



HAL
open science

L'exercice de "Revue de littérature" à l'aide d' IRAMUTEQ

Gaël Plumecocq

► **To cite this version:**

Gaël Plumecocq. L'exercice de "Revue de littérature" à l'aide d' IRAMUTEQ. Séminaire du Lereps, Mar 2017, Toulouse, France. 21 p. hal-01603181

HAL Id: hal-01603181

<https://hal.science/hal-01603181>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

L'exercice de "revue de littérature" à l'aide d'IRaMuTeQ :



1 Lexicométrie

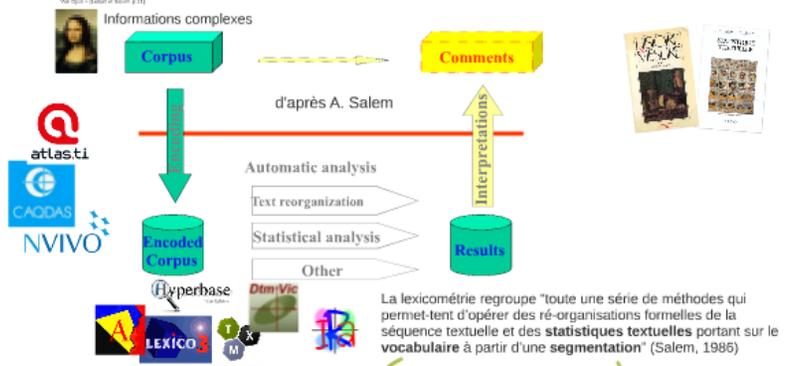


Table lexicale

Segmenter (nominal) : unités de contenu : mots, phrases, sections de texte.

Vocabulaire	Nombre d'occurrences														
	Absence (0) / Présence (1)														
Paris	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Parisien	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Le pont	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Château	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Le pont	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Château	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Le pont	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Château	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Le pont	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Château	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Chaque "segment" est alors rattaché à un vocabulaire spécifique.

Le reste : Classe 1 (17) + 178
Classe 2 (11) + 116

⚠ CE SONT LES SEGMENTS QUI SONT REGROUPÉS (ET PAS LE VOCABULAIRE)
Regroupement de textes qui utilisent des vocabulaires semblables

2 IRaMuTeQ

3 WEB OF SCIENCE

- permet de travailler sur un gros corpus
 - permet d'introduire des variables
 - permet de regrouper des abstracts en fonction de thématiques
- Se poser la question de la requête
- Se poser la question des bonnes variables
- Renvoyer à un nombre restreint de modalités
 - Etre pertinentes par rapport à la question de la recherche
 - Année : Certaines thématiques apparaissent-elles (ou disparaissent-elles) avec le temps ?
 - Nb de citations : certaines thématiques sont-elles plus pertinentes que les autres ?
 - Journaux/Discipline : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?
- Se poser la question de l'unité pertinente d'analyse

4 Un exemple



- Intérêts pour nous**
- Situer notre approche dans le paysage des EMC
 - Cibler une littérature précise pour identifier des trous dans la littérature
 - Caractériser les outils en lien avec les approches
 - Permet de montrer qu'il n'existe pas d'outils complètement adaptés à notre approche.



De toute façon : toujours revenir aux textes



L'exercice de "revue de littérature" à l'aide d'IRaMuTeQ :



1 Lexicométrie

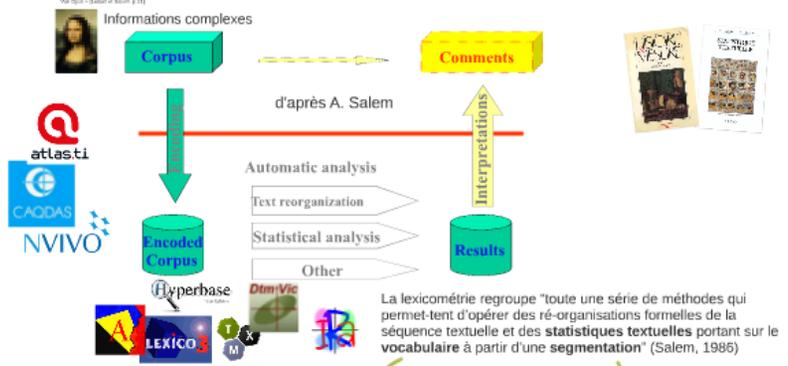


Table lexicale

Segmenter (nominal) : unités de contenu : mots, phrases, sections de texte.

Vocabulaire	Nombre d'occurrences														
	Absence (0) / Présence (1)														
Paris	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Parisien	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Le petit chat est blanc	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Chaque dimanche	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Il est temps de partir	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Il est si agréable	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Il est si agréable	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Chaque "segment" est alors rattaché à un vocabulaire spécifique

Le reste (17) / 173

Le mot "chat"

Chaque à (17) / 173

⚠ CE SONT LES SEGMENTS QUI SONT REGROUPÉS (ET PAS LE VOCABULAIRE)

Regroupement de textes qui utilisent des vocabulaires semblables

2 IRaMuTeQ

3 WEB OF SCIENCE

- permet de travailler sur un gros corpus
 - permet d'introduire des variables
 - permet de regrouper des abstracts en fonction de thématiques
- Se poser la question de la requête
- Se poser la question des bonnes variables
- Renvoyer à un nombre restreint de modalités
 - Etre pertinentes par rapport à la question de la recherche
 - Année : Certaines thématiques apparaissent-elles (ou disparaissent-elles) avec le temps ?
 - Nb de citations : certaines thématiques sont-elles plus pertinentes que les autres ?
 - Journaux/Discipline : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?
- Se poser la question de l'unité pertinente d'analyse

4 Un exemple



- Intérêts pour nous**
- Situer notre approche dans le paysage des EMC
 - Cibler une littérature précise pour identifier des trous dans la littérature
 - Caractériser les outils en lien avec les approches
 - Permet de montrer qu'il n'existe pas d'outils complètement adaptés à notre approche.



