



HAL
open science

La translittératie, ALCESTE et les Humanités Numériques. Présentation du Projet ANR Translit, Tâche 1, “ Epistémologie et circulation des savoirs ” : une analyse de corpus avec ALCESTE.

Alexandre Serres

► **To cite this version:**

Alexandre Serres. La translittératie, ALCESTE et les Humanités Numériques. Présentation du Projet ANR Translit, Tâche 1, “ Epistémologie et circulation des savoirs ” : une analyse de corpus avec ALCESTE.. 2015. sic_02525106

HAL Id: sic_02525106

https://archivesic.ccsd.cnrs.fr/sic_02525106v1

Preprint submitted on 30 Mar 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 - NonCommercial - NoDerivatives 4.0 International License

La translittératie, ALCESTE et les Humanités Numériques

**Présentation du Projet ANR Translit, Tâche 1
« Epistémologie et circulation des savoirs » :
une analyse de corpus avec ALCESTE
8 avril 2015**

**Alexandre Serres, MCF,
Université Rennes 2, PREFics,
co-responsable de l'URFIST de Rennes**

**“Qu'est-ce qu'un projet scientifique en humanités numériques ?”
Séminaire MSH-B / Humanités numériques - 2014-2015**

Qu'est-ce qu'un projet scientifique en Humanités Numériques ?

- Difficultés de la question
- **6 critères possibles** ; avoir le numérique comme :
 - **matériau principal de la recherche**
 - **outil de la recherche**
 - **support de diffusion**
 - **vecteur transdisciplinaire**
 - **regroupement de compétences**
 - **objet de recherche lui-même**
- **Mais peut-être 2 autres conditions nécessaires** :
 - **Combiner simultanément majorité de ces critères**
 - **Se revendiquer des Humanités Numériques**

En guise de plan

- **Translittératie, Projet ANR TRANSLIT**
- **Un projet d'analyse de corpus**
- **ALCESTE**
- **Et les Humanités Numériques ?**

Translittératie ?

l'origine linguistique

- De la « transliteration » à la « transliteracy »
- *« le mot translittératie est un dérivé du verbe to transliterate, qui signifie écrire ou imprimer une lettre ou un mot en utilisant les lettres les plus proches provenant d'un autre alphabet ou d'une autre langue. » (Sue Thomas)*

Translittératie : la définition canonique

- « *l'habileté à lire, écrire et interagir par le biais d'une variété de plateformes, d'outils et de moyens de communication, de l'iconographie à l'oralité en passant par l'écriture manuscrite, l'édition, la télé, la radio et le cinéma, jusqu'aux réseaux sociaux* » (Sue Thomas and al.)

La translittératie : un concept très large

- Un ensemble de compétences d'interaction :



D'où vient ce concept de *transliteracy* ?

