



# Intelligence artificielle et indexation des images

## Projets en cours et perspectives

Journées du patrimoine écrit : “L’image aura-t-elle le dernier mot ?  
Regards croisés sur les collections iconographiques en bibliothèques”

Jean-Philippe Moreux

**Bibliothèque Nationale de France**  
Département de la Coopération,  
service Coopération numérique et Gallica

Aix-en-Provence, 15-16 juin 2023



# Plan

1980-2023 : panorama

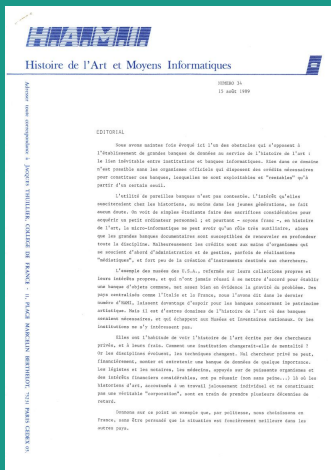
Techniques et projets

IA et histoire de l'art

Perspectives

1980-1990

# La vision



Jacques Thuillier, *Editorial, dans H.A.M.I. Histoire de l'art et moyens informatiques*, 34 (15 août 1989), p. 3, INHA, Archives 51/148. Cliché INHA

pris en considération cette source biographique) et qui exigent aujourd'hui une longue consultation à Paris, au Cabinet des manuscrits de la Bibliothèque Nationale...

Un second projet n'a pas été seulement mis au point et testé, mais une partie de son programme est dès maintenant couvert par Melle Marie-Claude Chaudonneret, grâce à la fois au C.N.R.S. et au ministère de la Culture, et surtout grâce à son dévouement et son érudition. Il s'agit d'une banque concernant les Salons parisiens du XIXe siècle, ou plutôt de la Révolution à la guerre de 1914. Quel historien de l'art français, quel documentaliste, quel marchand, n'a pas désespérément feuilleté les livrets de ces Salons pour identifier quelque tableau ou confirmer l'activité de quelque peintre ? Or c'est un travail fastidieux et dont les résultats sont toujours aléatoires. D'autre part il est utile de compléter les indications très sommaires des livrets par les albums qui paraissent conjointement, ceux de Landon de Béraud et de leurs nombreux émules, et par les comptes-rendus publiés à l'époque, riches de renseignements précis et de jugements significatifs, mais réclamant une recherche infinie.

Ici encore, l'informatique offre le recours miraculeux. Une banque relativement sommaire (exploitant les seuls livrets officiels) peut être montée très rapidement ; elle offrirait déjà des renseignements fort précieux. Les données fournies par les comptes-rendus pourraient y être jointes à mesure des dépouillements. **Mais surtout l'image devient ici essentielle.** Or elle existe, dès l'Empire, pour toute une partie des oeuvres, et il serait d'autant plus facile de la joindre qu'il s'agit d'une image en noir, et souvent au trait. Des crédits réduits et quelques vacataires suffiraient à monter un "thesaurus" de l'art du XIXe siècle comme il n'en existe nulle part encore, et qui, nous croyons pouvoir l'affirmer, modifierait fondamentalement l'idée que l'on se fait aujourd'hui de la production artistique du siècle passé.

On pouvait penser que l'Institut national mènerait à bien ces entreprises ; on pouvait espérer aussi qu'il permettrait d'animer des projets informatiques de grande envergure et ferait de la France un partenaire solide pour une collaboration internationale. Ici encore, citons deux exemples.

L'un concerne la banque CATART, dont le système a été monté et présenté dès 1982, et dont les dépouillements continuent depuis cette date, mais à un rythme sans rapport avec l'ampleur de la tâche. Rappelons que cette banque vise à permettre l'exploitation des catalogues de ventes d'art, de l'origine à nos jours. Il s'agit, non pas d'informatiser des éléments tirés de ces catalogues (parti qu'a choisi, quelque temps après la banque PROVENANCE INDEX du J. Paul Getty Museum), mais de laisser quiconque, selon ses intérêts particuliers, chercher à travers cette masse énorme les éléments qui le concernent, à propos de l'identification d'un tableau particulier, de l'établissement du catalogue raisonné d'un artiste, ou de l'histoire des collectionneurs et des collections, du goût d'une époque, du rôle du commerce, etc. Tout un pan de l'histoire de l'art s'est inscrit dans ces catalogues de ventes : Frits Lugt l'avait bien senti, qui en avait dressé le répertoire. Mais leur nombre

Salons parisiens du XIXe  
Banque de données  
Albums Landon Béraud

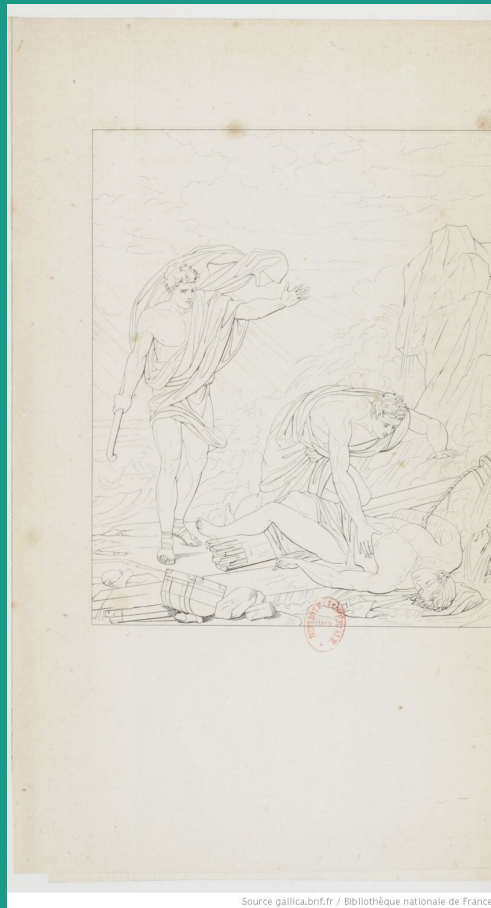
“Ici encore, l'informatique  
offre le recours miraculeux.

... Mais surtout l'image  
devient ici essentielle.”

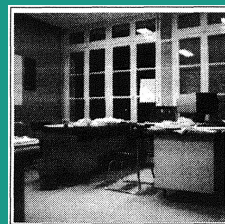
# 1980-1990

## La vision

Annales du Musée et de l'école moderne des beaux-arts. Salon de ... : recueil de morceaux choisis parmi les ouvrages de peinture et de sculpture exposés au Musée du Louvre ... par C. P. Landon, ... Landon, Charles-Paul (1760-1826). Béraud, Antony (1792-1860). Éditeur : chez C. P. Landon (Paris) Éditeur : Pillet aîné Date d'édition : 1808



Source gallica.bnf.fr / Bibliothèque nationale de France




Institut National Polytechnique de Lorraine  
Centre de Recherche en Informatique de Nancy Inria-Lorraine

77 INTL-088N  
1991 ZIOU, D.

### La détection de contours dans des images à niveaux de gris : mise en œuvre et sélection de détecteurs

## THÈSE

présentée et soutenue publiquement le 11 octobre 1991 pour l'obtention du Doctorat de l'Institut National Polytechnique de Lorraine (Spécialité Informatique) par Djemel ZIOU<sup>1</sup>



devant le jury composé de :

**Président :** Jean-Paul HATON  
**Rapporteurs :** Serge CASTAN  
René HUSSON  
**Examineurs :** Gérard GIRAUDON  
Gérald MASINI  
Roger MOHR

<sup>1</sup>Adresse de l'auteur à partir du 1/2/92 : EERIE-LERI, Parc Scientifique Jorge Besse, 30000 Nîmes



# 1990-2000

## Recherche par le contenu (CBIR)

The QBIC Project: Querying Images By Content Using Color, Texture, and Shape. w. Niblack R. Barber W. Equitz M. Flickner E. Glasman D. Petkovic P. Yanker, C. Faloutsos G. Taubin  
IBM Research Division, San Jose, 1993



The State Hermitage Museum: Digital Collection

HOME HELP SITE MAP FEEDBACK SERVICES CALENDAR TICKETS SHOP

INFORMATION COLLECTION HIGHLIGHTS EXHIBITIONS HERITAGE HISTORY CHILDREN& EDUCATION DIGITAL COLLECTION

### QUICK SEARCH

BROWSE •  
QBIC SEARCHES •  
ADVANCED SEARCH •

### QBIC COLOUR AND LAYOUT SEARCHES

Imagine finding a Gauguin masterpiece simply by recalling the organisation of his subjects or locating a Da Vinci painting by searching for its predominant colours. IBM's experimental Query By Image Content (QBIC) search technology offers this unique ability. Search for artwork visually using tools that an artist would use. For an overview of the QBIC searches, take a look at our animated demonstrations.

#### QBIC COLOUR SEARCH

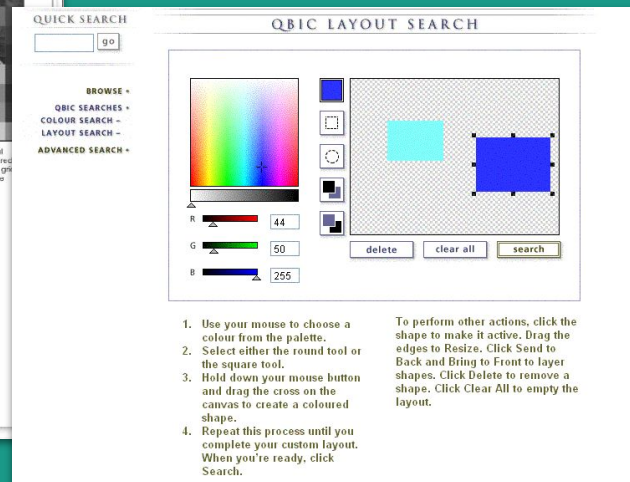
The QBIC Colour Search locates two-dimensional artwork in the Digital Collection that match the colours you specify. You select colours from a spectrum, define proportions, then execute the search. It really is that simple. Go to the [QBIC Colour Search Demo](#) to view a step by step demonstration of this search.

#### QBIC LAYOUT SEARCH

With the QBIC Layout Search, you become the artist. Using geometric shapes, you can arrange areas of colour on a virtual canvas to approximate the visual organisation of the work of art for which you are searching. Go to the [QBIC Layout Search Demo](#) to view a step by step demonstration of this search.



