Module de formation doctorale en informatique textuelle

Séance 4 - De la lexicométrie au traitement automatique des langues (TAL)

16/02/2019 – Créteil

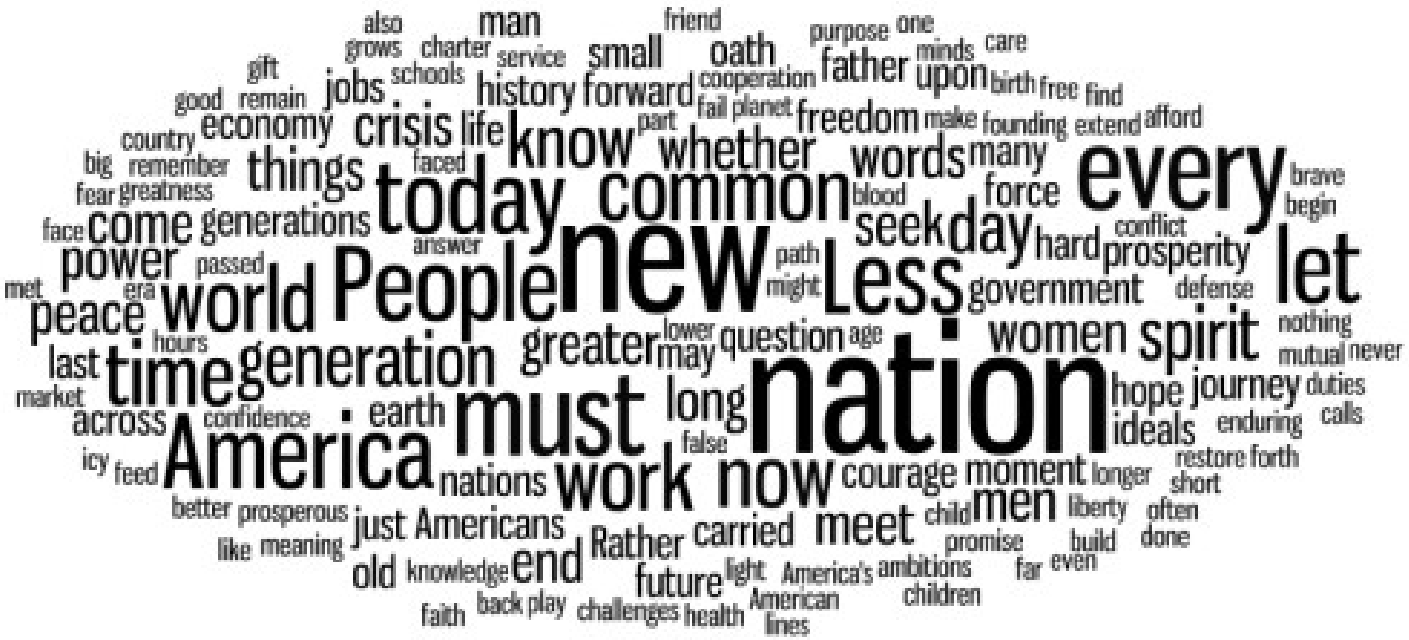
TreeCloud pour la visualisation et l'analyse de données textuelles

Philippe Gambette

LIGM
Université Paris-Est
Marne-la-Vallée

Le « nuage arboré », une information double

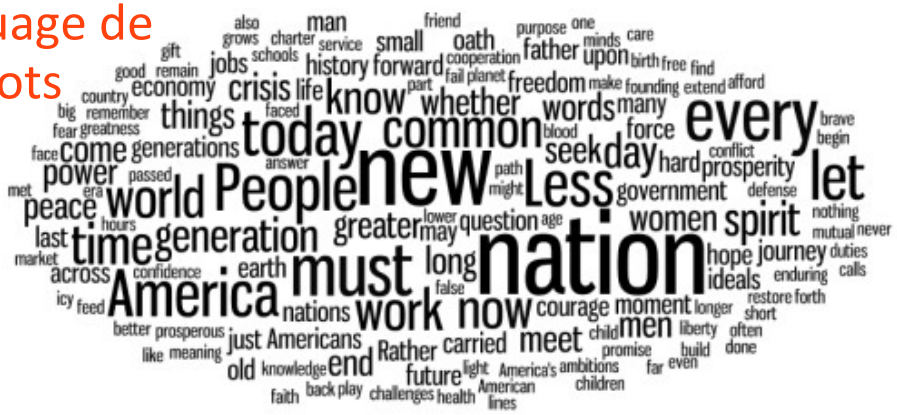
nuage de mots



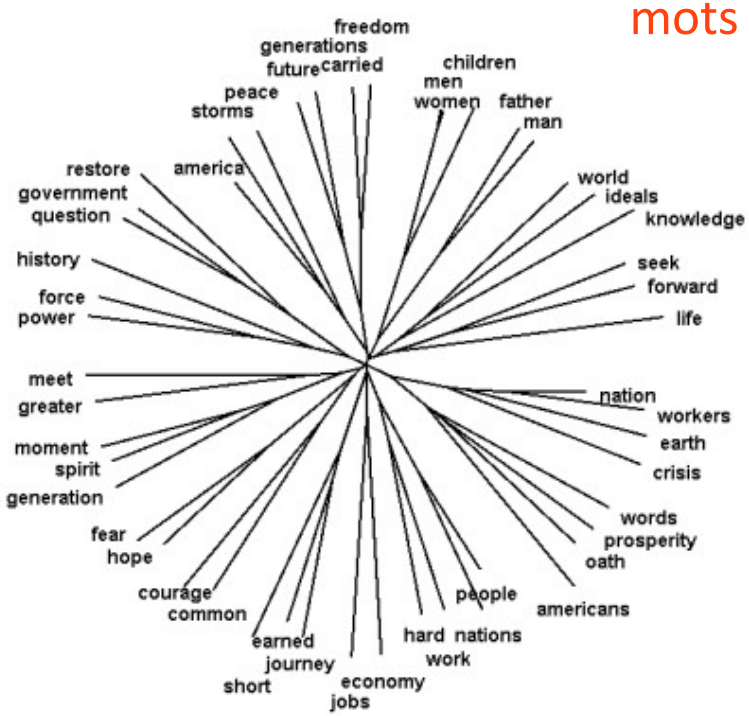
Discours inaugural de Barack Obama en 2008, Wordle

Le « nuage arboré », une information double

nuage de mots



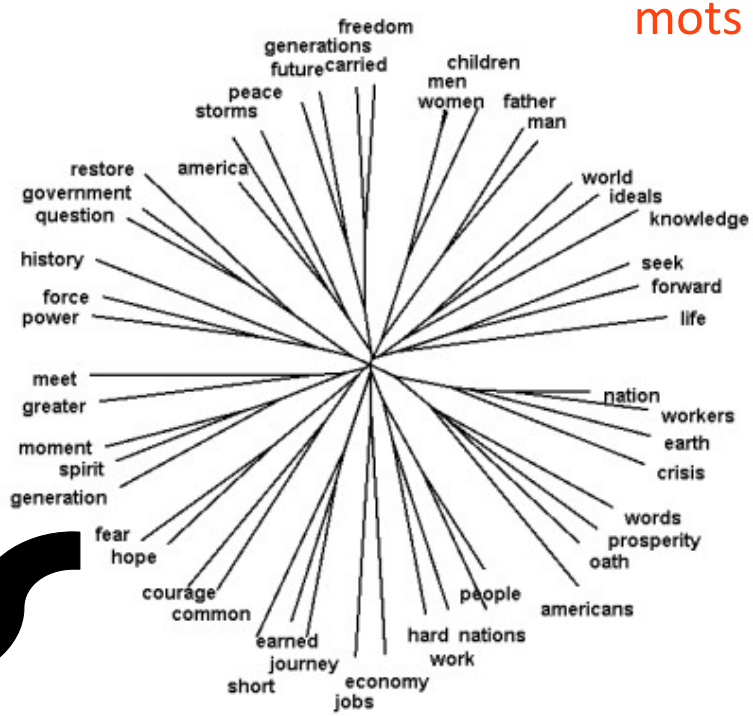
arbre de mots

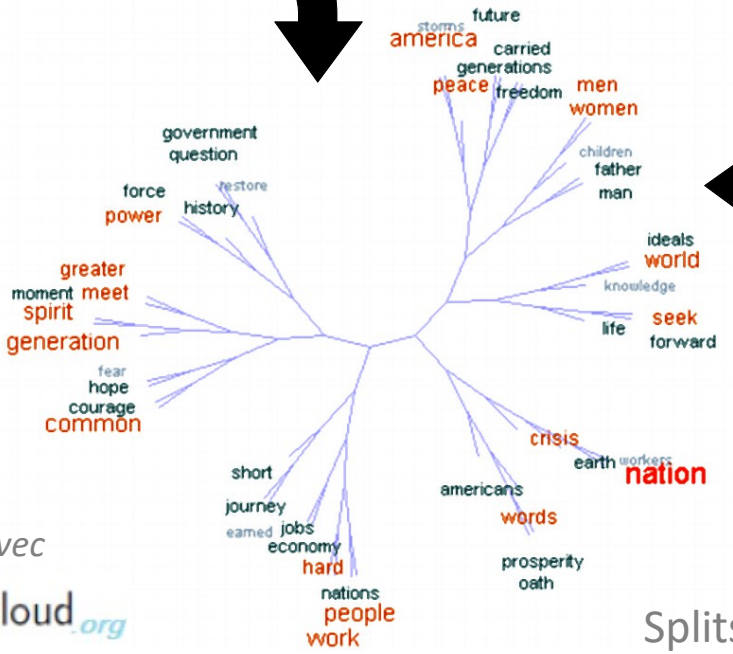


Le « nuage arboré », une information double

nuage de mots

arbre de mots





Discours inaugural de Barack Obama

construit avec

TreeCloud.org

SplitsTree4

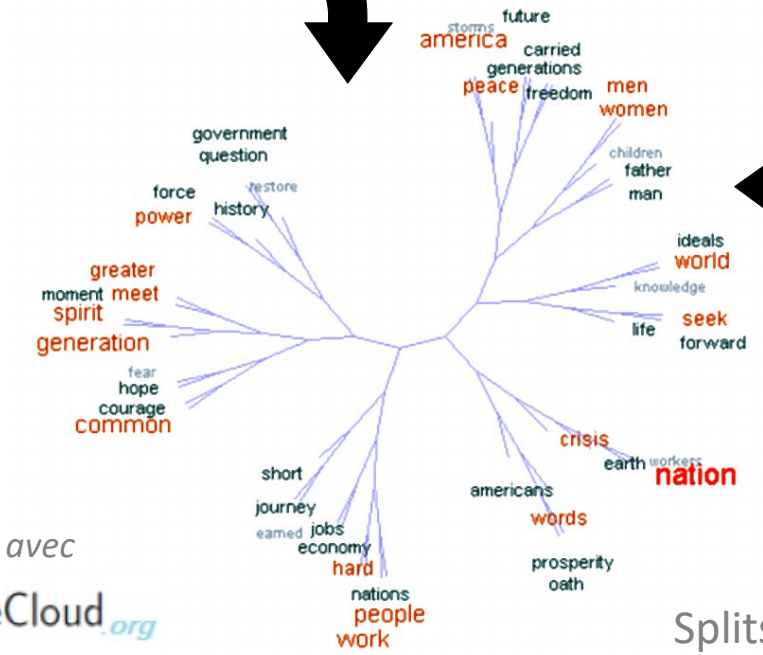
SplitsTree : Huson & Bryant, *Bioinformatics*, 2006

TreeCloud : Gambette & Véronis, *IFCS'09*

Le « nuage arboré », une information double

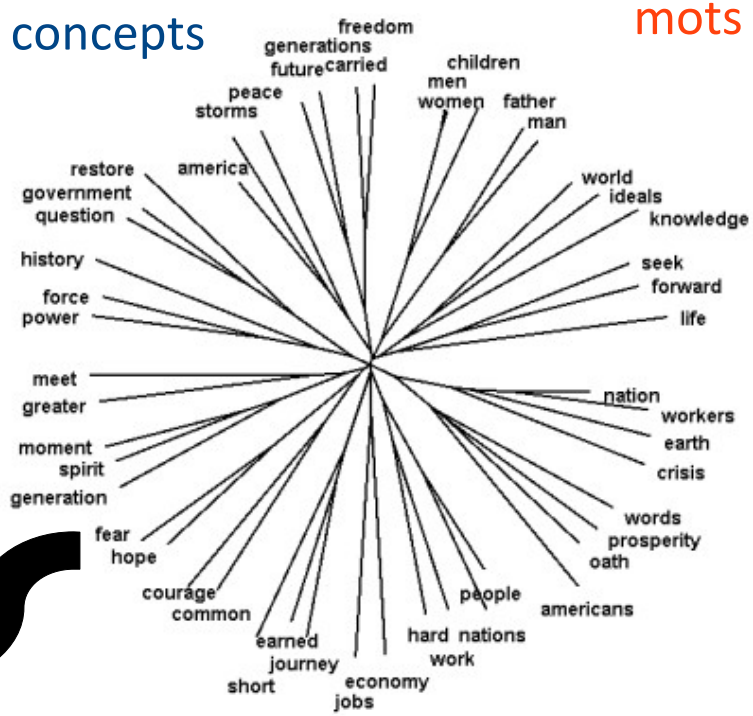
nuage de mots

hiérarchie des mots



hiérarchie des concepts

arbre de mots



Discours inaugural de Barack Obama

construit avec

TreeCloud.org

SplitsTree4

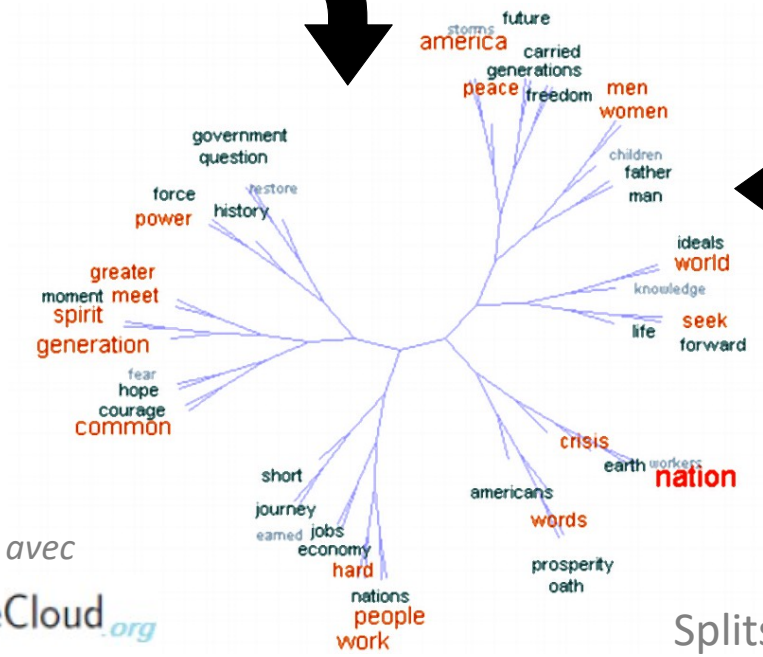
SplitsTree : Huson & Bryant, *Bioinformatics*, 2006

TreeCloud : Gambette & Véronis, *IFCS'09*

Le « nuage arboré », une information double

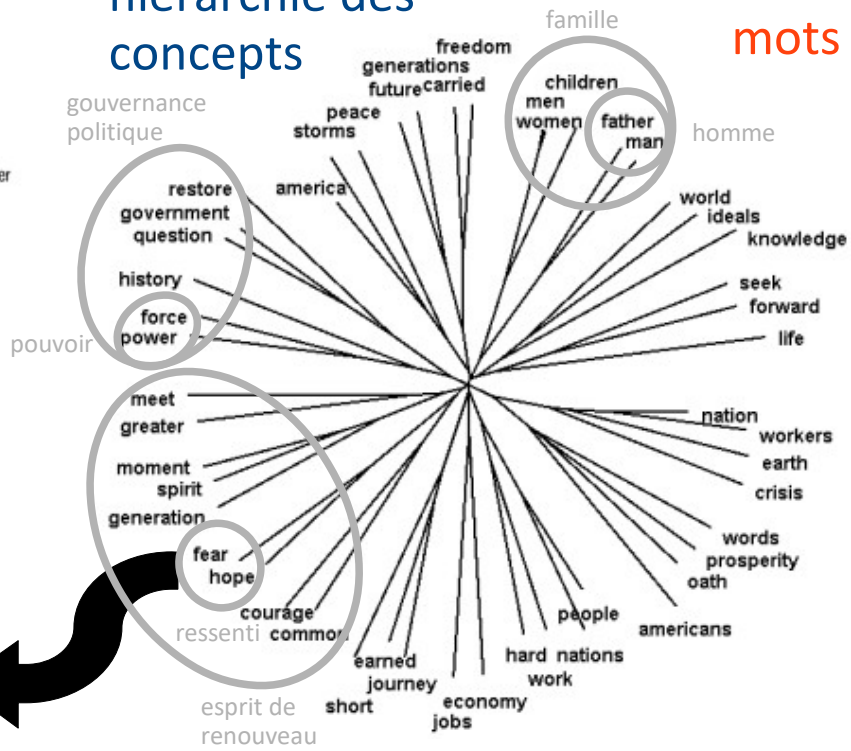
nuage de mots

hiérarchie des mots



hiérarchie des concepts

arbre de mots



Discours inaugural de Barack Obama

construit avec

TreeCloud.org

SplitsTree4

SplitsTree : Huson & Bryant, *Bioinformatics*, 2006

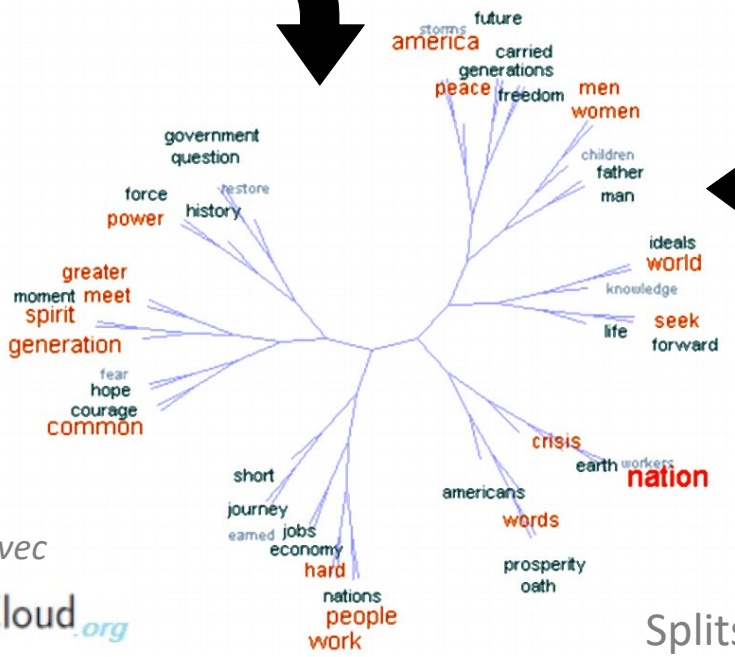
TreeCloud : Gambette & Véronis, *IFCS'09*

Le « nuage arboré », une information double

nuage de mots

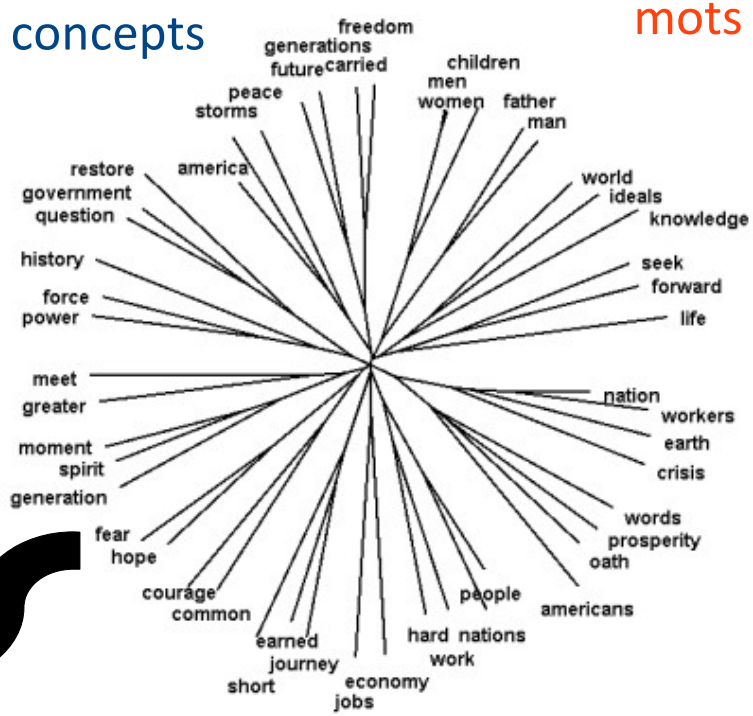
hiérarchie des mots

occurrences



hiérarchie des concepts

arbre de mots



cooccurrences

Discours inaugural de Barack Obama

construit avec

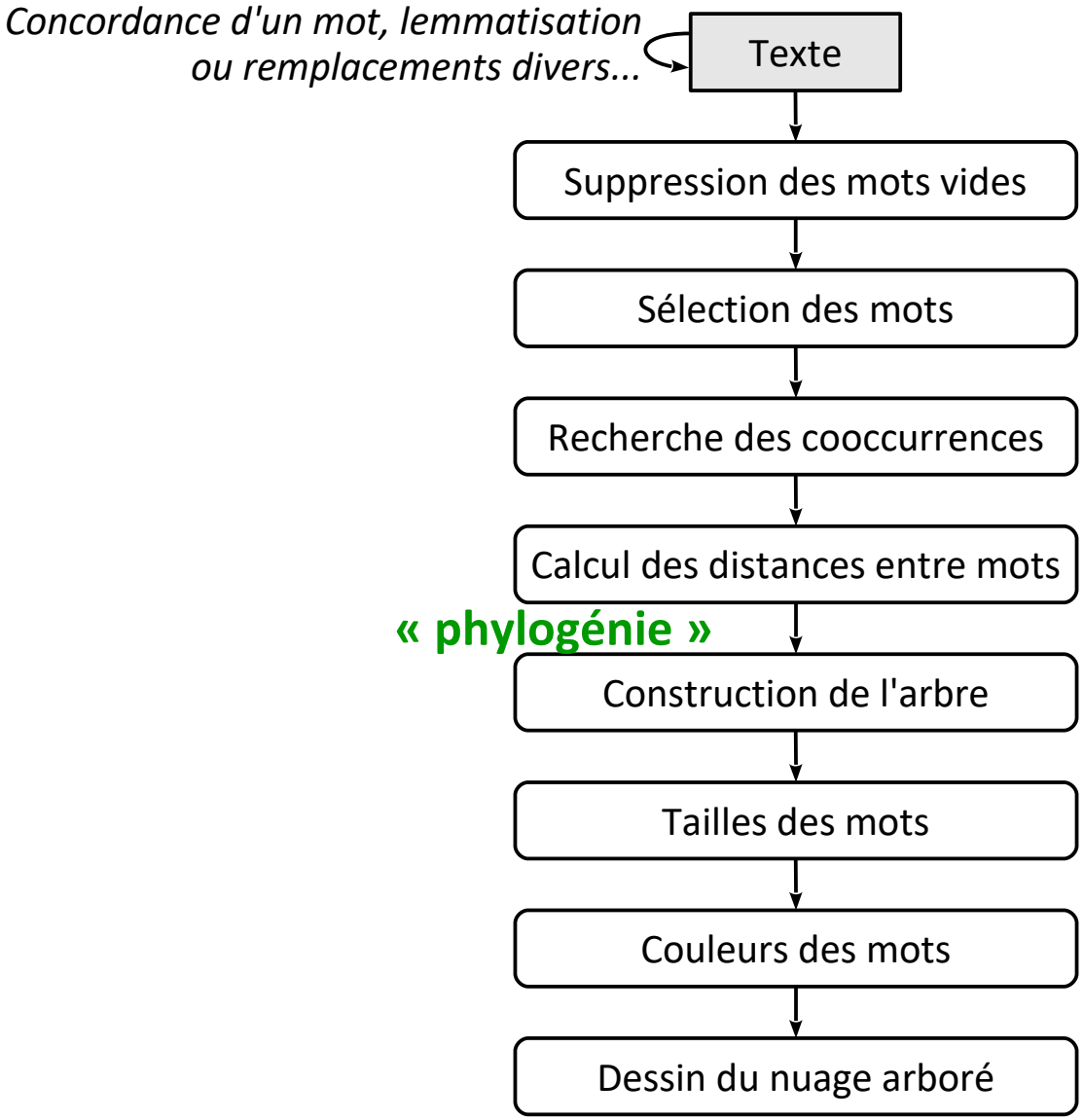
TreeCloud.org

SplitsTree4

SplitsTree : Huson & Bryant, *Bioinformatics*, 2006

TreeCloud : Gambette & Véronis, *IFCS'09*

Processus de construction



Proposé dans la version téléchargeable de TreeCloud

antidico anglais, français

n mots les plus fréquents, mots apparaissant plus de n fois, ou liste personnalisée

