

hollande_blog voynet_blog2
 chevenement_blog
 roubaud_public
 olin_public
 rochard_public
 chirac_assemblees
 alliotMarie_public
 d-bisq_public
 royalet_public
 besancon_blog
 lepen_blog2
 sarkozy_public

en , la liberté , la société , et donné à ceu
 ire , elle a miné la société sous tous ses an
 e nombreux cas où la société civile ne peut l
 l ' état mais par la société civile cela r'
 eut - on dire que la société civile s ' est r
 ue , au motif que la société spectaculaire tr
 a télévision ? si la société se délite , elle
 t du terrorisme , la société civile italienne
 d , niché dans notre société civile à nous
 ettoisement de notre société . nous ne sauro
 on radicale de notre société qui , en se tran
 réveille , et notre société illuminée s ' ét
 torial - la première société avancée qui en

2
 7
 2
 2
 2
 2
 1
 3
 2
 2
 3
 4
 2
 2
 2

bleu marine
 bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs en
 couleurs fraiches
 couleurs sourdes
 couleurs vives
 gris bleu
 gris mouette
 jaune fluo



Corpus issus du web : enjeux, méthodes et outils

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile gale 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine et
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle nuit
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sordides de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo

Corpus issus du web : enjeux, méthodes et outils

- Bienvenue à toutes et à tous et merci à...
- **Florence CLAUD**, conservateur en chef du patrimoine, responsable des référentiels documentaires aux Archives nationales, membre de l'EA 3624 (Centre Jean-Mabillon) de l'École nationale des chartes :
Le métalangage XML : apports pour traitement et analyse de corpus textuels et multimodaux
- **Valérie BEAUDOUIN**, Enseignante-chercheuse en sciences sociales, Télécom ParisTech :
Les corpus web et l'approche textométrique : nécessaire combinaison d'approches quantitatives et qualitatives

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mène par la société civile ... cela n
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme , la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société , nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine et
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de
couleurs de terre
couleurs de
couleurs de
couleurs fraîches
couleurs sourdes
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



En guise d'introduction...

- **1/ Constitution de corpus et statistiques textuelles (lexicométrie, textométrie...) :**
contextes, filiations et enjeux
- **2/ Corpus, web et « humanités » :**
enjeux terminologiques et scientifiques

en, la liberté, la société, et donné à ceu
ire, elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l' état mais par la société civile. Cela n
eut - on dire que la société civile s' est r
ue, au motif que la société spectaculaire
a télévision ? si la société se délite, elle
t du terrorisme la société civile italienne
d, niché dans notre société civile ? peut
ettoisaison de notre société nous ne sauro
on radicale de notre société qui, en se tran
réveille, et notre société illuminée s' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine
bleu marine et
bleu marine et
bleu marine rouges
bleu marine
bleu marine
couleur de sable
couleurs de terre
couleurs en
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



1/ Constitution de corpus, statistiques textuelles et big data : Séminaires et actualité de la question

- **Très rapidement, à titre indicatif, rappel :**
- **Plusieurs séminaires...**
- **Céditec** : Veille numérique, analyse du discours et outils statistiques (J-M Leblanc et C Barats), la visualisation des données textuelles (J-M Leblanc)...
- **Et équipes de chercheurs et/ou revues...**
- ✓ **UMR 7320 : Bases, Corpus, Langage** (site : <http://logometrie.unice.fr/>), revue en ligne Corpus, plateforme Hyperbase, base de données PoliText (discours politiques français 1789-2002), voir D. Mayaffre ; **Cortext, INRA, Sens**, (site <http://www.inra-ifris.org/projet-scientifique-8/>), l'Institut francilien recherche, innovation et société (IFRIS)
- ✓ **revue Mots**, ENS Lyon, (site : http://icar.univ-lyon2.fr/revue_mots/), (ICAR Interactions, Corpus, Apprentissages, Représentations, UMR 5191 CNRS, Lyon 2, ENS Lyon, Triangle UMR 5206, Céditec, UPEC), ; **Revue Semen**, Elliad, Besançon ; **Les Carnets du Cediscor**, Paris 3-Sorbonne nouvelle, SYLED-Cediscor, JADT...
- **Et actualité** : Informatique et SHS rencontres INRA-INRIA, Cortext, 4 juillet 2014
- **Actualité liée aux masses de données (contexte big data) et puissance des algorithmes : le web comme lieu de production et de circulation de données, de « stockage » (base de données en ligne), de traitement (outil en ligne)**

