

Institut Supérieur d'Informatique et des Mathématiques de Monastir



TECHNIQUES D'INDEXATION ET RECHERCHE MULTIMÉDIA

IMEN CHEBBI

Introduction à la Recherche d'Information

- 1. Recherche d'Information?*
- 2. Recherche basée sur le contenu*
- 3. Applications de la recherche d'information multimédia*
- 4. Évaluation des techniques de recherche*

1. Recherche d'Information (*RI ou IR: Information Retrieval*) ?

- Consiste à trouver de données (***objets multimédias***) peu ou faiblement structurées, dans une grande collection (Base de données : BD), en fonction d'un ***besoin d'information***



- L'opération de la RI est réalisée (dans une **BD**) par des outils informatiques appelés **Systemes de Recherche d'Information** (SRI).
- Le but principal d'un SRI est de retrouver les **objets** (*multimédias*) **pertinents** en réponse à une **requête** utilisateur.



- Ces **objets** sont typiquement **retournés** sous **forme** d'une **liste ordonnée**, où l'ordre est basé sur des **estimations de la pertinence**.
- C'est à l'**utilisateur** d'**extraire** l'information demandée à partir des **résultats** fournis.



1. Recherche d'Information?

- **Exemples**

- Recherche sur votre **ordinateur**.
- Recherche dans votre **boîte mail**.
- Recherche sur le **Web** (Utilisée par des milliards d'utilisateurs).
- Recherche dans une **base documentaire**, *publique* ou *privée*.



- **Recherche:** choix entre les données objets, basé sur...
 - La **condition** de **SELECT** type "base de données" (correspondance **exacte**)
 - Ou une **liaison de similarité** définie (meilleure correspondance)



- **Exemple** de **requête** impossible avec un SGDB "classique":

Récupérer tous les images "qui ressemblent" à une image requête



- **Recherche basée sur le contenu RBC ou CBR** (Content-based retrieval)
 - Standardisation des métadonnées (par exemple, ***MPEG-7, MPEG-21***)

MPEG-7 spécifie comment décrire le contenu

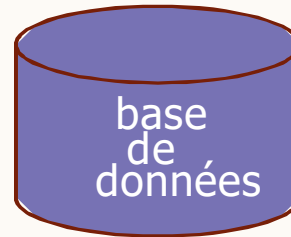
MPEG-21 fournit un cadre multimédia véritablement inter-opérable



Bases de données multimédias (BDMM)?

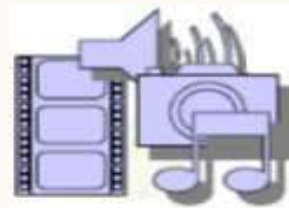
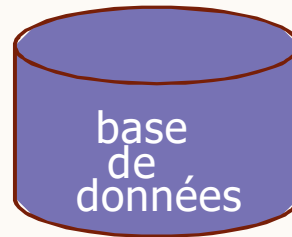
- ✎ **Base de données:** Ensemble structuré et organisé permettant de stocker des grandes quantités d'informations et d'en faciliter l'exploitation (ajout, mise à jour, **recherche d'Information RI**).

base
de
données

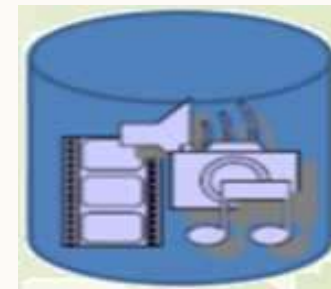


base
de
données

 **Media** : Tout support de diffusion ou d'expression d'informations
à l'attention d'un groupe



Multimédias



BDMM



Multimédias :

-Combine des médias de différents types :

 **Médias statiques** : Textes, images, objets 3D.

 **Médias dynamiques** : Sons, vidéos 2D, vidéos 3D

-Caractérisés par des besoins de **stockage très importants**.

Indexation et recherche multimédia (*Multimedia Retrieval*)

- **Besoins** pour un utilisateur → **Recherche** de documents (**images, audio, vidéo**).
- **Recherche** de documents **complets** ou quelques **passages** (des extraits)



- **Paradigmes** de recherche :

- Le texte environnant → peut être **manquant, inexact** ou **incomplet**.
- Recherche basée sur **le contenu** (à l'aide de **mots-clés** ou de **concepts**).