Fenêtre de cooccurrence paramétrée par taille et pas de glissement, ou caractère séparateur

12 formules de distance de cooccurrence

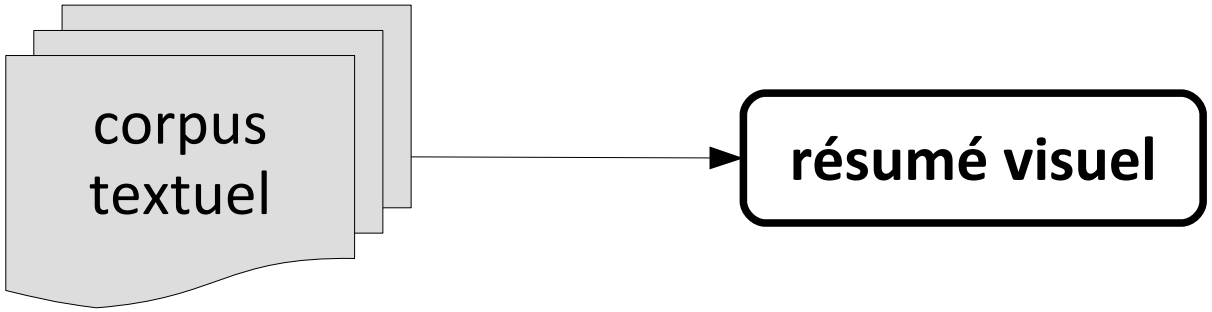
Appel transparent au logiciel SplitsTree

Fréquences ou valeurs personnalisées

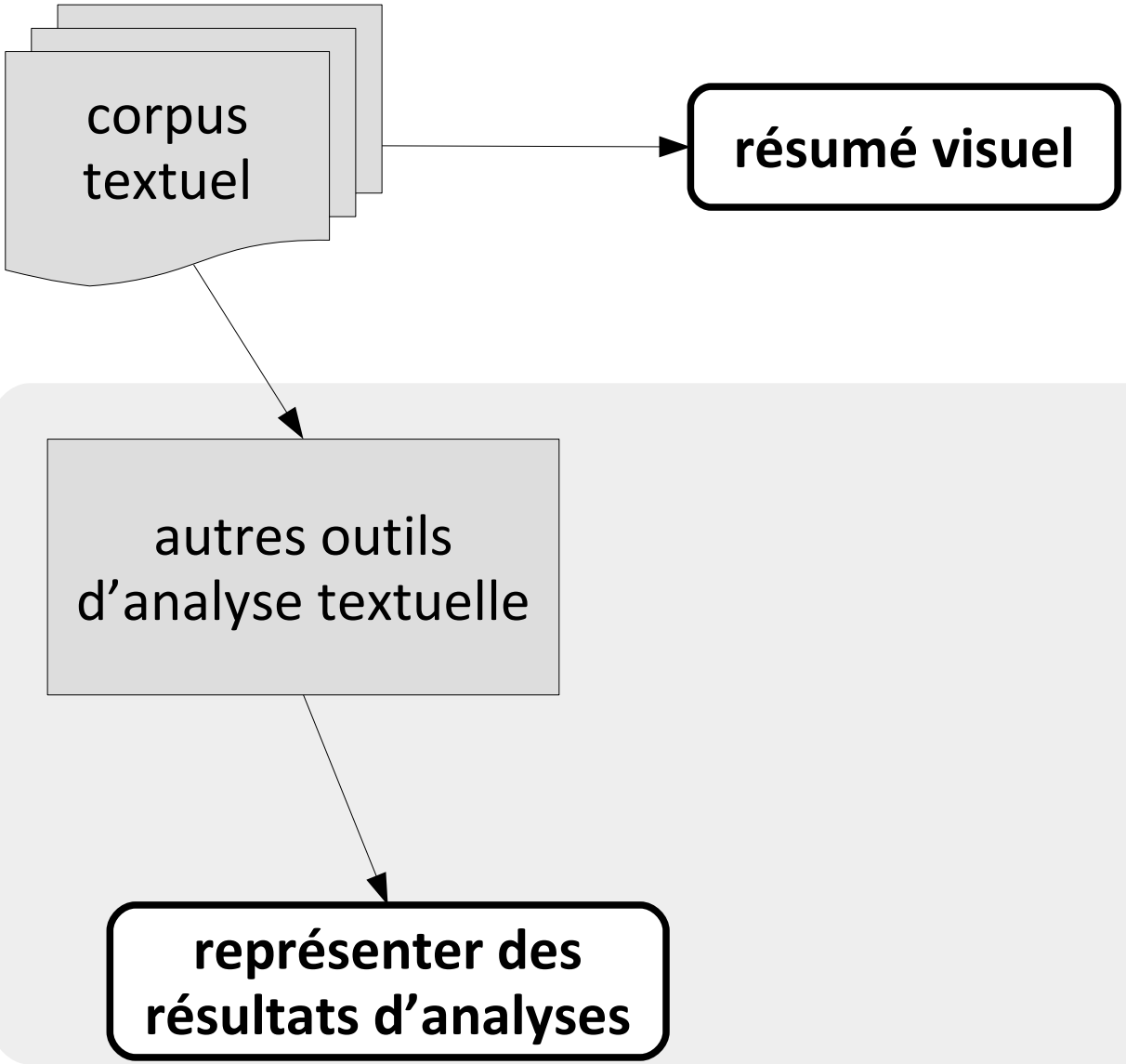
Fréquences, chronologie, dispersion, ciblées sur la cooccurrence d'un mot, ou valeurs personnalisées

Appel transparent au logiciel SplitsTree ou Dendroscope

Le « nuage arboré », pour quoi faire ?

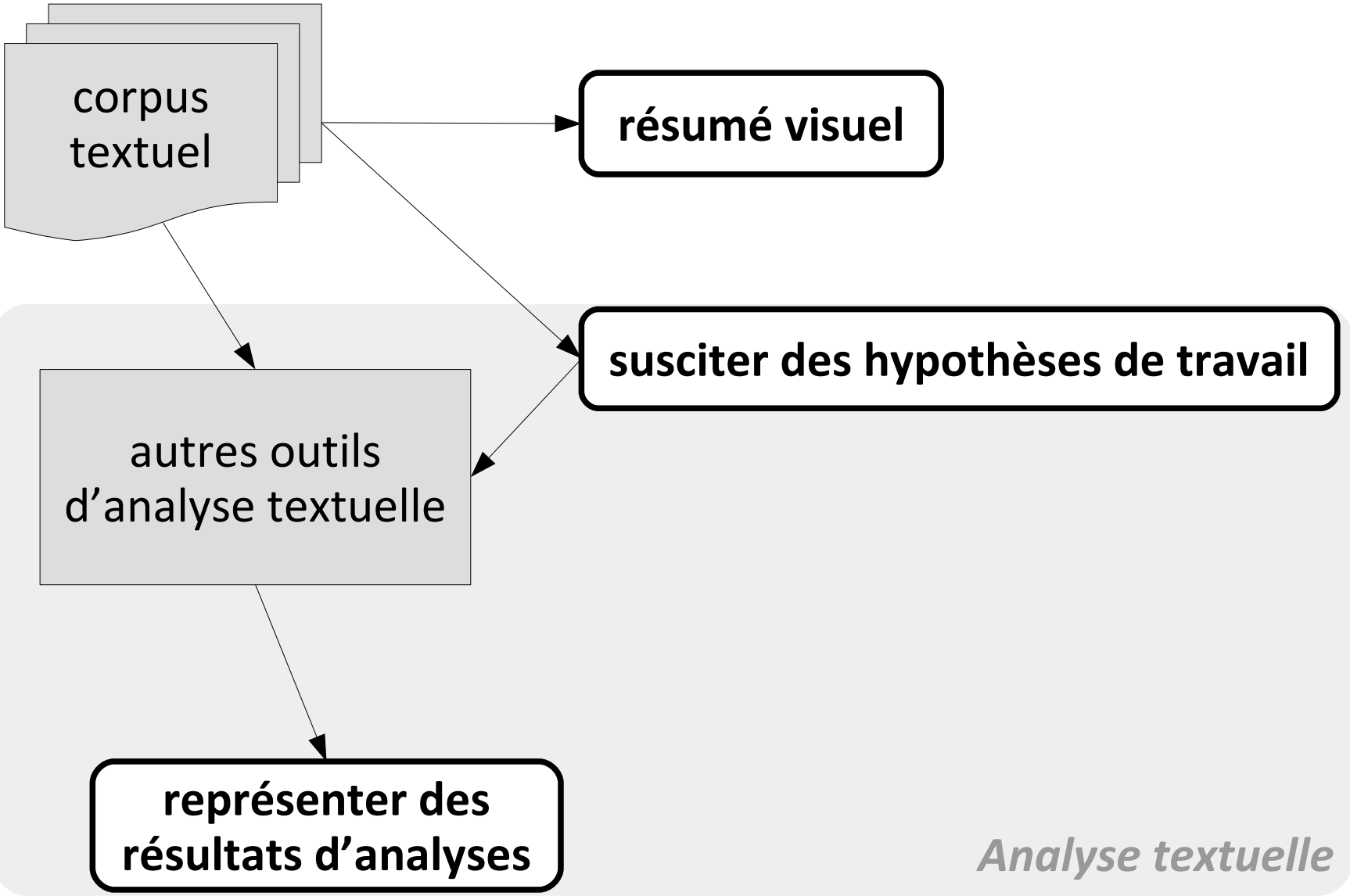


Le « nuage arboré », pour quoi faire ?



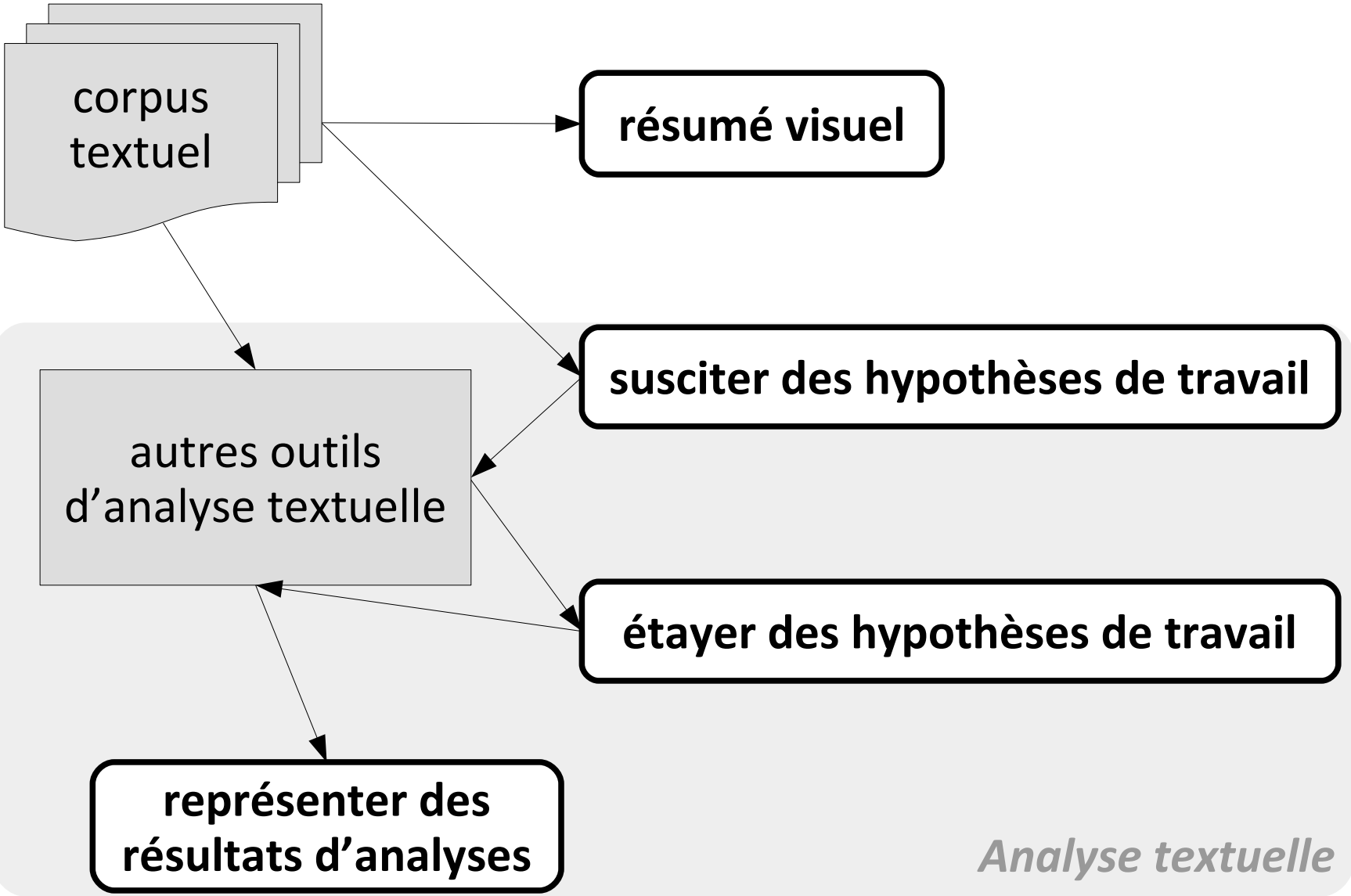
Amstutz & Gambette, *JADT 2010*

Le « nuage arboré », pour quoi faire ?



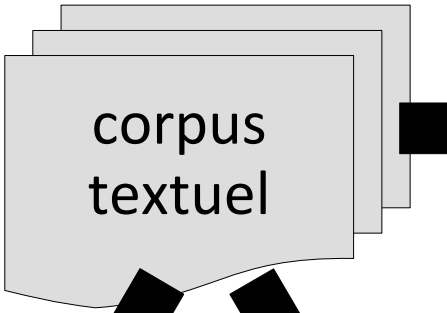
Amstutz & Gambette, *JADT 2010*

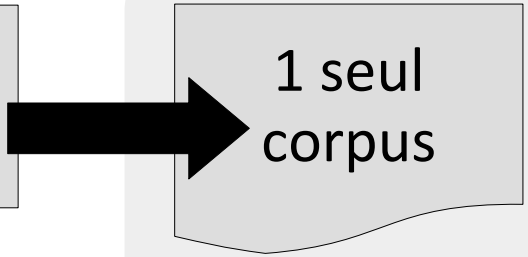
Le « nuage arboré », pour quoi faire ?



Amstutz & Gambette, *JADT 2010*

Exploration de corpus avec TreeCloud

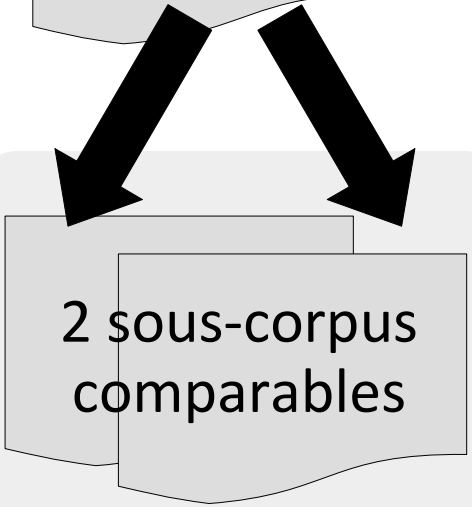




- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

Si le corpus est assez grand :

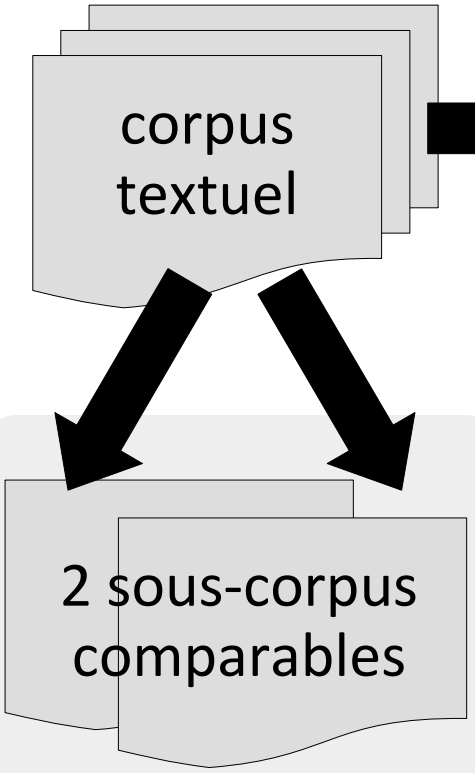
- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.



- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

Exploration de corpus avec TreeCloud



1 seul corpus

- **Interpréter les regroupements**
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

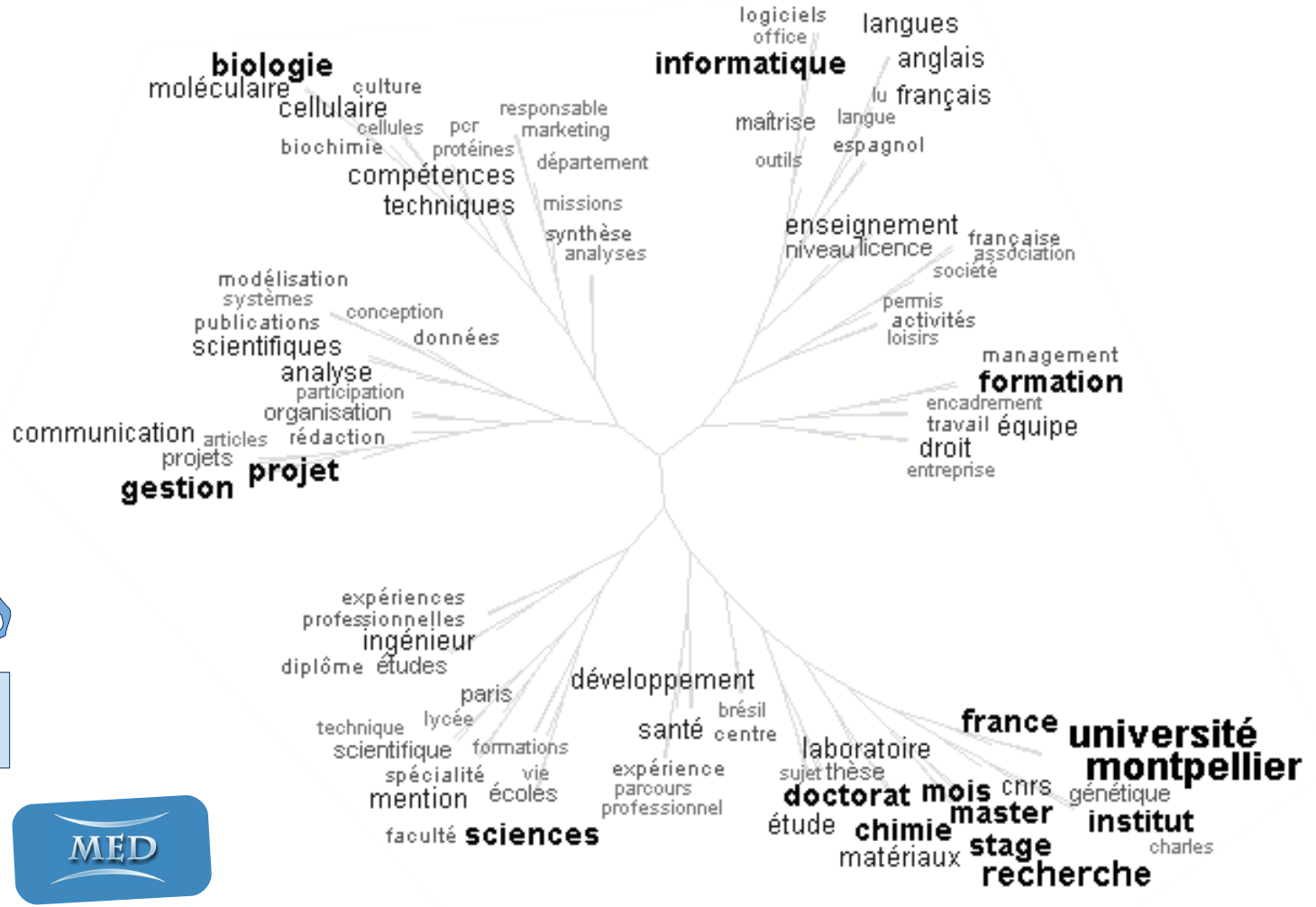
- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

Méthode : interpréter les regroupements

Dessiner des « patates »

Corpus : une centaine de CV soumis à une rencontre docteurs-entreprises



Rencontre Docteurs & Entreprises

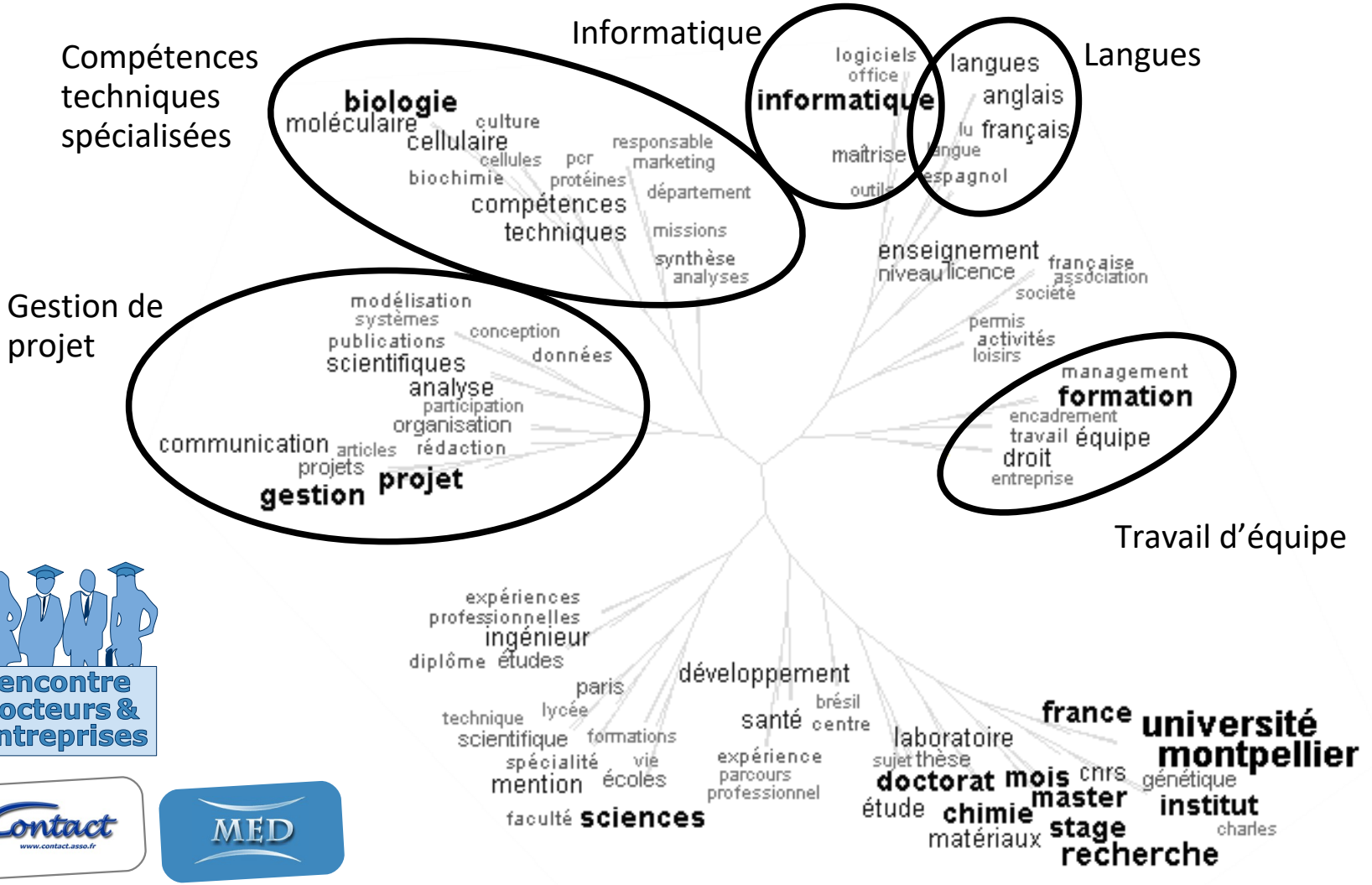
Contact www.contact.asso.fr

MED

Méthode : interpréter les regroupements

Dessiner des « patates »