De toute façon : toujours revenir aux textes



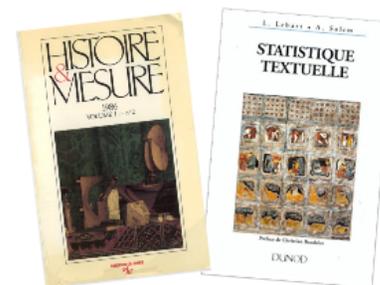


Lexicométrie

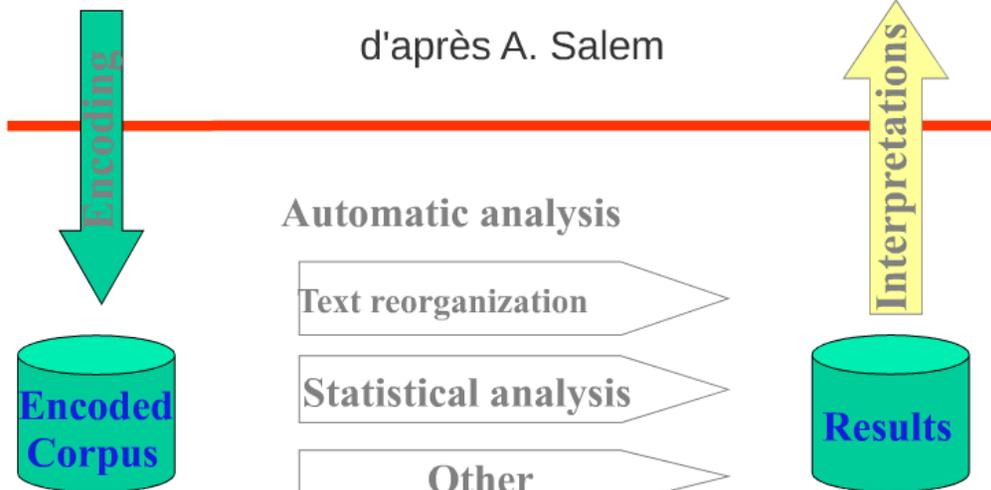
« Supposons (...) que l'on étudie les histogrammes des longueurs d'ondes correspondant aux couleurs d'un tableau de Rembrandt (pour chacun des pixels d'une reproduction). Il va de soi que l'on utilise une fraction dérisoire de l'information contenue dans l'image d'origine. Il est cependant possible que la forme de l'histogramme (ou d'une fonction plus élaborée des mêmes mesures et données de base) permette de distinguer un Rembrandt d'un Rubens ou d'un Van Dyck » (Lebart et Salem, p.21).



Informations complexes



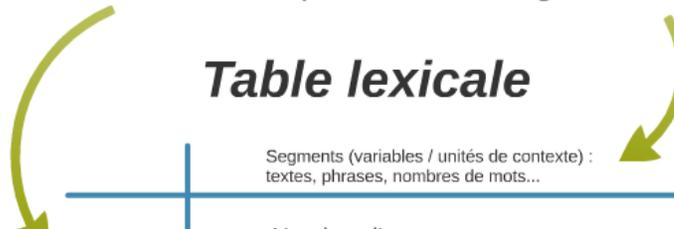
d'après A. Salem



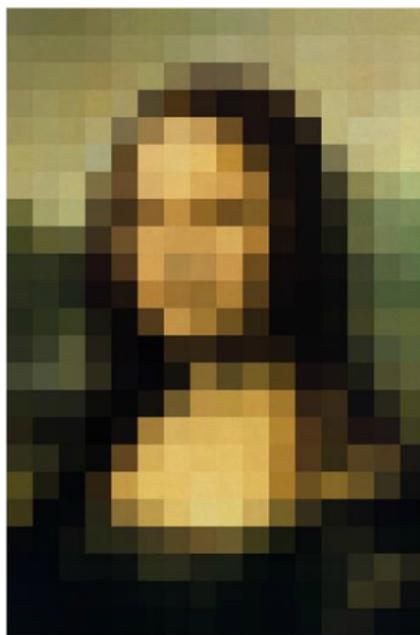
La lexicométrie regroupe "toute une série de méthodes qui permet-tent d'opérer des ré-organisations formelles de la séquence textuelle et des **statistiques textuelles** portant sur le **vocabulaire** à partir d'une **segmentation**" (Salem, 1986)

Table lexicale

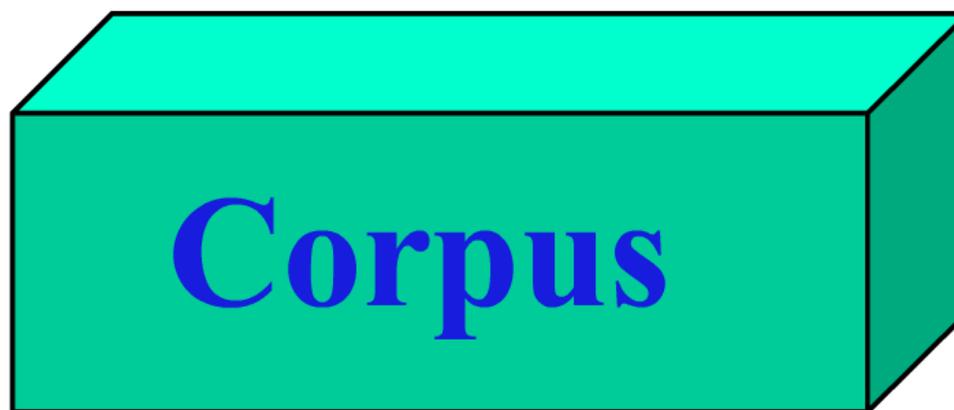
Segments (variables / unités de contexte) :
textes, phrases, nombres de mots...



« Supposons (...) que l'on étudie les histogrammes des longueurs d'ondes correspondant aux couleurs d'un tableau de Rembrandt (pour chacun des pixels d'une reproduction). Il va de soi que l'on utilise une fraction dérisoire de l'information contenue dans l'image d'origine. Il est cependant possible que la forme de l'histogramme (ou d'une fonction plus élaborée des mêmes mesures et données de base) permette de distinguer un Rembrandt d'un Rubens ou d'un Van Dyck » (Lebart et Salem, p.21).



