

La création d'une banque de terminologie et perspectives d'avenir en terminotique

Groupe de travail TERMIUM III

Volume 32, Number 2, juin 1987

Vers l'an 2000. La terminotique, bilan et prospectives

Objectives: Year 2000 Terminotics. State of the Art, Prospects for the Future

URI: <https://id.erudit.org/iderudit/004465ar>

DOI: <https://doi.org/10.7202/004465ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

Cite this article

Groupe de travail TERMIUM III (1987). La création d'une banque de terminologie et perspectives d'avenir en terminotique. *Meta*, 32(2), 111–123. <https://doi.org/10.7202/004465ar>

LA CRÉATION D'UNE BANQUE DE TERMINOLOGIE ET PERSPECTIVES D'AVENIR EN TERMINOTIQUE

GROUPE DE TRAVAIL TERMIUM III
ALDEAN ANDERSEN, MARIE BERGERON, GILBERT DUPUIS,
CHRISTINE LEONHARDT, JEAN-FRANÇOIS MARROTTE
ET RITA MORITZ

*Direction de la terminologie, Direction générale de la terminologie
et des services linguistiques, Secrétariat d'État du Canada,
Ottawa, Canada*

Qu'elle semble déjà lointaine l'époque où le traducteur peinait sur son texte, enseveli sous une tonne de dictionnaires et de vieilles boîtes à chaussures qui renfermaient le fruit de longues heures de recherches ! La traditionnelle boîte à souliers s'est transformée en un terminal à écran ou un micro-ordinateur donnant accès à des bases de données ou, en ce qui nous concerne, à des banques de terminologie.

Ces banques, on s'en doute, ne sont pas nées du jour au lendemain ni ne sont demeurées statiques. Depuis leurs modestes débuts, les banques de terminologie se sont, en effet, perfectionnées et spécialisées. On peut citer, à titre d'exemple, la transformation de BTUM¹ à TERMIUM² III que l'on connaît aujourd'hui. En effet, dès son acquisition par le Bureau des traductions, BTUM devient Termium I par la modification de son logiciel. Termium I se transforme ensuite rapidement par son transfert sur un logiciel de deuxième génération pour devenir Termium II. Finalement, après cinq années d'utilisation et de perfectionnement, est lancé le projet TERMIUM III, qui correspond au développement d'une banque de troisième génération, développement que les changements technologiques et les besoins des usagers ont rendu nécessaire. Ce nouveau logiciel a, entre autres choses, permis l'intégration au fonds terminologique de données multilingues, plus précisément des fiches en espagnol et en allemand couplées à l'une des deux langues officielles du Canada.

Évidemment, TERMIUM n'est pas la seule banque de terminologie. Chacune des autres grandes banques en exploitation, notamment la BTQ (Office de la langue française), EURODICAUTOM (Communauté économique européenne), LEXIS (Bundesprachenamt), NORMATEM (AFNOR) et TEAM (Siemens) possède ses caractéristiques propres et répond à des besoins précis en fonction de sa clientèle.

Dans la première partie de cet article, nous décrirons les étapes à franchir dans la mise sur pied d'une banque de terminologie. Nous traiterons ensuite de l'organisation et de la gestion des données en nous inspirant de notre expérience personnelle avec TERMIUM. Dans la seconde partie, nous jetterons un regard sur des techniques et des matériels informatiques qui auront, à notre avis, une influence sur le développement et l'évolution des banques de même que sur le travail dans les organismes de terminologie.

I. CRÉATION D'UNE BANQUE DE TERMINOLOGIE

A) ÉTAPES

Plusieurs étapes³ jalonnent le processus de création d'une banque notamment : le lancement du projet, l'exécution d'une étude de faisabilité, l'élaboration de la conception générale, le choix du matériel et du logiciel, l'élaboration de la conception détaillée, la mise en œuvre, la mise en place et la mise au point. Nous décrirons chacune de ces étapes en donnant dans certains cas des exemples tirés du projet TERMIUM.

Tout au long du processus, il importe de respecter une méthode rigoureuse, définie en fonction de critères précis qui répondent à des objectifs quantitatifs et qualitatifs bien déterminés.

1. Lancement du projet

Le lancement du projet consiste à en définir les grandes lignes et à établir les mesures de contrôle appropriées.

Cette étape comprend plusieurs phases dont la plus importante est l'*identification des besoins* qui peut représenter beaucoup plus de travail qu'on peut se l'imaginer au départ. Voici les aspects qu'il importe de considérer :

a) l'*identification du groupe client* :

La clientèle de TERMIUM est formée d'un groupe multidisciplinaire rassemblant principalement des traducteurs, des terminologues, des aides-terminologues, des rédacteurs, des opérateurs, des commis à l'interrogation et des gestionnaires ;

b) la *détermination des besoins particuliers* de chaque groupe :

Ces besoins doivent être définis par les usagers eux-mêmes ;

c) la *comparaison et le regroupement des principaux besoins*

d) l'*établissement des priorités des besoins* :

Cet exercice est facile à réaliser lorsque le groupe client est homogène. Nous avons affaire à un groupe client dont les besoins variaient considérablement. Nous avons résolu le problème en divisant les besoins en deux catégories : nécessaires et souhaitables. Il faut noter que les points c) et d) consistent à évaluer des besoins qui entrent en quelque sorte en conflit, ce qui n'est pas toujours facile à régler.

