TERMIUM sur CD-ROM

John Carey, Sylvie Dupont et Christine Leonhardt

Volume 36, numéro 1, mars 1991

La terminologie dans le monde : orientations et recherches

URI : <https://id.erudit.org/iderudit/002206ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

[Découvrir la revue](#)

Citer cet article

Carey, J., Dupont, S. & Leonhardt, C. (1991). TERMIUM sur CD-ROM. *Meta*, 36(1), 174-181.

TERMIUM SUR CD-ROM

JOHN CAREY, SYLVIE DUPONT ET CHRISTINE LEONHARDT
Secrétariat d'État du Canada, Hull, Canada

INTRODUCTION

Le Secrétariat d'État du Canada abrite un centre de terminologie de haut calibre, dont les activités aident le gouvernement canadien à atteindre les objectifs qu'il s'est fixé en matière de langues officielles. Cette expertise linguistique a aussi un impact culturel certain : des services gouvernementaux, des organismes nationaux, des établissements d'enseignement ainsi que des entreprises privées, tant au Canada qu'à l'étranger, sollicitent son avis, son expérience et ses compétences.

Le Secrétariat d'État exploite, depuis 1974, une banque de données linguistiques, TERMIUM, déjà bien connue dans les milieux de la traduction et de la terminologie, puisqu'elle compte près de 3 000 usagers tant en Amérique du Nord qu'en Europe.

Tout comme la langue elle-même et la terminologie qu'elle emploie, TERMIUM est en constante évolution. À cet égard, TERMIUM, depuis longtemps accessible en ligne, est aujourd'hui disponible sur CD-ROM (Compact Disk-Read-Only Memory), ou disque optique compact.

HISTORIQUE DU PROJET CD-ROM

Cette nouvelle étape dans le développement de TERMIUM a été amorcée en 1985, lorsqu'une compagnie canadienne proposait au gouvernement fédéral de verser TERMIUM sur CD-ROM.

Au cours de l'automne 1987, le Secrétariat d'État a lancé un projet-pilote pour évaluer la réaction des utilisateurs de TERMIUM à cette nouvelle technologie. Pendant cet essai, la version expérimentale a été installée dans dix services du Bureau de la traduction et dans six entreprises privées canadiennes.

Les participants ont ensuite répondu à un questionnaire d'appréciation, fourni des données statistiques sur leurs séances de travail, et soumis périodiquement des listes de termes interrogés, pour fins de comparaison à la version en ligne, en vue d'évaluer la fréquence nécessaire de mise à jour du disque. Nous avons ainsi pu déterminer que seul 5 % du contenu du CD-ROM peut différer de la version du système en ligne entre chaque nouvelle copie si nous remettons aux abonnés un nouveau disque tous les quatre mois.

Bref, l'expérience s'est révélée un tel succès, qu'en octobre 1988 le Secrétariat d'État offrait à toutes les personnes et à toutes les entreprises déjà intéressées par TERMIUM, la possibilité de participer au deuxième volet du projet-pilote. Nous avons reçu 180 réponses à cette invitation.

Une autre évaluation des performances de cette version de TERMIUM est en cours. Les données recueillies permettront de valider les dernières modifications apportées à TERMIUM.

Il est déjà évident, d'après les commentaires des utilisateurs, que TERMIUM sur CD-ROM répond très bien à leurs besoins. Plusieurs facteurs contribuent à ce succès dont le plus important est sans aucun doute le contenu de la base de données, c'est-à-dire les trois millions de termes, d'appellations et de locutions constituant cet outil précieux

d'aide à la traduction et à la rédaction. Cependant, la technologie utilisée, qui permet de rendre ces informations accessibles de façon rapide et facile, joue elle aussi un rôle de premier plan dans cette réussite. Certains participants ont pu constater, par exemple, que le taux de réponse atteint 80 % alors que celui du système en ligne plafonne à 47 % depuis des années. Ils attribuent cette amélioration au fonctionnement du logiciel de consultation et aux modes d'indexation.