en , la liberté , la société , et donné à ceu
 ire , elle a mène la société sous ses an
 e nombreux où la société civile ne peut l
 l ' état n'est pas la société civile . cela n
 eut - on dire que la société civile s ' est r
 ue , au motif que la société spectaculaire
 a télévision ? si la société se délite , elle
 t du terrorisme la société civile italienne
 d , niché dans notre société civile ? peut
 ettoisation de notre société nous ne sauro
 on radicale de notre société qui , en se tran
 réveille , et notre société illuminée s ' ét
 torial - la première société avancée qui

bleu marine
 bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs en
 couleurs fraîches
 couleurs sourdes
 couleurs vives
 gris bleus de terre
 gris mouette
 jaune fluo



1/ Quelques repères et enjeux

- **Schématiquement** , émergence statistique textuelle :
 - **rencontre histoire sociale et linguistique** (Tournier, Mots, 2010 N°94, Fiala, 2013)
 - **Contexte années 60-70** : statistiques, développement informatique, SHS (Benzécri, AFC, Cibois, approches quantitatives d'enquête)
- **Liens** philologie, lexicographie, littérature, socio-histoire,
- Démarches herméneutiques nouvelles : tenir compte du contexte, des conditions de production, notion de discours vs texte (Pêcheux, 1969)
- **Impossible** de retracer une histoire complexe
- **Masse de données et SHS** (Bourdieu, socio, statistique textuelle)
- **Notion de corpus et de traitement automatisé** : diversité de principes méthodologiques et de méthode de collecte
- **Diversité de programmes scientifiques**
exemples : corpus de textes littéraires (Frantext, 1962, CNRS, Nancy, TLF), corpus de textes politiques et socio-politiques (revue Mots, 1980, Laboratoire de ENS-Saint Cloud), bcp de labo P3, Besançon..

en , la liberté, la société, et donné à ceu
 ire, elle a mène la société sous tous ses an
 e nombreux où la société civile ne peut l
 l' état mène par la société civile. Cela m
 eut - on dire que la société civile s' est r
 ue, au motif que la société spectaculaire
 a télévision ? si la société se délite, elle
 t du terrorisme la société civile italienne
 d, niché dans notre société civile ? peut
 ettoisation de notre société nous ne saurio
 on radicale de notre société qui, en se tran
 réveille, et notre société illuminée s' ét
 torial - la première société avancée qui

bleu marine
 bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs en de
 couleurs fraîches
 couleurs sourdes de sable
 couleurs vives
 gris bleus de terre
 gris mouette
 jaune fluo



1/ Terminologie et approches

Se familiariser avec :

- une **terminologie complexe** (lemme, hapax...), des **filiations plurielles** (philologie, herméneutique, linguistique, sciences sociales..),
- **Une diversité d'approches** (lexicale, énonciative, morphosyntaxique...) et de **choix méthodologiques** (quelle unité retenir ?)



en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile gala 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisaison de notre société nous ne sauzio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui ,

bleu marine
bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en
couleurs fraîches
couleurs sourdes
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



1/ Constituer un corpus

- ✓ **Actualité** des outils et des recherches: outils en ligne, accessibilité des données, nombreux outils de traitement
- ✓ **Différentes étapes** : critères de sélection et de clôture du corpus, modalités de collecte et d'archivage, choix de l'outil d'analyse
 - *Quelles sont les hypothèses de recherche ?*
 - *Quelle problématique ? Quel cadre d'analyse ? (Lexique socio-politique, usages socio-discursifs, notion de motif, circulation, notion de genre, phraséologie, sémantique textuelle ?)*
 - *Que mesure t'on ?*
 - *Quelle unité de calcul ? Lemmatisation vs forme graphique/unité textuelle*
 - *Quel étiquetage ? (Entités nommées)*
 - *Quelle partition ?*
 - *Quels observables ?*
 - *Quelle clôture ?*
 - *Quels types de résultats ? (listes lexicales, AFC, cooccurrences, collocation, catégories grammaticales, racines morphologiques, familles lexicales, spécificités)*
 - *quelle représentation visuelle ?*
 - *Quels types d'interprétation ?*