QBIC reinterprets the virtual canvas as a grid of coloured areas, then matches the grid to other images stored in the database.



### QUICK SEARCH

BROWSE •  
QBIC SEARCHES •  
COLOUR SEARCH -  
LAYOUT SEARCH -  
ADVANCED SEARCH •

### QBIC LAYOUT SEARCH

A color palette with a rainbow spectrum and sliders for Red (44), Green (50), and Blue (255). A virtual canvas with a grid and two colored shapes (a cyan square and a blue rectangle) with handles for resizing. Buttons for 'delete', 'clear all', and 'search'.

1. Use your mouse to choose a colour from the palette.
2. Select either the round tool or the square tool.
3. Hold down your mouse button and drag the cross on the canvas to create a coloured shape.
4. Repeat this process until you complete your custom layout. When you're ready, click Search.

To perform other actions, click the shape to make it active. Drag the edges to Resize. Click Send to Back and Bring to Front to layer shapes. Click Delete to remove a shape. Click Clear All to empty the layout.

# 2000-2010

## Le web et ses millions d'images

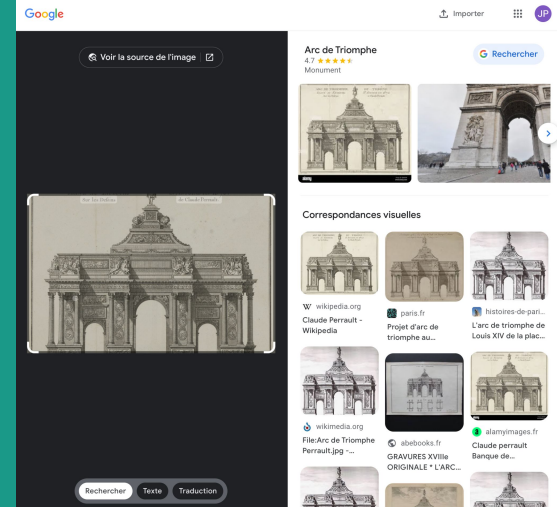
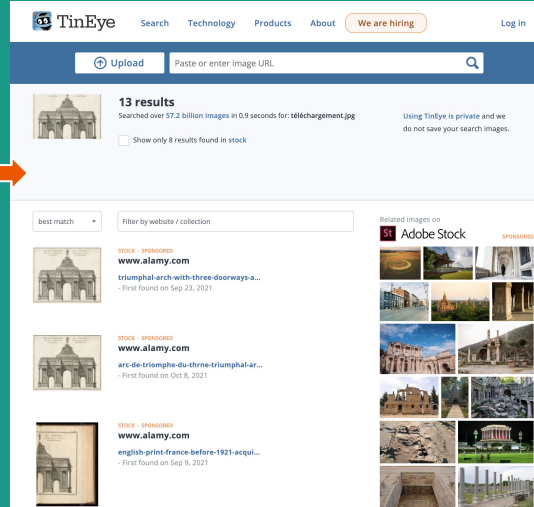
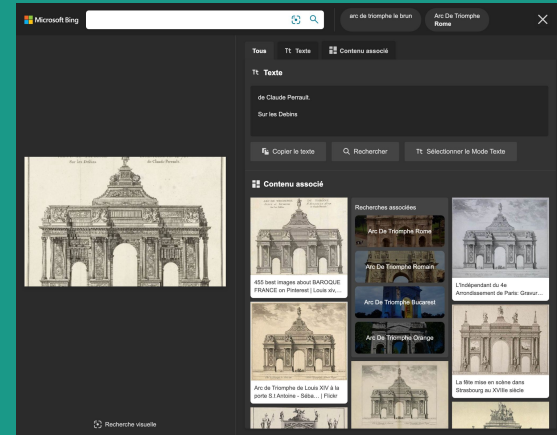
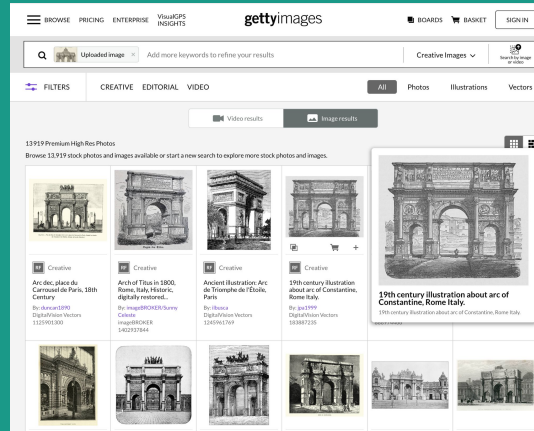
- Picsearch (2000)
- Google Product Search (2008)
- Microsoft Bing (MSN Search)

The screenshot shows the Picsearch homepage. At the top left is the Picsearch logo. To its right is a search bar with a "Search!" button and links for "Images Web", "Preferences", and "Advanced Search". Below the search bar is the tagline "Innovating image search" and a sub-tagline "Search the web for images. An image search service with more than 3,000,000,000 pictures." A section titled "Most popular image searches of the week" features four image thumbnails: Edgar Allan Poe, Gas Dispute, Somalia, and Cristiano Ronaldo. Below these thumbnails is a link that says "Click on an image to view the search results". At the bottom of the page are links for "Business Solutions", "About Picsearch", "Terms of Use", "Search Help", and "Image Directory", along with a copyright notice for ©2009 Picsearch.

This screenshot displays a product search interface for shoes. At the top, it says "Visually Similar Items". Below this are two rows of shoe images. The first row shows three pairs of shoes: Rocket 1 (\$71), New Balance (\$55), and New Balance (\$103). The second row shows three more pairs: Salomon (\$35), New Balance (\$90), and New Balance (\$95). To the right of the shoe images is a sidebar with various filters: "SIZE", "LAYOUT", "COLOR" (with options for "All", "Full color", "Black & white", and "Specific color" with a color palette), "STYLE", and "PEOPLE". The main area to the right of the filters shows a grid of image thumbnails, including landscapes and portraits.

2012-

# Deep learning, web (et ses milliards...)



TinEye  
Google Image, Goolge Lens  
Yahoo Image Search  
Bing Image Search  
Getty Images

# Typologie de techniques

>Segmentation

Classification

Modèles texte-image

Recherche par similarité

International Journal of Otolaryngology

**Figure 100%**

**Table 100%**

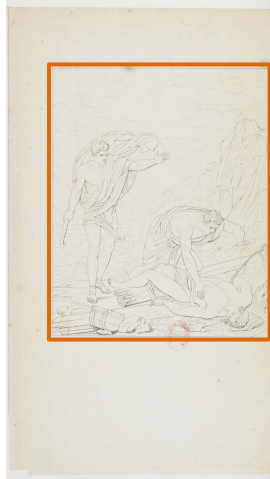
Item	Rate (%)
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14
Both allergic rhinitis and asthma	118.14

**Text 100%**

...

LayoutML v3 (Microsoft)

# Segmentation Imprimés



Annales du Musée et de l'école  
moderne des beaux-arts (1808).

“Où sont les images ?”



Vogue éd. française



Au Bon Marché, maison A. Boucicaut, Paris, blanc, 1917



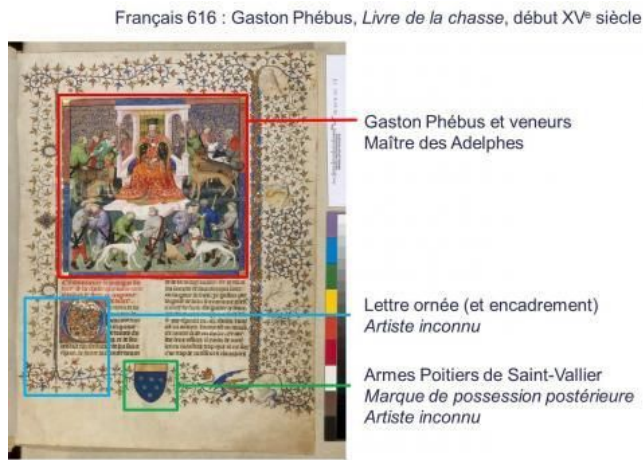
# Segmentation

## Photographies, manuscrits... et dossiers d'œuvre



Source gallica.fr / Bibliothèque nationale de France

Gallica, recueil de photographies



Gaston Phébus et veneurs  
Maître des Adelpes

Lettre ornée (et encadrement)  
Artiste inconnu

Armes Poitiers de Saint-Vallier  
Marque de possession postérieure  
Artiste inconnu

F. 13

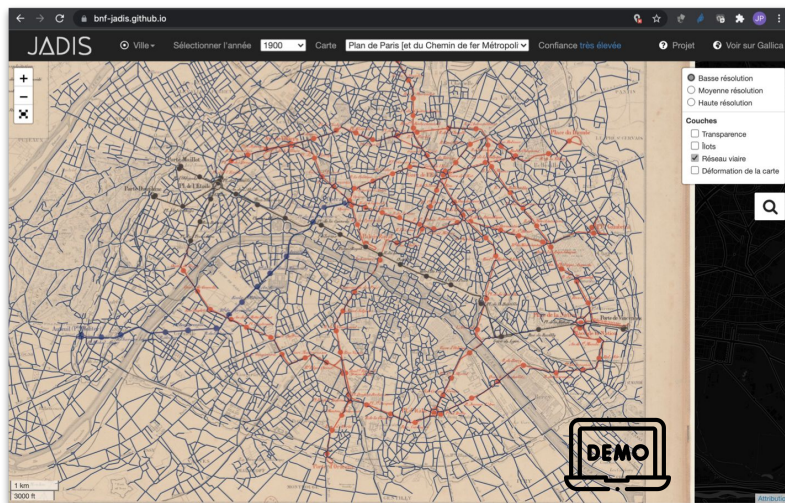
Mandragore, manuscrit enluminé



Quand l'archive incarne l'institution : le rôle des dossiers d'œuvre dans la fabrique documentaire de la muséologie d'un musée des Beaux-arts, Marie Després-Lonnet et Maryse Rizza

