


Künstliche Intelligenz in der Bilderschliessung: Praxisbeispiel Bildarchiv ETH-Bibliothek

Presentation

Author(s):

Graf, Nicole 

Publication date:

2024-01-30

Permanent link:

<https://doi.org/10.3929/ethz-b-000656804>

Rights / license:

[Creative Commons Attribution-ShareAlike 4.0 International](#)



Künstliche Intelligenz in der
Bilderschliessung: Praxisbeispiel
Bildarchiv ETH-Bibliothek
Nicole Graf

Einführung ins Bildarchiv

Zahlen & Fakten



- Gegründet in **2001** als Bildagentur
- Open Data in **2015**
- 6 Festangestellte (**470 %**) und 7 stud. Hilfskräfte (180 %)
- **3,6 Mio.** physische Bilder, **1,3 Mio.** digitalisiert
- Jährlich Digitalisierung: **60'000** bis **80'000**
- **947'422 Bilder** online auf [E-Pics Bildarchiv](#) (15 Jahre in 2021 vgl. [Blogpost](#))
- **Bestände:**
 - ETH [Professor:innen](#) und [Institute](#)
 - [Ansichten](#) & [Porträts](#)
 - [Luftbilder](#)
 - [Swissair](#), [Comet Photo AG](#) usw.

Erschliessung – mehrere Wissensschichten

Mehrere Wissensschichten...



Maschinen



Crowd



Expert:innen



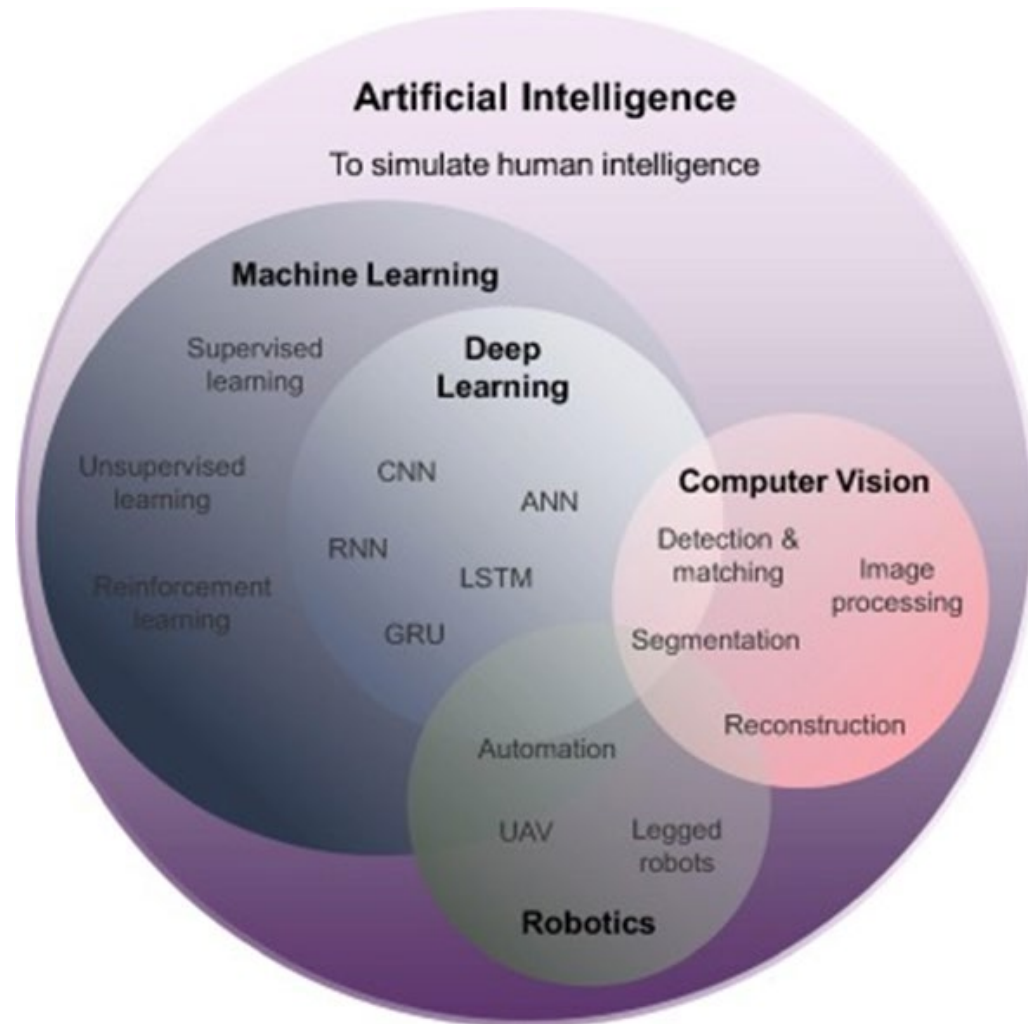
Bibliothekar:innen

Mehrere Wissensschichten



Maschinen

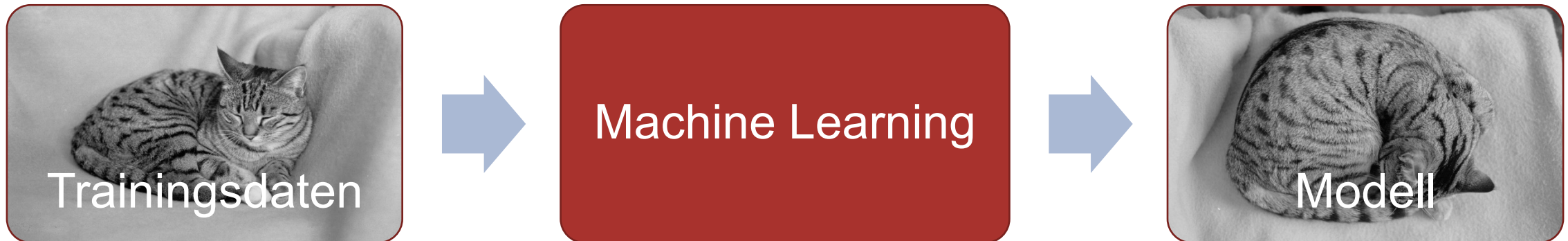
Was ist Computer Vision?