La deuxième phase du lancement, facultative dans les cas où le nouveau système ne vient pas remplacer un système déjà existant, consiste à *identifier les points forts et les points faibles de l'ancien système*. Dans le cas de Termium II, nous avons identifié les faiblesses suivantes : manque de documentation, aucune possibilité d'expansion et de raffinement des caractéristiques, utilisation trop poussée des données codées, système pas suffisamment convivial, présentation quelque peu inélégante et validation presque inexistante. Du côté des points forts, mentionnons simplicité d'utilisation, disponibilité et rapidité.

La troisième phase repose sur l'*identification des contraintes* susceptibles d'influencer le déroulement du projet, c'est-à-dire le temps ainsi que les ressources humaines et budgétaires. La plupart des groupes chargés du développement de systèmes ont tendance à en sous-estimer certains aspects. Les points suivants méritent une attention particulière :

- a) l'évaluation du temps d'exécution (on est toujours trop optimiste) ;
- b) les prévisions budgétaires (souvent insuffisantes) ;

- c) la dépendance du projet par rapport aux facteurs extérieurs (consultants, fournisseurs, etc.) ;
- d) la résistance au changement (généralement très élevée chez certains groupes d'utilisateurs) ;
- e) la définition des besoins (fréquemment mal cernés) ;
- f) l'évaluation de la puissance du système (souvent sous-estimée par rapport aux besoins).

La quatrième phase a trait à l'*établissement des prévisions et des échéances relatives à l'étude de faisabilité*. Idéalement, cette étude de faisabilité devrait être effectuée par des experts et comporter une étude de marché.

L'avant-dernière phase porte sur la sélection *du groupe de travail* qui doit être formé d'analystes (experts en informatique) et de représentants des utilisateurs. Le groupe de travail TERMIUM III, par exemple, était formé de deux chefs de projet (l'un pour l'aspect terminologique, l'autre pour l'aspect informatique), d'un représentant pour chaque groupe d'utilisateurs (traducteurs, terminologues, documentalistes, personnel de la saisie des données et du service à la clientèle) et enfin, de deux analystes et d'un programmeur. Le groupe de travail devrait, en principe, être formé aux techniques de développement de systèmes.

La dernière phase du lancement du projet consiste à *identifier les ressources* qui seront *nécessaires* ainsi que les responsabilités de chaque intervenant et les échéanciers.

2. Étude de faisabilité

Le lancement du projet est suivi par l'exécution de l'*étude de faisabilité* qui permet de déterminer la solution la plus appropriée (logiciel et matériel) au problème donné (besoins et objectifs) en fonction de la compatibilité, de la rentabilité et de la conformité technique du système avec l'organisation.

3. Conception générale

L'élaboration de la conception générale est nécessaire pour expliquer les résultats de l'étude de faisabilité. Elle sert aussi à produire une description fonctionnelle du cheminement des données et à élaborer une conception d'ensemble du système nécessaire pour obtenir l'autorisation de poursuivre. Cette étape comprend de multiples facettes qu'il serait trop fastidieux d'expliquer ici, mais que l'on peut résumer en deux grands points, soit la formation de l'équipe de projet (dont les représentants ont été identifiés à la cinquième phase du lancement de projet) et l'examen détaillé de certains éléments comme :

- ◆ les exigences relatives aux fichiers ou aux bases de données (sans oublier la conversion, le cas échéant),
- ◆ les exigences techniques,
- ◆ les critères de mesure de la performance d'un système et
- ◆ les besoins de formation.

Il existe un aspect de la conception générale que l'on a tendance à oublier : il s'agit de la présentation, aux utilisateurs, des besoins tels qu'ils ont été perçus ainsi que des contraintes existantes, tant sur le plan technique que sur le plan des ressources. Cette présentation permet d'une part aux utilisateurs de contribuer à l'élaboration du nouveau système, réduit leur résistance au changement et d'autre part décourage les attentes non réalistes.

4. Choix du matériel et du logiciel

Soulignons que le *choix du matériel et du logiciel* doit précéder l'élaboration d'une conception détaillée. Un avertissement est de mise ici en ce qui concerne le matériel : les groupes de travail se laissent souvent attirer par les dernières nouveautés et techniques. Nous estimons qu'il importe de procéder avec la plus grande prudence puisque, à la tâ-

che déjà fort complexe et délicate de développer et de mettre en place un système, va s'ajouter celle d'expérimenter les nouveaux logiciels et matériels et les options qui s'y rattachent.

Que l'on se dote d'une banque de terminologie ou que l'on remplace celle qui existe déjà, deux solutions sont offertes en ce qui concerne les programmes : le développement d'un nouveau logiciel ou l'achat d'un progiciel. En théorie, il est recommandé de choisir le logiciel avant le matériel. En pratique, le choix est souvent restreint par la nécessité d'amortir l'équipement disponible. Le Secrétariat d'État du Canada n'a pas connu ce problème, la majeure partie de l'équipement utilisé par Termium II ayant été loué.

L'espace dont nous disposons ne nous permet pas d'analyser la gamme des possibilités qui existent aujourd'hui sur le marché à cet égard et qui ne cessent de se multiplier. Nous nous contenterons donc de parler de la solution que nous avons adoptée, au moment de la conception de TERMIUM III, après avoir examiné les raisons qui nous ont guidés dans notre choix.

