

LA GESTION DES CONNAISSANCES EN GEOTECHNIQUE : TRENTE ANNEES D'ESSAIS ET PERSPECTIVES

THE MANAGEMENT OF GEOTECHNICAL KNOWLEDGE: THIRTY YEARS OF TRIALS AND PERSPECTIVES

René-Michel FAURE¹, Jean-François THIMUS²

¹*Ecole nationale des Travaux publics de l'Etat, Lyon, France*

²*Université catholique de Louvain, Louvain-la-Neuve, Belgique*

RÉSUMÉ – Parmi les enjeux sociétaux, la transmission du savoir, tant dans l'entreprise qu'à l'université, a été bouleversée ces dernières années par l'informatique et ses réseaux. En rappelant les travaux en géotechnique et informatique depuis trois décennies, des tentatives de systèmes experts, de bases de données, et plus récemment d'ontologies, l'article propose une approche nouvelle dédiée aux bureaux d'étude.

ABSTRACT – Knowledge management, either at office or university, is strike by networks. Reminding shortly past works dealing with geotechnics, data-bases, expert systems and ontologies, the paper presents a new approach sized for offices.

1. Introduction

L'enjeu du recueil des connaissances est primordial pour l'évolution de nos sociétés, malmenées par une gestion froide des ressources humaines et un transfert des connaissances souvent inhibé par le choix d'un point de départ. Faire du recueil de connaissances c'est donner au futur les bases nécessaires à son existence.

La transmission des savoirs fut d'abord orale, puis livresque et notre enseignement est encore basé sur le livre. Mais comment retrouver dans un livre l'information dont on a besoin pour traiter le problème posé par un contexte particulier? On a cru arriver à ce but avec l'aide de moteurs de recherche. Ceux qui utilisent les moteurs de recherche comme Google en connaissent les limites, car la contextualisation des connaissances n'y est pas incluse. Pour bien utiliser les moteurs de recherche, il faut savoir ce que l'on cherche, connaître en partie la réponse.

L'extension des domaines du savoir dans l'enseignement conduit à des programmes surchargés et qui ne peuvent fournir simplement le savoir suffisant et nécessaire pour traiter une situation donnée.

Le projet MKD, présenté ci-après, fera du contexte le point de départ de la recherche du savoir nécessaire.

2. Brefs rappels historiques

2.1. Années 80 et les bases de données

Avec le développement de l'informatique, le stockage de l'information devient possible et les premières tentatives de bases de données apparaissent.

Les axes de développement sont :

- Les essais de laboratoire : Modelisol (Favre et al., 1992)
- Les sites expérimentaux (Greenwood, 1988)
- Les glissements de terrain (Lacube et Durville, 1989)
- La surveillance des sites
- Les logiciels intégrés capables de fédérer des logiciels de calcul, comme le projet CASTOR au Canada (Soulié et al., 1992)
- L'information étant spatiale, les premiers SIG côtoient les bases de données et les premiers blocs-diagramme sont édités (Faure, 1999), (Faure et al., 2001, 2002)

2.2. Années 90 : Les systèmes experts

La communauté géotechnique française a connu un vaste engouement pour les systèmes à base de connaissance dans les années 1980 et 1990. Utilisant règles de production et moteur d'inférence, de nombreux systèmes experts ont donc été mis à l'étude comme XPENT (Faure et al. 1992) (Laouini et al., 1992) mais rarement utilisés en pratique. Les causes en sont multiples et ont parfois été analysées (Magnan 1992), mais trois points d'achoppement apparaissent :

- le besoin pour l'utilisateur de comprendre l'outil qu'il manipule, au moins ses principes fondamentaux, et le rejet des intermédiaires pour le traitement de la connaissance,
- l'investissement souvent trop important, au moins en temps et en énergie, pour alimenter une base de connaissances en mesure de fournir des réponses satisfaisantes.
- l'impossibilité de tout prévoir dans un monde ouvert, comme celui de la géotechnique.

De plus, comme l'écrit (Magnan, 2002) le travail du géotechnicien est souvent unique, à la façon d'un artisan, peu répétitif et de ce fait peu automatisable.

2.3. Années 2000 : Les ontologies

Si les ontologies trouvent leurs origines dans la philosophie (et notamment Aristote), ce n'est pas avant 1991 que l'on trouve dans la littérature spécialisée la première définition des ontologies dans le cadre des systèmes d'information:

- Neches et al. (1991) mettent en avant l'approche terminologique.
- Gruber (1993) voit dans les ontologies une conceptualisation du monde.
- Guarino (1998) privilégie une approche logicienne mettant en relation un vocabulaire et un agencement conceptuel (un modèle).
- Sowa (2001) se veut synthétique et fait le lien entre l'approche forcément terminologique du concept (il est difficile sinon impossible d'exprimer une idée autrement que par support linguistique) et la conceptualisation d'un domaine
- Aussenac-Gilles (2004) est beaucoup plus pragmatique en visant à la construction d'un référent de connaissances plutôt qu'à une représentation du monde.