Corpus : une centaine de CV soumis à une rencontre docteurs-entreprises

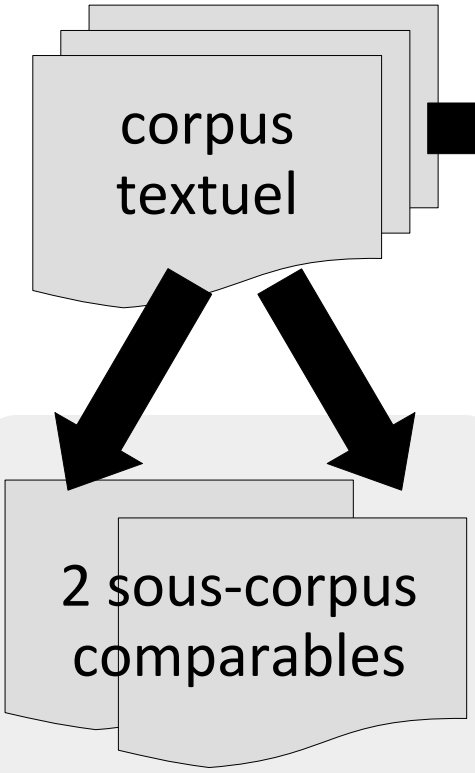


Rencontre Docteurs & Entreprises

Contact
www.contact.asso.fr

MED

Exploration de corpus avec TreeCloud



1 seul corpus

- Interpréter les regroupements
- **Examiner les voisinages des mots fréquents dans l'arbre**
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

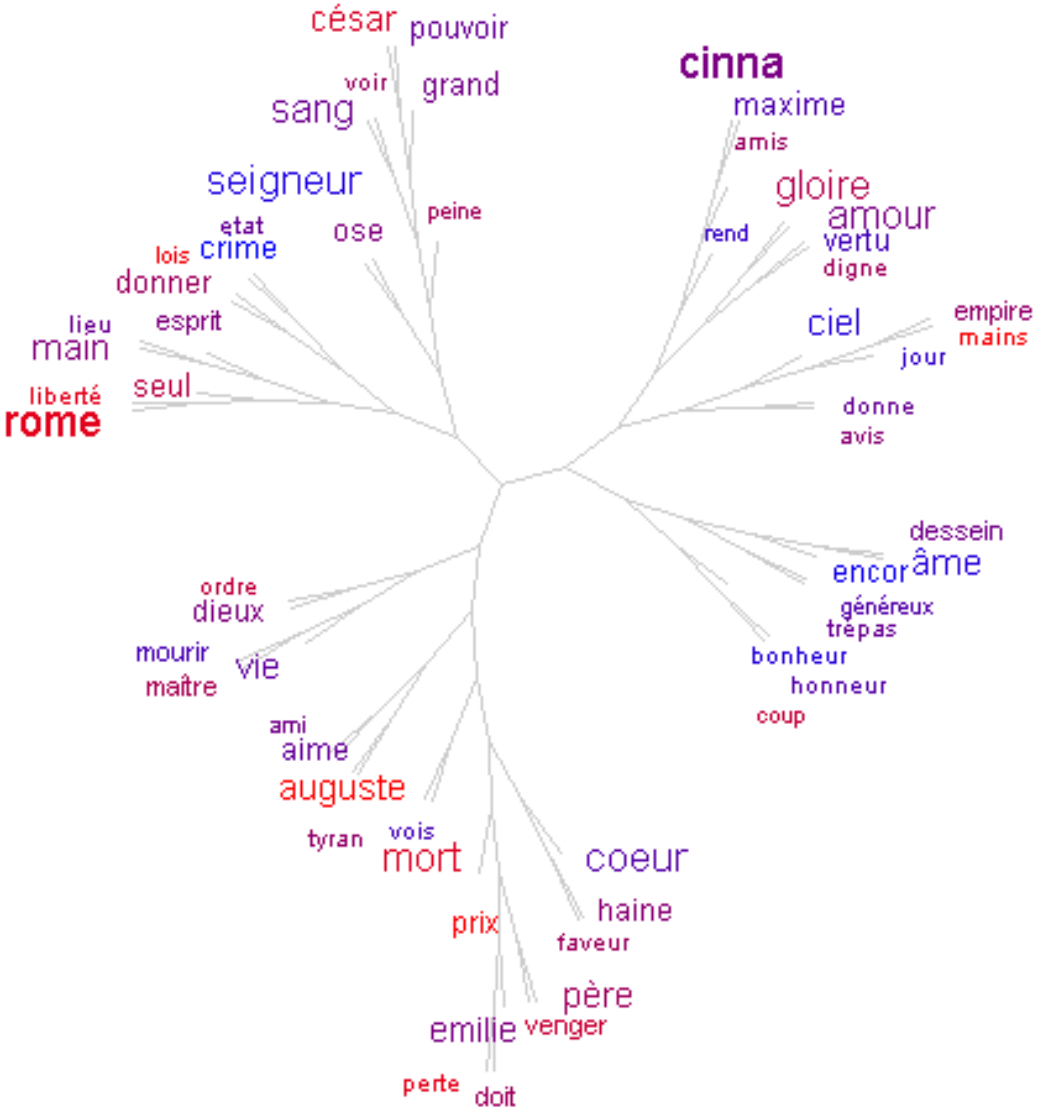
Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

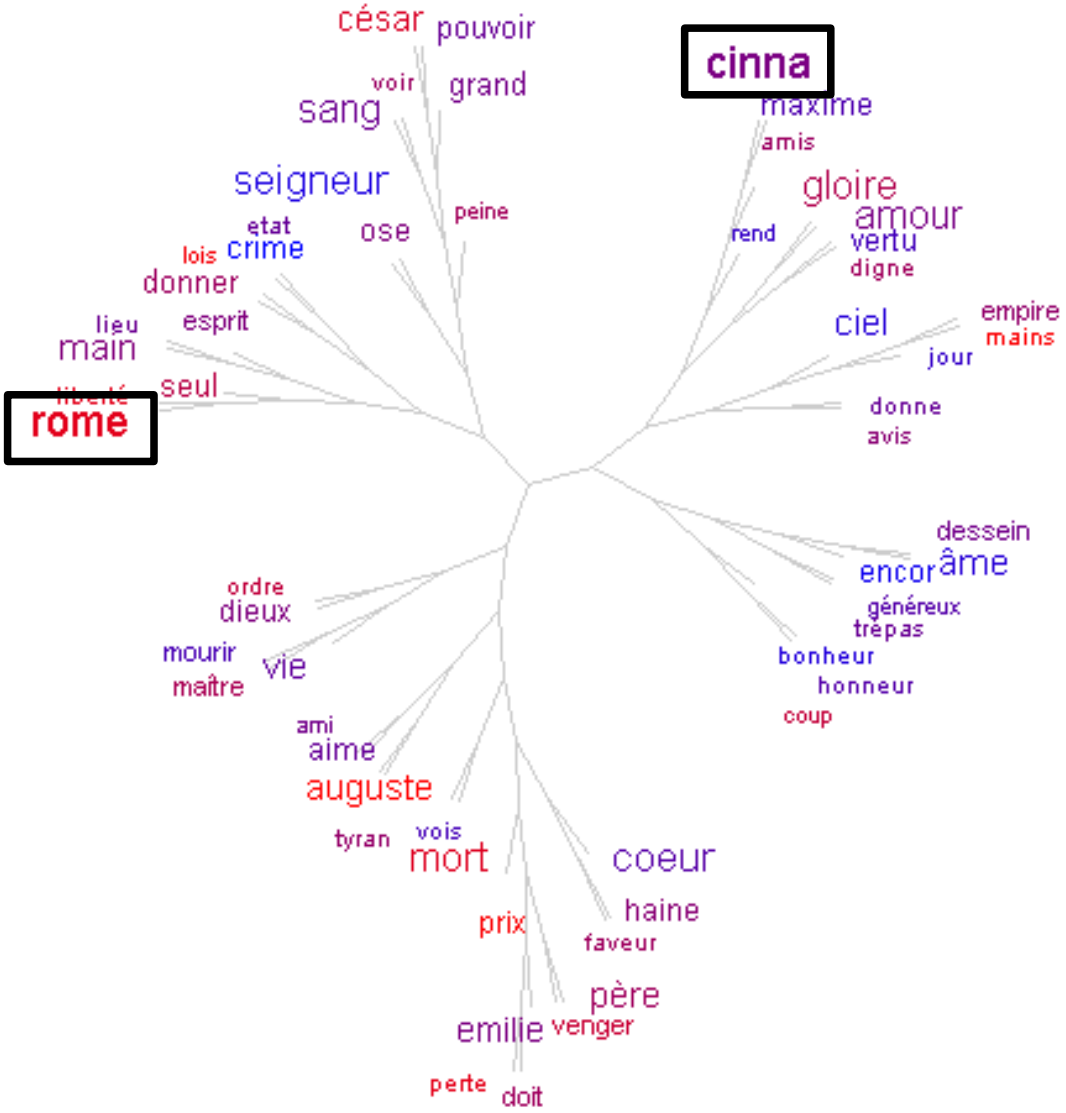
Méthode : voisinage des mots fréquents



Amstutz & Gambette, JADT 2010

Nuage arboré globaux des 60 mots les plus fréquents dans *Cinna* de Corneille (distance Liddell, fenêtre de largeur 20), colorés chronologiquement (rouge au début, bleu à la fin)

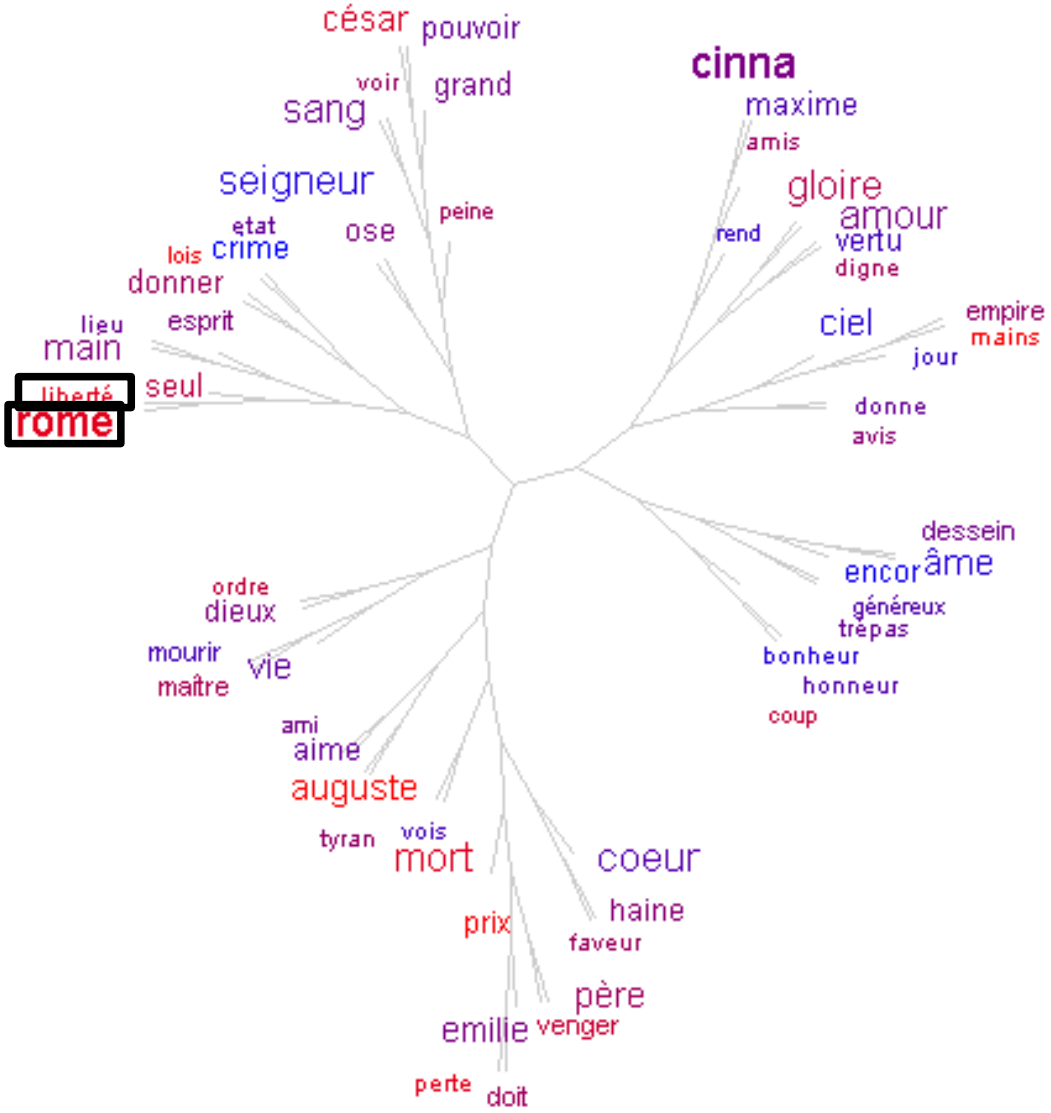
Méthode : voisinage des mots fréquents



Amstutz & Gambette, JADT 2010

Nuage arboré globaux des 60 mots les plus fréquents dans *Cinna* de Corneille (distance Liddell, fenêtre de largeur 20), colorés chronologiquement (rouge au début, bleu à la fin)

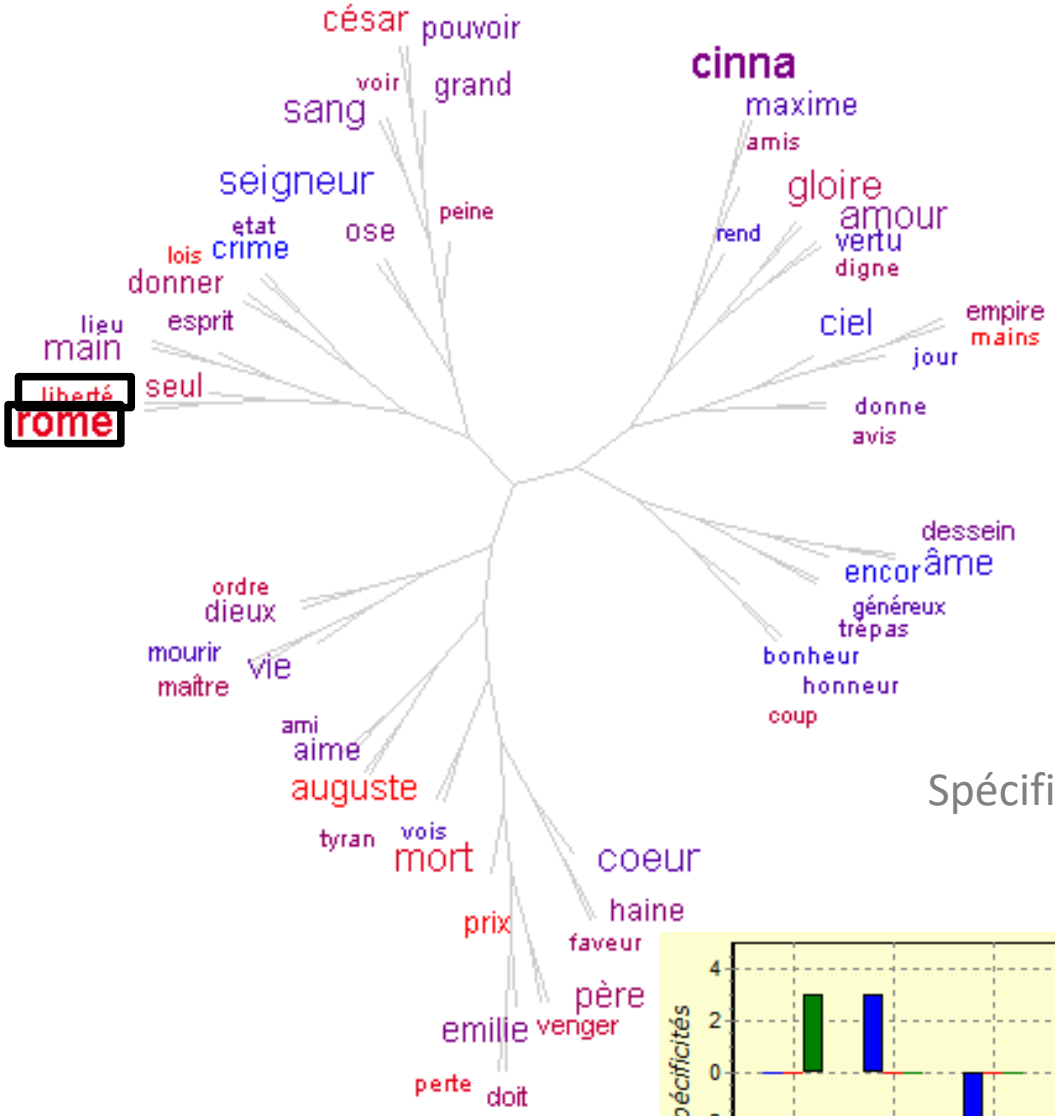
Méthode : voisinage des mots fréquents



Amstutz & Gambette, JADT 2010

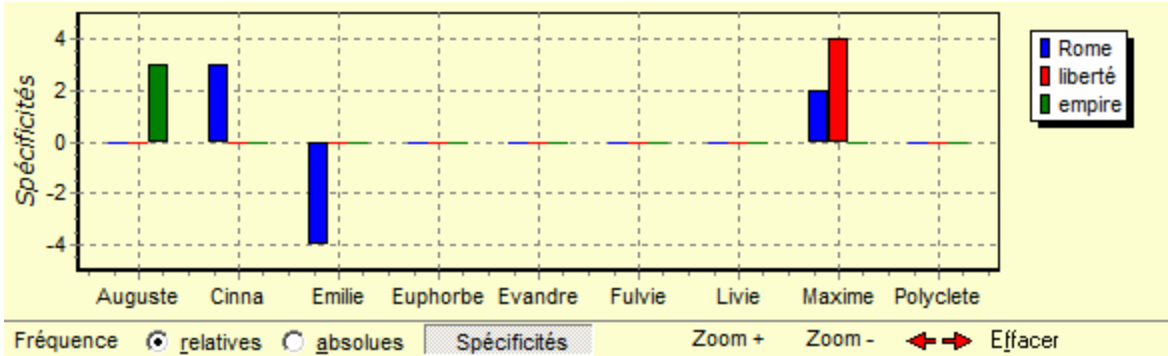
Nuage arboré globaux des 60 mots les plus fréquents dans *Cinna* de Corneille (distance Liddell, fenêtre de largeur 20), colorés chronologiquement (rouge au début, bleu à la fin)

Méthode : voisinage des mots fréquents

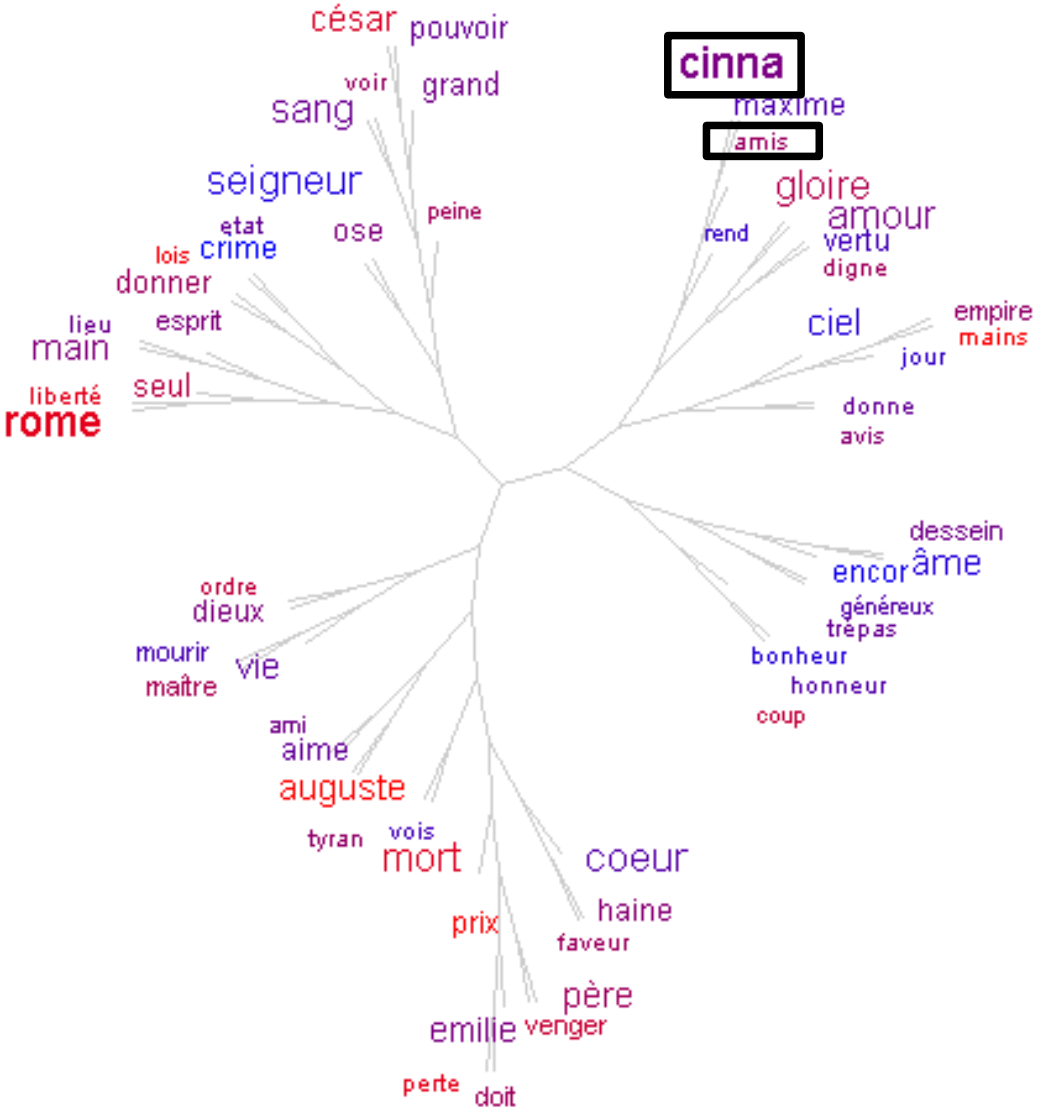


Amstutz & Gambette, JADT 2010

Spécificités d'emploi de « Rome », « liberté » et « empire », chez les différents personnages de Cinna, selon Lexico3

