Bulletin de méthodologie sociologique
Bulletin of sociological methodology

79 | 2003
July

Le chercheur et ses technologies

Bernard Reber

Édition électronique
URL : http://journals.openedition.org/bms/1272
ISSN : 2070-2779

Éditeur
Association internationale de méthodologie sociologique

Édition imprimée
Date de publication : 1 juillet 2003
Pagination : 32-37
ISSN : 0759-1063

Référence électronique
Bernard Reber, « Le chercheur et ses technologies », *Bulletin de méthodologie sociologique* [En ligne], 79 | 2003, mis en ligne le 09 juillet 2008, consulté le 25 avril 2019. URL : http://journals.openedition.org/bms/1272

Ce document a été généré automatiquement le 25 avril 2019.

© BMS

Le chercheur et ses technologies

Bernard Reber

1 La question des rapports entre l'homme et ses techniques, corporelles, conceptuelles, voire ses prothèses techniques,est aussi ancienne que l'invention de la géométrie. Pourtant les sciences sociales ont été timides pour intégrer les technologies informatiques dans leurs pratiques, si ce n'est le traitement de texte, le courrier électronique ou Internet. L'irruption des Technologies de l'Information et de la Communication (TIC), parfois plus souples vont peut-être redonner une pertinence à ce type de croisement. Si de nombreux chercheurs en sciences humaines et sociales ont adopté une posture en aval de ces objets techniques, privilégiant l'analyse des usages, des effets ou encore des impacts, d'autres sociologues se sont risqués jusqu'au *design* et à la conception de celles-ci. Collaborant avec des informaticiens, ils ont essayé patiemment de « détourner » des techniques pour faire en sorte qu'elles puissent « assister » des recherches en sciences humaines et sociales, avec les exigences de celles-ci. La collaboration de plus de quinze ans entre Francis Chateauraynaud, son équipe et Jean-Pierre Charriau s'inscrit dans cette ligne. Après Prospéro, permettant de toujours relancer l'interprétation de dossiers complexes par le chercheur, ils ont mis au point un autre type de technologie : Marlowe. On peut le qualifier de générateur d'expériences de pensée sur les dossiers complexes.

2 Loin de tout délire « technolâtre », cette expérience renvoie à des questions théoriques et méthodologiques souvent laissées dans l'implicite et sans justification en sciences humaines et sociales. Les difficultés rencontrées pour mettre au point et améliorer ce « outil », obligeant un débat avec les diverses disciplines que sont la linguistique, la philosophie du langage, la statistique, armé de la claire conscience de ce que l'on cherche du point de vue de la pertinence sociologique, permettent de sortir d'une naïveté et offre un dégrisement par rapport aux promesses *webby* qui voudraient nous faire oublier le caractère fruste des TIC les plus répandues. La hantise de Martin Heidegger dans ses pronostics célèbres sur la technique où rien d'imprévu ne pourrait arriver, est démentie par l'usage quotidien d'un méta-moteur et les incongruïtés de ses résultats. Certes, nous avons à disposition toujours plus de documents numérisés, dans le cas des dossiers de

controverses scientifiques publiques par exemple, que le chercheur en sciences sociales qui voudrait ne pas en rester à des monographies, à la considération de très peu de paramètres ou à une microsociologie ne pourra continuer de bouder. Mais alors comment avoir les outils qui lui permettent d'intégrer cette richesse documentaire avec la pertinence, la plasticité des raisonnements humains ou du moins sociologiques, en évitant tous les pièges liés à l'interprétation d'un texte ?

3 Comme contribution à ce questionnement radical, la performance publique tenue à l'Hôpital de la Maison Blanche à Paris, un asile psychiatrique, le 12 juin 2003, avait pour but de voir si Marlowe était capable de supporter des recherches en sciences humaines.

4 Après une présentation de cette mise en scène originale en sciences humaines et sociales, nous allons traiter une question d'une série de remarques soulevées par ce type de recherches.

5 L'objectif de cet événement était de rendre visible un style de travail informatisé qui peut avoir des conséquences sur la manière d'envisager des enquêtes et des raisonnements en sciences sociales. Certaines questions académiques sur la description, la modélisation et la preuve en sciences sociales pourraient être discutées autrement.