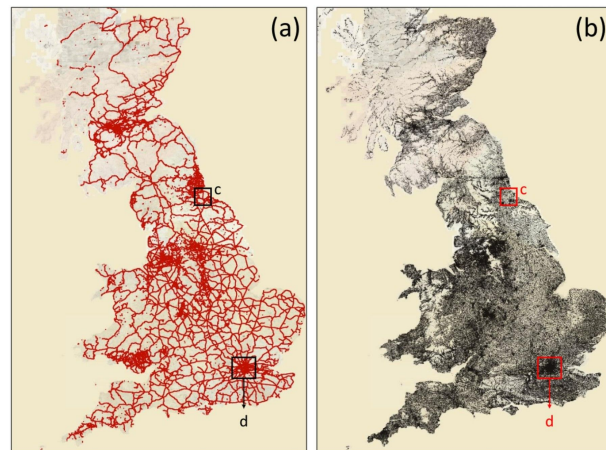
# Segmentation “sémantique”

## Cartes et plans



Projet JADIS (2019-2020)  
<https://bnf-jadis.github.io>

EPFL (BnF)



Projet Map Reader (2019-2023)  
<https://living-with-machines.github.io/MapReader/>

The Alan Turing  
Institute

BRITISH LIBRARY

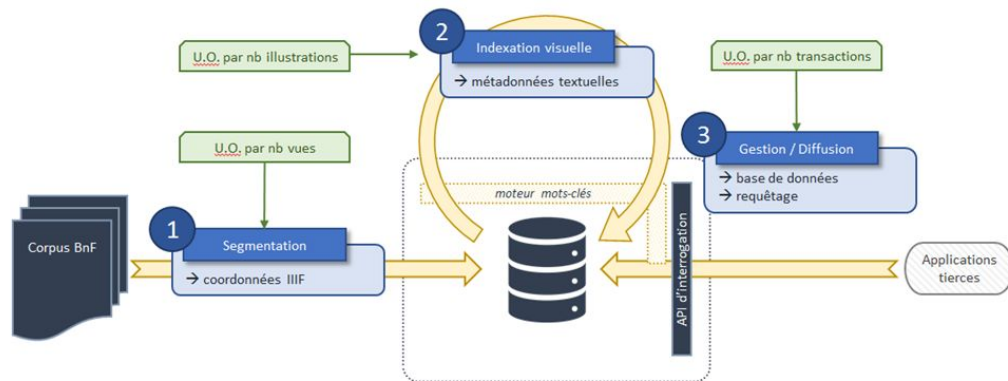
# Projet Gallica Images

## Segmenter, indexer

- Identifier les illustrations de la bibliothèque numérique
- Créer une indexation descriptive des illustrations
- Créer, administrer et mettre à disposition du public et des partenaires la base de données des illustrations
- 2023-2026
- IIIF de bout en bout



gallicapix.bnf.fr





# Outils

Yolo

Detectron2 (Meta AI)

LayoutLMv3

BertGrid

LayoutParser

docExtractor (LIGM)

dhSegment (EPFL)

Cordeep (Max Planck Institute)

Transkribus

CorDeep Extract Cluster

MAX PLANCK INSTITUTE FOR THE HISTORY OF SCIENCE

BIFOLD

téléchargement.png



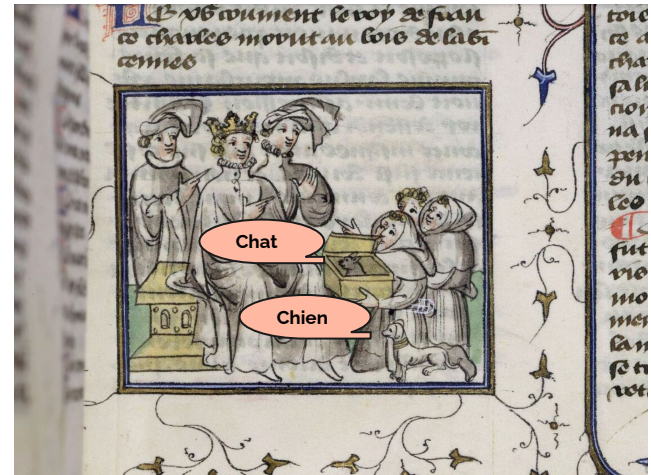
# Typologie de techniques

Segmentation

>Classification

Modèles texte-image

Indexation par similarité

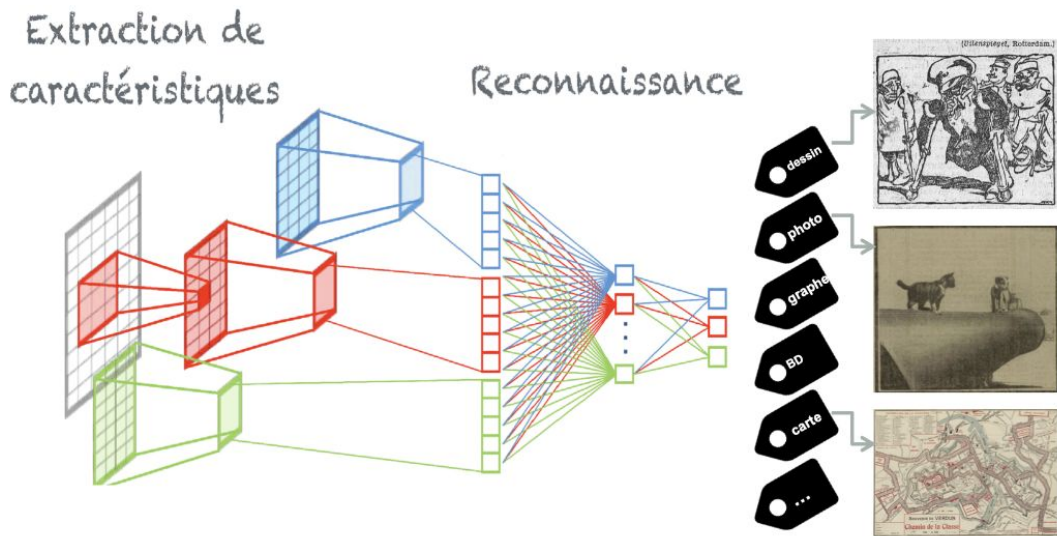


Charles IV le Bel et l'affaire du chat noir

# Classification d'image

Associer une “classe” à une image, à partir d'un modèle entraîné sur un jeu d'images classées (apprentissage par l'exemple, donc potentiellement laborieux...)

“Quelle est la nature de cette image (10 techniques/fonctions) ?”



Réseau de neurones artificiels entraîné, [gallicapix.bnf.fr](http://gallicapix.bnf.fr) (2017)

# GallicaCIP

## Classification d'images du patrimoine

INRIA équipe LinkMedia,  
 projet INRIA-MIC (2019)  
 Mandragore, corpus Zoologie :  
 24k images, 42k annotations,  
397 espèces

Défis : pas de localisation, classes  
 déséquilibrées, grande diversité  
 visuelle intra-espèce



Mandragore



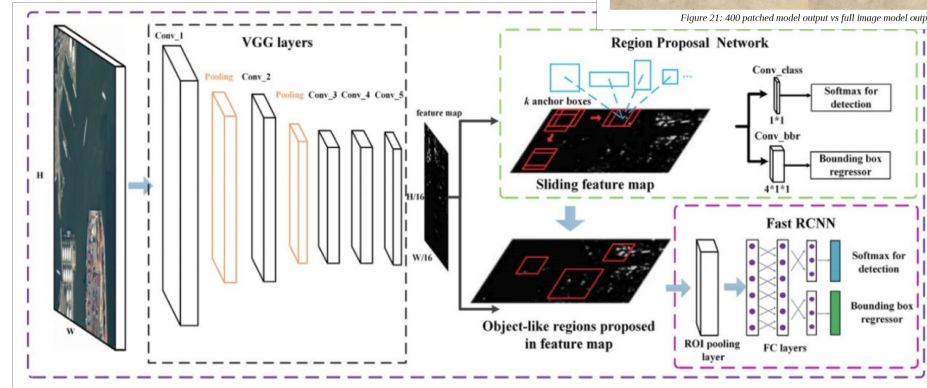
# GallicaCIP

## Classification d'images du patrimoine

- Réduction de complexité (397 espèces  $\Rightarrow$  30 genres)
- Supervision forte (Faster R-CNN)
- Augmentation, annotation manuelle des images (8k boîtes, 1,8k images, 100 images par classe a minima)
- Apprentissage par transfert sur modèle préentraîné (iNaturalist)
- Fenêtre glissante



Figure 21: 400 patched model output vs full image model output





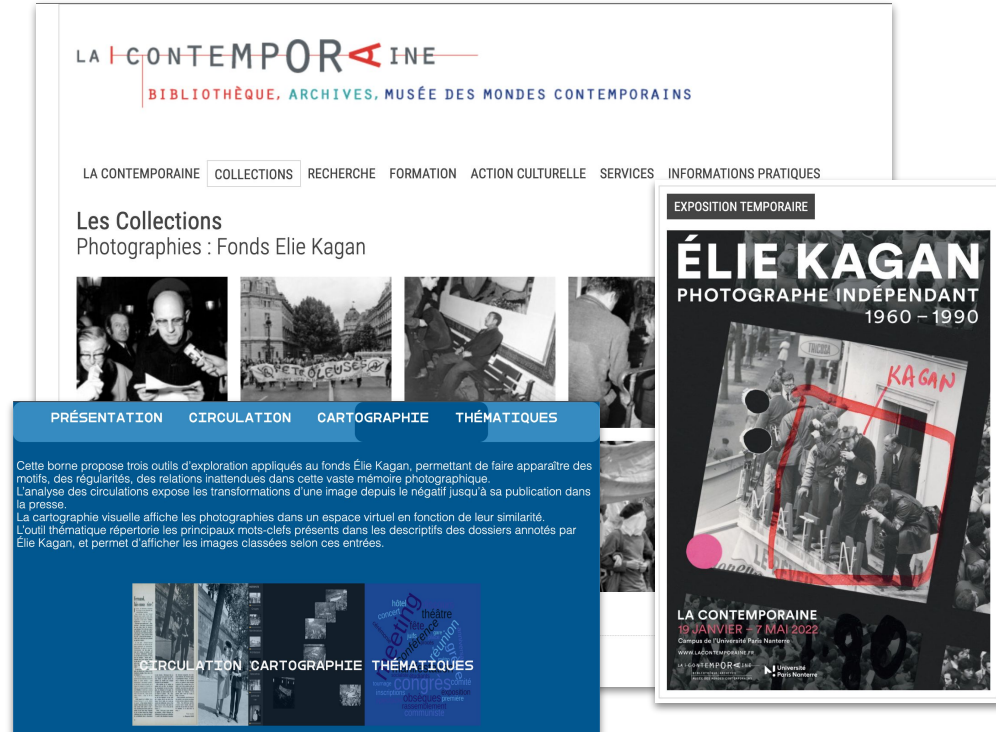
# La Contemporaine Muséographie

Projet MODOAP (2021, Labex “Les Passés dans le présent”, 2021) : fonds Elie Kagan (200 000), La Contemporaine

