

Les fonctionnalités textométriques à l'épreuve du WEB2 :

Recueil, visualisation,
ergonomisation, formation

Proposition de fonctionnalités

Pour une approche
ergonomique des dispositifs
lexicométriques





Contexte

- **Chantier Textopol**
- **Résonnance avec d'autres travaux**
 - Document numérique
 - Démocratie électronique
 - Dispositifs en ligne
 - TICE
 - Mémoire du Web
 - Web sémantique
- **Pratique courante et objet d'étude**

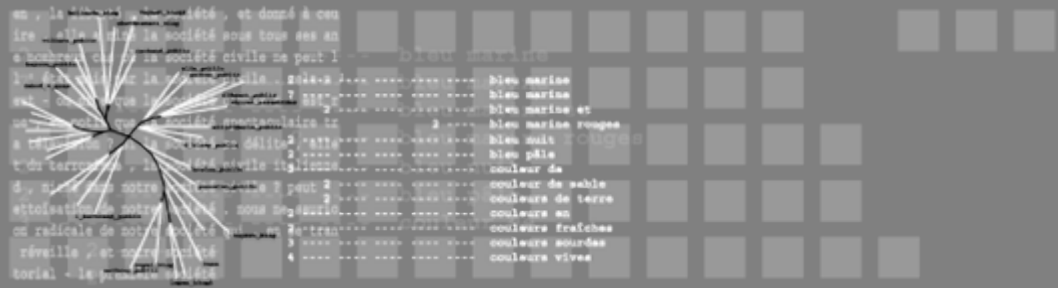




Contexte

- **Mutation sur le plan textométrique**
 - Réflexion sur des cartographies
 - Initiatives et nouveaux outils (Lexico3? Weblex? Hyperbase...) TXM





Etat du web

- Fonctionnalités et usages courants



Etat du web

GAPMINDER WORLD

HOME GAPMINDER WORLD DATA VIDEOS DOWNLOADS FOR TEACHERS GAPMINDER LABS

Open graph menu

CLIMATE CO2 emissions since 1820

In 1820, at the dawn of the Industrial Revolution, United Kingdom emitted most CO2 - both per person and in total emissions.

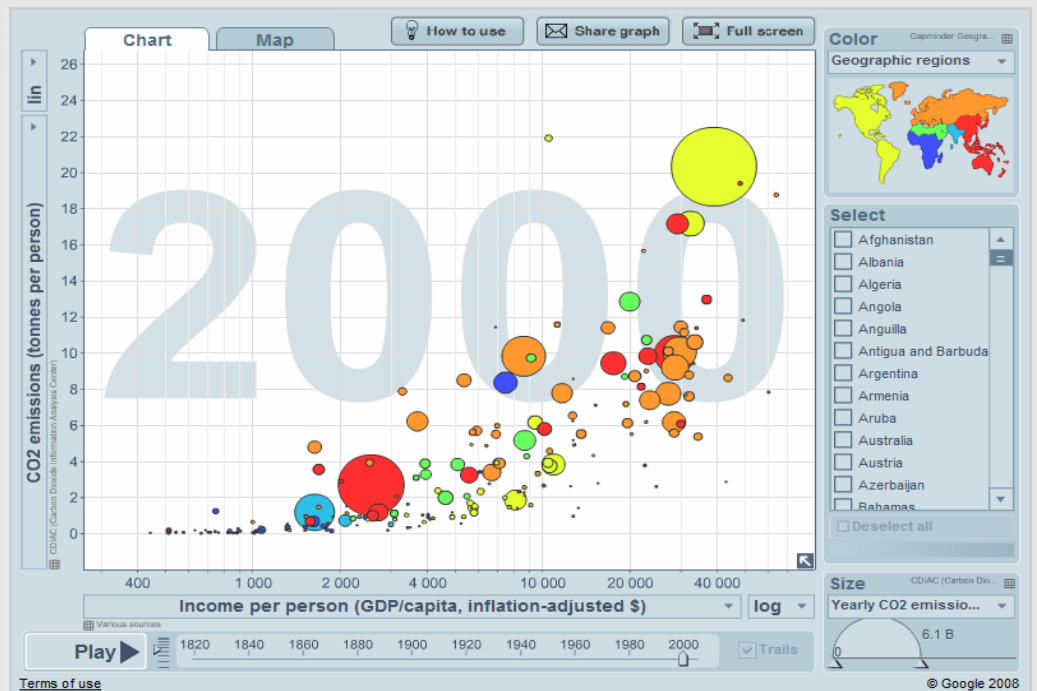
Click Play to see how USA becomes the largest emitter of CO2 from 1900 onwards.