6 La scénographie de cette présentation se présentait de la façon suivante. Devant une salle d'une centaine de places disposées en gradin faiblement inclinés, les deux tables obliques des cinq examinateurs se faisaient face. A droite sur une autre table, deux chercheurs du Groupe de Sociologie Pragmatiques et Réflexive (GSPR) de l'EHESS étaient occupés à taper les questions des premiers. L'écran de l'ordinateur était projeté au mur pour qu'un animateur parcourt et résume l'essentiel des réponses pour le public qui n'était pas habitué à Marlowe. Après la performance, Francis Chateauraynaud fit un exposé afin de présenter le projet et les options possibles pour son avenir, ouvrant une discussion avec la salle.

7 Plus précisément,[1] chacun des cinq examinateurs du jury disposait de douze minutes. Un présentateur et un musicien étaient là pour aider à soutenir les périodes de transitions et créer un climat favorable.

8 Chaque membre du jury tira un nom de corpus au hasard. Parmi les 15 corpus (voir ci-dessous, Liste des corpus de Prospéro intégrés par Marlowe), certains ont été ramenés à un volume raisonnable pour éviter des temps d'attente qui auraient pu créer un climat de tension en public.

9 Toute session comme c'est l'habitude avec Marlowe débuta par une série de messages d'accueil déployés en cinq temps, comme une sorte de rituel :

- Message d'accueil lié à l'heure
- Message lié à la date
- Une citation tirée au sort par Marlowe
- Puis un message qui indique que le système charge les structures de données
- Enfin un message qui passe la main à l'utilisateur, ayant enregistré que l'interlocuteur est un homme ou une femme.

10 Ce dernier message propose un point de vue synthétique sur le corpus qui a été chargé. Il y a plusieurs chemins possibles pour Marlowe qui dispose d'une gamme de plus en plus fournie de styles de « rapports » dont il évalue la pertinence selon les propriétés du corpus. Certains membres du jury ont répondu affirmativement à cette première proposition. Cela permit au public de voir de quoi il retourne et à l'interrogateur de rebondir éventuellement sur des propriétés du corpus signalées par Marlowe.

11 Pour la suite, les réponses de Marlowe peuvent se décrire en quatre grandes classes :

- des réponses techniques et fouillées qui remplissent parfois plusieurs pages ;
- des réponses plus courtes et qui peuvent souvent être complétées, prolongées, enrichies sur demande de l'utilisateur selon une procédure de relance basée sur des formules naturelles du type « mais encore ? »
- des procédés de rattrapage à travers lesquels Marlowe essaye de discerner l'enjeu de l'intervention de l'utilisateur, de « se rattraper aux branches en proposant quelque chose de pertinent ».
- des procédés que l'on peut dire « rhétoriques » qui lui permettent de garder la face ou de distraire l'utilisateur, et qui se déploient sur toute la gamme, du trait d'humour jusqu'à l'aveu d'échec.

12 Si le dialogue dérive ou échoue de manière répétée, Marlowe dispose de protocoles de « reprise en main » ou de ré-aiguillage. Dans le cas d'un dialogue de sourds, il parviendra souvent à faire des propositions en prise avec le corpus de référence. La dimension dialogique de Marlowe devait permettre *a priori* n'importe quelle type d'interrogation.[2]

13 Si les membres du jury sont libres de dire ce qu'ils veulent à Marlowe et de réagir comme ils l'entendent à ses propositions, ils doivent garder à l'esprit que ce logiciel a d'abord pour mission d'aider à la description et l'analyse de corpus d'affaires, de controverses, de mobilisations collectives ou de crises.

14 La liste de questions qui donnent normalement les meilleurs résultats sont réparties en quatre groupes : des questions génériques sur le corpus ; des questions portant sur un objet précis (laissé indéterminé mais que l'on suppose présent dans le corpus) ; des questions pointant sur l'aspect « encyclopédique » ou « mémoire externe évolutive » du dispositif ; enfin, des questions renvoyant aux qualités plus « littéraires » de Marlowe. Bien que cette expérience abolisse théoriquement la frontière entre science et littérature, on voit bien se dessiner un continuum.