→ besoin d'indexation basée sur le contenu → « problème d'***écart sémantique*** ».

– Combinaisons, y compris le *feedback*.



- Besoin d'**interfaces spécifiques**.

✍ **L'écart sémantique:** Absence de coïncidence entre les informations visuelles et leur interprétation par un utilisateur dans une situation donnée.

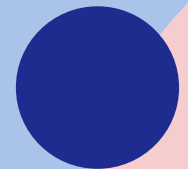
Problème



Recherche multimédia (*Multimedia Retrieval*)

Niveau « Signal »

- **Signal :**
 - **Variable** dans le **temps**, dans l'**espace** et/ou dans d'**autres dimensions physiques**,
 - **Analogique** : phénomène physique (pression d'une onde acoustique ou distribution d'intensité lumineuse) ou sa modélisation par un autre (électronique ou chimique par exemple),



- **Numérique:** même contenu mais "discrétisé"
 - de valeurs,
 - du temps,
 - de l'espace,
 - et/ou autres (fréquence lumineuse par exemple).

Recherche multimédia (*Multimedia Retrieval*)

Niveau « Signal »

- **Exemples :**

- **Son** (monophonique) :

- valeurs échantillonnées à 16 kHz sur 16 bits

- (une dimension temporelle, zéro dimension spatiale),



– **Image fixe** (monochrome) :

valeurs échantillonnées sur une grille 2D sur 8 bits

(dimension temporelle nulle, deux dimensions spatiales ;

la fréquence d'échantillonnage spatiale dépend du capteur),

20



- **Son stéréo, image couleur** : multiplication des canaux (dimension supplémentaire),
- **Vidéo** (séquence d'images) : comme une image fixe mais en plus échantillonnée dans le temps (24-30 Hz ; une dimension temporelle, deux dimensions spatiales, une dimension chromatique),
- **Images 3D** (scanners), séquences 3D, ...

Recherche multimédia (*Multimedia Retrieval*)

Niveaux « signal » et « sémantique »

- **Sémantique** (par opposition au **signal**) :
 - Concepts et relations « abstraites »,
 - Représentations symboliques (aussi signal),
 - Des niveaux d'abstraction successifs **du niveau** « signal / physique / concret / objectif » **au niveau** « sémantique / conceptuel / symbolique / abstrait / subjectif »,

- Distinction peu artificielle: **niveaux intermédiaires** difficiles à comprendre,
- Recherche au niveau **signal**, niveau **sémantique** ou une **combinaison** des deux.

Recherche multimédia (*Multimedia Retrieval*)

Requête Par Exemple (RPE)



***Indexation basée sur le contenu par
apprentissage supervisé***

***Concept
d'annotations***

***Documents
d'entraînements
(training)***

Extraction

Descripteurs

Entraînement (Train)

Modèle

Documents tests

Extraction

Descripteurs

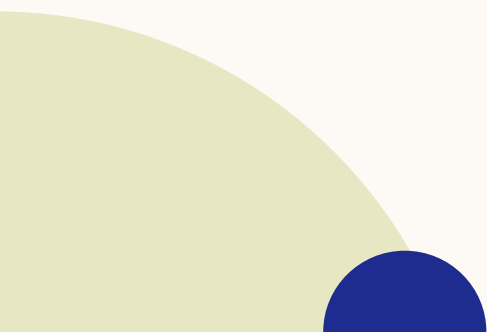
Prédire

Scores (par exemple probabilité de présence du concept)

25

Recherche multimédia (*Multimedia Retrieval*)

Recherche basée sur le contenu

- Aspects :
 - ***Signal*** : tableaux de nombres (« niveau bas »),
 - ***Sémantique*** : concepts ou mots-clés (« haut niveau »).
- 

- Recherche :

- **Sémantique** → **sémantique** : classique pour le texte,
- **Sémantique** → **signal** : images correspondant à un concept ?
- **Signal** → **signal** : image contenant une partie d'une autre image ?
- **Signal** → **sémantique** : concepts associés à une image ?