Réalisation d'une borne numérique d'exploration au sein de l'exposition E. Kagan :

- exploration thématique (classification d'images par modèle entraîné)
- nuage d'images similaires
- repérage des circulations dans la presse

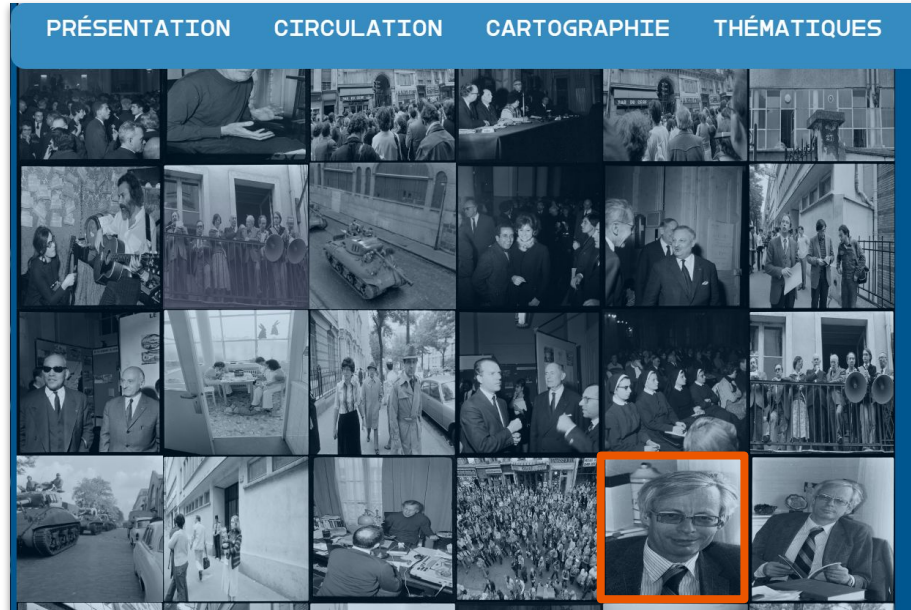
[https://modoap.huma-num.fr/Kagan\\_Contemporaine/index.php](https://modoap.huma-num.fr/Kagan_Contemporaine/index.php)



# “MODOAP”

## Muséographie

Manifestation



# “Exploring Ed Ruscha’s archive” Catalogage automatique

Archives Ed Ruscha, Getty (2020)  
Exposition virtuelle de 65 000  
photographies (sur 1 M) numérisée  
et indexées

Google Cloud Vision : classification  
et OCR ; qualité “brute” assumée

<https://12sunsets.getty.edu/>

The screenshot displays the '12 SUNSETS' website interface. At the top, it says 'EXPLORING ED RUSCHA'S ARCHIVE' with navigation links for 'DRIVE', 'SORT', 'ABOUT', and 'SHARE'. A search bar contains the text 'palm tree'. Below the search bar, a grid of image thumbnails is shown, with several images featuring palm trees. A red speech bubble with the word 'Palmier' points to one of the palm tree images. On the right side, a vertical list of search results is visible, including 'palm tree', 'parking lot', 'passenger car', and 'photograph'. The bottom of the page includes a footer with copyright information: '© J. Paul Getty Trust | Privacy Policy | Terms of Use | Trademarks | All images © Ed Ruscha. Used with permission.'





# Musée des arts décoratifs

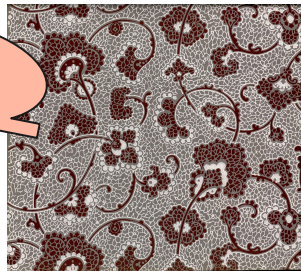
## Catalogage automatique

### Cataloguer en numérisant ?

1 M de documents, dont 400 000 échantillons de papiers-peints

- génération de “pré-notices”
- classification de types
- détection de motifs/objets

motif floral

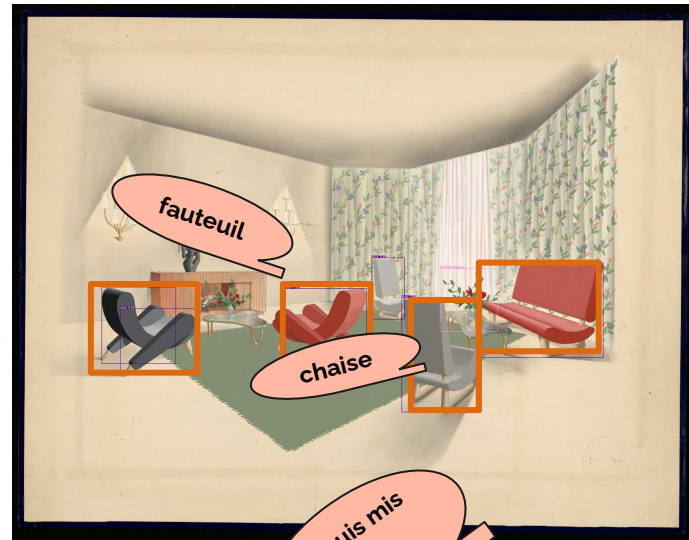


palette de couleurs



Fonds Leroy

### Expérimentation Yolo v5, Google Vision



Jean Royère

# Outils

---

Yolo

Faster-RCNN

Amazon, Google...

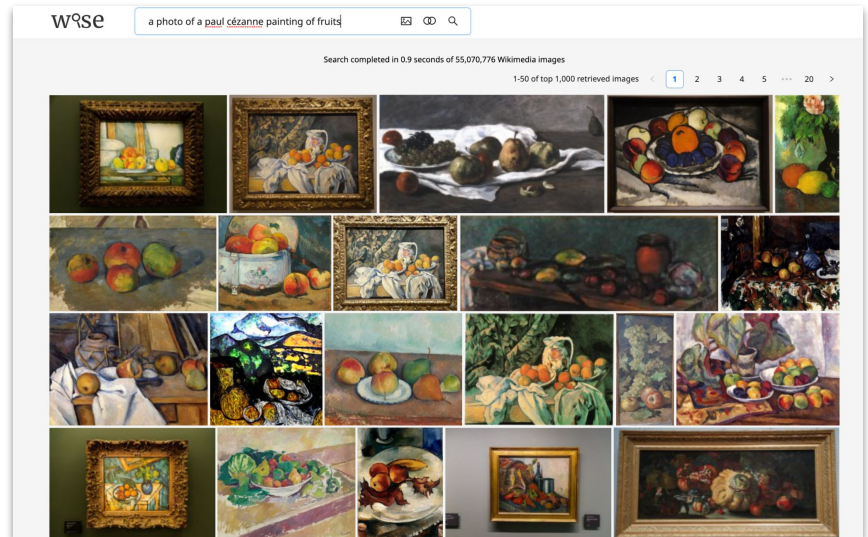
# Typologie de techniques

Segmentation

Classification

> **Modèles texte-image**

Indexation par similarité

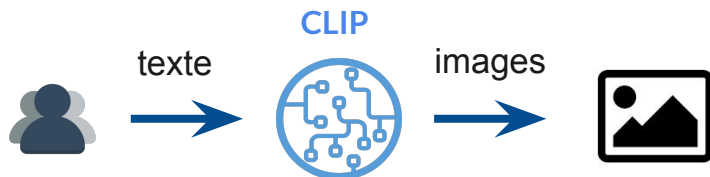


# CLIP

## Modèle texte-image

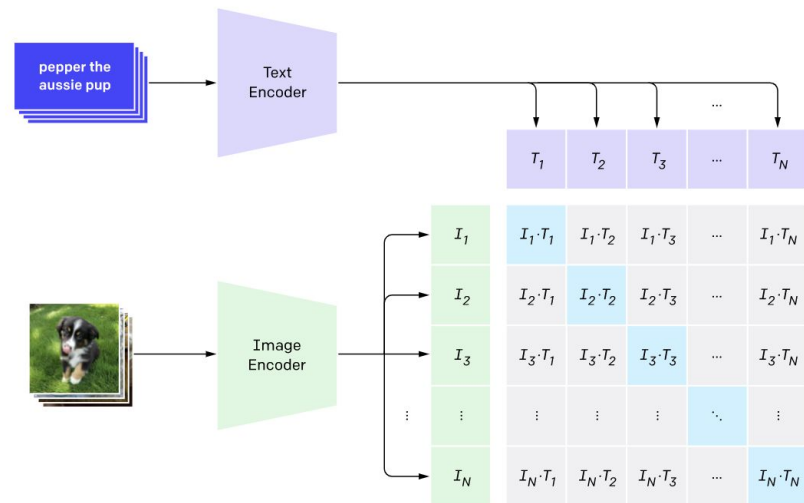
Modèle CLIP (OpenAI, 2021) : apprentissage d'un espace de représentation commun à partir de 400 M d'images du web et leurs légendes

- le modèle peut être opéré avec des requêtes en langage naturel
- large "connaissance visuelle" (-> pas d'entraînement local)



- les textes sémantiquement identiques sont proches dans cet espace
- les images visuellement similaires sont proches dans cet espace
- idem pour les textes/images