Informations complexes



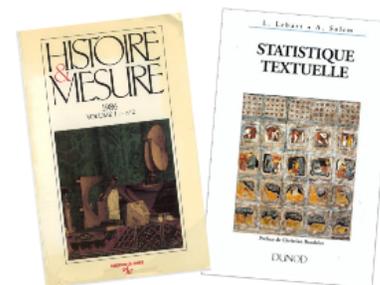


Lexicométrie

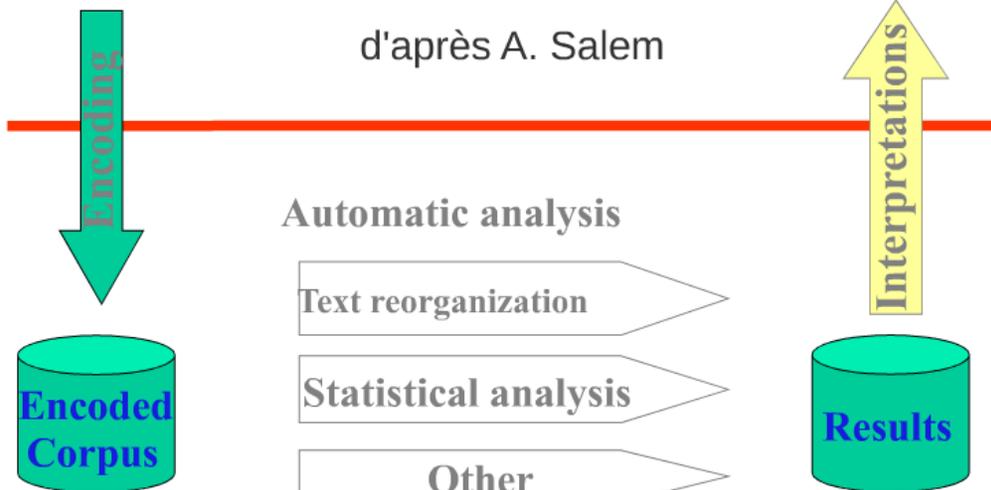
« Supposons (...) que l'on étudie les histogrammes des longueurs d'ondes correspondant aux couleurs d'un tableau de Rembrandt (pour chacun des pixels d'une reproduction). Il va de soi que l'on utilise une fraction dérisoire de l'information contenue dans l'image d'origine. Il est cependant possible que la forme de l'histogramme (ou d'une fonction plus élaborée des mêmes mesures et données de base) permette de distinguer un Rembrandt d'un Rubens ou d'un Van Dyck » (Lebart et Salem, p.21).



Informations complexes



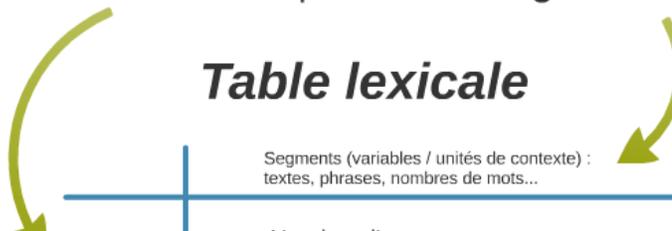
d'après A. Salem

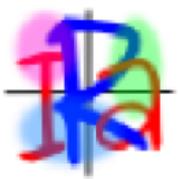


La lexicométrie regroupe "toute une série de méthodes qui permettent d'opérer des ré-organisations formelles de la séquence textuelle et des **statistiques textuelles** portant sur le **vocabulaire** à partir d'une **segmentation**" (Salem, 1986)

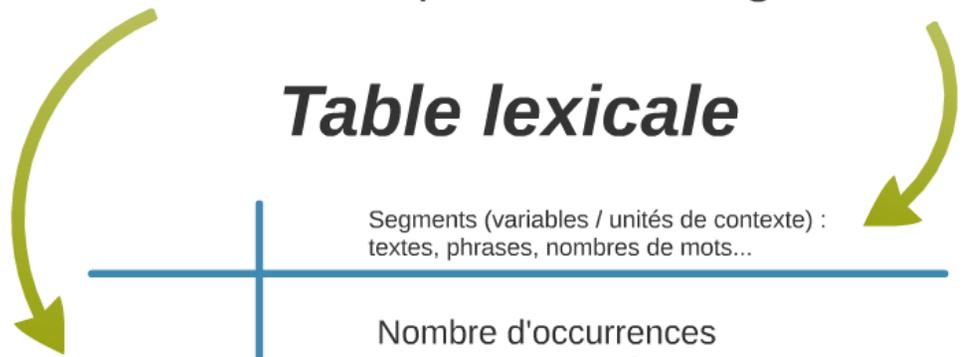
Table lexicale

Segments (variables / unités de contexte) :
textes, phrases, nombres de mots...





La lexicométrie regroupe "toute une série de méthodes qui permet-tent d'opérer des ré-organisations formelles de la séquence textuelle et des **statistiques textuelles** portant sur le **vocabulaire** à partir d'une **segmentation**" (Salem, 1986)



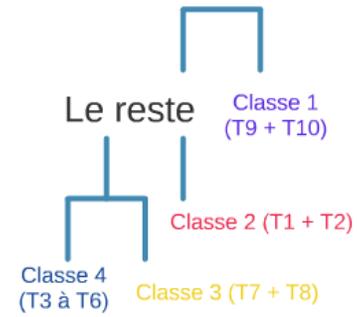
Vocabulaire :
mots
(lemmatisation)

Le petit chat est mort.
C'est dommage, il
était sympa le petit
chat. On le voyait
souvent au village...

Nombre d'occurrences
Absence (0) / Présence (1)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
petit	0	0	0	0	0	0	1	1	1	1
chat	0	0	0	0	0	0	0	0	1	1
être	0	0	1	1	1	1	1	1	1	1
mort	0	0	0	0	0	0	0	0	1	1
dommage	1	1	1	1	1	1	1	1	0	0
sympa	1	1	0	0	0	0	0	1	0	0
voir	1	1	1	1	1	1	0	0	0	0
souvent	1	1	0	0	0	0	0	0	0	1
village	0	1	0	0	0	1	1	0	1	0

Chaque "classe" de textes renvoie à un **vocabulaire spécifique**



Test du chi-deux pour chaque mot :

- ce mot est-il sur-représenté dans la classe en question, ou dans le reste du corpus ?