LE CONTENU DE TERMIUM SUR CD-ROM

Parmi les nombreux champs de la fiche TERMIUM originale, nous n'avons retenu pour copie sur CD-ROM que ceux nécessaires à la diffusion de l'information linguistique; les champs utiles à la gestion des données n'ont pas été saisis sur le CD-ROM. Chaque fiche comprend donc, pour chacune des deux langues, l'entrée principale, ses synonymes, variantes et abréviations, les termes clés, les définitions, les contextes et les observations.

Le CD-ROM contient également certains autres renseignements de la fiche originale permettant à l'utilisateur de juger de la pertinence des données linguistiques dans son contexte particulier. Les dates de création et de la dernière mise à jour de la fiche, par exemple, sont particulièrement utiles dans le cas des appellations et des termes des domaines de pointe. Les mentions de source des entrées et des justifications identifient les documents que l'utilisateur peut consulter à des fins de traduction ou de rédaction, et attestent de la crédibilité des informations linguistiques. Les domaines aident l'utilisateur à déterminer les contextes dans lesquels l'entrée et son équivalent peuvent être employés pour désigner la notion traitée sur la fiche, tandis que le fichier indique le degré de fiabilité des renseignements présentés. Le fonds identifie la nature du «problème» faisant l'objet de la fiche (terminologie, appellation ou difficulté de traduction).

LA DISPOSITION DES FICHES

La présentation des enregistrements est tout aussi importante que les méthodes d'accès. Puisque TERMIUM est le plus souvent consultée à des fins de traduction et que la langue de départ des demandes est le français ou l'anglais, la disposition des renseignements linguistiques en colonnes parallèles répond très bien aux exigences des utilisateurs. La langue d'arrivée des recherches étant le plus souvent le français, les renseignements dans cette langue apparaissent dans la colonne de droite.

Dans chaque langue, l'entrée principale est suivie, si la fiche en contient, de ses synonymes et abréviations. Viennent ensuite les justifications (définitions, contextes, observations), chacune commençant à la ligne et étant précédée d'un code qui identifie son type (DEF, CONT, OBS). Dans l'en-tête de la fiche figurent, en anglais et en français, les descripteurs de fonds, de fichier et de domaines. Les mentions de source et les dates de création et de mise à jour de la fiche apparaissent à la suite des données linguistiques. (Voir l'exemple de fiche joint en annexe.) Chacun des champs non linguistiques est explicitement identifié.

Tous les éléments mentionnés ci-dessus figurent dans le format d'affichage implicite des fiches sur CD-ROM. Le logiciel de consultation offre à l'utilisateur la possibilité de créer d'autres formats en fonction de ses besoins. Bien que la disposition ne soit pas modifiable, l'utilisateur peut sélectionner les champs qu'il désire afficher. Une fois les nouveaux formats créés, il est possible de passer d'un format à l'autre pendant l'affichage des fiches.

L'INDEXATION

La création d'index constitue une partie importante du traitement, étant donné que ce sont les index qui permettent aux utilisateurs d'extraire de la banque informatisée les données voulues. Dans le cas de TERMIUM, il faut fournir les moyens nécessaires pour donner accès de façon logique et efficace aux 915 000 fiches qui, après une compression préliminaire, occupent sur le disque optique compact les deux tiers de sa capacité. Le mode d'indexation a été choisi selon la nature et la structure des données et en fonction de l'utilisation qui en est faite.

TERMIUM a d'abord été conçue pour répondre aux besoins de la traduction, et ce principe fondamental oriente encore son évolution. Les usagers consultent la base de données dans le but de trouver l'équivalent d'une unité terminologique, d'une appellation ou d'une locution, qu'elle se compose d'un seul mot ou de plusieurs. Le critère de recherche le plus utile est évidemment cette unité. Par conséquent, il est essentiel de permettre aux usagers de taper tout simplement le terme dont il est question dans la langue de départ. Grâce au principe de création de fiches et aux index créés, TERMIUM peut répondre à ce besoin.

La création des fiches respecte le principe de l'uninotionnalité, ce qui veut dire que les terminologies partent d'une notion pour aller au terme désignant cette notion (et non d'une appellation de la notion dans une langue pour aller à une autre appellation dans une autre langue). Une seule fiche regroupe donc tous les renseignements, en anglais et en français, concernant une notion donnée. Dans leurs champs respectifs figurent l'entrée principale, ses synonymes et variantes, ses abréviations, ses équivalents et les explications relatives à la notion. En recherchant une des entrées sur la fiche dans la langue de départ, l'utilisateur a accès non seulement à l'équivalent, mais aussi à d'autres renseignements utiles.

