


L'intelligence artificielle dans la description d'images: exemple pratique d'Archive d'images de l'EPFZ

Presentation

Author(s):

Graf, Nicole 

Publication date:

2024-03-05

Permanent link:

<https://doi.org/10.3929/ethz-b-000663069>

Rights / license:

Creative Commons Attribution-ShareAlike 4.0 International



L'intelligence artificielle dans la description d'images : exemple pratique d'Archive d'images de l'EPFZ
Nicole Graf

Introduction à l'Archive d'images

Faits et chiffres



- Créée en 2001 en tant qu'agence d'image
- Open Data en 2015
- 6 salariés permanents (470%) et 7 assistants étudiants (180%)
- 3,6 millions d'images analogiques, dont 1,3 millions numérisées
- Numérisation par an : 60'000 à 80'000
- 947 422 images en ligne sur [E-Pics Image Archive](#)
- Collections:
 - Professeurs et instituts de l'EPF,
 - Vues & Portraits,
 - Photographie aérienne,
 - Swissair, Comet Photo AG, etc.

Catalogage – plusieurs couches de compétences

Plusieurs couches de compétences



Machines



Crowd



Experts



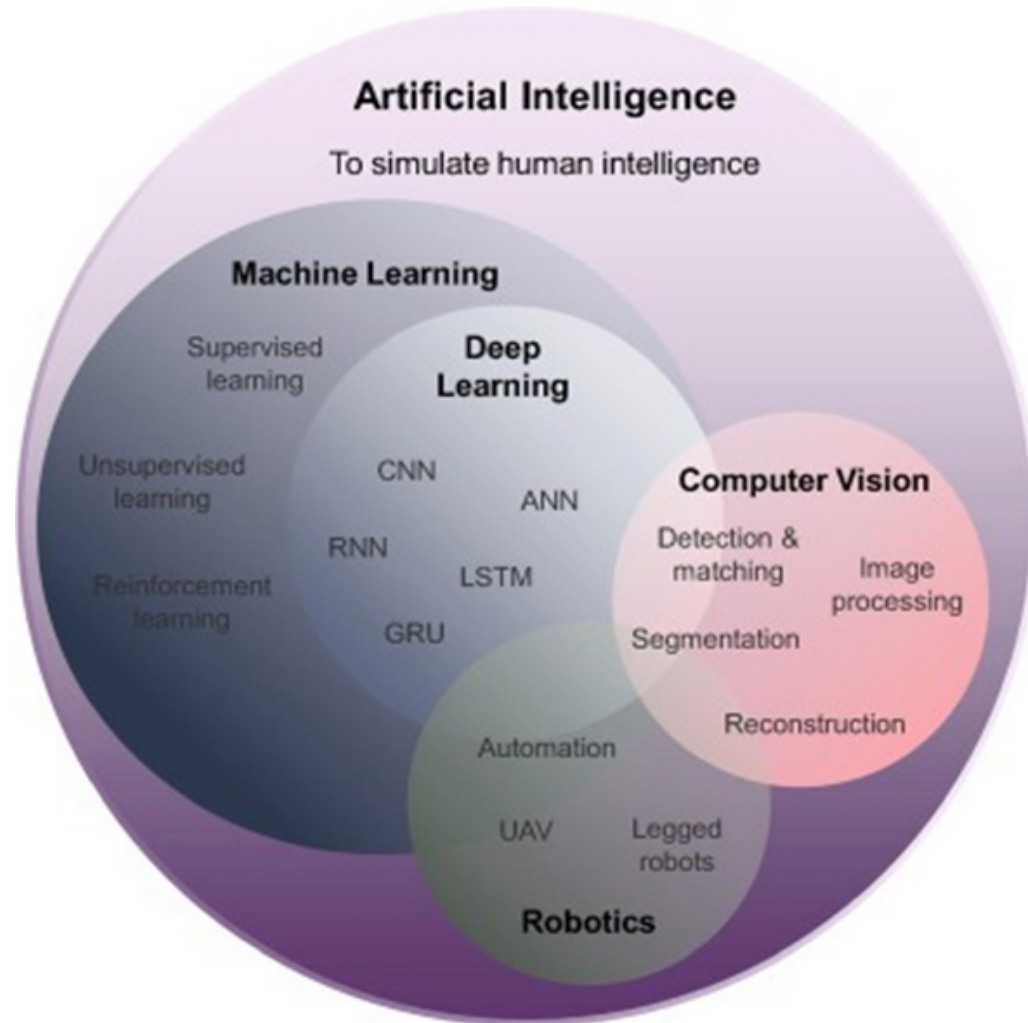
Bibliothécaires

Plusieurs couches de compétences



Machines

Qu'est-ce que la Computer Vision (CV)?