Le mot "chat" :

- présent dans 100 % des textes de la classe 1 (et dans 25 % du reste)
- présent dans 0 % des textes de la classe 2 (et dans 50 % du reste)
- présent dans 100 % des textes de la classe 3 (et dans 25 % du reste)
- présent dans 0 % des textes de la classe 4 (et dans 66 % du reste)



CE SONT LES SEGMENTS QUI SONT REGROUPES (ET PAS LE VOCABULAIRE)

Regroupement de textes qui utilisent des vocabulaires semblables



IRaMuTeQ

Table lexicale

Segments (variables / unités de contexte) :
textes, phrases, nombres de mots...



Nombre d'occurrences
Absence (0) / Présence (1)

Vocabulaire :
mots
(lemmatisation)

Le petit chat est mort.
C'est dommage, il
était sympa le petit
chat. On le voyait
souvent au village...

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
petit	0	0	0	0	0	0	1	1	1	1
chat	0	0	0	0	0	0	0	0	1	1
être	0	0	1	1	1	1	1	1	1	1
mort	0	0	0	0	0	0	0	0	1	1
dommage	1	1	1	1	1	1	1	1	0	0
sympa	1	1	0	0	0	0	0	1	0	0
voir	1	1	1	1	1	1	0	0	0	0
souvent	1	1	0	0	0	0	0	0	0	1
village	0	1	0	0	0	1	1	0	1	0

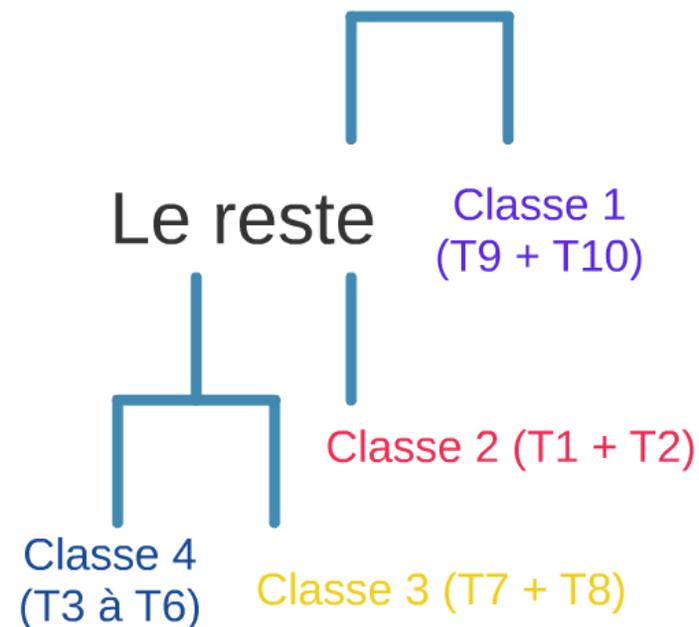


CE SONT LES SEGMENTS QUI

Nombre d'occurrences
Absence (0) / Présence (1)

Chaque "classe" de textes ren

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
petit	0	0	0	0	0	0	1	1	1	1
chat	0	0	0	0	0	0	0	0	1	1
être	0	0	1	1	1	1	1	1	1	1
mort	0	0	0	0	0	0	0	0	1	1
dommage	1	1	1	1	1	1	1	1	0	0
sympa	1	1	0	0	0	0	0	1	0	0
voir	1	1	1	1	1	1	0	0	0	0
souvent	1	1	0	0	0	0	0	0	0	1
village	0	1	0	0	0	1	1	0	1	0



CE SONT LES SEGMENTS QUI SONT
REGROUPES (ET PAS LE VOCABULAIRE)

Regroupement de textes qui utilisent des vocabulaires semblables

AMUTEQ

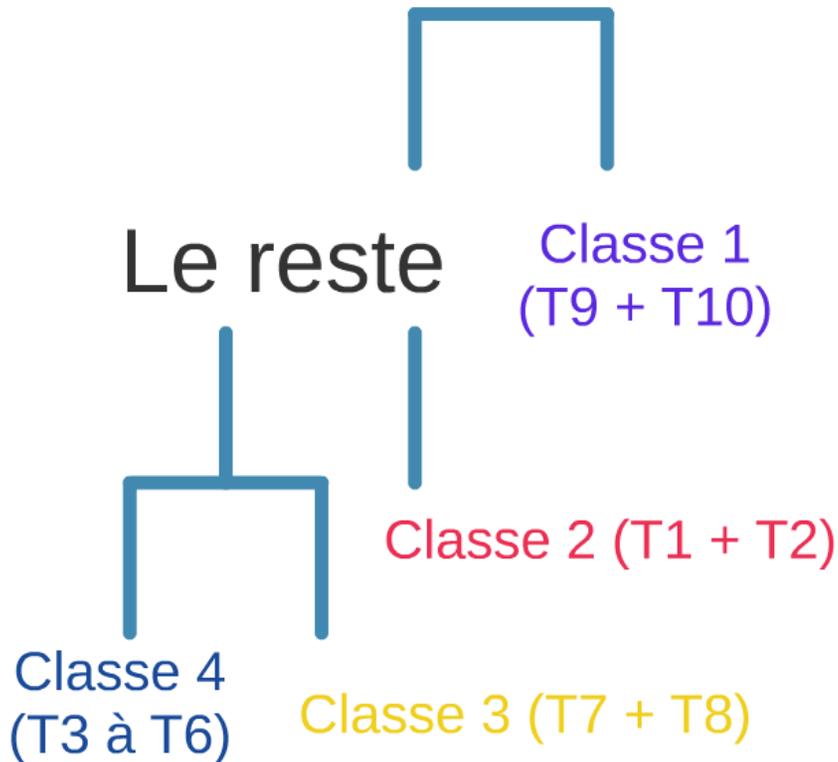
Chaque "classe" de textes renvoie à un vocable spécifique

Test du chi-deux pour chaque mot :

- ce mot est-il sur-représenté dans la classe en question, ou dans le reste du corpus ?

