



HAL
open science

Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques

Patrick Trabal

► **To cite this version:**

Patrick Trabal. Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques. Bulletin de Méthodologie Sociologique / Bulletin of Sociological Methodology, 2005, 85, p.10-43. <hal-03016322>

HAL Id: hal-03016322

<https://hal.parisnanterre.fr/hal-03016322v1>

Submitted on 20 Nov 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.



HAL Authorization

Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques

Patrick Trabal

- 1 Lors du premier congrès de l'Association Française de Sociologie (AFS), fut distribué un ouvrage contenant les résumés des 1068 communications annoncées dans quelque 43 réseaux thématiques¹. Ce rassemblement apparaît comme une opportunité sans précédent d'avoir un état de la sociologie française, même si demeurent des incertitudes quant à la représentativité de ce corpus par rapport à la discipline.
- 2 Voici, à première vue, un corpus typique de ceux que l'on analyse à l'aide du logiciel Prospéro : il s'agit d'un grand nombre de textes, tous écrits dans la même langue, par des acteurs de la même discipline, partageant un vocabulaire et des concepts communs même si leurs préoccupations diffèrent. On peut donc soumettre cet ensemble à des opérations comparatives et faire apparaître des différences selon les auteurs ou leurs appartenances institutionnelles, ainsi que des rapprochements autour de certains objets, certaines questions, certains concepts. Deux propriétés fondamentales des séries textuelles à l'analyse desquelles est voué le logiciel sont cependant absentes de cette collection de résumés :
 - *L'articulation autour d'un même objet ou d'une problématique commune* aux différents auteurs, sur la nature, l'importance, la réalité, les enjeux desquels on cherche à produire un accord : en effet, Prospéro a vocation à aider le chercheur à identifier des propriétés sur des objets d'alertes et de controverses, sur la construction des jugements et sur les argumentations déployées dans des dossiers marqués par de longues séries de textes et de discours hétérogènes². Or, les résumés présentés au colloque de l'AFS, rédigés sans référence les uns aux autres, ne visent pas à argumenter autour d'un objet de dispute, mais à présenter des travaux réalisés séparément – qui plus est sous une forme résumée ou programmatique, c'est-à-dire dans un format qui s'exprime notamment par une contrainte de longueur et se prête donc peu au déploiement d'argumentations caractérisant les dossiers habituellement étudiés à l'aide du logiciel.

- *L'épaisseur temporelle de la collection* : nombreuses sont les fonctions du logiciel qui visent à voir *apparaître, disparaître* ou *se transformer* des personnages et des objets dans le dossier dont rend compte la série textuelle, et à décrire la façon dont les arguments y sont repris et critiqués, les accords remis en question ou au contraire renégociés, les preuves rediscutées ou stabilisées. Or, dans notre collection de résumés, la production des discours est non seulement indépendante, mais en outre simultanée.

- 3 Que Prospéro peut-il alors apporter au chercheur désireux d'analyser ce corpus ?³ Si l'on veut mobiliser dans les meilleures conditions ce logiciel sur cet ensemble de textes, il est nécessaire de construire un objet se rapprochant de ceux qui préoccupent les développeurs-utilisateurs de Prospéro. Dans cette perspective, on va regarder le corpus comme un *dossier*, c'est-à-dire comme un ensemble de textes dans lesquels se discutent des notions, s'opposent des points de vue et se déploient des argumentations.
- 4 Sans proposer une histoire de ce congrès, il est sans doute utile de préciser qu'il fut précédé de différentes réunions au cours desquelles des réseaux thématiques en formation (RTF) furent constitués. En tenant compte de cette phase au cours de laquelle les sociologues se sont regroupés, on identifie un autre moment, antérieur à la rédaction des résumés, celui qui a débouché sur la rédaction des « appels à communication » de chaque RTF. On peut formuler l'hypothèse selon laquelle ces « appels » ont pesé sur les sociologues dans la rédaction des résumés sous des formes différentes. Il s'agissait en effet de rédiger leur texte en s'ajustant au mieux à la présentation du réseau choisi pour que la proposition de communication soit acceptée. Même si cela n'est qu'une formalité pour les sociologues confirmés, la soumission d'un texte dans un RTF, plutôt que dans un autre où la même proposition aurait pu être acceptée, peut être sujette à des incertitudes et à des enjeux. Ceux-ci peuvent porter sur la notoriété du RTF, liée en partie aux animateurs du groupe, ou sur le public escompté : faut-il assurer une communication dans un RTF où « l'on connaît bien les gens », ou faut-il au contraire essayer de se faire connaître ou reconnaître dans un réseau dans lequel il y aura X ou Y ? On peut donc penser ce corpus comme un ensemble de productions d'acteurs tentant de franchir une épreuve, de difficulté variable selon les personnes, pour être retenus dans le réseau de leur choix. Joindre à ce corpus de résumés les appels à communication qui, pourrait-on dire, les ont générés, permet alors d'introduire une variabilité portant non seulement sur la forme (le texte d'un appel à communiquer est rédigé généralement de façon œcuménique afin d'obtenir de nombreuses réponses), mais aussi sur le moment de rédaction.
- 5 Nous chercherons donc à évaluer ces effets d'ajustement aux appels à communication – ce qui nous conduira à recenser la proximité entre les différents textes – et, d'une façon plus générale, à repérer les contraintes et les ressources des auteurs dans l'écriture de leurs propositions. Mais auparavant, il semble utile de rendre compte du travail nécessaire pour utiliser le logiciel.

Travailler avec le logiciel Prospéro

- 6 Le logiciel fonctionne à partir des textes en langage naturel. Parce que le chercheur bénéficie des dictionnaires et autres ressources des précédents utilisateurs, l'on peut immédiatement recenser des listes d'éléments présents dans le corpus. Notre attention est alors portée sur les objets dont les poids sont importants. Le tableau des premiers éléments de la liste des entités de notre série de résumés figure en annexe (Tableau 1). On

voit qu'il est avant tout question de travail, d'analyse et de sociologie... On pourrait se contenter de noter cela, en remarquant que ces mots traduisent assez correctement notre activité. Toutefois, il est clair que le mot « travail » peut intervenir dans des expressions comme « sociologie du travail », expression que l'on peut construire pour la distinguer de celle-ci : « travail en sociologie ». Cette activité de codage des expressions est une des modalités centrales du travail avec le logiciel. Il fonctionne en effet sur l'idée selon laquelle le chercheur doit assumer la responsabilité des opérations de codage et d'interprétation, et que ce travail se réalise tout au long de l'exploration du corpus et non, comme le suggère une tradition de l'analyse des données, à l'issue du calcul informatique.

Tableau 1 : Liste des 30 premières entités, obtenue avant la construction des expressions

ENTITE	Poids (Nb d'occurrences)
travail	621
analyse	501
sociologie	382
processus	355
femmes	341
communication	336
recherche	320
étude(s)	306
acteurs	306
question	271
formes	271
LA-FRANCE@	261
relations	259
jeunes	252
rapports	239

Tableau 1 (suite)

ETAT-CENTRAL@	233
TRAVAILLEURS@	233
CHERCHEURS-ENSEIGNANTS@	229
ENFANTS@	227
LA-PLANETE@	225
construction	224
action	217
cadre	201
EUROPE@	196
objet	193
ECCOLE@	189
wie	189
professionnels	186
production	184
rapport	183

- 7 Cette position, largement développée par Francis Chateauraynaud dans son ouvrage sur Prospéro⁴, apparaît en rupture avec les attentes habituellement générées par l'informatique. En effet, l'on attribue souvent à la machine une mission d'objectivation et une capacité d'administration de la preuve. Cela revient symétriquement à assigner au chercheur un rôle subalterne alors que plusieurs traditions (en particulier compréhensive, pragmatique, phénoménologique et ethno-méthodologique) invitent à penser qu'il n'y a guère de textes sans lecteurs et de lectures sans interprétations. Toutefois, confier au chercheur le soin de fixer ce qui fait sens, selon lui, dans un ensemble de textes, ne conduit pas pour autant à une forme de relativisme. En effet, le logiciel est un moyen d'éprouver ses lectures, d'en suggérer d'autres, de réformer les interprétations qui ne résistent pas à la confrontation avec les éléments présents ou absents dans le corpus, et de faire des expériences sur les textes pour enrichir l'analyse.
- 8 Inutile donc de mobiliser ici les algorithmes visant à établir des co-occurrences entre les éléments du tableau 1 pour coder des expressions comme "sociologie du travail". Doter le chercheur de cette capacité à coder "sociologie du travail" pose néanmoins une série de problèmes. Il s'agit, à un premier niveau, de réserver le même sort à la "sociologie de l'éducation", la "sociologie de la famille", la "sociologie des religions"... Pour s'assurer de ne pas en oublier, on peut rechercher à l'aide de *formules*, des segments du type < « sociologie » - quelques mots outils (comme « de », « du », « des », « de la ») - n'importe quelle entité >. Dans le corpus des résumés du congrès de l'AFS, cette requête donne 136 occurrences sous 57 formes différentes, présentées dans le tableau 2. Cette liste fait apparaître que c'est la sociologie de la connaissance qui réalise le meilleur score (19 apparitions réparties dans 11 textes), alors que la sociologie du travail n'arrive qu'en deuxième position avec 15 occurrences⁵. D'une façon plus générale, si cette liste donne à lire les principaux objets de la discipline (au moins pour les formes ayant reçu plus d'une

seule occurrence), il reste à définir le statut que l'on donne à ces expressions. Faut-il toutes les coder, c'est-à-dire indiquer au logiciel de prendre l'ensemble de chaque expression comme un tout, ou faut-il au contraire traiter « sociologie » et « connaissance » comme deux entités séparées ? On peut répondre à cette question de différentes façons. Par exemple, on peut tenter de donner une réponse sociologique, épistémologique ou philosophique en interrogeant l'ontologie de la « sociologie de la connaissance » ou de la « sociologie du travail ». On peut s'interroger sur l'unité de la sociologie du travail en questionnant l'existence de théories, de méthodes, mais aussi de colloques, de revues qui rassemblent des communautés de chercheurs. Si la question ne semble pas se poser pour la sociologie du travail ou de la famille, qu'en est-il de la sociologie du bénévolat ? des prisons ? du voisinage ? Qui doit faire le choix ? Dans la tradition pragmatique, il semble possible de résoudre cette question en évaluant les conséquences des concepts déployés. En effet, coder une expression comme « sociologie du bénévolat » en indiquant au logiciel de prendre ce segment comme un tout, c'est poser une lecture qui permet à la fois de montrer un certain nombre de propriétés, mais aussi d'en masquer d'autres.