Read more ...

Share

See also:

- USA or China, who emits most CO2



Etat du web

fnac.com Agitateur de curiosité

Aide | Flux RSS | Newsletter | Mon compte

Mon panier
Votre panier est vide

Produits

My Fnac

Spectacles

Téléchargements et MP3

fnacLive

labo fnac

Adhérents

Livres

Musique

DVD, Blu-ray

Jouets

Jeux vidéo

Logiciels

Informatique

Téléphones, MP3, GPS

TV, vidéo, hi-fi

Photo, caméscopes

Voyages

Tirages photo

Rechercher

Livres

OK

» Livre » Les linguistiques de corpus



visuel non disponible

Les linguistiques de corpus

Benoît Habert, Adeline Nazarenko, André Salem

(donnée non spécifiée). Paru en 12/1997

Tirage épuisé, indisponible.

Je l'ai déjà | Je ne suis pas intéressé

Ma note : | Ajouter un avis

Vendez le vôtre

Les internautes ayant acheté Les linguistiques de corpus ont également acheté :



Introduction à...
Jacques Moeschl...
★★★★☆



Les mots franca...
Henri Mitterand



Arts et science...
François Rastie...



La variation so...
Françoise Gadet

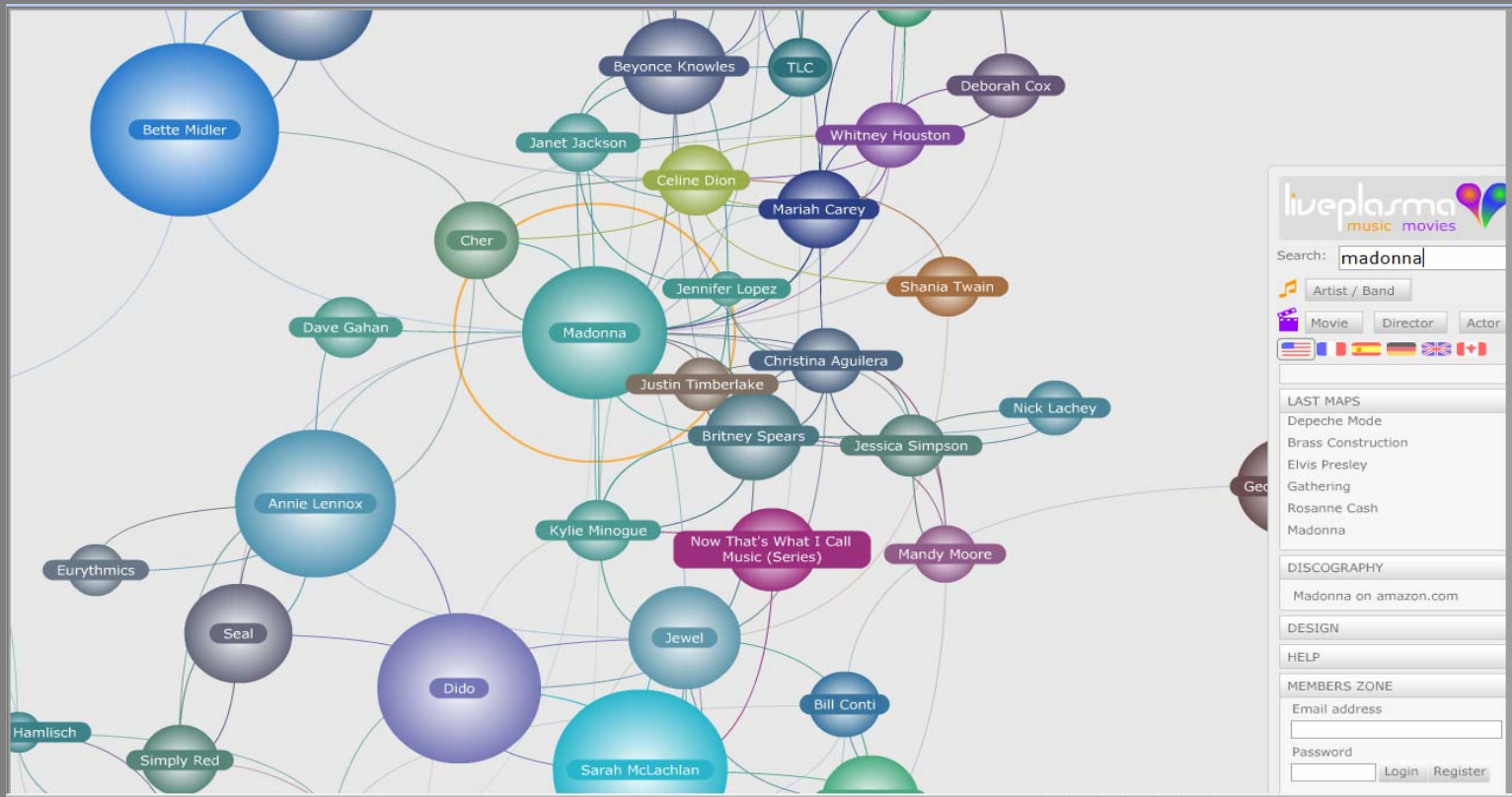


Pragmatique pou...
Dominique Maing...
★★★★☆



Comprendre la l...
Robert Martin

Etat du web





Etat du web

- Contamination d'un domaine sur l'autre?
- Que peut-on tirer du web pour l'appliquer à nos outils?
- Nécessité de repenser les interfaces



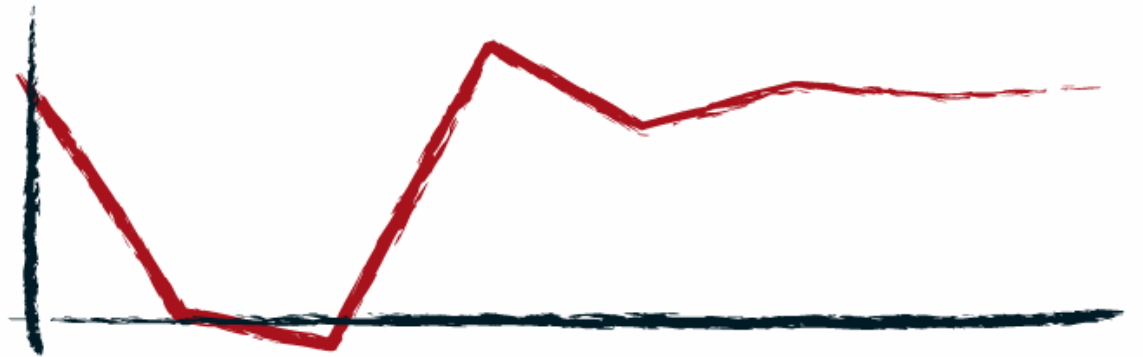
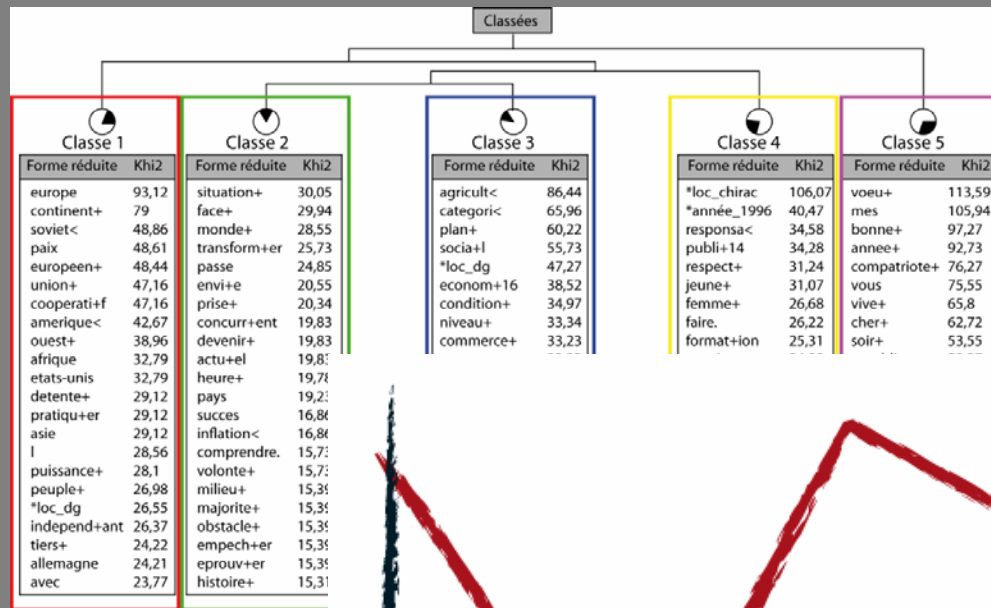


