

# Développement et applications de la base de données terminologiques de Matra Marconi Space

## 1 Introduction

L'équipe d'ingénierie documentaire et linguistique de Matra Marconi Space et du laboratoire Aramihs<sup>(1)</sup> est pluridisciplinaire à deux niveaux, par le statut des personnes qui travaillent (ingénieurs et chercheurs), et par la pluralité des compétences de ces personnes (linguistes, informaticiens, experts du domaine spatial et ergonomes).

Un travail de recueil et de traitement de la terminologie utilisée à Matra Marconi Space a été entrepris depuis deux ans pour constituer une base de données terminologiques bilingues (BDT). Il a été initialisé à l'occasion de la participation de l'équipe dans le projet Eurolang (projet Eureka). Ces travaux vont être approfondis dans le cadre du projet Eureka de mémoire d'entreprise (Mnemos) dont Matra Marconi Space est le coordinateur et pour le développement d'applications internes.

## 2 Besoins

En tant que maître d'œuvre de systèmes spatiaux, Matra Marconi Space, société internationale, doit maîtriser des quantités importantes d'informations le plus souvent véhiculées par des textes rédigés en

français et en anglais. Par exemple, le développement d'un satellite de télécommunication engendre la création et la gestion d'environ 50 000 documents qui sont transmis aux partenaires et clients, référencés et consultés tout au long de la durée de vie du produit spatial et modifiés par la prise en compte de nouvelles versions. Une utilisation correcte et appropriée des termes techniques réduit les ambiguïtés au sein d'une équipe et vis-à-vis des clients et sous-contractants. La maîtrise de l'utilisation du langage technique permet de gagner du temps et de l'argent, accroît la productivité des ingénieurs qui rédigent la documentation, et améliore ainsi le processus global de travail.

L'objectif de notre activité est d'améliorer la qualité du contenu de l'information et des traitements associés, d'optimiser le processus de traitement de l'information, et de contribuer à la mémoire technique d'entreprise. L'étendue de nos activités est ciblée sur les différents acteurs de l'entreprise et leur documentation associée afin de leur fournir des outils et des ressources utilisant les technologies basées sur l'ingénierie documentaire et linguistique.

Le développement d'une base de données terminologiques bilingues (français et anglais) du domaine spatial permet d'envisager un langage commun homogène entre les utilisateurs afin d'améliorer la communication et la compréhension de l'information, de former les ingénieurs sur le domaine et de capitaliser ainsi une des richesses de l'entreprise.

<sup>(1)</sup> NDLR : Aramihs = Action, recherche et application Matra/Irit en interface homme système.

### 3 Développement de la base de données terminologiques

#### 3.1 Modèle terminologique

Le modèle terminologique a été défini après une enquête dans la société afin de faire un état de l'existant, d'analyser et définir les besoins pour permettre l'intégration des données dans des systèmes plus complexes.

La principale caractéristique de ce modèle vient du fait qu'il distingue nettement trois types d'informations:

- L'entité appelée *terme* qui permet de stocker les informations concernant le terme en tant que mot, c'est-à-dire les informations sur son fonctionnement linguistique. Elle contient les informations suivantes: langue, variantes, pluriels irréguliers, forme U.S., nature, genre, nombre, forme abrégée, forme elliptique (attestées ou interdites), collocations, notes, synonymes et traduction (en anglais ou en français).
- L'entité *concept* qui stocke les informations concernant la sémantique désignée par le terme (définition et liens avec d'autres concepts). Elle contient les informations suivantes: métier, usage, recommandations, diachronie et notes.
- L'entité *lien terme-concept* qui donne des informations sur l'usage, c'est-à-dire le contexte dans lequel un terme désigne un concept. Elle contient les informations suivantes: identifiant, technique, définition, source, concepts reliés et notes.

#### 3.2 Méthode de constitution de la BDT

Nous avons mis au point une méthode originale de constitution de la BDT (sélection des termes et recueil des données) qui fait

intervenir les méthodes de la linguistique et celle de l'intelligence artificielle. L'étude linguistique vise à étudier des textes rédigés par les experts du domaine. Un volume suffisant de textes permet au linguiste-terminologue d'examiner un grand nombre de contextes et d'émettre des hypothèses sur les termes retenus mais aussi sur leur sens et les liens conceptuels. Un entretien avec l'expert permet de confirmer ou non les hypothèses émises par le linguiste et de remplir les champs restés vides. Chaque étape du recueil et du traitement des données est assistée par ordinateur: extraction des termes, sélection des contextes, saisie et vérification de la cohérence des données.

#### 3.3 Système de gestion de la terminologie

Nous avons développé un système de gestion de la terminologie qui permet de créer une fiche terminologique et de contrôler la cohérence entre les informations. Les fiches terminologiques sont stockées au format SGML afin de faciliter les traitements sur les données.

