



INFORMATION NUMÉRIQUE

Enjeux et Pratiques

Licence 3 / UE602 – Accès expert à l'information – 2025-2026

Bases de données bibliographiques

valerie.bonvallot@inist.fr

Sommaire : bases de données bibliographiques

1. Bases de données bibliographiques

- Définitions
- Historique
- Recherche bibliographique et requêtage

2. Interrogation de sources en libre accès

- exemple PubMed
 - HAL
 - OpenAlex
 - Istex
- } A vous de jouer

3. Conclusion

1. Définitions et historique

1.1 Définitions

- ▶ **Base de données** : « ensemble d'informations structurées accessibles au moyen d'un logiciel » (*Petit Robert*)
- ▶ **Base de données bibliographiques** : Des bases de données bibliographiques sont des outils structurés, complets et performants, en accès libre (= gratuit et ouvert à tous, ou *open access*) ou payant (= base propriétaire ou éditeur). Elles proposent des **références bibliographiques** (ou **notices bibliographiques**), qui décrivent de façon détaillée les publications qui ont été sélectionnées. Elles sont organisées en tables, composées de **métadonnées** ou de **champs** (exemples : auteur, titre, résumé, mot-clé, éditeur, source, etc.).
<https://coop-ist.cirad.fr/trouver-l-information/choisir-les-sources/3-les-bases-de-donnees-bibliographiques>
- ▶ **Recherche bibliographique** : Ensemble des méthodes, procédures et techniques ayant pour objet de retrouver les références bibliographiques de documents pertinents (sur un sujet défini en amont).
<https://adbs.fr/publications/glossaire>

1.1 Définitions

- ▶ **Archive ouverte** : Plateforme numérique institutionnelle ou thématique contenant des documents issus de la recherche scientifique, généralement déposés par leurs auteurs, et permettant au grand public d'y accéder gratuitement et sans contraintes. <https://skosmos.loterre.fr/TSO/fr/page/-ZNX5Z1VV-S>
- ▶ **Moteur de recherche internet** : Un moteur de recherche est un programme en ligne qui utilise un ensemble complexe de commandes, appelé algorithme, pour aider les utilisateurs à trouver des informations sur Internet
- ▶ **Agent conversationnel** : assistants virtuels alimentés par l'IA conçus pour interagir avec les utilisateurs en langage naturel, simulant des conversations similaires à des conversations humaines.

1.1 Définitions

- **Accès ouvert (AO) - open access (OA)** : L'accès ouvert signifie le **libre accès** à l'information et l'utilisation sans restriction des ressources électroniques pour tous. Tout type de contenu numérique peut être en libre accès, qu'il s'agisse de textes, de données, de logiciels, de fichiers audio ou vidéo ou de supports multimédias. Si la plupart d'entre eux ne concernent que le texte, un nombre croissant intègre le texte avec des images, des données et du code exécutable.

<https://www.unesco.org/fr/open-access>

- **C'est à vous** :

Citez une base de données bibliographiques

Citez un moteur de recherche scientifique

Citez un chatbot

1.1 Définitions

Tableau comparatif

	BdD	Moteur	Chatbot
Contenu			
Traitement			
Recherche			

1.2 Historique

► Des bases de données payantes (propriétaires)

- **1964 Web of Science** Eugène Garfield
 - Etats-Unis, gérée par Clarivate
 - Multidisciplinaire : Physical Sciences, Health Sciences, Life Sciences, Social Sciences & Humanities.
 - Plus de 21 000 revues scientifiques dépouillées, sélectionnées pour leur « qualité ». Plus de 79 millions d'enregistrements, plus de 11 9000 ouvrages, plus de 220 000 conférences couvertes.
 - Les citations pour une aide à la recherche d'information (RI)

1.2 Historique

► Des bases de données payantes (propriétaires)

- **1964 Web of Science** Eugène Garfield
 - **1998** : en ligne
 - Une base de données « classique » : champs/métadonnées pour RI
 - Les citations pour l'évaluation
 - Facteur d'impact
 - H-index
- **2014** : Incites outil d'analyse de la performance de la recherche développé par Clarivate Analytics. Il utilise des (méta)données de Web of Science pour produire des tableaux, des graphiques et des rapports. InCites comprend le Journal Citation Reports (JCR), InCites Essential (Essential Science Indicators)