en , la liberté, la société, et donné à ceu
ire , elle a mané la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile, cela n
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisiation de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu nuit
couleur de sable
couleur de sable
couleurs de terre
couleurs de terre
couleurs fraîches
couleurs fraîches
couleurs sourdes
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



1/ Constituer un corpus : un moment clé de la recherche, construire « sa » méthode

- Lien entre critère de constitution d'un corpus et hypothèses de travail du chercheur (définir les critères de collecte et de clôture du corpus, processus itératif)
- **Types de corpus et diversité d'analyse et de mesure : diversité des outils** (Lexico, Alceste, Tropes,...)
- **Exemples** : circulation d'énoncés (« majorité plurielle ») ou de motifs, approche socio-historique (« classement de Shanghai », cf. formule Krieg-Planque), usages socio-discursifs d'énoncés ou syntagmes « travailleurs immigrés », clandestins », analyse de controverses...

Quelques questions...

- Quel type d'analyse : Mesure d'itération graphico-lexicale, effets rhétoriques, identifier des thématiques, des phraséologies...?
- Quelle utilisation des seuils ? quel type d'algorithme ?
- Quelle interprétation des fréquences, sous ou sur représentation endogène ou non au corpus ?
- Quelle documentation du corpus ?
- Quel « refroidissement » du corpus ?
- Quelle mise en contexte ?

en , la liberté , la société , et donné à ceu
 ire , elle a mène la société sous tous ses an
 e nombreux où la société civile ne peut l
 l ' état mais par la société civile . cela n
 eut - on dire que la société civile s ' est r
 ue , au motif que la société spectaculaire
 a télévision ? si la société se délite , elle
 t du terrorisme la société civile italienne
 d , niché dans notre société civile ? peut
 ettoisation de notre société , nous ne sauzio
 on radicale de notre société qui , en se tran
 réveille , et notre société illuminée s
 torial - la première société avancée qui ,

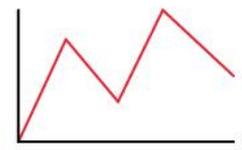
Principaux outils
 (Voir aussi CAQDAS, computer assisted qualitative data analysis)



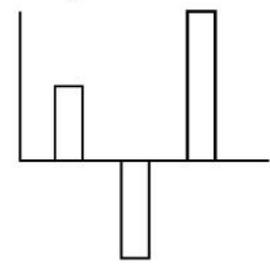
Typologies logicielles

Lexico3, Hyperbase, Txm...

Approche contrastive Longitudinale



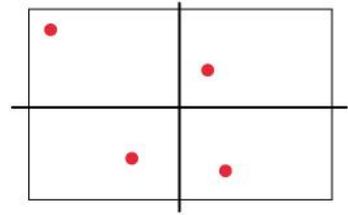
Spécificités



Concordances

rorisme , la société civile ital
 hé dans notre société civile ? pe
 tion de notre société . nous ne s
 cale de notre société qui , en se
 le , et notre société illuminée s
 - la première société avancée qui

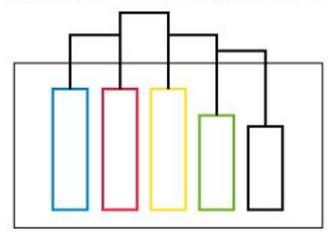
AFC



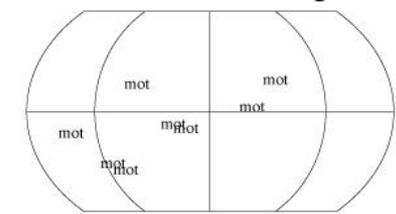
sur la base de partitions (norme endogène)

Approche structurante

Alceste : cooccurrences entre énoncés

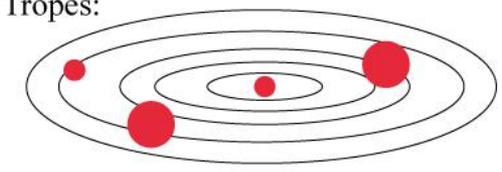


Astartex : cooccurrences généralisées



Ontologies sémantiques

Tropes:



Outils grammaticaux, aide à la rédaction

Cordial:



Alceste, Astartex...