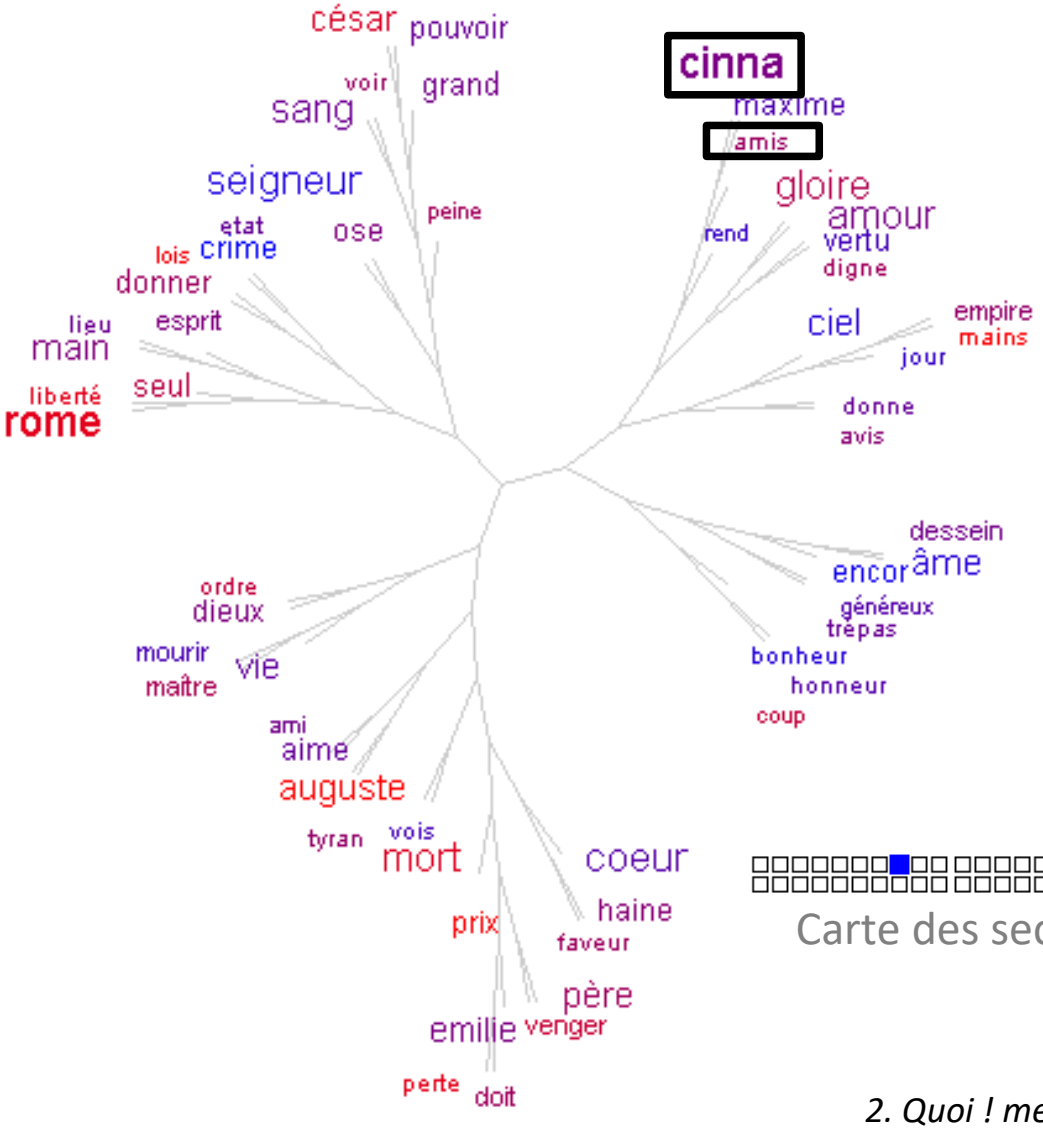


Méthode : voisinage des mots fréquents

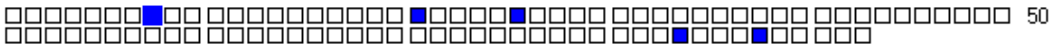


Amstutz & Gambette, JADT 2010

Méthode : voisinage des mots fréquents



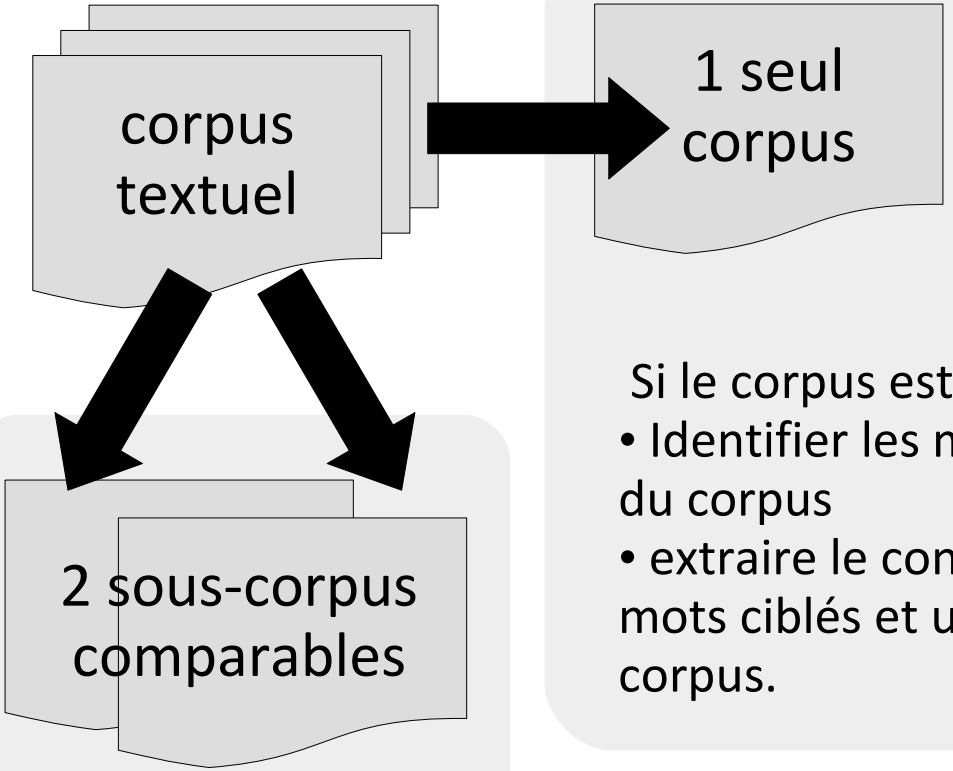
Amstutz & Gambette, JADT 2010



Carte des sections Lexico3 et contextes de « amis » dans les paroles d'Auguste dans *Cinna*

*1. Voilà, mes chers **amis**, ce qui me met en peine.*
*2. Quoi ! mes plus chers **amis** ! quoi ! Cinna ! quoi ! Maxime !*
*3. Reprenez le pouvoir que vous m'avez commis, Si donnant des sujets il ôte les **amis***
*4. Soyons **amis**, Cinna, c'est moi qui t'en convie*
*5. Il nous a trahis tous ; mais ce qu'il a commis Vous conserve innocents, et me rend mes **amis**.*

Exploration de corpus avec TreeCloud



- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- **Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)**

Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

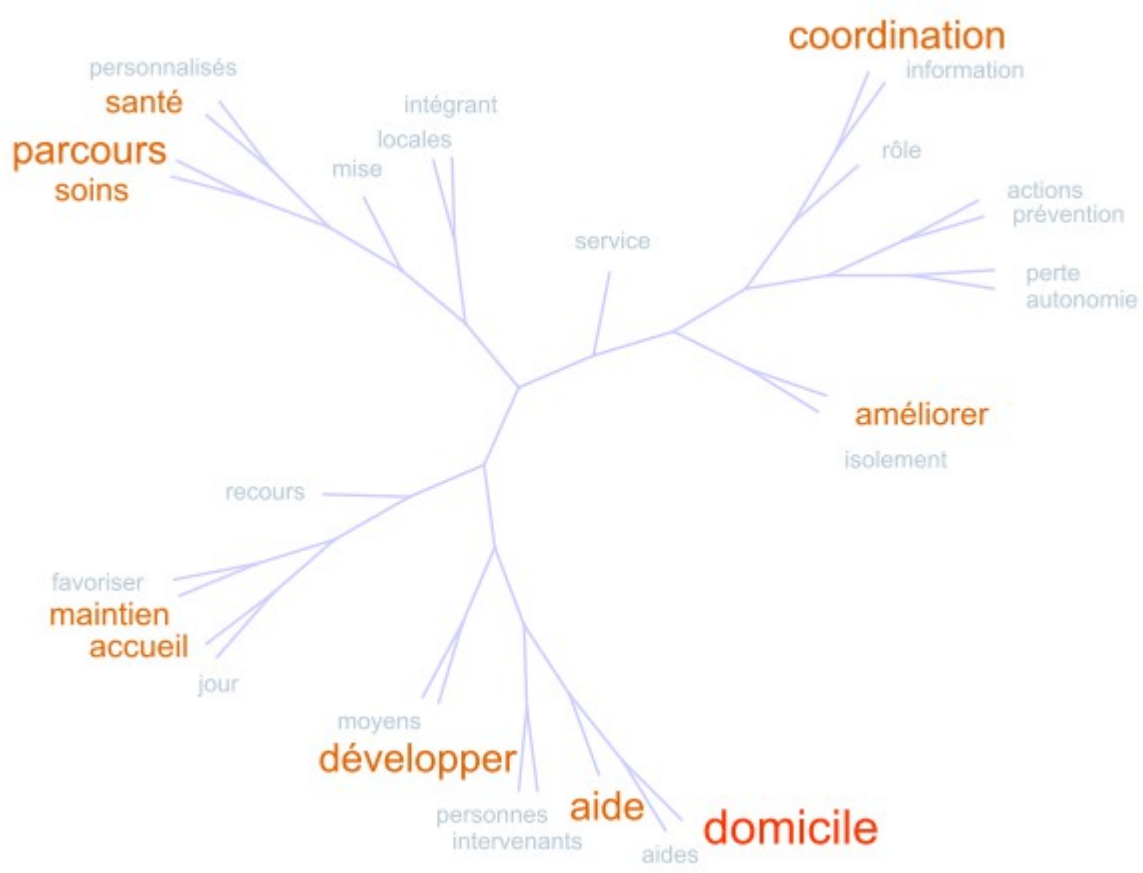
- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

Méthode : voisinage des verbes

Corpus : réponses à des questions ouvertes à des professionnels de la santé sur le parcours de santé des personnes âgées dans les Alpes de Haute-Provence

Suggestions d'améliorations :

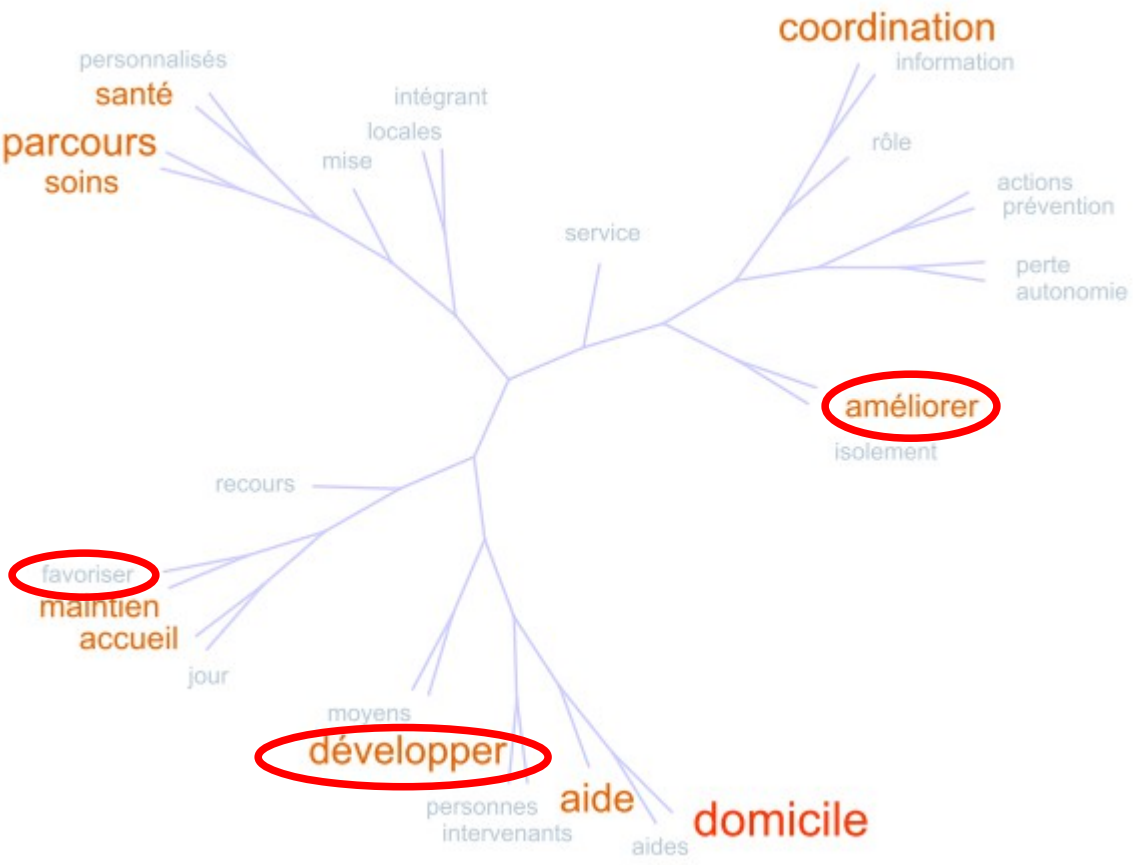


P. Grenier-Tisserand et
l'équipe projet de la DT 04

Méthode : voisinage des verbes

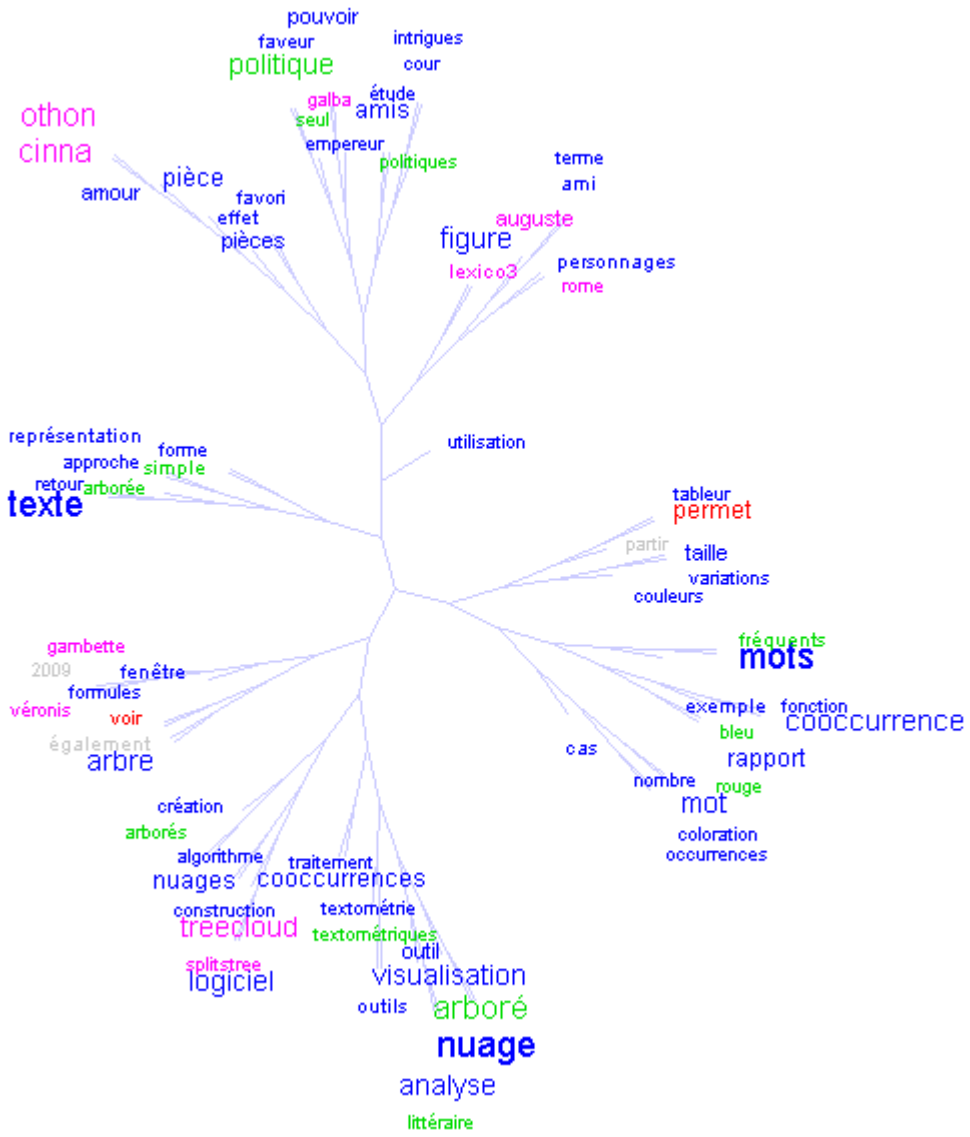
Corpus : réponses à des questions ouvertes à des professionnels de la santé sur le parcours de santé des personnes âgées dans les Alpes de Haute-Provence

Suggestions d'améliorations :



P. Grenier-Tisserand et
l'équipe projet de la DT 04

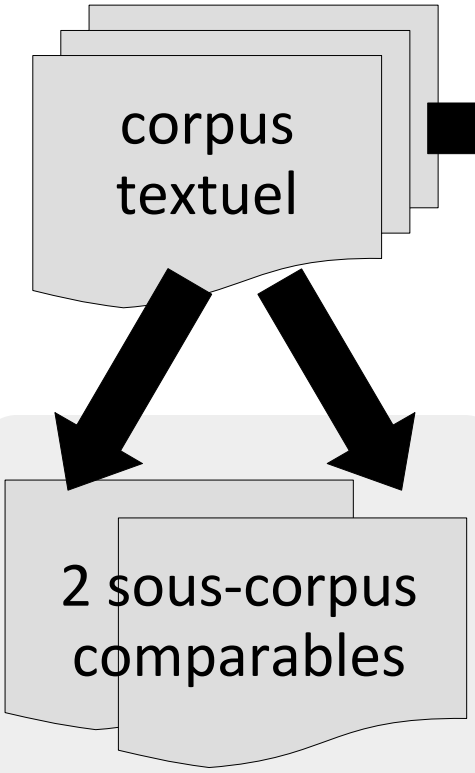
Perspective : coloration grammaticale



noms
adjectifs
verbes
noms propres

Nuage arboré des mots apparaissant 5 fois ou plus dans l'article d'Amstutz & Gambette, JADT 2010, distance Liddell, fenêtre de 20 mots, coloration personnalisée à partir d'un étiquetage TreeTagger

Exploration de corpus avec TreeCloud



1 seul corpus

- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

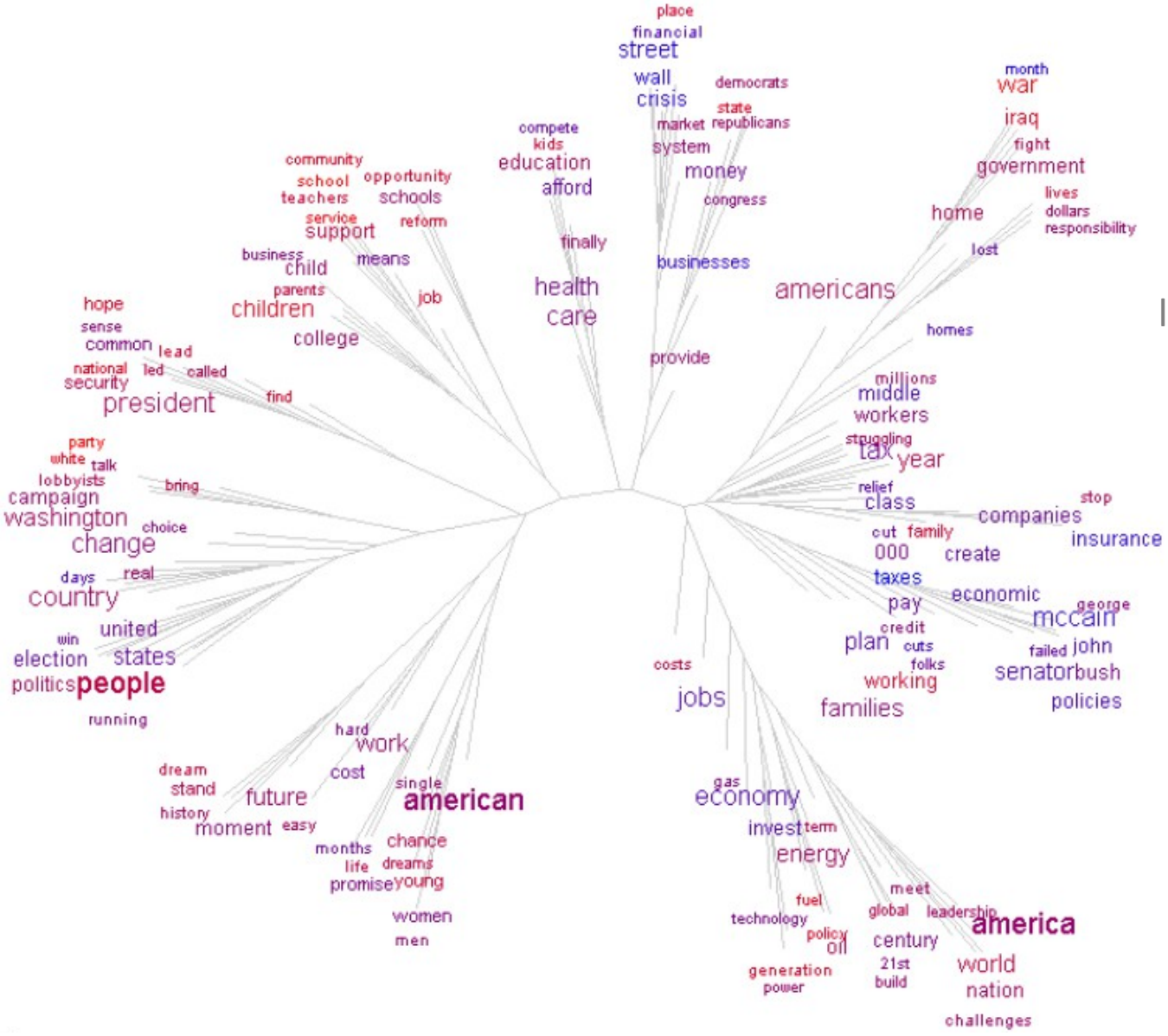
Si le corpus est assez grand :

- **Identifier les mots plutôt au début ou à la fin du corpus**
- extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

Méthode : mots au début ou à la fin

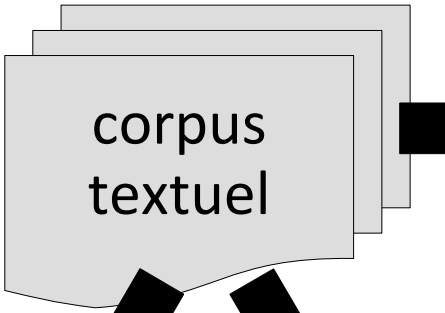


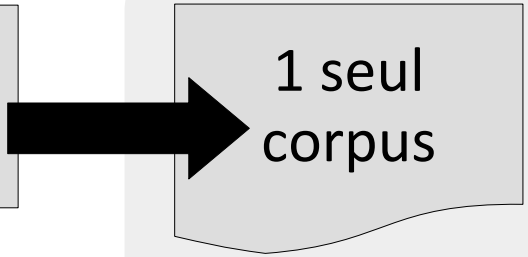
Nuage arboré de l'ensemble des discours de campagne de 2008 de Barack Obama, coloration chronologique

début de la campagne
fin de la campagne

Gambette & Véronis, IFCS 2009

Exploration de corpus avec TreeCloud

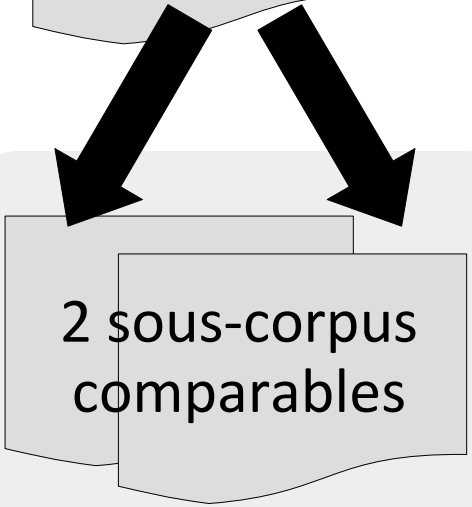




- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

Si le corpus est assez grand :

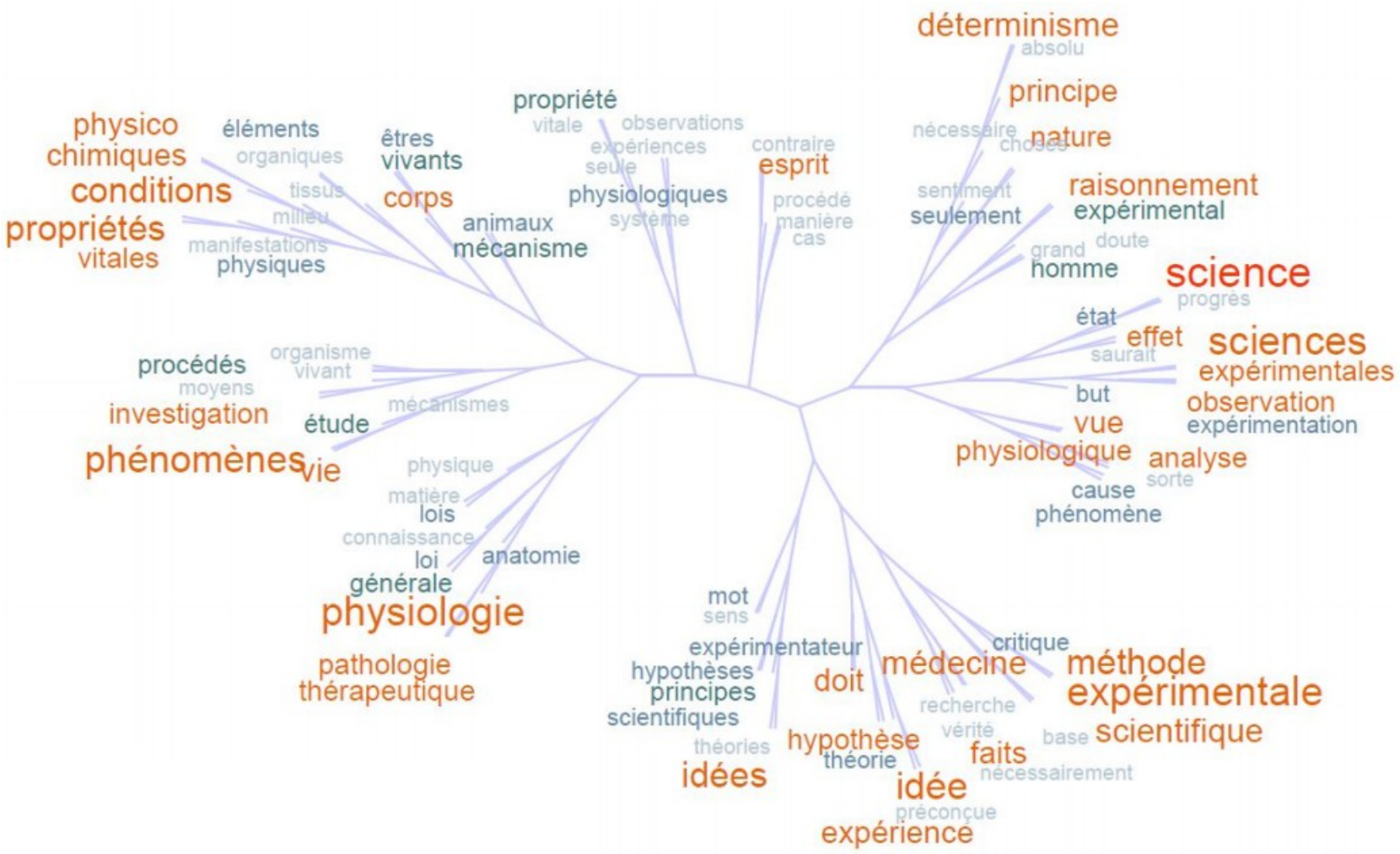
- Identifier les mots plutôt au début ou à la fin du corpus
- **extraire le contexte dans le corpus de certains mots ciblés et utiliser ces données comme corpus.**



- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

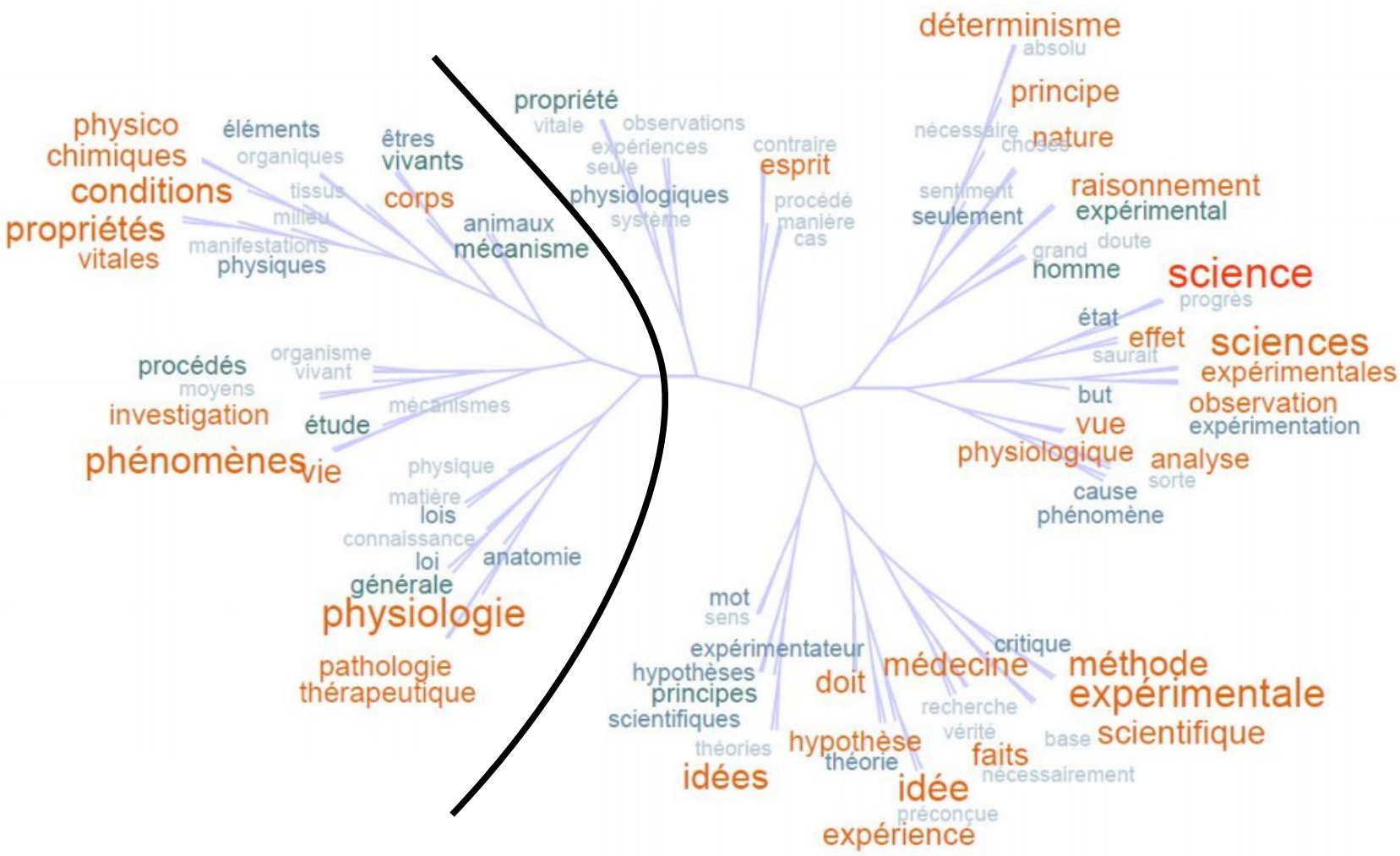
Travail mené avec Edna Hernandez

Nuages arborés des contextes de « étude »



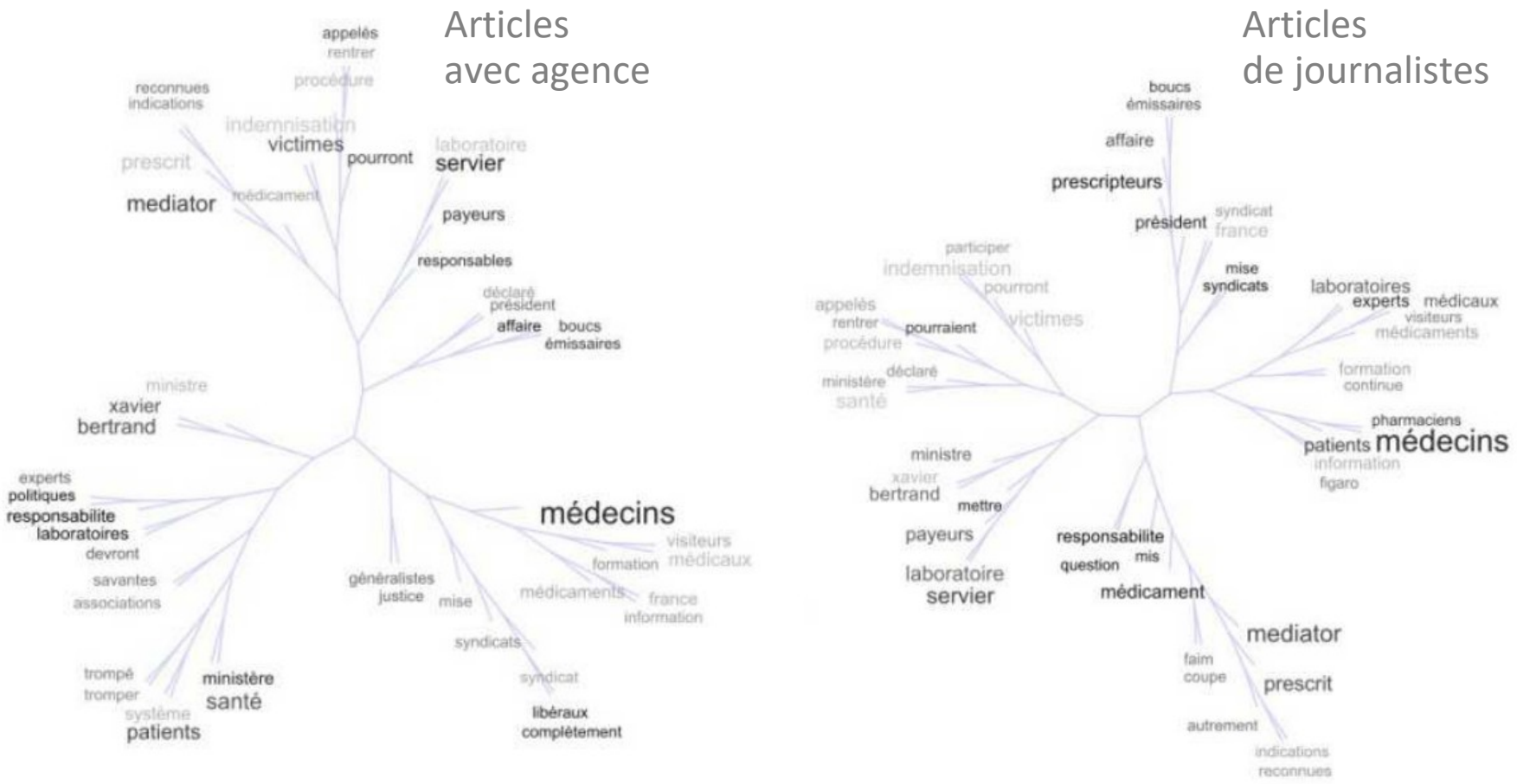
Nuage arboré des 100 mots les plus fréquents dans les contextes (10 mots avant, 10 mots après) des mots de la catégorie "étude" dans un corpus de textes scientifiques et littéraires sur la science (projet AnimalHumanité)

Nuages arborés des contextes de « étude »



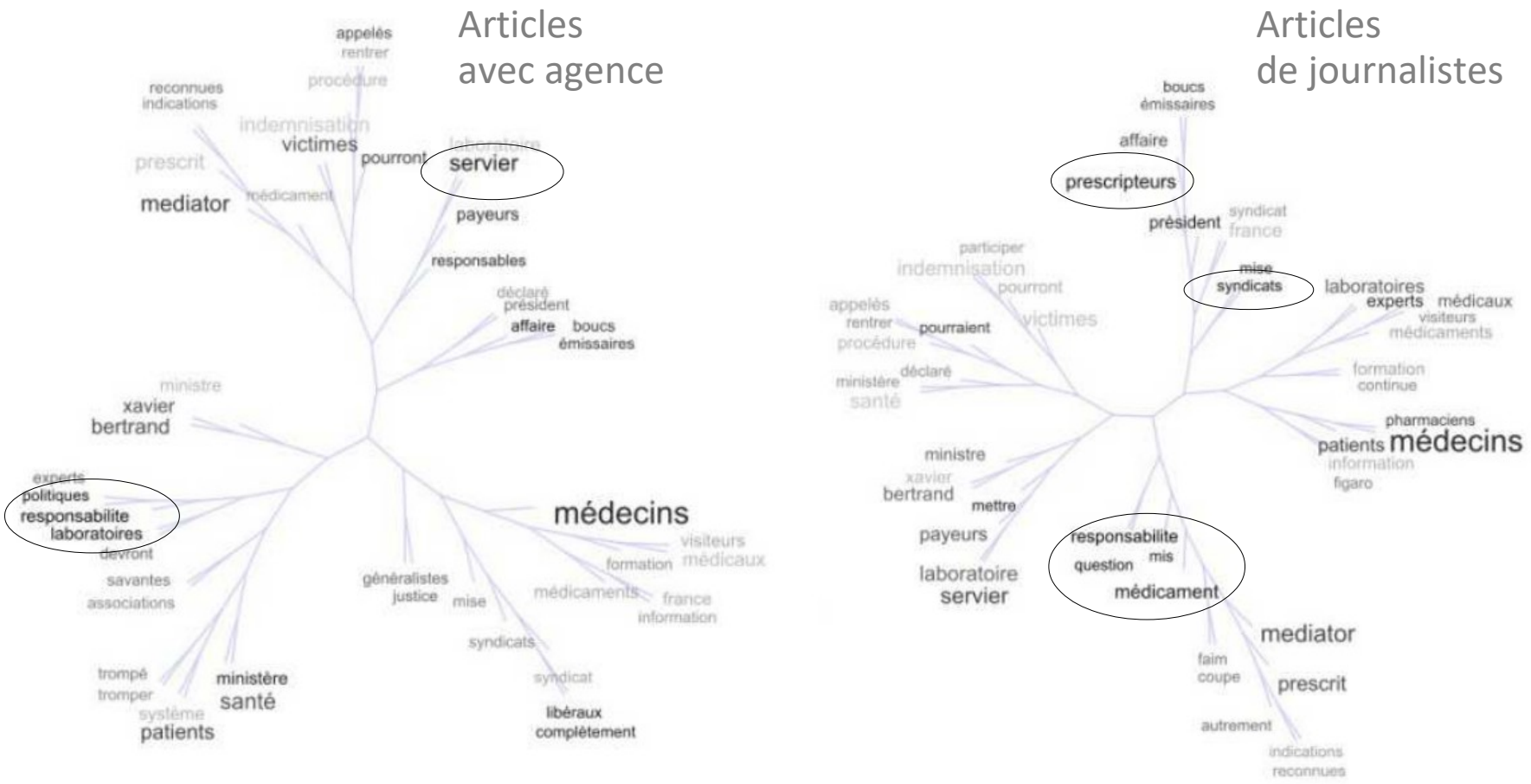
Nuage arboré des 100 mots les plus fréquents dans les contextes (10 mots avant, 10 mots après) des mots de la catégorie "étude" dans un corpus de textes scientifiques et littéraires sur la science (projet AnimalHumanité)

Nuages arborés des contextes de « médecins »



Nuage arboré des 50 mots les plus fréquents des contextes (10 mots avant et 10 mots après) du mot médecins dans le sous-corpus des articles sur le Mediator, colorés par le degré de cooccurrence avec le mot responsabilités (en noir pour les mots les plus cooccurrents), construit par TreeCloud avec la formule Liddell, et des fenêtres glissantes de 20 mots

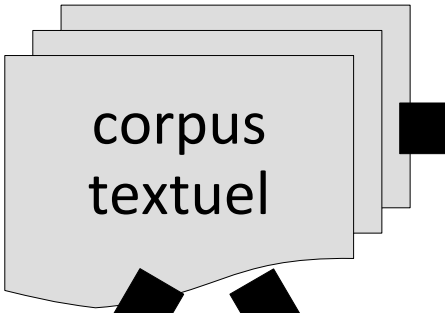
Nuages arborés des contextes de « médecins »

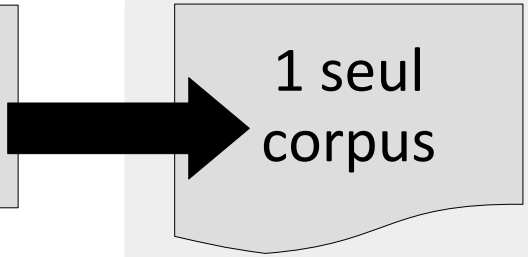


Nuage arboré des 50 mots les plus fréquents des contextes (10 mots avant et 10 mots après) du mot médecins dans le sous-corpus des articles sur le Mediator, colorés par le degré de cooccurrence avec le mot responsabilités (en noir pour les mots les plus cooccurrents), construit par TreeCloud avec la formule Liddell, et des fenêtres glissantes de 20 mots

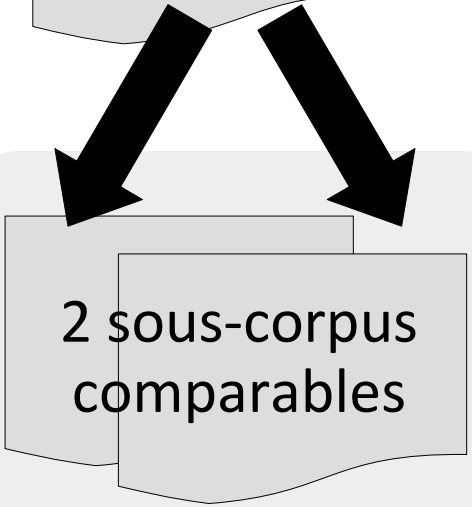
Martinez & Gambette, *Corpus,* 2013

Exploration de corpus avec TreeCloud





- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)



Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le voisinage dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

- **Identifier des regroupements communs ou différents**
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail mené avec Edna Hernandez

Illustration sur le corpus Mediator

Comparer les articles d'agences et articles de journalistes

Corpus : 595 articles d'agences contre 1496 articles de journalistes de 2011 évoquant l'affaire du Mediator dans la presse française.

Ensemble des articles

Gambette & Martinez, *Texto!*, 2013

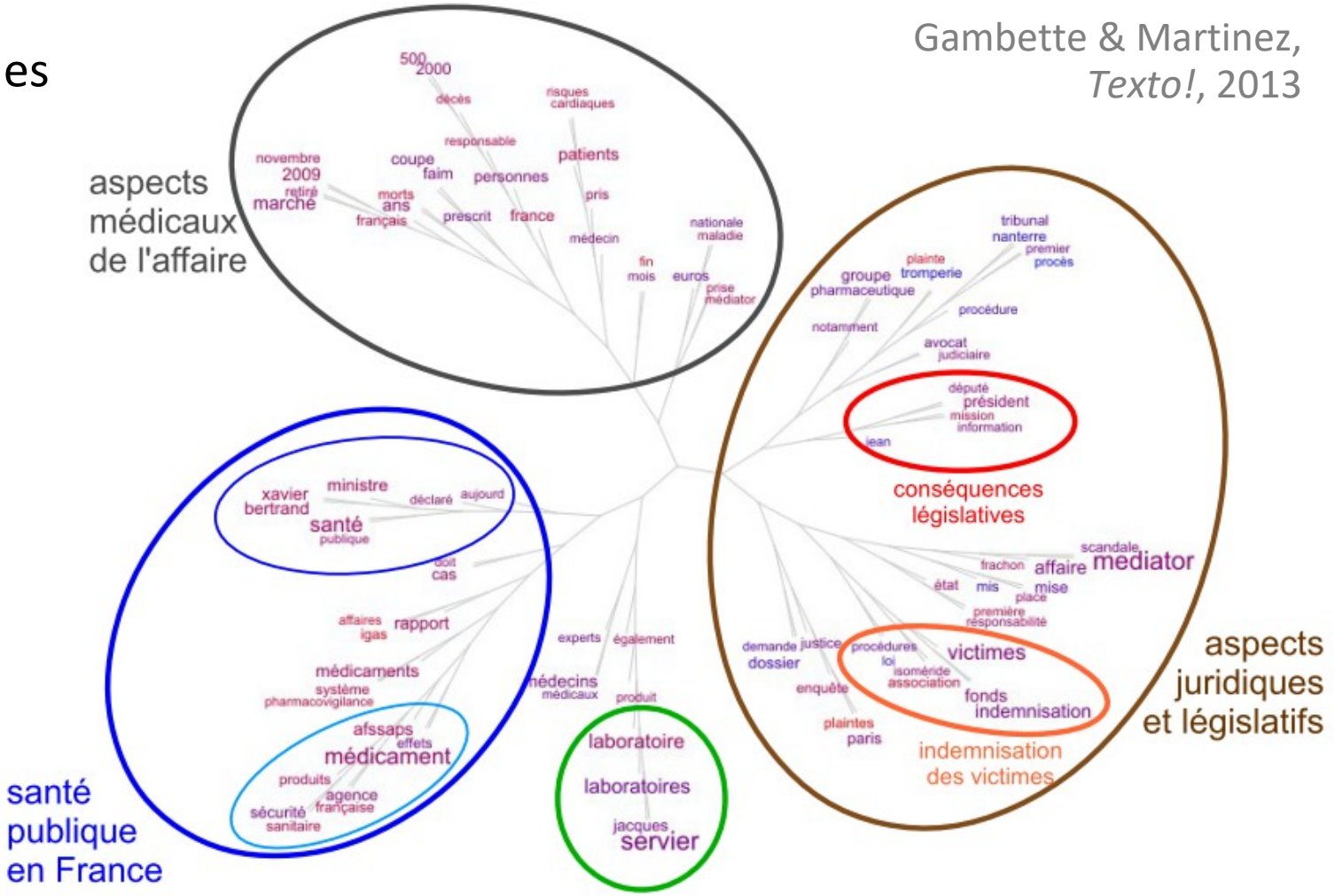
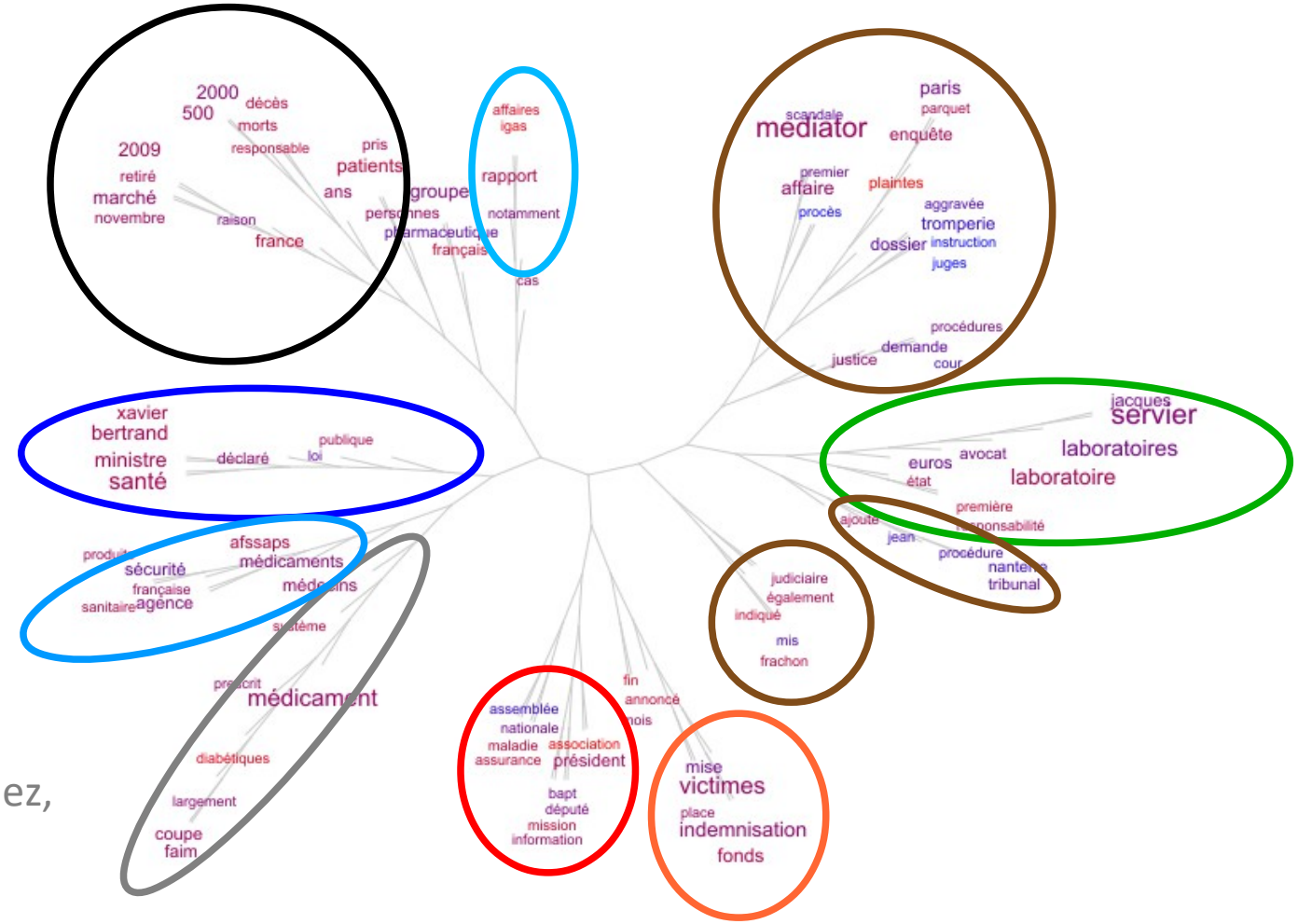


Illustration sur le corpus Mediator

Comparer les articles d'agences et articles de journalistes

Corpus : 595 articles d'agences contre 1496 articles de journalistes de 2011 évoquant l'affaire du Mediator dans la presse française.

