

Présentation du Dictionnaire Électronique des Synonymes (DES)

Laurette CHARDON

Laboratoire CRISCO - EA 4255 - Université de Caen

Mai 2023

- 1 Qu'est-ce que le DES ?
- 2 La maintenance du DES
- 3 Les statistiques d'accès
- 4 Les éléments composant le DES
- 5 L'espace sémantique
- 6 Quelques points complémentaires

Historique^[1] I

- **Conception 1996-1999** : Bernard Victorri, mathématicien et linguiste CNRS et Sabine Ploux, ingénieure CNRS
- **Principe de base** : Représenter un dictionnaire aussi complet que possible sous forme de **graphe** dont les **sommets** sont les entrées et les **arêtes** les liens synonymiques pour créer un espace sémantique en traitant mathématiquement divers aspects de la polysémie d'une vedette.
- **La base de données** : issu de l'INALF constituée de 7 dictionnaires avec un travail d'homogénéisation et de symétrisation.

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

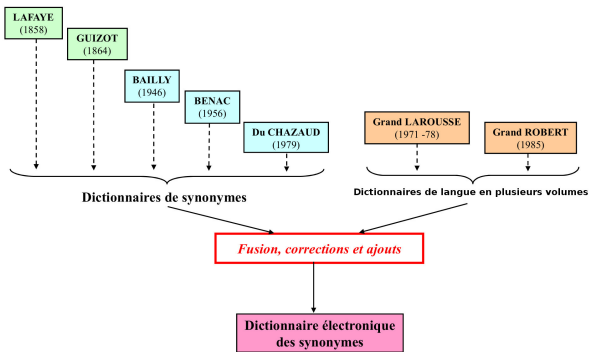
Quelques points complémentaires

Références

L'historique

L'interface actuelle

Historique^[1] II



Historique^[1] III

- La **symétrisation** signifie que si l'entrée A est synonyme de l'entrée B alors l'entrée B est synonyme de l'entrée A. La synonymie est définie ainsi : s'il existe **au moins un contexte** dans lequel on peut substituer B à A sans en modifier significativement le sens, alors il en existe au moins un dans lequel on peut substituer A à B
→ augmentation du nombre de liens.
- Le fondement des calculs sur le graphe pour l'étude de la polysémie est la notion de **clique**.

Interface actuelle I

http ://crisco.unicaen.fr/des



Accès DES - Accès direct DES - KMJ

Dictionnaire Electronique des Synonymes (DES)

Dernières Actualités : la lettre d'actualité n°6 du DES vient de paraître !

Tapez l'unité lexicale recherchée puis cliquez sur Valider ou tapez sur Entrée (données à jour du 26 août 2019)

Rechercher : Valider

curieux définition  espace sémantique 

55 synonymes

étonnant, étrange, amateur-trice, amusant, anxieux, attachant, attentif, avide, badaud, beau, bizarre, bizarroïde, chercheur, collectionneur-euse, déconcertant, désireux, dilettante, drolatique, drôle, extraordinaire, extravagant, fâneux, foule-au-pot, fouilleur, fouinard, fouineux, fureteur, fureteur, inconcevable, incroyable, indiscret, inouï, inquisiteur, insolite, intéressé, intéressant, intrigué, inusité, investigateur, juge, original, paradoxal, piquant, pittoresque, plaisant, rare, regardant, singulier, soucieux, spectateur-trice, surprenant, truculent, unique, voyeur

18 antonymes

étourdi, acrotopéage, anodin, banal, blasé, comman, courant, discret, habituel, incurieux, indifférent, insouciant, ordinaire, quelconque, quotidien, réservé, usuel, vulgaire

Interface actuelle II

Classement des premiers synonymes

étrange	■
bizarre	■
original	■
drôle	■
fureteur	■
fouineur	■
déconcertant	■
inquisiteur	■
pittoresque	■
singulier	■
truculent	■
insolite	■
paradoxal	■
étonnant	■
amusant	■
chercheur	■

59 cliques

- étonnant, étrange, bizarre, curieux, drôle, extraordinaire, insolite, original, singulier, surprenant
- étonnant, étrange, bizarre, curieux, extraordinaire, insolite, inusité, original, rare, singulier
- étonnant, étrange, bizarre, curieux, drôle, extraordinaire, original, singulier, truculent
- étonnant, étrange, bizarre, curieux, extraordinaire, original, paradoxal, singulier, surprenant
- étonnant, étrange, curieux, extraordinaire, inconcevable, incroyable, inouï, paradoxal, surprenant
- étrange, bizarre, curieux, drôle, extraordinaire, extravagant, insolite, original, singulier
- étonnant, étrange, bizarre, curieux, extraordinaire, incroyable, paradoxal, surprenant
- bizarre, curieux, drôle, extraordinaire, insolite, original, pittoresque
- bizarre, curieux, drôle, extraordinaire, original, pittoresque, truculent
- bizarre, curieux, extraordinaire, extravagant, original, singulier, unique
- bizarre, curieux, extraordinaire, original, rare, singulier, unique
- étonnant, étrange, bizarre, curieux, déconcertant, inusité, original

Les étapes mensuelles

- 1 Traitement et saisie des propositions des utilisateurs
- 2 Traitement et saisie des propositions calculées automatiquement par programme
- 3 Extraction de la base, traitements, vérifications, mise en ligne, sauvegardes
- 4 Mise à jour des statistiques (consultations, ajouts, modifications, contributeurs)
- 5 Traitement des mots les plus recherchés (journalier)
- 6 Mise à jour du DES autonome

Méthodologie suivie pour tout ajout dans la base de données

- Recherche d'une phrase dans laquelle l'un des termes peut remplacer l'autre et inversement \Rightarrow symétrie et contexte
- Même type d'objet (animé ou inanimé), même catégorie grammaticale
- Recherche des définitions dans le CNRTL, le Robert (en ligne via le SCD), wiktionary, parfois Frantext
- Exemples retenus : *référence - sommité ; essai-proposition ; facteur-paramètre ; hypothèse - supposition*
- Exemples non retenus :
 - pas dans les dictionnaires : *défour # déterrer, sérieux # agelaste*
 - pas de contexte : *récolte # enlèvement, engager # produire*
 - catégorie *grammaticale* différente : *frime # clinquant*

Statistiques BD I

Entre 2010 et 2022 :

- Le nombre d'entrées est passé de 49.147 à 50.379 ⇒ **plus de 1000 nouvelles entrées en 12 ans** : 2,2% par an
- Le nombre de liaisons synonymiques est passé de 200.601 à 212.232 → **plus de 10.000 nouvelles liaisons en 12 ans** : près de 5% par an
- Le nombre d'antonymes est passé de 12.036 à 12.651 → **500 de plus** : plus de 4% par an
- Le nombre de cliques est passé de 109.368 à 111.744 → **a évolué modérément : il y a eu de nouvelles cliques mais également beaucoup de regroupements**(cliques plus fournies)
- l'entrée qui a le plus grand nombre de synonymes : **BON avec 242**
- l'entrée qui a le plus d'antonymes : **CALME avec 192**

Statistiques BD II

Entre 2017 et 2022^[2] :

- 7457 propositions des utilisateurs regardées, 6920 valides et 5374 retenues soit 77 %
- 18627 modifications* sur le DES
⇒ **soit plus de 4650 par an**

* : synonymes et antonymes ajoutés, retirés, corrigés, modifications graphiques, variantes, ajout catégorie grammaticale

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

La mise à jour

Les statistiques d'évolution de la base

La base de données

La table des graphies :

id_graphie	int(5)	Identifiant
graphie	varchar(38)	Libellé
cnrtl	varchar(24)	Lien vers le dictionnaire CNRTL
date	timestamp	Date de création
nature	char(100)	Catégorie grammaticale