### 1. Contrastive pre-training



<https://openai.com/research/clip>

# CLIP : exemple

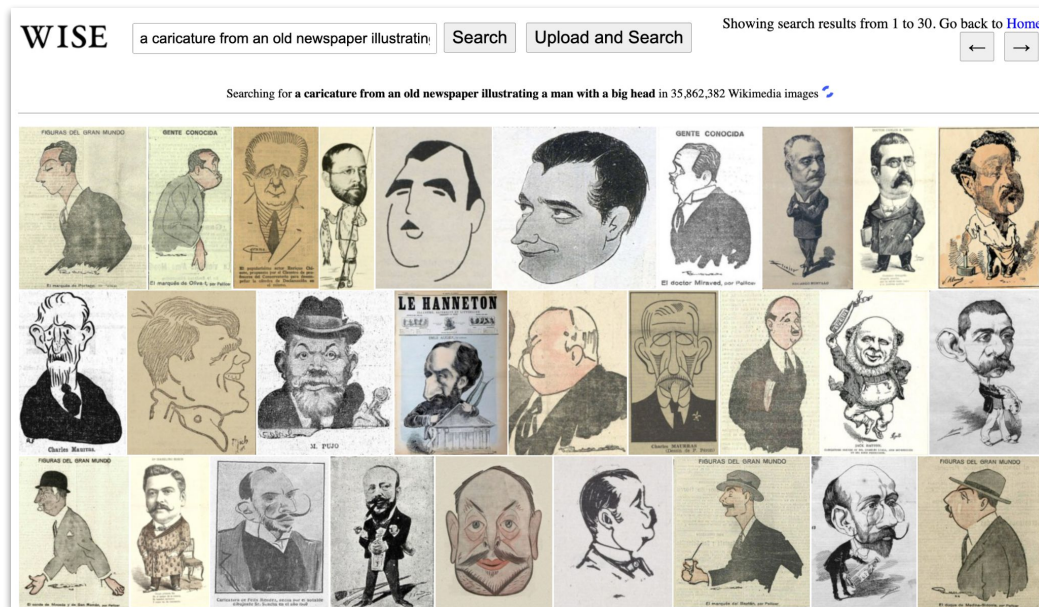
## WISE

OpenCLIP sur 50 M  
d'images de Wikimedia

<https://www.robots.ox.ac.uk/~vgg/>

Visual Geometry Group  
(Giles Bergel et al., université  
d'Oxford, 2023)

“a caricature illustrating a  
man with a big head from an  
old newspaper”



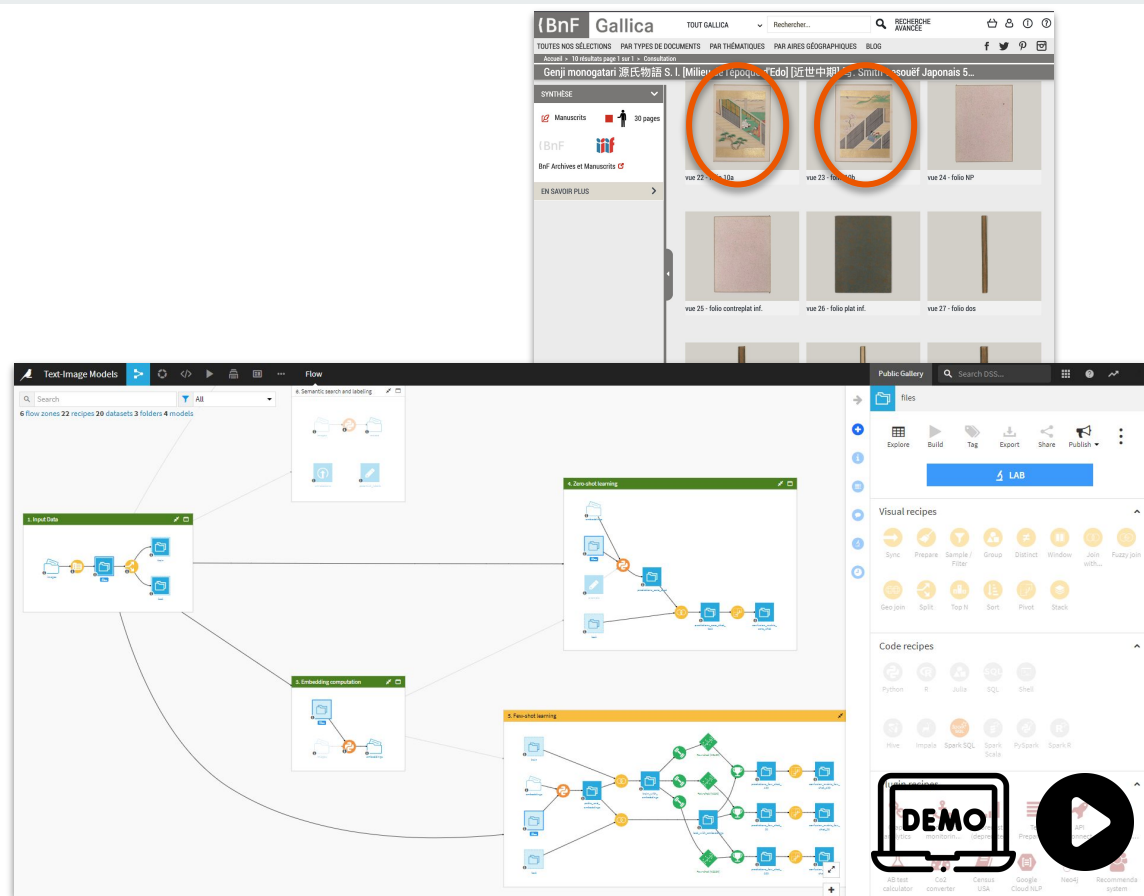
<https://meru.robots.ox.ac.uk/wikimedia/>

# CLIP : exemple

## Classification de documents

Filterer les feuillets illustrés :

- “estampe japonaise”
- “idéogrammes japonais”
- “reliure de livre”
- “page blanche”



The image displays a Dataiku project interface. The top part shows a 'Public Gallery' view of search results for 'Genji monogatari' (源氏物語). Two image thumbnails are circled in orange, representing the 'estampe japonaise' and 'idéogrammes japonais' categories. The bottom part shows a 'Flow' view of the CLIP workflow, which includes steps for 'Image Data', 'Embedding generation', and 'Classification'. A 'DEMO' button and a play icon are visible in the bottom right of the gallery.



# CLIP : exemple

Classification de photos  
pour des usages internes

Indexation d'une banque d'images  
(ateliers de restauration internes,  
50 000 images) avec CLIP :

- recherche de sujets d'intérêt  
(reliure, nerf, cuir...)
- recherche par texte libre



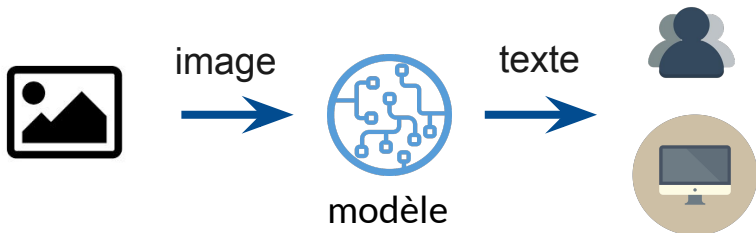
[https://github.com/altomotor/CLIP\\_test](https://github.com/altomotor/CLIP_test)

# CLIP, CoCa, Pix2Struct...

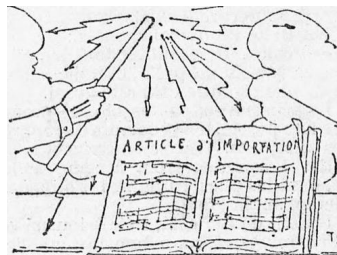
## Génération de légendes

Catalogage automatique

Accessibilité aux contenus  
iconographiques



A **map** of the region of France with the cities of Arrivee and Montfort.



A **drawing** of a person's face with the words Article and Importation written on it.

[PIX2STRUCT: SCREENSHOT PARSING AS PRETRAINING FOR VISUAL LANGUAGE UNDERSTANDING](#)

Kenton Leet al. (2022)



# Outils

---

CLIP

CoCa

Pix2Struct

# Typologie de techniques

Segmentation

Classification

Modèles texte-image

> **Indexation par similarité**

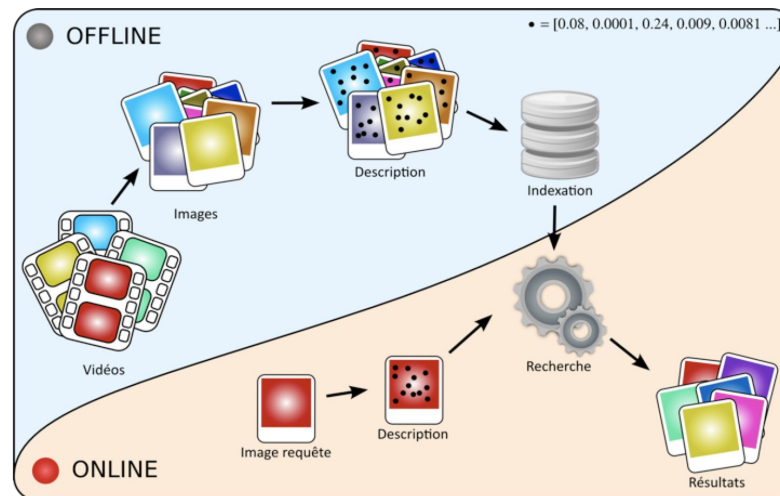


# Snoop

## Moteur de similarité visuelle

### Moteur Snoop (INA, Inria/équipe Zenith)

- Utilisé par l'application de science participative Pl@ntNet
- INA : aide à l'indexation de contenus par les documentalistes
- Projets R&D (média, sciences de l'information)



# Gallica Snoop

## Recherche avec interaction utilisateur

- **Projet R&D Gallica Snoop, 2019-2020** (convention cadre INRIA-MiC) : 1,2 M images, *“humain dans la boucle”*
- Expérimentations internes, projets hébergés par le BnF Datalab
- Modèles CNN ou SIFT

<https://snoop.inria.fr/bnf/>

The screenshot shows the Gallica Snoop search interface. At the top, a reference image of a classical building is shown with an orange arrow pointing to a search bar. The search results are displayed in a grid of 20 images, each with a green checkmark or a red X indicating user interaction. A 'DEMO' icon is visible in the top right corner.