1.2 Historique

► Des bases de données payantes (propriétaires)

■ 2004 Scopus

- Europe, gérée par Elsevier
- Multidisciplinaire
- 29 000 revues, 426 000 ouvrages, 102 600 000 enregistrements
- Une base de données « classique » : champs/métadonnées pour RI
- Les citations pour l'évaluation
 - Indicateurs de Scopus
- **2014** Scival : outil analytique et stratégique, créé par Elsevier à l'intention des agents et organismes responsables de la planification de la recherche à partir des données de Scopus

1.2 Historique

► Aux archives

► Archive nationale Hal



2001 Plateforme nationale d'archive ouverte multidisciplinaire en ligne développée par le **CNRS**, destinée au dépôt et à la diffusion d'articles scientifiques de chercheur.se.s, publiés ou non, et de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. L'accès aux données est libre, mais pas nécessairement leur utilisation ou réutilisation.

1.2 Historique

► A l'open access

► PubMed 1996 en ligne



40 millions d'enregistrements

- interface d'interrogation permettant l'accès à la base de données bibliographiques en libre accès *MEDLINE* (base de données bibliographiques produite par la National Library of Medicine, bibliothèque américaine spécialisée en médecine et dans les sciences et techniques associées).
- Gratuite, domaines de la médecine et de la biologie. La plupart des références contiennent un lien vers l'article entier (texte intégral). biologie, biochimie, médecine clinique, psychiatrie, santé publique, pharmacologie, toxicologie, odontologie, soins infirmiers, médecine vétérinaire, histoire de la médecine, économie, éthique..
- permet aussi d'interroger le **thésaurus** (liste structurée et hiérarchisée des termes d'un domaine) spécialisé dans le domaine biomédical *MESH* (Medical Subjects Headings) de la base *MEDLINE*.

1.2 Historique

► A l'open access



► OpenAlex

- Base de données bibliographiques en libre accès créée en 2022 (par l'organisme à but non lucratif OurResearch, qui détient aussi la base *Unpaywall*). Elle compte plus de 250 millions de documents et moissonne plusieurs référentiels et bases de données en libre accès.
- Elle contient des notices bibliographiques qui renvoient au texte intégral des publications en libre accès.
- Elle est présentée comme une alternative ouverte et gratuite aux bases de données commerciales WoS (Clarivate) et Scopus (Elsevier).
- Pour favoriser la transparence et l'accès libre à la connaissance scientifique, le Ministère de l'enseignement supérieur et de la recherche a établi en 2024 un partenariat avec *OpenAlex* (afin d'améliorer la qualité des données de la base). *OpenAlex* est donc considérée comme une infrastructure cruciale de science ouverte.

<https://www.ouvrirlascience.fr/partenariat-du-ministere-de-lenseignement-superieur-et-de-la-recherche-avec-openalex-pour-le-developpement-dun-outil-bibliographique-entierement-ouvert/>

1.2 Historique

► A l'open access



► Istex

- Infrastructure française (gérée par l'Inist) regroupant un ensemble de services et de ressources. 32 M de documents
- Gratuite pour tous : accès aux métadonnées
- Gratuite pour l'ESR : accès au texte intégral
- Réservoir de données éditeurs (payantes-budget) et de données ouvertes
- Multidisciplinaire - Multilinguisme
- Ressources terminologiques : Loterre <https://loterre.istex.fr/fr/>
- Services de fouille de textes : [Istex TDM](#) et [TDM Factory](#)

1.3 Requêtage - recherche bibliographique

Méthodologie

1. Identification du sujet

- Prise en main du sujet, choix des mots-clés pertinents (question : qui - quoi - où - quand - comment). Synonymes. Traductions
- Validation des mots-clés auprès du commanditaire

2. Interrogation des bases de données bibliographiques

- Choix des sources d'information pertinentes
- Rédaction de l'équation de recherche + tests. Opérateurs booléens, opérateurs de proximité, parenthésage, expression exacte ...