Articles d'agences



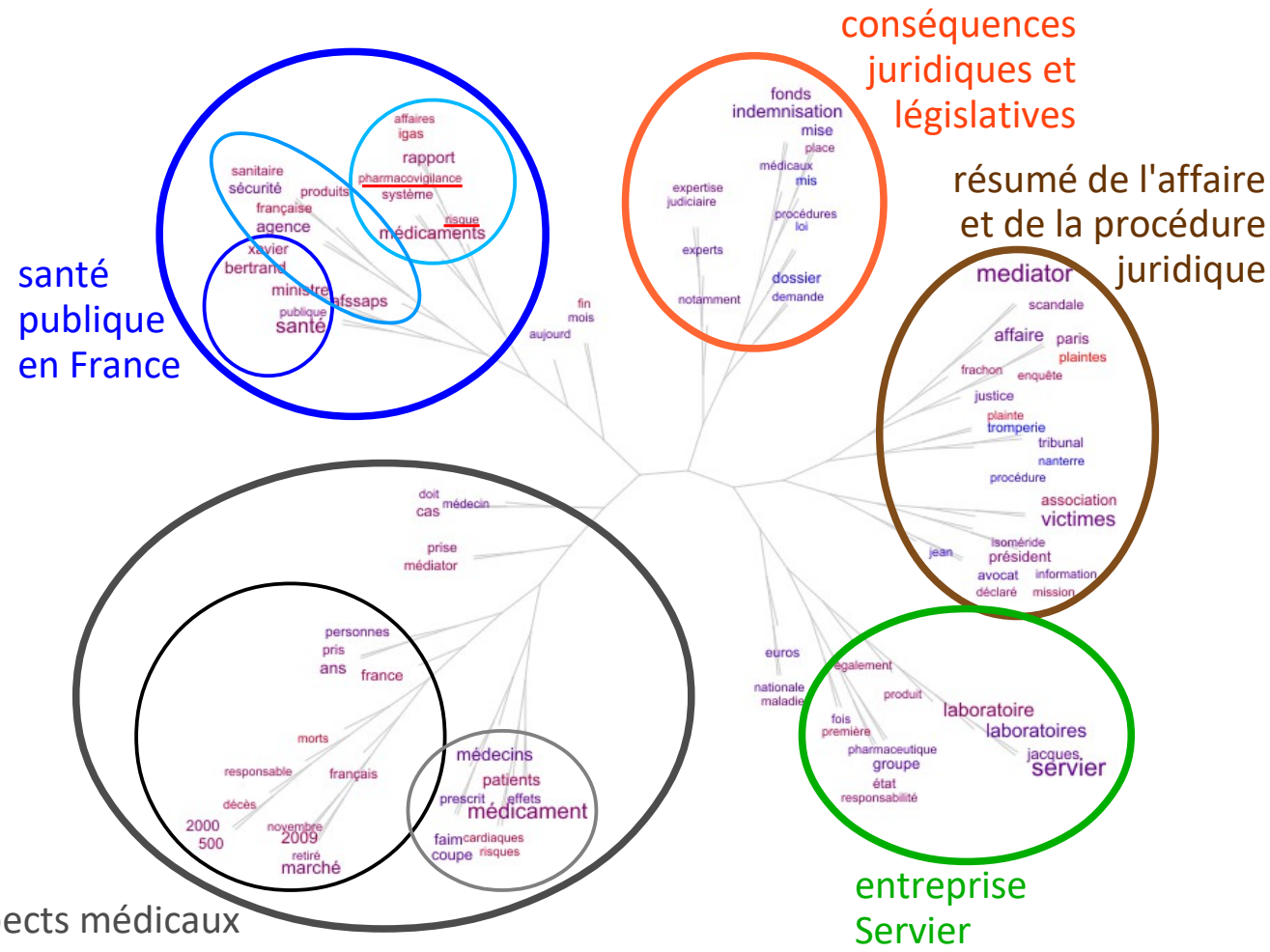
Gambette & Martinez, *Texto!*, 2013

Illustration sur le corpus Mediator

Comparer les articles d'agences et articles de journalistes

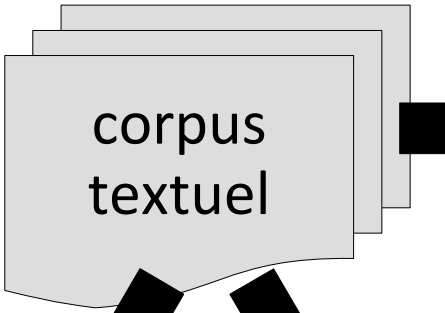
Corpus : 595 articles d'agences contre 1496 articles de journalistes de 2011 évoquant l'affaire du Mediator dans la presse française.

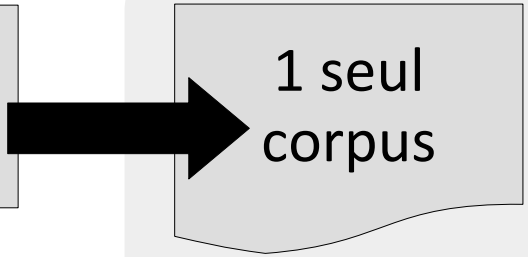
Articles
de journalistes



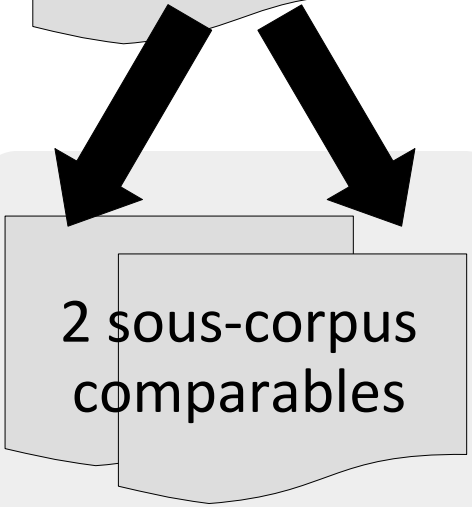
Gambette & Martinez,
Texto!, 2013

Exploration de corpus avec TreeCloud





- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)



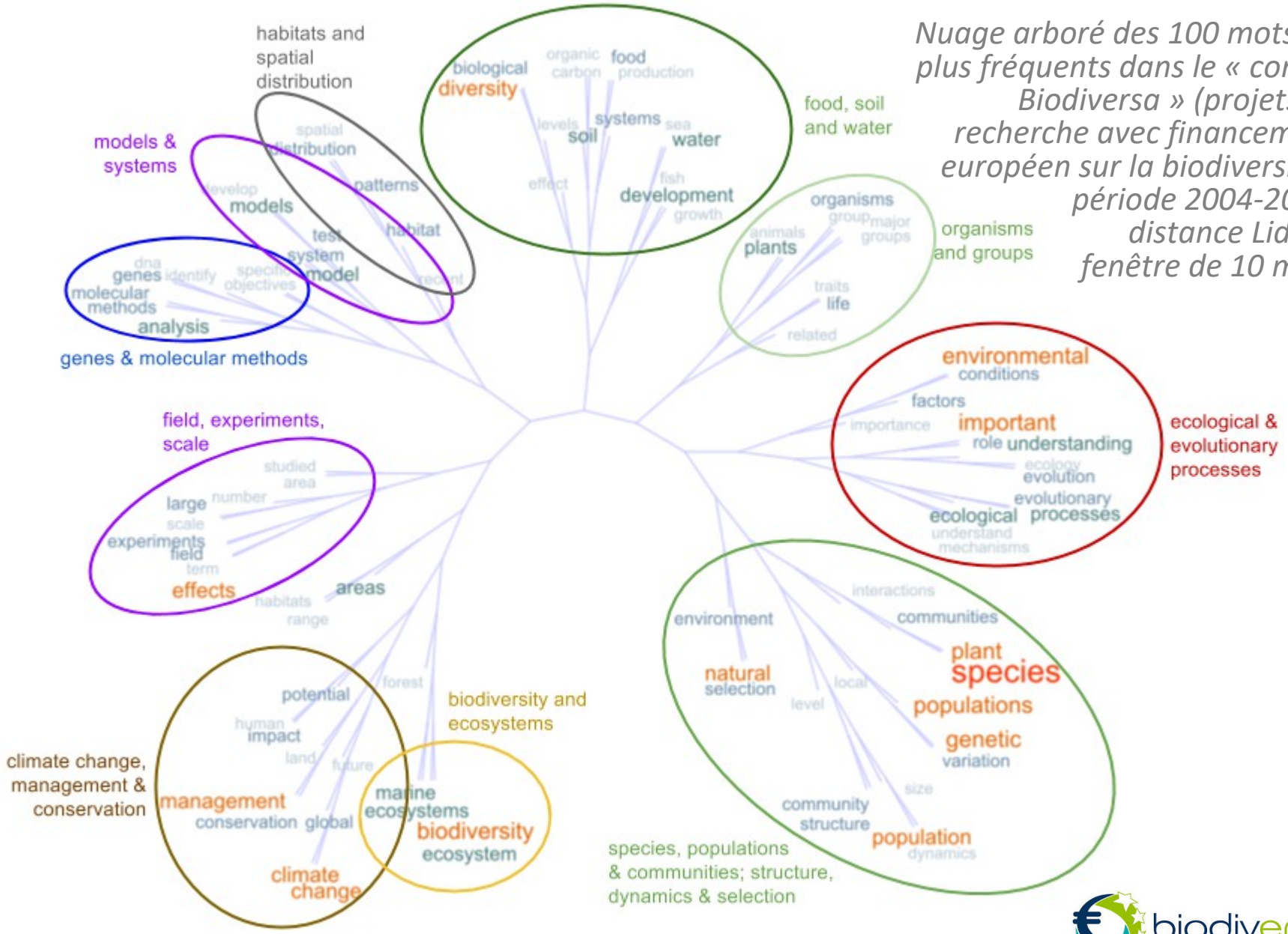
Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le voisinage dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- **Comparer (quanti+quali) les regroupements dans les deux arbres**
- Construire le nuage arboré des mots spécifiques de chaque sous-corpus

Travail en cours avec Edna Hernandez

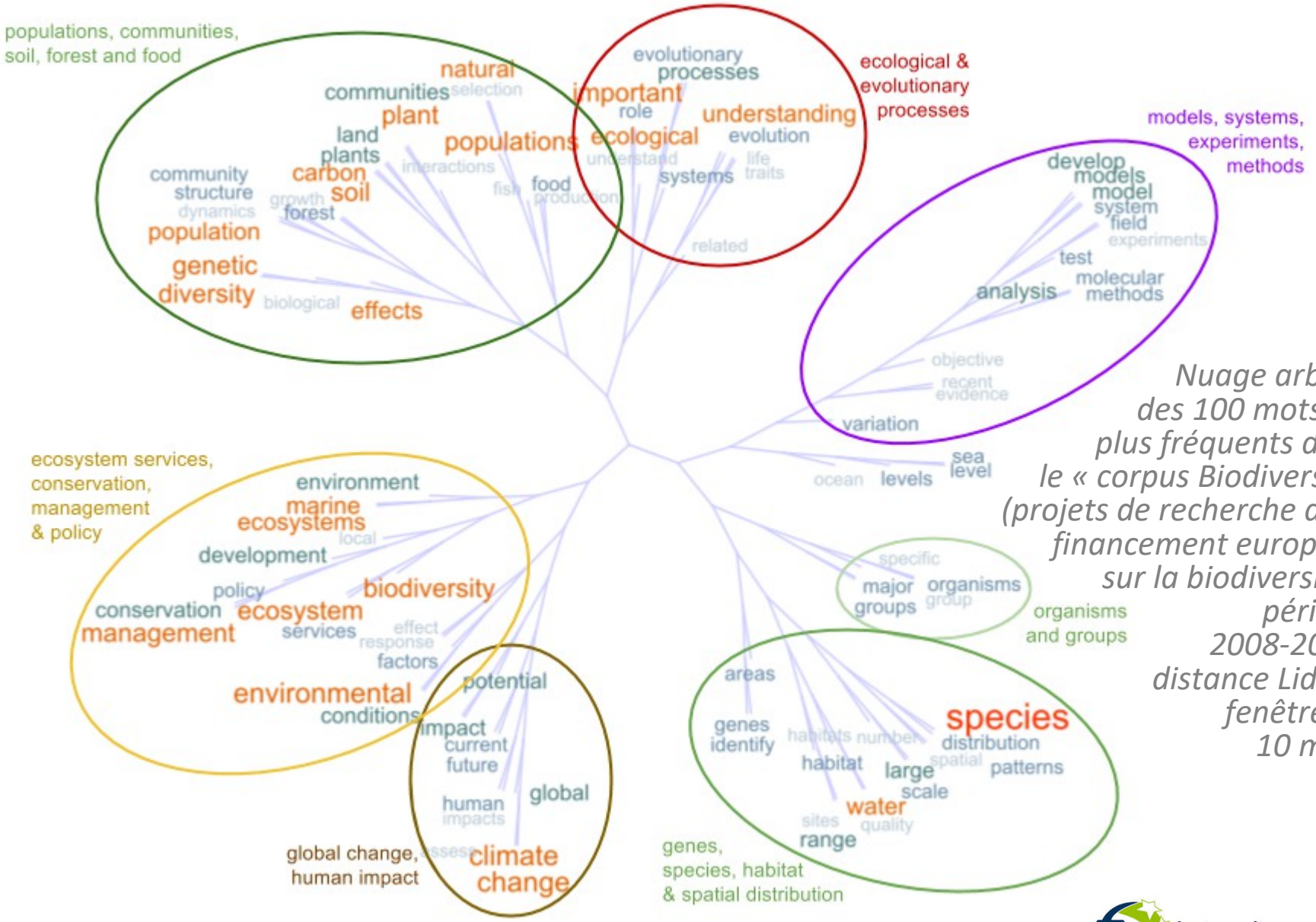
Méthode : comparaison de voisinages dans l'arbre



Nuage arboré des 100 mots les plus fréquents dans le « corpus Biodiversa » (projets de recherche avec financement européen sur la biodiversité), période 2004-2007, distance Liddel, fenêtre de 10 mots

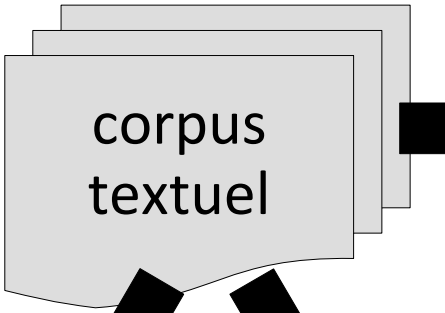
biodiversa

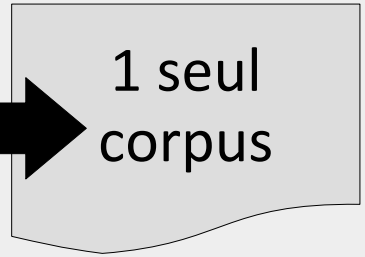
Méthode : comparaison de voisinages dans l'arbre



Nuage arboré des 100 mots les plus fréquents dans le « corpus Biodiversa » (projets de recherche avec financement européen sur la biodiversité), période 2008-2011, distance Liddel, fenêtre de 10 mots

Exploration de corpus avec TreeCloud

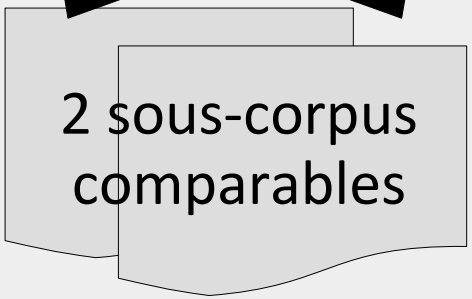




- Interpréter les regroupements
- Examiner les voisinages des mots fréquents dans l'arbre
- Examiner le voisinage de mots ciblés dans l'arbre («acteurs», verbes, etc.)

Si le corpus est assez grand :

- Identifier les mots plutôt au début ou à la fin du corpus
- extraire le voisinage dans le corpus de certains mots ciblés et utiliser ces données comme corpus.

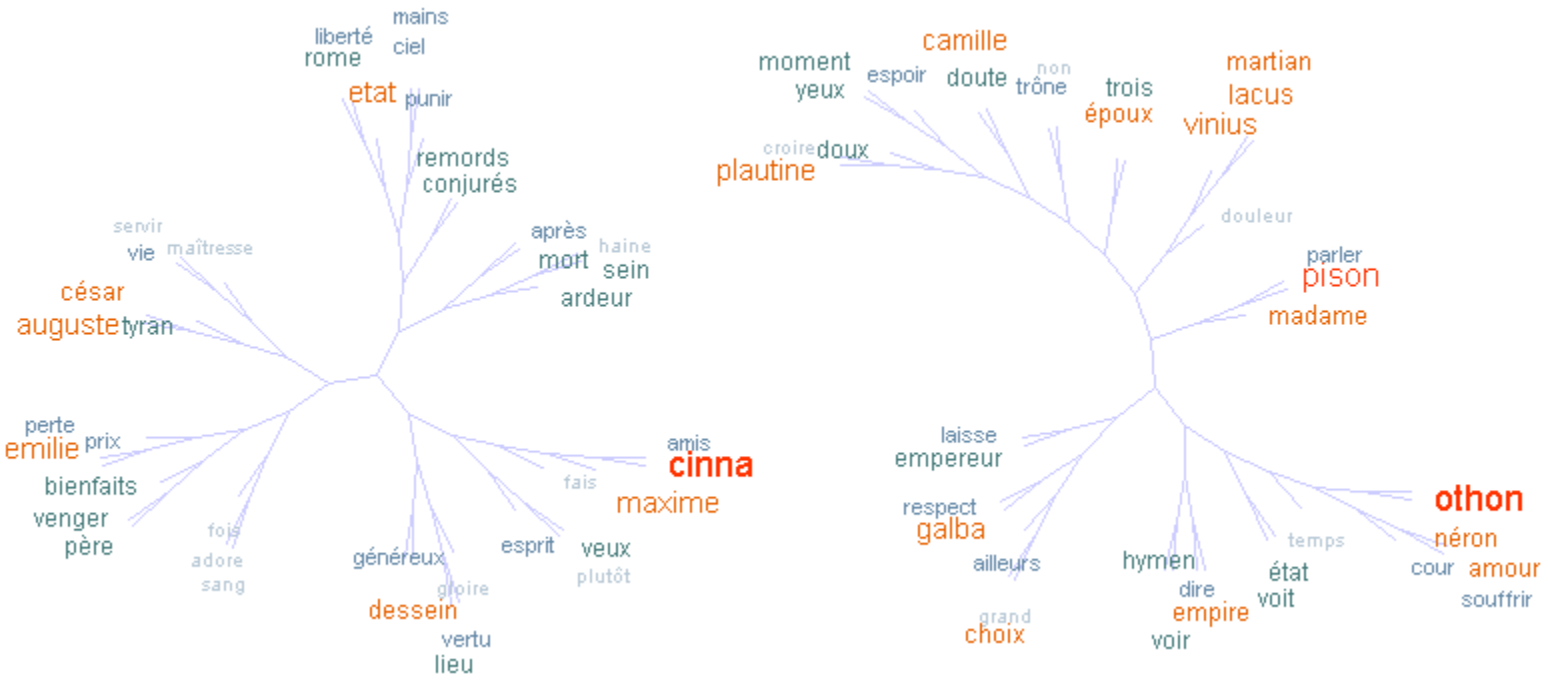


- Identifier des regroupements communs ou différents
- Comparer la fréquence des mots fréquents dans les deux arbres
- Comparer (quanti+quali) les regroupements dans les deux arbres
- **Construire le nuage arboré des mots spécifiques de chaque sous-corpus**

Travail en cours avec Edna Hernandez

Méthode : comparaison des spécifiques

Amstutz & Gambette,
JADT 2010

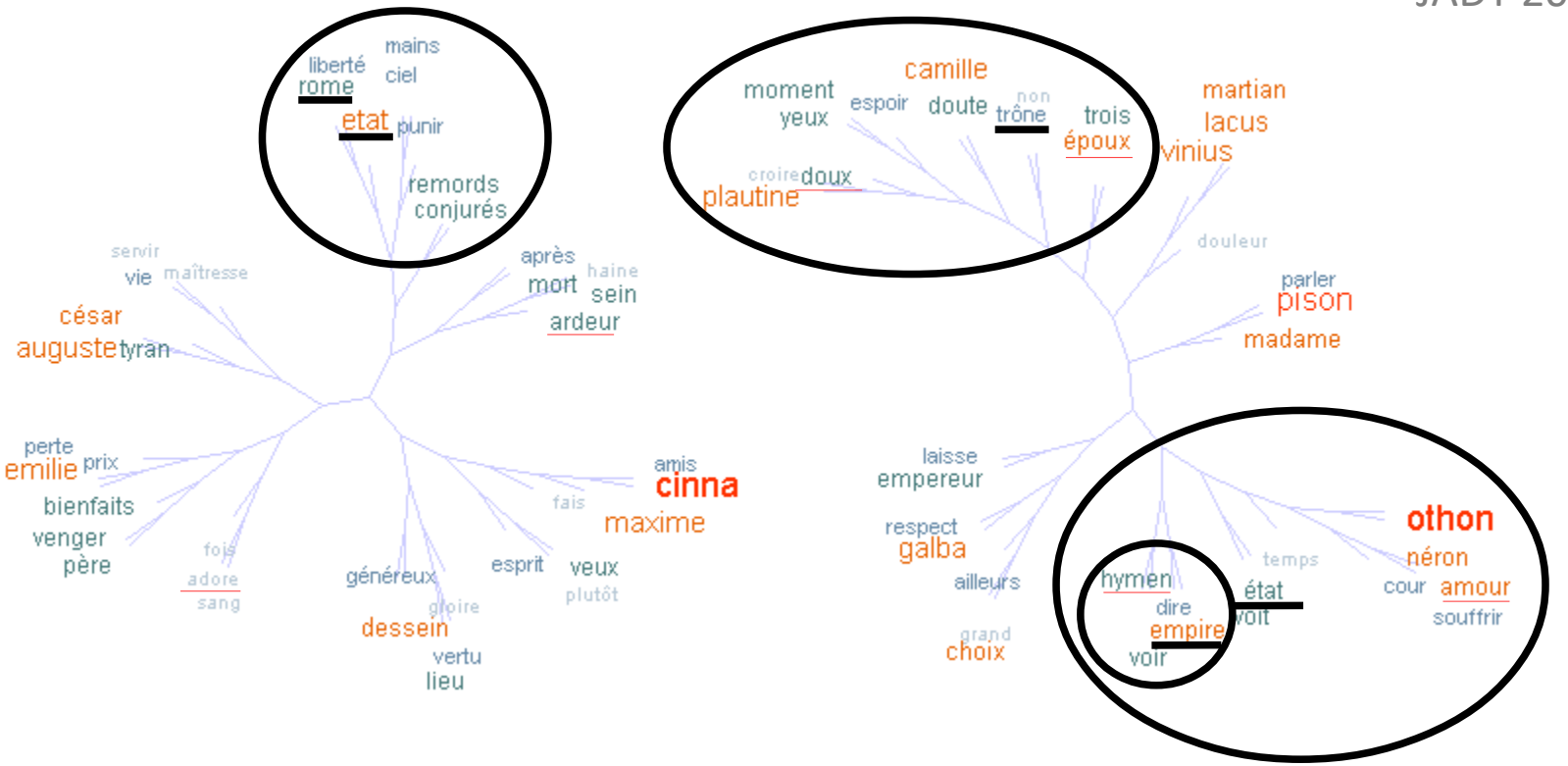


*Nuages arborés des **mots spécifiques** de Cinna et Othon, dimensionnés et colorés d'après leur spécificité calculée dans Lexico3.*

Quels moyens au service de la cause politique ?