# La Contemporaine

Recherche de  
reproductions

Recherche de  
reproductions  
de photo d'Elie Kagan  
dans la presse  
contemporaine avec  
le moteur Snoop / SIFT  
(détails, recadrage)



*Droit et Liberté*, mai 1961



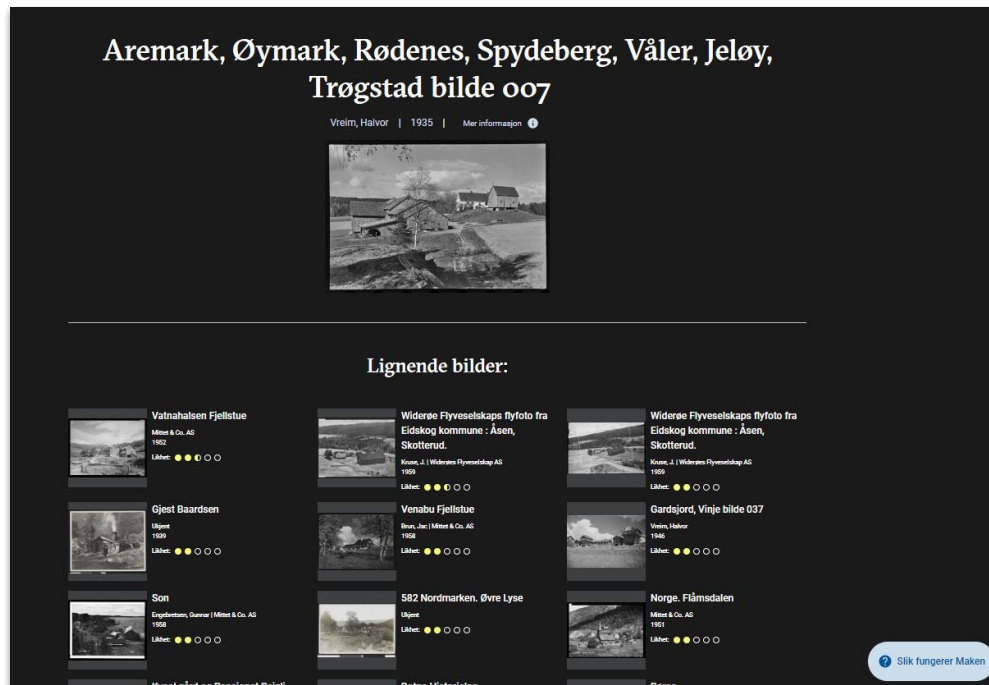
# Maken

## Moteur de similarité multimodale

Bibliothèque nationale de Norvège

Moteur de recommandation :

- pour les imprimés (500k notices)
- pour les images (500k, *embeddings* CLIP)



<https://www.nb.no/maken/>



# BnF : mySnoopy

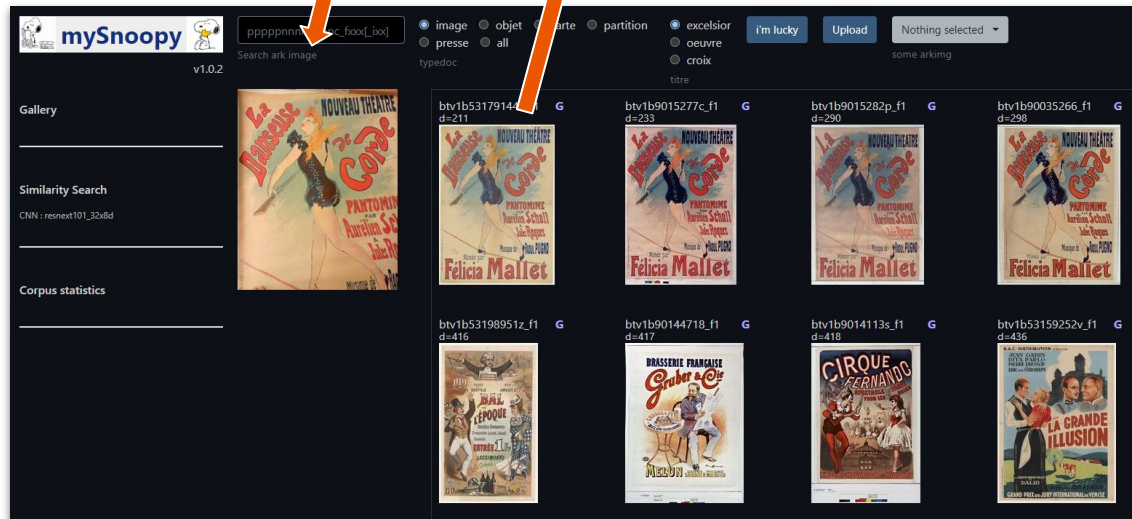
## Aide au catalogage

Aide au catalogage avec un démonstrateur interne (réseau CNN, Milvus, Towhee, Python)

(Dpt des Arts du spectacle, BnF, 2023)

1. Photo prise en magasin (smartphone)

2. Déjà numérisé dans Gallica ?



# Pixplot

## Visualisation de grandes collections

La Contemporaine :  
visualisation du fonds  
Elie Kagan par nuage  
d'images agrégées par  
similarité avec PixPlot  
(DHLab, Yale)



[https://modoap.huma-num.fr/pixplot\\_kagan/](https://modoap.huma-num.fr/pixplot_kagan/)

Visualizing Image Fields

modoap.huma-num.fr/pixplot\_kagan/index.html#

DH Lab

Retour à la borne

Représentation des photographies d'Elie Kagan dans un espace de similarité

TAILLE DES PHOTOS

ALIGNER LES PHOTOS

GROUPES

Groupes

Manifestations

Architecture

Rencontres

Scènes de rue

Enfants

A table

Portraits

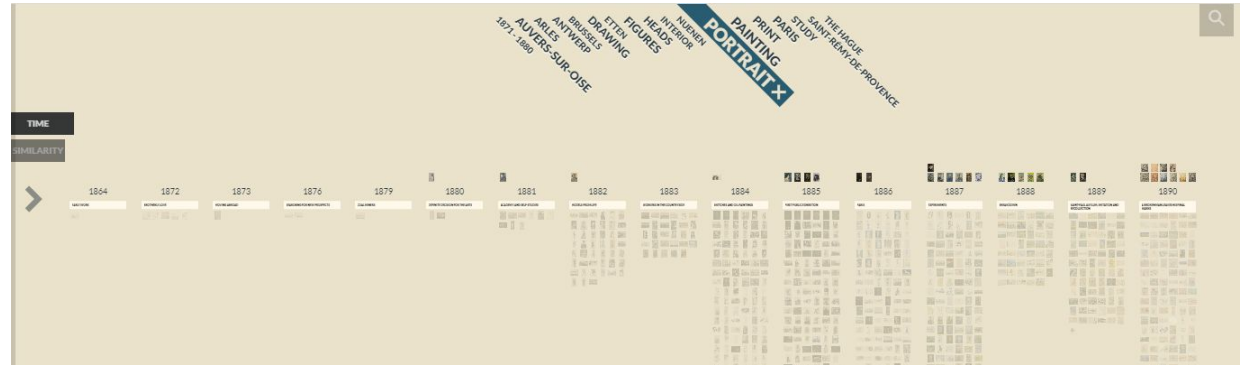
manifestations



# Vikus

## Visualisation de collections IIF

Visualiseur de collections IIF par date, mot-clé, proximité visuelle (modèle CLIP)



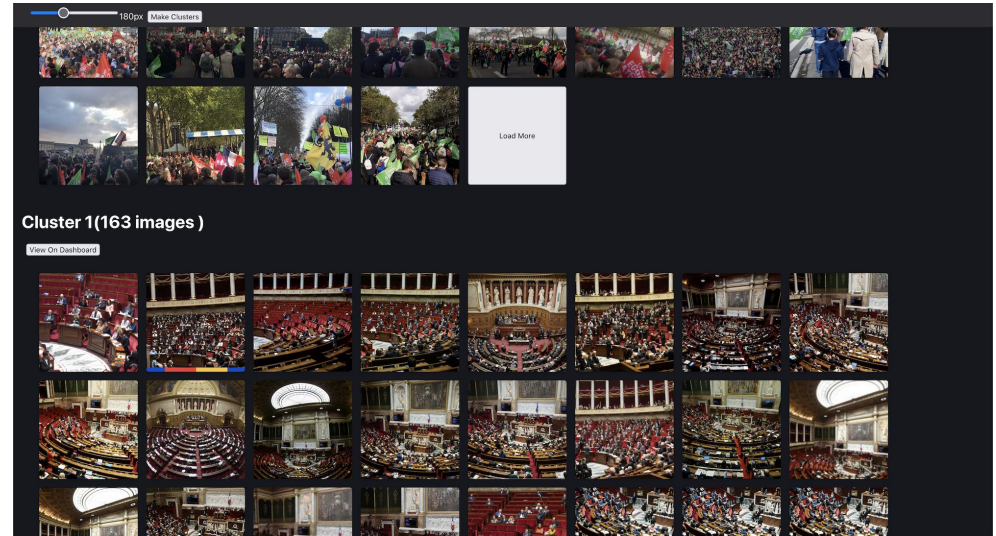
<https://vikusviewer.fh-potsdam.de/>

# Panoptic

## Annotation et exploration de corpus massifs

CERES, Sorbonne Université

Logiciel (en développement)  
d'annotation et d'exploration  
de gros corpus d'images SHS



# Outils

---

**CLIP, OpenCLIP**

**SNOOP (Ina-Inria)**

**Vikus (univ. Postdam)**

**Pixplot (univ. Yale)**

et aussi :

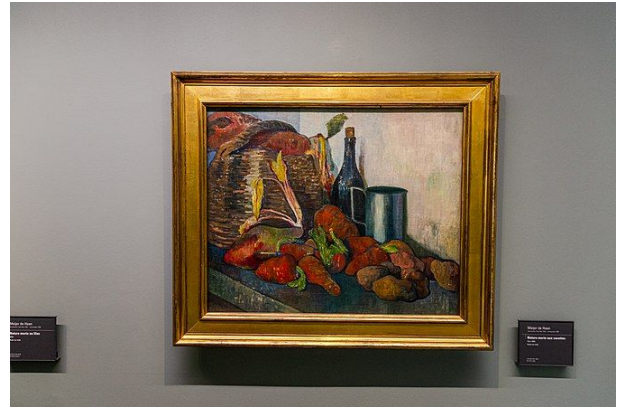
**Pimmi (Ina)**

**Collection Space Navigator (univ. Tallinn)**

**Panoptic (CERES)**

---

# IA et histoire de l'art



# imgs.ai

## Moteur de similarité

Université de Marbourg

Collections MET et Risjkmuseum :  
modèles CLIP (texte-image) ou CNN (image)

- Qu'est-ce que la similarité ?
- ... en particulier pour l'histoire de l'art (patrimoine, ESR)
- Prise en compte des niveaux d'interprétation requis ? (E. Panofsky)



“a triumphal arch with three doors”

imgs.ai is a fast, dataset-agnostic, deep visual search engine for digital art history based on neural network embeddings. imgs.ai utilizes modern approximate k-NN algorithms via Spotify's Annoy library to deliver fast search results even for very large datasets in low-resource environments, and integrates the OpenAI CLIP model for text-based visual search. Try it below on the complete Rijksmuseum and Metropolitan Museum of Art collections (CC0) or sign up for an account to access more functions/datasets (institutional email address and approval required.) imgs.ai is developed by Fabian Ofert, with contributions by Peter Bell and Oleg Harlamov. Get in touch at hi@imgs.ai.