Une ontologie apparaît comme un ensemble de syntagmes (mot ou groupe de mots) nommant des concepts avec un ensemble de relations reliant ces concepts. On obtient un graphe 3D avec des milliers de nœuds (les concepts) et des centaines de liaisons différentes (les relations entre deux concepts).

Il existe d'autres définitions des ontologies, mais elles vérifient toutes les assertions suivantes:

- Les ontologies sont des outils de référence pour une communauté,
- Les ontologies doivent permettre de dépasser les phénomènes terminologiques et donc corollairement, les ontologies traitent avant tout de concepts.

Une nouvelle structure d'ontologie est présentée dans cet article, structure qui permet de faciliter sa mise en œuvre et de résoudre les problèmes terminologiques comme la synonymie et la métonymie. Elle permet aussi le glissement sémantique pour l'élargissement des questions, des statistiques lexicales et l'obtention des thésaurus. (Faure et al., 2014a) (Faure et al., 2015)

3. Information, connaissance et méta connaissance

Pour l'ingénieur la connaissance est issue de quatre formes d'information:

- La dénomination des objets et des concepts. En effet connaître l'existence et la définition de quelque chose (un objet, une théorie, une action, un état..) est une information.
- La seconde forme est l'ensemble des relations, des faits reliant deux ou plusieurs concepts. Un savoir-faire comme la composition d'une formule de béton fait partie de ce second type de connaissance.
- La troisième forme est l'algorithme qui permet une traduction numérique des relations entre objets.
- La dernière forme est la donnée, généralement chiffrée.

Ces quatre formes conduisent à la connaissance quand l'ingénieur les utilise dans le but de construire. « *Pour obtenir la quantification de tel objet, quels concepts sont nécessaires pour utiliser ce logiciel, avec quelles précautions et quelles données ?* » est la démarche usuelle de l'ingénieur dans son acte de construire.

Il se sert aussi de sa méta-connaissance des choses, à savoir où et comment il peut avoir accès à telle information. C'est par l'expérience qu'il acquiert cette méta-connaissance. (Faure, 2007)

4. L'impact Internet pour l'enseignement

L'enseignement est un raccourci pour l'acquisition et le maniement de l'information. Il se heurte aujourd'hui à des difficultés et à une concurrence comme :

- Le choix des matières à enseigner, alors que les nouvelles technologies apportent chaque jour de nouvelles façons d'aborder le projet, car le temps imparti est limité.
- Comment apporter à l'ingénieur, dans sa carrière, de nouvelles informations ?
- Comment enseigner le savoir-faire, le savoir apprendre, le savoir-faire faire ?

Internet propose deux solutions :

- L'enseignement en ligne et les MOOC (Massive Open Online Course). Ces cours diffusés via Internet contraignent déjà les universités à mettre plus d'attractivité dans leur cursus, pour résister dans le combat économique qui s'annonce.
- Des moteurs de recherche de plus en plus performants. Le projet Watson d'IBM promet une analyse très performante de milliers de pages dans divers domaines, mais avec une machine superpuissante.

Si ces approches sont très utiles, elles ne sont réservées qu'à un petit nombre, autodidactes ou formés à ce type de recherche. Et elles sont dévoreuses de temps. Cependant des projets moins ambitieux rapprochant élèves et professeur ont toujours leur place, mais il apparaît aussi que d'autres approches sont à inventer, en triant l'information de façon plus pertinente (Faure et al., 2012).

5. Approche téléologique

Téléologique traduit la notion de but. On cherche quelque chose pour mieux résoudre un problème posé par la réalisation du projet de construction.

Une encyclopédie est peu utile si on ne sait pas la page où se cache l'information, le moteur de recherche fournira trop de références qu'on ne saura exploiter et l'enseignement fournira bien peu de fois une réponse précise.

5.1. Les projets Ramcesh et MKD

Le projet Ramcesh (Recueil Assisté et Maniement des Connaissances des Espaces Souterrains Habités) a été initialisé au CETu au début des années 2000 (Faure, 2004) (Faure et al., 2006 et 2007). Il a été suivi par le projet MKD (*Modelizing Knowledge by Domain*) montrant que l'approche est facilement transposable à tout domaine d'activité.

Le but de MKD est la mise en forme quasi automatique de toute l'information contenue dans des publications afin de répondre à la question : « *Que dois-je savoir pour traiter ce problème et son contexte ?* »

La réponse, si précise soit-elle, ne donne pas de solution, mais les bons éléments pour que l'utilisateur énonce une solution. La justesse de cette solution dépend de la précision du contexte décrit. Et quel avantage pratique de voir sur son écran tout ce qu'il faut savoir sur ce contexte. L'art de l'ingénieur n'est que facilité.