L'avantage du logiciel maison est, évidemment, d'être bâti sur mesure en fonction des spécifications des utilisateurs et avec leur collaboration. L'inconvénient majeur, cependant, est que le développement d'un nouveau logiciel entraîne de longs délais et des dépenses importantes. En revanche, le recours au progiciel permet d'implanter un système rapidement et à des coûts proportionnellement plus bas. Un autre avantage de ce genre de programme réside, de façon générale, dans sa fiabilité puisqu'il est implanté sur plusieurs sites. Toute chose comportant aussi des inconvénients, il faut souligner que, outre le fait qu'il est toujours possible de se tromper dans le choix d'un progiciel, un tel programme est créé pour répondre à des besoins standards et non pas à ceux, très spécifiques, d'un acheteur donné. Les modifications qu'il faut souvent y apporter viennent compliquer la maintenance du système.

C'est un progiciel que le Secrétariat d'État a adopté il y a quatre ans, car il était urgent de changer le logiciel, et aussi parce que son équipe était convaincue qu'il existait sur le marché des logiciels adéquats. Le ministère était à la recherche d'un progiciel bilingue, simple à utiliser et à modifier, et déjà implanté sur plusieurs sites ; le système recherché devait également comprendre une interface vidéotex TÉLIDON/NAPLPS⁴ afin de permettre, entre autres choses, l'emploi de graphiques, d'un jeu de caractères multilingues et d'une structure de menus.

À la suite d'un appel d'offres, le progiciel BASIS, créé par la société *Battelle Development Corporation* (Columbus, Ohio, É.-U.), fut acheté et modifié en fonction des spécifications de TERMIUM. Après plus d'un an d'exploitation, TERMIUM s'est avéré un outil simple à modifier et à utiliser. Toutefois, il nous est apparu que les modules de mise à jour sont instables et que celui qui sert au formatage des données est trop lent. Des programmes spéciaux en langage PASCAL ont dû être rédigés pour pallier ce dernier problème. L'instabilité des modules de mise à jour a été corrigée dans les dernières versions du progiciel ; toutefois, ces dernières n'ont pas encore été implantées sur TERMIUM en raison, justement, de ce que ce système n'utilise pas le progiciel BASIS standard. Ceci est d'ailleurs un aspect important dont il faut tenir compte puisque tout écart par rapport au progiciel standard peut entraîner des délais considérables dans la livraison des versions mises à jour du logiciel type.

La version de BASIS qui a été retenue fonctionne sur ordinateurs VAX de la société *Digital Equipment Corporation*. Deux exemplaires de ces superminiordinateurs, de modèle VAX 11/785, remplacent ainsi le gros ordinateur CYBER de la société CDC, précédemment utilisé pour Termium II. De plus, le parc complet de terminaux a été renouvelé : des terminaux à écran programmables couplés à des imprimantes asservies ont

pris la relève des vieux téléimprimeurs. Cette dernière mesure était nécessaire en raison de l'adoption d'une structure de consultation par menus et du codage des données selon le format NAPLPS.

5. Conception détaillée

Après le choix du matériel et du logiciel vient l'étape de la *conception détaillée* qui consiste à élaborer, à partir des prescriptions fonctionnelles produites lors de la conception générale, un tableau détaillé du système dont chaque élément est examiné et approuvé et qui correspond, idéalement, à un système figé. Brièvement, il s'agit d'abord de mettre au point les procédures de travail, le cheminement de l'information (diagrammes fonctionnels) et la façon de présenter les documents d'entrée et de sortie. Il faut ensuite définir les données et concevoir les fichiers ou les bases de données (ou les deux) ainsi que les procédures de contrôle du système. Ces activités sont suivies de la mise au point de la logique de traitement et de la liste des exigences techniques, puis de la définition des contrôles de validité des entrées, des fichiers et des sorties. Les trois derniers aspects de la conception détaillée sont d'envergure plus globale et constituent en l'établissement de la liste des manuels nécessaires à la formation, à l'exploitation et à la maintenance du système, en l'élaboration des plans détaillés de l'implantation et de la mise en place et, finalement, en la présentation aux gestionnaires pour leur permettre d'examiner et d'approuver cette partie du projet.

Encore une fois, il est préférable de présenter la conception détaillée au groupe client de façon à s'assurer à la fois de sa collaboration et de la justesse des perceptions initiales en ce qui concerne les besoins.

Il convient de mentionner qu'à cette étape du projet, il est fortement souhaitable de reporter toute modification au système jusqu'à ce qu'il devienne opérationnel. Trop souvent, on perd un temps précieux à tenter d'apporter des changements de dernière minute, souvent relativement accessoires. Ce qui importe, à ce stade du projet, c'est de s'assurer que le groupe technique chargé de développer le système en comprenne parfaitement les fonctions et l'application. Il est important que les experts chargés du développement et les représentants des usagers travaillent en très étroite collaboration et sur un pied d'égalité.

6. Mise en œuvre

L'étape de *mise en œuvre* suit la conception détaillée. Compte tenu des prescriptions détaillées du système, la mise en œuvre comprendra la création de formules, de manuels, de programmes et de la documentation relative à la formation, lesquels sont nécessaires à la mise sur pied d'un système opérationnel. Le déroulement de cette étape est le suivant :

1. établir les normes qui garantiront la qualité à respecter au cours de l'implantation ;
 2. élaborer des rapports détaillés programme-modules ;
 3. établir la logique du programme ;
 4. construire le dictionnaire des données ;
 5. coder le logiciel d'application ;
 6. produire tout le matériel imprimé et la documentation à l'appui du système et
 7. vérifier toutes les conditions de la logique de chaque programme et module.
- On ne saurait trop insister sur l'importance de la documentation.