- **Aux Etats-Unis** : travaux d'Alan Liu

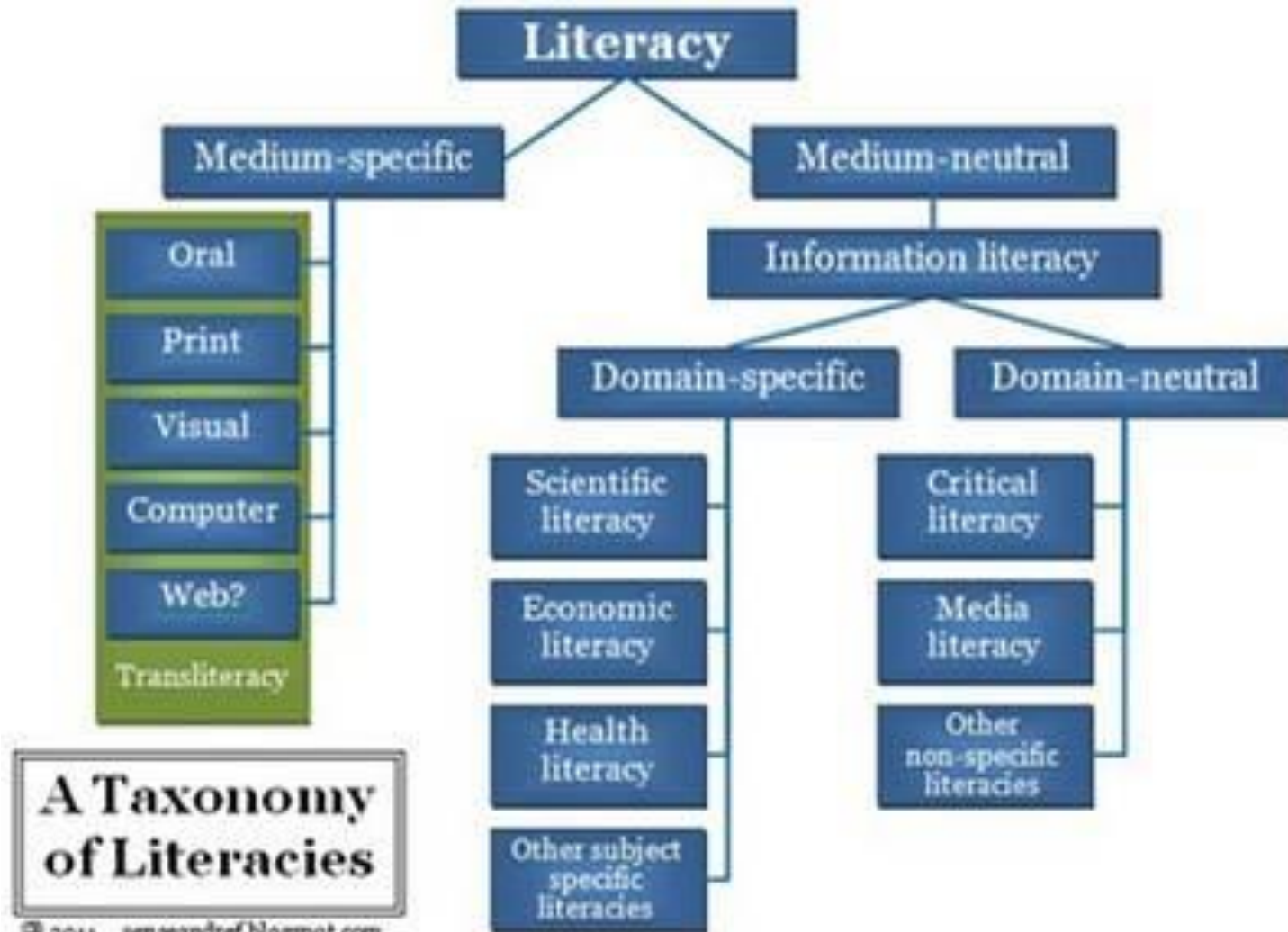
- Departement of English, UCSB (University of California at Santa Barbara)
- [Transliteracies Project](#), 2005
- Recherche sur la lecture en ligne

- **En Grande-Bretagne** : travaux de Sue Thomas

- Faculty of Arts, Design and Humanities, De Monfort University, in Leicester
- 2006 : création du [PART group](#) (Production and Research in Transliteracy)
- Disciplines littéraires
- Impact des TIC sur l'écriture et la lecture

> **Un concept au cœur des Digital Humanities...**

Comment penser les multiples *litteracies* ?



Approche globale ou analytique ?

- Comment penser l'hybridation des littératies ?
- Approche globale :
 - Translittératie, mais aussi la « *digital literacy* » (la culture numérique)
 - "*transliteracy does not replace, but rather contains, "media literacy" and also "digital literacy"*" (S. Thomas)
 - « *A unifying ecology* » : référence à la convergence des médias (Jenkins), à l'écologie des médias (McLuhan)
 - Translittératie comme la “mère de toutes les littératies” ?
- Approche plus analytique :
 - Identifier, définir, distinguer toutes les littératies
 - Etudier leurs relations, leurs convergences/divergences...
 - Plutôt l'approche française, notamment pour les chercheurs du projet ANR TRANSLIT

Aspects scientifiques du projet :

les objectifs

- Prend la suite de projets précédents (ERTé...)
- **Projet ANR**  **Translit**
 - Université Sorbonne Nouvelle, Divina Frau-Meigs
 - « *La translittératie : vers la transformation de la culture de l'information* »
 - 2013-2016
- **4 équipes de recherche, structurées autour de 4 "Tâches" :**
 - **Tâche 1 "Epistémologie et cartographie"**, Eric Delamotte, Université de Rouen, EA GHRIS
 - Tâche 2 "Dispositifs et usages", Vincent Liquète, Université de Bordeaux, IMS-CIH
 - Tâche 3 "Traitement des données", Eric Bruillard, ENS Cachan, STEF
 - Tâche 4 "Modélisation et retombées", Divina Frau-Meigs, Sorbonne Nouvelle, CREW
- Au total, plus d'une vingtaine de chercheurs mobilisés.
- Point commun : exploration des trois cultures (informatique, médiatique, informationnelle), de leurs enjeux, de leurs usages et des problématiques liées