Tropes, Cordial...

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile . cela n
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisaison de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



1/ Evolution des pratiques

➤ De la catégorisation des données à la visualisation des résultats

✓ Evolution des pratiques de la lexicométrie à la textométrie :

- catégorisations quasi-systématiques,
- annotations,
- lectures AFC, ...

✓ Des résultats à l'interprétation

- Outils de visualisation
- Quid interprétations sans retour au texte (Alceste, toujours des résultats)

en, la liberté, la société, et donné à ceu
ire, elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l' état mais par la société civile. Cela m
eut - on dire que la société civile s' est r
ue, au motif que la société spectaculaire
a télévision ? si la société se délite, elle
t du terrorisme la société civile italienne
d, niché dans notre société civile ? peut
ettoisation de notre société, nous ne saurio
on radicale de notre société qui, en se tran
réveille, et notre société illuminée s' ét
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



2/ Corpus, web et « humanités » : enjeux terminologiques et scientifiques

- **Numérisation et analyse automatique** ont précédé le **web**
- Cependant, web amplifie emploi de l'expression « **humanités numériques** » (« *digital humanities* », quid traduction et recouvrement sémantique, voir ci-après)
- **Spécificités du web :**
 - Logiques d'archivage changent (a priori vs a posteriori, Biblio Congrès, Twitter, 14/04/10, Ertzscheid),
 - Collecte en temps réel + fonctionnalité (retweet)
- quid notion traces explicites/implicites (Simonnot), quid notion de documents (Salaün), de données, d'observables (Twitter : matérialité textuelle + graphes relationnels), conditions de collecte (profondeur)
- Volume des corpus (« gooffre », Brunet)
- Algorithmes, big data et visualisation (cartographies)

en , la liberté , la société , et donné à ceu
 ire , elle a mène la société sous tous ses an
 e nombreux où la société civile ne peut l
 l ' état mène par la société civile gala 7
 eut - on dire que la société civile s ' est r
 ue , au motif que la société spectaculaire
 a télévision ? si la société se délite , elle
 t du terrorisme la société civile italienne
 d , niché dans notre société civile ? peut
 ettoisation de notre société nous ne saurio
 on radicale de notre société qui , en se tran
 réveille , et notre société illuminée s ' ét
 torial - la première société avancée qui ,

bleu marine
 bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs en
 couleurs fraîches
 couleurs sourdes
 couleurs vives
 gris bleus de terre
 gris mouette
 jaune fluo



2/ Corpus, web et « humanités » : enjeux terminologiques et scientifiques

De la **lexicométrie**, la statistique textuelle, la textométrie aux data sciences ou aux **digital humanities** (humanités numériques/ digitales)

De nombreux courants...

- **Digital humanities** (JC Plantin, Manuel analyse du web, 2013) : pas de définition unique, d'abord appelées **humanities computing**, collaboration père Roberto Busa et IBM (1949, index Thomisticus, textes Saint Thomas d'Aquin)
- **Objectifs** : traitement informatique pour données numérisées (corpus de textes et documents, domaine: Lettres, humanités)

- Importance de l'archivage , de l'indexation et des outils de fouille de données
- Explosion d'outils et ressources en ligne : pour Presner, *digital humanities 2.0*

En découle...