Au risque de commettre une lapalissade, il faut souligner que si, pour une raison ou une autre, on prévoit des difficultés sérieuses telles que des retards dans le développement du logiciel, des malentendus relatifs aux spécifications ou à la puissance du matériel, les responsables devront accepter de repousser les échéances et informer les usagers de la situation. Au risque d'en décourager plusieurs, il faut comprendre que, lors du dé-

veloppement des systèmes complexes, les retards sont la règle. Il faut donc au départ prévoir une très bonne marge de sécurité. En cas de retards, il faut plutôt faire contre mauvaise fortune bon cœur et parler de « rajustements » par rapport au programme initial. En d'autres termes, il faut prévoir une certaine souplesse pour l'étape de mise en œuvre.

À titre d'exemple, certains retards de livraison dans le cadre du développement de TERMIUM ont bousculé des échéanciers par ailleurs réalistes, ce qui a provoqué une très grande frustration pour le groupe de travail, les utilisateurs et par la suite, chez les gestionnaires.

7. Mise en place

Avec la *mise en place*, le système passera de l'élaboration à l'exploitation et l'entretien en passant, s'il le faut, par un projet pilote. Cette avant-dernière étape du développement d'un système comprend sept volets :

1. la conversion, le contrôle avant traitement et la vérification des fichiers existants :

Ce volet est fondamentalement centré sur l'utilisateur puisque celui-ci est le mieux placé pour déterminer les éléments de données qui doivent être conservés ou écartés.

2. la formation des utilisateurs et du personnel informatique à l'exploitation du système.

À la lumière de notre expérience, il appert que la formation d'un groupe restreint d'utilisateurs est amplement suffisante puisqu'il s'agit finalement d'un projet pilote. À ce stade-ci encore, on doit faire preuve de souplesse pour être en mesure de laisser tomber le système expérimental dans l'éventualité où il ne répond pas aux exigences de l'utilisateur.

3. la distribution des manuels et des instructions appropriées.

Il demeure que, dans la réalité, il est préférable de ne distribuer qu'une documentation sommaire, ce qui offre une plus grande marge de manœuvre.

4. l'établissement des procédures de contrôle.

5. la réalisation des opérations parallèles ou pilotes, au besoin.

6. l'examen et l'approbation de tous les résultats de conversions et de mise en place.

Lorsque le temps et les ressources le permettent, le « nettoyage » des conversions devrait se faire sur le champ afin de conserver toutes les données nécessaires et de partir du bon pied. Toute tâche remise à plus tard requiert deux fois plus d'efforts en raison du cumul des activités courantes et de la diminution des priorités.

7. le passage du système à l'état d'exploitation et d'entretien.

La mise en opération d'un système partiel comporte deux conséquences qu'il faut prendre en considération. Elle entraîne le cumul des fonctions suivantes : terminer le développement du nouveau système et en assurer les opérations normales d'entretien et de maintenance. De plus, le climat peut changer et il devient plus difficile d'obtenir les fonds et les ressources nécessaires à l'achèvement du projet.

8. Mise au point

La création d'une banque se termine enfin par la *mise au point* qui comprend les activités d'entretien et d'examen périodique nécessaires au bon fonctionnement du système et à l'évaluation de sa performance en fonction des objectifs et des critères fixés à la conception. Il s'agit ici d'établir les répercussions de toute exigence d'entretien sur les utilisateurs et les éléments du système, d'apporter les modifications pertinentes, d'en faire l'essai et de mettre à jour la documentation en conséquence.

Cette dernière étape constitue une activité continue puisqu'il est essentiel que le système évolue au rythme des techniques nouvelles. Elle permet de passer plus facilement de l'ancien système au nouveau.

Ce qui précède illustre de façon sommaire toutes les étapes, à notre avis nécessaires, du développement d'une banque de terminologie. De tout cela, il se dégage ce qui suit :

- ◆ La planification est l'élément primordial à tout projet et la mise sur pied d'un système n'y fait pas exception.
- ◆ Il faut toujours être prêt à réviser la situation et à remédier à tout problème sur-le-champ.
- ◆ Tout bon système repose sur une documentation adéquate.
- ◆ La participation des usagers est le gage de garantie que leurs besoins seront respectés et que, en retour, la banque bénéficiera de l'acceptation intégrale de la part du groupe client.

Maintenant qu'ont été esquissées les différentes étapes, nous croyons bon de discuter de deux aspects qui touchent à la fois le volet informatique et le volet terminologique d'une banque, c'est-à-dire l'organisation et la gestion des données.

B) ORGANISATION DES DONNÉES

L'une des tâches les plus délicates dans la création d'une banque est la définition des bases de données. Alors que les petites banques ne requièrent généralement qu'un fichier, les banques importantes en demandent plusieurs, des données de plusieurs types y étant versées. La création de bases de données distinctes permet de respecter la nature et l'homogénéité des informations. Il faut noter que, dans le cas de fichiers importants, cette opération est délicate ; si une restructuration se révélait nécessaire, le changement peut être long et coûteux. Il est donc fortement recommandé, au moment de la mise en place, de créer des fichiers échantillons et de procéder à des essais avant de transférer l'ensemble des données sur le nouveau système.