CV ist «AI für Bilder» - maschinelles Sehen

- Teilgebiet von KI und MLs
- Technologie, die Computern «helfen» soll, den Inhalt digitaler Bilder – Fotos und Videos – zu «sehen» und zu verstehen
- Fähigkeit des menschlichen Sehens reproduzieren und automatisieren
- Ein Bild auf Pixelebene analysieren, Pixel mit jeweils eigenen Farbwerten
- Mithilfe von Sensoren und Algorithmen, visuelle Informationen abzurufen, zu verarbeiten und zu verstehen
- Algorithmen, die auf Mustererkennung basieren

Wie wird die Maschine intelligent? Durch Training!

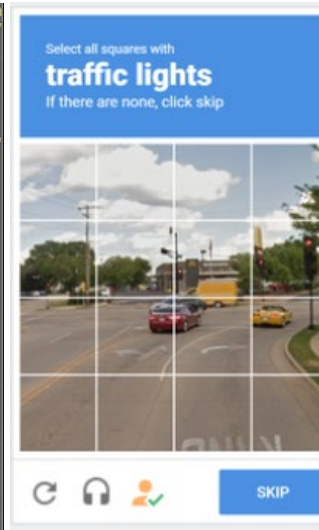


- 1000 repräsentative Bilder pro Objekt (Konzept)
- Händisch von Menschen kategorisiert (Daten werden trainiert → MTurks)
- Gute **Trainingsdaten** (also nicht nur schwarze Katzen auf grünem Gras)

- KI trainieren oder Maschinen lernen
- **Programmierung** des **ML-Algorithmus** und der entsprechenden **Regeln**
- Anschliessend wird der ML-Algorithmus **mithilfe der Daten trainiert**

- ML-Algorithmus beginnt auf Basis der Daten **Muster und Strukturen** erkennen und **eigenständig Modelle** zu bilden
- Die **Modelle** bilden anschliessend die Grundlage, um **Wahrscheinlichkeiten** und Werte vorherzusagen.

Anwendungen im Alltag



«Die KI ist ein Werkzeug, damit Menschen effizienter sind»

Richard Socher

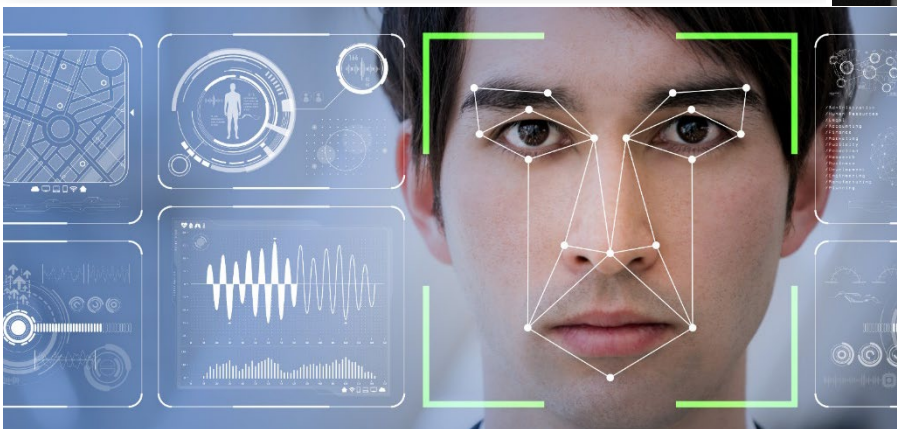
(01:25:50)

Richard Socher, was denken Maschinen?

Podcast «Alles gesagt»,

26.11.2020 (08:16)

[\(Link\)](#) [\(Transkript\)](#)



Anwendungsgebiete von Computer Vision

- **Bildentstehung**: z. B. Verbesserung der Helligkeits- und Kontrastunterschiede (Fotofilter)
- **Zeichen- und Schrifterkennung** (OCR, Handschrifterkennung HCR)
- **Objekterkennung**: komplexes Zusammenspiel von Merkmalsextraktion, Mustererkennung, Farbextraktion und selbst lernenden Entscheidungsalgorithmen der KI
- **Personenerkennung**: Gesichtserkennung (Alter, Kultur, Geschlecht), Mimik-Erkennung, Biometrie
- **Objektklassifizierung**: Ausspielen von Bildbeschreibungen oder von Klassifikationen/Tags («Autotagging»)
- **Umgekehrte Bildersuche**: Reverse Image Search, Content-based Image Retrieval
- **Generative AI: Bildergenerierung** (Dall-e, Midjourney, Bing Image Creator) und **Objekterkennung** (ChatGPT-4, Bing)

1. Durchlauf: Clarifai und DeepL

Februar 2021 bis Mai 2022


Experimentieren mit Clarifai: Autotagging von 1 Million Bildern

Neue, automatisch generierte Tags

- Bis 2021 basierten die Metadatenebenen hauptsächlich auf menschlichem Wissen
- Ab 2021: automatisch generierte Schlagwörter durch künstliche Intelligenz (KI), so genanntes Autotagging
- Autotags als **Ergänzung** zu den "intellektuellen" Schlagwörtern
- KI-Software: **Clarifai**, "Allgemeines" Modell
- Warum Clarifai? Man muss nehmen, was man bekommt → **Autotagging-Modul ist Teil der Katalogsoftware**

Autotagging mit Clarifai im Backend Cumulus

Autotagging Control Panel

Feldname	Feldinhalt
Bildcode	Dia_385-14415
Miniatur	
Wertung
▼ AutoTagging Control	
AutoTagging	AutoTagging noch nicht gestartet
Autotag Model	kein Wert
AutoTagging Color	<input type="checkbox"/>
AutoTagging Klassifizie...	<input type="checkbox"/>
AutoTagging Face	<input type="checkbox"/>

➔ Status setzen, Modell wählen, Farbe aktivieren




▼ AutoTagging Control	
AutoTagging	AutoTagging noch nicht gestartet
Autotag Model	AutoTagging noch nicht gestartet
AutoTagging Color	Wartend
AutoTagging Klassifizie...	Autotagging erfolgreich
AutoTagging Face	Autotagging fehlgeschlagen
	AutoTagging geprüft
	kein Wert