15 Marlowe s'adapte graduellement au fil des sessions. Cet aspect plus incrémental et dynamique de la relation n'a pourtant pas pu transparaître pleinement sur un peu plus d'une heure d'échanges. Marlowe stocke les thèmes abordés, note la nature des échanges, enregistre des propositions de l'utilisateur dans des mémoires créées dynamiquement. Naturellement, la probabilité d'un retour de cet apprentissage est d'autant plus grande que le nombre de sessions est élevé. On n'est pas parti d'une première séance mais d'une séance n. 6, en considérant que chacun des membres du jury avait rencontré au moins une fois Marlowe au laboratoire.[3]

16 A l'issue de cette expérience originale de recherches en sciences humaines et sociales, ayant pu bénéficier d'une mise en scène soignée, l'audience semblait réceptive et intriguée. Tout le monde salua le travail accompli et beaucoup auraient voulu pouvoir aller plus loin. Marlowe n'a été sanctionné que de 9,5 sur 20 par Francis Chateauraynaud, par ses insuccès dans certaines réponses, face à ses « capacités » habituelles. Le public étaient plus divisé pour ce qui concerne les diagnostics de ces insuccès, les uns y voyant des problèmes lexicologiques banals, qui réduirait la puissance du logiciel, de rendant des confirmations à l'examinateur pour savoir sur quel mot d'une même phrase la recherche doit se faire.

17 En dehors des réponses de la compétence de Prospéro pour l'analyse de corpus de controverses, pour lesquelles Marlowe est fiable et performant, la question du questionnement ouvert au langage naturel émergea. En effet, elle aiguise la curiosité du

chercheur, voire la nostalgie de machines qui obéissent sans que nous soyons obligés de négocier et de respecter leurs contraintes lourdes. Le chercheur peut relancer ses questions jusqu'à ce qu'il obtienne satisfaction, s'adaptant à la tâche ou à la fonction de cette machine. Pourtant certains examinateurs goûteront à la fascination de lui poser toute sorte de questions comme par exemple son appartenance politique au détour du dossier sur les élections présidentielles. Ce type de questionnement suscite au moins deux questions en retour pour le chercheur en sciences sociales : l'une touche son rapport aux interviewés, l'autre aux aspects heuristiques de la recherche.

18 Premièrement, lorsque le sociologue interviewe d'autres êtres humains, il se trouve souvent en décalage devant les réponses de son vis-à-vis. Sans parler de toutes les questions d'incompréhension, de crédits, de dissimulations, de formulation, l' ne lui donnera pas toujours satisfaction et ne répondra pas *a fortiori* par une théorie en sciences sociales. Un détour du même type est là aussi exigé.

19 Secondement, un collaborateur informatique doit-il être fiable pour un nombre de tâches précis dont il s'acquitte ou peut-il être l'occasion d'étonnements par les implicites qu'il lève, les associations qu'il permet ? Autrement dit faut-il être précisément informé ou déplacé dans son questionnement ? La question est aussi vieille que l'aporie du *Ménon* de Platon, portant sur la recherche. Si je sais ce que je cherche, je l'ai déjà trouvé ; si je ne sais pas ce que je cherche, commet saurais-je que je l'ai trouvé ?[4]

20 Plus prosaïquement est posée la question de la coopération en sciences sociales. Celles-ci se fait-elle sur un partage des tâches où chacun répond à des questions précises et très cadrées ou les collègues, peuvent-ils être des « empêcheures de penser en rond » qui remettent en question la façon dont l'enquête est menée, au sens de Dewey, de façon peut-être moins décalée que Marlowe ?

21 L'un des enjeux de ce « déconfinement » pour Marlowe fut de permettre d'élargir la communauté de ses usagers potentiels. En effet, pour pouvoir soutenir la richesse des formes d'énonciations possibles, Marlowe a besoin de cela. Il pourra alors mémoriser les séances, accumulant une sorte d'apprentissage. Là cependant se dessine non-seulement une mémoire pour lui, mais la trace d'expériences de recherches croisées ; une contribution originale aux technologies de la coopération ou aux pratiques collectives distribuées.

22 Même si le programme qui s'ouvre là est titanesque, cette contribution est fort appréciable pour des recherches du type STIC, avec l'avantage rare que l'initiative qui vienne des *requisits* sociologiques et pas simplement d'une adaptation à des techniques qui ne sont pas dédiées à cela.