Les outils textométriques

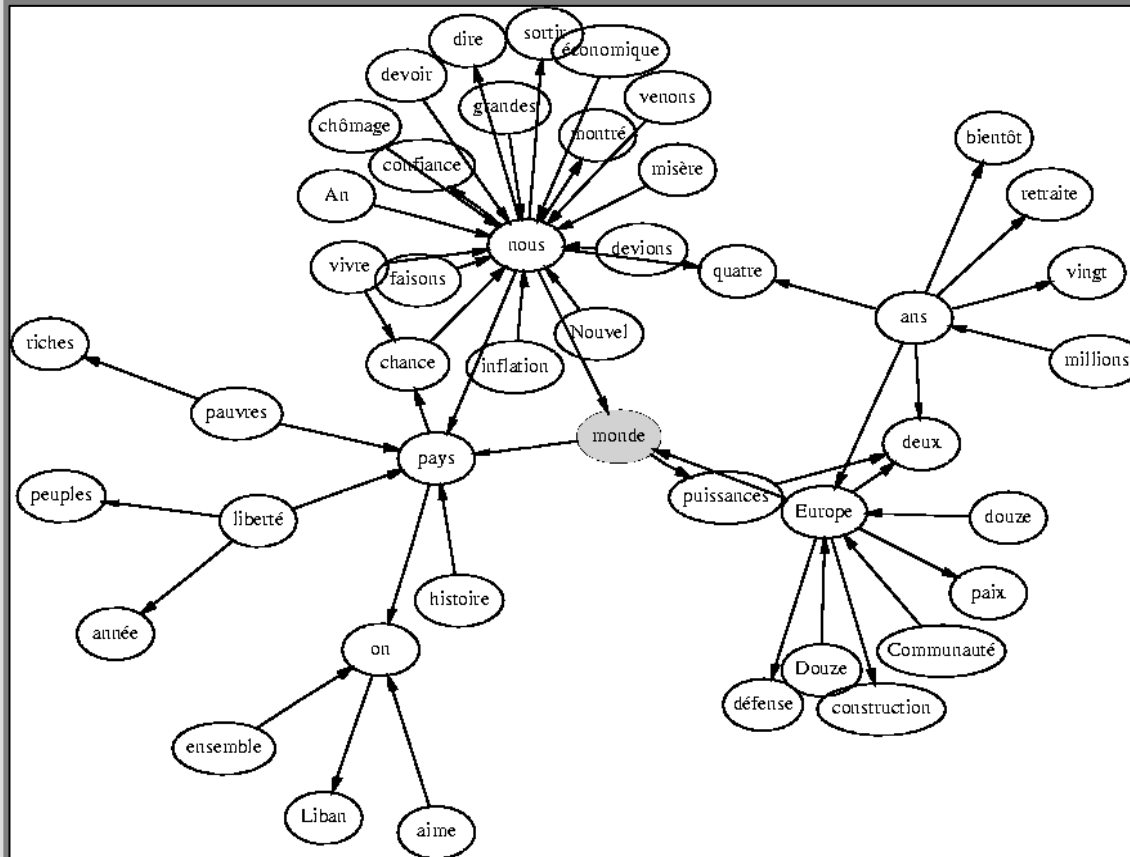
- Comparer les outils?
 - Approches longitudinales et contrastives
 - Approches structurelles



Les outils textométriques



Les outils textométriques



Terme	Frq Tot.	Frq P...	Spécif
pense	29	29	42
pensée	14	14	21
ceux	121	26	13
penser	7	7	11
je	411	41	8
qui	623	52	8
souffrent	12	7	8
ou	130	18	7
maladie	12	6	7
pensées	4	4	7
victimes	15	6	6
solitude	13	6	6
soldats	14	6	6
vont	15	5	5
seuls	16	5	5
ma	25	6	5
personn	3	3	5
celles	25	6	5
vers	44	9	5
fêter	7	4	5
connaissent	3	3	5
deuil	3	3	5
à	1071	68	5
vos	47	9	5
pensant	2	2	4
aussi	116	12	4
inondations	2	2	4
l'europe	2	2	4
pensais	2	2	4
injustice	6	3	4
soir	48	7	4
abord	35	7	4
soirée	8	3	4
nouvel	20	5	4
du	341	25	4
cruellement	2	2	4

