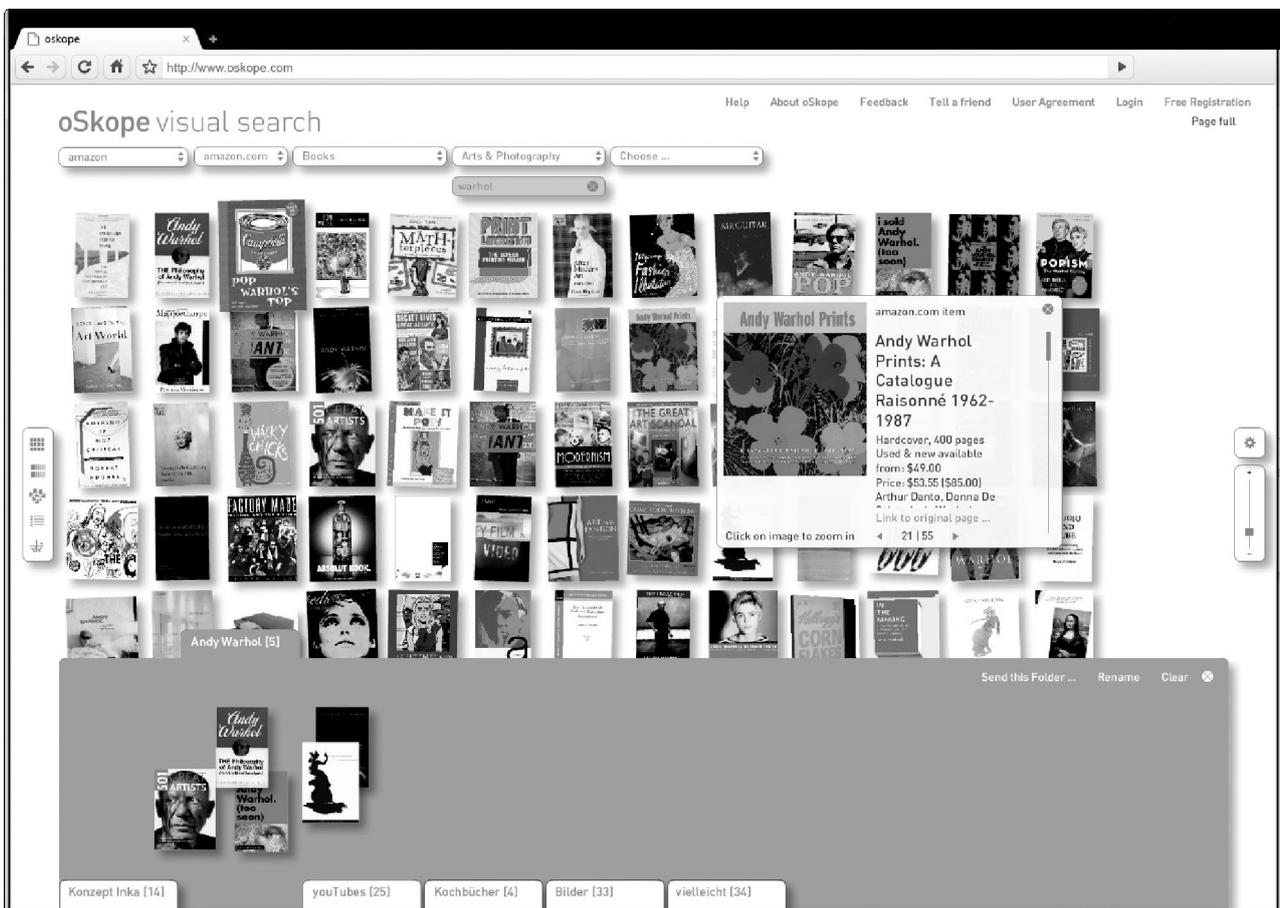


Abb. 1: Grundelemente der visuellen Suche bei oSkope visual search

Die von uns realisierten Projekte illustrieren das breite Einsatzspektrum dieses Suchparadigmas: Ausgehend von einem Prototypen für die visuelle Suche in den Produktdaten von Amazon, die von

² P. Morville, J. Callender. *Search Patterns: Design for Discovery*. O'Reilly, Sebastopol, Calif. 2010.