Pour permettre l'accès tel que décrit, il a fallu créer des index des «vedettes» de la fiche, dans lesquels sont rangés, en ordre alphabétique et au complet, tous les termes, toutes les appellations et expressions figurant sur les fiches de la base de données. Nous avons choisi de créer deux index distincts pour les parties anglaise et française de la fiche. Il aurait été possible de mettre dans un seul index toutes les vedettes, sans égard à la langue, mais un tel index ne permettrait pas à la base de données de répondre de façon efficace aux demandes réelles des usagers, qui savent dans quelle langue ils cherchent un équivalent et qui ne veulent pas perdre du temps à regarder des fiches non pertinentes. De même, nous aurions pu établir un index pour chaque champ contenant une vedette; cependant, dans le cas de TERMIUM, ce mode d'indexation ne permettrait pas d'optimiser l'exploitation des données.

Grâce à la structure de la fiche et aux deux index de vedettes, il est possible d'utiliser les mêmes données pour la traduction de l'anglais vers le français et vice-versa. De plus, les fiches sont accessibles par les synonymes et les abréviations aussi bien que par la vedette principale; lors de la recherche, le système traite toutes les vedettes de la même façon.

Afin de rendre les index plus performants, quelques règles spéciales ont été suivies. Par exemple, certains caractères ou mots faisant partie des syntagmes ne sont pas indexés. Il s'agit de signes de ponctuation et de mots-outils (prépositions, conjonctions, articles) qui risquent d'empêcher l'accès aux fiches. Ces caractères et mots «vides» sont remplacés par un espace lors de l'indexage. Une liste plutôt limitée de mots vides a été définie pour chaque langue. Il existe une exception à la règle: dans le cas des abréviations, les mots vides *a*, *d* et *l* en français et *a* et *s* en anglais figurent dans l'index.

Les index des vedettes conviennent parfaitement à l'interrogation de la base de données dans le but de trouver l'équivalent d'un terme donné: si une fiche est repérée, la

pertinence de la notion peut être vérifiée et l'utilisateur peut confirmer que les termes dans la langue d'arrivée conviennent à son contexte. Cependant, pour certaines applications, d'autres modes d'indexation sont nécessaires afin de permettre aux utilisateurs d'exploiter tous les renseignements linguistiques compris sur les fiches. Par exemple, si le terme demandé ne correspond pas exactement à l'une des entrées de l'index des syntagmes, l'utilisateur devrait avoir la possibilité de formuler sa demande d'une autre façon.

Pour optimiser l'accès aux fiches, des index de mots ont également été créés. Il s'agit de deux listes alphabétiques, l'une correspondant à la partie anglaise de la fiche, l'autre à la partie française, de tous les «mots» (selon les règles d'indexation, un mot est une chaîne de caractères précédée et suivie d'un espace) qui constituent les vedettes ou les justifications des fiches. Grâce à ces index, l'utilisateur peut formuler des demandes de recherche par mot simple ou par combinaison de mots spécifiques qui servent à repérer toutes les fiches traitant de près ou de loin le problème de traduction auquel il fait face. Une recherche approfondie de la base de données est par conséquent possible.

Certains renseignements sur les fiches n'ont pas été indexés et ne peuvent donc pas servir de point d'accès à la base de données. Par exemple, l'indexation des champs «fonds», «fichier» et «dates» n'est tout simplement pas jugée intéressante du point de vue de l'exploitation de la base de données sur CD-ROM. Dans le cas des domaines, un développement aurait été nécessaire pour rendre utilisable l'index, d'autant plus que pour bien s'en servir, l'utilisateur devrait connaître à fond le système de classement TERMIUM.

Certains lecteurs trouveront peut-être intéressant le fait que les quatre index requièrent un total d'environ 100 mégaoctets (millions de caractères) sur le CD-ROM, tandis que les données elles-mêmes, après une compression préliminaire, en occupent 400.