Objectifs de la Tâche 1: « Épistémologie et Cartographie »

- **Objectifs :**
 - Analyse transversale des trois littératies à travers des corpus (textes, sites...) ;
 - révéler la circulation des savoirs impliqués et en proposer des cartographies dynamiques.
 - Ces savoirs concernent les champs scientifiques, professionnels et institutionnels.
 - « Observer l'évolution des niveaux scientifiques, éducatifs et sociaux autour de l'hypothèse de la convergence des domaines de l'info-médias, l'info-doc et de l'info-data, dans une perspective translittéracique »

Aspects méthodologiques : méthodologie générale du projet ?

- **Méthodologie croisée, avec deux types d'outils :**
 - **Analyse lexicale d'un corpus de textes, avec un outil reconnu, Alceste**
 - logiciel d'aide à l'analyse des discours,
 - **Analyse cartographique d'un corpus de ressources sur le web (sites, textes divers...), menée avec des outils :**
 - de collecte de données (ALADOC)
 - de cartographie automatique (Hyphe)
 - de visualisation (Gephi)
 - Action encore en projet, non présentée ici

Aspects méthodologiques : quel corpus ?

- **De l'importance du corpus :**
 - objet et matériau de recherche
- **2 problèmes méthodologiques :**
 - Représentativité
 - Subjectivité
- **3 principes simples :**
 - Comparabilité des 3 sous-corpus :
 - des trois domaines : media, informatique, infodoc
 - Cohérence interne :
 - Même type de textes
 - Des variables et des critères communs

Aspects méthodologiques :

Quels critères pour le corpus ?

- **Sept variables :**

- **La thématique :**

- les 3 éducations à...
(informatique, médias, infodoc)

- **L'influence :** choix de textes « d'influence »

- Combinaison d'indicateurs et du « regard expert »

- Quelles influences ?

- Scientifique
- Institutionnelle
- Professionnelle

- **Le nombre :**

- 15 textes par domaine ; total de 45

- **La nature des textes :**

- Textes scientifiques, académiques
- Textes institutionnels
- Textes professionnels, associatifs

- **La langue :**

- Textes francophones, ou traduits en français

- **La variable temporelle :**

- Plutôt après 2000, mais aussi textes antérieurs

- **La variable technique :**

- Des textes numérisés

Aspects méthodologiques : Quel traitement sur les données avec quel outil ?

- **Traitement d'analyse lexicale, avec ALCESTE**
 - **Analyse des Lexèmes Co-occurrents dans les Enoncés Simples d'un Texte**
 - Outil d'analyse lexicale de données textuelles, créé en 1986 :
 - « *outil d'aide à l'interprétation d'un corpus textuel* » (Max Reinert)
 - Société Image, soutien du CNRS
 - Objet d'Alceste :
 - de quoi et comment parle-t-on dans un corpus donné de textes ?
 - Principes de fonctionnement :
 - Segmentation de textes et regroupement des formes
 - Regroupement des éléments de discours selon des liens de proximité lexicale
 - Mise en évidence de mondes lexicaux (« classes »), dont la signification doit être donnée par le regard du chercheur

Repères sur le fonctionnement d'ALCESTE

- **Trois types de segmentations du corpus :**
 - **un découpage en UCI (Unités de Contexte Initiales) :**
 - une UCI = un texte, un chapitre, toute partie bien identifiée du corpus ;
 - l'utilisateur définit lui-même les UCI ;
 - **une segmentation en UC (Unités de Contexte) :**
 - une UC = un paragraphe, un segment de texte de taille réduite, une phrase, un énoncé ;
 - **une segmentation en UCE (Unités de Contexte Élémentaire) :**
 - Une UCE = ensemble de formes, de chaînes de caractères séparées par des espaces

Quels premiers résultats ?