Searching 392630 images in Rijksmuseum

Embedding clip\_vit  angular  128  30  a triumphal arch with thi

<https://zentralwerkstatt.org/blog/ten-years-of-image-synthesis>

# Indexation IA et vocabulaires contrôlés ? imgs.ai

- à intégrer ?
- complémentaires ?
- irréconciliables ?

Subject

What

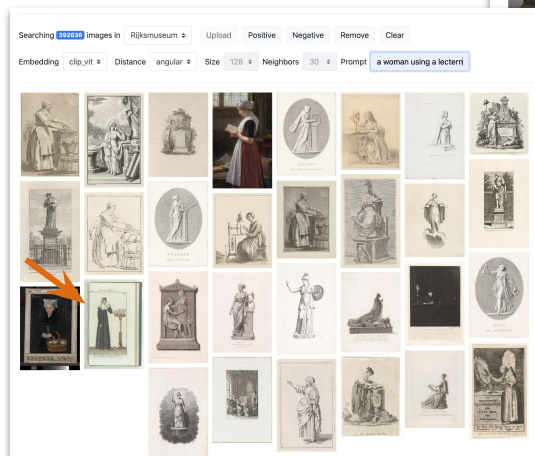
- fashion plates
- head\_gear\_touque (+ women's clothes)
- head\_gear\_hat (+ women's clothes)
- dress\_@own\_redingote (+ women's clothes)
- handkerchief (+ women's clothes)
- accessories - music.metronome (+ music.tenor\_tuning-fork\_etc.

When

- 1808-12-18 - 1808-12-18

femme,  
pupitre

“une femme utilisant un pupitre”



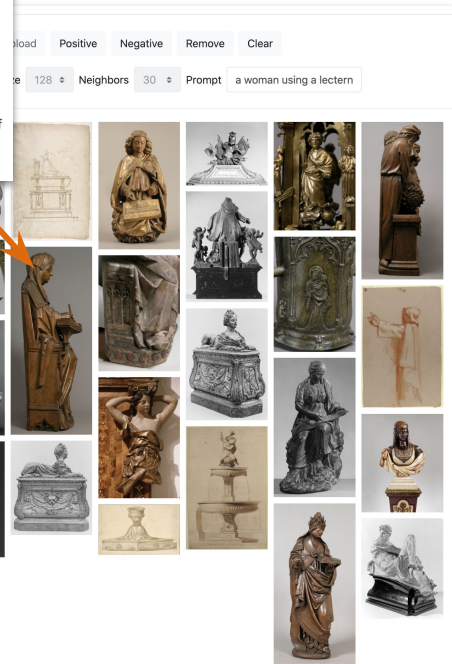
## Saint Bridget of Sweden

ca. 1470

Master of Soesterbeeck

On view at The Met Fifth Avenue in Gallery 305

Saint Bridget of Sweden (ca. 1303–1373) is shown in the act of writing her “Revelations,” which, according to legend, were dictated to her by Christ himself. Shown in the monastic habit of the most prominent Saviour of the Brigittines, she was venerated in northern Europe.





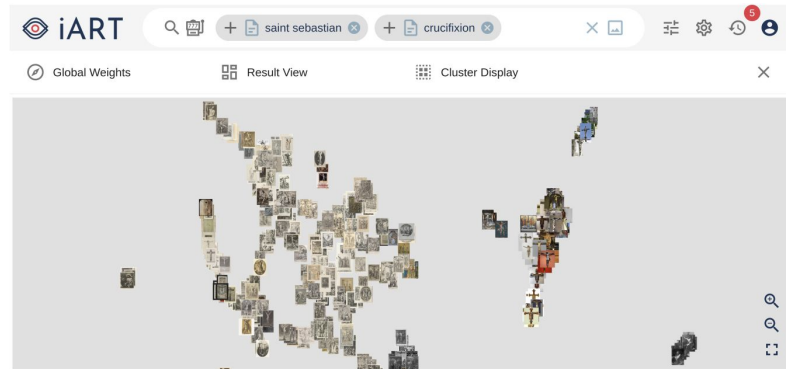
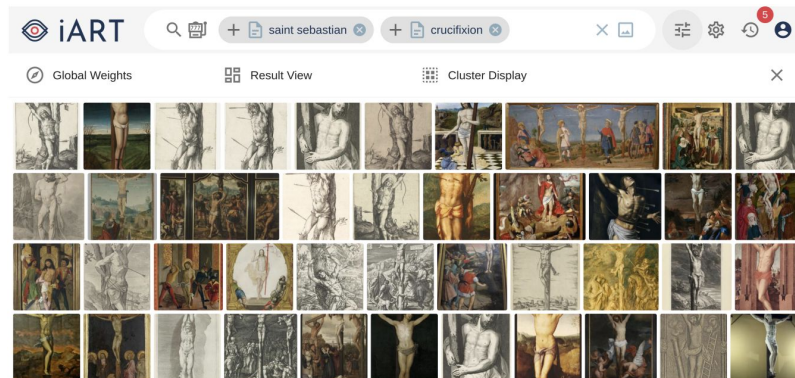
# iART

## Moteur pour la recherche en histoire de l'art

Universités Leibnitz, Paderborn, Munich

Plateforme paramétrable pour l'étude comparative de corpus d'images :

- plusieurs modèles d'indexation image
- métadonnées (notices)
- navigation multimodale



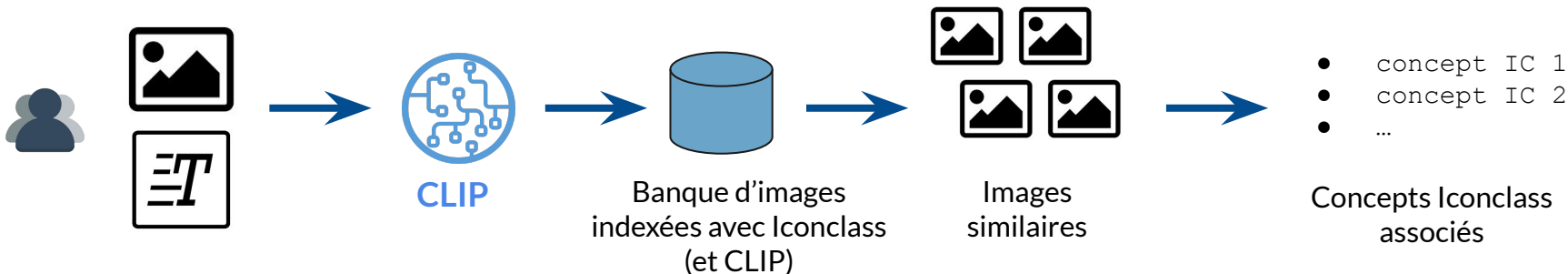
<https://www.iart.vision/>

<https://arxiv.org/pdf/2108.01542.pdf>

# Indexation IA et vocabulaires contrôlés ? Le cas d'Iconclass

Catalogage automatique : entraîner un modèle d'indexation Iconclass (à partir d'un corpus images/indexations Iconclass)... *difficile*

Aide à l'utilisation d'Iconclass : *aisé*



GT: church , water course, city-view, and landscape with man-made constructions, street, clouds, rowing-boat, canoe, sailing-ship, sailing-boat, windmill, cow.  
Caption: ships .

BLEU 1:  $1.12 \times 10^{-22}$  ROUGE: 0.0  
BLEU 2:  $3.55 \times 10^{-18}$  CIDEr: 0.0  
BLEU 3:  $1.12 \times 10^{-16}$  CLIP-S: 0.6218  
BLEU 4:  $4.51 \times 10^{-16}$  ReCLIP-S: 0.5998  
METEOR: 0.0



GT: plants and herbs: marjoram, hill, potted plants, container of ceramics: jar, jug, pot, vase, pig, scholastic education, tuition, Contrariety; 'Contrarietâ', Vulgarity, proverbs, sayings  
Caption: 'hoofed animals : boar, container of metal : bucket, can, canister, drum, tin, proverbs, sayings.

BLEU 1: 0.1425 METEOR: 0.079  
BLEU 2: 0.1049 ROUGE: 0.211  
BLEU 3:  $7.7407 \times 10^{-7}$  CIDEr: 0.057  
BLEU 4:  $2.153 \times 10^{-9}$  CLIP-S: 0.884  
METEOR: 0.079 ReCLIP-S: 0.817



GT: apostle, unspecified, key.  
Caption: head turned to the right, historical persons.

BLEU 1:  $3.43 \times 10^{-16}$  ROUGE: 0.0  
BLEU 2:  $1.54 \times 10^{-16}$  CIDEr: 0.0  
BLEU 3:  $3.68 \times 10^{-16}$  CLIP-S: 0.5221  
BLEU 4:  $1.85 \times 10^{-16}$  ReCLIP-S: 0.6427  
METEOR: 0.0

# Projets

---

**imgs.ai (univ. Marbourg)**

**iART (univ. Leibnitz)**

**EyeCON (univ. Nanterre)**

“Transfer Learning for the Visual Arts: The Multi-Modal Retrieval of Iconclass Codes”,  
Nikolay Banar, Walter Daelemans, Mike Kestemont

“Towards Generating and Evaluating Iconographic Image Captions of Artworks”,  
Eva Cetinic

“Multimodal Search on Iconclass using Vision-Language Pre-Trained Models”, Cristian Santini et al.