LE LOGICIEL DE CONSULTATION

La convivialité et la souplesse du logiciel de consultation de TERMIUM sur CD-ROM ont contribué de façon importante à la réaction positive des utilisateurs. La nouvelle version de ce logiciel sera, nous l'espérons, aussi bien acceptée.

En ce qui concerne l'interface avec l'utilisateur, le logiciel fait appel à la technologie des fenêtres, permettant ainsi l'affichage de plusieurs types de renseignements à la fois. Par exemple, pendant la formulation d'une demande de recherche, la liste des index, les entrées de l'index sélectionné et la demande elle-même sont affichées simultanément dans trois fenêtres, ou zones de l'écran, distinctes. (Voir l'exemple d'écran de sélection joint en annexe.) Les fiches sont présentées, une à la fois, dans une seule fenêtre qui occupe tout l'écran.

Le logiciel offre plusieurs possibilités d'interrogation intéressantes, comme le balayage de l'index. Pendant la frappe du terme recherché, le curseur se déplace dans l'index et la fenêtre des entrées affiche en son centre l'entrée qui correspond aux caractères tapés, entourée des entrées précédentes et subséquentes. De cette manière, d'autres termes ou expressions qui pourraient servir à l'interrogation de la base de données sont automatiquement présentés. La sélection de ces entrées à partir de l'index est très simple : il suffit d'appuyer sur une touche.

L'interaction entre l'utilisateur et le logiciel de consultation est optimisée au moyen d'une combinaison de menus déroulants, de touches de fonction, de fenêtres de dialogue et de commandes données au clavier.

PERSPECTIVES D'AVENIR

Nous aimerions profiter de l'évolution technologique dans le domaine de la micro-informatique et rendre TERMIUM sur CD-ROM plus performante. Parmi les

améliorations que nous avons l'intention d'apporter au contenu, mentionnons l'ajout d'une base de données de décodage des codes de source figurant sur les fiches linguistiques et d'une base de données multilingues (espagnol et allemand). Dans ce dernier cas, il est cependant possible qu'un deuxième CD-ROM soit produit à cette fin, étant donné que les deux bases de données linguistiques ne viseraient peut-être pas les mêmes clients.

Du point de vue fonctionnel, une amélioration très importante que nous aimerions apporter est la possibilité d'utiliser le logiciel de consultation en mode résident, permettant ainsi à l'utilisateur de consulter le CD-ROM à partir de son logiciel de traitement de texte au moyen d'une commande au clavier. L'utilisateur pourrait couper une partie de la fiche repérée (par exemple, l'équivalent), retourner au traitement de texte et coller l'information copiée du CD-ROM dans la traduction.

La possibilité d'interroger la base de données en mode hypertexte serait aussi très intéressante. De cette manière, l'utilisateur pourrait indiquer, dans une fiche à l'écran, un mot ou un syntagme, puis donner une seule commande qui lancerait une deuxième recherche de la base de données avec comme critère de recherche le terme désigné.

Il faudrait aussi offrir une version de TERMIUM sur CD-ROM qui tourne sur les micro-ordinateurs Apple.

Les quelques améliorations mentionnées ici rendraient TERMIUM sur CD-ROM beaucoup plus intéressante. D'autres avenues seront aussi étudiées — ce produit n'est qu'aux toutes premières étapes de son évolution.

ASPECTS TECHNIQUES

Puisque TERMIUM sur CD-ROM tourne sur micro-ordinateur, cette version de la banque de données linguistiques offre certains avantages par rapport au système en ligne. Ainsi, l'accès à la base de données est beaucoup plus facile : il n'est pas nécessaire de passer par un réseau de télécommunications; aucun code d'utilisateur ni mot de passe ne sont requis. La disponibilité de la banque de données n'est limitée que par celle du micro-ordinateur auquel le lecteur de CD-ROM est relié. De plus, l'utilisateur peut interroger à volonté la base de données : il n'y a pas de frais de branchement risquant de limiter le nombre de termes interrogés, la durée de la séance d'interrogation ou la fréquence des séances de travail.