- Différentes analyses conjointes :
 - des trois corpus : 45 textes
 - des deux corpus EAM et EAI : 30 textes
- Objet :
 - identification de classes thématiques lexicales, constituant des unités de discours
 - Établissement de cartographies par Alceste :
 - interprétation des logiques discursives des corpus
 - Identification par Alceste de 6 classes lexicales :
 - présentées dans une Classification Descendante Hiérarchique

Taille des points = pourcentage de pertinence = importance des UCE



Liste des mots significatifs pour chacune des classes avec leur coefficient d'association Phi

Classe 4		Classe 5		Classe 1		Classe 6		Classe 7		Classe 2		Classe 3	
Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi
documentali	.44	national	.28	informatique	.40	media	.38	examiner	.24	ordinateur	.30	savoir	.28
etablisseme	.30	presse	.26	enseigne	.33	enfant	.22	film	.24	algorithm	.29	information	.19
documentair	.29	semaine	.24	lycee	.21	mediatique	.19	image	.24	logiciel	.29	connaissanc	.16
documenter	.23	regional	.23	option	.19	critiquer	.18	photograph	.24	machine	.26	apprenti	.16
pedagog	.22	ministere	.20	disciplin	.19	jeune	.17	publicite	.23	utilisation	.21	informationn	.15
politique	.18	centre	.19	college	.19	monde	.15	genre	.19	programmer	.20	competent	.15
personnel	.16	journal	.18	universit	.18	social	.15	public	.18	donnee	.18	culture	.14
ressource	.16	stage	.17	science	.18	consommati	.13	television	.18	informatique	.16	savoir-faire	.14
mise	.14	academi	.17	brevet	.18	message	.12	productif	.17	langage	.16	pertinent	.14
circulaire	.13	commission	.16	secondaire	.16	expressif	.12	monter	.17	systeme	.16	processus	.13
mission	.13	education	.16	primaire	.16	cultur	.12	different	.17	interface	.15	informationn	.13
disposition	.13	stagiaire	.16	niveau	.14	economique	.11	sequence	.16	outils	.15	identifier	.12
professeur	.13	partenari	.15	specialite	.14	societe	.11	produit	.16	traitement	.15	cognitiv	.12
fond	.13	local	.15	enseigner	.14	influencer	.11	texte	.16	manipuler	.15	capacite	.12
projet	.13	liaison	.15	concours	.14	realite	.10	eleve	.16	souris	.14	problem	.11
travail	.12	inspection	.14	terminal	.14	education	.10	mediatique	.16	calcul	.14	source	.11
espace	.11	equipe	.14	pays	.13	adulte	.10	jeu	.15	objet	.13	construire	.11
enseigner	.10	beneficier	.13	introduction	.13	democratie	.10	publicitaire	.15	commande	.13	notion	.11
service	.10	coordonnate	.13	programme	.13	esprit	.10	facon	.15	robot	.13	usage	.10
accueil	.10	lyceen	.13	licence	.13	individu	.10	textuel	.15	voiture	.13	literacy	.10
Absence	Phi	Absence	Phi	Absence	Phi	Absence	Phi	Absence	Phi	Absence	Phi	Absence	Phi
media	-.13	informatique	-.11	documentair	-.08	informatique	-.13	informatique	-.11	media	-.15	media	-.13
informatique	-.13	documentair	-.08	media	-.08	documentair	-.10	formation	-.10	education	-.13	etablisseme	-.13
mediatique	-.08	savoir	-.07	action	-.08	documentali	-.09	documentali	-.08	formation	-.09	documentali	-.11
communic	-.08	connaissanc	-.07	information	-.07	formation	-.08	etablisseme	-.07	etablisseme	-.08	academi	-.10
monde	-.07	notion	-.07	productif	-.06	disciplin	-.08	savoir	-.07	documentali	-.08	national	-.10

Classes liées à

Dispositifs d'Enseignement

Savoirs et culture

EMI

Communication
médias

Maîtrise de(s)



Classe 4

Classe 5

Classe 1

Classe 6

Classe 7

Classe 2

Classe 3

Documentation

CLEMI

Informatique

Enjeux
sociétaux

Objets des
médias

Objets de
l'informatique

Savoirs
informationnels

Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi	Présence	Phi
documentati	.44	national	.28	informatique	.40	media	.38	examiner	.24	ordinateur	.30	savoir	.28
etablissem	.30	presse	.28	enseigne	.33	enfant	.22	film	.24	algorithm	.29	information	.19
documentair	.29	semaine	.24	lycee	.21	mediatique	.19	image	.24	logiciel	.29	connaissanc	.18
documenter	.23	regional	.23	option	.19	critiquer	.18	photograph	.24	machine	.28	apprenti	.16
pedagog	.22	ministere	.20	disciplin	.19	jeune	.17	publicite	.23	utilisation	.21	informationn	.15
politique	.18	centre	.19	college	.19	monde	.15	genre	.19	programmer	.20	competent	.15
personnel	.16	journal	.18	universit	.18	social	.15	public	.18	donnee	.18	culture	.14
ressource	.16	stage	.17	science	.18	consommat	.13	television	.18	informatique	.16	savoir-faire	.14
mise	.14	academi	.17	brevet	.18	message	.12	productif	.17	langage	.16	pertinent	.14