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

La mise à jour

Les statistiques d'évolution de la base

La base de données

La table des synonymes :

id	int(5)	Identifiant de la liaison
id_graphie1	int(5)	Identifiant renvoyant à la table des graphies
id_graphie2	int(5)	Identifiant renvoyant à la table des graphies
litigieux	tinyint(1)	1 : litigieux ; 0 : non litigieux (par défaut)
commentaire	int(10)	Identifiant renvoyant à la table des commentaires
date	timestamp	Date de création
observation	char(255)	commentaire interne
contexte	char(255)	phrase d'exemple

La table des antonymes :

id_graphie1	int(5)	Identifiant renvoyant à la table des graphies
id_graphie2	int(5)	Idem avec comme contrainte d'être supérieur au précédent
date	timestamp	Date de création

La table des commentaires :

id	int(11)	Identifiant
intitulé	varchar(160)	libellé du commentaire

Les statistiques de consultation - Année 2022

Récapitulatif des visites

 **4 641 953** visites

 **11 min 57s** durée moyenne d'une visite

 **22 %** visiteurs ayant survolé (quitté le site après une page)

 **9** actions (pages vues, téléchargements, liens sortants et recherches internes) par visite

 **0,14s** temps de génération moyen

 **40 537 094** pages vues, **34 622 962** pages vues uniques

 **196** recherches totales sur votre site, **126** mots-clés uniques

 **829** téléchargements, **680** téléchargements uniques

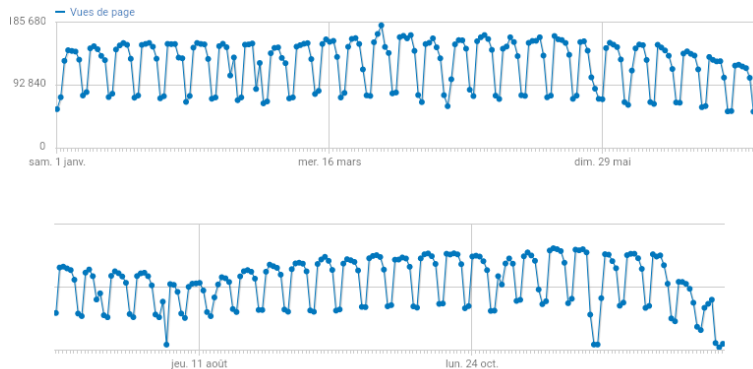
 **762 251** liens sortants, **724 367** liens externes uniques

 **6 522** actions max en une visite

Qu'est-ce que le DES ?
La maintenance du DES
Les statistiques d'accès
Les éléments composant le DES
L'espace sémantique
Quelques points complémentaires
Références

Les statistiques globales
Les statistiques journalières
Le profil des utilisateurs
Les mots les plus recherchés

Les statistiques de consultation - Année 2022



Qu'est-ce que le DES ?
La maintenance du DES
Les statistiques d'accès
Les éléments composant le DES
L'espace sémantique
Quelques points complémentaires
Références

Les statistiques globales
Les statistiques journalières
Le profil des utilisateurs
Les mots les plus recherchés

Les statistiques de consultation - Année 2022

▼ CRISCO

📅 DE 2022-01-01 À 2022-12-31

🏠 TOUTES LES VISITES



4 641 953 visites




















Les statistiques de consultation - Année 2022

Continent

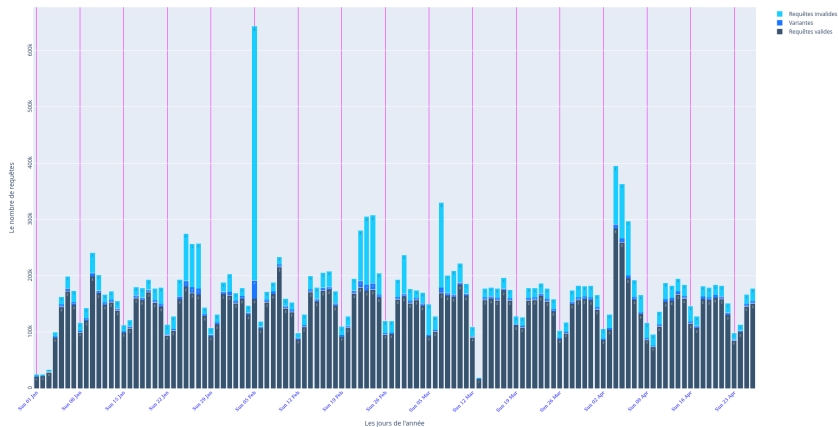
CONTINENT	▼ VISITES
Europe	3 937 545
Amérique du Nord	610 480
Asie	51 260
Afrique	14 272
Inconnu	13 781
Amérique du Sud	11 446
Océanie	5 555

Les statistiques de consultation - Année 2022

Pays		▼ VISITES	
 France	74,9 % 3 475 315	 Espagne	1,1 % 50 537
 États-Unis	8,6 % 401 457	 Italie	0,7 % 31 159
 Canada	4,5 % 208 174	 Corée du Sud	0,6 % 29 705
 Royaume-Uni	3,2 % 150 816	 Inconnu	0,3 % 13 781
 Suisse	1,6 % 74 393	 Japon	0,2 % 10 819
 Belgique	1,3 % 60 866	 Brésil	0,2 % 10 285
 Allemagne	1,2 % 55 968	 Pays-Bas	0,2 % 7 979
		 Russie	0,1 % 5 294
		 Chine	0,1 % 5 171
		 Maroc	0,1 % 4 919

Les statistiques journalières - Année 2023

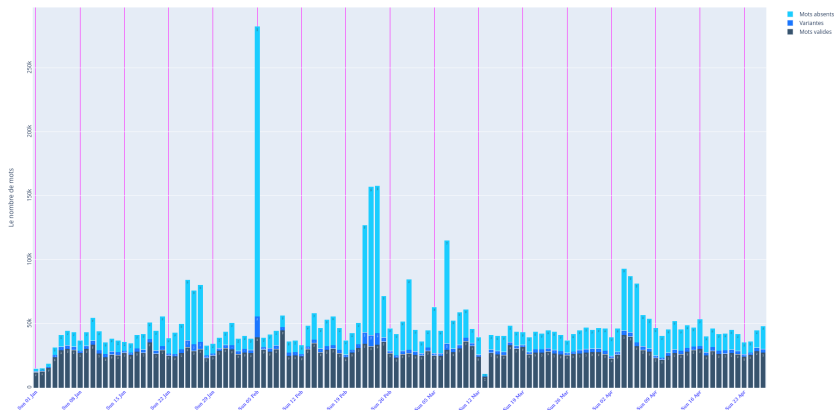
Requêtes en 2023 dans le Dictionnaire Électronique des Synonymes (DES) - CRISCO - UNICAEN



[Accès direct au graphe interactif](#)

Les statistiques journalières - Année 2023

Nombre de mots les plus recherchés en 2023 dans le Dictionnaire Électronique des Synonymes (DES) - CRISCO - UNICAEN



[Accès direct au graphe interactif](#)

16/71

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] [4] :



Dictionnaire Electronique des Synonymes (DES)

En ce qui concerne le dictionnaire de synonymes, je peux vous dire d'ores et déjà, puisque je l'utilise en ligne depuis plusieurs années, qu'il est remarquable par la richesse de l'éventail des synonymes qu'il propose et par la facilité avec laquelle on peut passer une autre nuance de vocabulaire en cliquant sur l'un des mots proposés... Mon seul regret était donc de ne pouvoir l'utiliser lorsque j'étais loin de tout, y compris une connexion Internet au Népal en particulier. Cette lacune est donc maintenant comblée avec le logiciel autonome. C'est donc un outil qui m'est très précieux et m'avait été signalé il y a quelques années par Patrick Carré, l'un des mes amis traducteurs du Comité de Traduction Padmakara, qui est aussi écrivain.