Tableau 2 : Formes du type "sociologie [de, du, de la, des] ENTITE"

sociologie de la connaissance (19) ; sociologie du travail (15) ; sociologie des professions (8) ; sociologie de la culture (7) ; sociologie des médias (6) ; sociologie du sport (6) ; sociologie des organisations (5) ; sociologie de la vieillesse (5) ; sociologie du droit (3) ; sociologie de la famille (3) ; sociologie de la littérature (3) ; sociologie de la consommation (3) ; sociologie de la jeunesse (2) ; Sociologie des arts (2) ; sociologie des usages (2) ; sociologie des œuvres (2) ; sociologie des activités (2) ; sociologie des relations (2) ; sociologie des mouvements sociaux (2) ; Sociologie de la connaissance (2) ; sociologie de Bourdieu (1) ; sociologie de Durkheim (1) ; sociologie de la critique (1) ; sociologie des arts (1) ; sociologie des formes (1) ; sociologie des cadres (1) ; sociologie des publics (1) ; sociologie des savoirs (1) ; sociologie des rapports (1) ; Sociologie des croyances (1) ; sociologie des homicides (1) ; sociologie des religions (1) ; sociologie des inégalités (1) ; sociologie des associations (1) ; sociologie des dispositions (1) ; sociologie des intellectuels (1) ; sociologie des représentations (1) ; psycho-sociologie des organisations (1) ; sociologie des politiques publiques (1) ; sociologie du risque (1) ; sociologie du cinéma (1) ; Sociologie du travail (1) ; sociologie du chômage (1) ; sociologie du racisme (1) ; sociologie du bénévolat (1) ; sociologie du voisinage (1) ; sociologie du vieillissement (1) ; sociologie de la prison (1) ; sociologie de la musique (1) ; Sociologie de la Jeunesse (1) ; sociologie de la peinture (1) ; sociologie de la réception (1) ; sociologie de la production (1) ; Sociologie de la concurrence (1) ; sociologie de la délinquance (1) ; sociologie de la communication (1) ; sociologie de la juridicisation (1)

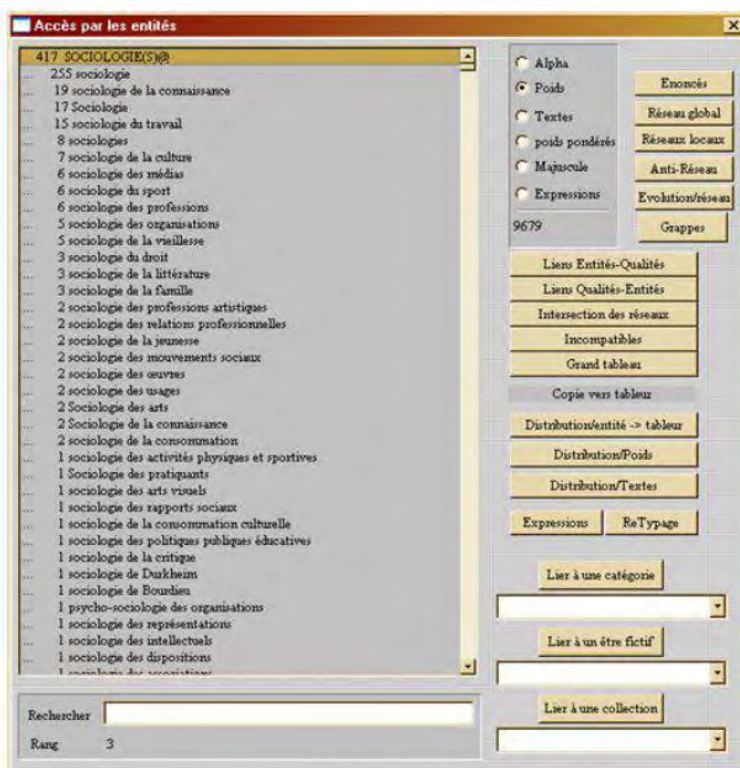
- 9 Nous proposons d'analyser la pertinence de ces lectures en mobilisant conjointement plusieurs jeux de dictionnaires et de concepts. L'un d'entre se caractérisera par l'absence totale d'agrégation, « sociologie », « Sociologie », « sociologies », « Sociologies », « connaissance », « travail », « famille », « bénévolat »... étant considérés comme des atomes (appelons « lecture n°1 » le jeu de concepts traduisant ce point de vue). Une deuxième lecture (avec son jeu de concepts n°2) considérera au contraire l'ensemble de ces expressions comme des entités à part entière et distinctes : la sociologie du travail et la sociologie du sport seront deux choses différentes. On peut imaginer une troisième lecture basée sur la construction d'un gros ensemble contenant l'ensemble de ces formes, ainsi que toutes les façons de désigner « la sociologie » (au singulier, au pluriel, en déclinant toutes les formes graphiques) : ce type de regroupement, qui sous Prospéro prend le nom d'être fictif⁶, permet de traiter comme un tout la sociologie au sens

générique et les différents courants thématiques ou théoriques qui la composent, sans perdre la trace de chaque élément.

Figure 1 : On voit le score de SOCIOLOGIE(S)@ dans l'ensemble des premiers éléments de la liste des entités



Figure 2 : déploiement de l'être fictif SOCIOLOGIE(S)@⁷



10 Commençons par mesurer les effets de ces différents codages sur l'ordre d'apparition des entités du corpus (cf. Tableau 3). Si, dans la première colonne, nous retrouvons sans surprise le début du tableau 1, nous apercevons que la deuxième lecture renvoie la sociologie beaucoup plus bas dans la liste au point qu'elle n'apparaît plus dans les 10 premières entités, le travail arrivant quant à lui toujours en tête malgré la diminution de son score liée à la création de l'expression. La troisième lecture rétablit la sociologie dans le haut de la liste, sans toutefois parvenir à lui faire occuper la première ou la deuxième place.

Tableau 3 : Premières entités en fonction du jeu de codage utilisé

Interprétation 1	Score	Interprétation 2	Score	Interprétation 3	Score
travail	621	travail	605	travail	605
analyse	501	analyse	501	analyse	501
sociologie	382	processus	355	SOCIOLOGIE(S)@	417
processus	355	femmes	341	processus	355
femmes	341	communication	335	femmes	341
communication	336	recherche	320	communication	335
recherche	320	étude(s)	306	recherche	320
étude(s)	306	acteurs	306	étude(s)	306
acteurs	306	question	271	acteurs	306
question	271	formes	270	question	271
formes	271	LA-FRANCE@	261	formes	270
LA-FRANCE@	261	relations	257	LA-FRANCE@	261
relations	259	sociologie	255	relations	257
jeunes	252	jeunes	252	jeunes	252
rapports	239	rapports	238	rapports	238

- 11 L'intérêt de cette mise en variation des lectures réside moins dans les positions des éléments des listes que dans ce que ces différentes approches rendent lisible. Il est possible notamment de travailler sur la façon dont les chercheurs désignent la sociologie. Une piste assez frayée revient à recenser les qualités qui lui sont attribuées. On peut en effet supposer que les entités les plus qualifiées sont considérées comme importantes. Quels que soient les jeux de dictionnaires mobilisés, la sociologie arrive parmi les entités assez qualifiées, derrière des noms comme « forme », « travail », « analyse », « action », avec des scores assez proches. Les qualificatifs associés à « sociologie » sont assez semblables mais interrogent, une fois encore, le chercheur sur les expressions à créer.



- 12 La lecture de cet écran nous plonge dans les mêmes problèmes de codage que ceux recensés plus haut. Ne faut-il pas créer des expressions comme « sociologie économique », « sociologie pragmatique » ou « sociologie compréhensive » ? Certes, si l'on pense qu'il est nécessaire de créer « sociologie du travail », il convient assurément de construire « sociologie économique ». Mais faut-il coder les expressions apparaissant plus bas dans cette fenêtre, comme « nouvelles sociologies », « sociologie politisée », « sociologie reconnue » ? Il semble que dans ces deux derniers cas, la sociologie est plus qualifiée que décrite par une spécialisation. Si l'on veut analyser la façon dont la sociologie est qualifiée par les sociologues – et donc discutée, critiquée, mise en tension –, il devient pertinent de créer les expressions dans les seuls cas où l'on précise l'orientation de la discipline (compréhensive, explicative, pragmatique, interactionniste...). On retrouve alors des oppositions sur les types de sociologie en voyant apparaître la façon dont les chercheurs tentent de critiquer un type de sociologie et de préciser la spécificité de leur travail. Apparaissent en effet des qualités comme une sociologie « classique », « banale », « politisée », « enchantée » mais aussi « nouvelle », « fine », « riche », « reconnue ». C'est un travail similaire qu'il nous faut réaliser sur les qualifications. Les entités qualifiées de "sociales" ("social", "sociaux"...) posent des problèmes similaires : si la création d'expressions comme "psychologie sociale" ne pose pas de problème, faut-il créer "le mépris social" ?
- 13 Ce travail de codage met en évidence les possibilités de changements de focale : selon le niveau d'analyse que l'on privilégie, se donnent à lire les objets de la sociologie, les

familles de discipline ou encore le travail du sociologue cherchant à se démarquer d'autres approches.