À titre d'exemple, prenons à nouveau TERMIUM...

TERMIUM comprend 35 bases de données : la base principale (base de données linguistiques) et 34 bases de données secondaires. Ces dernières, dont certaines peuvent être consultées par les utilisateurs, servent principalement aux usages suivants :

- ◆ à la validation et au décodage des données au moment du chargement dans la base principale ;
- ◆ à titre de référence ;
- ◆ au stockage de l'information de gestion ;
- ◆ au contrôle de l'accès des utilisateurs au système.

Il s'avère, après plus d'un an d'utilisation, que le recours à un si grand nombre de bases ralentit quelque peu l'interrogation et surtout la mise à jour ; à cet effet, TERMIUM doit être réservé plusieurs heures par jour au chargement des nouvelles données. En contrepartie, la validation évite le versement de données techniquement incorrectes. Il ne faut pas hésiter à revoir le processus de validation après quelques mois d'exploitation et à en vérifier la pertinence.

Après avoir déterminé les sous-ensembles de la banque, il faut définir leurs éléments, c'est-à-dire les champs de la fiche et leurs caractéristiques (noms des champs, modes d'indexation et de sortie, etc.).

Il importe d'employer des noms de champs clairs et uniformes ; si des données sont communes à plusieurs bases de données, il faut adopter le même nom partout. On verra à éviter les mots abscons qui réclament des efforts de mémoire et qui entraînent souvent des fautes de frappe.

Chaque champ à indexer doit être étudié attentivement. Il peut être intéressant d'indexer certains champs sous des rubriques communes ou, au contraire, individuellement. Il est également possible d'employer des libellés distincts pour la recherche et la sortie (si l'on veut, par exemple, restreindre l'accès de certains champs). Il faut décider

si l'on aura recours à des listes de mots vides (par exemple, les articles) ou à des préférences (par exemple, les variantes orthographiques) puis en définir les valeurs. Il existe aussi la possibilité de créer des clés de recherche ou de sortie complexes, c'est-à-dire des clés qui regroupent plusieurs champs. Ainsi, d'une part, la recherche d'un critère peut être effectuée dans différents champs indexés séparément et, d'autre part, l'utilisateur peut, par la frappe d'un seul mot, déclencher l'affichage de plusieurs champs. Il est facile de modifier ces clés et d'en créer de nouvelles au besoin, puisque cela ne touche en rien l'indexation.

L'organisation des données doit faire l'objet d'une attention particulière, surtout dans le cas de fichiers importants ; une erreur d'indexation peut se révéler, en effet, désastreuse. Il importe donc de procéder avec précaution, les informaticiens travaillant en étroite collaboration avec les représentants des usagers.

Du côté de la consultation, les banques de terminologie actuelles sont accessibles sous l'une des deux formes suivantes : la façon traditionnelle qui requiert l'emploi d'un langage d'interrogation, et la forme plus conversationnelle d'une structure de menus où l'utilisateur opère une sélection dans les options qui lui sont présentées.

La première méthode d'accès permet généralement une utilisation plus souple d'une banque, adaptée à chaque séance, et implique que les usagers ont reçu au préalable une formation intensive et complète. L'utilisateur peut, par exemple, effectuer des interrogations avec opérateurs booléens, peut n'afficher que certains champs ou sauvegarder le résultat de recherches.

La seconde méthode permet, dans certains cas, de se passer de formation ou de la limiter au minimum. Les menus guident l'utilisateur au moment de la consultation et, s'il existe un module de dépannage incorporé au système, lui permettent même de se passer de manuel de référence. Le menu contrôle également l'emploi de la banque par l'utilisateur, ce dernier devant généralement se contenter des options affichées. Cependant, il se produit que, dans le cas d'utilisateurs réguliers, ces derniers se lassent de la répétition des menus ; c'est pourquoi plusieurs logiciels offrent la possibilité de les court-circuiter.

TERMIUM et NORMATERM sont des exemples de cette dernière méthode de consultation. TERMIUM appartient même aux deux types de banques décrites ci-dessus ; en effet, il est possible d'avoir accès aux données non seulement par les menus, mais aussi, si la permission est accordée à l'utilisateur, d'employer le langage d'interrogation standard du logiciel BASIS.

C) GESTION DES DONNÉES

Bien que ce ne soit pas à proprement parler une préoccupation directe de la majorité des usagers d'un système, l'aspect administratif joue néanmoins un rôle capital pour les « décideurs ».

Pour illustrer les différentes facettes de la gestion des données, nous nous servirons de la structure que nous avons adoptée pour TERMIUM, qui s'appuie principalement sur un système d'information de gestion dont on poursuit le développement. Soulignons cependant que la manière de gérer des données terminologiques varie d'un système à l'autre et que celle que nous décrivons ici n'est donnée qu'à titre d'exemple.

La base de données linguistiques de TERMIUM contient environ 800 000 fiches, soit près de 1 700 000 termes. Par ailleurs, les bases de données secondaires comportent environ 100 000 autres enregistrements. On estime que le total des mises à jour pour l'ensemble des bases de données se situe entre 150 000 et 200 000 transactions par an. La gestion d'un tel ensemble nécessite, on s'en doute, des ressources considérables.