Matthieu Ricard
Docteur en génétique cellulaire, moine
bouddhiste tibétain, auteur et photographe
24 juillet 2018

Contact : crisco.webmestre@unicaen.fr

Qu'est-ce que le DES ?
La maintenance du DES
Les statistiques d'accès
Les éléments composant le DES
L'espace sémantique
Quelques points complémentaires
Références

Les statistiques globales
Les statistiques journalières
Le profil des utilisateurs
Les mots les plus recherchés

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] ^[4] :



Dictionnaire Electronique des Synonymes (DES)

Bonsoir

Je m'associe tardivement au concert d'éloges à l'endroit des créateurs de cet outil riche et performant qu'est le dictionnaire des synonymes. Je suis fier d'avoir retrouvé mes divers pseudos dans la liste des contributeurs et d'avoir vu mes quelques propositions acceptées. je voulais dire qu'auteur compositeur amateur et animateur bénévole d'un atelier d'écriture, non seulement j'ai tiré parti de cet outil, mais que mes "élèves" l'utilisent en permanence.

En fait, il a d'abord une utilité pratique : trouver le synonyme avec le nombre de syllabes adapté lorsque celui qui vous vient à l'esprit est trop long ou trop court pour le nombre de syllabes d'un vers ; ensuite, il permet souvent d'emprunter des voies inattendues d'expression, notamment lorsqu'une proposition du dictionnaire est plus poétique, ou plus imagée, il ouvre en quelque sorte des parcours hors des sentiers battus et vous permet d'embarquer pour des contrées inconnues ou encore inexplorées de vous, bref il est source d'inspiration et de création !!

Merci encore !

JF (3 octobre 2018)

Université de Caen Normandie

Contact : crisco.webmestre@unicaen.fr

Laurette CHARDON

Présentation du Dictionnaire Électronique des Synonymes (DES)

Qu'est-ce que le DES ?
La maintenance du DES
Les statistiques d'accès
Les éléments composant le DES
L'espace sémantique
Quelques points complémentaires
Références

Les statistiques globales
Les statistiques journalières
Le profil des utilisateurs
Les mots les plus recherchés

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] [4] :



Dictionnaire Electronique des Synonymes (DES)

Chères et chers membres de Crisco,

*Je rejoins avec joie cette chaîne de remerciements et compliments !
Le français n'étant pas ma langue maternelle, je consulte votre site
maintes fois par jour. Ma thèse de doctorat (et tout ce que j'écris en
français) est pleinement et constamment nourri par votre site. Viva
Crisco !*

*Amitiés,
Alma (18 septembre 2018)*

Université de Caen Normandie

Contact : crisco.webmestre@unicaen.fr

Laurette CHARDON

Présentation du Dictionnaire Électronique des Synonymes (DES)

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] [4] :



Dictionnaire Electronique des Synonymes (DES)

Bonjour

Ce site, fruit de votre laborieux travail est tellement intégré à mon quotidien et ce depuis si longtemps qu'il ne m'est jamais venu à l'esprit de vous en féliciter !!! Un peu comme à une époque on n'aurait pas félicité Larousse pour son « dico ». Votre vidéo de remerciements m'a rappelé à l'ordre. C'est avec un immense plaisir que je me joins aux autres pour vous dire un grand merci pour cette œuvre salutaire pour de notre belle et complexe langue. Si vous l'ignorez, je me permets de vous préciser que votre site participe aussi au rayonnement du français à l'international.

Eric

Ile de la Réunion (17 septembre 2018)

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] [4] :



Dictionnaire Electronique des Synonymes (DES)

Bonjour,

Je viens de voir la vidéo de remerciements et je tenais également à vous écrire pour vous dire un grand merci. J'utilise votre dictionnaire depuis 2009-2010, et ce quotidiennement. Je suis traductrice dans la publicité et le marketing, et cet outil m'a permis, parfois par des moyens très alambiqués d'associations de termes, de trouver des solutions créatives pour mes traductions, y compris des slogans qui ont été retenus par mes clients. Encore merci, vous m'avez sauvé la mise plus d'une fois !

*Kind regards
Tina (17 septembre 2018)*

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] ^[4] :



Dictionnaire Electronique des Synonymes (DES)

Je suis journaliste rédacteur, et j'utilise presque quotidiennement votre dictionnaire dans le cadre de cette activité. Vous avez donc participé sans le savoir à l'écriture de nombreux articles parus dans Science & Vie, dans Médiacités, à la Cité des Sciences...

Cet outil simple, efficace et gratuit me manquerait s'il venait à disparaître. Je le recommande également à mes élèves de l'Ecole de Journalisme de Lille.

Pierre-Yves Bocquet
Journaliste
3 septembre 2018

Le profil des utilisateurs

Deux vidéos réalisées avec une centaine de remerciements entre juillet et octobre 2018^[3] ^[4] :



Dictionnaire Electronique des Synonymes (DES)

*Sur votre site, je vois que vous nous invitez à nous plaindre. Je ferai exactement l'inverse : pour moi, l'existence de Crisco suffit à elle seule à justifier l'existence d'Internet.
Merci d'être là.*

Guy-Patrick Sainderichin
4 septembre 2018

Les mots les plus recherchés

- Ils sont calculés depuis janvier 2019 par jour et par mois et sont accessibles sur le cloud de l'université ^[5] :
- Les plus courants : en effet, de plus, permettre, proposer, important, projet, utiliser, développer, découvrir, présenter, faire comprendre, responsable, intéressant, par exemple ...
⇒ Termes génériques liés au domaine professionnel avec quelques exceptions selon les périodes.

Qu'est-ce que le DES ?
La maintenance du DES
Les statistiques d'accès
Les éléments composant le DES
L'espace sémantique
Quelques points complémentaires
Références

Les statistiques globales
Les statistiques journalières
Le profil des utilisateurs
Les mots les plus recherchés

Sous forme de nuages de mots

Mots les plus recherchés en 2019 et 2020
Les 6000 premiers sur un total de 54.489

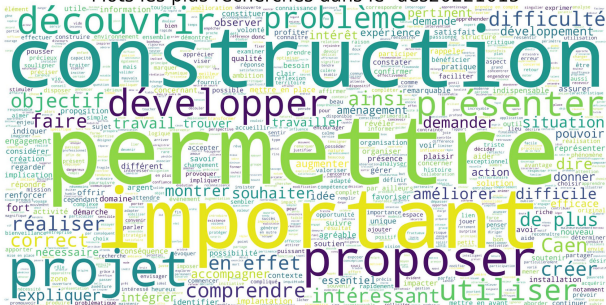


Sous forme de nuages de mots

un jour en particulier ...

Lié à un article dans La tribune le samedi 15/10/22 : Forum zéro carbone : les villes confrontées au défi de la sobriété ?

Mots les plus recherchés dans le -des20221015



Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

Les statistiques globales

Les statistiques journalières

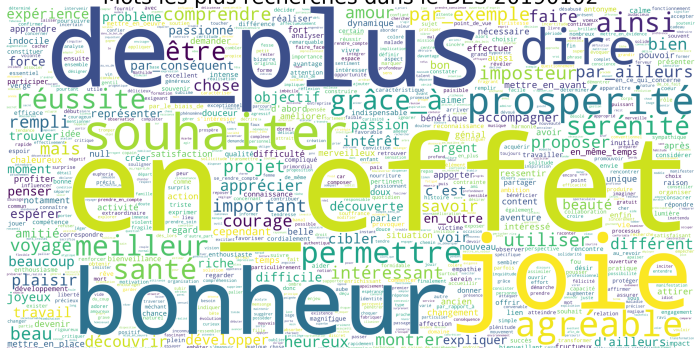
Le profil des utilisateurs

Les mots les plus recherchés

Sous forme de nuages de mots

Un autre jour en début d'année...