- 14 D'une façon plus générale, il semble intéressant de mobiliser simultanément deux démarches contradictoires. L'une vise à faire diminuer le poids des mots dont on fait l'hypothèse qu'il recouvre des sens et des usages différents. C'est notamment le cas de "communication" : ce terme désigne à la fois un objet de la sociologie et la contribution du sociologue au congrès. Il devient alors nécessaire, si l'on ne veut garder que l'objet sociologique, d'isoler par la création d'expressions appropriées des formules comme "dans cette communication", "l'objet de ma communication" ou "proposer une communication". Ce travail entraîne la baisse importante du nombre d'occurrences de l'entité « communication » puisqu'elle passe ainsi de la troisième position avec 323 occurrences à la 104^e place avec un score de 63 !
- 15 L'autre démarche vise à faire exister des notions présentes sous des formes trop différentes pour qu'ils apparaissent dans le haut des listes. A titre d'exemple, si l'on s'étonne que le mot "culture" ne réalise qu'un score de 133, on peut rechercher s'il n'existe pas d'autres moyens de désigner cette notion. Une recherche, à l'aide d'une formule visant à repérer les entités suivies de "culturel", "culturels" ou "culturelles", permet d'identifier 113 autres éléments comme "formes culturelles", "identités culturelles", ou "appartenance culturelle".
- 16 Ce travail effectué, il reste à s'interroger sur ce qui résiste aux différentes opérations de codage. Ainsi, quelques entités restent marquantes. Le fait que l'on parle beaucoup de "travail", de "question" ou de "recherche" n'est pas surprenant. On peut en revanche être surpris du score important d'entités comme "femmes" (6^e rang) ou "militaire" (69^e rang). L'on ne s'attendait pas à une place aussi importante puisqu'elles rivalisent en terme de rang avec des objets largement travaillés par la tradition sociologique comme l'école (24^e rang), les élites (83^e rang) ou les syndicats (90^e rang). Que dire de ces éléments au score surprenant ? Indiquent-ils un renouveau des objets de la sociologie ? Nous interrogeons-ils sur la représentativité du corpus, c'est-à-dire des sociologues présents au colloque de l'AFS, par rapport à l'ensemble de la discipline ? Sans tirer de conclusions hâtives, on va vouloir regarder à quoi ces éléments sont liés. Leur poids indique cependant qu'ils risquent d'être connectés à de nombreuses autres entités. Prospéro permet de repérer les réseaux spécifiques, c'est-à-dire les liens restant importants après la suppression virtuelle des éléments statistiquement dominants. Dans ces conditions, on découvre que les femmes sont connectées à "marché du travail", "genre", "compétences", "violences", "immigration", ce qui rend lisibles quelques façons d'évoquer les "femmes" dans les résumés. Le même travail nous amène à identifier que le réseau spécifique des militaires se compose de "armées", "forces", "leadership", "ethos", "concertation", "profession"...
- 17 En travaillant sur les petits objets qui sont liés aux éléments importants, on peut décrire ces derniers de façon plus fine. Pour pousser plus loin encore cette logique qui nous conduit à abandonner les objets à fort score, nous pouvons envisager de repérer les entités absentes des réseaux respectifs de "femmes" et de "militaires". Dans le premier cas, on note que "femmes" n'apparaît jamais dans le même énoncé que "transformation", "perspective", "individualisation", "politiques sociales", "économie", "subjectivité", "décision"... et jamais dans le même texte que "expériences", "observation", "évaluation" "expertise", "management", "science". De même, les militaires n'apparaissent jamais dans le même énoncé que "comparaison", "socialisation", "formation", "espace public", et jamais dans le même texte que "acteurs", "individus", "stratégies", "dispositifs"...

- 18 Ce travail pointe à la fois sur des propriétés des objets et sur des façons de rédiger les résumés.

Comment rédiger une proposition de communication ?

- 19 Au-delà du travail permettant de repérer les entités du corpus et d'identifier la façon dont elles sont connectées, penchons-nous maintenant sur les conditions de production de ces textes et cherchons à saisir la façon dont les chercheurs abordent l'épreuve du choix concernant le sort de leur proposition de communication.
- 20 Le format d'une proposition de communication est peu établi. En dehors d'une contrainte de longueur, les règles supposées de production d'un texte varient. Une lecture du corpus à l'œil nu suffit pour s'apercevoir que certains s'efforcent de synthétiser leur argument quand d'autres rédigent un texte qui reste très programmatique, annonçant un travail parfois non réalisé lors de la soumission du texte⁸. On peut repérer avec le logiciel des formules comme "recherche en cours", "travail de terrain en cours", "enquête en cours" ou encore porter l'attention sur les verbes conjugués au futur. Dans des corpus moins spécifiques, l'accord verbal est un faux ami puisque l'on peut trouver des énoncés comme "et il *faudra* attendre des années avant d'avoir vu apparaître...". Néanmoins ici, dans le cadre de résumés, on ne trouve guère de description historique de terrain qui utiliserait le futur, et les marques du futur et les verbes renvoient bien à une projection de ce que sera la communication. De même, une lecture attentive du corpus sur papier conduit à repérer que certains ont proposé leur contribution dans plusieurs réseaux thématiques. Le logiciel peut nous aider à repérer systématiquement ces textes identiques, ou pratiquement identiques. On peut interpréter ces pratiques de différentes façons. Peut-être s'agit-il de multiplier le nombre de communications pour étoffer des CV ? Le grand nombre de doctorants inscrits dans ce congrès nourrit cette hypothèse. Cependant, parmi les 80 auteurs (sur 974) ayant proposé plus d'une communication, certains sont dans des positions institutionnelles dont on peut supposer qu'elles ne nécessitent pas d'utiliser cette stratégie pour consolider leur dossier. Dans ce cas, il s'agit le plus souvent de textes différents, ce qui invite à penser que l'argument de leurs auteurs ne pouvait tenir dans une courte communication ou que l'auteur fut sollicité par différents réseaux thématiques, ou encore qu'il s'agit d'une stratégie pour « occuper le terrain » (il est important de se faire entendre et d'assurer – autant que faire se peut – sa présence dans des lieux donnés). La présentation du même résumé dans plusieurs sessions peut aussi s'interpréter comme une tentative de multiplication des chances de voir sa communication acceptée.

Des contraintes de production d'un résumé sociologique

- 21 Les incertitudes liées au sort des propositions de contribution au moment où elles sont rédigées nous invitent à travailler sur les contraintes de contenu pesant sur les auteurs. S'agissant d'un congrès de sociologie, il semble que celles-ci se soient posées pour certains, notamment pour ceux dont la qualité de sociologues est potentiellement discutée : qu'il s'agisse bien sûr des doctorants commençant leurs recherches, des collègues évoluant en dehors des laboratoires ou départements universitaires labellisés en sociologie (en sciences de la communication, en sciences de l'éducation, en STAPS, ou dans des institutions de recherche ne précisant pas explicitement la discipline), ou encore

de ceux travaillant dans des disciplines voisines (l'anthropologie, l'ethnologie, la démographie, la linguistique, l'histoire, la psychologie sociale), des chercheurs peuvent penser que leur légitimité à communiquer au congrès de l'Association Française de Sociologie n'est pas acquise et qu'il convient de s'ajuster aux attendus implicites d'une proposition de contribution dans la discipline⁹.

- 22 Partons d'un modèle selon lequel la qualification de sociologue ne pose plus problème si l'on parvient à répondre aux exigences suivantes :
- inscrire son propos dans la discipline
 - faire référence à des concepts ou à des auteurs de la sociologie
 - mentionner un travail théorique ou s'appuyer sur des modèles
 - montrer que l'on s'appuie sur des données empiriques ou que l'on se réfère d'une façon ou d'une autre au "réel"
 - mobiliser des méthodes sociologiques
 - proposer quelque chose de nouveau.
- 23 Ces conditions sont évidemment discutables ; les chercheurs proches de la philosophie sociale ou de l'histoire de la discipline pourraient contester, par exemple, la nécessité d'un terrain pour produire une réflexion sociologique. Le propos n'est pas ici de démontrer la validité absolue de ce modèle mais plutôt de montrer comment le logiciel peut être mobilisé pour évaluer la façon dont les résumés répondent (ou non) à ces différents critères.

L'inscription dans la sociologie

- 24 Les formules ayant permis de coder des expressions comme "sociologie du travail" nous ont conduits à créer un être fictif rassemblant toutes les références directes à la discipline (Tableau 2) avec un score de 417 occurrences réparties dans 214 textes. Mais l'inscription dans la sociologie peut aussi passer par des expressions comme "le travail sociologique" ou "l'approche sociologique", ou encore par des mots comme "socio-historique" ou "socio-culturel". Ces remarques invitent à regrouper l'ensemble des résumés mentionnant d'une façon ou d'une autre la discipline, soit 267 résumés utilisant à 548 reprises l'une ou l'autre de ces expressions.
- 25 L'on peut discriminer ces usages par auteur mais aussi par réseau thématique. En tenant compte du volume produit, on peut ainsi noter que les trois réseaux utilisant le plus les représentants de cet être fictif sont, par ordre décroissant, "sociologie de la connaissance", "sociologie des arts et de la culture", "théories et paradigmes sociologiques". Symétriquement, ceux qui y ont le moins recours sont (toujours par ordre décroissant) "gestion politique du corps et des populations", "identité, subjectivité, revendication" et "sociologie de la gestion".

La référence à des concepts

- 26 Faire allusion à la discipline est une chose, mais c'est la mobilisation de ses notions-clés qui assure que l'activité est sociologique. Nous avons distingué deux dispositifs de rassemblement. D'une part, on peut utiliser une catégorie nommée "vocabulaire sociologique" dont les représentants indiquent la présence d'un discours sociologique dans des corpus habituels, lesquels contiennent des contributions d'auteurs variés. Par exemple, parler d'"acteurs", de "représentation" ou de "mouvements sociaux" marque la référence aux sciences sociales mais pas de façon suffisamment discriminante pour

assurer que l'auteur est effectivement sociologue, puisque essayistes, journalistes et autres commentateurs de "la société" font largement usage de ces termes. Il semble donc nécessaire de créer une collection de notions forgées par la tradition sociologique et moins utilisées par les non-sociologues. Plus précisément, il s'agit de concepts qui se sont imposés au fil de l'histoire de la discipline en se stabilisant, s'explicitant, se diffusant. On peut supposer que les auteurs de ce corpus mobilisant des concepts comme "idéal-type", "habitus", "mondes sociaux" et autres représentants de la collection que nous avons créée (cf. Tableau 4), s'inscrivent plus spécifiquement dans la sociologie. Ils sont 750 dans ce cas à utiliser 165 notions dans 810 résumés. Ces textes sont présentés plutôt dans des RTF comme "théories et paradigmes sociologiques", "normes, déviances et réactions sociales" et "classes, inégalités fragmentations" (pour citer dans l'ordre décroissant les trois réseaux qui utilisent le plus cette collection), alors que "sociologie du travail et activité", "sociologie de la consommation", "méthodes" sont ceux qui la mobilisent le moins¹⁰.