#### 3.4 Exemples<sup>(1)</sup>

L'utilisateur peut faire une requête en langage naturel sur la base afin d'afficher la fiche terminologique correspondant à sa requête.

L'utilisateur peut visualiser le réseau sémantique entre les informations du domaine à l'aide de l'interface graphique.

(1) Figure 1: Recherche terminologique. (Voir page 25).

Figure 2: Réseau sémantique. (Voir page 26).

### 4 Applications de la BDT

Actuellement la base de données terminologiques contient environ 2 000 concepts et informations terminologiques associées. Par son utilisation et intégration à d'autres environnements, elle constitue le noyau d'applications diverses:

#### 4.1 Consultation

La consultation des informations terminologiques contenues dans la base de données peut se faire directement à partir d'un terme, sigle ou définition. Grâce aux capacités hypermédia du système, l'utilisateur pourra naviguer d'une fiche terminologique à une autre en passant par les concepts reliés ou l'interface de visualisation du réseau sémantique. De plus, des images peuvent être associées aux concepts.

#### 4.2 Enseignement assisté par ordinateur

Une personne non initiée au domaine pourra se former en naviguant dans le réseau sémantique et en interrogeant la base. Ces informations terminologiques peuvent être intégrées dans un didacticiel pour former les ingénieurs à des systèmes complexes.

#### 4.3 Aide à la rédaction de documents techniques

Le rédacteur d'un document technique peut générer automatiquement la liste des sigles et le glossaire correspondant au document en cours de rédaction et aux informations contenues dans la base.

La base de données terminologiques peut être couplée à

Figure 1 : Recherche terminologique.

Matrimarconi/Projets/Frame Builder/Documents/DOC-DOCSTEP

Recherche

File Edit Format View Special Graphics Table Structure



**DOCSTEP**

Ref. : DOC-MA-3909  
Issue : 00  
Revision : 00  
Date : May 95  
Page : 5

**1 - GENERAL OVERVIEW OF THE SATELLITE**

The satellite - C 1 payload offers high EIRP performance on 20 active channels over 2 areas. The frequency of the 20 active channels ( 10 per polarisation) can be selected total of 26 installed ( 13 per polarisation). All channels are individually switchable to widebeam or the superbeam coverage. A single receive coverage area is defined for and BSS channels.

The following figures illustrates :

- the receive coverage areas and performances,
- the transmit Widebeam coverage areas,
- the transmit Superbeam coverage areas.

The receive antenna is a (dual gridded) fixed antenna with a 1.2 m aperture diameter on the top floor. This antenna includes a diplexer which separates the 14 GHz and the 18 GHz. The transmit coverages are provided by two deployable reflector antennas with a 2.3m. These antennas use shaped reflector technology and a timing mechanism to optimize the pointing performance after in orbit calibration of the biases.

The repeater block diagram is shown in figure 2. 1/1. It is a single down conversion active with separate chains for FSS channels and BSS channels up to the channelized sections, which are common to FSS and BSS frequency bands.

The CLE (the ADCS computer) also manages the ADCS sensors and **actuators** through its own data bus : the MACS which is fully dedicated to the attitude control function.

**CRITERES DE SELECTION**

Terme :

Langue :  Métier :

Définition :