Christoph Schifferli konzipiert und von Peter Chylewski programmiert wurde, entwickelten die Autoren eine alternative Suchoberfläche für Amazon, Ebay und YouTube und andere Webservices (www.oskope.de) . Eine visuelle Suche für Bauprodukte (www.baunetzprodukte.de) sowie ein wissenschaftliches Archiv für Materialien (www.materialarchiv.ch) bauen darauf auf und wurden sowohl von Architekten und Künstlern als auch im Museum gut angenommen.

Grundelemente der visuellen Suche

Die wiederkehrenden Grundelemente dieser Anwendungen stellen wir in der Folge kurz vor:

Single-Page User Interface

Die visuelle Suche bricht mit dem Paradigma des Seitenaufbaus herkömmlicher Web-Anwendungen. Sowohl für die Detail-Ansicht als auch beim Blättern werden nur diejenigen Daten nachgeladen, die tatsächlich benötigt werden, ohne den Bildschirm komplett neu aufzubauen. Die schnellen Reaktionszeiten ohne die Angst, bestehende Ergebnisse zu verlieren, animieren die Nutzer zu zusätzlichen Aktionen. Damit kann ein „Flow“ aufgebaut werden, der die Zeit vergessen lässt und die Besucher zu umfangreicheren Suchen animiert. Die Nutzeranalyse von oSkope weist darauf hin, dass dies tatsächlich funktioniert. Der durchschnittliche Nutzer bleibt über vier Minuten auf der Seite und tätigt dabei über zwölf Klicks innerhalb der Site (Suchanfragen, Detailinformationen und Weiterblättern). Diese Werte liegen fast doppelt so hoch (Besucherzeit) bzw. fünfmal höher (Zahl der Klicks) als bei vergleichbaren Web-Sites. Die neue Version von Google Images zeigt, dass es auch bei HTML-basierten Anwendungen zielführend sein kann, über AJAX-Techniken gezielt Daten nachzuladen und einzublenden, statt bei jedem Klick die Seite komplett neu aufzubauen.

Resultat-Ansichten

Bei den vertikal angeordneten Trefferlisten herkömmlicher Suchen passen in der Regel zwischen 5 (Amazon/eBay) und 10 Treffer (Google/Bing) gleichzeitig in den sichtbaren Bereich des Bildschirmsfensters. Danach muss gescrollt werden oder ein Wechsel auf die nächste Resultatseite erfolgen. Da die Vorschaubilder bei der visuellen Suche nicht nur vertikal, sondern auch horizontal nebeneinander dargestellt werden können, passen hier deutlich mehr Treffer auf eine Seite. Je nach Bild- und Bildschirmgröße können 50 bis 150 Datensätze gleichzeitig betrachtet werden. So erhält der Nutzer einen Gesamtüberblick und findet im besten Fall bereits auf dem ersten Bildschirm Ergebnisse, zu denen er sich sonst nie durchgeklickt hätte.

Alternative Anordnungen der Vorschaubilder unterstützen unterschiedliche Ansätze des Findens sowie individuelle Vorlieben der Anwenderinnen und Anwender: Tabellarische Anordnungen in Form einer linearen Liste oder einer zweidimensionalen Tabelle orientieren sich am stärksten an herkömmlichen Ergebnislisten. Dadurch erleichtern sie neuen Nutzern den Umstieg auf die visuelle Suche und ermöglichen zudem die Sortierung der Ergebnisse nach zusätzliche Informationen wie etwa Name, Preis oder Verkaufsrang. Bei der Materialdatenbank lassen sich die Ergebnisse nach vorgegebenen Kategorien wie z.B. Holz oder Stein gruppieren. Damit lassen sich zusammengehörige Resultate auf einen Blick erfassen. Gänzlich neue Einsichten eröffnet die Anordnung der Bilder auf zwei Achsen, bei eBay zum Beispiel nach Angebotsdauer und Preis oder nach Festigkeit und Dichte in der Materialdatenbank. Im Quadranten unten rechts liegen hier leichte Materialien mit hoher Festigkeit oder die „Schnäppchenecke“ von bald auslaufenden Angebote mit niedrigem Höchstgebot.

Diese Liste lässt sich je nach Datenbestand durch zusätzliche Visualisierungen erweitern. So wäre etwa eine Anordnung der Vorschaubilder auf einer geografischen Landkarte oder die Andeutung einer dritten Dimension durch Staffelung in die Tiefe denkbar.