Tableau 4 : Représentants de la collection NOTIONS-SOCIOLOGIQUES*

acteurs (306) ; champ (136) ; normes (130) ; culture (126) ; la société (120) ; représentations (120) ; socialisation (110) ; réseau (71) ; interactions (68) ; légitimité (55) ; rapports sociaux (53) ; rationalisation (48) ; citoyenneté (42) ; conflits (42) ; interaction (42) ; domination (37) ; légitimation (33) ; acteur (31) ; cultures (25) ; champs (24) ; groupes sociaux (23) ; paradigme (23) ; reproduction (20) ; capital social (19) ; lien social (18) ; division du travail (17) ; habitus (17) ; classes sociales (16) ; déviance (16) ; rationalité (16) ; action collective (15) ; mouvements sociaux (15) ; le social (13) ; action sociale (12) ; représentations sociales (12) ; ethos (12) ; groupe social (11) ; mobilité sociale (10) ; identité professionnelle (10) ; temps sociaux (10) ; espace social (10) ; changement social (10) ; contrôle social (10) ; régulation sociale (9) ; intervention sociale (9) ; réalité sociale (8) ; activité sociale (8) ; catégorie sociale (8) ; insertion sociale (8) ; déviances (8) ; les élites (7) ; position sociale (7) ; classe sociale (7) ; individualisme (7) ; Etat Social (6) ; statut social (6) ; normes sociales (6) ; réseaux sociaux (6) ; environnement social (6) ; reconnaissance sociale (6) ; liens sociaux (5) ; cadres sociaux (5) ; mandes sociaux (5) ; processus sociaux (5) ; inscription sociale (5) ; expériences sociales (5) ; épistémologie sociale (5) ; effets pervers (5) ; fait social (5) ; individualisme méthodologique (5) ; mouvement social (5) ; positivisme (5) ; l'élite (4) ; univers sociaux (4) ; utilité sociale (4) ; histoire sociale (4) ; structures sociales (4) ; hiérarchies sociales (4) ; contraintes sociales (4) ; reproduction sociale (4) ; trajectoires sociales (4) ; objet sociologique (4) ; champ social (4) ; changements sociaux (4) ; déterminisme (4) ; rite (4) ; rites (4) ; structuralisme (4) ; norme sociale (3) ; règle sociale (3) ; rôles sociaux (3) ; conflit social (3) ; intégration sociale (3) ; la bureaucratie (3) ; PCS (3) ; dominations (3) ; Etat providence (3) ; historicité (3) ; idéaux-types (3) ; identité collective (3) ; marxisme (3) ; differentialisme (3) ;

Tableau 4 (suite)

protestantisme (3) ; interactionnisme (2) ; rôle social (2) ; objet social (2) ; sujet social (2) ; temps social (2) ; cadre social (2) ; faits sociaux (2) ; lutte sociale (2) ; lois sociales (2) ; espaces sociaux (2) ; origines sociales (2) ; pré socialisation (2) ; réalités sociales (2) ; mécanismes sociaux (2) ; trajectoire sociale (2) ; actions collectives (2) ; anomie (2) ; catégories socioprofessionnelles (2) ; classe ouvrière (2) ; rationalités (2) ; science sociale (2) ; sociétés industrielles (2) ; controverse socio-technique (2) ; inégalités socio-spatiales (2) ; organisation bureaucratique (1) ; effets sociaux (1) ; justice sociale (1) ; « mondes sociaux » (1) ; traditions sociales (1) ; stéréotypes sociaux (1) ; dominations sociales (1) ; composantes sociales (1) ; déterminismes sociaux (1) ; catégorisation sociale (1) ; constructionnisme social (1) ; « singularité sociale » (1) ; dimension socialisatrice (1) ; manière d'être socialement (1) ; dimensions socialisatrices (1) ; perspective social-constructiviste (1) ; « ethnographie sociologique » (1) ; CSP (1) ; catégorie socioprofessionnelle (1) ; champs sociaux (1) ; déterminismes (1) ; exclusions (1) ; fonctionnalisme (1) ; idéal-type (1) ; prénotions (1) ; sociation (1) ; sociétés post-industrielles (1) ; sociétés traditionnelles (1) ; stratification sociale (1) ; investigation socio-anthropologique (1) ; caractéristiques socio-économiques (1) ; contextualisation socio-historique (1) ; expérimentations socio-culturelles (1) ; temporalité socio-professionnelle (1) ; déterminations socio-économiques (1) ; groupes socio-professionnels (1) ; méthode socio-ethnographique (1) ; réseaux socio-techniques (1) ; cadres socio-cognitifs (1) ; champs socio-cognitifs (1) ; Habitus (1) ; constructivisme (1) ; asynchronisme (1) ; behaviorisme (1) ; falsificationnisme (1) ; subjectivisme (1) ; a-théisme (1) ; néo-malthusianisme (1) ; kantisme (1) ; interactionnisme (1) ; Multiculturalisme (1) ; utilitarisme (1) ; constructionnisme (1) ; relativisme (1)

La convocation des auteurs de la discipline

- 27 Si la référence à des auteurs est très conseillée dans une communication, sa présence dans une proposition de contribution est plus discutable, compte tenu de la longueur réduite imposée. Pour évaluer la façon dont les personnes de l'AFS ont géré cette contrainte, il semble pertinent de distinguer deux collections : l'une constituée des sociologues classiques ou assez connus pour être cités (on pense à Durkheim, Weber, mais aussi des philosophes, historiens, anthropologues ou autres acteurs des sciences sociales non présents lors de ce congrès) ; l'autre constituée des auteurs-acteurs du corpus, c'est-à-dire des auteurs des résumés qui peuvent aussi apparaître dans le corpus. A un premier niveau, on peut repérer les auteurs le plus souvent cités : par exemple le score important de Bourdieu (34) qui dépasse ceux de Weber (26) et de Durkheim (9). Toutefois, au-delà de la curiosité de savoir qui est cité et qui ne l'est pas, il est intéressant de discriminer ces usages en fonction du réseau thématique. Il apparaît que la collection des auteurs des sciences sociales est largement utilisée par le RTF "Théories et paradigmes sociologiques" (plus de 10 fois plus que la valeur théorique représentant l'indépendance), suivi des groupes sur "la sociologie de la connaissance" et les "approches plurielles du sujet" (score plus de 5 fois supérieur à la valeur théorique). L'on retrouve donc parmi les contributions retenues dans les réseaux thématiques les plus théoriques celles qui mentionnent le plus les auteurs des sciences sociales. Symétriquement, trois réseaux ne mentionnent aucun des auteurs de la collection : "famille, vie privée, vie publique", "vieillesse", "sciences de l'expertise des sciences humaines"¹¹.
- 28 Mais lorsqu'on considère la collection des auteurs-acteurs¹², les choses s'inversent. Les réseaux "théories et paradigmes sociologiques" et "sociologie de la connaissance", tout comme 12 autres RTF, ne citent aucun sociologue présent au congrès, alors que l'on ne recense qu'une seule référence dans le réseau "approches plurielles du sujet" qui s'était

également distingué dans l'utilisation fréquente de l'autre collection. Celle des auteurs-acteurs est en revanche très utilisée par le réseau "sociologie du sport" (plus de 10 fois plus que la valeur théorique). Il convient de noter que les sociologues de ce réseau font souvent référence à des auteurs de ce même RTF. L'on peut expliquer cela par le fait que les sociologues participant à ce RTF évoluent souvent dans les STAPS (Sciences et Techniques de Activités Physiques et Sportives)¹³, qui rassemble des chercheurs en sciences expérimentales utilisant très largement les références bibliographiques ; on peut donc supposer que les sociologues du sport ont intégré ces usages au contact de leurs collègues et, alors même qu'ils se sont réunis dans le cadre de l'AFS, ont gardé cette habitude. Viennent ensuite les réseaux "pluralisme religieux et citoyenneté", "sociologie de l'engagement", "sociologie de la consommation" et "sociologie de l'éducation".

Le travail théorique

- 29 A défaut de mentionner des auteurs, on peut montrer que le propos s'inscrit dans un cadre théorique ou, en tout cas, fait allusion à des débats de ce type. La recherche d'éléments comme "théorie" ou "modèle" suppose d'être attentif à ne pas inclure les modèles des acteurs et de recenser des expressions contenant des entités qualifiées de "théorique(s)". Ce faisant, l'on dénombre 220 occurrences de cet être fictif se distribuant dans 152 résumés.
- 30 Sans surprise, ce sont les groupes "sociologie de la connaissance", "approches plurielles du Sujet" et "théories et paradigmes sociologiques" qui réalisent les meilleurs scores (avec des coefficients multiplicatifs par rapport à l'indépendance valant respectivement 6.6, 6.3 et 5.7). L'on dénombre cinq réseaux n'utilisant jamais ces représentants : "Vieillesse", "parcours de vie et dynamiques sociales", "pluralisme religieux et citoyenneté", "identité - subjectivité - revendication", "sociologie de l'expertise".

Faire référence à des données empiriques et/ou au monde réel

- 31 Cet impératif est sans doute l'un des plus discutés puisqu'il est lié à un débat assez classique sur la place de la philosophie sociale. On attend souvent d'un doctorant qu'il "fasse du terrain", réalise des enquêtes et apporte des données empiriques originales. Du coup, il semble pertinent d'évaluer la présence d'éléments comme "données", "matériau empirique" ou "travail de terrain" dans notre corpus en construisant une catégorie rassemblant ce type de représentant (cf. Tableau 5). Elle obtient un score de 651 occurrences réparties dans 386 textes. C'est le réseau "méthodes" qui l'utilise le plus. Les RTF "théories et paradigmes sociologiques", "mouvements sociaux", "sociologie de la connaissance" et "politiques sociales" y font rarement référence alors que le groupe "approches plurielles du sujet" apparaît comme le seul à ne jamais mentionner l'un de ces éléments.