3. Préparation du corpus de notices bibliographiques

- Export des notices bibliographiques
- Envoi au commanditaire (chercheur) pour validation

1.3 Interrogation de bases de données bibliographiques

Exemple de notice bibliographique (PubMed)

PMID- 17833902
 OWN - NLM
 STAT- PubMed-not-MEDLINE
 DCOM- 20100702
 LR - 20100608
 IS - 0036-8075 (Print)
 IS - 0036-8075 (Linking)
 VI - 143
 IP - 3611
 DP - 1964 Mar 13
 TI - Pleistocene Wood Rat Middens and Climatic Change in Mohave Desert: A Record of
 Juniper Woodlands.
 PG - 1171-3
 AB - Leafy twigs and seeds of juniper are abundant in nine ancient Neotoma middens
 discovered in low, arid, desert ranges devoid of junipers, near Frenchman Flat,
 Nevada. Existing vegetation is creosote bush and other desert shrubs. Twelve
 radiocarbon dates suggest that the middens were deposited between 7800 to more
 than 40,000 years ago. Dominance of Utah juniper and absence of pinyon pine in
 most deposits indicates a local Pleistocene woodland climate more arid than the
 usual pinyon-juniper climate.
 FAU - Wells, P V
 AU - Wells PV
 FAU - Jorgensen, C D
 AU - Jorgensen CD
 LA - eng
 PT - Journal Article
 PL - United States
 TA - Science
 JT - Science (New York, N.Y.)
 JID - 0404511
 EDAT- 1964/03/13 00:00
 MHDA- 1964/03/13 00:01
 CRDT- 1964/03/13 00:00
 PHST- 1964/03/13 00:00 [pubmed]
 PHST- 1964/03/13 00:01 [medline]
 PHST- 1964/03/13 00:00 [entrez]
 AID - 143/3611/1171 [pii]
 AID - 10.1126/science.143.3611.1171 [doi]
 PST - ppublish
 SO - Science. 1964 Mar 13;143(3611):1171-3. doi: 10.1126/science.143.3611.1171.



Champ «PMID»



Champ « Date de publication » DP



Champ «Titre» TI



Champ « Résumé » AB



Champs «Auteur» FAU / AU



Champ « Titre de la source » JT

1.3 Interrogation de bases de données bibliographiques

Champs (PubMed)

PubMed Advanced Search Builder

Add terms to the query box

All Fields

- Affiliation
- All Fields
- Author
- Author - Corporate
- Author - First
- Author - Identifier
- Author - Last
- Book
- Conflict of Interest Statements
- Date - Completion
- Date - Create
- Date - Entry
- Date - MeSH
- Date - Modification
- Date - Publication
- EC/RN Number
- Editor
- Filter

- Grants and Funding
- ISBN
- Investigator
- Issue
- Journal
- Language
- Location ID
- MeSH Major Topic
- MeSH Subheading
- MeSH Terms
- Other Term
- Pagination
- Pharmacological Action
- Publication Type
- Publisher
- Secondary Source ID
- Subject - Personal Name
- Supplementary Concept
- Text Word

- Text Word
- Title
- Title/Abstract
- Transliterated Title
- Volume

1.3 Interrogation de bases de données bibliographiques

Modes d'interrogation de bases de données bibliographiques

Recherche simple Recherche avancée (avec la rédaction d'une équation de recherche)

Equation de recherche

Organisation de termes (=mots-clés) et d'opérateurs de recherche, selon une syntaxe préétablie et spécifique au langage d'interrogation utilisé, pour répondre à une demande d'information.