En raison de l'ampleur d'un tel système, il est essentiel d'exercer un contrôle strict tant du processus de mise à jour (aspect quantitatif) que du contenu (aspect qualitatif). Seule une gestion bien informée permet de maintenir l'efficacité et l'efficience des opéra-

tions et de fixer ou de revoir les objectifs de planification en fonction de la demande (besoins des usagers).

TERMIUM a donc recours, pour ce faire, à un système d'information de gestion (SIG), c'est-à-dire un système dont l'objectif est d'assurer la gestion de la ressource informationnelle. On y distingue les éléments suivants : les intrants, les processus de traitement, les extrants, les mécanismes de contrôle et les boucles de rétroaction. Le SIG a pour fonction d'emmagasiner et de produire les données de gestion et d'assister le décideur dans sa prise de décision.

Les données de nature diachronique sont stockées dans deux bases de données.

La première, dont la mise à jour est totalement automatique, reçoit les statistiques d'utilisation du système. Elle permettra de connaître, par exemple, le nombre de termes interrogés par groupe d'utilisateurs. La seconde, dont la mise à jour est manuelle, contient l'information relative aux opérations de modification de la base de données linguistiques. Elle permet, par exemple, d'obtenir le nombre de fiches modifiées durant une période donnée.

Les données de nature synchronique sont obtenues au moyen de programmes soumis régulièrement, qui procurent sous différents angles des vues instantanées de la base de données linguistiques. Par exemple, un de ces rapports présente le nombre total de fiches ventilé selon les grandes subdivisions (fonds et fichiers) alors qu'un second offre une ventilation des occurrences par domaines.

Les données statistiques dont nous disposons sont regroupées sous les quatre catégories suivantes :

◆ *Les productions*

Elles permettent d'évaluer et de comparer les extrants en regard des objectifs de planification ; ces données fournissent l'information nécessaire au contrôle et aux ajustements des objectifs de planification à la lumière de l'expérience acquise (établissement d'objectifs plus réalistes).

◆ *Les arriérés*

Ils permettent de prendre des mesures à court terme pour corriger les situations conflictuelles (les goulots ou les accumulations le long de la chaîne d'alimentation) ; ces données sont établies pour chaque étape afin de réaffecter les ressources nécessaires au redressement des situations inacceptables.

◆ *Le rendement*

Il permet d'évaluer la quantité et la qualité de travail accompli et de fixer les objectifs de travail des employés à tous les niveaux de la hiérarchie en fonction des grands objectifs identifiés précédemment.

◆ *Les rapports d'analyse*

Ils permettent d'orienter les efforts à moyen et à long termes ; ces rapports sont faits *a posteriori* et donnent une vue d'ensemble dans une perspective macroscopique.

On peut noter que nous avons ainsi des rapports sur les différentes étapes du processus : en amont, les intrants ; en aval, les extrants ; et au centre, les traitements comme tels. Les rapports d'arriérés, de production, de rendement et d'analyse *a posteriori* sont ensuite évalués sous l'éclairage des différents niveaux de la hiérarchie (haute direction, cadres intermédiaires et opérations) pour faire le lien avec les objectifs de planification.

II. PERSPECTIVES D'AVENIR

Notre expérience avec TERMIUM nous a permis de constater à quel point la terminotique, tout au moins en ce qui a trait aux grandes banques de données terminologiques, était tributaire de l'informatique la plus traditionnelle. En effet, les banques ne servent aujourd'hui essentiellement qu'au stockage et à la consultation au moyen de

langages d'interrogation plus ou moins évolués. De plus, la consultation se fait au moyen de terminaux dont c'est la fonction unique et qui, habituellement, ne donnent accès qu'à une seule banque.

Cette constatation nous pousse à conclure qu'il est nécessaire, pour des raisons autant économiques que professionnelles, de se tourner rapidement vers de nouvelles techniques logicielles et matérielles d'information et d'améliorer sans tarder la coordination des efforts en terminologie.

Nous faisons allusion, comme on l'aura deviné, à la bureautisation de l'espace de travail (tâche possible dès maintenant) ainsi qu'aux récents développements en logiciels et matériels informatiques (qui sont disponibles ou qui le seront d'ici quelques années).

A) BUREAUTIQUE

La bureautique constitue le premier point que nous voulons aborder. Cette application de l'informatique vise à automatiser le travail de bureau et, plus particulièrement, le traitement des textes. Elle fait le pont entre l'équipement de bureau classique (photocopieurs, téléphones, machines à écrire et à calculer, etc.) et les moyens informatiques traditionnels (archivage et recherche de l'information, transfert électronique, etc.) en partie au moyen des télécommunications. Bureautiser signifie essentiellement informatiser l'espace de travail. Les burovisseurs, c'est-à-dire les postes de travail informatisés, remplacent, entre autres choses, les machines à écrire sur les bureaux. Les burovisseurs consistent soit en terminaux reliés à un réseau informatique, soit en microordinateurs reliés en réseau local ou intégrés à un réseau informatique.

Chaque employé a ainsi accès à toute une gamme de fonctions (par exemple traitement et correction de texte, messagerie électronique, aide à la rédaction, gestion de mémoires individuelles, gestion du temps au moyen d'agendas électroniques, archivage électronique, etc.) jusque-là réservés à des spécialistes.