---

**RESULTATS DE LA RECHERCHE** nbre de termes trouvés

---

**TERMES TROUVES**

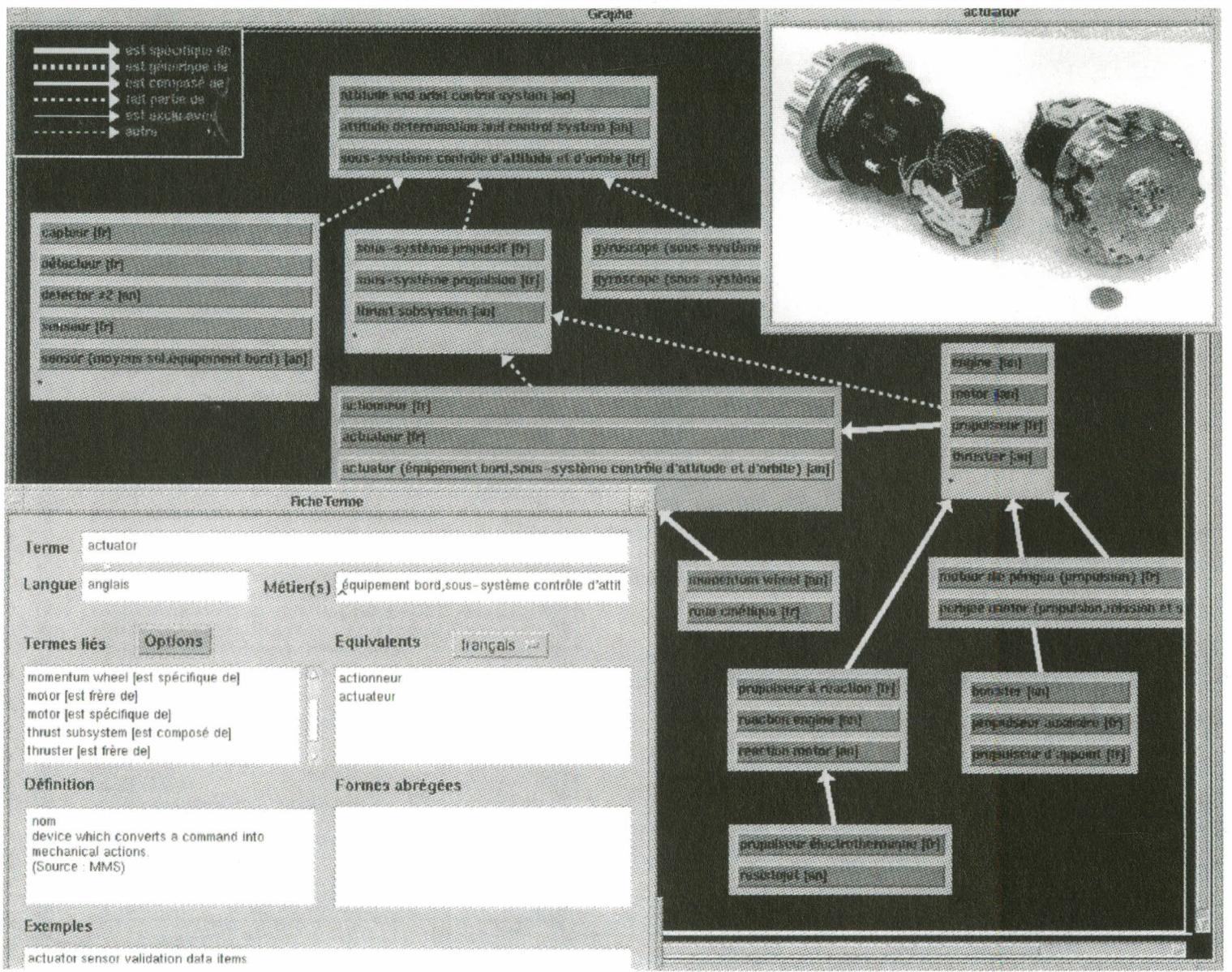
**actuator (équipement bord sous-système contrôle d'attitude et d'orbite)**

actuator model (sous-système contrôle d'attitude et d'orbite, moyens sol)

RECHERCHER
AFFICHER TERME
AFFICHER GRAPHE
ETAPES
FERMER

Row: A E PARACOURANT 5 of 12 100%

Figure 2 : Réseau sémantique.



un système de recherche d'information afin d'améliorer l'indexation automatique des documents et le processus de recherche d'information qui prendra en compte la terminologie du domaine et les mots composés.

L'utilisation d'un correcteur orthographique et grammatical adapté au domaine technique permettra d'améliorer la qualité linguistique des documents techniques.

#### 4.4 Aide à la consultation de documents électroniques

La base de données terminologiques permet de générer des liens dynamiques entre les informations. Ainsi, un rédacteur de documents électroniques pourra connaître la définition d'un terme ou la forme développée d'un sigle utilisé dans le document.

#### 4.5 Mémoire d'entreprise

La mémoire d'entreprise est un outil qui étend le système d'information à la gestion des connaissances sous toutes ses formes (hommes, groupe d'hommes, documents, graphiques, bases de données). Elle s'intéresse aussi bien à la valeur des informations qu'aux liens qui les unissent. Elle offre la possibilité de retrouver, au bon moment, avec le bon niveau et sous la bonne forme, les informations qui sont dispersées dans l'entreprise. La mémoire d'entreprise s'appuie sur les trois fonctions suivantes :

- La formalisation qui transforme l'expérience et le savoir-faire sous une forme réutilisable;
- La modélisation qui permet de structurer l'information et s'exerce sur la terminologie utilisée et les objets et processus de l'entreprise;
- L'échange d'information entre personnes qui permet de décloisonner les métiers, les structures

hiérarchiques et plus généralement les hommes de l'entreprise.

Les informations contenues dans la base de données terminologiques permettent de capitaliser une des richesses de l'entreprise et interviennent en support de l'application mémoire d'entreprise.

### 5. Conclusion

Le développement de cette base de données terminologiques bilingues et du système de gestion de la terminologie est le fruit d'une coopération étroite entre le laboratoire Aramihs et les ingénieurs de Matra Marconi Space. Les résultats ont été obtenus grâce à la composition d'une équipe multidisciplinaire et bilingue qui associait des informaticiens, linguistes, ergonomes et experts du domaine spatial. La consultation de la base de données terminologiques est en cours d'évaluation par les ingénieurs et le développement des différentes applications présentées ci-dessus est en cours de finalisation.

*Amaury Simon,  
Matra Marconi Space - Aramihs,  
Toulouse.*