<https://adbs.fr/publications/glossaire>

Indexation

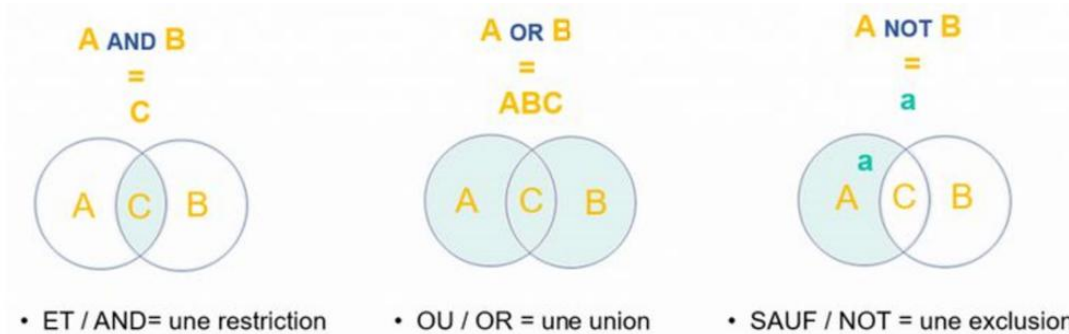
Vocabulaire libre / vocabulaire contrôlé

Thésaurus : liste organisée de termes contrôlés, relation hiérarchique, sémantique entre les termes (mots-clés, descripteurs). Terme préférentiel - Terme générique - Terme spécifique -
Synonymes - Traductions

TG Véhicule
TS Quatre roues
TS Camion
TS Voiture
syn. : *bagnole*
TS Autocar
syn : *bus, car*

1.3 Interrogation de bases de données bibliographiques

Opérateurs booléens de recherche principaux



Exemples : climate AND change climate OR change climate NOT change

Opérateurs de proximité

Pour réduire l'espace entre 2 termes par rapport au AND

Souvent NEAR(n) où n est le nombre de termes séparant les 2 termes recherchés

Exemple : climate NEAR(2) change, il peut y avoir jusqu'à 2 termes entre climate et change

1.3 Interrogation de bases de données bibliographiques

Troncatures et expression exacte

Troncature pour remplacer des lettres

* en fin de mot : change* => changer, changement ...

Joker pour remplacer 1 lettre

? au sein d'un mot ou à la fin : r?chauffement => réchauffement, rechauffement

Expression exacte (« phrase » en anglais)

changement climatique # « changement climatique »

par défaut souvent AND (mais à vérifier)

changement climatique = changement AND climatique

Parenthésage

Pour mettre des priorités et faire des regroupements

(change* OR variation*) AND (climat* OR temp?rature*) AND (France OR Italie*)

2. Interrogation de sources en libre accès

Sujet de recherche

Les effets du réchauffement climatique chez les jeunes.

Un historique sur 10 ans

1.3 Interrogation de bases de données PubMed

Quels mots-clés ?

Synonymes

Traduction

1.3 Interrogation de bases de données

PubMed

Recherche simple :