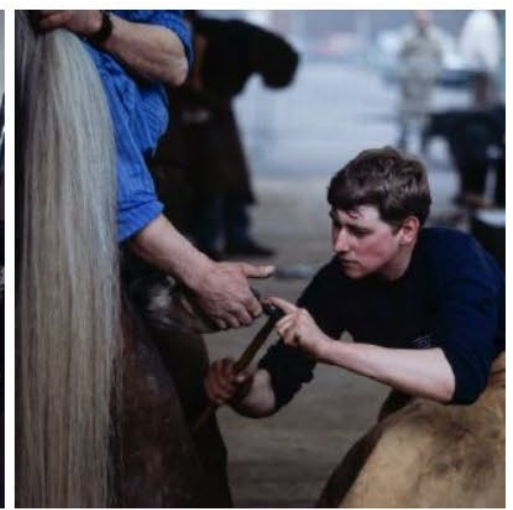
▼ AutoTagging Control	
AutoTagging	AutoTagging noch nicht gestartet
Autotag Model	General
AutoTagging Color	General
AutoTagging Klassifizie...	Wedding
AutoTagging Face	Travel
	Food
	Apparel
	kein Wert

▼ AutoTagging Control	
AutoTagging	AutoTagging noch nicht gestartet
Autotag Model	General
AutoTagging Color	<input checked="" type="checkbox"/>
AutoTagging Klassifizie...	<input type="checkbox"/>
AutoTagging Face	<input type="checkbox"/>

➔ Autotagging Ergebnis

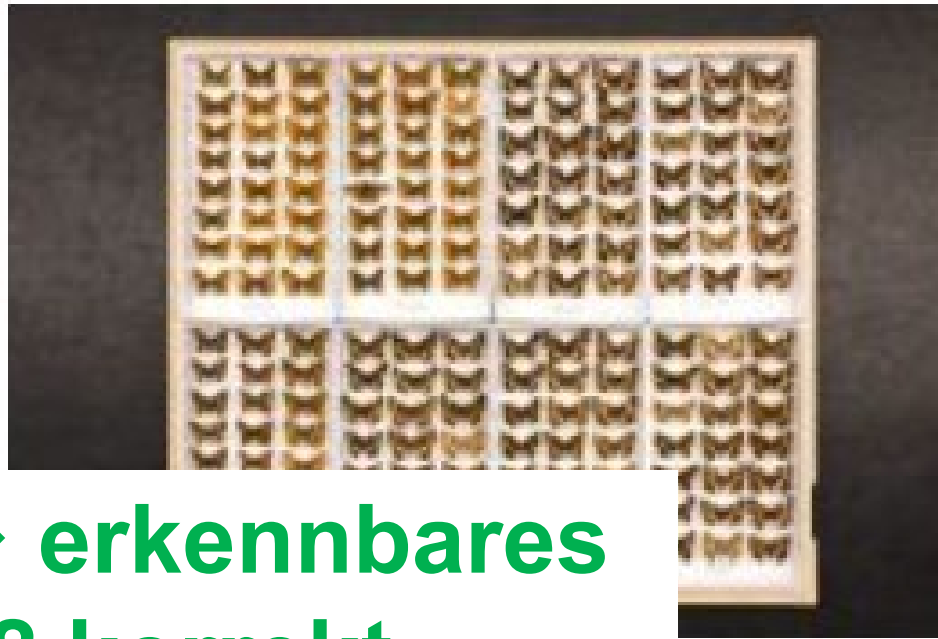
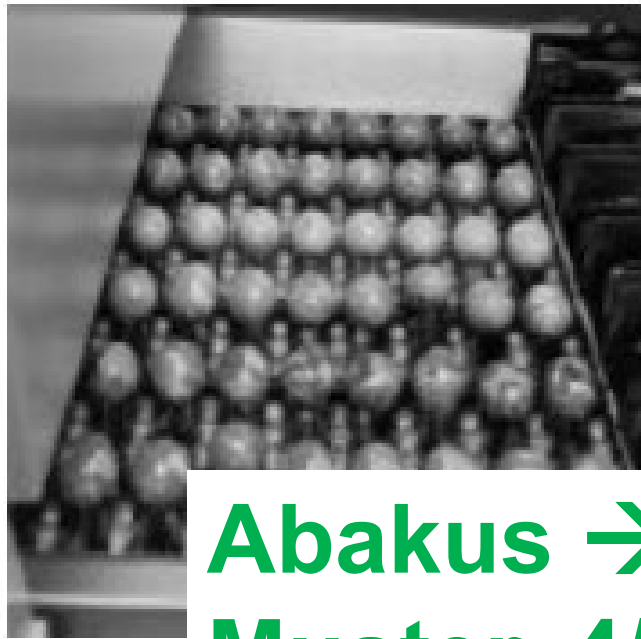


Feldname	Feldinhalt
Bildcode	Dia_385-14415
Miniatur	
Wertung
▼ AutoTagging Control	
AutoTagging	Autotagging erfolgreich
Autotag Model	General
AutoTagging Color	<input checked="" type="checkbox"/>
AutoTagging Klassifizie...	<input type="checkbox"/>
AutoTagging Face	<input type="checkbox"/>
▼ AutoTagging Result	
Kategorien	Schlagwörter
	abenteuer Abwechslung berg Eis
	gletscher kalt keine Person landschaft
	menschen Reise schnee skifahrer
	winter
AutoTaggingColor 1	<input type="color"/>
AutoTaggingColor 2	<input type="color"/>
AutoTaggingColor 3	<input type="color"/>
AutoTaggingColor 4	kein Wert 30.01.2024 13
AutoTaggingColor 5	kein Wert



Schlagzeuger → top





**Abakus → erkennbares
Muster, 4/8 korrekt**



Kategorien ^

schreiben ⊗

ETHBIB.Bildarchiv

Schreiben ≡

Slate (Schreiben)