CV ou « AI for images » - que voient les machines

- Sous-domaine de l'IA et du ML
- Technologies conçues pour « aider » les ordinateurs à « voir » et à décrypter le contenu des images numériques – photos et vidéos
- Reproduire et automatiser la capacité de la vision humaine
- Analyser une image au niveau des pixels, à partir de leurs propres valeurs de couleur
- Avec l'aide de capteurs et d'algorithmes pour récupérer, traiter et comprendre des informations visuelles
- Algorithmes basés sur la reconnaissance de formes

Comment la machine devient-elle intelligente ? Par l'entraînement!



Machine Learning



- 1000 images représentatives par objet (concept)
- Catégorisation manuelle par des humains (les données sont entraînées → MTurks)
- Bonnes données d'entraînement (pas seulement des chats noirs sur de l'herbe verte)
- Entraîner l'IA ou Machine learning
- Programmation de l'algorithme (ML) et des règles correspondantes
- L'algorithme ML est ensuite entraîné à l'aide des données
- Sur la base des données, l'algorithme d'apprentissage automatique commence à reconnaître les modèles et les structures et à former des modèles de manière indépendante
- Les modèles constituent alors la base de la prédiction des probabilités et des valeurs.

Applications au quotidien

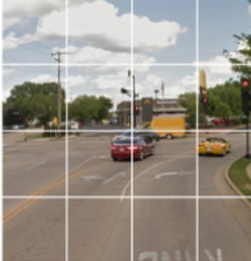






Bildersuche
Suche auf Google mit einem Bild statt mit Text. Ziehe das Bild hierher.

Bild hier ablegen



Select all squares with
traffic lights
If there are none, click skip



    **SKIP**



Domaines d'application de la CV

- **Retouche d'images** : par exemple, amélioration des différences de luminosité et de contraste (filtres photo)
- **Écriture et reconnaissance de caractères** (OCR, Handwriting Recognition HCR)
- **Reconnaissance d'objets** : interaction complexe entre l'extraction de caractéristiques, la reconnaissance de formes, l'extraction de couleurs et les algorithmes décisionnels auto-apprenants de l'IA
- **Reconnaissance des personnes** : reconnaissance faciale (âge, culture, sexe), reconnaissance des expressions faciales, biométrie
- **Classification d'objets** : Descriptions d'images ou classification /balisage (« autotagging »)
- **Recherche inversée d'images** : Reverse Image Search, identification basée sur le contenu
- **IA générative** : génération d'images (Dall-e, Midjourney, Bing Image Creator) et reconnaissance d'objets (ChatGPT-4, Bing)

1^{ère} phase avec: Clarifai et DeepL

février 2021 à mai 2022


Expérimenter avec Clarifai : Autotagging d'1 million d'images

Nouveaux Tags générés automatiquement

- Jusqu'en 2021, les classes de métadonnées étaient principalement basées sur les connaissances humaines
- À partir de 2021 : mots-clés générés automatiquement par l'intelligence artificielle (IA), ce que l'on appelle l'autotagging
- Autotags en complément des mots-clés « intellectuels »
- Logiciel d'IA : Clarifai, modèle « général »
- Pourquoi choisir Clarifai ? Il faut prendre ce que l'on reçoit. Le module d'auto-tagging fait partie de notre logiciel de gestion documentaire.

Autotagging avec Clarifai dans le 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"/>

➔ Définir l'état, sélectionner le modèle, activer la couleur




▼ 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

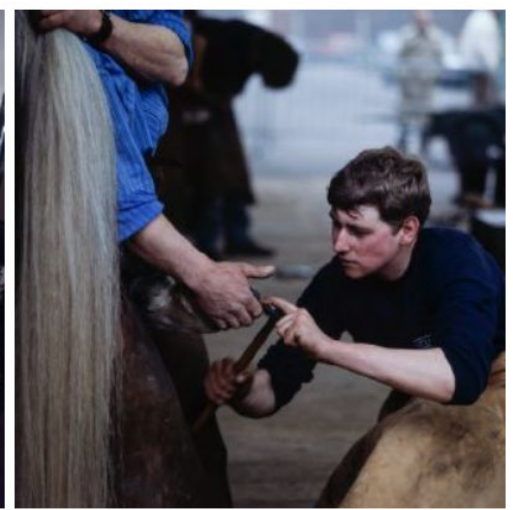
▼ AutoTagging Control	
AutoTagging	AutoTagging noch nicht gestartet
Autotag Model	General
AutoTagging Color	General
AutoTagging Klassifizie...	Wedding
AutoTagging Face	Travel
	Food
	Apparel
	kein Wert