Tableau 5 : Liste des représentants de la catégorie "Données empiriques"

enquête (147) ; entretiens (113) ; enquêtes (46) ; terrain (46) ; corpus (39) ; des données (33) ; terrains (18) ; questionnaires (18) ; archives (18) ; enquête de terrain (15) ; statistiques (15) ; sur le terrain (14) ; matériau (7) ; recherche de terrain (6) ; travail de terrain (4) ; une analyse empirique (4) ; recherches de terrain (4) ; le réel (3) ; études empiriques (3) ; matériau empirique (3) ; base empirique (2) ; travaux empiriques (2) ; pertinence empirique (2) ; investigation empirique (2) ; observations empiriques (2) ; observations de terrain (2) ; base de données (2) ; situations empiriques (1) ; objet empirique (1) ; cadre empirique (1) ; examen empirique (1) ; impasse empirique (1) ; problème empirique (1) ; raisons empiriques (1) ; enquêtes empirique (1) ; validité empirique (1) ; limites empiriques (1) ; exemples empiriques (1) ; émergence empirique (1) ; éléments empiriques (1) ; effets empiriquement (1) ; modalités empiriques (1) ; évaluation empirique (1) ; matériaux empiriques (1) ; références empiriques (1) ; conditions empiriques (1) ; recherches empiriques (1) ; difficultés empiriques (1) ; hétérogénéité empirique (1) ; ordre juridique empirique (1) ; minimum de données empiriques (1) ; collecte de données empiriques (1) ; problème théorique et empirique (1) ; territoire biopolitique empirique (1) ; question sera abordée empiriquement (1) ; institution disciplinaire empirique (1) ; observation de situations empiriques (1) ; réflexion théorique et mesure empirique (1) ; méthodes de recueil de données empiriques (1)

- 32 Toutefois, il est possible de mentionner le réel, sous des formes très variées. Si de nombreux énoncés évoquent le "travail de terrain" du chercheur, on n'est pas surpris de lire que les sociologues évoluant dans le réseau "théories et paradigmes sociologiques", lorsqu'ils font allusion à cette dimension, en parlent de façon très différente, comme en témoignent ces deux exemples :

Une connaissance herméneutique critique au regard des connaissances qui se donnent pour vraies, une sociologie de la connaissance procédant à des « imputations » croisées par lesquelles les savoirs examinés sont sans cesse référés à leurs conditions empiriques de possibilité, peut-elle échapper à une imputation récurrente de même nature ? [F. Farrugia, Réseau "sociologie de la connaissance"]

Max Weber scande dans toute son œuvre et en particulier dans les Essais sur la théorie de la science que le réel est infini en intention et en extension. [G. Bronnier, RTF "Théories et paradigmes sociologiques"]

- 33 Si les données empiriques et le réel sont ici convoqués, c'est dans des débats épistémologiques et non dans une perspective où on confronterait effectivement des considérations théoriques à "des faits". Ce point conduit à effectuer un choix. Soit on considère que "le réel" dont il est question ici ne répond pas aux exigences du modèle ; dans ce cas, il suffit de retirer de la catégorie des représentants comme "le réel" ou "conditions empiriques", car on considère que ces expressions marquent une distanciation – et donc ne concerne pas directement le terrain. Soit on défend une conception selon laquelle il est difficile de retirer ce type d'expressions sans prendre le risque de perdre des textes qui les utilisent pour introduire la construction du travail empirique. Face à ce type d'interrogations – semblables à celles que nous évoquons plus haut à propos de la construction de SOCIOLOGIE(S)@-, il est souvent conseillé de garder deux versions de cette catégorie pour précisément évaluer les conséquences de nos interprétations. Pour des raisons de lisibilité, dans la suite de l'article, nous développerons le travail en gardant la version incluant toutes ces expressions et nous la discuterons.

La mobilisation des méthodes

- 34 Généralement, la production et l'analyse de ces "faits" supposent d'utiliser des méthodes qui doivent être mentionnées. Lorsqu'elles sont considérées comme "classiques", il suffit

généralement de les nommer. "Analyse thématique", "entretiens semi-directifs", "questionnaires auto-administrés", "observation participante" sont des exemples de méthodes que nous avons rassemblés dans une collection. Lorsque celles-ci ne sont pas très diffusées, elles sont généralement explicitées et on peut supposer qu'apparaissent des marques comme "méthode", "méthodologie", que nous avons ajoutées à cette collection qui, ainsi constituée, nous permet d'isoler 297 résumés empruntant l'un ou l'autre de ces termes.

- 35 C'est logiquement le réseau "Méthodes" qui utilise le plus les représentants de cette collection (relativement, le score relatif est près de 9.6 fois plus important que la moyenne). Trois groupes n'y ont jamais recours : "pluralisme religieux et citoyenneté", "mouvements sociaux", "approches plurielles du Sujet". Dans ce dernier cas, l'importance accordée aux considérations théoriques dans ce réseau ne laisse aucune place à des considérations de méthodes. Les autres RTF privilégiant la valence théorique ("théories et paradigmes sociologiques" et "sociologie de la connaissance") parviennent à évoquer respectivement à quatre et deux reprises des questions de méthodes. Mais les représentants de la collection utilisés par ces auteurs marquent des considérations davantage épistémologiques que techniques, et renvoient toujours aux statuts des méthodes chez des auteurs classiques ou dans approches théoriques :

On ne peut confondre les synthèses métathéoriques parsonienne ou habermassienne avec des paradigmes unificateurs proposant des méthodes de résolution exemplaires s'étendant à des objets toujours plus étendus. [A. Bouvier, Réseau "théories et paradigmes sociologiques"]

- 36 Cette collection de méthodes apparaît du coup plus discriminante que la catégorie "données empiriques" pour évaluer la distance des réseaux préoccupés par la théorie, avec ceux accordant une importance décisive aux terrains et aux méthodes d'investigation. Par ailleurs, cette collection permet de repérer les méthodes les plus souvent mentionnées dans le corpus. Ce sont les entretiens (parfois qualifiés de "semi-directifs", d'"approfondis", ou de "compréhensifs) qui obtiennent le meilleur score (236 occurrences). Les méthodes plus quantitatives, repérées avec des éléments comme "questionnaire(s)", "question ouverte", "statistique", ne réalisent qu'un score de 77, tandis qu'il est fait allusion à 22 reprises à des observations (participante(s), ethnographique(s)). Notons que les questions d'analyse de données qualitatives, identifiables par des mots comme "corpus", "archives", "analyse de contenu"... réalisent un score de 64. Nous avons isolé de ces derniers regroupements des méthodes informatisées. On parle de logiciel(s) à 24 reprises (mais pas forcément pour désigner des logiciels pour la sociologie), et l'on cite 21 fois des noms de logiciels utilisés par les sciences sociales : Prospéro (12), Alceste (8), PAM (3), SPSS (1), RESO (1).

Proposer quelque chose de nouveau

- 37 Un congrès scientifique est théoriquement un lieu de présentation de nouvelles connaissances. Résulte de ce postulat, une contrainte d'originalité pesant sur les sociologues souhaitant communiquer, contrainte qui peut les conduire à souligner l'apport de leurs travaux. Celui-ci peut se décliner sous plusieurs formes : description d'un terrain peu exploré, mise en évidence d'une inégalité ou d'une évolution ayant échappé à l'investigation sociologique, renouvellement de discussions théoriques, présentation ou mobilisation de méthodes originales.

- 38 Une catégorie de qualités nous permet de recenser ce qui est nouveau. Il s'agit le plus souvent de nouveaux acteurs, de nouvelles formes d'action, de nouvelles normes caractérisant un changement dans la société, que le sociologue doit décrire. Prospéro permet d'identifier des formes assez pures de la mise en exergue de la nouveauté. Parmi celles-ci, on peut repérer les séquences du type < /CAT_QUALITE=Nouveauté /ENTITE / T=3 /ENTITE >¹⁴ qui repère des expressions comme "nouvelles formes de participation", "nouvelle génération de bénévoles", ou "une nouvelle culture de sexe".
- 39 A un premier niveau, ce type de résultat conduit à recenser les nouveautés et les changements préoccupant les sociologues. Il s'agit principalement de populations qui auraient changé, car si l'on parle d'une "nouvelle génération de jeunes", d'une "nouvelle classe ouvrière" ou de "nouveaux retraités", c'est bien que ces catégories sociales ont subi des transformations. Il s'agit aussi de nouvelles pratiques ou de nouveaux comportements, ou encore de nouveaux collectifs. On peut repérer des énoncés dans lesquels ces nouveautés foisonnent :
- L'objet de cette communication est d'attirer l'attention sur la transformation du champ vitivinicole en France, eu égard aux transformations de la structure du marché mondial et à l'autonomisation de la profession d'œnologue, en insistant sur les changements survenus dans les formes organisationnelles, les modes de légitimation et le rapport à l'innovation. [M.F. Garcia-Parpet]
- 40 A un second niveau, ce recensement conduit à penser que les auteurs de ces énoncés considèrent l'activité sociologique comme une description des changements sociaux, des transformations ou des mutations sociales. Il est beaucoup plus rare de trouver des expressions marquant l'émergence de nouveaux modèles sociologiques, de nouvelles méthodes d'investigation ou de renouvellement de discussion.
- 41 Travailler sur la nouveauté suppose symétriquement de s'intéresser à ce qui est ancien, car cette qualité peut être une marque de la nouveauté. Une catégorie de qualités marquant l'ancienneté permet de repérer des "modèles familiaux traditionnels", des "goûts classiques", des "anciennes configurations socio-politiques". Mais le plus souvent ces qualificatifs s'appliquent à des modèles ou approches sociologiques. Il s'agit parfois de se référer à l'inscription dans une tradition pour marquer la référence à la discipline (comme nous l'avons vu plus haut, il s'agit de mobiliser une méthode classique d'analyse de données ou de mentionner une thèse classique). Cependant, le plus souvent on cherche à dépasser cette tradition en proposant quelque chose de nouveau. On essaye par exemple de "reposer la question classique de 'la transformation des passions en intérêts' ", de prolonger "des études classiques sur la socialisation professionnelle", ou de souligner les faiblesses d'un ensemble de "problématiques classiques de la sociologie" pour proposer son approche.
- 42 On peut aussi mobiliser un autre outil de Prospéro permettant de repérer les énoncés qui mobilisent des couples de représentants appartenant aux catégories "Nouveauté" et "Ancienneté". Cela rend possible l'identification de formes dialectiques opposant le nouveau et l'ancien.