mit Bild v

Lizenz v

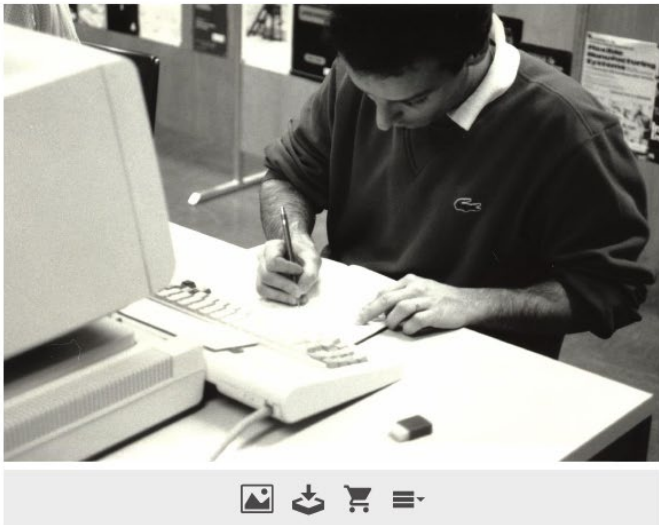
Form v

Farbe v

Farbenauswahl v

Zeitraum v

schreiben → neue Treffer



Kategorien: ⓘ

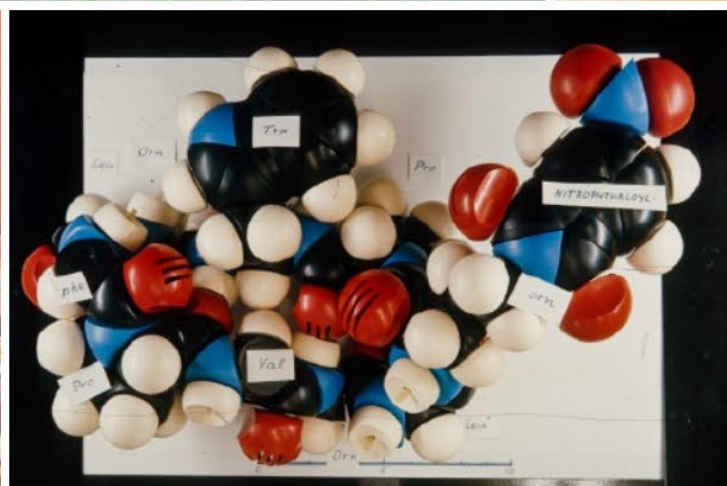
- Ansichtensammlung
- Archiv + Bibliothek
- ETHICS
- Geräte der Informationstechnik (Computer)
- Reportagefotografien
- Unbekannt
- Zürich-Oberstrass

Schlagwörter:

- Büro
- Computer
- Erwachsener
- Innenaufnahme
- Konzentration
- Mann
- Menschen
- Möbel
- Schreiben
- Schreibtisch
- Technologie
- Unternehmen



Colour: FireBrick



DeepL (2022–)

Neue, maschinell übersetzte

Metadatenfelder

1. Titel (English)
2. Beschreibung (English)
3. Enthalten in (English)
4. Kommentare (English)

Ergebnisse

- Überwiegend gute Übersetzungen
- Problematisch: Schweizer Ortsnamen wie "Brunnen« (fountain), "Baden« (bathing) oder "Speicher« (memory).
- Abkürzungen wie "v. S." [von Süden], "v. N. O." [aus Nord-Ost] werden nicht anerkannt
- Schweizer Begriffe wie ARA (Abwasserreinigungsanlage) werden nicht erkannt
- Übersetzen Sie jede Sprache einzeln

Autotranslate with DeepL Pro in Cumulus backend

Autotranslate Control Panel: Status setzen

Field Name	Field Content
Record Name	Com_L23-0914-0001-0003
AutoTranslate Start Beschreibung	erledigt
Caption (German)	ja nein erledigt no value
AutoTranslate Start Titel	erledigt
Title (German)	! London, Chelsea, 352 King's Road
AutoTranslate Start Enthalten in	erledigt
Is part of (German)	! London, 1974. Reportage mit 68 Bildern (Auswahl digitalisiert, keine zusätzlichen Motive vorhanden)
AutoTranslate Start Kommentare	erledigt
Comments (German)	! Anonym: (wahrscheinlich) Chelsea, 313 King's Road (Chelsea Rare Books)

Autotranslate result



Field Name	Field Content
Caption	! Natural science antique shop with instruments
Title	! London, Chelsea, 352 King's Road
Is part of	! London, 1974. reportage with 68 pictures (selection digitized, no additional motifs available)
Comments	! Anonymous: (probably) Chelsea, 313 King's Road (Chelsea Rare Books)

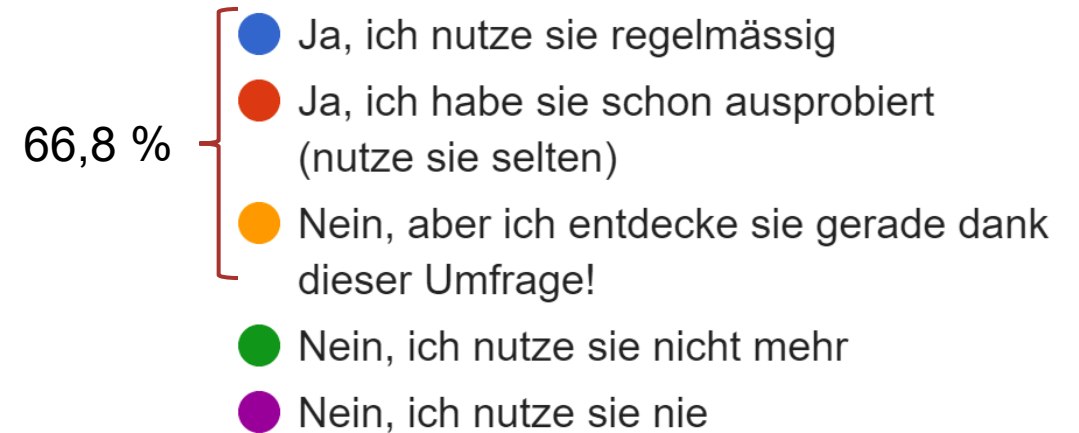
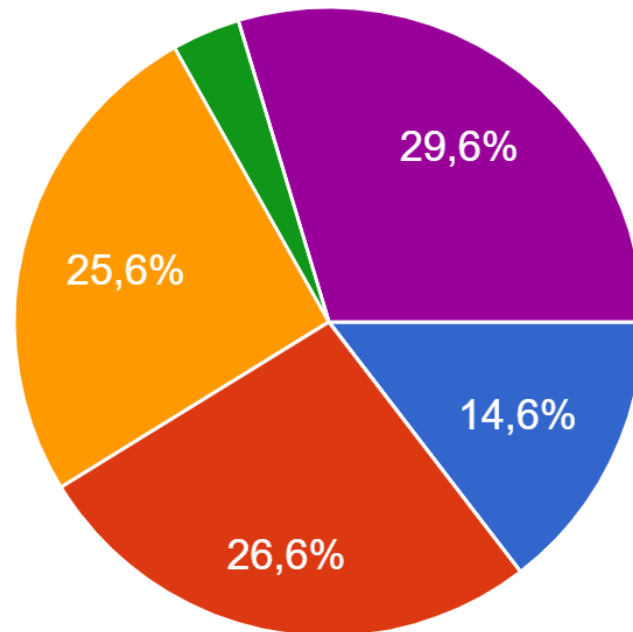
Aufwand & Kosten