Mots les plus recherchés dans le DES 20190102



Qu'est-ce que le DES ?
 La maintenance du DES
 Les statistiques d'accès
 Les éléments composant le DES
 L'espace sémantique
 Quelques points complémentaires
 Références

Les statistiques globales
 Les statistiques journalières
 Le profil des utilisateurs
 Les mots les plus recherchés

Sous forme de nuages de mots

Une recherche particulière ...

Mots les plus recherchés dans le DES 20190427



Classement des premiers synonymes [1] |

- **Méthode de calcul avant 2015 :**

Soit S synonyme de V , vedette

$$\text{score} = \frac{\text{Nb de Cliques de } V \text{ contenant } S}{\text{Nb total cliques de } V} \quad (1)$$

- **Constat :**

→ Ce score fait apparaître les hyperonymes ou des mots plus polysémiques que V

→ Or les utilisateurs recherchent plutôt des mots précis

Classement des premiers synonymes [1] II

• Nouvelle méthode depuis 2015 :

- Ne plus centrer le calcul sur V
- La proximité sémantique entre S et V devrait être la même que celle entre S' et V' en prenant S' comme V et V' comme S → symétriser la formule

$$\text{score} = \frac{\text{Nb de cliques communes } S \text{ et } V (= c)}{(\text{Nb total cliques } V) + (\text{Nb total cliques de } S) - c} \quad (2)$$

- Au dénominateur, on ne compte qu'une seule fois les cliques communes

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

Le classement des premiers synonymes

Le calcul des liens probables

Les liaisons litigieuses

La féminisation

La catégorie grammaticale

Le calcul des cliques

Classement des premiers synonymes [1] III

Tableau 4. *Premiers essais du nouveau classement hiérarchique des synonymes.*

<i>rapport</i>				<i>souple</i>			
Classement ancien		Classement nouveau		Classement ancien		Classement nouveau	
<i>relation</i>	25 %	<i>relation</i>	20 %	<i>malléable</i>	20 %	<i>malléable</i>	15%
<i>liaison</i>	20 %	<i>liaison</i>	13 %	<i>maniable</i>	17 %	<i>maniable</i>	14%
<i>correspondance</i>	15 %	<i>correspondance</i>	12 %	<i>doux</i>	14 %	<i>flexible</i>	12%
<i>affinité</i>	12 %	<i>affinité</i>	10 %	<i>flexible</i>	14 %	<i>docile</i>	8%
<i>union</i>	12 %	<i>corrélacion</i>	8 %	<i>adroit</i>	12 %	<i>coulant</i>	8%
<i>accord</i>	11 %	<i>analogie</i>	8 %	<i>léger</i>	12 %	<i>soumis</i>	8%
<i>analogie</i>	9 %	<i>similitude</i>	8 %	<i>agile</i>	11 %	<i>élastique</i>	8%
<i>similitude</i>	9 %	<i>connexion</i>	7 %	<i>coulant</i>	11 %	<i>adroit</i>	7%
<i>corrélacion</i>	9 %	<i>lien</i>	6 %	<i>mou</i>	11 %	<i>agile</i>	7%
<i>lien</i>	9 %	<i>union</i>	6 %	<i>soumis</i>	11 %	<i>accommodant</i>	6%

Classement des premiers synonymes [1] IV

Tableau 4. *Premiers essais du nouveau classement hiérarchique des synonymes.*

<i>rapport</i>				<i>souple</i>			
Classement ancien		Classement nouveau		Classement ancien		Classement nouveau	
<i>relation</i>	25 %	<i>relation</i>	20 %	<i>malléable</i>	20 %	<i>malléable</i>	15%
<i>liaison</i>	20 %	<i>liaison</i>	13 %	<i>maniable</i>	17 %	<i>maniable</i>	14%
<i>correspondance</i>	15 %	<i>correspondance</i>	12 %	<i>doux</i>	14 %	<i>flexible</i>	12%
<i>affinité</i>	12 %	<i>affinité</i>	10 %	<i>flexible</i>	14 %	<i>docile</i>	8%
<i>union</i>	12 %	<i>corrélation</i>	8 %	<i>adroit</i>	12 %	<i>coulant</i>	8%
<i>accord</i>	11 %	<i>analogie</i>	8 %	<i>léger</i>	12 %	<i>soumis</i>	8%
<i>analogie</i>	9 %	<i>similitude</i>	8 %	<i>agile</i>	11 %	<i>élastique</i>	8%
<i>similitude</i>	9 %	<i>connexion</i>	7 %	<i>coulant</i>	11 %	<i>adroit</i>	7%
<i>corrélation</i>	9 %	<i>lien</i>	6 %	<i>mou</i>	11 %	<i>agile</i>	7%
<i>lien</i>	9 %	<i>union</i>	6 %	<i>soumis</i>	11 %	<i>accommodant</i>	6%

Pour *rapport*, *corrélation* (plus adapté) remonte ; *union* et *accord* reculent

Classement des premiers synonymes [1] V

Tableau 4. *Premiers essais du nouveau classement hiérarchique des synonymes.*

<i>rapport</i>				<i>souple</i>			
Classement ancien		Classement nouveau		Classement ancien		Classement nouveau	
<i>relation</i>	25 %	<i>relation</i>	20 %	<i>malléable</i>	20 %	<i>malléable</i>	15%
<i>liaison</i>	20 %	<i>liaison</i>	13 %	<i>maniabale</i>	17 %	<i>maniabale</i>	14%
<i>correspondance</i>	15 %	<i>correspondance</i>	12 %	<i>doux</i>	14 %	<i>flexible</i>	12%
<i>affinité</i>	12 %	<i>affinité</i>	10 %	<i>flexible</i>	14 %	<i>docile</i>	8%
<i>union</i>	12 %	<i>corrélation</i>	8 %	<i>adroit</i>	12 %	<i>coulant</i>	8%
<i>accord</i>	11 %	<i>analogie</i>	8 %	<i>léger</i>	12 %	<i>soumis</i>	8%
<i>analogie</i>	9 %	<i>similitude</i>	8 %	<i>agile</i>	11 %	<i>élastique</i>	8%
<i>similitude</i>	9 %	<i>connexion</i>	7 %	<i>coulant</i>	11 %	<i>adroit</i>	7%
<i>corrélation</i>	9 %	<i>lien</i>	6 %	<i>mou</i>	11 %	<i>agile</i>	7%
<i>lien</i>	9 %	<i>union</i>	6 %	<i>soumis</i>	11 %	<i>accommodant</i>	6%

Pour *souple*, *élastique* (plus adapté) remonte ; *doux* et *léger* reculent

Calcul des liens probables^[1] |

- La proximité sémantique entre 2 candidats est liée au taux de synonymes partagés^[6]
- La formule actuelle :
 - Soit E1 et E2 synonymes d'ordre 2
 - Soit n1 le nombre de synonymes+antonymes de E1
 - Soit n2 le nombre de synonymes+antonymes de E2
 - Soit c le nombre de synonymes+antonymes communs à E1 et E2
 - La 1ère formule considère le taux de synonymes+antonymes communs par rapport à l'ensemble des synonymes+antonymes des 2 entrées

$$s1 = \frac{c}{n1 + n2 - c} \quad (3)$$

Calcul des liens probables^[1] II

- La 2^{de} formule tient compte d'un phénomène souvent observé : lorsque l'une des entrées possède moins de synonymes+antonymes que l'autre mais que presque tous sont communs, le lien est souvent validé "meilleur côté"

$$s3 = \max\left(\frac{c}{n1}; \frac{c}{n2}\right) \quad (4)$$

- La 3^{ème} formule est une variante de la 2^{de}

$$s2 = \text{moy}\left(\frac{c}{n1}; \frac{c}{n2}\right) \quad (5)$$

- Introduction d'une constante pour favoriser un peu (s+), moyennement (s++) ou beaucoup (s+++)
l'aspect quantitatif
→ Total de 9 formules

Calcul des liens probables^[1] III

- Conclusion :
 - Taux de succès de plus de 80%
 - On remarque que les types de synonymes proposés diffèrent selon les formules :
 - “meilleur côté” : plus d’hyponymie c-a-d l’un des termes est précis ou rare, l’autre moins. Ex : *mystérieux / finanderie, finesse / mirobolant, ennuyeux/barjo*
 - score global pris en compte : les domaines sémantiques sont de taille voisine. Ex : *déchiffrage / décryptage, bafouillant / balbutiant, bichonné / pomponné*
 - aspect qualitatif / quantitatif : l’aspect qualitatif va favoriser les termes précis ou rares, l’aspect quantitatif, les termes plus généraux et fréquents.