En quoi les dynamiques de participation font-elles émerger de nouveaux types de représentants (lui ne se confondent pas avec les représentants élus traditionnels) et de nouvelles conceptions de la représentation et de la représentativité ? Il s'agira aussi de mesurer jusqu'à quel point ces formes et ces conceptions s'inscrivent en continuité ou en décalage avec l'évolution de la représentation classique, et dans quelle mesure elles contribuent à une modification de celle-ci. [Y. Sintomer]

- 43 Cela permet également de rechercher des figures dans lesquelles de nouvelles pratiques sociales engendreraient une obligation de renouveler les approches sociologiques. Les seules formes identifiables concernent de nouvelles pratiques qui créent une contrainte pour le travail sociologique ou qui l'obligent à compléter ses répertoires de description :

Le pouvoir économique et symbolique des think tanks finit par imposer « un nouveau langage » et de nouvelles pratiques aux sciences sociales. [D. Lavergne]

Nous analyserons comment les acteurs européens ont élaboré cette stratégie de réforme des systèmes de retraite, en montrant comment ils ont contourné les obstacles institutionnels en élaborant une politique cognitive, qui même si elle n'est pas contraignante, délimite de nouveaux principes d'action et une nouvelle conception des régimes de retraite. [C. Mendin]

Des façons de gérer ces contraintes

- 44 Nous avons donc cherché des textes exemplaires – au double sens du mot – en repérant les résumés qui tiendraient l'ensemble de ces contraintes, c'est-à-dire qui contiennent une inscription dans la discipline (soit en parlant explicitement de sociologie, soit en utilisant des concepts de la discipline), font référence à des auteurs de sciences sociales (qu'ils figurent parmi les sociologues de l'AFS ou non), mentionnent des questions théoriques, une allusion au réel ou à des données empiriques, mobilisent une des méthodes de la collection tout en soulignant quelque chose de nouveau (en qualifiant quelque chose d'ancien ou de nouveau – au sens de la catégorie présentée plus haut). Un seul texte répond à l'ensemble de ces critères :

Les plateformes technologiques en sciences de la vie, entre bricolage et innovation. La présente communication s'intéresse au phénomène d'émergence de plateformes technologiques dans les sciences de la vie, et à ses conséquences scientifiques et organisationnelles. L'enquête est basée sur l'analyse du recensement des plateformes technologiques en sciences de la vie entrepris en 2001 par le CEA, l'INSERM, l'INRA et le CNRS, ainsi que sur des entretiens approfondis menés auprès de responsables des plateformes recensées, et des dirigeants d'entreprises privées ayant une activité de plateforme technologique dans le domaine des biotechnologies. Les options théoriques retenues sont à la croisée d'une sociologie de la production scientifique en termes de « reconfigurations » (Knorr-Cetina, Hilgartner, Callon, etc.) et de cycles de vie des assemblages instruments-connaissances-compétences (Cambrosio, Fujimura, etc.) ; d'une analyse des dispositifs de « recherche-technologie » (Joerges, Shinn) ; ainsi que d'une réflexion sur les conséquences en termes de design organisationnel des micro-régulations et des activités de « bricolage organisationnel ». On montre que selon l'état de la technique, le degré de stabilisation des connaissances, et le caractère routinisé ou ésotérique des compétences mises en jeu, la forme organisationnelle et le type de biens produits par une plateforme sont appelés à se modifier, confirmant ainsi l'idée d'un « cycle de vie » de ces assemblages. On assiste également à l'émergence de collectifs nouveaux, rassemblant scientifiques, industriels et intermédiaires, qui s'attachent à résoudre les problèmes de gestion des plateformes en inventant de nouvelles règles et procédures d'évaluation, et en proposant également des innovations organisationnelles. Les analyses montrent également l'influence des « cultures épistémiques » sur les négociations et la construction des règles d'accès. Cette influence s'exprime à travers le rôle structurant des petites décisions (micro-règles) pendant l'exploration des options organisationnelles, et a un impact plus large sur le modelage des relations science-industrie, et des politiques de la science. [A Peerbaye, RTF "Sciences, innovations technologiques et société"]

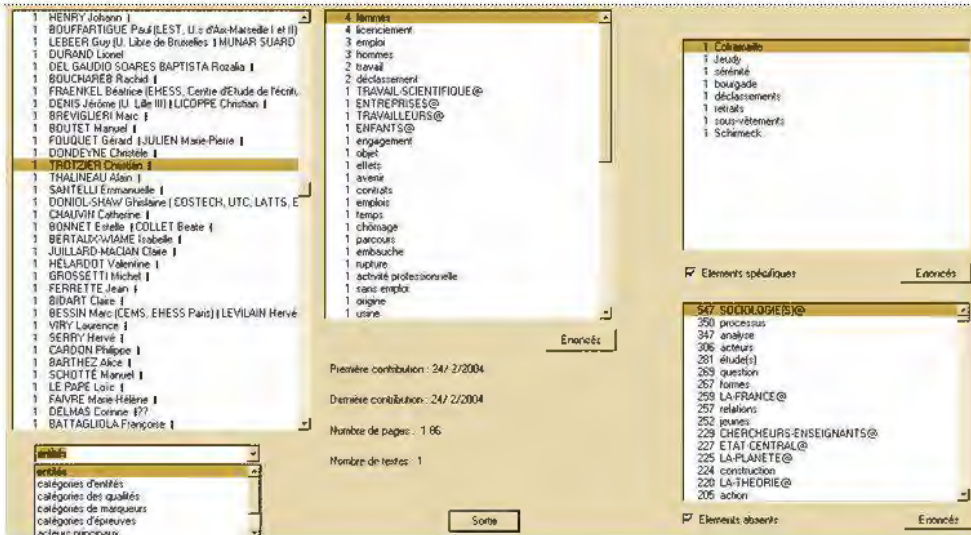
- 45 Symétriquement, l'on peut rechercher des textes qui ne vérifient aucun de ces critères. 27 textes ne répondent à aucune des conditions citées plus haut. Le propos n'est pas d'être normatif en séparant les "bons" résumés de ceux qui n'auraient pas dû être acceptés car trop à la marge de la discipline. Il s'agit au contraire d'identifier des façons de présenter un résumé qui, sans se soumettre au jeu de contraintes présenté ci-dessus, satisfont toutefois les exigences d'un appel à communication. Celui-ci est en fait tout aussi exemplaire que le précédent.

L'avenir inégal. Trajectoires de femmes et d'hommes après un licenciement collectif.

En 1981, les cent soixante-dix derniers salarié-e-s de Coframaille, usine de sous-vêtements de Schirmeck, sont licencié-e-s. Dans la même bourgade située à 45 kilomètres de Strasbourg, Jeudy, une entreprise de mécanique, supprime cent quatre-vingt-dix emplois en 1983. L'objet de la recherche est d'étudier le devenir de ces femmes et de ces hommes au cours des vingt ans qui suivent la rupture de leur contrat de travail. Les retraits du travail des femmes de plus de cinquante ans, par le chômage ou l'inactivité, sont massifs. Les ouvrières qui retrouvent au moins un emploi après le licenciement ont subi, dans leur majorité, un déclassement qui trouve son origine dans une longue instabilité professionnelle et/ou des conditions de travail dégradées. La détermination à retrouver une activité professionnelle reste forte. Toutefois, durant les premières années qui suivent le licenciement, des ouvrières reconsidèrent leur engagement dans le travail rémunéré. Ainsi, les mères d'un enfant en bas âge demeurent sans emploi durant plus de cinq ans. Les employées retrouvent plus rapidement un emploi que les ouvrières, en subissant parfois un déclassement. Confiantes dans leurs chances d'embauche, nombreux sont les ouvriers à réagir avec sérénité au licenciement. La trajectoire comportant un ou deux contrats à durée indéterminée et à temps plein est dominante. Toutefois, les parcours plus tumultueux ne sont pas rares. Ainsi, la perte d'emploi associée à l'alcoolisme a des effets dévastateurs. Chez les hommes, les déclassements, nettement moins fréquents que chez les femmes sont souvent dus à une baisse de rémunération. [C. Trotzier Session mixte "Parcours de vie et dynamique sociale /"Travail (productif et reproductif), rapports sociaux, rapport de genre"]

- 46 Ce texte décrit largement un terrain mais n'utilise pas le vocabulaire dont nous avons supposé qu'il indiquait la marque d'une allusion au réel. De plus, le résumé donne des informations qui témoignent d'une probable enquête quantitative. Enfin, ce texte rend compte d'un travail à partir de récits biographiques qui figurent parmi les méthodes sociologiques. Toutefois, ne mobilisant pas explicitement les représentants de nos collections et catégories, il ne répond pas aux conditions que nous avons fixées. Comment éviter cette série d'erreurs ?
- 47 D'une part, il faut bien se souvenir que nos jeux de concepts ne sont jamais figés : on peut à tout moment ajouter ou retirer un élément d'une catégorie ou d'une collection, que cela s'applique généralement parce qu'il manquait ou que sa présence n'est pas pertinente à nos yeux, ou que cela s'applique à tel corpus dans lequel un terme ou un objet est mobilisé d'une façon atypique par rapport aux corpus précédemment étudiés. Ainsi, nous avons hésité à mettre le mot "trajectoire" dans la collection de nos notions sociologiques car il ne caractérisait pas, à nos yeux, de façon suffisamment décisive, le vocabulaire de la discipline. Il demeure que sa répétition marque un discours sociologique et, nous avons tenu compte de cette propriété pour la suite¹⁵.
- 48 D'autre part et comme suite logique de cette première remarque, un principe de travail avec Prospéro invite à ne pas conclure sur un seul indice en raison de l'infinie variété de la langue. Si nous souhaitions soutenir l'hypothèse que ce texte se situe à la marge de la

discipline – ce qui n'est pas notre avis –, il faudrait trouver d'autres indicateurs. Nous pourrions alors mobiliser d'autres outils pour confirmer – ou plutôt ici infirmer – l'hypothèse. Par exemple, nous pouvons utiliser cette fenêtre qui rend lisibles les éléments spécifiques et les éléments absents (sur cette photo d'écran, les entités) chez un auteur.



- 49 Sans même ouvrir le résumé (bien que tous les boutons "énoncés" invitent à garder le contact avec le texte), nous pouvons nous apercevoir que des éléments spécifiques renvoient vraisemblablement à un travail empirique. En effet, il semble difficile, dans le résumé d'une communication en sociologie, de parler de "sous-vêtements" sans que ces objets ne soient associés à une description de terrain.
- 50 Ceci nous invite à modifier les critères qui avaient conduit à écarter le texte pour un défaut de matériau empirique. Pour cela nous pouvons nous appuyer sur la création d'une nouvelle collection, ou compléter la collection nommée "objets usuels" qui contient des éléments trop familiers pour que, apparaissant dans ce corpus, ils aient des chances de renvoyer à autre chose que la description d'un terrain. Il suffit alors de modifier la règle en associant la présence d'objets usuels à un indicateur marquant l'existence de données empiriques. Ce travail invite à la réflexivité et Prospéro constitue, dans cette perspective, un appui décisif.

Comment répondre à un « appel à communication » ?