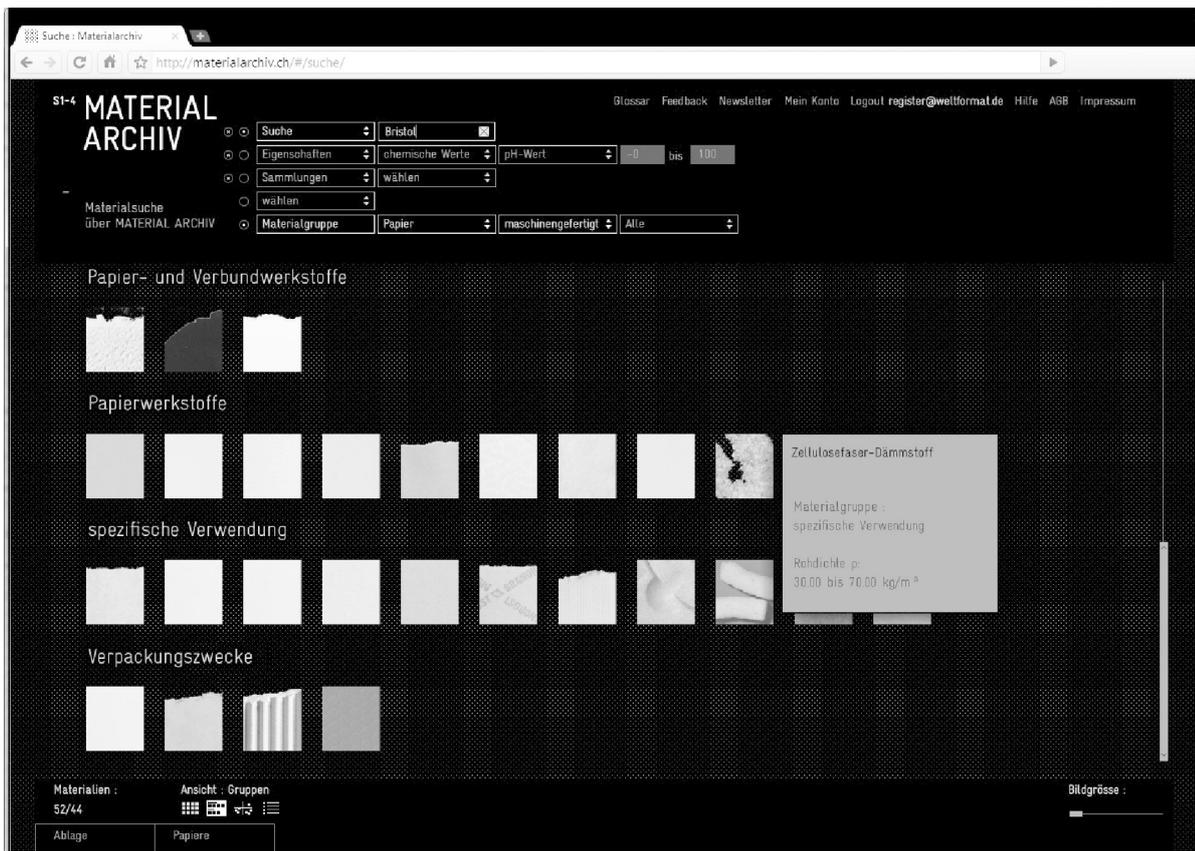


Abb. 2: Gruppenansicht mit erweiterter Suche (MATERIAL.ARCHIV)

Preview und Detailansicht

Nach einer groben visuellen Sichtung der Ergebnisse sind für den weiteren Ablauf zusätzliche Informationen relevant: Wer ist der Autor des Buchs mit dem attraktiven Umschlag? Ist das günstige Angebot ein gebundenes Buch oder ein Taschenbuch, gebraucht oder neu? Dieser Prozess ist bei uns zwei- bis dreistufig aufgebaut. Je nach Anwendung werden erste Infos – der Name des Materials etwa – von Anfang an direkt im oder unter dem Bild angezeigt. Beim Darüberfahren mit der Maus werden in einem kleinen Fenster mehr oder weniger ausführliche Zusatzinformationen eingeblendet. Möchte der Anwender sich detailliert informieren, verharret er mit dem Mauszeiger länger über dem Bild oder klickt darauf. Danach öffnet sich ein Fenster mit im Hintergrund nachgeladenen Detailinformationen. Diese Ansicht ist nicht auf Texte beschränkt: Es können auch weitere Bilder, Audio-Daten oder Videos direkt in der Applikation abgespielt werden.