- Approches :
 - **Bottom-up** : *signal* → *sémantique*,
 - **Top-down** : *sémantique* → *signal* ,
 - **Combinaison** des deux.

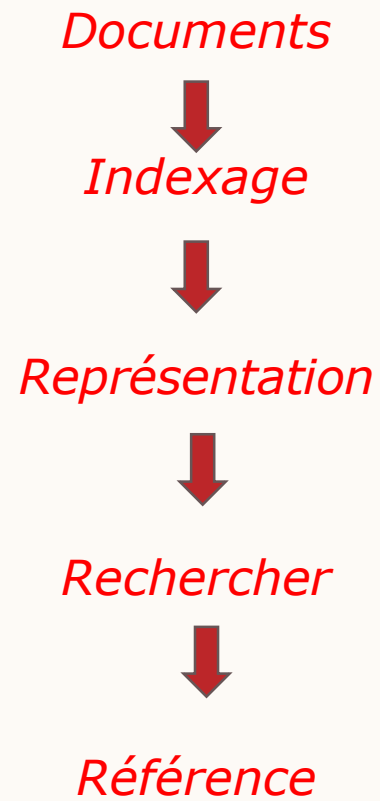
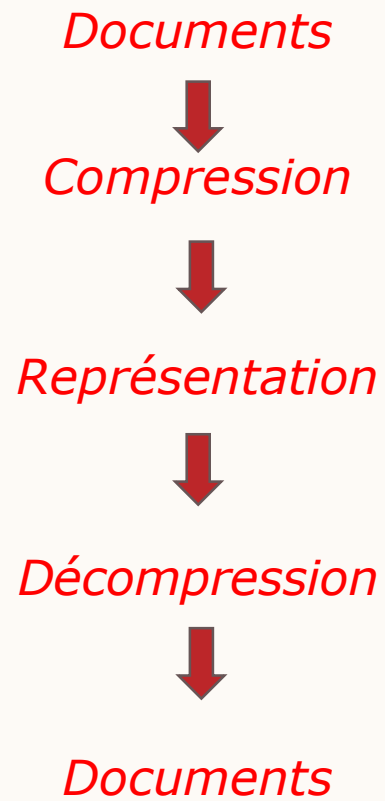
Recherche multimédia (*Multimedia Retrieval*)

Représentation de documents

- **Compression** : *encodage et décodage*
- **Indexation** : caractérisation du contenu

JPEG DV
GIF MPEG-1
PNG MPEG-2
MJPEG
MPEG-4

Base de données vs. Moteur de recherche



Problèmes

- Choix du **modèle de représentation**,
- **Méthode d'indexation** et organisation des index,
- **Choix** et mise en place du **moteur de recherche**,
- **Volume de données** très élevé,
- Nécessité d'une **intervention manuelle**.

Recherche multimédia (*Multimedia Retrieval*)

Modèles de représentation

- Niveau **sémantique** :
 - mots-clés, groupes de mots, concepts (i.e. l'outil linguistique **thésaurus**),
 - Graphes conceptuels (concepts et relations),
- Niveau **signal** :
 - Vecteurs de caractéristiques,
 - Ensembles de points d'intérêt,



- Niveau ***intermédiaire***:
 - Transcription d'une piste audio,
 - Jeux d'images clés,
 - Représentations mixtes et structurées des détails,
 - Spécificités du domaine d'application,
- Normes (***MPEG 7***: outils standardisés de description des contenus)

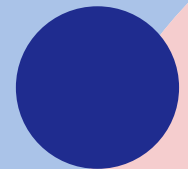
Recherche multimédia (*Multimedia Retrieval*)

Méthodes d'indexation et organisation des index

Après construction des représentations à partir du contenu du document,

- Extraire les fonctionnalités pour chaque document ou partie de document:
 - Niveau **signal** : traitement **automatique**,
 - Niveau **sémantique** : (plus complexe) de **manuel** à **automatique**.

- Organiser globalement les fonctionnalités de la recherche :
 - Trier, classer, pondérer, tabuler, formater, ...
- Spécificités du domaine d'application,
- Problème du compromis **qualité/coût**.