- 51 Il est une autre contrainte qui mérite d'être analysée séparément car elle semble de nature différente. La proposition de communication rédigée en fonction des éléments précédents devait être soumise à un réseau thématique qui, s'il l'acceptait, intégrait l'exposé du travail dans une session. C'est une chose que son résumé soit considéré comme suffisamment bon pour être présenté à l'AFS ; c'en est une autre qu'il le soit dans le réseau choisi. Or, les comités scientifiques des RTF ne se consultent pas pour envisager quel serait le lieu le plus approprié pour chaque proposition. Du coup, pèse encore sur les auteurs une contrainte de proximité entre leur résumé et l'appel à communication diffusé quelques mois plus tôt par le RTF dans lequel ils souhaitent s'inscrire.
- 52 Pour évaluer le degré de rapprochement entre ces deux textes, nous avons « détourné » un algorithme du logiciel qui permet de contrôler la proximité entre les textes¹⁶. A notre

ensemble de 1068 résumés¹⁷, nous avons donc ajouté les textes de présentation des 42 RTF. En lançant un algorithme qui calcule sur la base d'indices la proximité de chaque texte avec tous les autres, nous pouvions attendre, d'une part, que nombre des résumés proposés dans un RTF donné soient proches les uns des autres et, d'autre part, que ceux-ci soient proches de la présentation du réseau. Si tel était le cas, cela signifierait assurément que les sociologues se seraient ajustés à une « demande » qui prenait la forme d'un « appel à communication ». Toutefois, une simple consultation à l'œil nu des actes du congrès ou de simples discussions avec ces acteurs particuliers que sont nos collègues suffit pour prendre la mesure que beaucoup auraient pu, le plus souvent, communiquer dans un autre RTF que celui finalement choisi.

- 53 Si l'on veut décrire les processus par lesquels un résumé présenté dans un RTF donné semble plus proche d'un autre RTF, il nous faut avant tout préciser cette notion de proximité. L'algorithme prend tous les textes deux à deux et calcule un certain nombre d'éléments¹⁸ qu'ils ont en commun (à un pourcentage près). Il s'agissait dans un premier temps de définir la proximité pour fixer ce pourcentage. Nous avons procédé de façon empirique¹⁹ afin que l'algorithme retienne comme proches deux textes qui ont le même objet (par exemple, une étude de la formation des tailleurs de pierre) ou la même approche théorique (quand, dans les deux textes, on parle d'interactionnisme), mais en réformant des couples de textes qui n'ont en commun que la sociologie. Nous avons ensuite évalué la proximité entre les résumés et les appels à communication du RTF dans lesquels ils avaient été présentés, puis la proximité des résumés entre eux. Le tableau 6 livre l'ensemble des résultats. Il se lit de la façon suivante. Parmi les 49 résumés du RTF 4, 13 sont proches d'un appel à communication mais 4 seulement sont proches de l'appel du réseau RTF 4 ; par ailleurs, on identifie 96 couples de textes proches impliquant ceux de ce réseau – les autres éléments de ces couples se déployant dans l'ensemble des RTF que l'on trouve en colonne²⁰.

Tableau 6 : Feuille Excel

Tableau 6 bis : Libellés des réseaux

1	Savoirs, travail, professions
2	Migrations et relations interethniques
3	Normes, déviances et réactions sociales
4	Sociologie de l'éducation et de la formation
5	Classes, inégalités, fragmentations
6	Politiques sociales, protection sociale et solidarités
7	Vieillesse, vieillissement et parcours de vie
8	Sociologie du militaire, sécurité, armées et société
9	Sociologie de l'urbain et des territoires
10	Sociologie de la connaissance
11	Sociologie de la consommation
12	Sociologie économique
13	Sociologie du droit
14	Sociologie des arts et de la culture
15	Jeunesse, Ages de la vie, Générations
16	Sociologie clinique
17	Gestion politique du corps et des populations
18	Approches plurielles du Sujet
19	Santé, Médecine, Maladie et Handicap
20	Méthodes
21	Mouvements sociaux
22	Parcours de vie et dynamiques sociales

Tableau 6 (suite)

23	Sociologie du travail et activités
24	Travail (productif et reproductif), rapports sociaux, rapport de genre
25	Travail, organisations, emplois
27	Sociologie des intellectuels
28	Recherches en sciences sociales sur la sexualité
29	Sciences, innovations technologiques et société
30	Sociologie de la gestion
31	Sociologie du sport et des activités physiques
33	Famille, vie privée, vie publique
34	Les recompositions de l'espace politique dans la société contemporaine
35	Sociologie de l'engagement, de la vie associative et du bénévolat
36	Théories et paradigmes sociologiques, philosophie de la sociologie
37	Sociologie des médias
38	Sociologie de l'environnement
39	Identité - subjectivité - revendication - changement social
40	Sociologie des institutions
42	Sociologie de l'expertise des sciences humaines et sociales
43	Pluralisme religieux et citoyenneté

On trouve par ailleurs des sessions mixtes dont les thématiques sont au croisement de deux réseaux. C'est le cas pour les RTF, 22 & 24, 29 & 31, 29 & 33, 4 & 5, 5 & 9, 2 & 16.

54 La lecture de ce tableau est difficile du fait du nombre important de RTF que l'on ne peut désigner que par leur numéro. Malgré tout, nous le reproduisons dans son intégralité car

il fournit des informations qui peuvent intéresser les acteurs de ces réseaux. Nous ne commenterons ici que quelques points saillants permettant de discuter notre propos.

- 55 Nous repérons 108 résumés proches d'un appel à communication, mais dans 29 cas seulement il s'agit du RTF choisi. Cette distance apparente peut s'expliquer par le fait que les formats d'un appel à communication et d'un résumé coïncident mal. Toutefois, on peut s'étonner du fait que l'on trouve plus de résumés proches d'un appel à communication d'un autre réseau : cela nourrit l'hypothèse selon laquelle les sociologues ont choisi leur réseau pour d'autres raisons que la proximité de leurs préoccupations avec celles affichées par les animateurs des différents RTF.
- 56 Par ailleurs, la diagonale de la matrice pointe sur des résumés proches et présentés dans le même réseau. Certains RTF apparaissent du coup comme particulièrement « cohérents » au sens d'une proximité des préoccupations communes des chercheurs qu'ils rassemblent. C'est notamment le cas pour les RTF 1 (Savoirs, travail et professions), RTF 4 (Sociologie de l'éducation), RTF 6 (politiques sociales), RTF 24 (Travail productifs et reproductifs, rapport sociaux...) et RTF 30 (Sociologie de la gestion). Il convient de noter que certains RTF n'ont rassemblé aucun couple de contributions proches.
- 57 Enfin, ce travail permet d'évaluer la proximité entre les réseaux. La lecture des libellés des différents réseaux thématiques fait apparaître des objets proches. La sociologie du travail est, à titre d'exemple, éclatée dans trois RTF différents. Faut-il considérer que les sociologues étudiant « le travail » partagent des objets ou des paradigmes ou, au contraire, qu'ils restent attachés à des préoccupations très éloignées ? L'on note sur le tableau l'existence de 10 couples de textes proches dont l'un appartient au RTF 1 (Savoirs, travail et professions) et l'autre au RTF 24 (Travail productifs et reproductifs, rapport sociaux...). En revanche, le RTF 25 (Travail, organisation et emplois) apparaît comme très décentré par rapport à ceux-ci. Ceci nous invite à penser que plusieurs communications des RTF 1 et 24 auraient pu être présentées dans l'autre réseau²¹ et, partant, à interroger la pertinence de l'organisation de deux réseaux distincts. Mais cela relève de l'histoire de la constitution de ces réseaux, qui hélas ne figure pas dans le matériau analysé.

Conclusion

- 58 Même si, pour céder un instant au plaisir de la personnification, Prospéro a dû s'adapter à un matériel inhabituel et notamment dépourvu des dimensions temporelles qu'il affectionne particulièrement, il nous a néanmoins permis d'explorer cet ensemble de textes en relevant quelques propriétés intéressantes. On a pu voir la singularité de quelques réseaux thématiques dans leur utilisation de collections et catégories marquant le rapport à la sociologie, à ses concepts, ses auteurs, ses matériaux empiriques, ses méthodes. L'exploration de cette masse textuelle, largement facilitée par l'outil informatique, conduit à repérer des façons de désigner des terrains sans mentionner des mots et expressions que l'on croyait pourtant incontournables. L'enjeu de ce type de considération réside dans la construction de modèles plus précis pour tester les hypothèses du chercheur et ses interprétations. A un troisième niveau, le logiciel permet d'évaluer la proximité de l'ensemble des textes d'un corpus. L'algorithme nécessite sans doute d'autres réglages, mais il laisse apparaître des pistes intéressantes tant sur ses applications à d'autres corpus que sur celui étudié.

- 59 Pour proposer une analyse plus fine de la façon dont les sociologues rédigent leurs résumés, il faudrait assurément compléter ce corpus. "La méthode comparative est la seule qui convienne à la sociologie"²². A la forme radicale du précepte durkheimien, nous préférons cette formulation de Dewey : "sans collection de phénomènes observés en des endroits et en des temps différents, l'enquête fondée, qu'elle soit de sens commun ou scientifique, ne peut progresser".²³ En tout état de cause, il est nécessaire de faire varier notre matériel qui, en l'état, n'est constitué que d'un seul type de documents, qui plus est, uniformément datés. Nous avons tenté d'introduire de la variation en ajoutant les appels à communications qui furent rédigés quelques mois avant ces résumés. On pourrait ajouter une série de documents rendant compte de la constitution des réseaux thématiques – voire de transcriptions d'entretiens réalisés à cet effet –, ce qui permettrait d'en faire l'histoire et donc d'identifier des discussions, des arguments, des accords qui ont été nécessaires pour fixer le titre du réseau et l'appel à communication. Cela éclairerait notamment les hypothèses sur la proximité de certains réseaux.
- 60 Mais l'on peut aussi rajouter des textes d'autres congrès de sociologues, par exemple ceux de l'Association Internationale des Sociologues de Langue Française (AISLF) ou les futures rencontres de l'AFS. Ainsi, pourrait-on évaluer des différences entre les deux associations et, à l'avenir, repérer les évolutions des communications en sociologie, identifier les concepts, les auteurs cités, les objets, les méthodes qui apparaissent ou disparaissent. Dans ces conditions, Prospéro serait sur son terrain de prédilection.
-

BIBLIOGRAPHIE

Chateauraynaud F., 2003a *Prospéro: une technologie littéraire pour les sciences humaines*, Paris: Editions du CNRS.

Chateauraynaud F., 2003b *Pour un observatoire informatisé des alertes et des crises environnementales – une application des concepts développés lors des recherches sur les lanceurs d'alerte*, Rapport de recherche (Programme EPR 2000), Paris: CEMAGREF/GSPR (EHESS), pp 150-154.