Le mot "chat" :

- présent dans 100 % des textes de la classe 1 (et dans 25 % du reste)
- présent dans 0 % des textes de la classe 2 (et dans 50 % du reste)
- présent dans 100 % des textes de la classe 3 (et dans 25 % du reste)
- présent dans 0 % des textes de la classe 4 (et dans 66 % du reste)



QUI SONT
(VOCABULAIRE)



WEB OF SCIENCE™

...à ça

Revue

Titre

Auteurs

Partition

Vocabulaire / Formes

Passer de ça...

- permet de travailler sur un gros corpus
- permet d'introduire des variables
- permet de regrouper des abstracts en fonction de thématiques

Se poser la question de la requête

Se poser la question des bonnes variables

- Renvoyer à un nombre restreint de modalités
- Être pertinentes par rapport à la question de la recherche
 - **Année** : Certaines thématiques apparaissent-elles (ou disparaissent-elles) avec le temps ?
 - **Nb de citations** : certaines thématiques sont-elles plus porteuses que les autres ?
 - **Journaux/Discipline** : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?

Se poser la question de l'unité pertinente d'analyse



WEB OF SCIENCE™

Revue

Incorporating stakeholder preferences for land conservation: Weights and measures in spatial MCA

By: Strager, MP (Strager, MP); Rosenberger, RS (Rosenberger, RS)

View ResearcherID and ORCID

ECOLOGICAL ECONOMICS
Volume: 58 Issue: 1 Pages: 79-92
DOI: 10.1016/j.ecolecon.2005.05.024
Published: JUN 10 2006
View Journal Information

Abstract
Spatial multicriteria models may provide an equitable and efficient means for incorporating people's preferences in social decisions. However, in order for these tools to be effective, they should include criteria that are locally relevant and measurable in a spatial framework. This paper integrates measures of stakeholder preferences with GIS data in a spatial multicriteria framework for identifying high priority areas for land conservation. Individual participants' preference weights were measured using the Analytical Hierarchy Process. Individual preferences were aggregated into groups representing outside experts and local stakeholders. Aggregate preferences differed across groups, illustrating an affinity for local knowledge of stakeholders vs. universal broader issues by outside experts. The mapping of priority areas for conservation was relatively unaffected by the weights, mostly due to the lack of spatial measures for locally relevant criteria. (c) 2005 Elsevier B.V. All rights reserved.

Keywords
Author Keywords: analytic hierarchy process; experts vs. local stakeholders; preference weights; spatial multicriteria analysis
KeyWords Plus: ATTRIBUTE DECISION-ANALYSIS; MULTICRITERIA EVALUATION; ECOSYSTEM MANAGEMENT; PUBLIC PREFERENCES; PRESERVATION; FARMLAND; SYSTEM; GOALS

Titre

Auteurs

...à ça

EE-discourse - Bloc-notes

Fichier Edition Format Affichage ?

that taste plays an important role in consumer acceptance, and the provision of nutrition information does translate into substantial premiums for the biofortified variety. there is a substantial hypothetical bias in the wtp for the new varieties, and while cheap talk mitigates this bias, it does not eliminate it.

**** *PY_2011 *P_6 *SO_AJAE

Partition

function by obtaining measures of precision of elasticities and consumer welfare loss estimates due to reductions in commercial landings in the u.s. great lakes, even with curvature constraints imposed, confidence intervals are obtained by using the bootstrap, subsample bootstrap, and subsample jackknife.

Vocabulaire / Formes

there is a broad understanding that intergenerational equity is a necessary, if not sufficient, condition for sustainability. likewise, there is a growing consensus that sustainability science requires stakeholder engagement to be successful. these two ideas demand some meaningful way of engaging the future as a stakeholder if sustainability is to be operationalized. rawls' theory of justice provides a model for how this might be accomplished, yet there are both conceptual and practical problems with a rawlsian approach. we propose using retrospective assessment as a means of learning how to approach future stakeholder engagement in sustainability.

**** *PY_2012 *P_7 *SO_EE

"the restitution of parkland to the khomani san ""bushmen"" and mier ""agricultural"" communities in may 2002 marked a significant shift in conservation in the kgalanadi transfrontier park and environs in south

Passer de ça...

V1	D	E	F	G	H	I	J	K
1	DE	GP	AF	BF	CA	TI	SD	EE
2	Cuk, M	Duhosova, Dana; Zmeskal, Zdenek	Company for MANAGING AND MODE					
3	Baharem, F; Yusuf, Y; Yusuf, M; Wan Fariz Saiz; Baharem, Fauz	ALGORITHM PROCEEDING OF KNOWI						
4	Schless, AJ; DeCoster, G; Hfan, A; Marek, A; Mousavi, Seyf, S; Kite selector	RIVER FLOW 2014						
5	Schless, AJ; DeCoster, G; Dimoutafis, A, M	(Evaluation of RIVER FLOW 2014						
6	Soltz, J	Along, Bowen, Xiang, Huijie, JJ; Ming, YA	THE THIRTEENTH WUHAN IN					
7	Chen, A; Frangopol, DM; R Dong, Y.; Frangopol, D. M.; Sabadino, S.	Sustainable BRIDGE MAINTENANCE,						
8	Chen, A; Frangopol, DM; R Mohammed, A. W.; Altanayako, U.; Aki	National suaj BRIDGE MAINTENANCE,						
9	Ames, PM; Bagkalis, J; B Simoes-Mangano, M. J.; Nunes, I. L.	Contributions OCCUPATIONAL SAFETY						
10	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; Leal, J. C., Jr.	GIS multiform SAFETY, RELIABILITY AND						
11	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; de Sousa, G. F. M.	A multi-criteria SAFETY, RELIABILITY AND						
12	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; de Sousa, G. F. M.	Two sided or SAFETY, RELIABILITY AND						
13	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; de Sousa, G. F. M.	Multidimens SAFETY, RELIABILITY AND						
14	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; de Sousa, G. F. M.	Multidimens SAFETY, RELIABILITY AND						
15	Steenbergen, RDM; Vancil Silva de Sousa, P. P.; de Sousa, G. F. M.	Assessment SAFETY, RELIABILITY AND						
16	Meers, TH; Pinar, SO; Lam, S.; Chungto; Chen, Ken T-H.; Fan, Kuo	A novel hyper INNOVATION, COMMUNO						
17	Yoto, G; Ueki, E	Okuhara, Koji; Shibata, Junko; Oomoto, Consensus © 2012 INTERNATIONAL O						
18	Oraggio, S; Kavloka, A	Braskers, Mfs; Janssens, Gerrit K.	SHORTTEST BY EUROPEAN SIMULATION					
19	IEE	Sonegata, Akiyoshi	A Methodical 2012 INTERNATIONAL O					

tion de la requête



WEB OF SCIENCE™

...à ça

Revue

Auteurs

Titre

Partition

Vocabulaire / Formes

Passer de ça...