L'évolution de la technologie de l'interface lecteur de CD-ROM et micro-ordinateur rend très facile l'installation de notre produit. L'absence de normes techniques imposait des paramètres d'installation non standard aux éditions de TERMIUM utilisées au cours du projet-pilote; les normes établies dernièrement seront respectées pour les prochaines éditions.

Il suffira alors que le lecteur de CD-ROM soit relié au micro-ordinateur à l'aide du logiciel MS-DOS CD ROM Extensions¹ de Microsoft et d'un programme de commande de périphérique approprié. Le logiciel Extensions est normalement vendu avec le lecteur de CD-ROM mais il est aussi disponible comme produit distinct. Si les programmes nécessaires sont installés sur le micro-ordinateur, tout CD-ROM respectant la norme internationale ISO-9660, y compris TERMIUM, peut être lu. Ces développements rendent la consultation des CD-ROM facile et presque transparente. Les CD-ROM peuvent même être consultés à partir de micro-ordinateurs reliés en réseau local.

Les développements techniques touchent non seulement les utilisateurs de CD-ROM, mais aussi les éditeurs. L'évolution technologique a déjà résulté en la création de produits — matériel et logiciel — qui permettent même à des non informaticiens de préparer des données en vue de leur stockage sur CD-ROM. Le Secrétariat d'État étudie présentement la possibilité d'utiliser certains de ces produits afin de créer sur CD-ROM

plusieurs bases de données, y compris TERMIUM, sans passer par un bureau de service. La seule étape qui se ferait à l'extérieur du Ministère serait le pressage des disques.

L'ÉTAPE DE LA COMMERCIALISATION

La réaction favorable des participants au projet-pilote nous a encouragé à poursuivre la diffusion de TERMIUM sur CD-ROM. Par contre, et conformément à la politique gouvernementale de récupérer une partie des coûts des services offerts au public, le Secrétariat d'État a décidé de mettre en place un programme de recouvrement des coûts pour TERMIUM. C'est ainsi que la prochaine version du CD-ROM sera offerte par voie d'abonnement.

Le coût de l'abonnement² a été fixé à 1 500 \$ CA par année et comprendra :

1. *CD-ROM et logiciel de consultation* : une édition à jour du disque sera distribuée tous les quatre mois;
2. *Documentation* : guide d'interrogation, instructions d'installation, manuel de référence;
3. *Formation* : des séances d'initiation et de suivi seront offertes à intervalles réguliers dans tous les grands centres (Montréal, Toronto, Vancouver, Winnipeg, Moncton, Région de la capitale nationale, etc.).

L'utilisation de TERMIUM sur CD-ROM nécessite l'équipement suivant : un micro-ordinateur IBM PC/XT/AT/386/PS2 ou compatible avec au moins la version 3.1 de DOS et un minimum de 640 kilooctets de mémoire centrale; la version 2.0 ou une version plus récente du logiciel MS-DOS CD ROM Extensions de Microsoft (au moins la version 2.1 est nécessaire si vous utilisez la version 4.x de DOS); un lecteur de CD-ROM compatible avec Extensions.

Toute demande d'abonnement à TERMIUM sur CD-ROM doit être adressée à la Division de la promotion et de la coordination; un contrat de licence et un bon de commande seront alors expédiés. Voici nos coordonnées :

Division de la promotion et de la coordination
Direction de la terminologie et des services linguistiques
Langues officielles et Traduction
Secrétariat d'État du Canada
Ottawa (Ontario) Canada
K1A 0M5
Téléphone : (819) 997-3712
Télécopieur : (819) 994-3670

NOTES

1. Les droits d'auteur sont la propriété de Microsoft Corporation.
2. À noter que les ministères du gouvernement fédéral ainsi que les organismes figurant aux annexes 1 et 2 de la *Loi sur la gestion des finances publiques* sont, à l'heure actuelle, exclus du programme de recouvrement des coûts.
3. FindIt, le logiciel qui permet la consultation de TERMIUM sur CD-ROM, a été créé par la société Reteaco Inc., qui en détient les droits d'auteur.