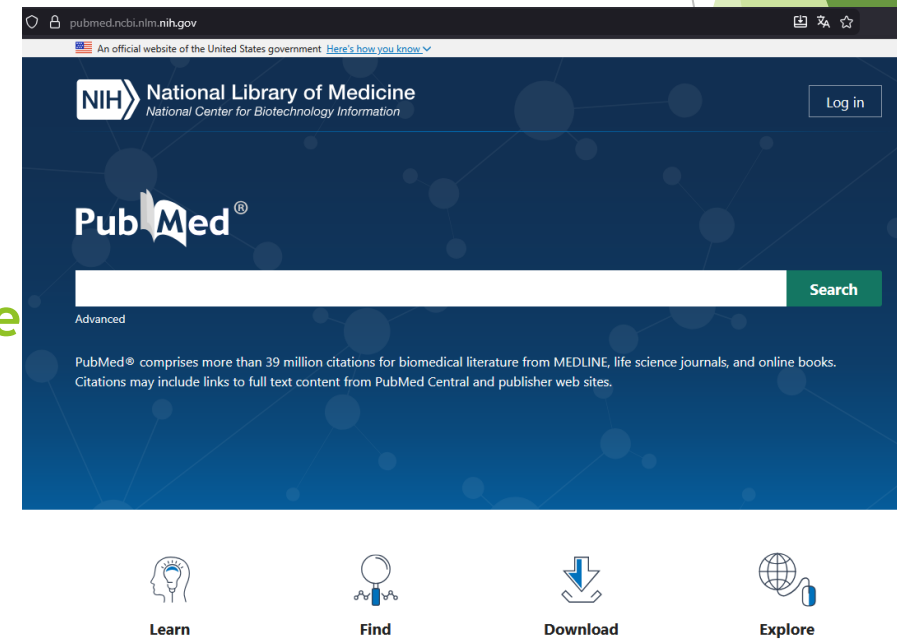
<https://pubmed.ncbi.nlm.nih.gov/>

- Un multiterme sans puis avec guillemets : modalités de recherche d'expression exacte
- Un multiterme sans puis avec opérateur booléen : détection de l'opérateur par défaut
- Un multiterme en français (avec puis sans accent) puis en anglais : modalités de recherche du multilinguisme

→ traduction : Google, Deeple,

Hetop : <https://www.HeTOP.eu/hetop/rep/fr/MESH/>

Chercher un guide d'interrogation de la base étudiée



1.3 Interrogation de bases de données PubMed

Hetop : <https://www.HeTOP.eu/hetop/rep/fr/MESH/>

The screenshot shows the HeTOP website interface. The top navigation bar includes the HeTOP logo and a search bar. Below the search bar, there are tabs for 'Description', 'Hiérarchies', 'Relations', and 'PubMed / DocCISMeF'. The main content area displays search results for 'réchauffement climatique'. On the left, there is a sidebar with 'Vos recherches' and 'Meilleurs candidats'. The main results section shows 'Réchauffement de la planète' (Descripteur MeSH) with various details including 'Identifiant d'origine', 'Libellé préféré', 'Définition du MeSH', 'Synonyme MeSH', and 'Synonyme CISMef'.

This screenshot shows the detailed view of the 'Réchauffement de la planète' (Descripteur MeSH) page. The page is divided into several sections: '1. Qualificatif(s) utilisable(s) pour ce mot clé :', '2. Options', and '3. Effectuer la recherche :'. The '1. Qualificatif(s)' section lists various qualifiers like 'histoire', 'organisation et administration', 'économie', and 'législation et jurisprudence'. The '2. Options' section includes checkboxes for 'seulement les principales' and 'sans explosion'. The '3. Effectuer la recherche' section features logos for 'DocCISMeF', 'LiSSa', and 'PubMed'. The right sidebar contains a list of related terms and a 'tous les types' dropdown menu.