Sammelbox

Stets sichtbar am unteren Bildschirmrand sind ein oder mehrere Ordner als Merk- oder Wunschlisten. Beliebige Suchergebnisse können über Drag-and-Drop darin abgelegt werden. So können Einträge kurzzeitig zurückgelegt oder zu persönlichen Sammlungen geordnet werden. Je nach Anwendung können mehrere Ordner gleichzeitig angelegt, benannt und für weitere Sitzungen gespeichert werden. Ein Sammelordner kann per E-Mail verschickt werden, um die Ergebnisse der Recherche anderen Personen zukommen zu lassen oder dauerhaft zu archivieren.

Kombination mit weiteren Search Patterns

Diese für die visuelle Suche spezifischen Grundelemente lassen sich frei mit weiteren Search Patterns kombinieren. So lässt sich etwa die Eingabe im Suchfeld durch Autovervollständigung ergänzen. Suchanfragen lassen sich über einen Kategorienbaum einschränken oder erweitern. Mithilfe einer erweiterten Suche können Volltextsuche auf bestimmte Felder bzw. numerische Parameter auf vorgegebene Wertebereiche eingegrenzt werden. Die Entscheidung für solche

zusätzlichen Funktionen hängt sowohl vom Datenbestand als auch von der Zielgruppe der Anwendung ab.

RFID-Tags zur Einbindung der Dingwelt

Für die Arbeit im Archiv und die Präsentation im Rahmen von Ausstellungen sollten bei MATERIAL.ARCHIV zudem die Materialmuster mit den Inhalten der Datenbank verknüpft werden. Dies erfolgte über die Anbringung von RFID-Tags an den Materialmustern. Sobald ein oder mehrere Materialien in den Scan-Bereich neben das Terminal gelegt werden, erscheinen die entsprechenden Einträge im Browser als Ergebnisse einer visuellen Suche, die den unmittelbaren Zugriff auf das Datenblatt und die Personalisierungsoptionen zur Speicherung und zum Versand ermöglicht.³

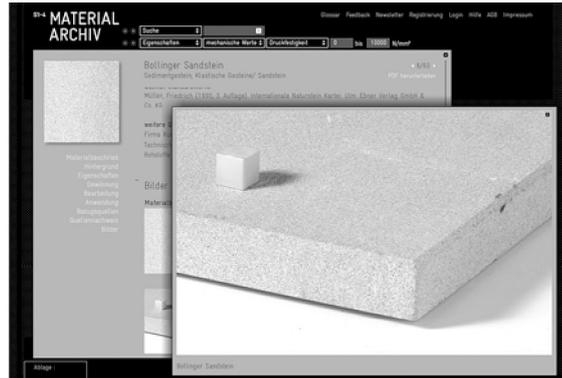
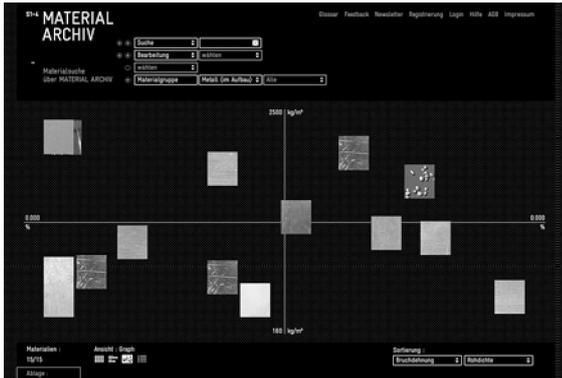


Abb. 3: Graphansicht mit erweiterter Suche (MATERIAL.ARCHIV)

Abb. 4: Detailansicht (MATERIAL.ARCHIV)

Abb. 5: Kiosk mit RFID-Anbindung im Gewerbemuseum Winterthur, © Materialarchiv

Abb. 6: Kiosk mit RFID-Anbindung im Archiv der Hochschule Luzern, © Materialarchiv

Technische Umsetzung

Frontend

Das Frontend ist für die Präsentation der Suchergebnisse zuständig. Es wurde als Rich Internet Application (RIA) in Flash programmiert. Dies ermöglicht flüssige Animationen, Drag-and-Drop-Operationen sowie die Integration vieler Medienformate. Zudem können die meisten Berechnungen, etwa zur Rotation und Skalierung von Bildern, Client-seitig ausgeführt werden. Dies minimiert die Zahl der Serveranfragen und gewährleistet kurze Antwortzeiten auch bei vielen gleichzeitigen Besuchern der Anwendung.