ANNEXE 1

EXEMPLE DE FICHE

DATA COLLECTION

Terminology

FILE

Main File (1)

SUBJECT FIELD(S)

Radar and Radio Guidance

EN

sidelooking radar *a *CORRECT,
CORRECT
side-looking radar *b;c *CORRECT,
CORRECT *S.L.R.;SLR
side-looking antenna radar *j
*CORRECT,CORRECT *SLAR

S.L.R. *b *CORRECT,CORRECT
SLR *c *CORRECT,CORRECT
SLAR *j *CORRECT,CORRECT

DEF*A ground mapping radar used
aboard aircraft involving the use
of a fixed antenna beam pointing
out the side of an aircraft either
abeam or squinted with respect to
the aircraft axis. The beam is
usually a vertically oriented fan
beam having a narrow azimuth width.*

a

SOURCES:

a*INELE*1984***836
b*META*1988*33*3
c*OTANA*1977***S-11
d*ICAO-9294*1986*1**280
e*INDAH-F*1987*01**27
f*3ZIG*1989
g*LERAS*1987***37
h*THMOD*1982***285
i*AVIA-F*1985*40*3*259
j*AVIA-E*1985*40*1*47

DATE

890220

DATE OF UPDATE / DATE DE MISE À JOUR

890907

FONDS

Terminologie

FICHER

Fichier principal(1)

DOMAINE(S)

Radar et radioguidage

FR

radar à visée latérale *b *CORRECT,
CORRECT;MASC,MASC *R.V.L.
radar à vision latérale *c;g;h
*CORRECT,CORRECT;MASC,MASC
radar à balayage latéral *d;i
*CORRECT,CORRECT;MASC,MASC
radar à antennes latérales *e *A
VERIFIER,UNCONFIRMED;
MASC,MASC
R.V.L. *b *CORRECT,CORRECT
RVL *f *CORRECT,CORRECT

DEF*Radar permettant une
observation au sol à partir d'un
avion en vol horizontal.*h

ANNEXE 2

EXEMPLE D'ÉCRAN DE SÉLECTION³

COMMANDES : Recherche Fiches Copie Options Sortie Aide
 Base de données : TERMIUM Éditeur : Canada

<p style="text-align: center;">4 Index Plus :</p> <p>TEN TFR English words / Mots anglais French words / Mots français</p> <p style="text-align: center; border: 1px solid black; display: inline-block; padding: 2px;">FENÊTRE DES INDEX</p>	<p style="text-align: center;">Fréq. — 934837 Entrées Plus : ↑ ↓</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 10px; text-align: right;">1</td><td>side looking airborne multi miss</td></tr> <tr><td style="text-align: right;">1</td><td>side looking airborne radar</td></tr> <tr><td style="text-align: right;">1</td><td>side looking antenna</td></tr> <tr><td style="text-align: right;">1</td><td>side looking antenna radar</td></tr> <tr><td style="text-align: right;">1</td><td>side looking radar</td></tr> <tr><td style="text-align: right;">1</td><td>side looking radar system</td></tr> <tr><td style="text-align: right;">1</td><td>FENÊTRE DES ENTRÉES</td></tr> <tr><td style="text-align: right;">1</td><td>side looking system aperture</td></tr> <tr><td style="text-align: right;">2</td><td>side lot line</td></tr> </table>	1	side looking airborne multi miss	1	side looking airborne radar	1	side looking antenna	1	side looking antenna radar	1	side looking radar	1	side looking radar system	1	FENÊTRE DES ENTRÉES	1	side looking system aperture	2	side lot line
1	side looking airborne multi miss																		
1	side looking airborne radar																		
1	side looking antenna																		
1	side looking antenna radar																		
1	side looking radar																		
1	side looking radar system																		
1	FENÊTRE DES ENTRÉES																		
1	side looking system aperture																		
2	side lot line																		

Recherche : 7 fiches repérées Plus :

(TEN : side looking airborne modular mu ou side looking airborne multi miss ou side looking airborne radar ou side looking antenna ou side looking antenna radar ou side looking radar ou side looking radar system ou side looking sonar [ou] ...

FENÊTRE DE RECHERCHE

Enter=Sélection F1=Aide F2=Bases de données F3=Sortie F4=Et F5=Ou
 F6=Fenêtres F8=Nouv. demande F9=Fiches F10=Commandes Del=Annul. sélection