1.3 Interrogation de bases de données PubMed

Recherche avancée : rédaction d'équation de recherche guidée

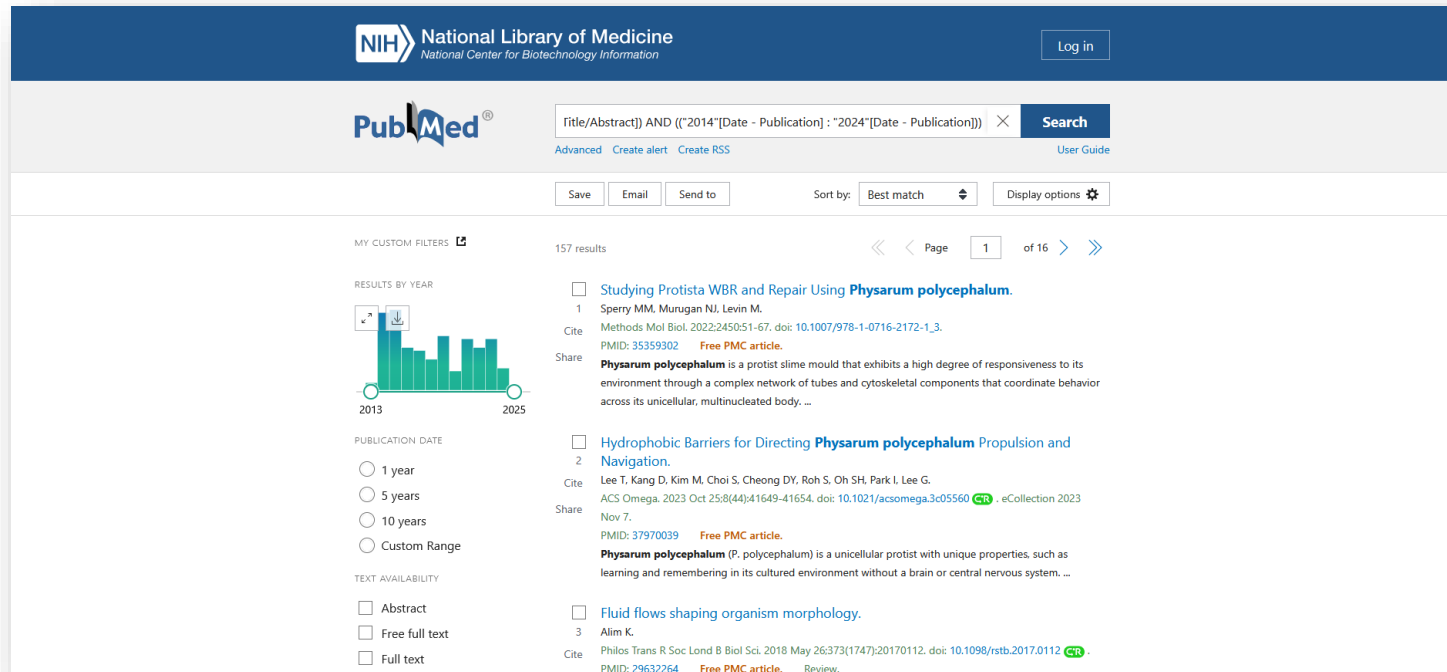
The screenshot shows the PubMed Advanced Search Builder interface. At the top, there's a header with the NIH logo and 'National Library of Medicine' text. Below this, the 'PubMed Advanced Search Builder' title is visible. The main area contains a 'Query box' with a dropdown menu for 'Add terms to the query box' (currently set to 'Title/Abstract'). A list of search terms is shown below the dropdown, including 'physarum polycephalum' and its variations with counts. To the right of the query box, there are buttons for 'ADD', 'Show Index', and 'Search'. Annotations with arrows point to these elements: 'liste des champs d'interrogation disponibles' points to the dropdown menu; 'saisie manuelle ou semi-automatique d'une équation de recherche' points to the query box; 'opérateurs booléens AND OR NOT disponibles' points to the 'ADD' button; and 'références indexées dans le thésaurus' points to the 'Show Index' button. At the bottom, there's a 'History and Search Details' section and a footer with links to NCBI Literature Resources, MeSH, PMC, Bookshelf, and Disclaimer.

➡ Tester sur différents champs : All fields / Title / Mesh

➡ Combiner les termes : opérateurs booléens, parenthésage, troncatures, expression exacte
Réchauffement climatique, global warming, adolescent(s), adolescence, jeune(sse), young, youth

1.3 Interrogation de bases de données PubMed

Recherche avancée : page de résultats



➔ Possibilité de filtrer les résultats grâce aux facettes : 10 dernières années

1.3 Interrogation de bases de données PubMed

Consultation du thésaurus MESH (utile pour définition / arborescence)

- Lien disponible dans le pied de page de *PUBMED* MeSh ou via l'URL <https://www.ncbi.nlm.nih.gov/mesh/>

The screenshot shows the PubMed Search Builder interface for the term "Global Warming". On the left, there is a description of "Global Warming" and various search options. On the right, there is a "PubMed Search Builder" box with a search input field and buttons for "Add to search builder", "AND", and "Search PubMed". Below this, there is a "Related information" section with links to "PubMed", "PubMed - Major Topic", "Clinical Queries", and "NLM MeSH Browser". At the bottom, there is a "Recent Activity" section showing a list of search terms and their corresponding MeSH counts.

Global Warming
Increase in the temperature of the atmosphere near the Earth's surface and in the troposphere, which can contribute to changes in global climate patterns.
Year introduced: 2010

PubMed search builder options
[Subheadings:](#)

- ☐ economics
- ☐ history
- ☐ legislation and jurisprudence
- ☐ mortality
- ☐ prevention and control
- ☐ statistics and numerical data

☐ Restrict to MeSH Major Topic.
☐ Do not include MeSH terms found below this term in the MeSH hierarchy.