▼ 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"/>

➔ Résultat Autotagging

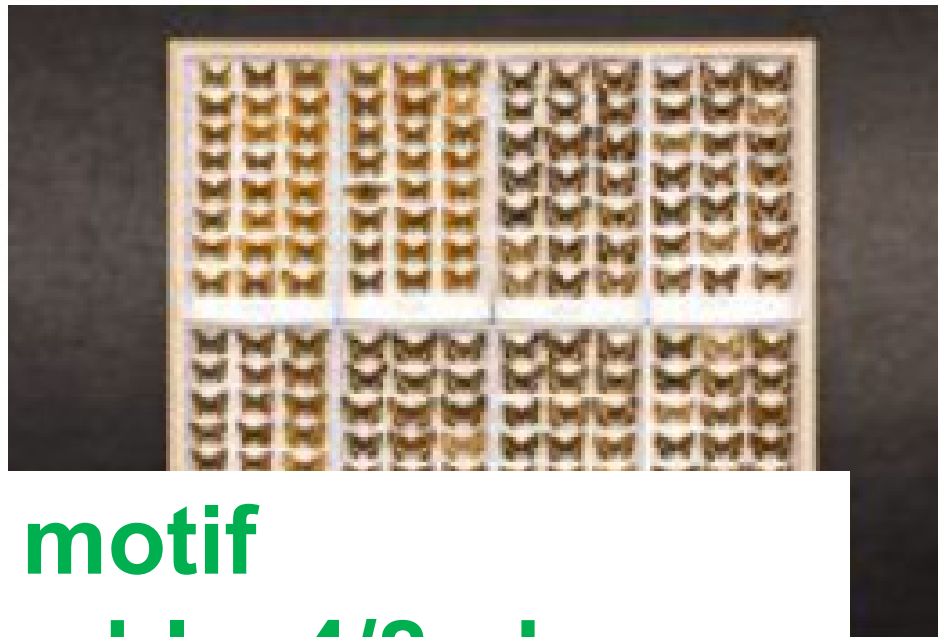
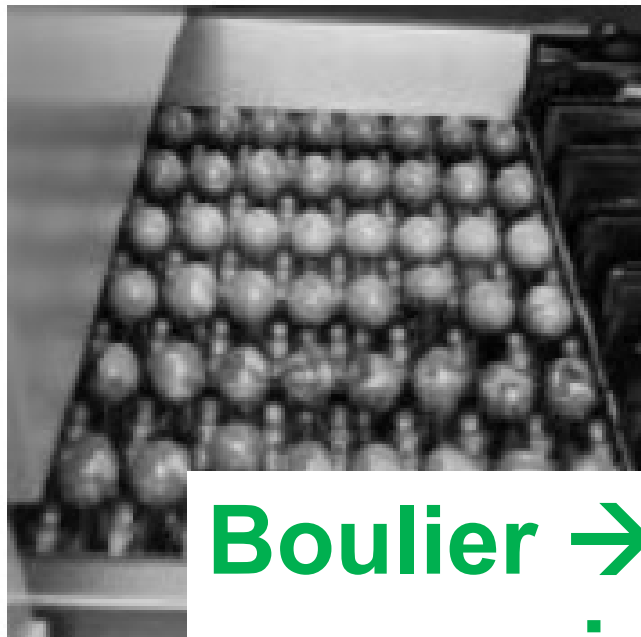


Feldname	Feldinhalt
Bildcode	Dia_385-14415
Miniatur	
Wertung
▼ AutoTagging Control	
AutoTagging	Autotagging erfolgreich
Autotag Model	General
AutoTagging Color	<input checked="" type="checkbox"/>
AutoTagging Klassifizie...	<input type="checkbox"/>
AutoTagging Face	<input type="checkbox"/>
▼ AutoTagging 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 05.03.2024 13
AutoTaggingColor 5	kein Wert