Calcul des liens probables^[1] IV

- Les liens refusés relèvent quasiment tous de la co-hyponymie :
hêtraie # sapinière, paillotte # yourte

Les liaisons litigieuses

- Correspondent à des liaisons synonymiques péjoratives principalement
- Le “litige” porte sur la relation : le champ litigieux est à 1 dans la table des synonymes (Cf page 12)
- **La machine ne sait pas quel mot de la relation est litigieux**
- Total de 166 liaisons litigieuses au 31 mars 2023 (Voir transparent 60 et suivants) sur 212.313 relations au total.

La féminisation

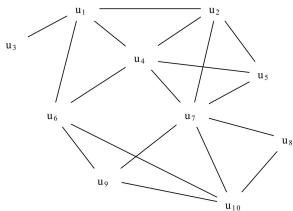
- Sur la base du guide d'aide à la féminisation des noms de métiers, titres, grades et fonctions de 1999^[7]
- 1ère méthode consistant à ajouter un suffixe :
collaborateur → collaborateur-trice ; conducteur → conducteur-trice ;
cuisinier → cuisinier-ère
⇒ méthode abandonnée car elle rend les traitements automatiques compliqués
- Nouvelle méthode en ajoutant les formes féminines comme variantes.
collaboratrice → collaborateur ; conductrice → conducteur ; cuisinière → cuisinier
⇒ en cours de changement

La catégorie grammaticale

- Le projet a duré 2 ans (janvier 2021 à novembre 2022)
- A partir des données de l'ATILF (*Analyse et Traitement Informatique de la Langue Française, UMR7118*)
- Traitement informatique en langage Python et la librairie Spacy par recherche de schémas (*-ais, -euse, -ale, -ande, . . .*)
- Puis traitement manuel avec insertion automatique
- Traitement informatique de vérification (par exemple si un verbe synonyme d'un "non verbe" → à vérifier)
- Voir le rapport [halshs-03956407](#) et le [dépôt git correspondant](#)

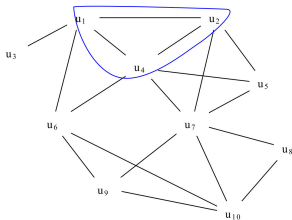
Calcul des cliques^[8]

- Une clique (sous-graphe complet maximal) est un ensemble le plus grand possible de sommets du graphe tous reliés deux à deux, « le plus grand possible »
- En d'autres termes, une clique est un ensemble d'unités lexicales qui ont la propriété d'être toutes synonymes les unes des autres, aucune autre unité ne pouvant être rajoutée à l'ensemble sans que l'on perde cette propriété.
- Depuis juin 2019, le calcul des cliques est réalisé avec une librairie^[9] python implémentant l'algorithme de Bron-Kerbosch



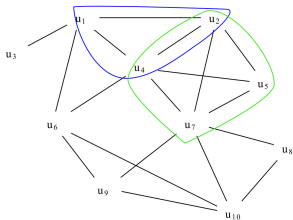
Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



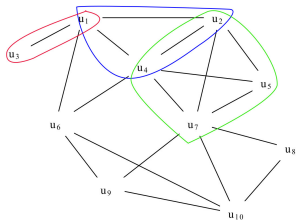
Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



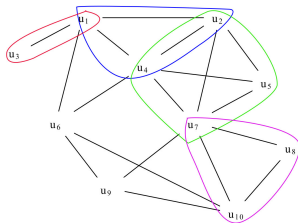
Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



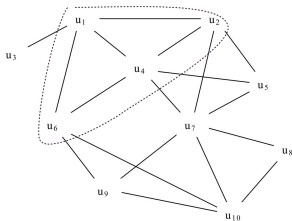
Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



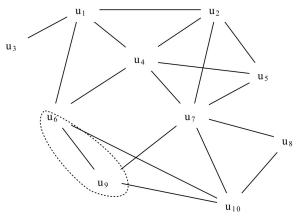
Le principe :

- Les ensembles $\{u_1, u_2, u_4\}$, $\{u_2, u_4, u_5, u_7\}$, $\{u_1, u_3\}$ et $\{u_7, u_8, u_{10}\}$ sont des exemples de cliques, alors que l'ensemble $\{u_1, u_2, u_4, u_6\}$ n'est pas une clique, puisque u_2 et u_6 ne sont pas synonymes.
- De même $\{u_6, u_9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de u_{10} et c'est donc $\{u_6, u_9, u_{10}\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.



Le principe :

- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

Le classement des premiers synonymes

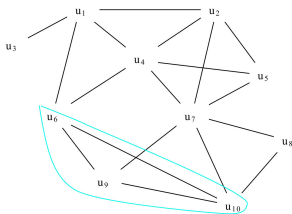
Le calcul des liens probables

Les liaisons litigieuses

La féminisation

La catégorie grammaticale

Le calcul des cliques



Le principe :

- Les ensembles $\{u_1, u_2, u_4\}$, $\{u_2, u_4, u_5, u_7\}$, $\{u_1, u_3\}$ et $\{u_7, u_8, u_{10}\}$ sont des exemples de cliques, alors que l'ensemble $\{u_1, u_2, u_4, u_6\}$ n'est pas une clique, puisque u_2 et u_6 ne sont pas synonymes.
- De même $\{u_6, u_9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de u_{10} et c'est donc $\{u_6, u_9, u_{10}\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

Le classement des premiers synonymes

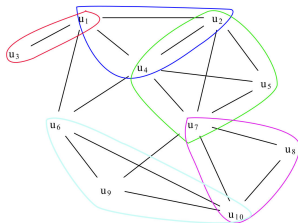
Le calcul des liens probables

Les liaisons litigieuses

La féminisation

La catégorie grammaticale

Le calcul des cliques



Le principe :

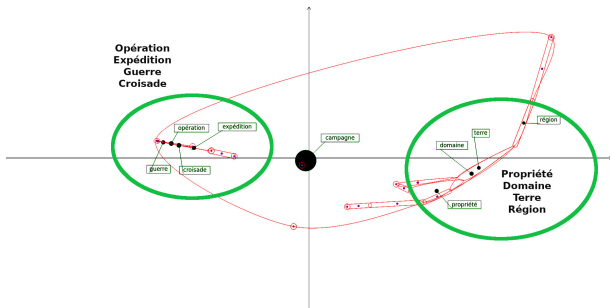
- Les ensembles $\{u1, u2, u4\}$, $\{u2, u4, u5, u7\}$, $\{u1, u3\}$ et $\{u7, u8, u10\}$ sont des exemples de cliques, alors que l'ensemble $\{u1, u2, u4, u6\}$ n'est pas une clique, puisque $u2$ et $u6$ ne sont pas synonymes.
- De même $\{u6, u9\}$ ne forme pas une clique parce qu'elle n'est pas maximale : ces deux unités sont toutes deux synonymes de $u10$ et c'est donc $\{u6, u9, u10\}$ qui constitue une clique.
- Comme on peut le constater, deux cliques peuvent être disjointes, ou avoir une ou plusieurs unités en commun.