Recherche multimédia (*Multimedia Retrieval*)

Choix et mise en place du moteur de recherche

- Rechercher la « meilleure correspondance » entre une **requête** et les **documents**,
- **Sémantique** → **Sémantique**:
 - Modèles **logiques**, **vectoriels** et **probabilistes**,
 - Mots-clés, groupes de mots, graphes conceptuels, ...

- **Sémantique** → **Signale** :

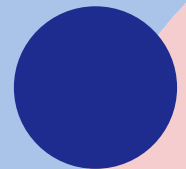
- Correspondance évaluée lors de la phase d'indexation (en général).

- **Signale** → **Signale** :

- Couleur, texture, points d'intérêt, ...

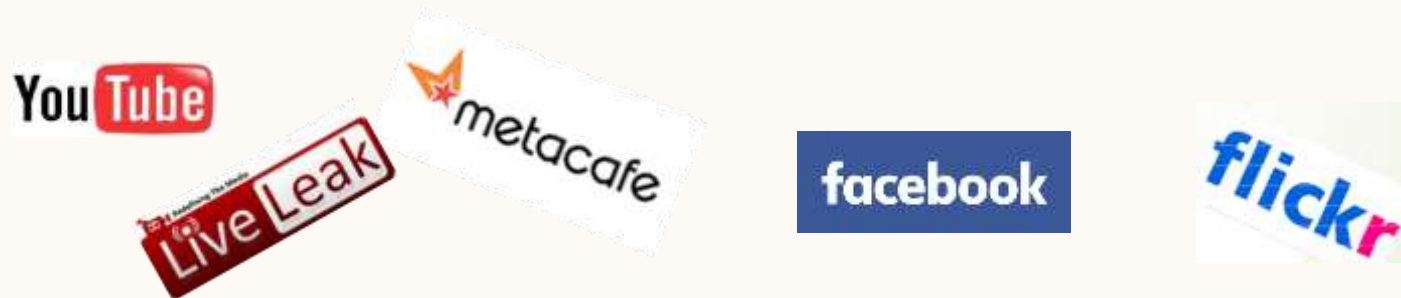
- Images, morceaux d'images, croquis, ...

- Recherche avec des requêtes **mixtes**.



3.Applications

- Beaucoup de contenu multimédia sur le **Web**
 - **Réseaux sociaux**, par exemple *Facebook, MySpace, Hi5*, etc.
 - **Partage de photos**, par exemple *Flickr, Photobucket, Instagram, Google Photos*, etc.
 - **Partage de vidéos**, par exemple, *YouTube, Metacafe, blip.tv, Liveleak*, etc.



3.Applications

Google Photos: partage de photographies et vidéos



- Analyse et classe les photos selon les :

- **Personnes** qui apparaissent dessus,
- Les **endroits**,
- Les **objets** présents,





- Classe les photos en **trois groupes** : *Lieux*, *Contacts* et *Thèmes* (ex: ciel, chat, panorama)
- Fonctionne avec un système de **reconnaissance** automatique des **visages** (désactivé dans certains pays).





Google Lens : technologie de reconnaissance d'image

- **Affiche** les informations pertinentes liées aux **objets identifiés** à l'aide d'une **analyse visuelle** reposant sur le **réseau neurones**
- **Intègre** les applications des **caméras** standards **d'Android**,
- **Identifie** les objets en lisant les **codes-barres**, les **codes QR**, les **étiquettes**, le **texte**, etc...

4. Caractéristiques des Systèmes de recherche multimédia

- **Statique:** nombre élevé de requêtes de recherche (***accès en lecture***), *peu de modifications* des données
- **Dynamique:** souvent des ***modifications des données***
- **Passif:** la base de données ne ***réagit qu'aux demandes*** de l'extérieur
- **Actif:** la fonctionnalité de la base de données conduit à des ***opérations*** au niveau de l'***application***