- permet de travailler sur un gros corpus
- permet d'introduire des variables
- permet de regrouper des abstracts en fonction de thématiques

Se poser la question de la requête

Se poser la question des bonnes variables

- Renvoyer à un nombre restreint de modalités
- Être pertinentes par rapport à la question de la recherche
 - **Année** : Certaines thématiques apparaissent-elles (ou disparaissent-elles) avec le temps ?
 - **Nb de citations** : certaines thématiques sont-elles plus porteuses que les autres ?
 - **Journaux/Discipline** : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?

Se poser la question de l'unité pertinente d'analyse



- **Nb de citations** : certaines thématiques sont-elles plus porteuses que les autres ?
- **Journaux/Discipline** : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?

Se poser la question de l'unité pertinente d'analyse



Un exemple

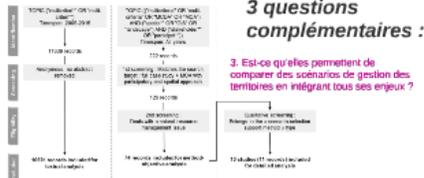


How Do Multi-criteria Assessments Address Landscape-level Problems? A Review of Studies and Practices

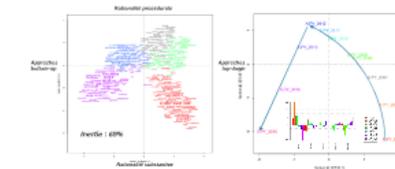
Sandrine Allain *, Gaël Plumeccocq, Delphine Leenhardt

2. Quelles sont les méthodes qui combinent le « multicritère », le « participatif » et le « spatial » ? Pour quoi faire ?

1. Quelle place pour les évaluations de durabilité « multicritère » ?



Résultats



Dans les EMC, les approches spatiales et participatives reposent sur des hypothèses de rationalités différentes.

On aurait pu faire :

Intérêts pour nous

Situer notre approche dans le paysage des EMC

Cibler une littérature précise pour identifier des trous dans la littérature



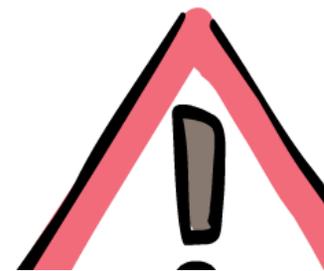
Analyse de contenu

Caractériser les outils en lien avec les approches



Permet de montrer qu'il n'existe pas d'outils complètement adaptés à notre approche.

revenir aux textes



How Do Multi-criteria Assessments Address Landscape-level Problems? A Review of Studies and Practices

Sandrine Allain *, Gaël Plumecocq, Delphine Leenhardt

2. Quelles sont les méthodes qui combinent le « multicritère », le « participatif » et le « spatial » ? Pour quoi faire ?

1. Quelle place pour les évaluations de durabilité dans la constellation « multicritère » ?

3 questions complémentaires :

3. Est-ce qu'elles permettent de comparer des scénarios de gestion des territoires en intégrant tous ses enjeux ?

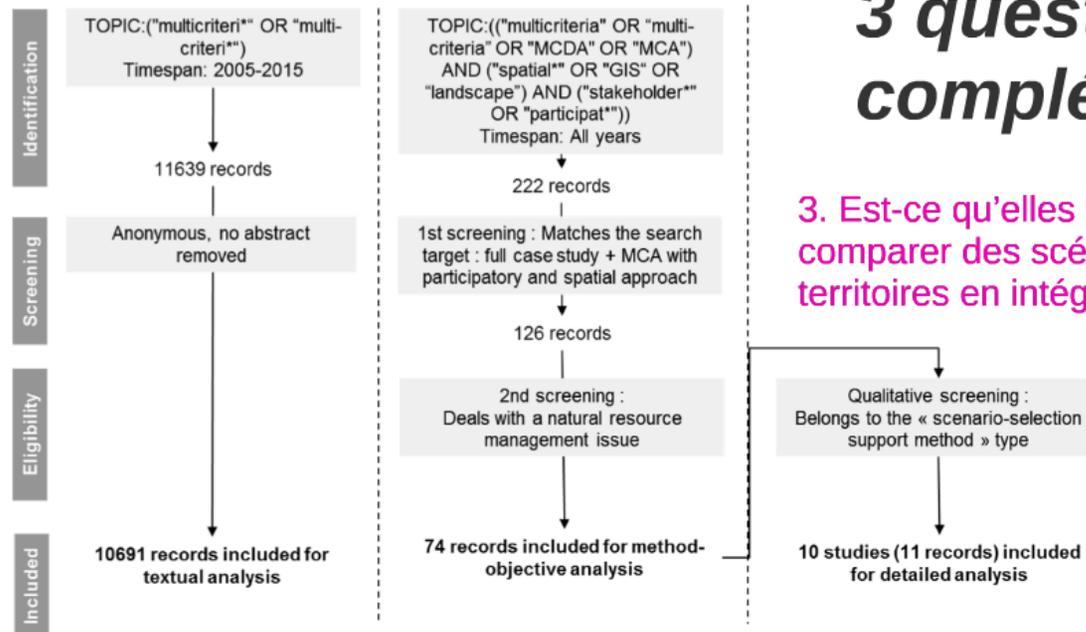


Fig. 1 PRISMA flow diagram for each objective addressed in the review (database search made in June 2015).

- **Nb de citations** : certaines thématiques sont-elles plus porteuses que les autres ?
- **Journaux/Discipline** : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?