	Autotagging	Autotagging Colour	Autotranslate
Anzahl Datensätze	825'804	503,262	1'142'019
Ø-Sekunden / Datensatz	20-30	2	3
Datensätze / Nacht	1'899	7,700	8'800
Startdatum	21.02.2021	17.05.2022	15.03.2022
Enddatum	16.05.2022	15.07.2022	23.12.2022
Anzahl produktive Ist-Nächte (Mo-So)	449	65	112
Kosten	USD 1.20 / 1000 unit = USD 1594		EUR 500 / 25 Mio. Zeichen = EUR 2700

Benutzer:innenbefragung im Mai 2022: «Nutzen Sie die Keywords?»

Nutzen Sie die Keywords (gelb markiert)?

199 Antworten



Online-Umfrage Mai 2022, Bericht

2. Durchlauf: Qualitätskontrolle

Juli 2023 bis heute

Die erste Kritik wächst...

Befund

- Alphabetische Stichwortliste mit 4'600 Tags
- Tags sind allgemeiner, geben viele Details an, sind teilweise sehr gut
- Tags wurden zunächst weder überprüft noch korrigiert... Das «verbieten» wir auch unserer Crowd!
- Wir wollten die Autotags so lassen, wie sie waren (zu grosser Aufwand)

Doch die Kritik wächst

- Bei genauerem Hinsehen fanden wir aber auch schlichtweg falsche, problematische oder widersprüchliche Tags
- Der Unterschied zwischen Keywords (KI) und Schlagwörtern ist im Web-Frontend nicht klar
- Mitarbeiter:innen kritisieren, unsere intellektuelle Arbeit würde verwässert
- Akademiker:innen kritisieren, wir würden dadurch Stereotype und Vorurteile zu reproduzieren
- Immer mehr Vorträge von GLAM-Institutionen zum Thema KI zeigen die gleichen Ergebnisse: Katzen und Hunde. Aber ist es das?

Kategorien

Kategorien durchsuchen...

ETHBIB.Bildarchiv

- + Bestände
- + Fotografen
- + Genre
- Keywords**

General

- Abakus
- Abbild
- Abbildung
- Abblocken
- Abdeckung
- Abend
- Abenddämmerung
- Abendessen
- Abenteuer
- Abfahrt
- Abfall
- Abfalleimer
- Abfluss



Bildinformationen

Bildcode: Ans_01693-005
Fotograf: Unbekannt
Titel: Zürich, ETH Zürich, Hauptgebäude (HG), Hauptbibliothek, Katalogsaal, ETHICS Online-Abfrage

Keywords vs. Kategorien vs. Schlagwörter

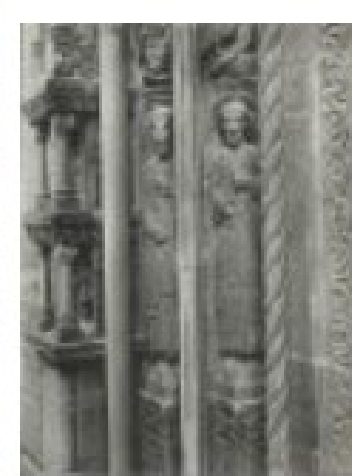
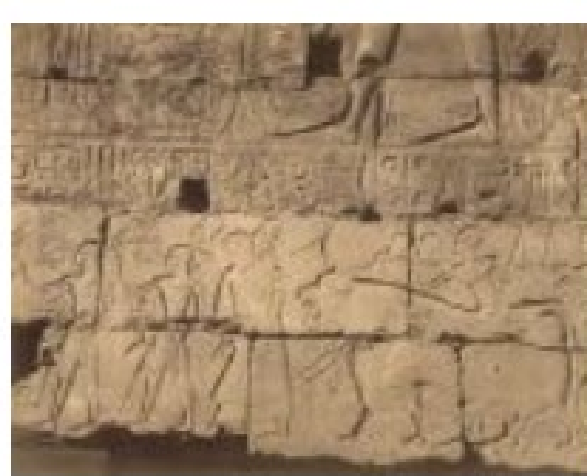
Farbe: schwarz/weiss

Form: Querformat

Mass: 9 x 13 cm

Kategorien: Ansichtensammlung Archiv + Bibliothek ETHICS Geräte der Informationstechnik (Computer) Reportagefotografien Unbekannt Zürich-Oberstrass

Schlagwörter: Büro Computer Erwachsener Innenaufnahme Konzentration Mann Menschen Möbel Schreiben Schreibtisch Technologie Unternehmen