Exemples

1. Recherche **statique passive**

– **Art**: Cas d'utilisation historique

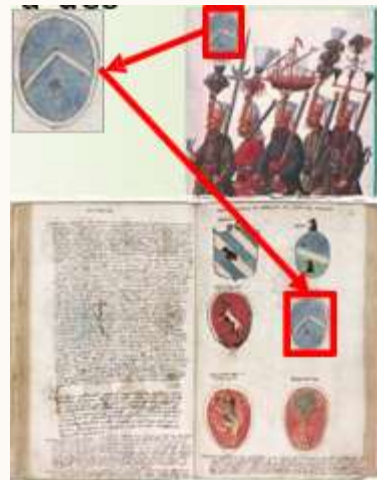
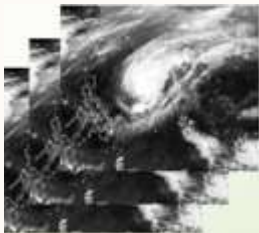
1. Recherche **dynamique active**

– Alerte pluies grâce à l'évaluation des photos satellites

Armoiries enregistrée dans
une base de données
multimédia

Extraction

↪ Signal d'avertissement
de **typhon**



5.Évaluation de la recherche

- **L'évaluation** des techniques de la recherche

- **Efficacité** du système

- **Utilisation efficace** des ressources système
- **Extensible** même sur de grandes collections
- Dépend de l'**environnement** (matériel): Utilisation de la mémoire, CPU-time, Nombre d'opérations E / S, Temps de réponse....



–**Effectivité** du processus de recherche

- **Haute qualité** du résultat
- Dépend de la **requête** et de la **pertinence** des résultats
- Il faut prendre en compte des besoins implicites d'information
 - ▶ Priorité à l'évaluation de l'utilité, de la facilité d'utilisation et de la convivialité du système

- Le **processus de recherche** doit être **efficace** ou **effective**?

Cela dépend de l'application!

1. Pertinence

- La **pertinence** en tant que mesure de **recherche**: chaque document sera classé binaire comme **pertinent** ou **non pertinent** par rapport à la requête.
 - Cette classification est effectuée manuellement par des «**experts**»
 - La réponse du système à la requête sera comparée à cette classification
 - Comparer la réponse obtenue avec le résultat «idéal»

Les **experts** disent: c'est **pertinent**



recherché (= pertinent)

Collection

**trouvé (= résultat de la
requête) trouvé (= résultat
de la requête)**



La **recherche automatique** dit:
pertinent

5.3 Faux positifs

Documents **non pertinents**, classés comme **pertinents** par le système: *Fausses alarmes (fa)*

- Augmenter inutilement le jeu de résultats
- Généralement inévitable (ambiguïté)
- Peut être facilement éliminé par l'utilisateur

5.4 Faux négatifs

Documents **pertinents**, classés par le système comme **non pertinents**: *Faux départ (fd)*

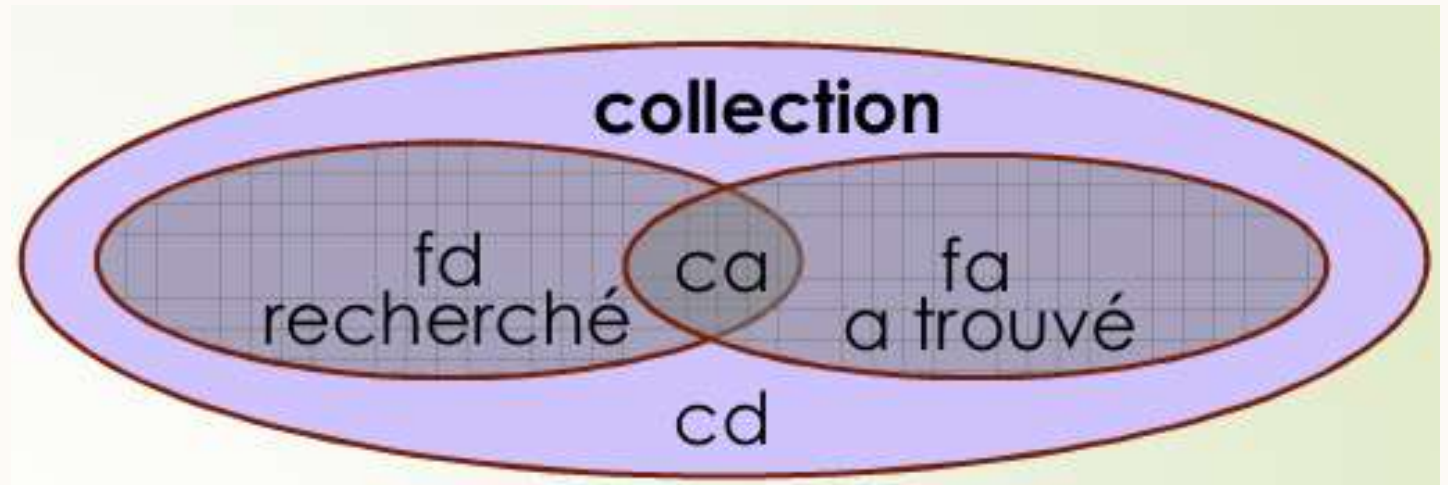
- **Dangereux**, car ils **ne peuvent pas être facilement détectés par l'utilisateur**
 - Y a t-il de «meilleurs» documents dans la collection que le système n'a pas renvoyés?
 - Les *fausses alarmes* ne sont généralement **pas aussi mauvaises** que les *faux départ*