³ Vgl. U. Honegger, L. Glanzmann. Von der Schublade ins Netz (Titelgeschichte). HOCHPARTERRE (3) (2009), p. 18-25.

Mit der zunehmenden Verbreitung des iPad, mit seinem großen Touch-Screen eine ideale Hardware-Plattform für visuelle Suchen, sollten aufgrund der fehlenden Flash-Unterstützung alternative Realisierungen ernsthaft geprüft werden. Mit der Unterstützung des canvas-Element von HTML5 bzw. der Scalable Vector Graphics Standard des W3C in den aktuellen oder angekündigten Versionen aller gängigen Web-Browser sowie JavaScript-Bibliotheken zur Browser-übergreifenden Realisierung flüssiger Animationen wie z.B. <http://scripty2.com/> sollte dies mittlerweile auch technisch möglich sein.

Middleware: Verbindung des Clients mit der Datenbank

Der Flash-Client greift sowohl für Suchanfragen als auch für die Anzeige von Detailinformationen auf eine Middleware auf dem Web-Server zu. Diese kapselt die Zugriffe auf eine relationale Datenbank bzw. auf die Web-Services von Drittanbietern. Nur Bilder werden vom Frontend nach Möglichkeiten direkt geladen. Diese zusätzliche Schicht war zunächst eine rein technische Notwendigkeit. Gemäß der sog. „same origin policy“, dem Sicherheitsmodells von Flash, kann eine in den Browser eingebettete Flash-Anwendung, die von <http://oskope.com> geladen wird, nicht direkt mit Diensten wie z.B. <http://rest.api.ebay.com> kommunizieren. Diese zusätzliche Abstraktionsschicht hat aber auch Wartung und Nachnutzung der Anwendung erleichtert. Da das Frontend keine konkreten Kenntnisse der einzelnen Web-Services benötigt, konnte die Integration zusätzlicher Dienste wie etwa YouTube ohne Code-Änderungen im Frontend erfolgen. Eine klar spezifizierte Schnittstelle zwischen Frontend und Middleware erlauben zudem eine arbeitsteilige Programmierung und erleichtert die Fehlersuche und testgetriebene Entwicklung.

Fazit

Die visuelle Suche schließt die zielgerichtete Suche nicht aus. Ihr Zielpublikum besteht aber primär aus Entdeckerinnen und Entdeckern, denen sie über ein intuitives und interaktives Interface das Finden erleichtert. Kataloge, Archive und Sammlungen mit einem attraktiven Bildbestand eignen sich besonders für einen solchen Zugang, der neben der gezielten Recherche eine attraktive Präsentation mit Möglichkeiten zum forschenden und spielerischen Entdecken bietet. Jüngere Studien haben gezeigt, dass aufgrund technischer und kultureller Veränderungen diese Art des Suchens zunehmend zu einem Zeitvertreib wird, dessen Motivation weniger einer konkreten Zielstellung als dem Wunsch nach Überraschung, „edutainment“ oder einfach nur Ablenkung oder Entspannung entspringt.⁴ Durch intuitive Bedienung und eine attraktive Präsentation der Resultate kann die visuelle Suche solche Anwendungsszenarien besser unterstützen, als herkömmliche Suchtechniken, die sich primär an quantitativen Bewertungen der Ergebnisse gemäß Genauigkeit (precision) und Trefferquote (recall) orientieren.

⁴ M. L. Wilson. D. Elsweiler. Casual-leisure Searching: the Exploratory Search scenarios that break our current models. In Proceedings of the Fourth Workshop on Human-Computer Interaction and Information Retrieval, Rutgers University, New Brunswick, NJ, 22 August 2010. <http://www.cs.swan.ac.uk/~csmaw/pubs/HCIR2010.pdf>