Se poser la question de l'unité pertinente d'analyse



Un exemple

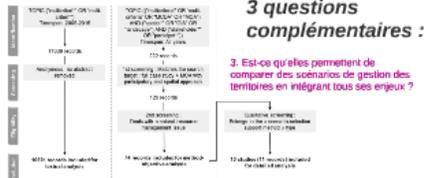


How Do Multi-criteria Assessments Address Landscape-level Problems? A Review of Studies and Practices

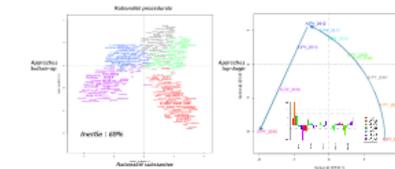
Sandrine Allain *, Gaël Plumelecq, Delphine Leenhardt

2. Quelles sont les méthodes qui combinent le « multicritère », le « participatif » et le « spatial » ? Pour quoi faire ?

1. Quelle place pour les évaluations de durabilité « multicritère » ?



Résultats



Dans les EMC, les approches spatiales et participatives reposent sur des hypothèses de rationalités différentes.

On aurait pu faire :

Intérêts pour nous

Situer notre approche dans le paysage des EMC

Cibler une littérature précise pour identifier des trous dans la littérature



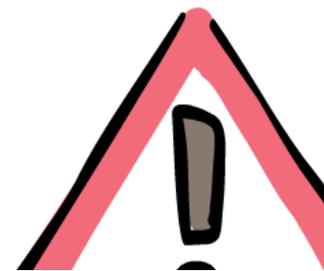
Analyse de contenu

Caractériser les outils en lien avec les approches



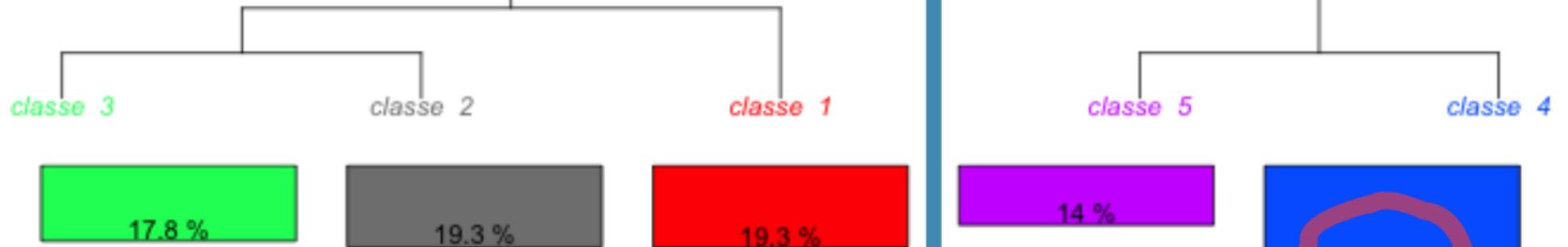
Permet de montrer qu'il n'existe pas d'outils complètement adaptés à notre approche.

revenir aux textes



PROBLEMES TECHNIQUES

DOMAINES D'APPLICATION



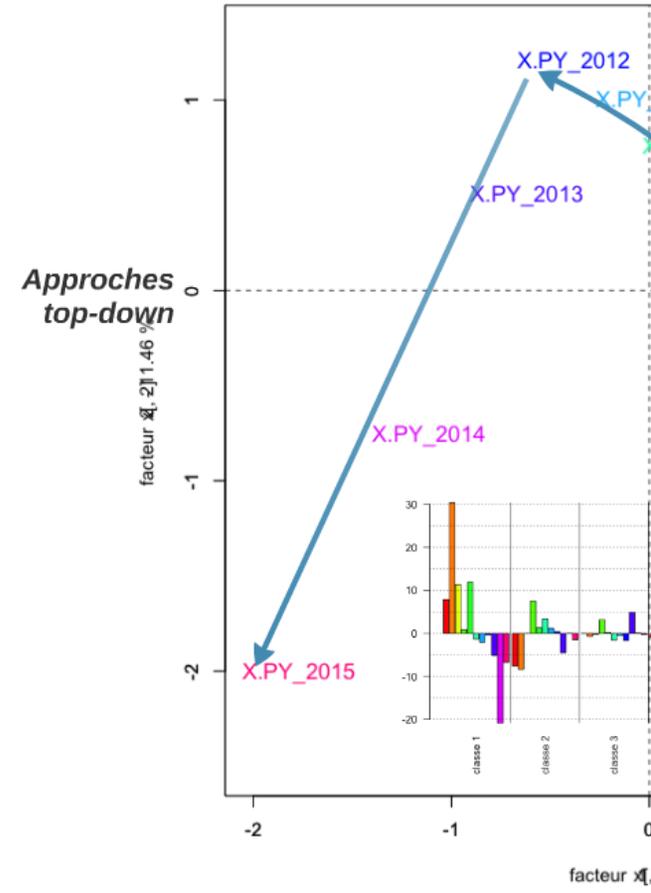
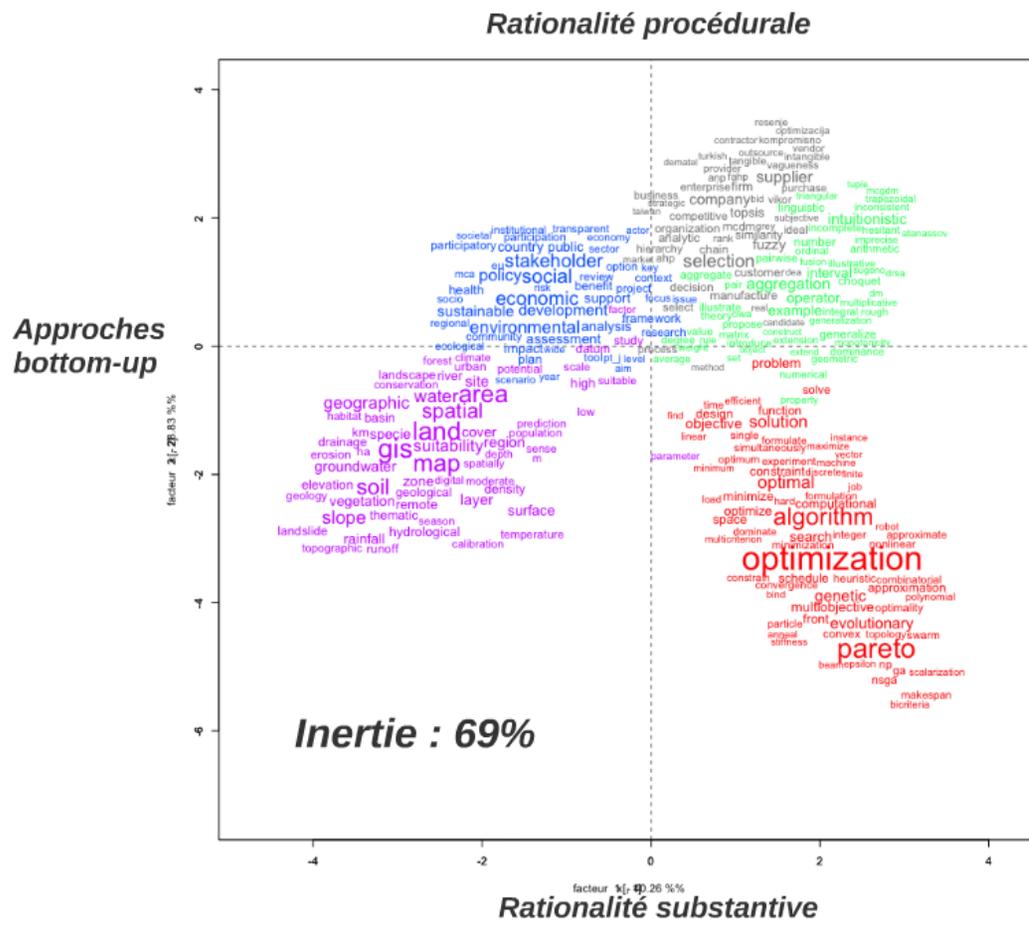
aggregation
intuitionistic
operator
example
interval
fuzzy
number
generalize
linguistic
choquet
aggregate
alternative
pairwise
membership
illustrate
arithmetic
theory
value
incomplete
propose
matrix
relation
dominance
numerical
ordinal
degree
introduce
integral