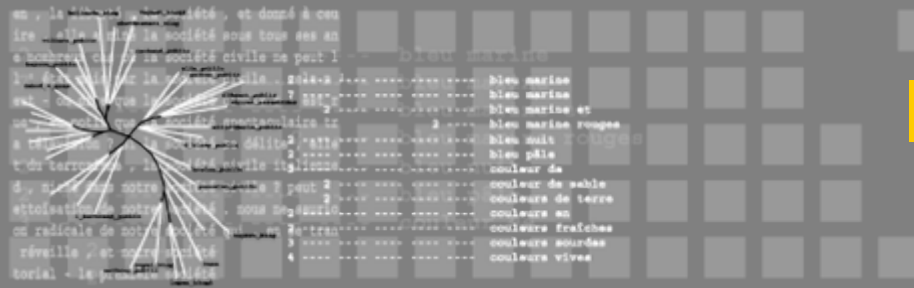
Les outils textométriques

Classe 1 : 255 U.C.E (27%)

Classe 2 : 714 U.C.E (73%)

Forme réduite	Khi2	Forme réduite	Khi2
je	969	votre	12,07
vous	148,97	soir+	12,07
souhait<	148,06	bonne+	11,34
suis	60	redire.	11,25
dire.	45,77	tiens	11,25
mes	44,38	present+er	11,25
voeu+	38,95	echang+er	11,25
dire+	38,14	metropole+	10,98
forme+	37,02	savoir.	10,9
me	27,54	m	10,52
annee+	27,44	soiree+	10,12
vouloir.	27,02	avais	10,12
heur+eux	23,87	coeur+	9,97
compatriote+	23,61	promettre.	9,86
francais+	23,31	parl+er	9,43
adress+er	22,45	année_1974	9,34
cher+	19,82	mon	9,34
ai	18,98	pas	9,32
nom+	17,59	ce	9,27
outr	15,11	ma	8,67
qu+	14,12	mer+	8,67
enjeu+	14,07	recu+	8,43
adresse+	13,15	aim+er	8,43
*loc_giscard	12,76	demand+er	8,09
propos+er	12,36	rappel+er	7,63

Forme réduite	Khi2	Forme réduite	Khi2
*loc_dg	52,38	trouv+er	4,1
nous	14,84	ailleurs	3,97
econom+	10,6	leur	3,97
année_1963	10,31	guerre+	3,93
devoir.	8,67	moderne+	3,73
année_1965	7,29	solid+e	3,61
année_1961	7,27	rapport+	3,61
année_1968	6,24	acquérir	3,61
année_1960	6,16	soi	3,61
moyen+	5,56	quel	3,61
developpement+	5,5	troubl+er	3,61
*année_1966	5,12	scientifi<	3,61
arme+	5,07	ensemble+	3,55
cooperati+f	5,07	*année_2000	3,55
techn+	5,07	*année_1962	3,44
peuple+	5,06	mondia+	3,44
face+	4,71	moment+	3,38
organisat+ion	4,71	cas	3,24
*année_2001	4,57	dehors	3,24
avons	4,52	telle	3,24
elle	4,46	multipli+er	3,24
jusqu+	4,34	monetaire+	3,24
plan+	4,34	but+	3,24
conflit+	4,34	route+	3,24
inflation<	4,34	aid+er	3,01



Implémentations plus ergonomiques

- Ergonomie vs esthétique
- L'ergonomie?
 - Industrie
 - Espace de travail
 - IHM (internet + multimédia)
 - Design



Références du domaine

- **Ergonomie**
 - Bastien et Scapin
 - A. Boucher
- **Didactique des écrans**
- **Cognition**



Ergonomie

- Accéder rapidement aux données
- De façon logique
- Qui fasse sens
- Sans surcharge cognitive
- Visualisation des données