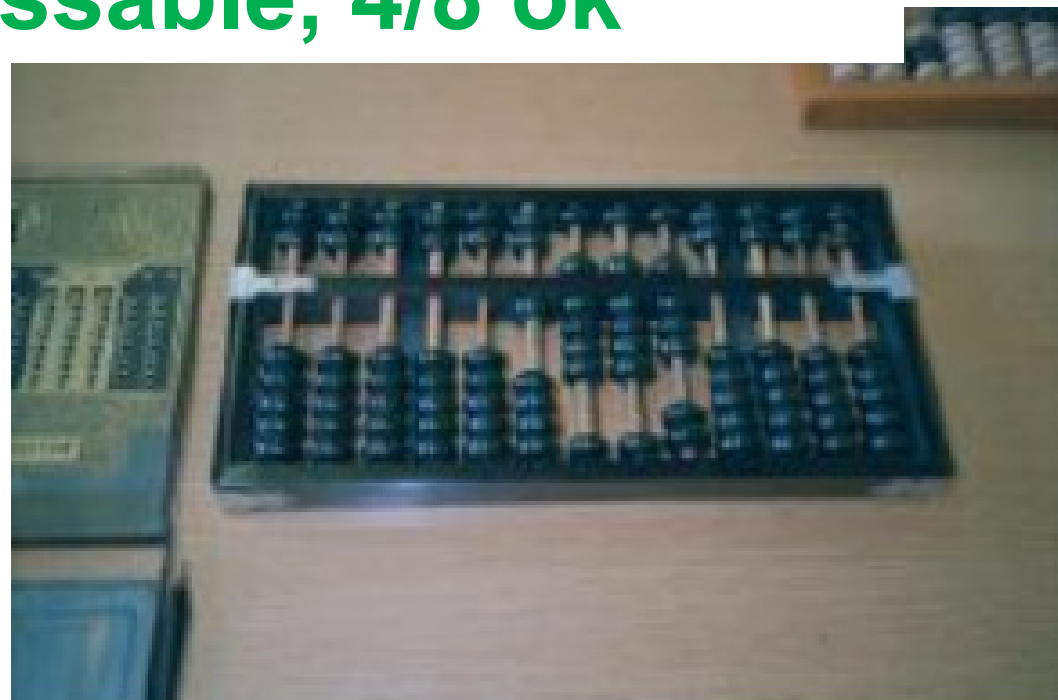


batteur → top





Boulier → motif reconnaissable, 4/8 ok



Kategorien ^

schreiben ⊗

ETHBIB.Bildarchiv

Schreiben ≡

Slate (Schreiben)

mit Bild v

Lizenz v

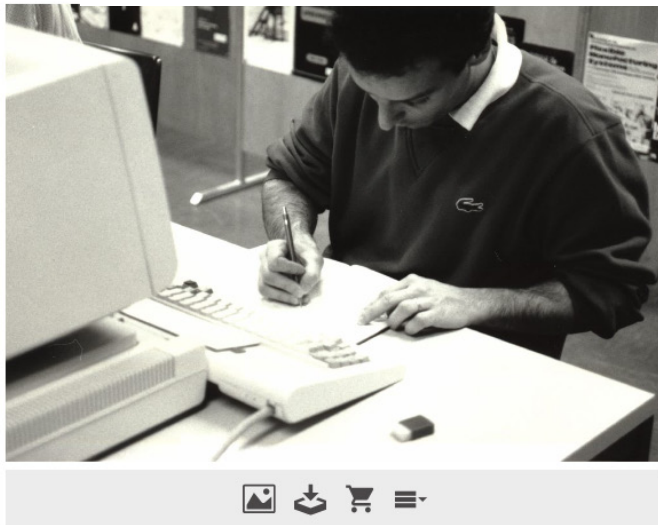
Form v

Farbe v

Farbauswahl v

Zeitraum v

Écrire (schreiben) → nouveaux résultats

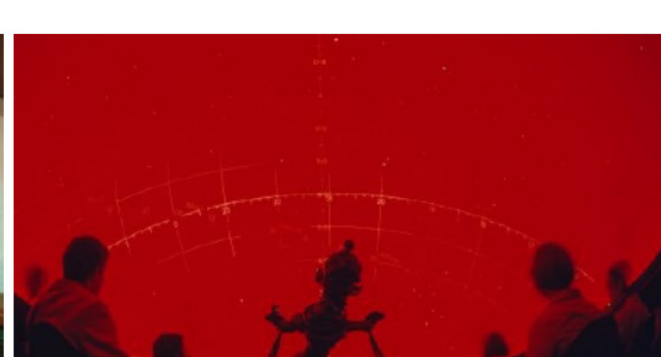
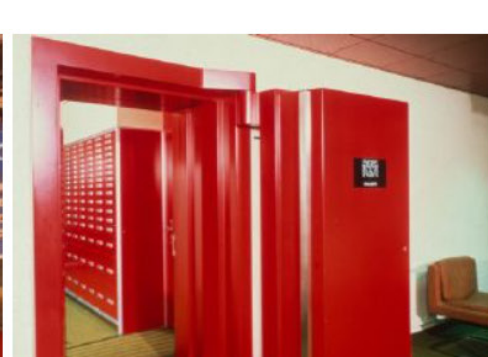