- **Cultural analytics ou computational humanities** (analyse des productions culturelles numérisées)
- **Digital methods** (objets et méthodes natifs du web)

en , la liberté , la société , et donné à ceu
ire , elle a mané la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile gala 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui ,

bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle nuit
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



2/ Corpus, web et « humanités » : enjeux terminologiques et scientifiques

➤ Cultural analytics ou computational humanities :

- articulation de méthodes computationnelles (*software studies*) et des corpus de ressources culturelles (couvertures du magazine Time),
 - analyse des productions culturelles numérisées (initié par Manovich, 2007)
- **Rappel : Digital humanities (humanités numériques)** l'application d'une méthode d'analyse informatique aux sciences humaines

➤ Digital methods :

- pas étude des objets numérisés mais **objets natifs du web** (hyperliens, wikis, moteurs de recherche (Richard Rogers, 2009)) ;
- controverses en ligne, visualisation des relations entre acteurs, applications (Like),
- méthodes natives du web : techniques de scraping (extraction contenu web en fichiers prêts à l'analyse)
- **Nota bene** : différence *digital method*, *virtual method* (Hine, 2005) où adaptation des méthodologies traditionnelles aux objets en ligne (entretiens par e-mail, analyse ethno. de communautés en ligne)

en , la liberté , la société , et donné à ceu
ire , elle a méné la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état méné par la société civile .
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme , la société civile italienne
d , niché dans notre société civile ? peut
ettoisiation de notre société nous ne sauzio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée .
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



2/ Corpus, web et « humanités » : visualisation des données

- Nouveaux outils, nouvelles visualisations, nouvel apprentissage (Gephi, linkurious...)
 - cartographie et « belles images » (Lien avec l'archéologie, quelques exemples...)
 - Quid des résultats à l'interprétation ?
 - Quid outils de visualisation ?
- **Prudence...**
 - facilité des outils techniques qui fabriquent des résultats (artefact)
 - Délégation à la machine, outils d'aide à la décision, jolis visuels, effet de réalité des visuels
- Difficultés culture commune : informaticien, SHS, SIC, Sciences du langage...

en, la liberté, la société, et donné à ceu
 ire, elle a mène la société sous tous ses an
 e nombreux où la société civile ne peut l
 l' état mais par la société civile. Cela n
 eut - on dire que la société civile s' est r
 ue, au motif que la société spectaculaire
 a télévision ? si la société se délite, elle
 t du terrorisme la société civile italienne
 d, niché dans notre société civile ? peut
 ettoisaison de notre société nous ne saurio
 on radicale de notre société qui, en se tran
 réveille, et notre société illuminée s' ét
 torial - la première société avancée qui

bleu marine
 bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs de
 couleurs fraîches
 couleurs sourdes
 couleurs de sable
 couleurs vives
 gris bleus de terre



2/ Corpus, web et « humanités » : représentations et interprétations

Que visualise-t'on ?
 Graphes ? Nœuds ?
 Liens ? Arcs ?

Algorithmes et
 interprétations des
 représentations

Jean-Christophe
 Plantin, Manuel
 d'analyse du web,
 Colin, 2013

The Fukushima online
 issue map: spheres of
 actors involved

Jean-Christophe Plantin, 2012
 Cartonomics.org



- governmental sources
- prefectures
- mapmaking sphere
- mapping mashup
- industries
- international organizations
- universities
- anti-nuclear activists
- children & mother defense group
- independant bloggers
- citizen groups



en, la liberté, la société, et donné à ceu
 ire, elle a mène la société sous tous ses an
 e nombreux où la société civile ne peut l
 l' état mène par la société civile. Cela n
 eut - on dire que la société civile s' est r
 ue, au motif que la société spectaculaire
 a télévision ? si la société se délite, elle
 t du terrorisme ? la société civile italienne
 d, niché dans notre société civile ? peut
 ettoisement de notre société, nous ne sauro
 on radicale de notre société qui, en se tran
 réveille, et notre société illuminée s' ét
 torial - la première société avancée qui

bleu marine
 bleu marine
 bleu marine et
 bleu marine rouges
 bleu nuit
 bleu pâle
 couleur de
 couleur de sable
 couleurs de terre
 couleurs de
 couleurs fraîches
 couleurs sordides de sable
 couleurs vives
 gris bleus de terre
 gris mouette
 jaune fluo



Pour conclure...