5.5 Correctes positifs

Tous les documents **correcte**, **classés** par le système comme **pertinent**: *correctes alarmes (ca)*

5.6 Correctes négatifs

Tous les documents **correcte**, **classés** par le système comme **non pertinents**: *corrects départ (cd)*



Tous les ensembles sont disjonctifs et leur réunion constitue la totalité de la collection de documents

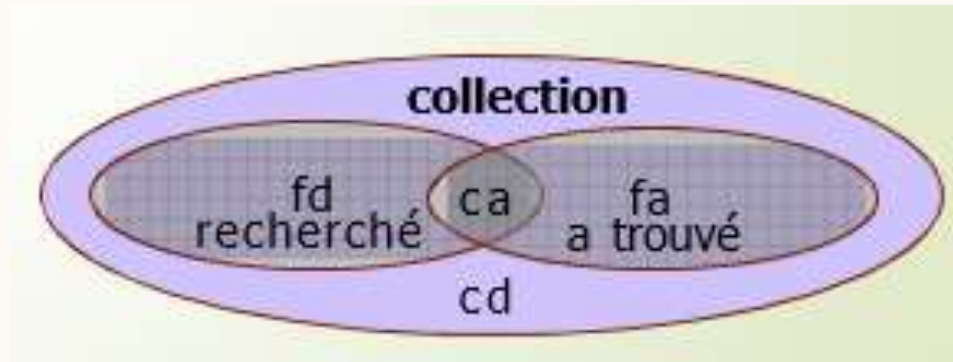
5.7 Vue d'ensemble

- **Matrice de confusion:** visualise l'efficacité d'un algorithme

Évaluation de l'utilisateur \ Évaluation du système	pertinent	impertinent
pertinent	<i>ca</i> ✓	<i>fd</i> ✖✖
impertinent	<i>fa</i> ✖	<i>cd</i> ✓

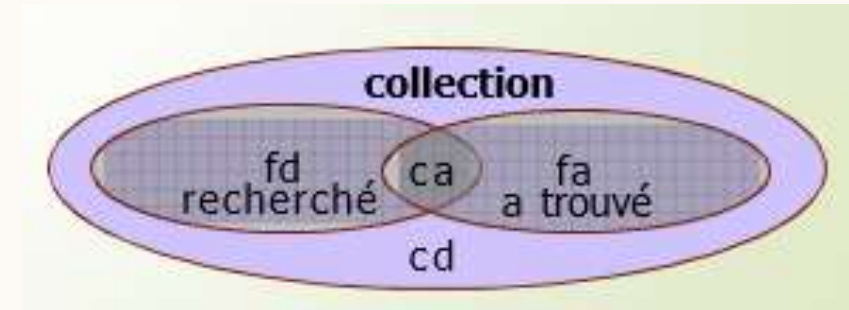
5.8 Interprétation

- **Résultats pertinents** = $fd + ca$
 - Trié à la main par des experts!
- **Résultats récupérés** = $ca + fa$
 - Récupéré par le système