Kategorien: ⓘ

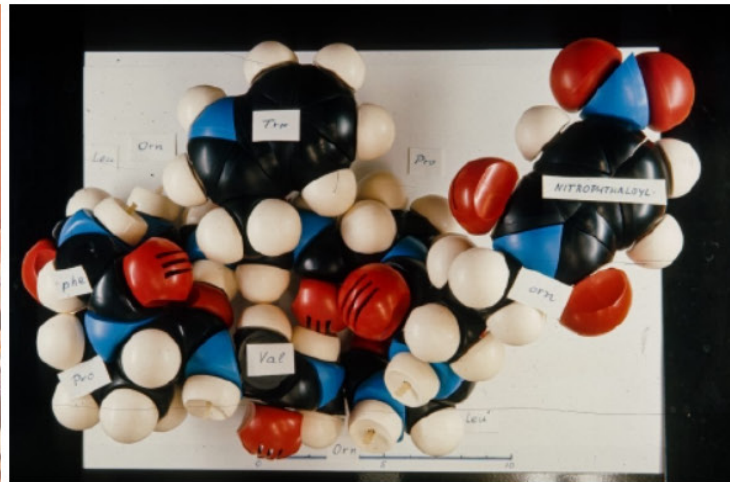
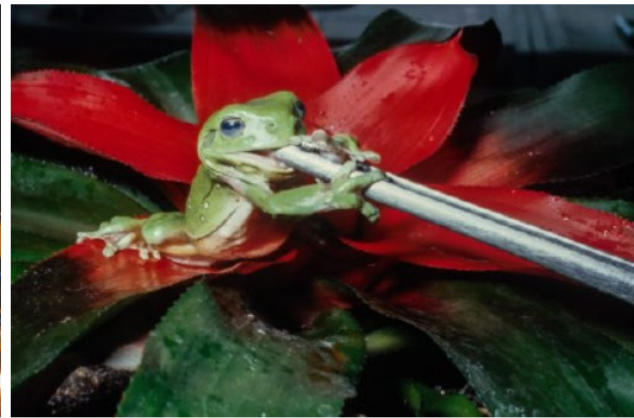
- Ansichtensammlung
- Archiv + Bibliothek
- ETHICS
- Geräte der Informationstechnik (Computer)
- Reportagefotografien
- Unbekannt
- Zürich-Oberstrass

Schlagwörter:

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



Couleur: FireBrick



DeepL (2022–)

Nouveaux champs de métadonnées traduits automatiquement

1. Titre (anglais)
2. Description (Anglais)
3. Inclus dans (Deutsch)
4. Commentaires

Résultats

- Les traductions sont généralement bonnes
- Problématique : interprétation des noms de lieux tels que Brunnen (« fontaine »), Baden (« bain ») ou Speicher (« mémoire »).
- Des abréviations telles que « v. S. » [du sud], « v. N. O. » [du Nord-Ouest] ne sont pas reconnues
- Les termes suisses tels que STEP (station d'épuration des eaux usées) ne sont pas reconnus
- chaque langue traduite séparément

Traduction automatique avec DeepL Pro dans le backend Cumulus

Panneau de configuration de traduction automatique : Définir le statut

Field Name	Field Content
Record Name	Com_L23-0914-0001-0003
AutoTranslate Start Beschreibung	erledigt
Caption (German)	ja nein
AutoTranslate Start Titel	erledigt no value
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 resultat



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)

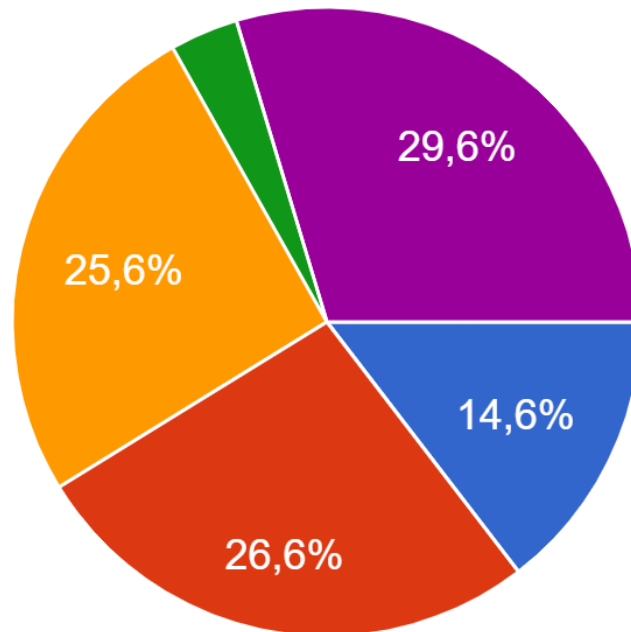
Efforts et coûts

	Autotagging	Autotagging Colour	Autotranslate
Nombre de données	825'804	503,262	1'142'019
Ø-secondes / enregistrement	20-30	2	3
Données / Nuit	1'899	7,700	8'800
Début	21.02.2021	17.05.2022	15.03.2022
Fin	16.05.2022	15.07.2022	23.12.2022
Nombre de nuitées productives (du lundi au dimanche)	449	65	112
Coût	USD 1.20 / 1000 unit = USD 1594		EUR 500 / 25 Mio. Signe = EUR 2700

Enquête auprès des utilisateurs en mai 2022 : « Utilisez-vous les mots-clés ? »

Nutzen Sie die Keywords (gelb markiert)?