Tree Number(s): G16.500.175.374.500
MeSH Unique ID: D057232
Entry Terms:

- Warming, Global

Previous Indexing:

- [Greenhouse Effect \(1994-2009\)](#)

See Also:

- [Carbon Footprint](#)
- [Greenhouse Gases](#)

[All MeSH Categories](#)
[Phenomena and Processes Category](#)
[Biological Phenomena](#)
[Ecological and Environmental Phenomena](#)
[Climatic Processes](#)
[Climate Change](#)
[Global Warming](#)

PubMed Search Builder

Add to search builder AND Search PubMed

Related information

PubMed
PubMed - Major Topic
Clinical Queries
NLM MeSH Browser

Recent Activity

Turn Off Clear

- Global Warming MeSH
- global warming (1) MeSH
- "global warming" (4) MeSH
- "rechauffement climatique" (0) MeSH
- rechauffement climatique (0) MeSH

See more...

Définition du terme spécifique (ou descripteur) « global warming»

Arborescence des descripteurs (relations hiérarchiques = du plus générique au plus spécifique) // organisation des connaissances

1.3 Interrogation de bases de données PubMed

Export du résultat

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

PubMed®

warming"[All Fields])) AND (("young"[All Fields]) OR ("adolescent"[All Fields]))

Search

Advanced Create alert Create RSS User Guide

Save • Email Send to Sort by: Most recent Display options

Save citations to file

Selection: All results on this page

Format: Summary (text)

Create file Cancel

All results on this page

All results on this page

All results

Selection

Summary (text)

Summary (text)

PubMed

PMID

Abstract (text)

CSV

1.3 Interrogation de bases de données PubMed

Alertes : besoin de s'identifier (gratuit)

2 A vous de jouer !

Rappel demande thématique de recherche :

Les effets du réchauffement climatique chez les jeunes. Un historique sur 10 ans

Méthodologie : répartition par petits groupes

- Prise en main du sujet et choix de **mots-clés en anglais et français**
- Validation des mots-clés auprès du chercheur
- Rédaction de l'équation de recherche
- Analyse des résultats

Se répartir en fonction des 3 bases

- HAL
- ISTEEX
- OpenAlex

2 A vous de jouer !

Critères	Réponses
Nom de la base	
Producteur de la base	
Année de création	
Nombre de documents à ce jour	
Conditions d'accès	
Mise à jour	
Domaines couverts	
Période couverte	
Types de documents recensés	
Langues traitées	
Modalités de recherche (simple, avancée ...)	
Vocabulaire libre / contrôlé	

2 A vous de jouer !

Critères	Réponses
Champs interrogeables	
Champs interrogés par défaut	
Opérateurs booléens - lequel est par défaut ?	
Opérateurs de proximité	
Expressions exactes	
Troncatures	
Requête	
Résultats : nb de documents	
Résultats : nb de documents / année	
Facettes	
Tri	
Téléchargement - Export	
Alerte	
Historique des recherches	

5. Conclusion

Derniers conseils méthodologiques lors d'une recherche bibliographique :

- **Bien choisir les sources d'information ou bases de données bibliographiques**, en fonction de sa recherche (bases propriétaires multidisciplinaires, bases en libre accès généralistes ou spécialisées dans un domaine d'activité, archives ouvertes...). Les résultats peuvent être très différents d'une base à l'autre.
- **Croiser plusieurs sources d'informations** si besoin.
- **Faire valider son corpus par le commanditaire** avant de démarrer toute analyse (comme dans le cas d'élaboration d'études bibliométriques, qui demandent de nombreux traitements de données (dédoublonnage, « curation », l'homogénéisation...).

Quelques remarques :

- Le mouvement de la **science ouverte** incite de plus en plus les chercheur.se.s à déposer leurs publications dans des archives ouvertes ou des revues en libre accès.
En savoir plus : <https://www.science-ouverte.cnrs.fr/>
- Actuellement, la recherche française et internationale se tourne de plus en plus vers des bases de données en libre accès.



Merci pour votre attention !