Le principe [10]

- L'espace sémantique est le résultat d'une méthode factorielle : l'Analyse Factorielle des Correspondances (AFC).
- L'AFC permet de résumer de manière synthétique l'essentiel d'un grand nombre de données, impossible à voir à l'oeil nu. Pour cela, on extrait quelques facteurs (de 2 à 4) les plus représentatifs.
- **L'espace sémantique d'un mot va donc permettre d'en dégager les principaux sens selon 2 ou 3 facteurs (en 2D ou 3D).**

Exemple 1 avec CAMPAGNE

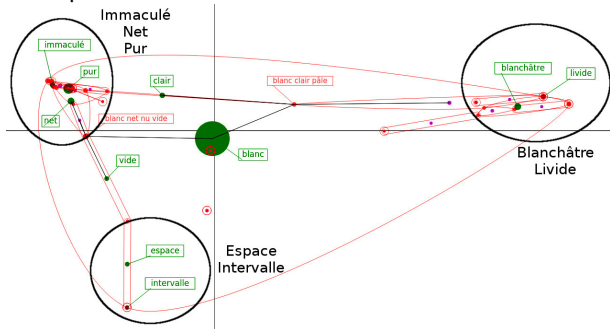
Deux composantes se dégagent : “stratégique” (opération, expédition, guerre, croisade) et “topographique” (propriété, domaine, terre, région) :



Accès en ligne : <https://crisco2.unicaen.fr/espsem/campagne>

Exemple 2 avec BLANC

Trois composantes se dégagent : sur l'axe X en tant qu'adjectif s'oppose le blanc-propreté et le blanc-couleur relié par la clique {blanc,clair,pâle} ; sur l'axe Y se dégage le blanc-typographie en tant que substantif :



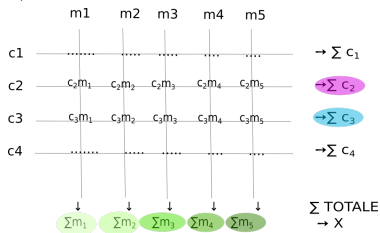
Accès en ligne : <https://crisco2.unicaen.fr/espsem/blanc>

Pour aller un peu plus loin ...^[10]

- Chaque facteur (= axe du graphique) va résumer un % de l'information : le premier en apporte le plus, le second un peu moins, etc ...
- Sur cet axe, vont s'opposer deux facettes du facteur représenté
- Selon le graphique 2D ou 3D, les deux ou trois facteurs vont se "combiner" et donner de deux à six regroupements de sens.
- L'AFC se base sur le calcul de la décomposition en valeurs singulières (cf [wikipedia](#)).

Pour obtenir l'espace sémantique c'est-à-dire les coordonnées sur 3 dimensions de chaque synonyme et de chaque clique :

Matrice d'adjacence : c → cliques; m → mots
 $C_i M_j = 1$ si le mot appartient à la clique, 0 sinon



Matrice (M) de p lignes et de n colonnes

Chaque coefficient est divisé par :

- 1 la somme totale X
- 2 la racine carrée de la somme des lignes (= nombre de mots qui composent la clique)

$$\sqrt{\sum_{i=1}^n C_i} \quad (6)$$

- ③ la racine carrée de la somme des colonnes (= nombre de cliques dans lesquelles ce mot apparaît)

$$\sqrt{\sum_{j=1}^p M_j} \quad (7)$$

- ④ On obtient donc une matrice Q issue de M divisée par la somme totale et les racines carrées des sommes marginales
- ⑤ Si on considère dans M les lignes comme des vecteurs avec les coordonnées suivantes :

$$\vec{C}_i = (c_i m_j)_{j=1}^p \quad (8)$$

alors dans Q sera composée des vecteurs :

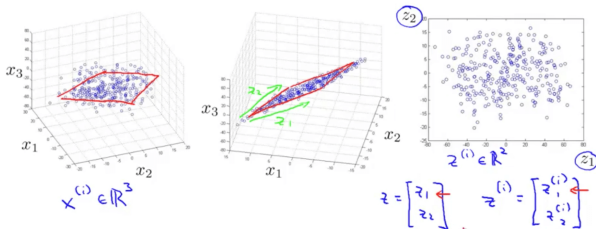
$$\vec{C}_i = (x_{ij})_{j=1}^p \quad (9)$$

avec

$$x_{ij} = \frac{c_i m_j}{X \cdot \sqrt{\sum_{i=1}^n C_i} \sqrt{\sum_{j=1}^p M_j}} \quad (10)$$

- 6 Le principe ensuite est de trouver le meilleur plan en 2D (ou espace 3D) sur lequel seront projetés de façon orthogonale les points du vecteur \vec{C}_i tout en minimisant la perte d'information comme schématisé par cet exemple [11] [12] :

Reduce data from 3D to 2D



- ⑦ Depuis février 2020, la librairie fanalysis^[13] en python est en cours d'étude pour simplifier les calculs ci-dessus

Qu'est-ce que le DES ?

La maintenance du DES

Les statistiques d'accès

Les éléments composant le DES

L'espace sémantique

Quelques points complémentaires

Références

DES et la science ouverte

Le graphe d'adjacence

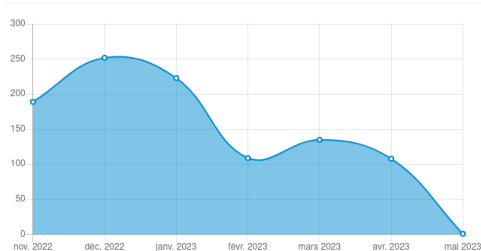
La liaison entre deux variables qualitatives - khi-deux

Visualisation de la distance du khi-deux dans l'espace sémantique

Autres ressources

Les données du DES
sont sur la plate-forme
ORTOLANG

Statistiques de consultation (1017 vues au total)



Statistiques de téléchargements (depuis novembre 2016 📄)

35 téléchargements complets
12 fichiers ou dossiers téléchargés

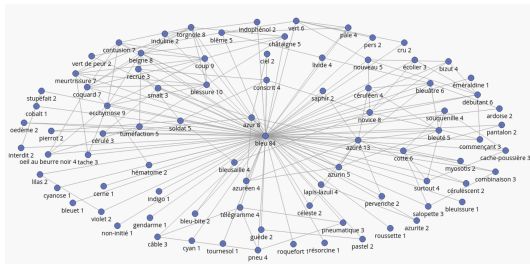
Le Plan de Gestion de Données : en cours de rédaction
Un **dépot git** avec les données des publications réalisées.

Le graphe d'adjacence

La liaison entre deux variables qualitatives - khi-deux
Visualisation de la distance du khi-deux dans l'espace sémantique
Autres ressources

Matrice d'adjacence de BLEU(84 synonymes) :
représentation graphique des liens synonymiques enregistrés dans la base de données sans aucun calcul

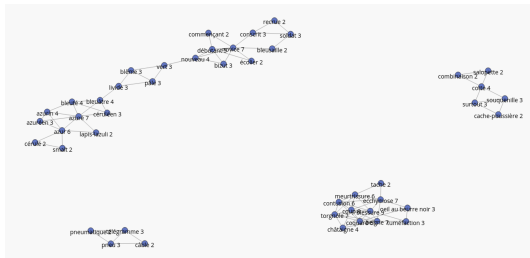
Graphe de BLEU



Le graphe d'adjacence

La liaison entre deux variables qualitatives - khi-deux
Visualisation de la distance du khi-deux dans l'espace sémantique
Autres ressources

Matrice d'adjacence de BLEU avec une présentation "améliorée" :
tous les liens avec la vedette ont été supprimés ainsi que les synonymes isolés



Le graphe d'adjacence

La liaison entre deux variables qualitatives - khi-deux

Visualisation de la distance du khi-deux dans l'espace sémantique

Autres ressources

Graphes d'adjacence du DES - CRISCO - Université de Caen

(fichiers d'avril 2020)

Le graphe d'adjacence d'une vedette V donnée est une représentation graphique de tous les liens synonymiques de V et également des liens entre les synonymes. Chaque sommet apparaît au moins une fois et qu'il représente suivi du nombre de liens que ce sommet possède avec les autres sommets.