199 Antworten



66,8 %

- Ja, ich nutze sie regelmässig
- Ja, ich habe sie schon ausprobiert (nutze sie selten)
- Nein, aber ich entdecke sie gerade dank dieser Umfrage!
- Nein, ich nutze sie nicht mehr
- Nein, ich nutze sie nie

Online-Umfrage Mai 2022, Bericht

2^{ème} phase : Contrôle qualité

juillet 2023 à aujourd'hui

Premières critiques...

Résultats

- Liste alphabétique de mots-clés avec 4 600 autotags
- Les tags sont plus généraux, donnent beaucoup de détails, sont parfois très bons
- Les tags n'ont pas été vérifiés ou corrigés dans un premier temps... Nous « interdisons » aussi au volontaires de le faire !
- Nous voulions laisser les autotags tels qu'ils étaient (trop d'efforts)

Mais les critiques se multiplient

- En y regardant de plus près, cependant, nous avons également trouvé des balises tout simplement erronées, problématiques ou contradictoires
- La différence entre les tags (IA) et les mots-clés n'est pas claire sur le frontend web
- Les employés critiquent l'édulcoration de leur travail intellectuel
- Les universitaires critiquent le fait que cela permettrait de reproduire des stéréotypes et des préjugés
- De plus en plus de conférences données par des institutions GLAM sur le thème de l'IA montrent les mêmes résultats : des chats et des chiens. Mais est-ce vraiment tout ?

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

Autoren: ETH Zürich, Hauptgebäude (HG), Hauptbibliothek, Katalogsaal, ETHICS Online-Abfrage

Keywords vs. Kategorien vs. Schlagwörter

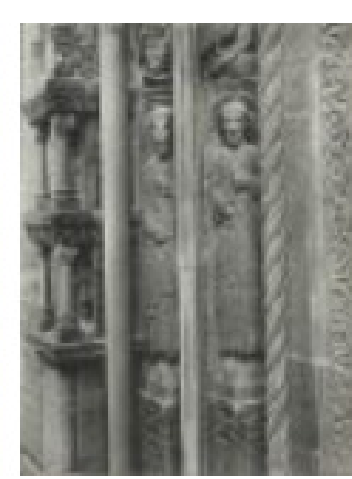
Farbe: schwarz/weiss

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

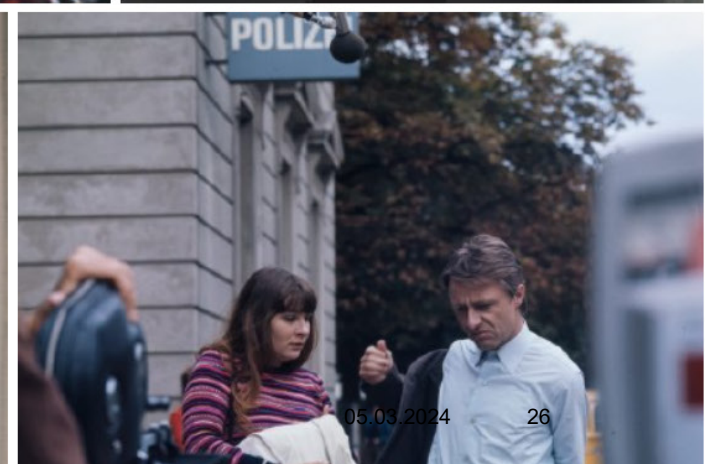


**Soulagement →
Erreur de traduction
de «relief»**





Crime → flop





Bildcode:
0494b-00...
Titel:
engebiet...
progr...

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:
Rhônetal, Mon...

Bildcode:
LBS_MH01-0...
Titel:
Rhônetal, 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. ...

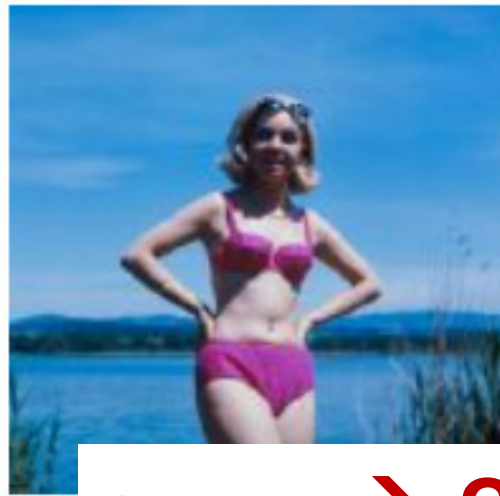
Bildcode:
LBS_R1-930220
Titel:
Mont Blanc