Les traducteurs, clients privilégiés des banques de terminologie, et les terminologues forment des cibles idéales de la bureautisation. En effet, ils consacrent l'essentiel de leurs activités à la manipulation de textes, problème auquel s'adresse d'abord la bureautique. Examinons brièvement quelques aspects qui pourraient influencer sur les traducteurs et les terminologues, en tâchant de ne pas perdre de vue les banques de terminologie.

Ainsi, les burovisseurs peuvent servir aux traducteurs à bien des fonctions. Ils leur permettent d'effectuer, au moyen de programmes de traitement de texte, leur traduction sur écran. Le texte peut alors être rapidement modifié et reproduit, ce qui élimine les frappes successives. De plus, les traducteurs peuvent définir des formules standards de lettres ou de notes auxquelles ils auront recours si besoin est. Le burovisseur peut également être employé pour la consultation de banques de terminologie, soit avant la traduction, soit en parallèle à cette opération.

Certains logiciels permettent notamment d'obtenir, après analyse du texte à traduire et consultation d'un dictionnaire automatique, un lexique de travail à utiliser en parallèle à la traduction, ainsi qu'une liste des mots qui n'apparaissent pas dans le dictionnaire consulté. Avant de commencer la traduction, le traducteur peut donc chercher les équivalents des mots absents et mettre à jour le lexique de travail ainsi, éventuellement, que le dictionnaire permanent.

Le texte traduit sera ensuite soumis à des logiciels de correction orthographique qui repéreront les coquilles et les fautes élémentaires. Depuis peu, des programmes de correction grammaticale ou stylistique ont fait leur apparition ; toutefois, il faut reconnaître que ces derniers sont encore rudimentaires et ne présentent pas un grand intérêt pour le moment.

Parmi les autres fonctions offertes qui viennent à l'esprit, mentionnons l'archivage électronique des textes et de leur traduction, la création de lexiques ou de vocabulaires maison avec mise à jour ultra-rapide.

Tout comme pour les traducteurs, la bureautique peut constituer un apport important au travail des terminologues. Ces derniers font beaucoup de rédaction (notes, articles, etc.). L'informatique permet de rendre ce travail plus facile. Ainsi, aux programmes de traitement et de correction de texte déjà mentionnés, s'ajoutent les « processeurs d'idées » (en anglais « *idea processors* » ou « *outline processors* ») ou programmes d'aide à la rédaction qui ont pour but d'informatiser les différentes étapes de la rédaction et plus particulièrement la préparation du canevas. De plus, ces programmes offrent certaines fonctions de traitement de texte ou créent des fichiers que l'on peut modifier au moyen d'autres logiciels.

La consignation des données sur différentes formules (fiches de dépouillement, fiches de terminologie ou formules de demandes de renseignements) occupe une grande partie du temps de travail du terminologue. Or, cette tâche peut facilement être informatisée au moyen de programmes de gestion de fichiers et de formules. Il faut toutefois prendre conscience que la consignation prend alors un caractère nettement plus procédurier.

Le remplacement des fichiers de dépouillement manuels par des fichiers électroniques assure ainsi un accès accru aux données (par exemple, la recherche dans les contextes devient facile). Cela élimine également l'emploi de fiches renvois : c'est l'ordinateur qui effectue les renvois nécessaires. Le partage des fichiers de travail entre terminologues devient possible. Les fichiers demeurent actifs et peuvent être restructurés au besoin pour faire face à des situations nouvelles.

L'informatisation de la fiche de terminologie accélère son chargement. En effet, elle en élimine les frappes successives dues à la correction d'erreurs, à des modifications de fond ou à la saisie par des mécanographes pour chargement dans la banque de terminologie (comme c'est le cas de TERMIUM). Seules les parties qui ont besoin d'être retouchées le sont effectivement et la fiche peut être envoyée électroniquement au service responsable de la mise à jour. On peut même envisager des programmes assez évolués pour reporter automatiquement les valeurs communes sur tout un lot de fiches.

La plupart des grands organismes de terminologie offrent un service de renseignements terminologiques par téléphone (service SVP). En règle générale, cette fonction est assurée à partir d'un bureau spécial muni d'un terminal, d'une table de travail et d'ouvrages de référence. La bureautisation des services de terminologie permettrait d'éliminer ces postes, les appels pouvant être automatiquement acheminés aux terminologues de quart. Il serait également possible d'informatiser les formules de demandes ce qui permettrait, par exemple, le report automatique des coordonnées des clients et de consulter en parallèle des banques de terminologie. Le système pourrait aussi fournir rapidement des données statistiques sur ce service.

Évidemment, l'équipement bureautique permet la consultation de multiples banques de données, soit terminologiques, soit documentaires. Qui plus est, il offre l'avantage de permettre l'extraction de fichiers de travail pour traitement ultérieur : par exemple, pour la préparation de lexiques ou de vocabulaires. Le fichier de manœuvre ainsi créé peut être fusionné avec des fichiers de dépouillement et manipulé au moyen de différents logiciels (tri, traitement de texte, etc.).

Ce ne sont que quelques aperçus des possibilités offertes par la bureautique. Bien sûr, la plupart des logiciels mentionnés, sinon tous, sont disponibles sur les équipements informatiques traditionnels.

Notons au passage que la bureautique cherche à intégrer et à partager l'information ainsi qu'à accélérer les échanges alors que l'informatique classique au contraire a plutôt tendance à compartimenter. Dans ce sens la bureautique est plus un modèle d'organisation qu'une technologie.