Acquisitions par la pratique

informel

activités

Examiner	24
Film	24
Image	24
Photograph	24
publicité	23

Objets des médias

SPME

CLEMI	32
National	28
Presse	26
Semaine régional	24
	23

Classe 7

Classe 5

Enjeux éducatifs

Média	36
Enfant	22
Critiquer	18
Monde	15
Société	11

Enjeux sociétaux

Classe 6

Documentation 26%

Savoir	26
Information	19
Connaissance	16
Apprenti	16
informationnelle	15

Savoirs

Classe 3

Documentaliste	44
Établissement	30
Documentaire	29
C.D.I.	28
documenter	23

Classe 4

Ordinateur	30
Algorithme	29
Logiciel	29
Machine	26
utilisation	21

Objets de l'informatique

Informatique

Informatique	40
Enseignement	33
Lycée	21
Option discipline	19
	19

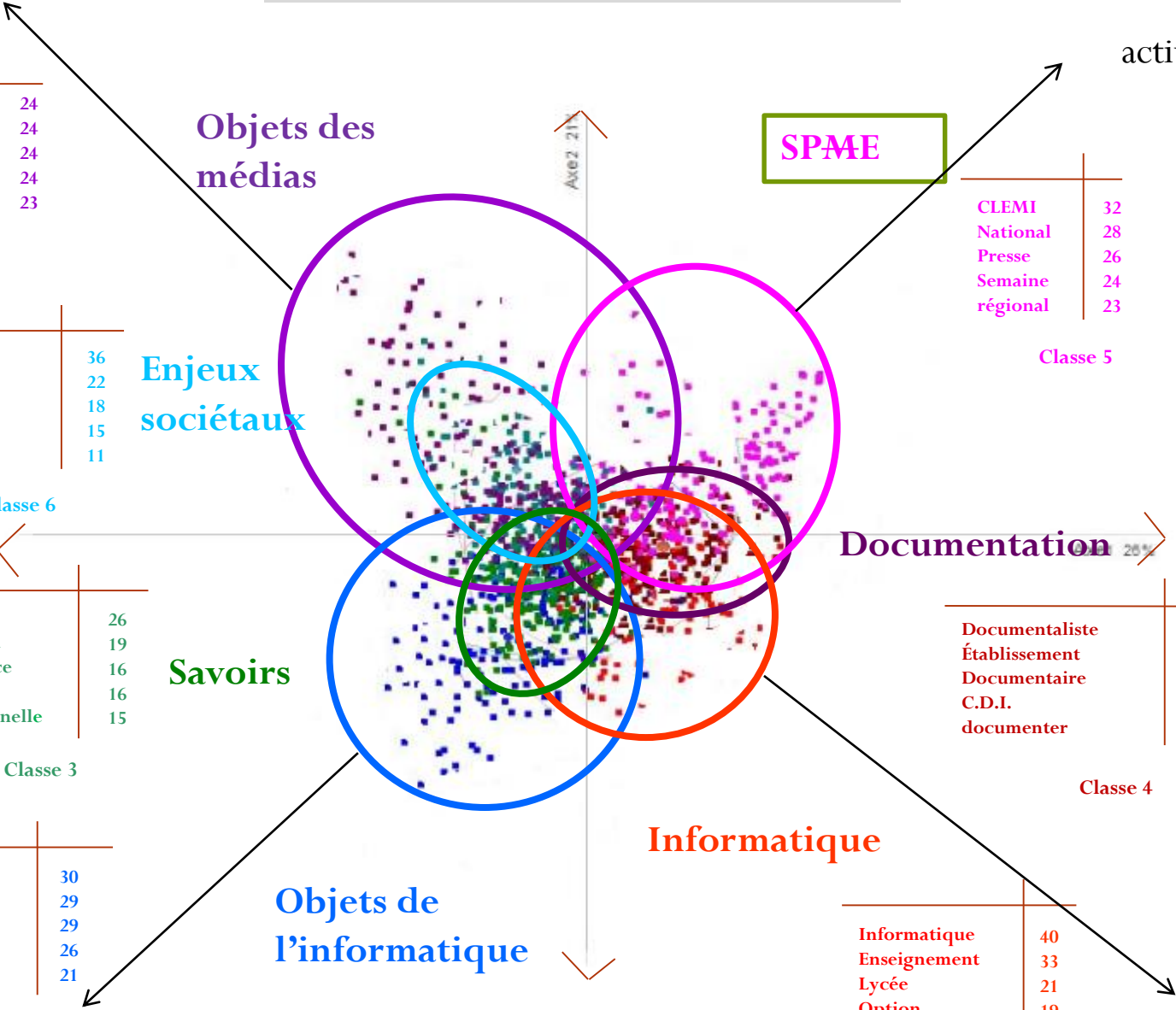
formel

Dispositifs scolaires

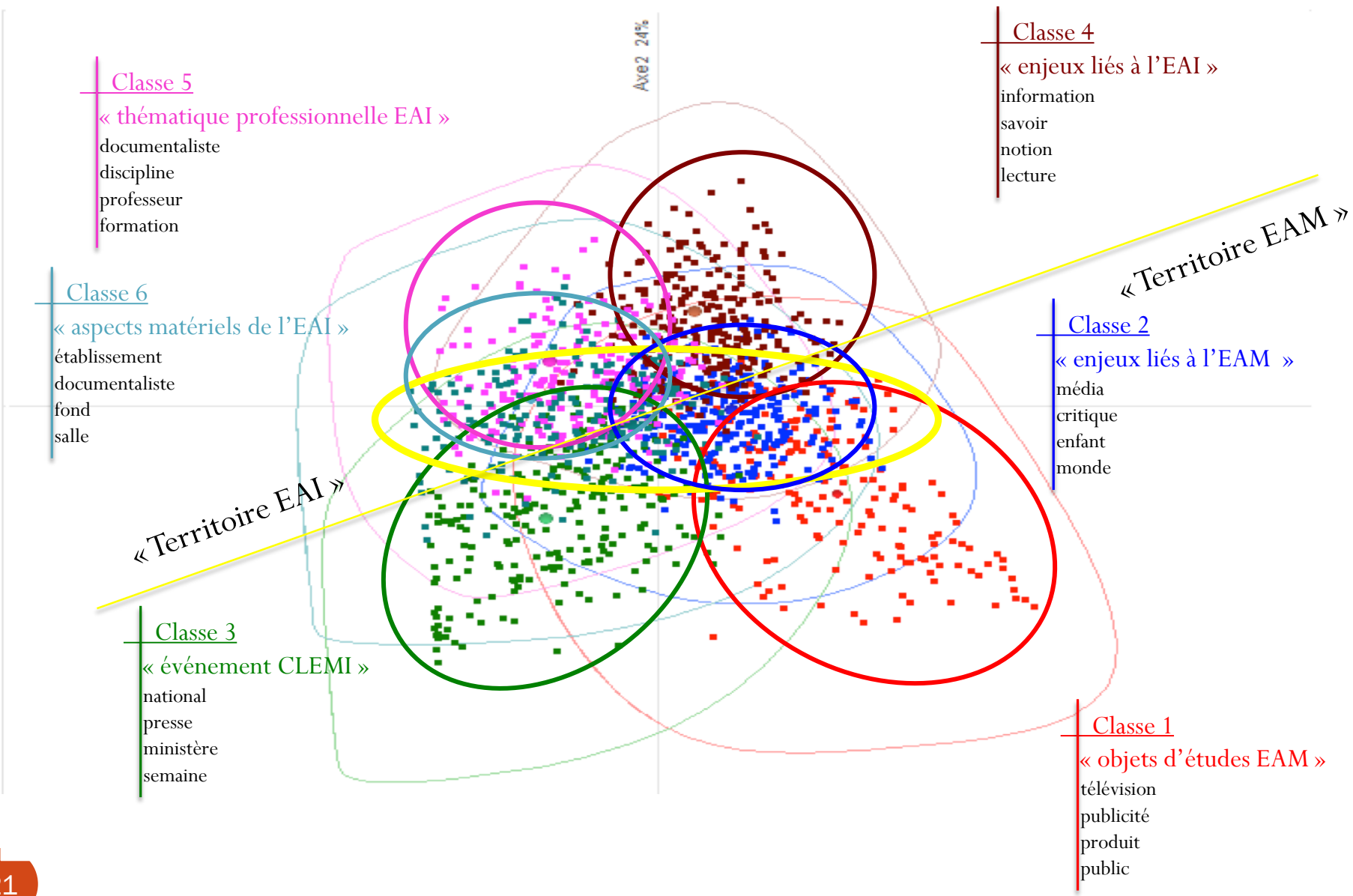
non formel

Enseignables

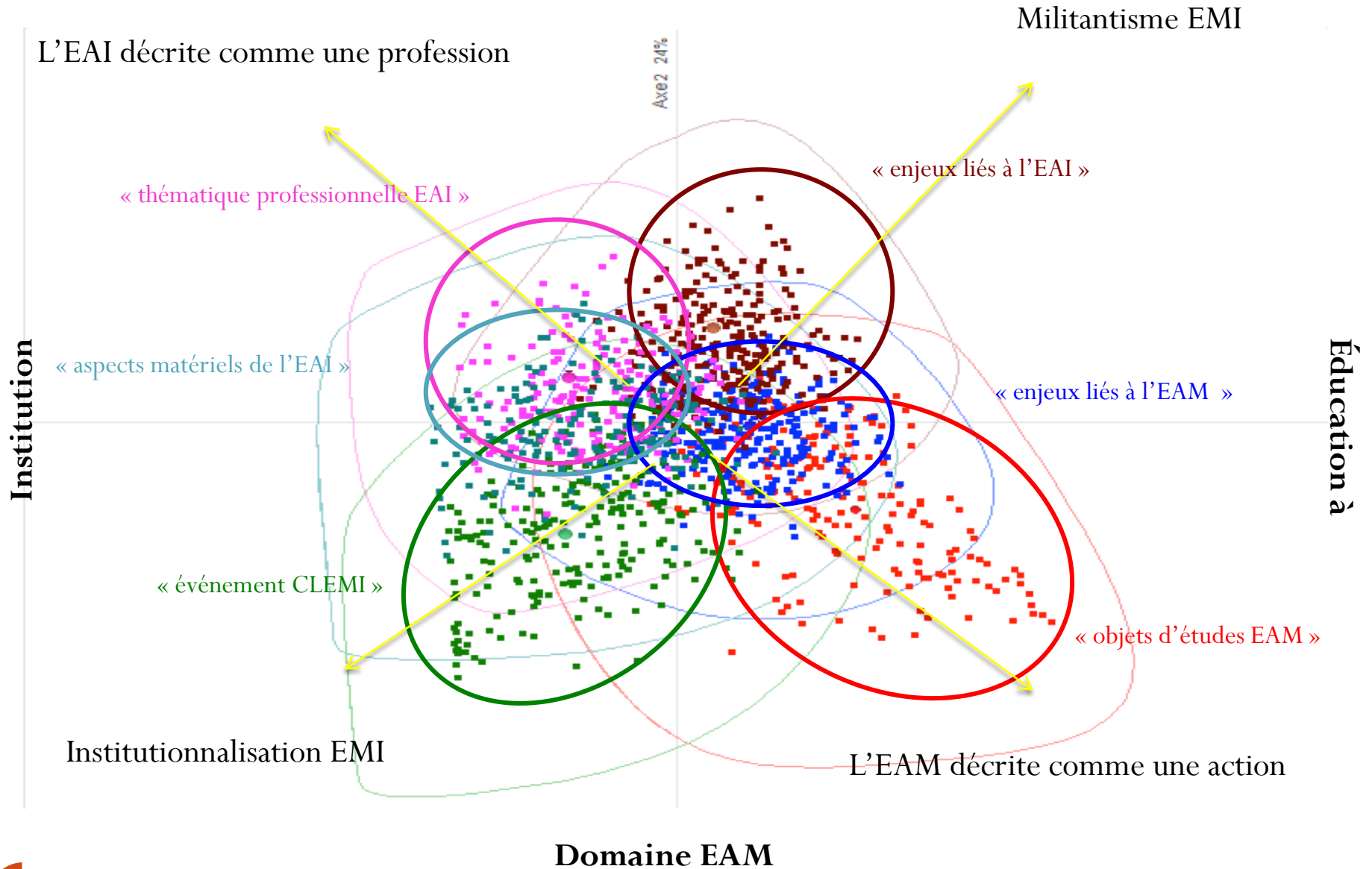
Classe 1



Analyse factorielle en coordonnées : sur 2 corpus EAM/EAI



Domaine EAI



Premières conclusions sur l'EMI (Education aux Médias et à l'Information)

- Deux domaines distincts construits selon des logiques et des particularités différentes
- Des points de convergence autour des enjeux et de l'opérabilité de l'EMI
- Des logiques professionnelles, institutionnelles et scientifiques distinctes à creuser
- De nouvelles questions à poser pour l'EMI...