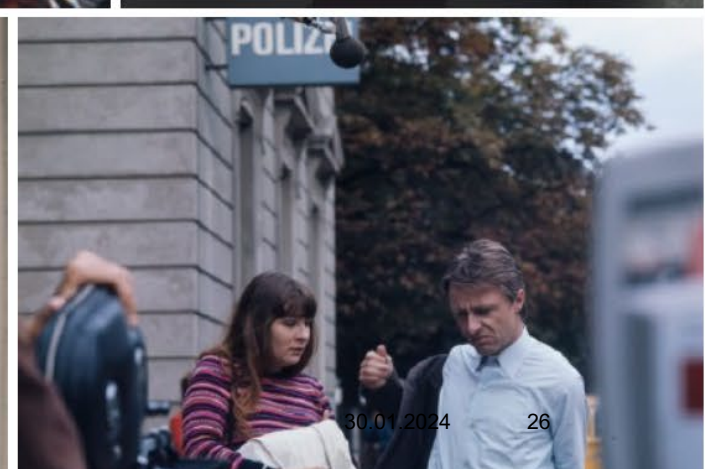


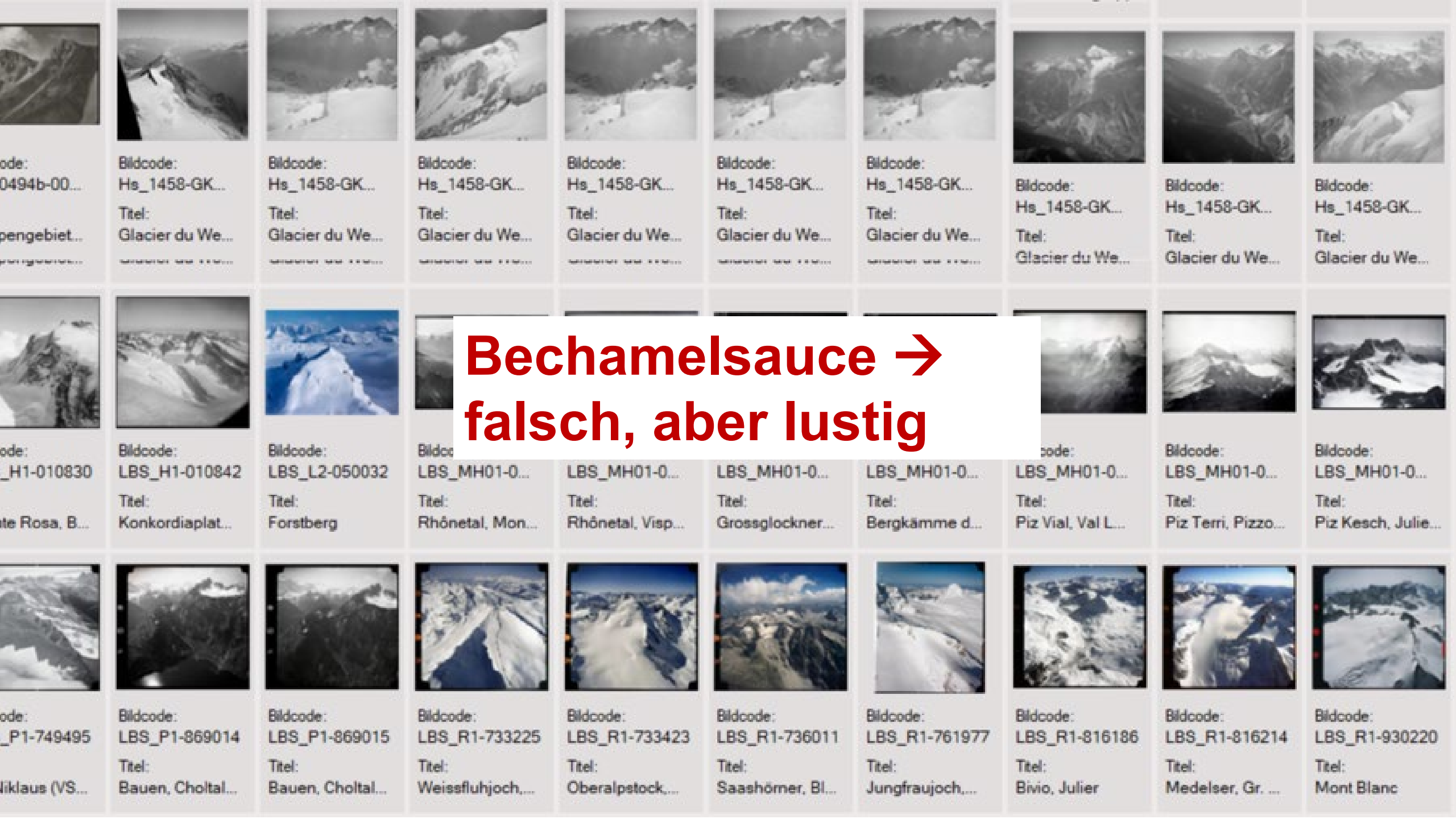
**Erleichterung →
Fehlübersetzung von
relief**





Verbrechen → flop





Bildcode:
0494b-00...

Titel:
engebiet...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
Hs_1458-GK...

Titel:
Glacier du We...

Bildcode:
_H1-010830

Titel:
te Rosa, B...

Bildcode:
LBS_H1-010842

Titel:
Konkordiaplat...

Bildcode:
LBS_L2-050032

Titel:
Forstberg

Bildcode:
LBS_MH01-0...

Titel:
Rhônental, Mon...

Bildcode:
LBS_MH01-0...

Titel:
Rhônental, Visp...

Bildcode:
LBS_MH01-0...

Titel:
Grossglockner...

Bildcode:
LBS_MH01-0...

Titel:
Bergkämme d...

Bildcode:
LBS_MH01-0...

Titel:
Piz Vial, Val L...

Bildcode:
LBS_MH01-0...

Titel:
Piz Terri, Pizzo...

Bildcode:
LBS_MH01-0...

Titel:
Piz Kesch, Julie...

Bildcode:
_P1-749495

Titel:
niklaus (VS...

Bildcode:
LBS_P1-869014

Titel:
Bauen, Choltal...

Bildcode:
LBS_P1-869015

Titel:
Bauen, Choltal...

Bildcode:
LBS_R1-733225

Titel:
Weissfluhjoch...

Bildcode:
LBS_R1-733423

Titel:
Oberalpstock...

Bildcode:
LBS_R1-736011

Titel:
Saashörner, Bl...

Bildcode:
LBS_R1-761977

Titel:
Jungfrauoch...

Bildcode:
LBS_R1-816186

Titel:
Bivio, Julier

Bildcode:
LBS_R1-816214

Titel:
Medelser, Gr. ...