Chateauraynaud, F., Charriau, J.-P., 1992, « Hétérogénéité d'une machine sociologique », *Technologies, idéologies, pratiques*, Vol. X, n°2-4, pp. 337-349

Dewey J., 1938, *The Theory of Enquiry*, (trad. en Français, 1967), Paris : PUF (2ème édition, 1993), p. 257 (p. 183 dans la version anglaise).

Durkheim E., 1894, *Les règles de la méthode sociologique*, Paris : PUF, 23^e édition (1987), p. 124.

Latour B., 1984, *Les Microbes : guerre et paix*, Paris : Métailié

NOTES

1. Notons que six sessions mixtes (entre deux réseaux), furent organisées.

2. Sur l'inscription originelle de cette technologie littéraire au sein, notamment, d'une sociologie de la critique, voir Chateauraynaud et Charriau (1992).

3. La question n'est pas pertinente dans le seul cadre de cet exercice sur les résumés de l'AFS : les usages du logiciel tendent effectivement à se diversifier avec la multiplication de ses utilisateurs, certains d'entre eux souhaitant analyser un matériel textuel très varié et ne répondant pas toujours à ces propriétés, comme des transcriptions de « focus groups ».

4. Voir Chateauraynaud, 2003a

5. Si on cumule les expressions commençant par « sociologie » et celles avec « Sociologie », les deux premières de la liste obtiennent respectivement 21 (pour la connaissance) et 16 (pour le travail).

6. Appelons SOCIOLOGIE(S)@ cet être fictif. Le symbole @ sert simplement à distinguer ces êtres fictifs dans les différentes listes (comme par exemple sur le tableau 1 figurant en annexe) ; ainsi, on maintient visible à tout moment le fait qu'il s'agit d'un regroupement de plusieurs entités.

7. Pour lire les autres représentants, on peut se reporter au Tableau 2 (annexe).

8. Les délais entre la soumission de la proposition de communication et le congrès furent variables selon les réseaux thématiques, mais il s'agissait généralement d'envoyer un texte pendant l'été 2003. Les auteurs furent prévenus de l'acceptation de leurs contributions à l'automne pour un congrès fin février 2004. Ce calendrier, connu de tous, a permis à certains d'envisager avec plus ou moins de précision le type de travail qu'ils pourraient présenter plus de six mois plus tard.

9. Pour créer l'association et légitimer son fonctionnement, il a été question à plusieurs reprises d'organiser la discipline en réaction à l'affaire Teissier. Sans entrer dans le débat sur la pertinence contestée de cet argument, l'on peut considérer que pour certains, un enjeu de ce congrès était précisément de montrer ce qu'est la sociologie, enjeu qui a pu peser comme un impératif pour les étudiants et les chercheurs évoluant au marge de la discipline.

10. En fait, le plus bas score est réalisé par une session commune à deux RTF : "famille, vie privée, vie publique" / "sciences, innovations technologique et société". Toutefois, ces deux réseaux mobilisent individuellement beaucoup plus cette collection, que ces trois RTF.

11. Notons que deux réseaux mixtes figurent à nouveau parmi les groupes qui ne font pas référence à ces auteurs (les RTF séparément mobilisent cette collection).

12. Cette notion développée initialement dans la sociologie des sciences (Latour., 1984) permet de restituer aux acteurs la maîtrise de leurs arguments et de leurs dispositifs - ils en sont les auteurs - mais aussi à marquer une transformation dans l'analyse des textes et des discours : il n'y a plus de coupure entre le monde des idées et le monde de l'action. Les auteurs de textes sont traités comme des acteurs ou des actants intervenant pleinement dans la définition d'enjeux, d'intérêts, d'arguments et de jugements. Dans Prospéro, cette notion reçoit néanmoins un contenu plus technique : il s'agit des entités du corpus qui sont aussi déclarées comme des auteurs de textes et qui ont pour propriété d'apparaître au centre d'au moins un des textes du corpus.

13. Il s'agit d'une section universitaire (n° 74 au CNU).

14. Cette formule signifie que l'on recherche un segment constitué d'un représentant de la catégorie de qualité "Nouveauté" (contenant des éléments comme "nouveau", "nouvel", "nouvelle", "novateur"...) suivi d'une entité et d'une autre entité séparée de la première par un maximum de trois objets.

15. Si la méthode de travail n'était pas cumulative, le gain de que permet Prospéro en serait considérablement réduit. On reprend évidemment des dictionnaires et jeux de concepts d'un corpus à l'autre, mais il est toujours nécessaire d'en adapter les contenus au matériel dont on dispose et aux questions qu'on lui adresse. Ce mot "trajectoire", qui figure désormais dans cette catégorie, peut polluer une étude sur un corpus sportif dans lequel on parlerait de "trajectoire d'un ballon" ou sur un dossier sur la balistique. Le mot "trajectoire" peut aussi être utilisé pour désigner un parcours professionnel, mais sans que l'auteur ne l'appréhende de façon sociologique. C'est la raison pour laquelle nous avons également hésité pour le mot carrière car, en dépit de son importance dans une tradition sociologique bien connue, il reste souvent utilisé par la presse pour décrire le parcours d'un homme politique ou d'une personnalité connue.

16. Cet outil est particulièrement utile pour éprouver la pertinence des corpus, puisque l'on est parfois tenté de les compléter en ajoutant de nombreux textes qui restent très proches les uns des autres (par exemple, une série de dépêches d'agence de presse) sans que la valeur ajoutée au corpus soit significative. Sur ce point, on peut se reporter à Chateauraynaud, 2003b

17. 19 textes sont en effet des doublons, c'est-à-dire que pour chacun d'entre eux il existe dans le corpus un autre texte parfaitement identique (même auteur, même titre, même résumé) mais figurant dans un RTF distinct. Nous avons pris le parti de conserver ces doublons car ils témoignent de la proximité des RTF dans lesquels ils ont été présentés.

18. Il s'agit plus précisément du nombre d'objets des deux textes, et de la somme des éléments contenus dans les intersections de leurs entités, de leurs qualités et de leurs acteurs principaux. Pour une description plus précise de ces notions d'entités, de qualité ou d'acteurs principaux, on peut se reporter à : Chateauraynaud, 2003a.

19. Nous avons fait varier le pourcentage d'acteurs principaux que les deux textes devaient partager pour être considérés comme proches. Ce type de réglage mérite d'être testé sur des corpus assez divers pour être validé.

20. La matrice est donc, par nature, symétrique. La diagonale correspond à des textes proches d'autres résumés présentés dans le même RTF.

21. D'autant plus que les résumés furent rédigés sans références directes aux autres textes.

22. Voir Durkheim, 1894.

23. Voir Dewey, 1938.

RÉSUMÉS

Le but de cet article est d'éprouver l'intérêt du logiciel Prospéro pour l'analyse du corpus des résumés du premier congrès de l'Association Française de Sociologie (AFS). Le logiciel étant construit principalement pour étudier des corpus de textes marqués par des variations (en particulier dans les argumentations et dans le temps), nous avons considéré ces résumés comme un *dossier*, c'est-à-dire comme un ensemble de textes marqués par l'incertitude et les contraintes des auteurs proposant une communication dans un Réseau Thématique (RTF). Nous avons dans un premier temps montré combien les opérations de codage, qui dans Prospéro appartiennent au chercheur, sont décisives pour faire varier le niveau d'analyse et évaluer ses hypothèses. Dans un

second temps, nous avons réalisé un modèle tentant de décrire les attentes d'un lecteur d'une proposition de communication. Nous avons alors construit des catégories, des collections et des êtres fictifs afin d'éprouver ce modèle. Le logiciel apparaît dans ces conditions comme un outil permettant de repérer des propriétés singulières du corpus (par exemple de discriminer l'usage d'une catégorie ou d'une collection en fonction du RTF) mais aussi de pointer des textes atypiques qui invitent à réviser le modèle. Enfin dans un troisième temps, nous avons étudié la façon dont les auteurs se soumettent à une contrainte de proximité entre leur résumé et l'appel à communication diffusé quelques mois plus tôt par le RTF choisi. Cette approche montre que seuls certains réseaux font apparaître une certaine unité en rassemblant des textes partageant des propriétés communes. A l'issue de ce travail qui permet de repérer quelques aspects intéressants de la sociologie française, nous pensons surtout avoir montré l'intérêt et les limites de Prospéro. Celui-ci est apparu comme un outil au service du chercheur permettant d'éprouver ses interprétations, ses hypothèses et ses modèles. Le logiciel serait encore plus utile si le corpus pouvait être complété par des textes permettant d'obtenir une variation (notamment historique) et l'analyse de la production sociologique y gagnerait assurément.

The Program Prospero Tried Out Against A Corpus of Sociological Abstracts: The objective of this article is to test the potential of the program, Prospero, for analyzing the abstracts of the first congress of the French Sociological Association (AFS). Prospero was conceived as a program for the analysis of corpora of texts characterized by a certain variability (particularly in their use of argumentation and time). We consider these abstracts as a "dossier"; that is as a collection of texts characterized by the unknowns and the obligations of authors proposing a presentation in a Thematic Network (TN) session. To begin with, we show how much recoding — done by the researcher in the case of Prospero — is decisive in determining the level of analysis and of hypothesis testing. We then propose a model describing the expectations of a potential presenter and construct categories, collections and fictitious entities to test this model. The program, in this context, functions as an instrument for identifying singular properties of the corpus (for example, judging the use of a category or a collection to characterize a TN). But it also identifies atypical texts which can encourage the modification of the model. Finally, we study how presenters adapt to the requirements of a proximity between the abstract and the call for papers distributed a few months before by the TNs. This approach shows that only a few TNs bring together texts sharing similar characteristics. This work permits us to identify certain interesting aspects of French sociology and, in particular, the possibilities and the limits of Prospéro. The program functions as a research tool which can test a researcher's interpretations, hypotheses and models. The program would be even more useful if corpora could be complemented by texts permitting a variation (particularly historic) which would benefit the analysis of sociological scientific production.

INDEX

Keywords : Computer Programs, Prospero, Analysis of Corpora, Coding, Interpretation, Proximity, Sociology

Mots-clés : Logiciel, Analyse de corpus, Codage, Proximité, Sociologue

AUTEUR

PATRICK TRABAL

Lab. Sport & Culture (Université Paris X) et GSPR (EHESS) ; ptrabal@u-paris10.fr