- Dimension recherche vs applicative
- **Enjeux scientifiques** : formats, documentation des corpus, clôture, porosité du web, renouvellement problématique (Cf. S. Wojcik, « participation politique »)...
- Questions et conditions d'appropriation pour chercheurs ?



en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile ...
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme ? la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société nous ne sauzio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui ,

bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo



En guise de transition...

- Une nécessaire créativité et une articulation approches quantitatives et qualitatives pour documenter les corpus (comprendre les dispositifs, accès aux cahiers des charges) et interpréter les résultats (entretiens compréhensifs, observations ethnographiques...)
- Question du métalangage : **Florence CLAUD**
- Question de l'articulation approches quantitatives et qualitatives : **Valérie BEAUDOUIN**

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mais par la société civile gala
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisiation de notre société vous ne sauzio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui ,

bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en de
couleurs fraîches
couleurs sordides de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo

Quelques références bibliographiques

- BARATS C., LEBLANC JM, FIALA P., « Approches textométriques du web : corpus et outils », In *Manuel d'analyse du web*, Paris, Armand Colin, Collection U, janvier 2013, pp.100-124.
- BENZECRI J-P., *Histoire et préhistoire de l'analyse des données*, Paris, Dunod, 1982.
- BENZECRI, J-P., *Linguistique et lexicologie*, Paris, Dunod. BRUNET, E., 1981, *Le vocabulaire français de 1789 à nos jours, d'après les données du Trésor de la langue française*. Genève-Paris, Slatkine-Champion, 2007.
- BRUNET E., « Un corpus français de 44 milliards de mots. Plongée dans le GOOFRE », *Cahiers de lexicologie*, à paraître, 2014.
- BONNAFOUS Simone, TEMMAR Malika (éds, *Analyse de discours et sciences humaines et sociales*, Paris, Ophrys), 2007.
- CHARAUDEAU P., MAINGUENEAU D., *Dictionnaire d'analyse du discours*, Seuil, 2002.
- CHATEAURAYNAUD, F., *Prospéro - Une technologie littéraire pour les sciences humaines*. Paris, CNRS Editions, 2003.
- CIBOIS, Philippe, *L'analyse factorielle*, Paris, PUF, 1994.
- CIBOIS, Philippe. *Les méthodes d'analyse d'enquête*. Nouvelle édition [en ligne]. Lyon : ENS Éditions, 2014 (généré le 09 mars 2014). Disponible sur Internet : <<http://books.openedition.org/enseditions/1443>>. ISBN : 9782847885743.

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mène par la société civile gala 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée s ' ét
torial - la première société avancée qui ,

bleu marine
bleu marine
bleu marine et
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs de
couleurs fraîches
couleurs sordides de sable
gris bleus de terre
gris mouette
jaune fluo

Quelques références bibliographiques

- DESROSIERE A., *Gouverner par les nombres, L'argument statistique* (tomes I et II), Paris, Presses de l'Ecole des Mines, 2008.
- ERTZSCHEID, O., GALLEZOT, G., SIMONNOT, Br., « A la recherche de la « mémoire » du web : sédiments, traces et temporalités des documents en ligne », CH3 pp 53-68, In C. Barats (dir) *Manuel d'analyse du web*, Paris, Armand Colin, 2013.
- FIALA, P., *L'interprétation en lexicométrie. Une approche quantitative des données lexicales. Langue française n° 103*, p. 113-122, 1994.
- FIALA Pierre, 2007, « L'analyse de discours, analyse de contenu, statistique lexicale, approche sémantico-énonciative », dans BONNAFOUS Simone, TEMMAR Malika (éds), *Analyse de discours et sciences humaines et sociales*, Paris, Ophrys.
- FRANÇOIS, B., NEVEU, E., *Espaces publics mosaïques, acteurs, arènes et rhétoriques des débats publics contemporains*, Rennes, PUR, Res Publica, 1999.
- GOODY J., *La raison graphique. La domestication de la pensée sauvage*, Paris, Minit, 1979.
- KRIEG-PLANQUE A., *La notion de « formule » en analyse de discours, Cadre théorique et méthodologique*, Presses Universitaires de Franche-Comté, Besançon, 2009.