Méthode : comparaison des spécifiques

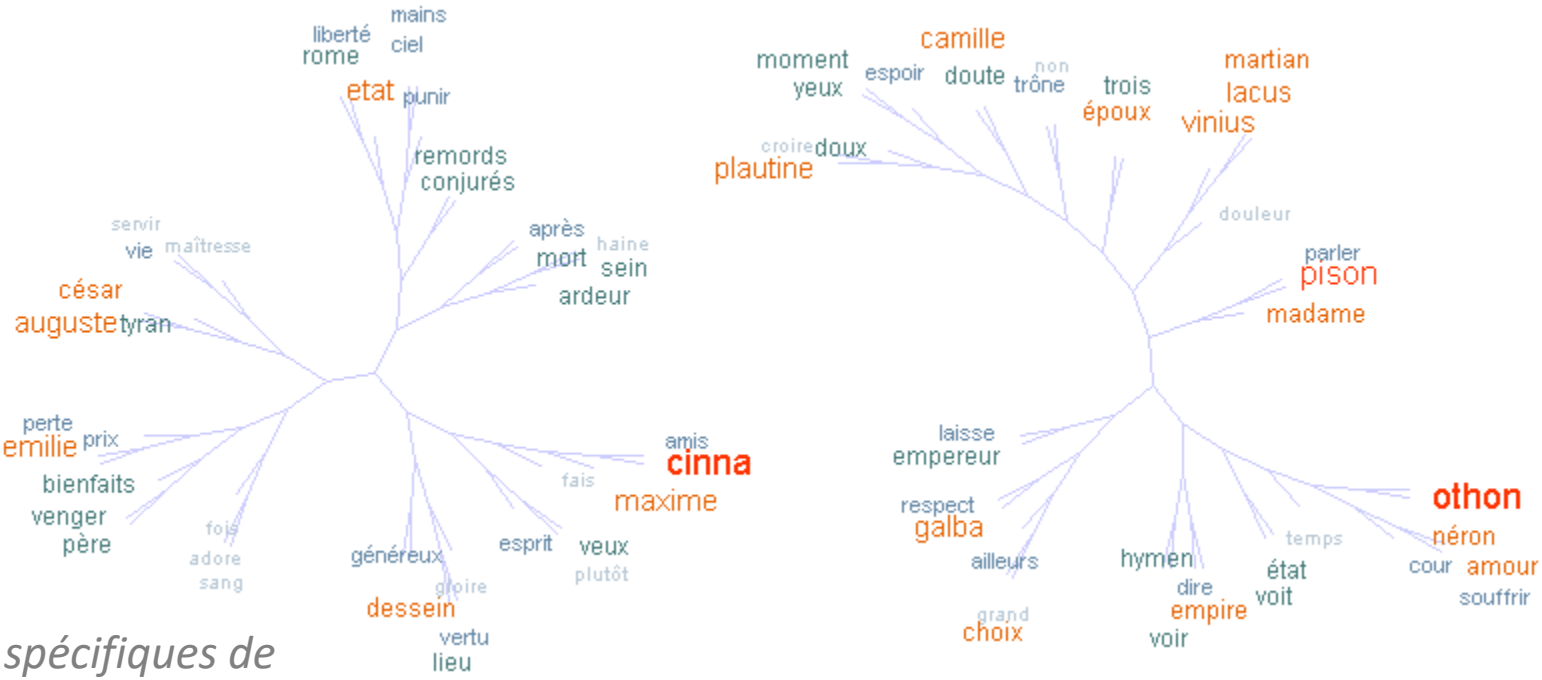
Amstutz & Gambette,
JADT 2010



*Nuages arborés des **mots spécifiques** de Cinna et Othon, dimensionnés et colorés d'après leur spécificité calculée dans Lexico3.*

Quels moyens au service de la cause politique ?

Méthode : comparaison des spécifiques

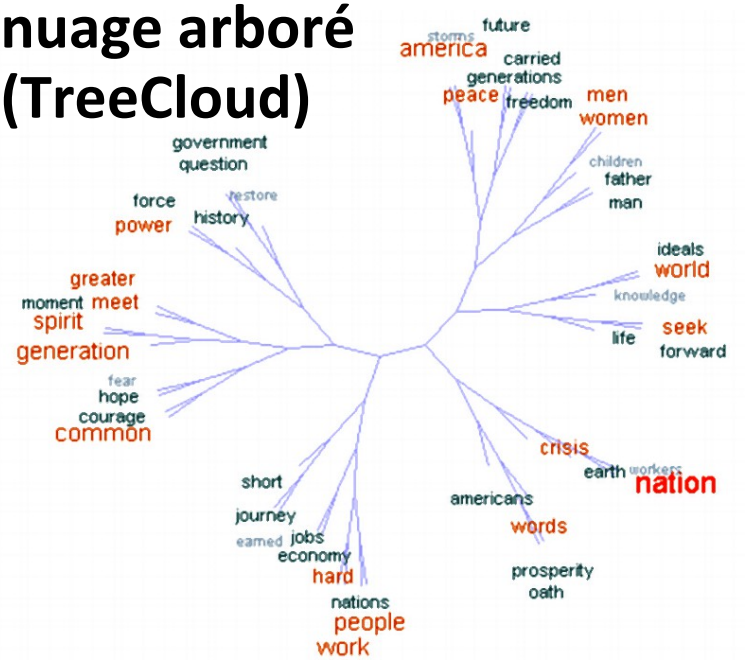


mots spécifiques de Cinna et Othon d'après Lexico3

	Cinna	*Othon*
Lieu du pouvoir et objet de la confrontation entre les personnages	Rome (« liberté »)	Empire (« trône »)
Souverain en place	tyran	Empereur
Membres du corps politique	amis	maîtres / seigneurs
Moyens au service de la cause politique	gloire	amour matrimonial (« amour », « hymen », « choix »)
Caractérisation de la pièce	Pièce de FONDATION	Pièce de SUCCESSION DYNASTIQUE

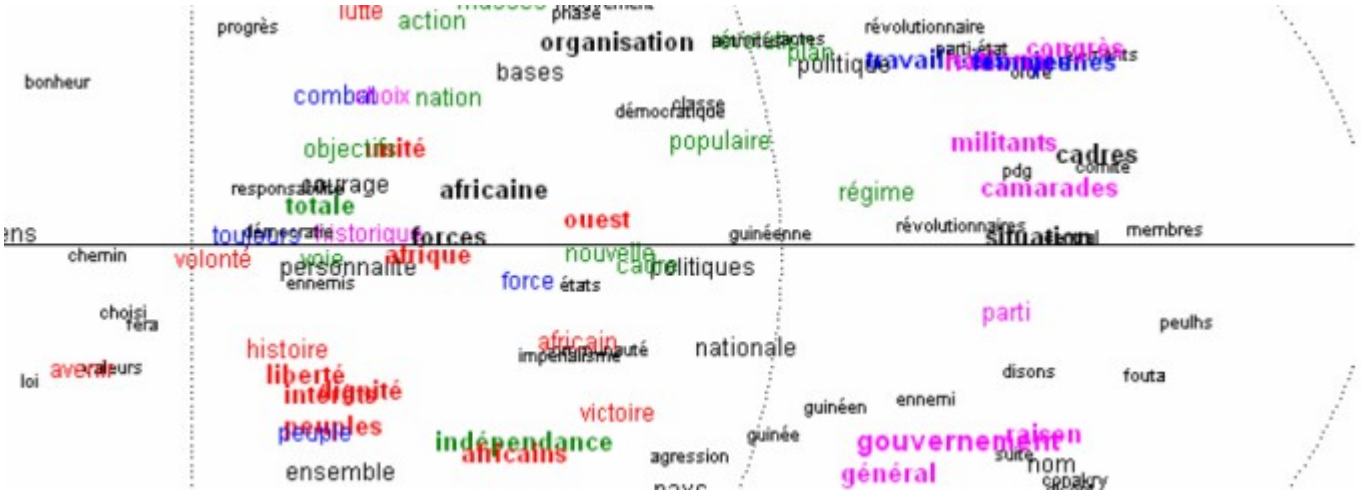
Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)



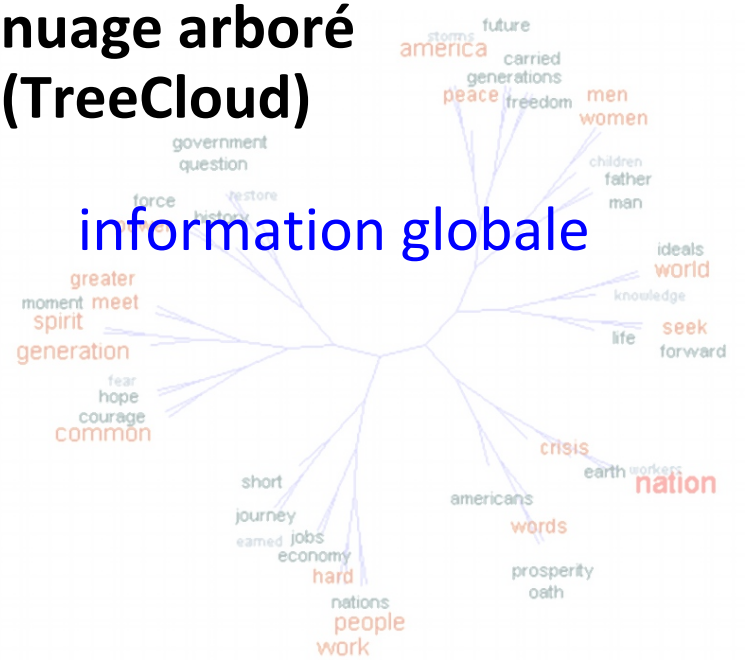
réseau de mots (PhraseNet d'IBM ManyEyes)

projection des mots (Astartex)



Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)

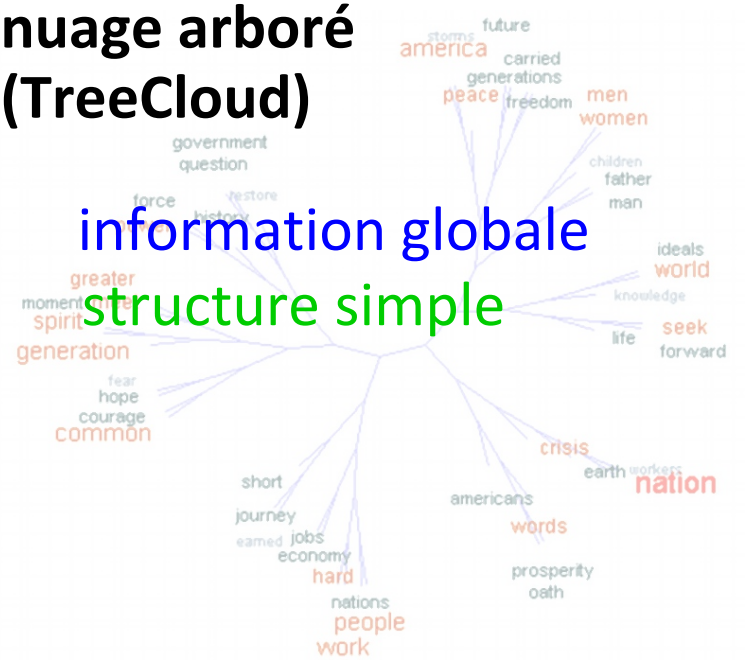


réseau de mots (PhraseNet d'IBM ManyEyes)

projection des mots (Astartex)

Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)



information globale

structure simple

réseau de mots (PhraseNet d'IBM ManyEyes)

information locale

structure complexe

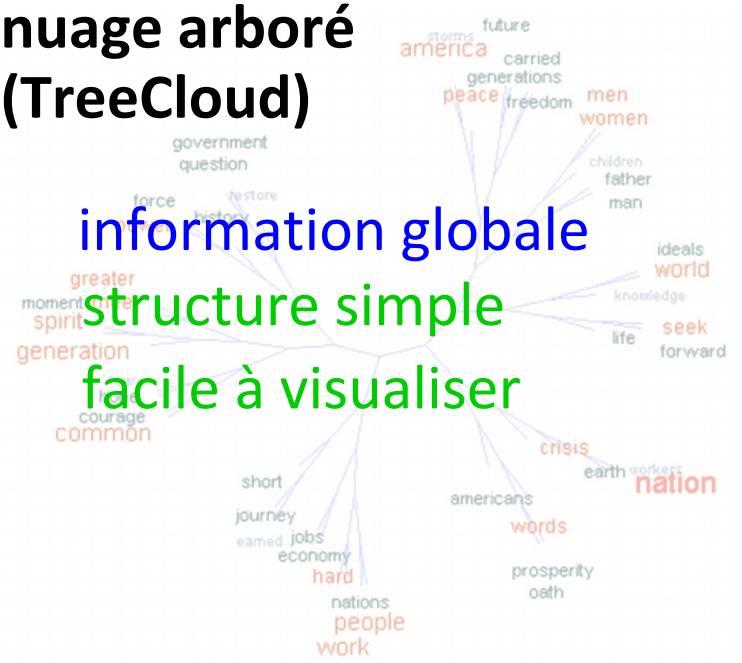
projection des mots (Astartex)

information globale

structure complexe

Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)



information globale

structure simple

facile à visualiser

réseau de mots (PhraseNet d'IBM ManyEyes)

information locale

structure complexe

difficile à visualiser

projection des mots (Astartex)

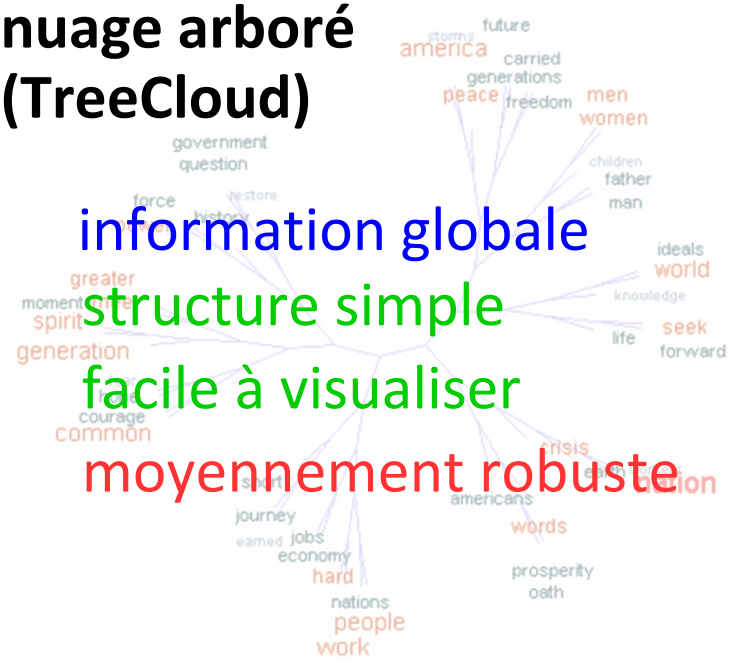
information globale

structure complexe

facile à visualiser mais chevauchements

Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)



information globale

structure simple

facile à visualiser

moyennement robuste

réseau de mots (PhraseNet d'IBM ManyEyes)

information locale

structure complexe

difficile à visualiser

robuste

projection des mots (Astartex)

information globale

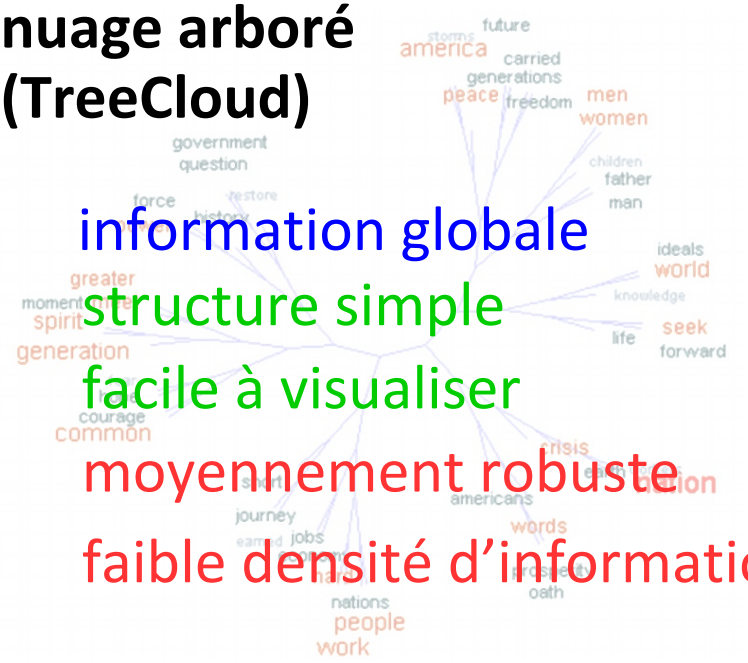
structure complexe

facile à visualiser mais chevauchements

robuste

Comparaison avec d'autres visualisations

nuage arboré (TreeCloud)



information globale

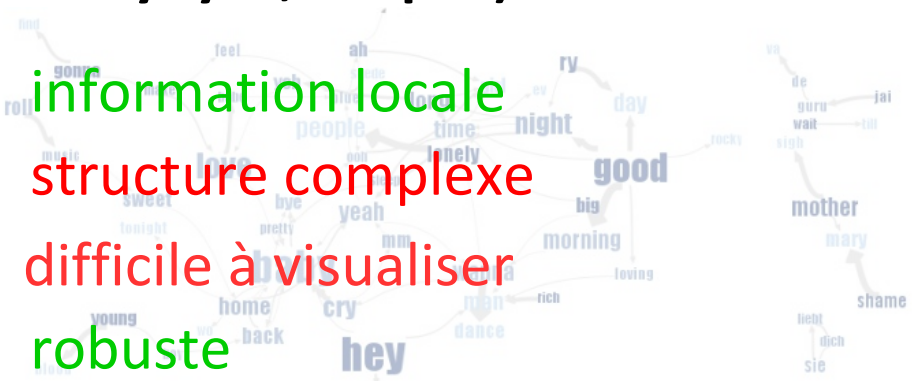
structure simple

facile à visualiser

moyennement robuste

faible densité d'information

réseau de mots (PhraseNet d'IBM ManyEyes, Tropes)



information locale

structure complexe

difficile à visualiser

robuste

forte densité d'information

projection des mots (Astartex, AFC sur Lexico, TextObserver, etc.)

information globale

structure complexe

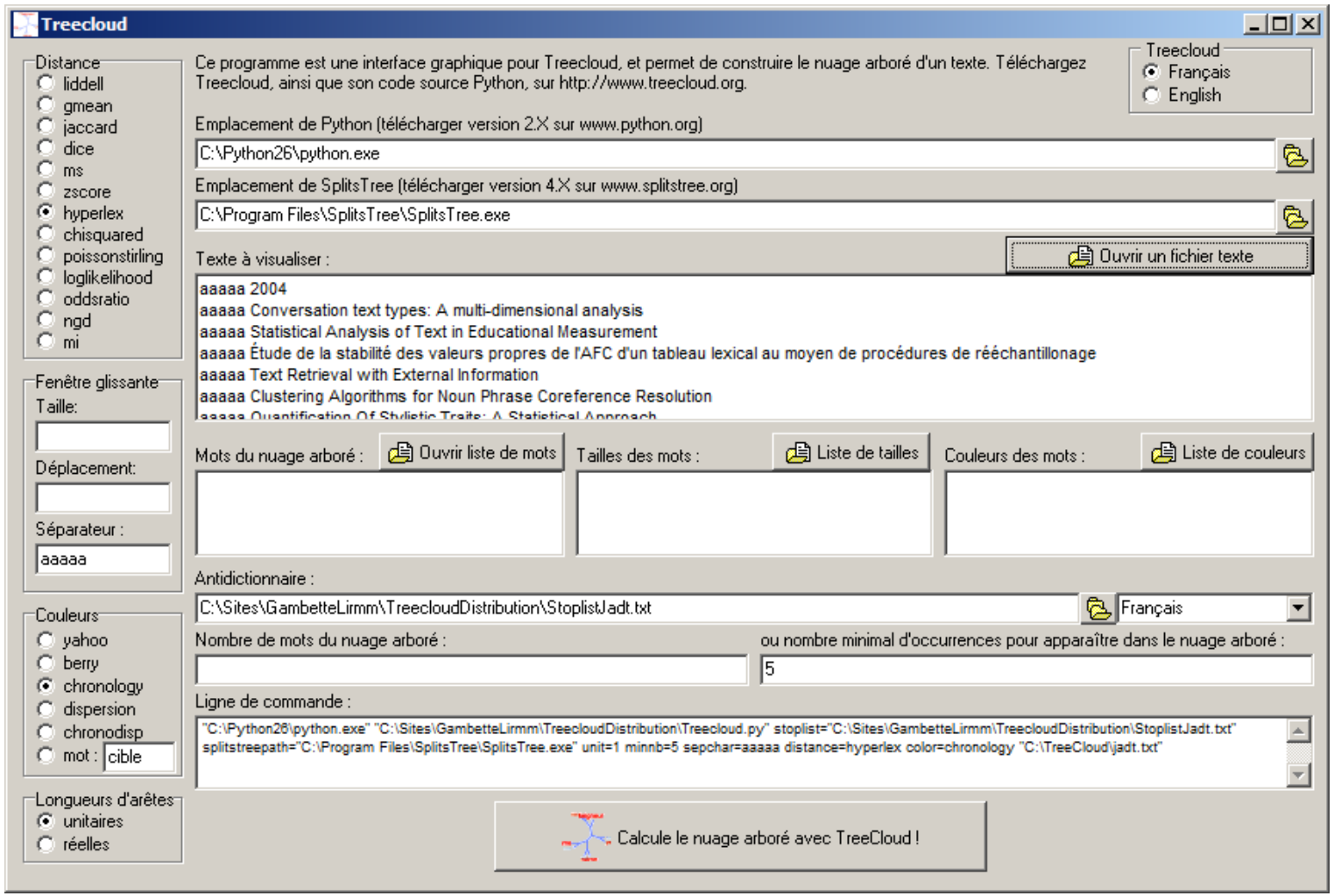
facile à visualiser mais chevauchements

robuste

forte densité d'information

Implémentations

Logiciel libre TreeCloud (Python/Delphi) + SplitsTree (Java)

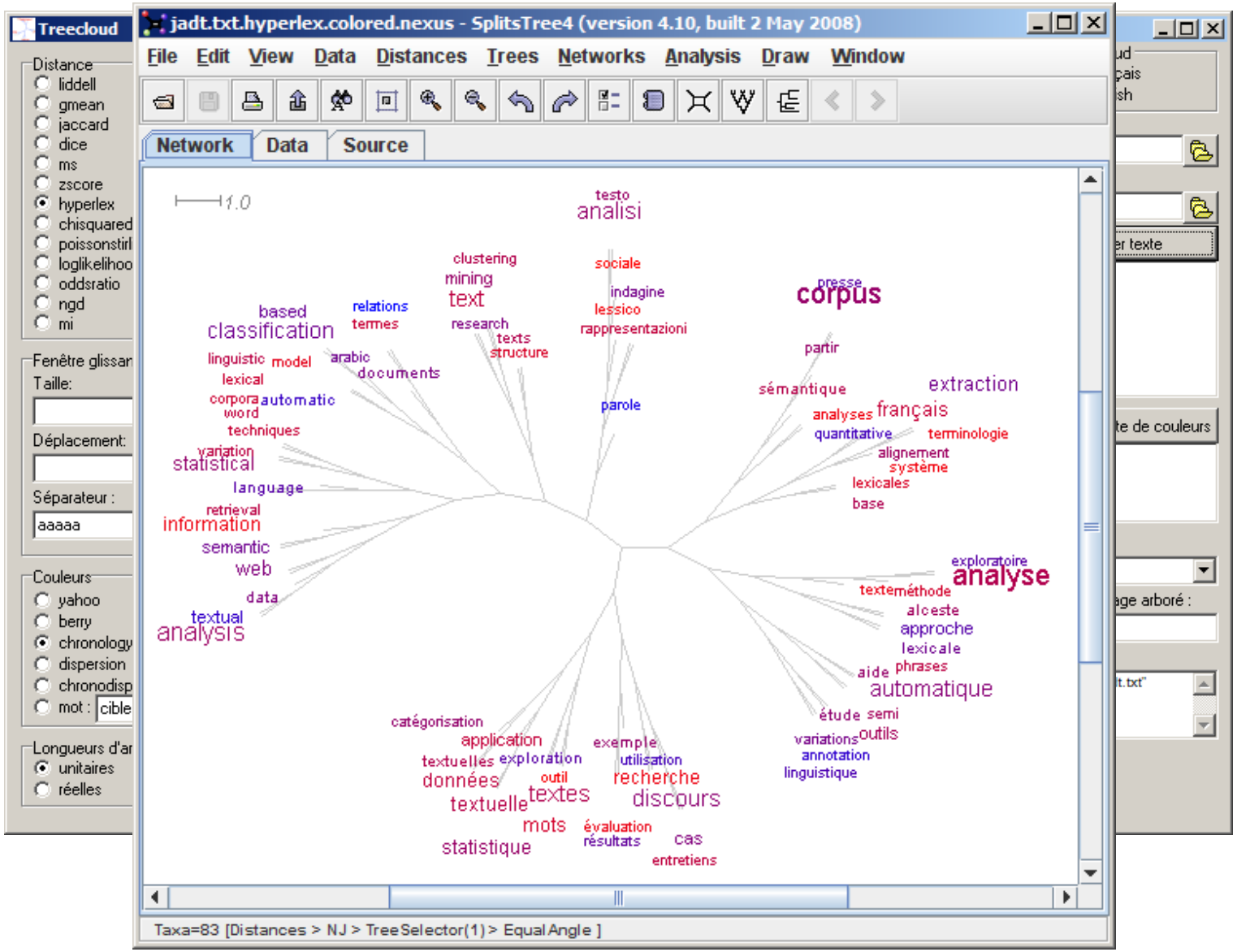


www.treecloud.org

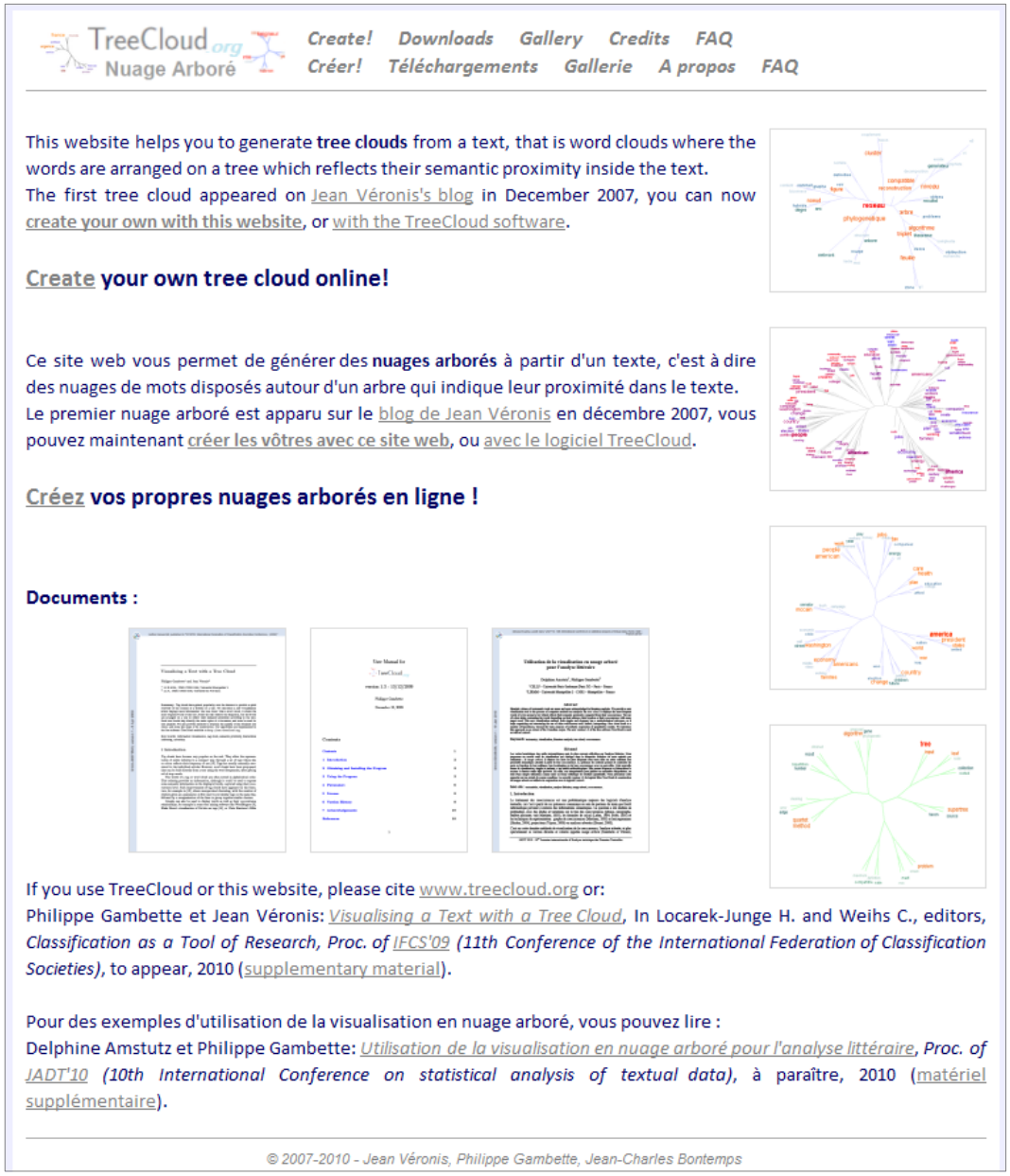
www.splitstree.org

Implémentations

Logiciel libre TreeCloud (Python/Delphi) + SplitsTree (Java)