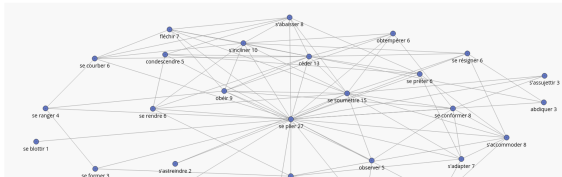
Vedette :
Vous avez saisi : "se plier"

Enlever l'affichage de la vedette

La suppression de l'affichage de la vedette V va supprimer tous les liens de V avec ses synonymes ; seuls restent les liens entre les synonymes eux-mêmes. Les synonymes qui n'étaient rattachés à aucune vedette vont apparaître isolés avec le chiffre 0 c'est-à-dire 0 liens avec les autres synonymes.

La suppression de l'affichage de V va à la fois rendre le graphe plus lisible si V a beaucoup de synonymes et permettre en même temps de repérer les synonymes isolés qui pourraient être mieux étudiés : soit en leur ajoutant de nouveaux liens soit en les supprimant (mot vieilli, peu utilisé, incorrect ...).

Graphe de SE PLIER



Accès à l'interface en ligne :
crisco3.unicaen.fr/desgraphe
- choix suppression de la vedette
- choix d'affichage des synonymes avec un nombre de liens supérieur à une limite choisie

Avec des données **quantitatives**, on va parler de **moyenne** et **d'écart type**. Avec des données **qualitatives**, on va parler de **fréquence** et **d'indices de liaison** ^[14]

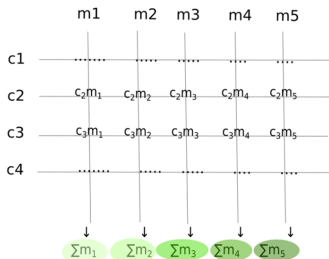
L'équation du khi-deux pour définir la distance entre deux cliques (S. Ploux et B. Victorri ^[8])

$$d^2(C_k, C_l) = \sum_{i=1}^n \frac{x}{x_{.i}} \left(\frac{x_{ki}}{x_k} - \frac{x_{li}}{x_l} \right)^2 \quad (11)$$

avec

$$x_{.i} = \sum_{j=1}^p x_{ji}, \quad x_k = \sum_{i=1}^n x_{ki}, \quad x = \sum_{i=1}^n \sum_{j=1}^p x_{ij} \quad (12)$$

Matrice d'adjacence : c → cliques; m → mots
 $C_i M_j = 1$ si le mot appartient à la clique, 0 sinon



→ ΣC_1

→ ΣC_2

→ ΣC_3

→ ΣC_4

Σ TOTALE
 → X

Le calcul du Phi entre deux cliques, C_2 et C_3 =

$$\left[\frac{X}{\Sigma m_1} \times \left(\frac{c_2 m_1}{\Sigma C_2} - \frac{c_3 m_1}{\Sigma C_3} \right)^2 \right]^+ +$$

$$\left[\frac{X}{\Sigma m_2} \times \left(\frac{c_2 m_2}{\Sigma C_2} - \frac{c_3 m_2}{\Sigma C_3} \right)^2 \right]^+ +$$

$$\left[\frac{X}{\Sigma m_3} \times \left(\frac{c_2 m_3}{\Sigma C_2} - \frac{c_3 m_3}{\Sigma C_3} \right)^2 \right]^+ +$$

$$\left[\frac{X}{\Sigma m_4} \times \left(\frac{c_2 m_4}{\Sigma C_2} - \frac{c_3 m_4}{\Sigma C_3} \right)^2 \right]^+ +$$

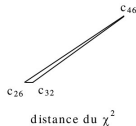
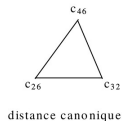
$$\left[\frac{X}{\Sigma m_5} \times \left(\frac{c_2 m_5}{\Sigma C_2} - \frac{c_3 m_5}{\Sigma C_3} \right)^2 \right]^+ +$$

Déductions sur la métrique du khi-deux :

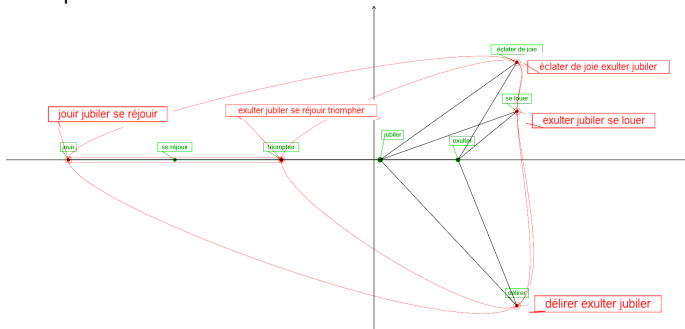
- Chaque synonyme intervient dans le calcul avec un “poids” plus faible s’il est présent dans un grand nombre de cliques : les synonymes les moins spécifiques jouent un rôle moins important dans la discrimination des sens de l’unité.
- La distance entre 2 cliques nombreuses qui possèdent beaucoup de synonymes “spécifiques” en commun sera plus faible que celle entre 2 cliques peu nombreuses qui diffèrent par des synonymes spécifiques.
- La métrique canonique ne permettra pas de mettre en avant ces 2 critères (voir exemple suivant)

Métrique canonique (exemple avec interdire) :

- Pour la métrique canonique, deux couples de cliques qui diffèrent par un même nombre de synonymes sont analogues : la distance entre le premier couple de cliques est égale à la distance entre le second couple de cliques.
- 3 cliques : c₂₆ : {défendre, inhiber, prohiber, proscrire} ; c₃₂ : {empêcher, inhiber, paralyser, supprimer} ; c₄₆ : {paralyser, pétrifier}
- Les 2 premières caractérisent des nuances du sens usuel de interdire ; la dernière fait référence à un sens différent : "Devant tant de mauvaise foi, Jean est resté tout interdit".
- Distance canonique : représentation aberrante avec c₃₂ qui est un peu plus proche de c₄₆ (différent par 4 autres synonymes) que de c₂₆ (différent par 6 autres synonymes) ! En revanche, la distance du Phi rectifie de façon très nette cette aberration



Exemple avec la vedette JUBILER

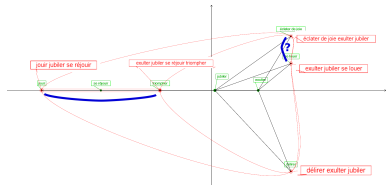


Voici son espace sémantique avec les cliques

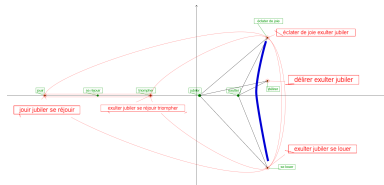
La distance du khi-deux entre les cliques :

cliques	c1	c2	c3	c4	c5
c1 :exulter, jubiler, se réjouir, triompher	0				
c2 :jouir, jubiler, se réjouir	1.76	0			
c3 :éclater de joie, exulter, jubiler	1.82	2.21	0		
c4 :exulter, jubiler, se louer	1.82	2.21	0		
c5 :délirer, exulter, jubiler	1.82	2.21	1.88	1.88	0

Selon le tableau précédent, la distance la plus courte est celle entre c1 et c2.



Sur la vue axe 1-2, cela n'apparaît pas clairement



Sur la vue axe 1-3, c'est évident

⇒ Le calcul du khi-deux tient compte des 6 premières dimensions, ce qui n'apparaît pas nécessairement sur la vue 1-2 par défaut.