**Sauce béchamel →
Faux, mais drôle**

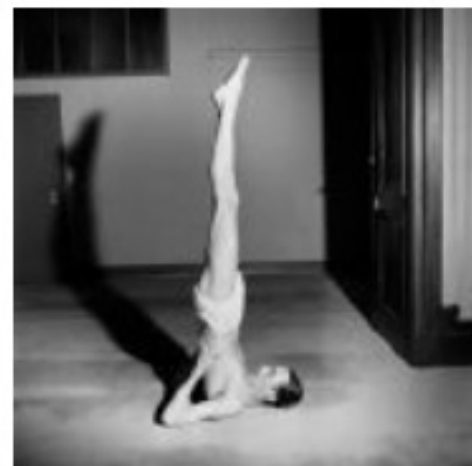


Rituel → biaisé





sexy → ???



Analyse des tags automatiques (en cours)

Après le 1er contrôle : 4 585 autotags de A à Z (von Aasfresser bis Zypressenholz). Les corrections suivantes seront apportées

1. **Fusion** de minuscules/majuscules : 3 914 autotags
2. **Analyse visuelle de chaque autotag et suppression des autotags incorrects/problématiques**, critères supplémentaires tels que
 - Tags erronés avec le mot résultats >10 images
 - Tags erronés avec peu de résultats, mais où beaucoup plus d'images seraient disponibles
 - Tags avec des sujets non pertinents pour l'ETH
3. **86% des tags analysés jusqu'ici (« S » en cours d'analyse) : 1 486 tags sont OK (44%), 1877 tags sont supprimés**
4. **Prochaines étapes : Évaluation sémantique des balises : quelles balises fonctionnent, lesquelles ne fonctionnent pas ?**
 - **Genre**: verbe, nom, adjectif, pronom, ...
 - **Thématiques**: architecture, nature, plantes, personnes, nourriture, concepts,...

3^{ème} phasee : Navigu

Janvier 2024

Expérimentation avec Navigu : recherche exploratoire avec 1 million d'images

Recherche de ressemblances et tri visuel d'images (travail en cours)

- En complément de la requête textuelle traditionnelle
- Télécharger des exemples d'images
- Les images sont affichées dans une grille 2D, triées visuellement, regroupées par similarité
- Une fonction de carte interactive permet d'explorer les concepts connexes