Interface web



www.treecloud.org

Interface basée sur le logiciel libre NuageArboré de Jean-Charles Bontemps, en C, CGI/Python, et JavaScript.

http://sourceforge.net/projects/nuagearbor/

Développements supplémentaires avec d3.js par Deepak Srinivas

Interface web

www.treecloud.org

TreeCloud org
Nuage Arboré

Create! *Downloads* *Gallery* *Credits* *FAQ*
Créer! *Téléchargements* *Gallerie* *A propos* *FAQ*

This website helps you to generate **tree clo**
words are arranged on a tree which reflects
The first tree cloud appeared on Jean Vé
create your own with this website, or with t

Create **your own tree cloud online**

Ce site web vous permet de générer des r
des nuages de mots disposés autour d'un ar
Le premier nuage arboré est apparu sur le
pouvez maintenant créer les vôtres avec ce

Créez **vos propres nuages arborés**

Documents :

If you use TreeCloud or this website, please
Philippe Gambette et Jean Véronis: *Visual*
Classification as a Tool of Research, Proc. c
Societies), to appear, 2010 (supplementary r

Pour des exemples d'utilisation de la visual
Delphine Amstutz et Philippe Gambette: *Ut*
JADT'10 (10th International Conference
supplémentaire).

TreeCloud org
Nuage Arboré

Créer! *Téléchargements* *Gallerie* *A propos* *FAQ*

Créez vos propres nuages arborés !

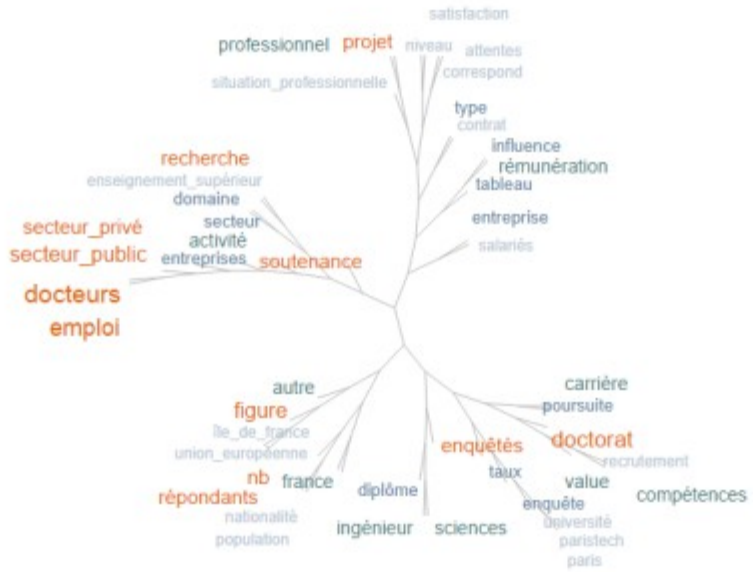
Collez votre texte dans le cadre ci-dessous, puis cliquez sur *Envoyer* ! Attention, l'utilisateur suivant verra votre texte quand il se connectera au site, si vous ne voulez pas faire apparaître vos textes, installez plutôt TreeCloud sur votre machine.

Texte :

[Texte extrait de http://www.adoc-tm.com/2013rapport.pdf]

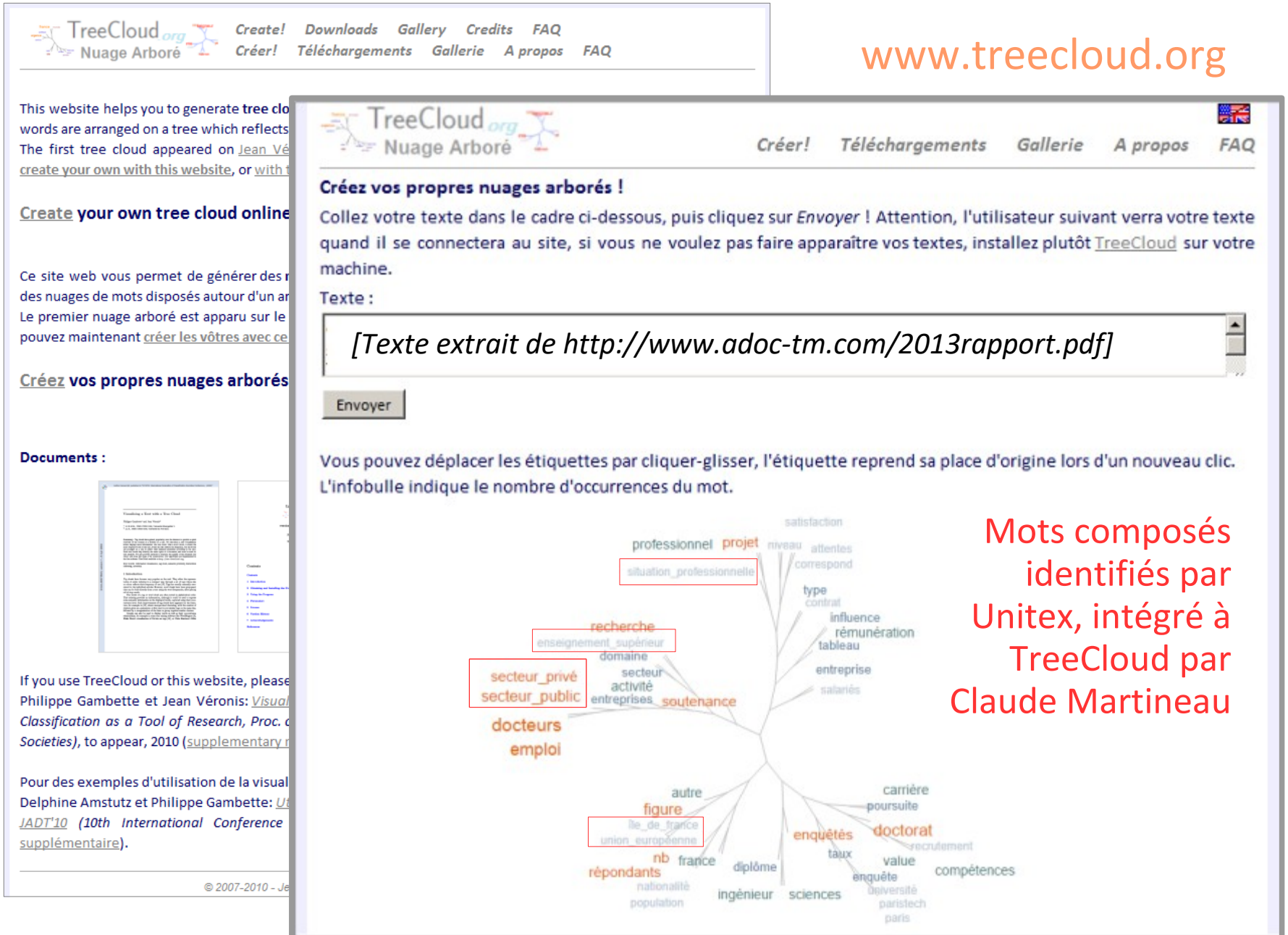
Envoyer

Vous pouvez déplacer les étiquettes par cliquer-glisser, l'étiquette reprend sa place d'origine lors d'un nouveau clic. L'infobulle indique le nombre d'occurrences du mot.



Interface web

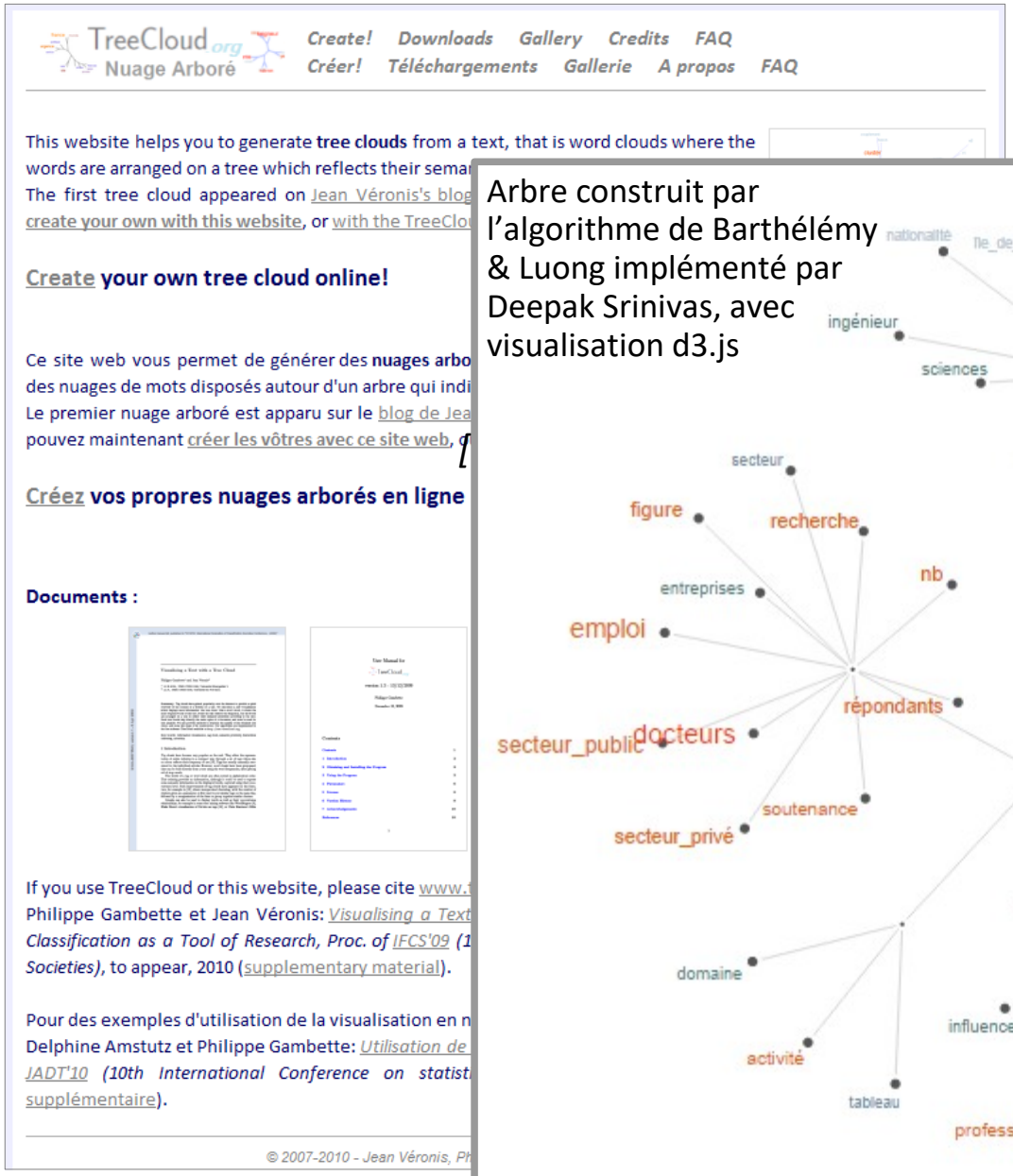
www.treecloud.org



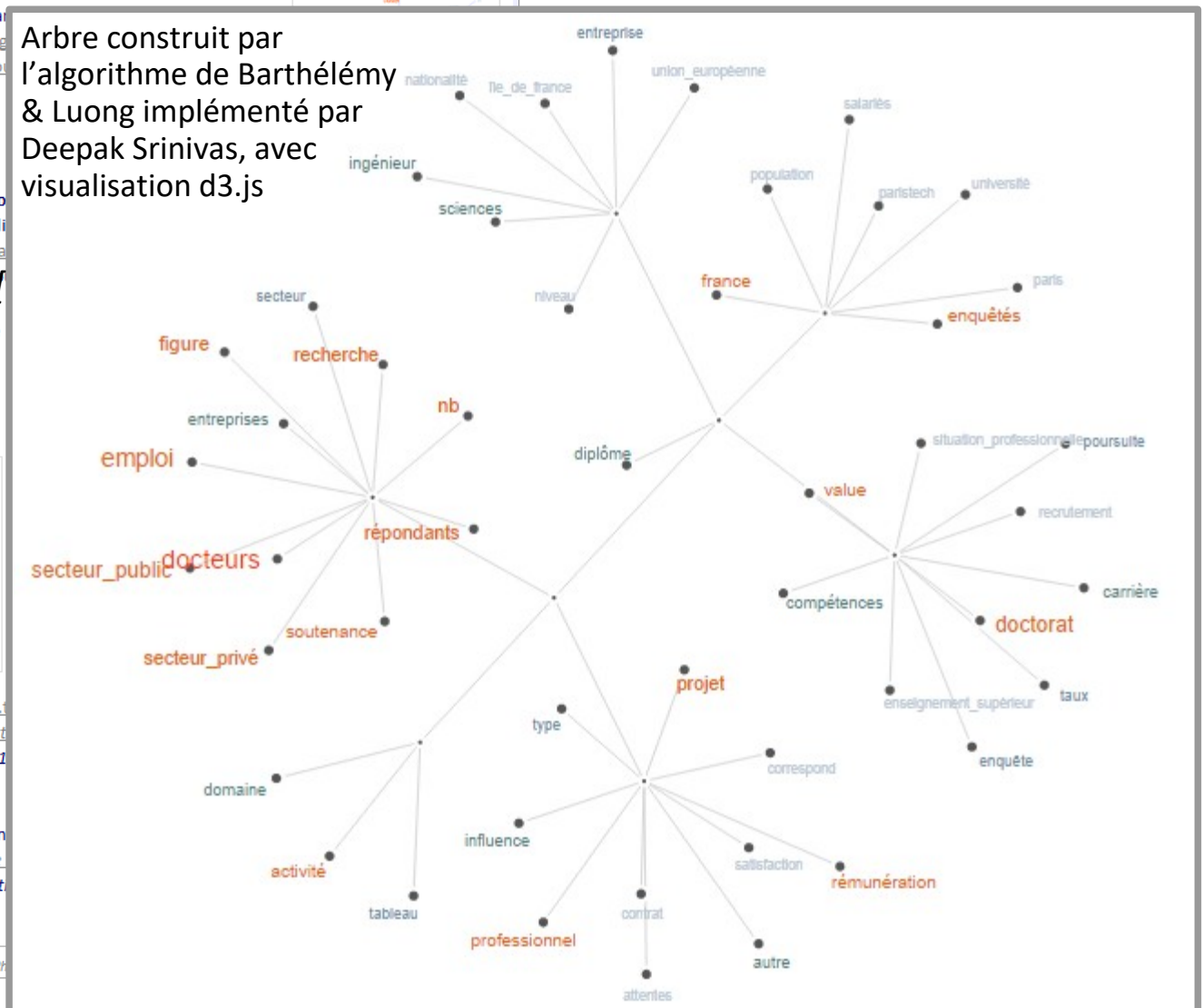
Mots composés identifiés par Unitex, intégré à TreeCloud par Claude Martineau

Interface web

www.treecloud.org

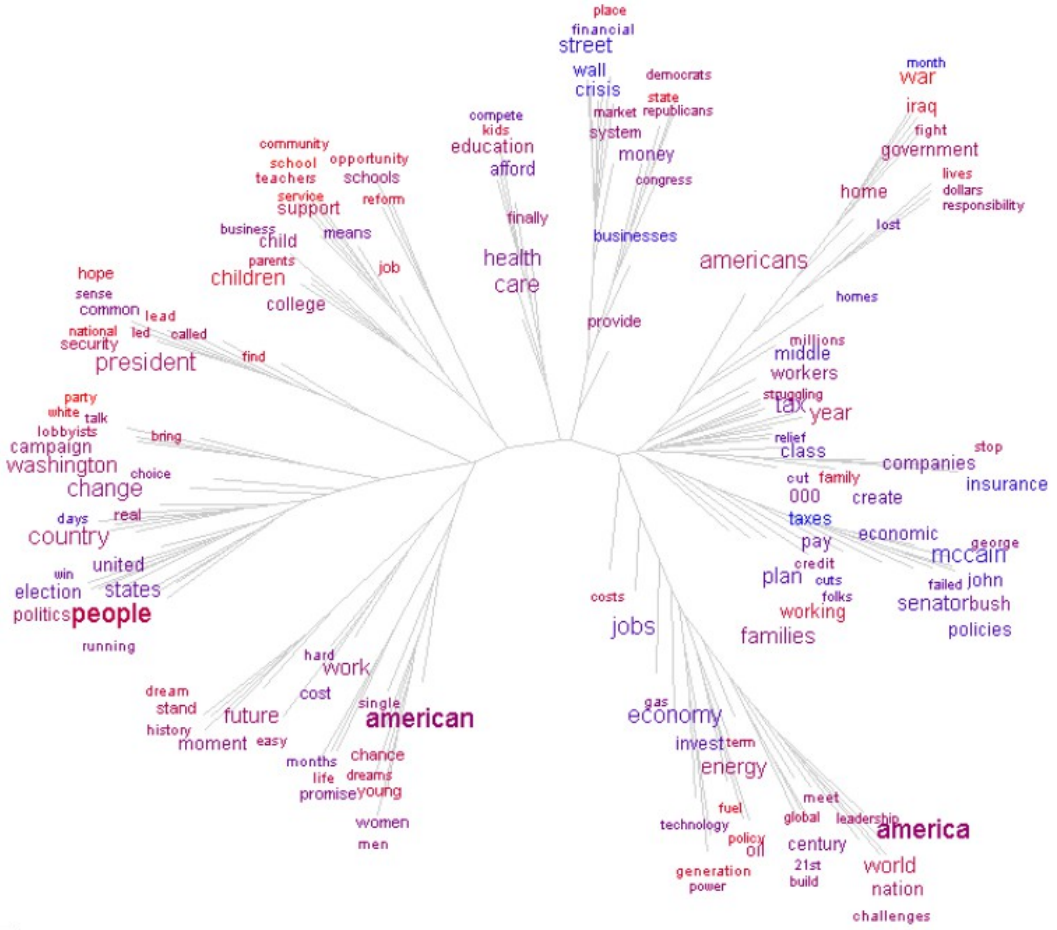


Arbre construit par l'algorithme de Barthélémy & Luong implémenté par Deepak Srinivas, avec visualisation d3.js



Temps d'exécution

Limites sur la taille du corpus pour utiliser TreeCloud ?



30 secondes pour la construction du nuage arboré de l'ensemble des discours de campagne de Barack Obama (>300 000 mots)

Références (*treecloud.org*)

Philippe Gambette, Jean Véronis (2009)
Visualising a Text with a Tree Cloud, *IFCS'09, Studies in Classification, Data Analysis, and Knowledge Organization* 40, p. 561-570
http://igm.univ-mlv.fr/~gambette/Re20090317.pdf

Delphine Amstutz & Philippe Gambette (2010)
Utilisation de la visualisation en nuage arboré pour l'analyse littéraire, JADT'10 (Proceedings of the 10th International Conference on statistical analysis of textual data), Statistical Analysis of Textual Data, p. 227-238
http://igm.univ-mlv.fr/~gambette/Re20100611.pdf

Philippe Gambette, Nuria Gala & Alexis Nasr (2012)
Longueur de branches et arbres de mots, *Corpus* 11:129-146
http://igm.univ-mlv.fr/~gambette/Re20120209.pdf

William Martinez & Philippe Gambette (2013)
L'affaire du Médiator au prisme de la textométrie, *Texto!* XVIII(4)
http://www.revue-texto.net/index.php?id=3318

Philippe Gambette, Hilde Eggermont & Xavier Le Roux (2014)
Temporal and geographical trends in the type of biodiversity research funded on a competitive basis in European countries, *rapport BiodivERsa*
http://www.biodiversa.org/700/download

Nadège Lechevrel & Philippe Gambette (2016)
Une approche textométrique pour étudier la transmission des savoirs biologiques au XIXe siècle, *Nouvelles perspectives en sciences sociales* 12(1):221-253
https://hal-upec-upem.archives-ouvertes.fr/hal-01408455

Philippe Gambette, Tita Kyriacopoulou, Nadège Lechevrel & Claude Martineau (2017)
Anatomie, animaux, vocabulaire de la vivisection : construire des ressources lexicales pour visualiser une thématique dans un corpus littéraire, *Colloque AnimalHumanité, Expérimentation et fiction : l'animalité au cœur du vivant*, décembre 2016
https://hal-upec-upem.archives-ouvertes.fr/hal-01609198

Claude Martineau (2017)
TreeCloud, Unitex: increased synergy, *ECLAVIT Workshop*, 24 novembre 2017
https://hal-upec-upem.archives-ouvertes.fr/hal-01702091

Atelier

Tutoriel :
https://docs.google.com/document/d/1OauE9EflJTyVR3gM7ZPc3cGJ3-N0Iq2ghPD5RNrb_YY/edit?usp=sharing

Possibilités de **réalisations numériques sur vos corpus** :

- en ajoutant des liens sur les mots du nuage arboré, avec *TreeCloud Linker* :
  - Mots-clés des publications de l'UPEM :
    http://treecloud.univ-mlv.fr/treecloud-linker/
  - Collections du musée Fragonard de l'école vétérinaire d'Alfort :
    http://treecloud.univ-mlv.fr/treecloud-linker/fragonard.html

- en les chargeant dans *TreeCloud Corpus* :
  - *Vœux présidentiels* rassemblés par Jean-Marc Leblanc, 1960-2018
    http://treecloud.univ-mlv.fr/treecloud-voeux/
  - *Lettres républicaines* de Daniel Stern (pseudonyme de Marie d'Agoult), 1848
    http://treecloud.univ-mlv.fr/treecloud-corpus/lettres-republicaines/