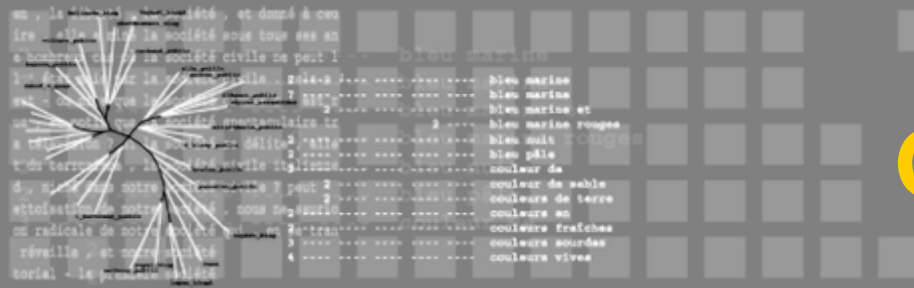


Références du domaine

- Grand public vs Spécialistes
- Interfaces efficaces & conviviales
- Domaine universitaire vs commercial (test utilisateur)

Critères ergonomiques (B & S)

- **Guidage** : incitation, groupement/distinction entre items (par le format, par la localisation), feed back immédiat, lisibilité
- **Charge de travail** : brièveté, (concision, actions minimales), densité informationnelle
- **Contrôle explicite** : actions explicites, contrôle utilisateur
- **Adaptabilité** : flexibilité, prise en compte de l'expérience de l'utilisateur
- **Gestion des erreurs** : protection contre les erreurs, qualité des messages d'erreur, correction des erreurs
- **Homogénéité/cohérence**
- **Signifiante des codes et dénominations**
- **Compatibilité**



Guidage : incitation

- Alternatives
- Etapes, Cheminement, temps...
 - Hyperbase
- Propositions
 - Curseur, vumètre
 - Adressage efficace des textes
 - Pages, paragraphe...



Guidage : Groupement /distinction entre les items

- Catégories
 - Topologie
 - visuel

Navigation: Classification double 121

Classification descendante

u.c.e. caractéristiques

Analyse factorielle

Formes représentatives

Corpus

Clés catégorielles

Classification ascendante

Dictionnaire des formes

Statistiques

ECRAN PANORAMIQUE

Classe 1 Classe 2 Classe 3 Classe 4 Classe 5 Classe 6 Classe 7

Formes réduites	Chi2
suffrage+	61,06
remerci+er	39,91
million+	39,91
ont	29,23
lui	24,38
*dat_1995_07_05	21,09
salu+er	16,67
presid+ent	16,67

Catégories "grammaticales"	Chi2
Nombres	36
Marqueurs de la personne (mots outils)	4
Formes reconnues mais non codées	4
Démonstratifs, indéfinis et relatifs (mots outils)	1
Marqueurs d'une intensité (mots outils)	0
Marqueurs d'une modalisation (mots outils)	0
Adjectifs et adverbes	0
Verbes	0

Formes réduites	Eff.	Catégories
et	200	Marqueurs d'une relation
a	198	Formes non reconnues
je	188	Marqueurs de la personne
qu+	132	Démonstratifs, indéfinis
qui	118	Démonstratifs, indéfinis
pour	86	Marqueurs d'une relation
france	82	Lieux, pays
dans	72	Marqueurs d'une relation

Mots analysés	Classes	Pourcentage
1	10	20,75%
2	10	20,75%
3	22	44,50%
4	12	24,00%
5	14	28,00%
6	18	36,00%
7	18	36,00%

en la société, et donné à ces
ire aller dans la société sous tous ses an
s'embrassent dans la société civile ne peut l
bien marine
bien marine
bien marine et
bien marine rouges
bien nuit rouges
bien pâle
couleur de
couleur de sable
couleur de terre
couleurs en
couleurs fraîches
couleurs sourdes
couleurs vives

Charge de travail : brièvement

- Concision 8/9 item / écran
- Densité informationnelle
 - Lien
 - Visualisation
 - Hyperbase
- Action minimale



Gestion des erreurs et protection contre les erreurs

- **Fonctionnement**
 - Exemple: localisation corpus
- **Qualité des messages**
 - erreur Baob2
 - erreur de lecture du fichier des facteurs
 - Error ; invalid



Proposition de fonctionnalités et de visualisations

- Intégrer la 3eme dimension dans les analyses factorielles
 - (F. Pierre et JM. Viprey)
- Naviguer entre les différentes partitions d'un corpus
- Implémenter le mouvement dans les analyse multidimensionnelles et longitudinales