B) LOGICIEL

Nous désirons ici souligner les progrès réalisés dans le domaine de l'intelligence artificielle, plus particulièrement les travaux sur les systèmes experts et le traitement automatisé des langages naturels qui sont appelés à révolutionner les banques de terminologie.

Dans un proche avenir, on doit s'attendre à ce qu'on puisse interroger ces banques au moyen des langages naturels (anglais, français, etc.). La consultation deviendra un véritable dialogue entre l'utilisateur et le système. Ce dernier ne cherchera plus seulement la chaîne de caractères tapée par l'utilisateur, mais aussi les variantes et les formes proches (par exemple, le pluriel). Mentionnons aussi que certains logiciels existants n'indexent plus la forme orthographique du mot, mais sa forme phonétique... Et pourquoi pas ? De tels systèmes suggéreront des réponses, si l'équivalent exact n'a pu être trouvé.

On peut également entrevoir l'intérêt que soulèvent les travaux actuels en analyse sémantique, par ordinateur, des textes étendus. Les programmes développés au cours de ces recherches transforment des textes spécialisés en tableaux, listes, résumés, etc. Il s'agit d'un premier pas vers le dépouillement et l'analyse automatiques des données terminologiques.

C) MATÉRIEL

Les travaux sur l'intelligence artificielle touchent non seulement le logiciel, mais aussi le matériel. Il se peut que, d'ici quelques années, de nouveaux matériels rendent le bloc-notes périmé. Nous pensons aux appareils de saisie optique ou de reconnaissance de la parole qui commencent à faire leur apparition dans l'industrie. Ainsi certaines sociétés ne font plus d'inventaire avec un bloc de formules et un crayon, mais emploient plutôt un micro-ordinateur portable avec dispositif de reconnaissance de la parole. L'employé dicte la liste des pièces à l'ordinateur qui les stocke sous forme numérique et qui dresse ensuite l'inventaire. On peut imaginer que cet équipement servira un jour aux terminologues dans le dépouillement de textes. Cela épargnerait beaucoup de temps actuellement consacré à la copie sur papier de l'information.

Ces derniers mois, un nouveau type de matériel de stockage a commencé à faire une percée : les disques optiques. Leur capacité de stockage surpasse de loin celle des disquettes et des disques rigides. Ils sont, de plus, quasiment inaltérables et d'une très grande flexibilité d'emploi. Leur inconvénient majeur toutefois réside dans ce qu'on ne peut, à l'heure actuelle, que les lire. Pour obtenir une version mise à jour de leur contenu, il faut donc acheter de nouveaux disques.

Il existe deux types de ces disques optiques : les vidéodisques (30 cm, 1 gigaoctet et plus) et les disques compacts ou CD-ROM (12 cm, 550 mégaoctets). Les premiers sont surtout destinés à des applications spécialisées qui nécessitent le stockage d'une quantité énorme d'information (le cas des grandes bibliothèques de référence) ; les seconds visent des marchés plus ouverts, celui des organismes et des individus ayant à leur disposition des micro-ordinateurs.

Nous effectuons actuellement des essais sur CD-ROM. Nous comptons ainsi pouvoir verser sur un disque ou deux une copie de TERMIUM et la diffuser sous cette forme chez certains clients. Cela assurerait à notre réseau une marge de manœuvre importante.

Ainsi que l'on peut en juger, l'informatique est appelée à jouer un rôle de plus en plus important dans le travail des terminologues. Ceci ne constitue rien d'étonnant : cela ne fait que suivre le phénomène d'informatisation qui bouleverse nos sociétés depuis quelques années.

La clientèle des banques de données terminologiques va changer. L'intégration croissante de ces systèmes aux réseaux de télécommunications et le développement de la télématique (et surtout du vidéotex) auront pour effet de donner accès à un public plus large, moins spécialisé. Les traducteurs demeureront encore longtemps les principaux clients des banques de terminologie, mais l'ouverture à un nouveau public influencera la nature et l'évolution de ces dernières.

Nous verrons au cours des années à venir la création ou le développement d'un grand nombre de petites banques spécialisées (sur micro-ordinateurs notamment) coexistant avec quelques superbanques. Cela exigera des efforts particuliers de coordination et de collaboration, tant techniques que terminologiques, entre organismes afin d'éviter le doublement des travaux. Il faudrait ainsi assurer un maximum de compatibilité des données afin de faciliter les échanges. L'accès interbanques devrait être encouragé. Plus particulièrement, on devrait s'attaquer à la mise au point d'un protocole d'accès standard ; de cette façon, les usagers ne devraient recevoir qu'une seule formation. Par ailleurs, les domaines de responsabilité de chaque organisme devraient être soigneusement définis et des mécanismes de collaboration mis au point afin que chacun puisse profiter des aires de compétence des autres services ; de cette façon, on évitera que des recherches entreprises chez l'un soient refaites chez l'autre.

C'est ainsi que nous voyons l'avenir de l'informatique en terminologie. Cette science ne sera plus restreinte aux banques de terminologie traditionnelles, mais s'intégrera de façon plus étroite à tout le processus de recherche terminologique.

Notes

1. La Banque de terminologie de l'Université de Montréal.
2. La Banque de terminologie du gouvernement canadien.
3. Il faut noter que la terminologie employée dans ce texte est celle en usage au sein du gouvernement canadien.
4. *North American Presentation Level Protocol Syntax.*