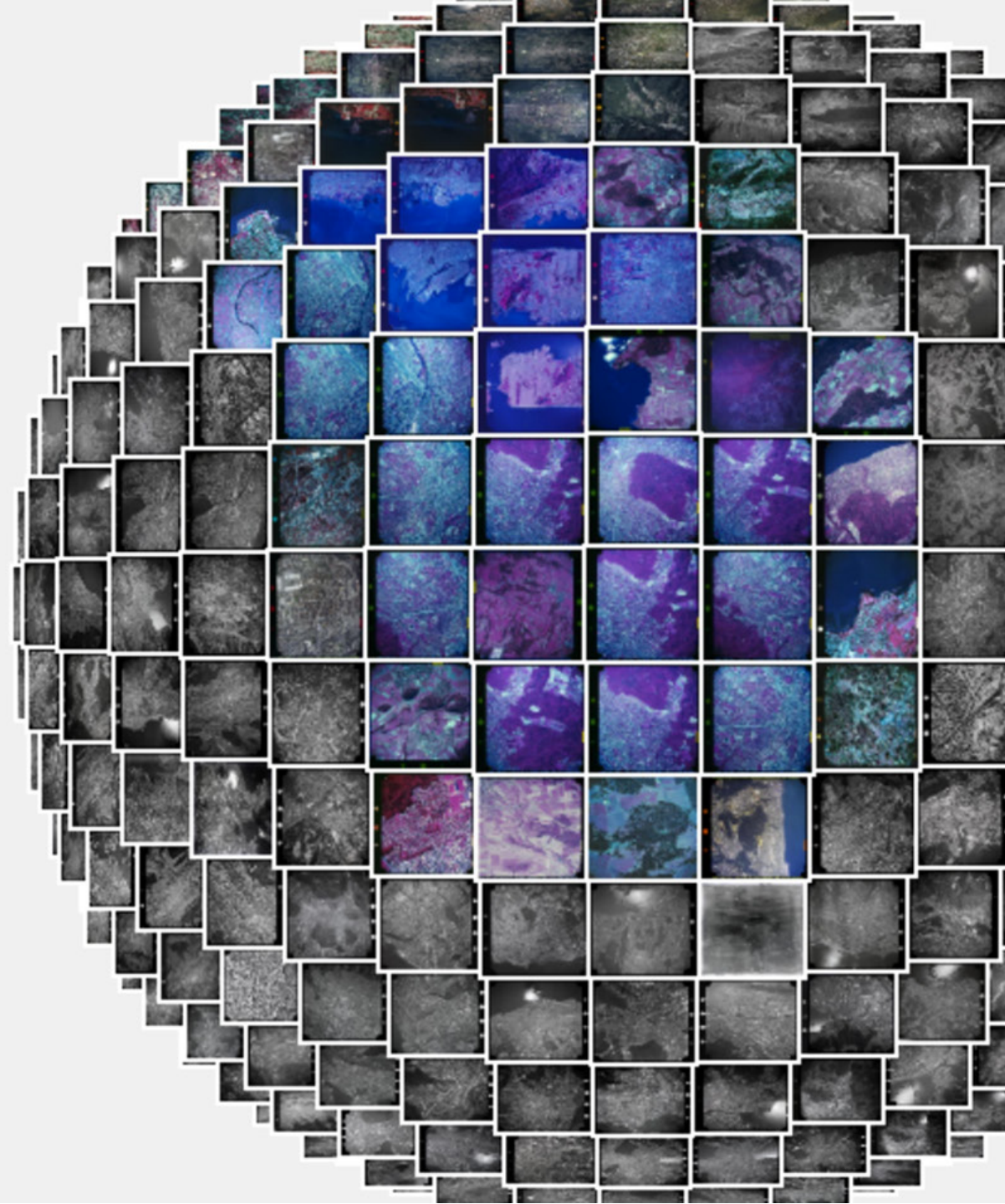
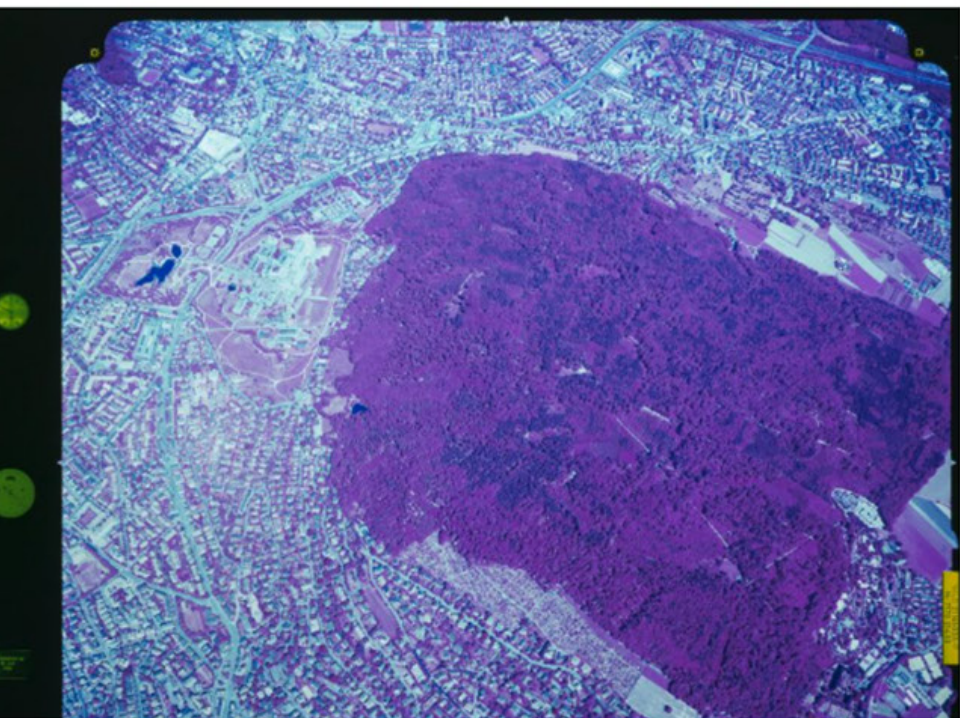
Technologie

- Calcul de similarité d'image à l'aide de vecteurs
- Pas de sortie par autotags
- Effort : < 2 semaines (!)

Select
Image Set

ETH Images (884878)

Text Search Input



Conclusion

Approche pragmatique

- « Il y a une opportunité, saisissons-la. »
- L'IA peut être un outil précieux pour l'indexation du contenu des images.
- Des couches de compétences supplémentaires peuvent aider les utilisateurs à découvrir notre contenu de nouvelles manières.
- Cependant, il est important de reconnaître les limites de l'IA et d'en tenir compte lors de leur application.

Il s'agit plus d'un processus que d'un projet

Étapes suivantes

- L'analyse qualitative fournit des informations précieuses pour le développement et l'application de méthodes de classification d'images assistées par l'IA.
- Définition de tâches et de procédures plus spécifiques, également en collaboration avec des chercheurs en informatique et/ou en humanités numériques.

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](#)