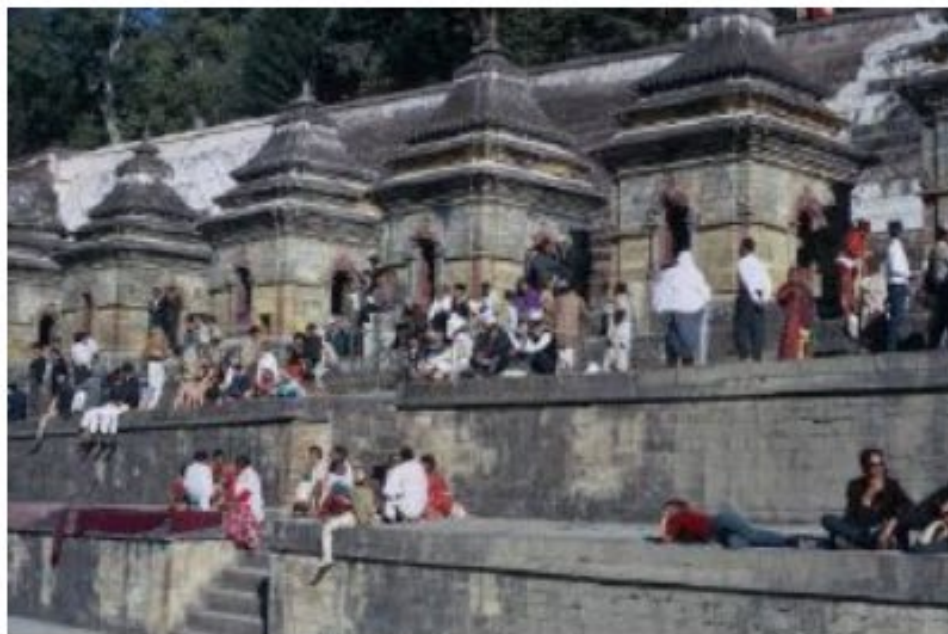
Bildcode:
LBS_R1-930220

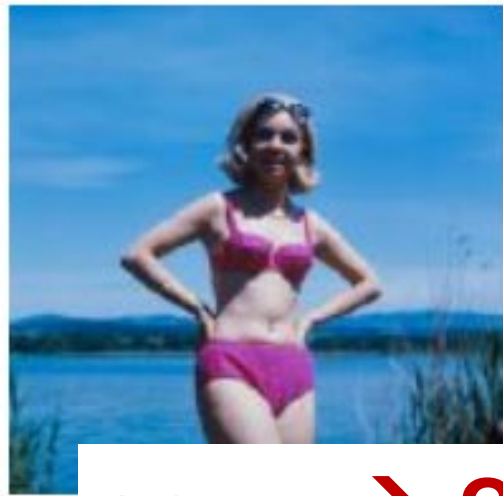
Titel:
Mont Blanc

Bechamelsauce → falsch, aber lustig

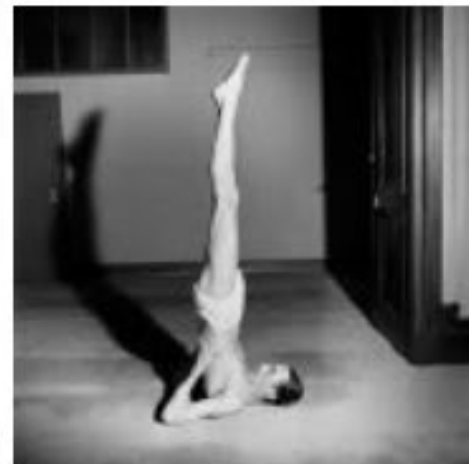
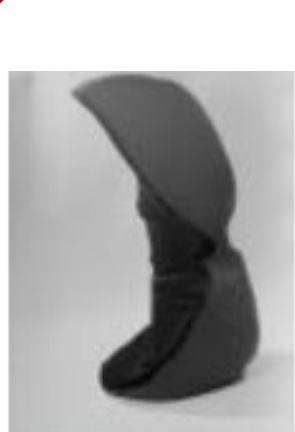


Ritual → biased





sexy → ???



Analyse der Autotags (in Arbeit)

Nach dem 1. Durchlauf: **4'585 Tags** von aasfresser bis Zypressenholz.

Folgende Korrekturen werden vorgenommen

1. **Merge** von Klein-/Grossschreibung: **3'914 Tags**
2. **Visuelle Analyse jedes Tags und Löschung** von falschen/problematischen Autotags, **zusätzliche Kriterien** wie z.B.
 - Hauptsächlich fehlerhafte Tags mit **Treffern >10 Bilder**
 - Hauptsächlich falsche Tags mit **wenigen Treffern**, bei denen aber viel mehr Bilder verfügbar wären
 - Hauptsächlich fehlerhafte Tags mit **nicht ETH-relevanten Themen**
3. **86 % der analysierten Tags** (nur «S» in Bearbeitung): 1'486 Tags sind in Ordnung (44 %), 1877 Tags werden gelöscht
4. **Nächste Schritte: Semantische Auswertung der Tags: Welche Tags funktionieren, welche nicht?**
 - **Wortart:** Verb, Substantiv, Adjektiv, Pronomen, ...
 - **Sachgruppen:** Architektur, Natur, Pflanzen, Menschen, Lebensmittel, Konzepte,...

3. Durchgang: Navigu

Januar 2024

Experimentieren mit Navigu: explorative Suche von 1 Million Bildern

Ähnlichkeitssuche und visuelle Bildsortierung (work in progress)

- Als Ergänzung zur traditionellen Textabfrage
- Beispielbilder hochladen
- Bilder werden in einem visuell sortierten 2D-Raster angezeigt, gruppiert nach Ähnlichkeit
- Interaktive Kartenfunktion zur Erkundung von verwandten Konzepten