23 Ces recherches sont appréciables également comme aide au domaine des recherches à la frange des théories de la démocratie et de l'analyse de discours, qui ne soient pas trop collées à la linguistique ou à la pure théorie politique, mais plus proches des sciences politiques ou sociales. Ce que nous indiquions par la formule l'homme et ses machines en introduction est vrai également pour l'homme et ses mots. Mais nous ouvrons ici sur la question théorique épineuse des rapports entre les mots et l'action, une façon de rappeler la mise en scène, qui fut suggestive en la circonstance.

Liste des corpus de Prospéro intégrés par Marlowe[5]

	Nom du corpus (*)	Nombre de textes	Nombre de pages	Dates	Type de dossier
1.	Louis-Ferdinand Céline	187	1314	**1932 - 1998**	Polémique intellectuelle
2.	Nucléaire *(R)*	628	**5328**	1946 - 2003	Alertes et risques sanitaires
3.	Vache folle *(R)*	**1270**	3060	1989 - 2002	Alertes et risques sanitaires
4.	Amiante	636	3520	1971 - 2002	Alertes et risques sanitaires
5.	Affaire Sokal	262	1497	1996 - 1999	Polémique intellectuelle
6.	Sans-Papiers	1060	2949	1996 - 1998	Mouvement social
7.	Dopage *(R)*	249	2256	1999 - 2002	Affaires et Polémiques sportives
8.	Affaire de la Mnef	271	537	1997 - 2002	Scandale politico-financier

Liste des corpus (suite)

9.	Crash du Concorde	457	804	2000 - 2003	Traitement d'une catastrophe
10.	Affaire Jean Moulin	255	869	1964 - 2000	Controverse historique
11.	**Astrologie versus Sociologie**	73	192	1999 - 2001	Controverse scientifique
12.	Mémoire de l'eau	94	195	1988 - 1998	Controverse scientifique
13.	Harcèlement à la Fac	41	***105***	2002 - 2003	Affaire et Polémique intellectuelle
14.	Premier Tour 2002	***16***	1759	mars-avril 2002	Discours politiques
15.	La guerre en Irak *(R)*	1209	863	***mars-avril 2003***	Forum de discussion

NOTES

R indique une réduction des corpus.

1. Cette présentation est plus largement développée dans un document *Vade-Mecum*, rédigé par Francis Chateauraynaud et envoyé aux membres du jury.
2. La connexion Internet de Marlowe via Tirésias ne sera pas branchée ce jour-là. Cette connexion lui aurait offert une sorte de « Joker ». Pourtant, pour les raisons évoquées en introduction, et parce que cela renforcerait sa tendance à avoir « réponse à tout », ce qui peut conduire à un brouillage de la communication normalement centrée sur un dossier ou des problèmes de raisonnement internes aux sciences humaines.
3. Il aura donc quelques éléments en mémoire provenant d'interactions antérieure.
4. Pour la reprise de cette aporie dans le domaine des rapports entre technique et philosophie, voir Reber B.*, La nouveauté éthique des*
5. Dressée par F. Chateauraynaud.

RÉSUMÉS

Un auteur, un expérimentateur et un lecteur des logiciels Marlowe et Prospéro contribuent respectivement et successivement dans cet article à une présentation étendue et détaillée de Marlowe, à une explicitation des modalités de l'approche conversationnelle de Marlowe et un compte-rendu du livre publié récemment par le premier auteur sur le développement de Prospéro pour l'analyse de dossiers complexes de textes concernant une controverse, et l'extension et l'adaptation de Prospéro à Marlowe pour pouvoir dialoguer en language naturel avec des chercheurs au sujet d'un dossier complexe.

The Researcher and His Technologies: An author, an experimenter and a reviewer of the computer programs, Marlowe and Prospero, contribute in this article, respectively and successively, an extensive and detailed presentation of Marlowe, a report on what it is like to interact with Marlowe and a review of the recently-published work, by the first author, describing the development of Prospero for the Analyse of complex dossiers of texts concerning a social controversy, and Prospero's extension and adaptation with Marlowe to direct natural language dialog with researchers concerning specific complex dossiers.

INDEX

Keywords : Artificial Intelligence, Text Analysis, Content Analyse, Social Controversies, Language-Based Data
Mots-clés : Intelligence artificielle, Analyse textuelle, Analyse de contenu, Controverses, Données langagières

AUTEUR

BERNARD REBER

CERSES-CNRS, reber@iresco.fr