- Page web sur le site du Crisco :
<http://crisco.unicaen.fr/dictionnaire-electronique-des-synonymes/>,
notamment :
- Vidéo de présentation
- Les actualités
- Les visualisations 3D
- Les publications scientifiques
- Les statistiques d'accès

- [1] Michel MOREL et Jacques FRANÇOIS. “Le Dictionnaire Electronique des Synonymes du CRISCO : un outil de plus en plus interactif”. In : Revue française de linguistique appliquée Vol. XX.39 (2015), p. 9-28. URL : <https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2015-1-page-9.htm#>.
- [2] CRISCO. Présentation du DES : le bilan des mises à jour. URL : <http://crisco.unicaen.fr/dictionnaire-electronique-des-synonymes/presentation-du-des/>.

- [3] CRISCO. Première vidéo de remerciements des utilisateurs. 2018. URL : <http://crisco.unicaen.fr/dictionnaire-electronique-des-synonymes/actualites-des/remerciements-des-utilisateurs-du-des-920615.kjsp?RH=1530619460865>.
- [4] CRISCO. Seconde vidéo de remerciements des utilisateurs. 2018. URL : <http://crisco.unicaen.fr/dictionnaire-electronique-des-synonymes/actualites-des/remerciements-des-utilisateurs-du-des-2nde-video--947671.kjsp?RH=1530619460865>.

- [5] CRISCO.
Mots les plus recherchés dans le DES et l'espace sémantique.
2019. URL : <https://unicloud.unicaen.fr/index.php/s/no7esiR7ZRHEeBo>.
- [6] Jean-Luc MANGUIN. "Transitivité partielle de la synonymie : applications aux dictionnaires de synonymes". In : CORELA (2004). URL : <https://journals.openedition.org/corela/611>.
- [7] CNRS- INLF.
Guide d'aide à la féminisation des noms de métiers, titres, grades et
URL : <https://www.vie-publique.fr/sites/default/files/rapport/pdf/994001174.pdf>.

- [8] Sabine PLOUX et Bernard VICTORRI. “Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes”. In : Traitement Automatique des Langues 39 (1998), p. 161-182. URL : <https://halshs.archives-ouvertes.fr/halshs-00009433>.
- [9] Abhinav SARKAR.
Finds all maximal cliques in a graph using the Bron-Kerbosch algorithm
URL : <https://gist.github.com/abhin4v/8304062>.
- [10] Céline POUDAT et Frédéric LANDRAGIN.
Explorer un corpus textuel. Champs linguistiques. De Boeck, 2017, p. 240. URL : <https://hal.archives-ouvertes.fr/hal-01373918>.

- [11] Ritchie NG.
Illustration d'une réduction de dimension de 3D en 2D.
URL : <https://www.ritchieng.com/machine-learning-dimensionality-reduction/>.
- [12] WIKIPEDIA. Décomposition en valeurs singulières. URL :
https://fr.wikipedia.org/wiki/D%C3%A9composition_en_valeurs_singuli%C3%A8res.
- [13] Olivier GARCIA.
Python module for Factorial Analysis : Simple and Multiple Correspondence Analysis
URL :
<https://github.com/OlivierGarciaDev/fanalysis>.

- [14] WIKI**STAT**. Statistique descriptive bidimensionnelle. URL :
<https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-l-des-bi.pdf>.

Liaisons litigieuses I

abâtardir , efféminer	1
affaiblir , efféminer	2
africain , nègre	3
agent , argousin	4
agent , bourre	5
agent , cogne	6
alanguir , efféminer	7
algèbre , chinois	8
algèbre , hébreu	9
amante , gonzesse	10
amante , grognasse	11
amollir , efféminer	12
anar , extrémiste	13
anarchie , chaos	14

Liaisons litigieuses II

anarchie , confusion	15
anarchie , désordre	16
anarchie , gâchis	17
anarchie , pagaille	18
anarchie , trouble	19
anarchie , émeute	20
anarchiste , extrémiste	21
arbi , bicot	22
argousin , flic	23
argousin , gardien	24
argousin , surveillant	25
argus , surveillant	26
avare , grigou	27
bamboula , mélanoderme	28
bicot , bougnoul	29

Liaisons litigieuses III

bicot , raton	30
boniche , bonne	31
boniche , domestique	32
boniche , employé de maison	33
boniche , gens de maison	34
boniche , servante	35
bougnoul , raton	36
bougre , sodomite	37
byzantin , chinois	38
cafard , flic	39
caisse , voiture	40
calbut , caleçon	41
candidater , postuler	42
chair à canon , soldat	43
chinois , difficile	44

Liaisons litigieuses IV

chinois , hébreu	45
chinois , jaune	46
chinois , pointilleux	47
cogne , flic	48
cogne , gendarme	49
commère , concierge	50
coureuse , fille	51
courtisane , fille	52
créole , nègre	53
donzelle , femme	54
débiliter , efféminer	55
efféminer , ramollir	56
efféminer , émasculer	57
exagération , gasconnade	58
exagérer , gasconner	59

Liaisons litigieuses V

fanfaronnade , gasconnade	60
femme , gonzesse	61
femme , nana	62
femme publique , fille	63
fillasse , fille	64
fille , furie	65
fille , gonzesse	66
fille , gouge	67
fille , gourgandine	68
flic , flicard	69
flic , gendarme	70
flic , poulet	71
flicaille , police	72
folle , gonzesse	73
franchouillard , français	74

Liaisons litigieuses VI

français , gaulois	75
français , gouère	76
gasconnade , hâblerie	77
gasconnade , rodomontade	78
gasconnade , vanterie	79
gasconner , hâbler	80
gay , pédale	81
gay , pédé	82
gendarme , mange—mil	83
gendarme , pandore	84
gendarme , virago	85
gonzesse , lope	86
gonzesse , lopette	87
gonzesse , mignon	88
gonzesse , pédale	89

Liaisons litigieuses VII

gonzesse , pédoque	90
gonzesse , pédé	91
gonzesse , pédé raste	92
gonzesse , salope	93
gonzesse , sodomite	94
gonzesse , tante	95
gonzesse , tantouse	96
gonzesse , tapette	97
gouge , servante	98
gougnotte , lesbienne	99
gouine , lesbienne	100
gousse , lesbienne	101
grigou , radin	102
homme , zig	103
homme , zigoto	104

Liaisons litigieuses VIII

homme, zouave	105
homme, zèbre	106
homophile, inversé	107
homosexuel, inversé	108
homosexuel, mignon	109
homosexuelle, tribade	110
judaïque, étroit	111
lesbienne, tribade	112
lope, lopette	113
lope, pédale	114
lope, pédoque	115
lope, pédé	116
lope, pédéraste	117
lope, sodomite	118
lope, tante	119

Liaisons litigieuses IX

lope , tantouse	120
lope , tapette	121
lopette , pédale	122
lopette , pédoque	123
lopette , pédé	124
lopette , pédéraste	125
lopette , sodomite	126
lopette , tante	127
lopette , tantouse	128
lopette , tapette	129
manoeuvre , nègre	130
midget , minus	131
moricaud , mulâtre	132
moricaud , noiraud	133
moricaud , négriillon	134

Liaisons litigieuses X

moukère , prostitué	135
mélanoderme , nègre	136
noir , nègre	137
nègre , prête-plume	138
né grille , pygmée	139
petit , pygmée	140
prostitué , traînée	141
pédale , pédé	142
pédale , sodomite	143
pédale , tapette	144
pédale , uraniste	145
pédoque , pédé	146
pédoque , sodomite	147
pédoque , tapette	148
pédoque , uraniste	149

Liaisons litigieuses XI

pédé ,pédéraste	150
pédé ,sodomite	151
pédé ,tante	152
pédé ,tantouse	153
pédé ,tapette	154
pédé ,uraniste	155
pédéraste , sodomite	156
pédéraste , tapette	157
sodomite , tante	158
sodomite , tantouse	159
sodomite , tapette	160
tante , tapette	161
tantouse , tapette	162
tapette , travesti	163
tapette , uraniste	164

Liaisons litigieuses XII

travelo , travesti	165
travelo , uraniste	166

liaisonslitigieuses_2023-5-2.csv