"L'éducation aux médias et à l'information est un domaine complexe qui pose encore la question de sa définition et de son périmètre". (Rapport ANR Translit, COST, Politiques d'éducation aux médias et à l'information en France.)

Et les Humanités Numériques ?

Question : Quelle est la part du numérique dans le projet ?

- Omniprésence du numérique dans le projet ANR TRANSLIT
- 7 critères remplis, sur les 8 énoncés au début, dans la Tâche 1
- **a) Le numérique comme matériau de la recherche :**
 - Projet de recherche portant en quasi-totalité sur des corpus de textes numérisés, ou des ressources web
- **b) Le numérique comme outil de la recherche :**
 - Les outils numériques au cœur du projet de la Tâche 1 ANR Translit : Alceste, Gephi, Hyphe...
- **c) Le numérique comme support de diffusion :**
 - *Question : Quelles modalités de diffusion et de publication des résultats ? Quel lien avec l'Open Access ?*
 - Site web du projet ANR, projet de diffusion de textes sur les Archives Ouvertes, et mise en ligne des données sur le site.

Et les Humanités Numériques ?

- **d) Le numérique comme vecteur transdisciplinaire :**
 - *Question : Quel est / quels sont le(s) champ(s) disciplinaire(s) d'origine : SHS, ALL, STIC, Informatique... ?*
 - Plusieurs disciplines représentées : SIC, Sciences de l'Education, Informatique, Psychologie, Linguistique...
 - **Deux grands champs disciplinaires concernés :**
 - **Les SHS** : trois sous-domaines spécifiques :
 - **les sciences de l'information-documentation** et la réflexion sur la culture de l'information-documentation, sur les usages ;
 - **l'éducation aux médias** et la réflexion sur la notion de média
 - **les sciences de l'éducation**, et la réflexion sur la didactique
 - **L'Informatique** : dimension surtout éducative de l'informatique.

Et les Humanités Numériques ?

- e) **Le numérique comme facteur de regroupement de compétences :**
 - *Question : Quels types de profils professionnels requis pour le projet ?*
 - Dans Tâche 1 :
 - des enseignants-chercheurs, formateurs en info-doc
 - un ingénieur linguiste, spécialisé sur Alceste
 - des informaticiens d'une SSII
 - *Question : Quelles sont les compétences requises ?*
 - Compétences théoriques, disciplinaires (infocom, linguistique), informatiques (maîtrise des outils de text mining) ;
 - **Principale difficulté dans la Tâche 1** : manque de compétence et de formation dans l'utilisation d'Alceste ;

Et les Humanités Numériques ?

- **f) Le numérique comme objet de recherche :**
 - Objet de la recherche du projet : la convergence des cultures de l'information, liées au numérique, la translittératie...
 - *Question : Quels sont les concepts transversaux ou disciplinaires en lien avec le numérique ?*
 - les cultures de l'information : définition, périmètre, contenus, acteurs, circulation des savoirs...
 - les pratiques informationnelles, médiatiques et numériques
 - la didactique de l'information
 - la notion de translittératie
 - la notion de culture numérique
 - l'EMI
 - les notions essentielles d'informatique

Et les Humanités Numériques ?

- **g) La combinaison des 6 critères :**
 - Présence simultanée des critères précédents
 - > Donc un projet en HN ?
- **h) Quelle revendication des Humanités Numériques ?**
 - Très faible voire inexistante :
 - Méconnaissance des Humanités Numériques par nombreux chercheurs
 - Pas « d'affiliation » explicite, revendiquée au mouvement des Humanités numériques, mais reconnaissance par certains que le projet fait partie de fait des HN
 - Question secondaire, en débat...

En conclusion

- **Retour sur la difficulté de caractérisation des projets en HN :**
 - Dimension réflexive importante : quel impact du numérique sur les SHS ?
 - Sur ce point, le Projet ANR Translit n'est pas un projet en HN
- **La question-clé des compétences et des cultures multiples :**
 - Disciplinaire, informatique, technique...
- **La dimension performative des Humanités Numériques :**
 - un projet en HN = un projet qui s'affiche et se revendique comme tel ?

Merci de votre attention !

Contact :

alexandre.serres@uhb.fr