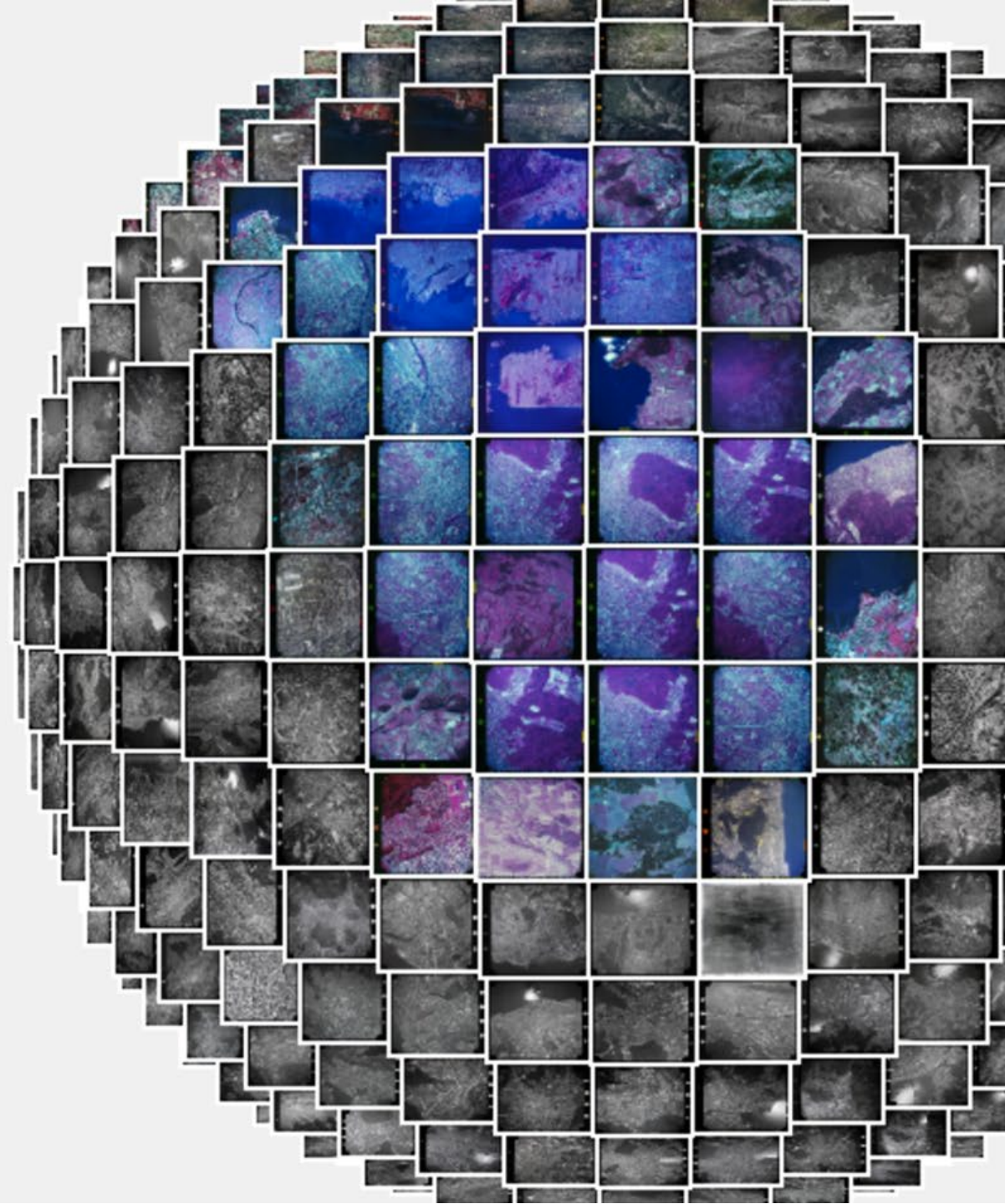
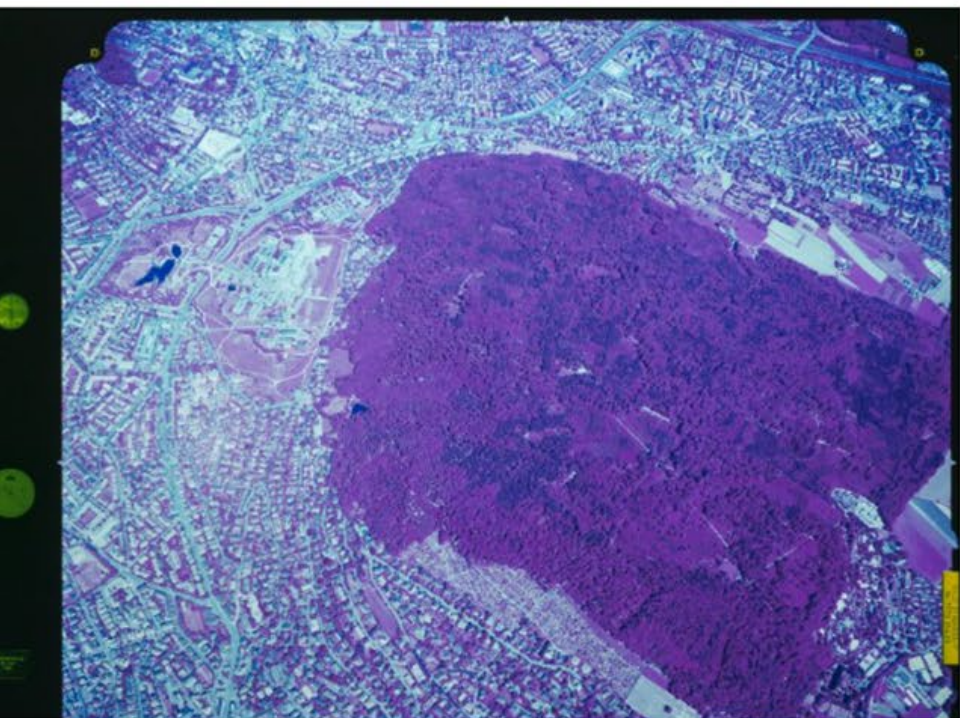
Technik

- Bildähnlichkeitsberechnung mittels Vektoren
- Keine Ausgabe von Autotags
- Aufwand: < 2 Wochen (!)

Select
Image Set

ETH Images (884878)

Text Search Input



Fazit

- **Pragmatischer Ansatz**
 - "Es gibt eine Chance, lasst sie uns nutzen."
 - KI kann ein wertvolles Werkzeug für die Indizierung von Bilddaten sein.
 - Zusätzliche Wissens Ebenen können den Nutzern helfen, unsere Inhalte auf eine neue Art und Weise zu entdecken.
 - Es ist jedoch wichtig, die Grenzen der KI zu erkennen und diese bei der Anwendung zu berücksichtigen.
- **Es ist eher ein Prozess als ein Projekt**
- **Nächste Schritte**
 - Die qualitative Analyse liefert wertvolle Erkenntnisse für die Entwicklung und Anwendung von KI-gestützten Bildklassifikationsmethoden.
 - Definition spezifischerer Aufgaben und Vorgehensweisen, auch in Zusammenarbeit mit Forschenden aus der Informatik und/oder den Digital Humanities.
 - Teilnahme an Diskussionen über ethische Rahmenbedingungen.

Nicole Graf
nicole.graf@library.ethz.ch

ETH Zürich
ETH-Bibliothek
Leitung Bildarchiv
Rämistrasse 101
8092 Zürich
Tel. +41 44 632 80 81

Twitter: [@niggegraf](https://twitter.com/niggegraf)

[E-Pics Bildarchiv](#)
[Navigu](#)
Google Bilder
[sMapshot](#)
[Wikimedia Commons](#)

[Arthistoricum.net](#)
[Luftbildindex](#)

[Blog Crowdsourcing](#)
[Twitter-Kanal](#)