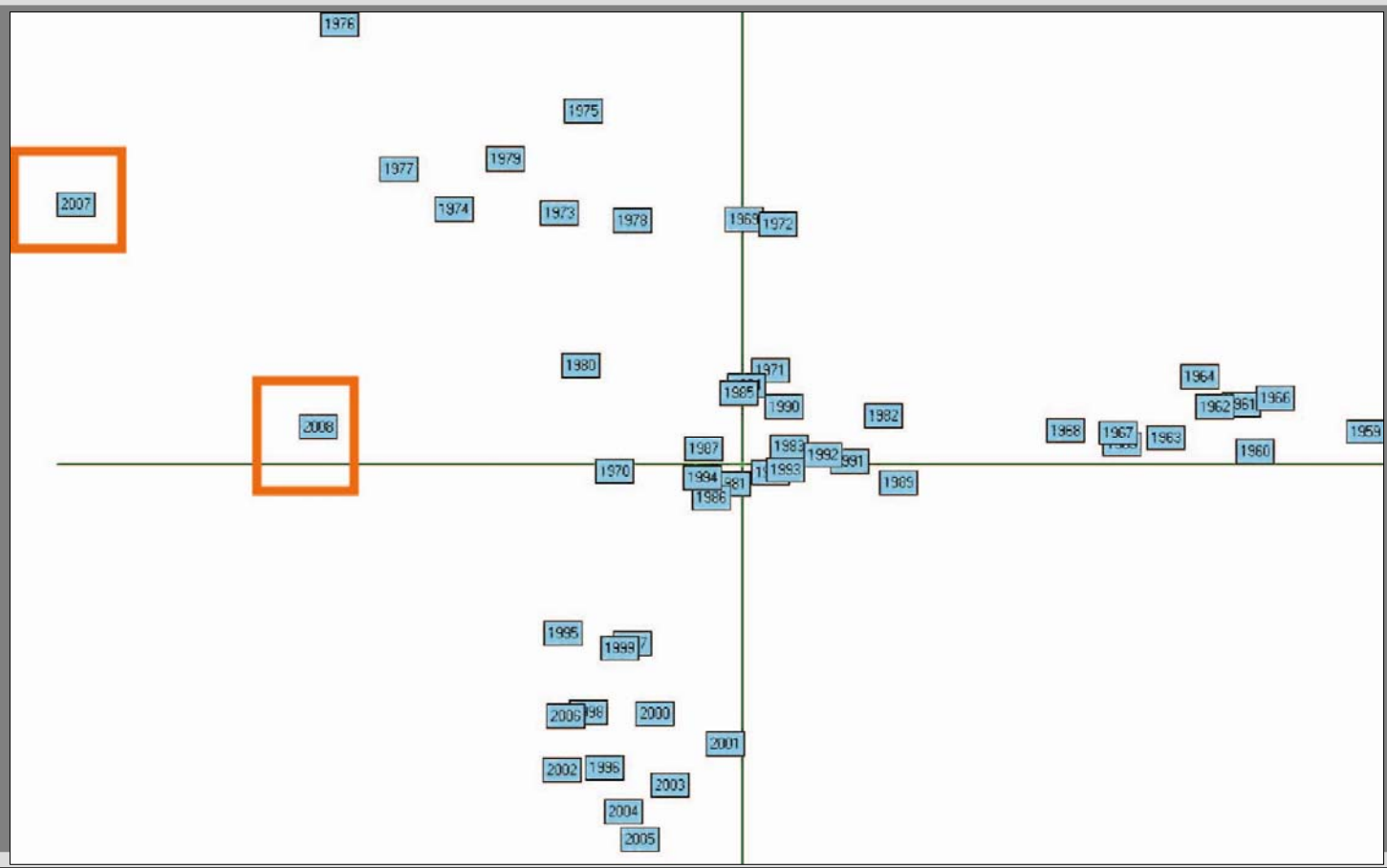
Implémenter le multimédia

- **Lier les domaines**
 - du multimédia, 3D, RVRA
 - des interfaces
 - de l’hypertextualité
 - de l’édition numérique
 - de l’ergonomie, de la visualisation
 - de la lexicométrie, textométrie, logométrie



en elle... et donné à ce...
 ire... la société sous ses an...
 e... la société civile se peut l...
 bien marine
 bien marine
 bien marine et
 bien marine rouges
 bien nuit rouges
 bien pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs en
 couleurs fraîches
 couleurs sèches
 couleurs vives

Vœux 1959-2008



en... la... société, et donné à ce...
 ire... la... société sous tous ses an...
 e... la... société civile se peut l...
 le... la... 2014...
 se... 7...
 ne...
 a...
 t...
 d...
 att...
 ce...
 ré...
 to...

Vœux 1959-2001