5.9 Analyse Précision-Rappel (Precision-Recall)

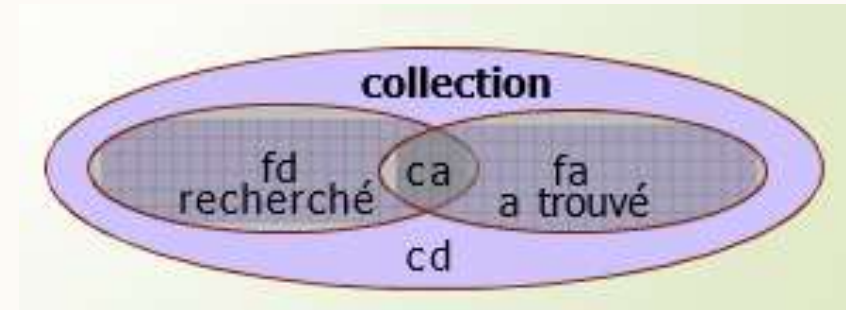
- **Précision**, mesure le **rapport** de documents *correctement renvoyés* par rapport à tous les *documents renvoyés*: $P = ca / (ca + fa)$
 - Valeur entre $[0, 1]$ (**1** représentant la **meilleure** valeur)
 - Un nombre élevé de fausses alarmes signifie **mauvais résultats**



• **Rappel**, mesure le **rapport** entre les documents *correctement renvoyés* et tous les *documents pertinents*:

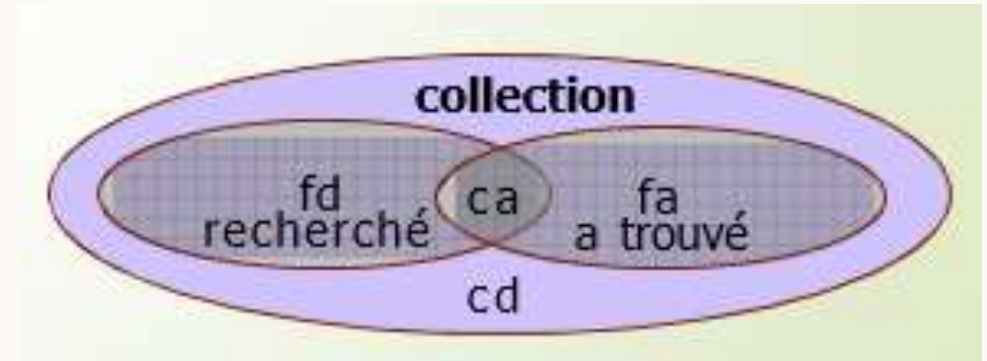
$$R = ca / (ca + fd)$$

- Valeur entre [0, 1] (**1** représentant la **meilleure** valeur)
- Un nombre élevé de faux départ signifie mauvais résultats

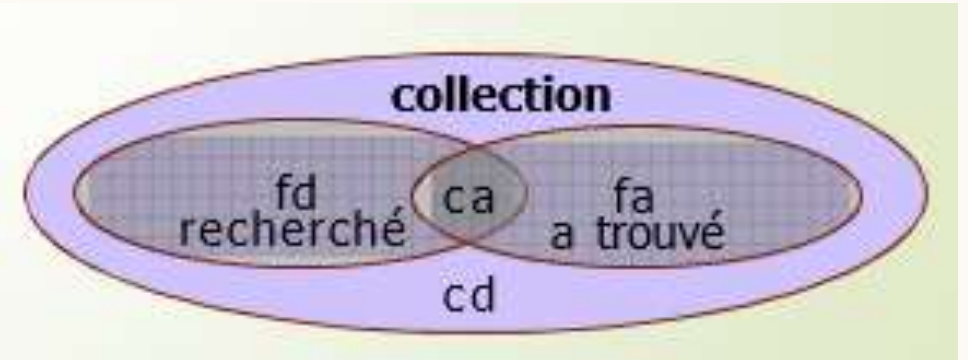


- **Précision-Rappel**, les deux mesures n'ont de sens que si elles sont considérées **en même temps**

- **Par exemple**, obtenez un **rappel** parfait en renvoyant simplement tous les documents, mais la **précision** est alors extrêmement basse...