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mène par la société civile gala 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisiation de notre société vous ne sauzio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine
bleu marine et
bleu marine rouges
bleu nuit
bleu pâle
couleur de
couleur de sable
couleurs de terre
couleurs en
couleurs fraîches
couleurs sourdes de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo

Quelques références bibliographiques

- LEBLANC JM, BARATS C, « Exploration de corpus multimodaux pour l'analyse d'un processus de médiatisation : l'exemple du « classement de Shanghai » dans la presse francophone (2003-2010) et de son incidence sur la présentation de soi des universités sur leur pages Web », in actes du colloque JADT 2012, Liège (Belgique), 13-15 juin 2012, pp. 81-93, 2012.
- LEMERCIER C. & ZALC C, *Méthodes quantitatives pour l'historien*, Paris, La Découverte, 2008.
- MAINGUENEAU D., *Analyser les textes de communication*, Paris, Nathan, 2000.
- MAYAFFRE Damon, « Analyses logométriques et rhétoriques des discours » dans OLIVESI Stéphane (dir), *Introduction à la recherche en SIC*, Grenoble, PUG, 2007.
- MAYAFFRE D., *Nicolas Sarkozy. Mesure et démesure des mots*, Grenoble, PUG, 2012.
- OLIVESI, S., *Sciences de l'information et de la communication. Objets, savoirs, discipline*, PUG, (2de édition, mise à jour), 2013.
- PERES, M, LEBLANC JM, « Modèles tridimensionnels pour la représentation de l'état des connaissances et propositions de visualisation pour l'analyse des corpus textuels, JADT, à paraître, 2014.
- PECHEUX M., *Analyse automatique du discours*, Paris, Dunod, 1969.
- PLANTIN, JC, « D'une carte à l'autre : le potentiel heuristique de la comparaison entre graphe du Web et carte géographique », in *Manuel d'analyse du Web en Sciences Humaines et Sociales*, Barats C. (dir.), Armand Colin, Paris, pp. 228-242, 2013.

en , la liberté , la société , et donné à ceu
ire , elle a mène la société sous tous ses an
e nombreux où la société civile ne peut l
l ' état mène par la société civile gala 7
eut - on dire que la société civile s ' est r
ue , au motif que la société spectaculaire
a télévision ? si la société se délite , elle
t du terrorisme la société civile italienne
d , niché dans notre société civile ? peut
ettoisation de notre société nous ne saurio
on radicale de notre société qui , en se tran
réveille , et notre société illuminée
torial - la première société avancée qui

bleu marine
bleu marine
bleu marine et
bleu marine et
bleu marine rouges
bleu nuit
bleu nuit
couleur de
couleur de sable
couleurs de terre
couleurs de terre
couleurs fraîches
couleurs sordides de sable
couleurs vives
gris bleus de terre
gris mouette
jaune fluo

Quelques références bibliographiques

- SALAÜN, JM, *Vu, lu, su, Les architectes de l'information face à l'oligopole du web*, Paris, La Découverte, 2012.
- SALEM, A., *Analyse factorielle et lexicométrie: synthèse de quelques expériences*. *Mots*, 1982.
- TOURNIER, M., *Spécificité politique et spécificité lexicale*, *Mots* n° 2, 1981.
- TOURNIER, 1986, *La lexicométrie socio-politique*. *Le courrier du CNRS*, 65, p.24-32, 1986.
- TOURNIER, M., *De France à Je. La traversée des emplois Cooccurrences et connexions*. Des sources du sens, École Normale Supérieure Lettres Sciences Humaines Lyon, Collection Langages , 2003.
- TOURNIER, M., « Mots et politique, avant et autour de 1980 Entretien », *Mots. Les langages du politique* [en ligne], N°94, 2010.
- VENIARD, M., *La nomination des événements dans la presse. Essai de sémantique discursive.*, Presses universitaires de Franche-Comté, 2013.
- VERON, E., *Construire l'événement, les médias et l'accident de three mile Island*, Paris, Editions de Minuit, 1981.