selection
supplier
company
fuzzy
topsis
firm
analytic
customer
similarity
chain
manufacture
organization
mcdm
decision
select
competitive
ideal
enterprise
ahp
hierarchy
process
business
vikor
product
purchase
propose
anp
tangible
vagueness

optimization
pareto
algorithm
optimal
solution
genetic
evolutionary
objective
search
problem
multiobjective
computational
optimize
space
front
minimize
constraint
design
schedule
function
approximation
solve
convex
nsga
ga
np
heuristic
optimality
nonlinear
minimization
swarm

gis
land
area
map
soil
spatial
slope
water
geographic
suitability
layer
site
region
groundwater
surface
geographical
zone
cover
vegetation
specie
locate
rainfall
remote
high
hydrological
river
basin
watershed

social
economic
stakeholder
environmental
policy
development
sustainability
sustainable
country
assessment
analysis
support
management
public
impact
health
indicator
benefit
plan
framework
european
participatory
review
future
project
national
resource
community



Dans les EMC, les approches spatiales et participatives reposent sur des hypothèses de rationalités différentes.

On aurait pu f

L'exercice de "revue de littérature" à l'aide d'IRaMuTeQ :



1 Lexicométrie

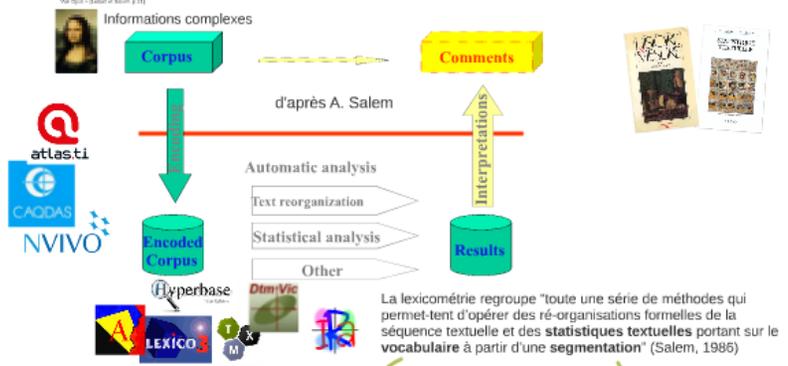


Table lexicale

Segmenter (nominal) : travail de contenu; textes, glissements, segments de mots.

Vocabulaire	Nombre d'occurrences															
	Absence (0) / Présence (1)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Paris																
Parisien																
Le petit chat est blanc																
Chat blanc																
Chat blanc et noir																
Chat noir																
Chat noir et blanc																
Chat blanc et noir et blanc																

Chaque "segment" est alors rattaché à un vocabulaire spécifique.

Le reste (17) = 173

Le mot "chat" (15) = 15

Chaque (15) = 15

Chaque (15) = 15

CE SONT LES SEGMENTS QUI SONT REGROUPÉS (ET PAS LE VOCABULAIRE)

Regroupement de textes qui utilisent des vocabulaires semblables

2 IRaMuTeQ

3 WEB OF SCIENCE

- permet de travailler sur un gros corpus
 - permet d'introduire des variables
 - permet de regrouper des abstracts en fonction de thématiques
- Se poser la question de la requête
- Se poser la question des bonnes variables
- Renvoyer à un nombre restreint de modalités
 - Etre pertinentes par rapport à la question de la recherche
 - Année : Certaines thématiques apparaissent-elles (ou disparaissent-elles) avec le temps ?
 - Nb de citations : certaines thématiques sont-elles plus pertinentes que les autres ?
 - Journaux/Discipline : certaines thématiques sont-elles plus liées à des journaux ou des disciplines ?
- Se poser la question de l'unité pertinente d'analyse

4 Un exemple



- Intérêts pour nous**
- Situer notre approche dans le paysage des EMC
 - Cibler une littérature précise pour identifier des trous dans la littérature
 - Caractériser les outils en lien avec les approches
 - Permet de montrer qu'il n'existe pas d'outils complètement adaptés à notre approche.



De toute façon : toujours revenir aux textes