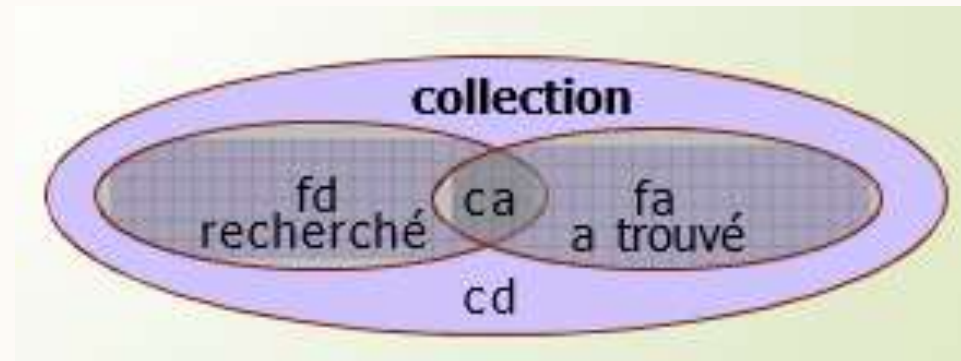


- Peut être **équilibré** en réglant le système
 - **Par exemple**, des *ensembles* de résultats *plus petits* entraînent de meilleurs taux de **précision** au prix du **rappel**
- Généralement, la **moyenne précision-rappel**, de davantage requêtes, est aussi **prise en compte** (macro évaluation)



4.12 Évaluation réelle

- **Alarmes** (éléments retournés) divisés en **ca** et **fa**
 - La précision est facile à calculer
- **Départ** (éléments non retournés) ne sont pas si faciles à diviser en **cd** et **fd**, car toute la collection doit être classée
 - Le rappel est difficile à calculer



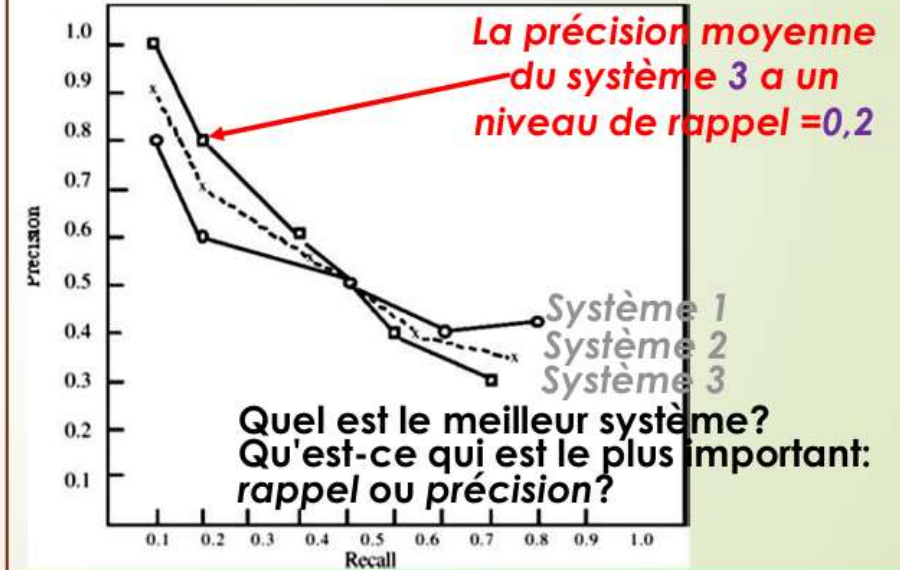
- **Benchmarks** standardisés
 - Connexions et requêtes fournies
 - Ensembles de résultats annotés

Exemple

Requête	fa	ca	fd	cd	P	R
Q1	8	2	6	4	0,2	0,25
Q2	2	8	2	8	0,8	0,8
Moyenne					0,5	0,525

Représentation

Courbes Précision-Rappel



Énoncé de l'exercice

Soit une collection de **10 images** (I_1 à I_{10}). Un expert a déterminé que pour une requête donnée, seules les images I_1, I_3, I_5 **et** I_7 sont pertinentes (soit un total de 4 images pertinentes).

Vous soumettez la requête au système, et celui-ci vous retourne la liste suivante (ordonnée par score de similarité) : **Résultats du système** : $[I_1, I_2, I_3, I_8, I_5]$

Question 1 : Identifier les valeurs ca, fa, fd, cd

Question 2 : Calculer la Précision et le Rappel

Question 3 : Calcul de la courbe Précision/Rappel

MERCI