---

# Perspectives



“A painting by Edvard Munch depicting a librarian fearing to be replaced by an artificial intelligence”  
(DALL.E 2)

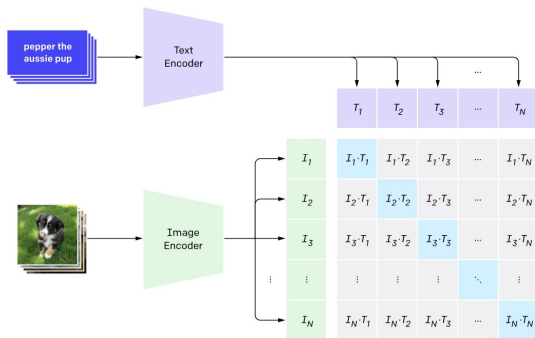
# L'âge d'or ?

## Modèles prêts à l'emploi

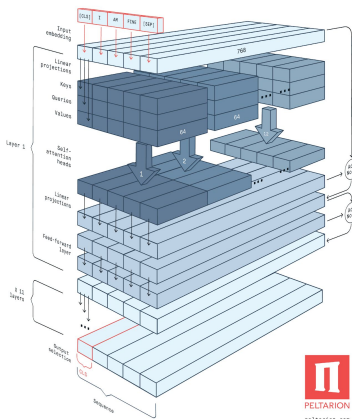
Ces modèles consomment les contenus du web...

**CLIP (2021)**  
450 M de paires légende/image

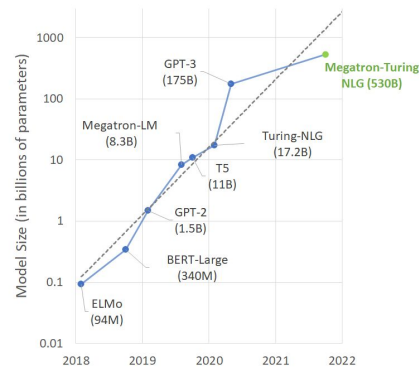
### 1. Contrastive pre-training



**BERT (2020)**  
110 M de paramètres



**GPT-3 (2021)**  
175 milliards de p.



**GPT-4 (2023)**  
1 trillion ?

Modèle texte-image

# Corpus/modèle

## Général/spécialisé

Pissenlit/pâquerette ?  
CLIP : 97% (précision)

Export Created  
4 months ago  
September 03, 2020

Export Size  
1821 images

Annotations  
flowers

### Preview

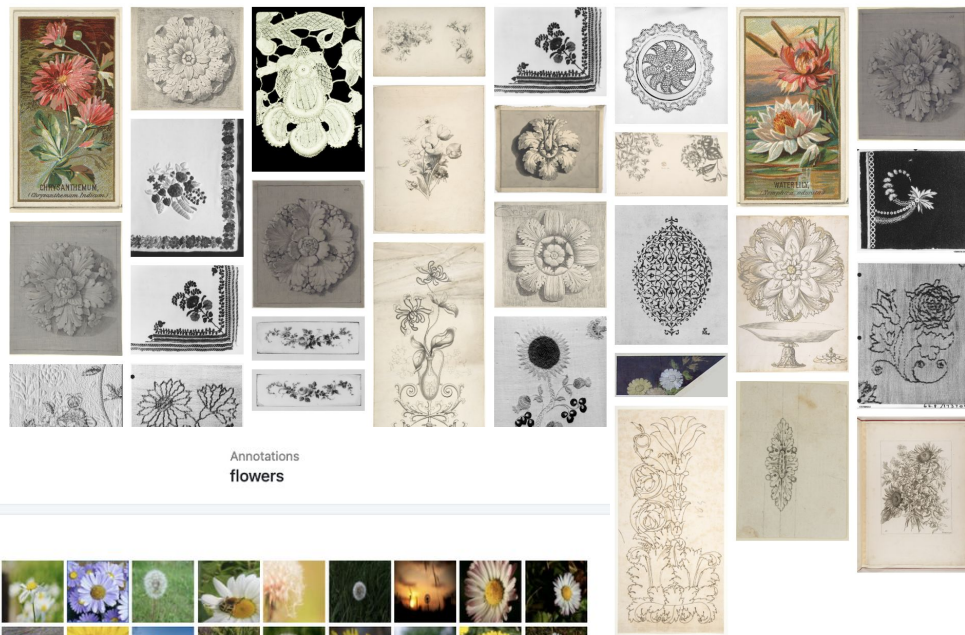


“un dessin d’une pâquerette”

<https://imgs.ai>

Searching **358426** images in Metropolitan

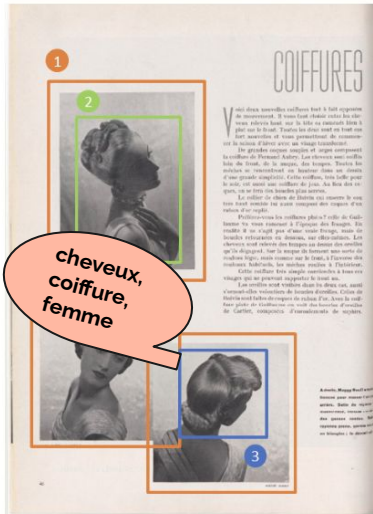
Embedding  Distance  Size  Neighbors





# Vocabulaire, modèle

## Contrôlé/libre, image/texte-image



Détection d'objets/concepts  
 ⇒ "sujets" ⇒ "catalogue" ⇒  
 indexation plein texte

Vogue (1920)

### "coiffure de femme"



Indexation  
 CLIP ⇒  
 interrogation  
 en "langage  
 naturel"

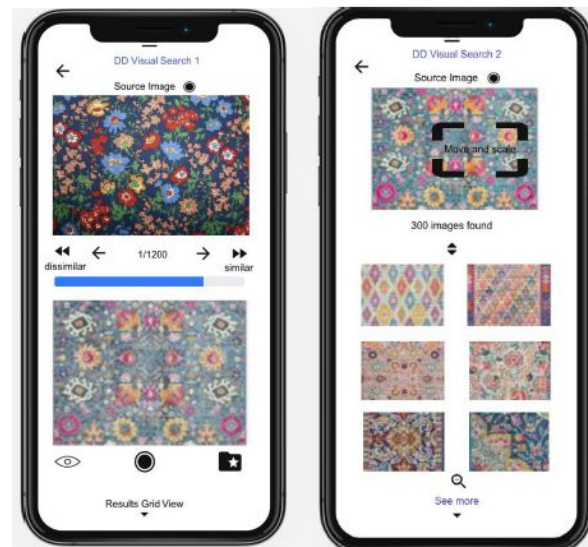
L'Oeuvre,  
 L'Excelsior (1930)

# Le souci de l'utilisateur

## Utilisabilité ?

Projet du programme “Towards a National Collection Foundation” (TNA, univ. of Surrey, V&A Museum...)

“Application de méthodes de vision par ordinateur (CV) et d'intelligence artificielle explicable (XAI) pour améliorer la capacité du grand public et des chercheurs spécialisés à découvrir des collections visuelles de manière nouvelle et/ou plus efficace”

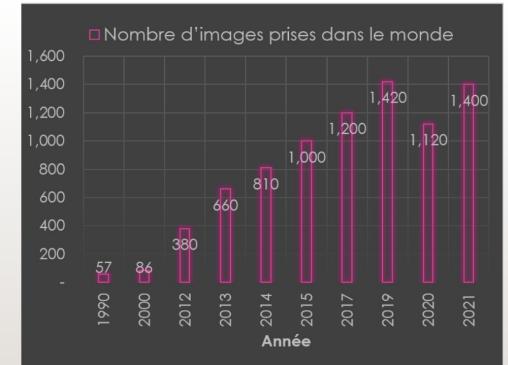
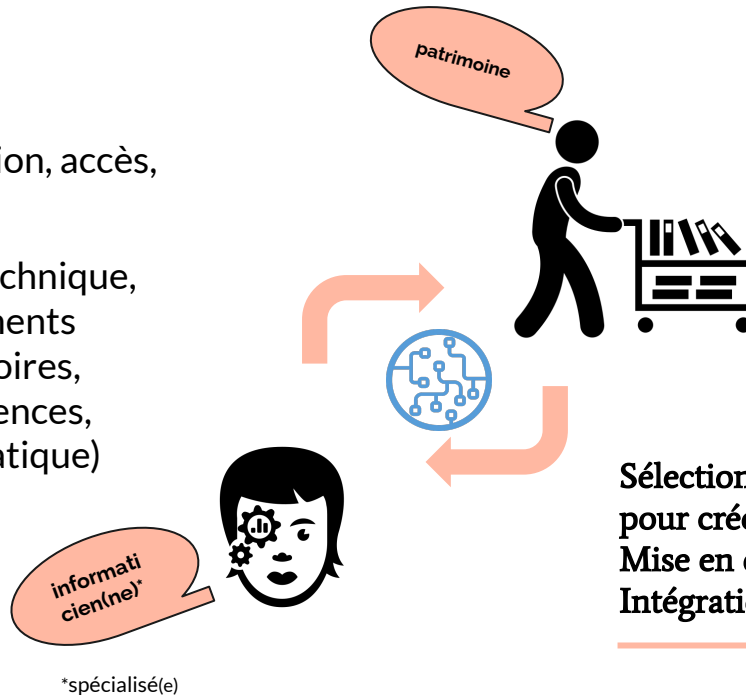


<https://www.nationalcollection.org.uk/sites/default/files/2021-10/Deep%20Discoveries%20Final%20Report%20.pdf>

# Conclusion

**Promesses** : indexation, accès, volumes

**Défis** : complexité technique, rythme des changements techniques, boîtes noires, intégration (compétences, architecture informatique)



Source : Mitran, 2014 ; Garoscio, 2017 ; Pantic, 2021

Shahrazad Rahmé (2022)

**Sélection d'une approche, d'un modèle,  
pour créer/améliorer un service  
Mise en œuvre (technique, design)  
Intégration avec les SI existants**



# Merci !

jean-philippe.moreux@bnf.fr

CENL “AI in libraries” network group :

<https://www.cenl.org/networkgroups/ai-in-libraries-network-group/>

AI4LAM : [ai4lam.org](http://ai4lam.org)