Dans ce projet, l'ontologie proposée est une ontologie de poids léger, pragmatique et distinguant deux relations, la relation de subsomption (est une sorte de) et la relation de voisinage (est voisin de) pour construire des arbres de sous-domaines faciles à définir. Le granule de connaissance, structure de petite dimension correspond aux règles de production des systèmes experts, augmenté d'un univers et d'un modèle d'univers. C'est le granule qui permet de résoudre les difficultés de représentation du langage. Une recherche plein texte des formes déductives (si alors) par des outils de traitement du langage fournit les fragments de textes sur lesquels seront bâtis les granules dont la structure permet la composition en des granules plus complets et de plus vaste portée (Faure et al., 2008) (Faure et al., 2015). La figure 1 montre la structure du granule et sa place dans l'ontologie du domaine géotechnique. L'annexe explique plus en détail ces notions.

5.2. Mise en œuvre

Le choix des textes représentant le domaine est la première action de la mise en œuvre de MKD. Ce corpus de textes pourra être augmenté à chaque moment (cas de la veille technologique par exemple)

Un module de MKD, construit classiquement une base documentaire de documents fournis.

Un autre module recherche tous les syntagmes (mots ou groupe de mots utilisés plusieurs fois) des textes, qui sont rangés dans les arbres des sous-domaines.

Avec des outils de traitement automatique du langage, MKD retrouve les fragments de texte (phrases déductives) qui sont transformés en granules. Au fil des textes ajoutés, la base de connaissances ainsi créée croît et se complète.

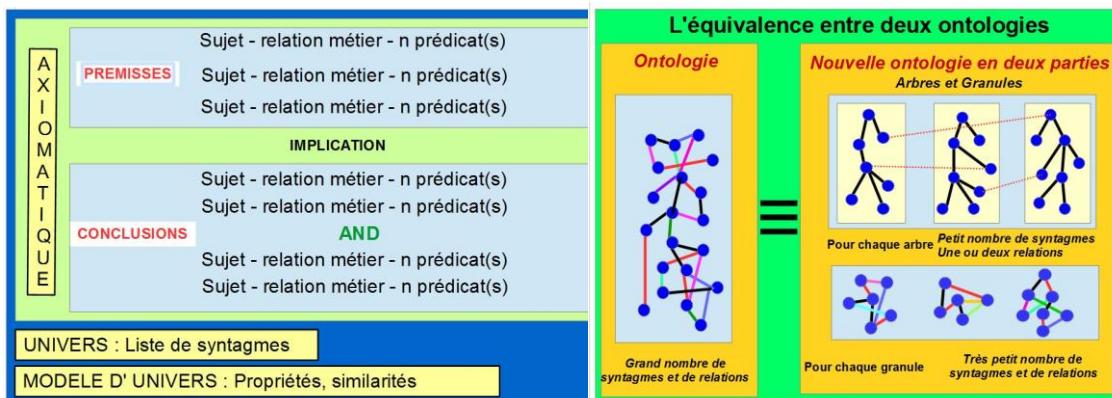


Figure 1 : la structure du granule et l'ontologie équivalente choisie

5.3.Applications

- L'application encyclopédique de la base de connaissances correspond à la recherche dans la base de tous les granules apportant une réponse totale ou partielle à une question.
- L'indexation automatique de documents se fait lors de la mise en base, la signature de l'article (syntagmes de l'article et de l'ontologie est calculée). Cette signature peut servir de mots clés pour toute recherche de document.
- La fouille de texte et le remplissage de formulaires servent, par exemple, à extraire automatiquement les informations pour alimenter une base de cas. La gestion d'un patrimoine est ainsi facilitée.
- La vérification de conformité et de cohérence entre textes ou règlements est simple par détection des granules contradictoires et l'utilisateur averti de ces incohérences devra les lever.
- Suivi de chantier et détection de risques : Les fiches de suivi d'un chantier peuvent être systématiquement analysées, dès leur réception, et des avertissements faits si certaines valeurs ne sont plus conformes aux règles, même numériques, explicitées dans les granules.
- Aide à la conception : Aspect encyclopédique, détection d'incohérence, liaison avec codes de calcul permettent d'envisager une application type SIAD, (Système Informatique d'Aide à la Décision) centrée sur la construction d'un type d'ouvrage.
- Aide au dépouillement ou à la rédaction d'appel d'offre : La détection d'incohérence et le remplissage de formulaire à des fins de comparaison sont une première étape dans le dépouillement d'appels d'offre en vue du choix de l'un d'entre eux en fonction de critères qui peuvent être décrits sous forme de granules.
- Mémoire d'un laboratoire, d'un bureau d'expertise ou de contrôle : Au fil des années les expertises s'accumulent et qui peut se souvenir de tous les détails des expertises afin d'assurer une cohérence entre les expertises faites dans des temps différents et part des collaborateurs différents. L'outil MKD permet le recueil et la recherche de ces détails en utilisant les granules de connaissance qui évitent la redondance et gardent le contexte. Pour un problème donné les nouveaux collaborateurs récupèrent rapidement le passé du bureau d'étude. La confidentialité des expertises (judiciaires) est assurée car la référence inscrite dans le granule n'est dans ce cas, qu'un numéro de dossier.

- Base de cas pour raisonnement par analogie : la recherche de cas semblables (ou d'articles semblables) est une simple comparaison de la signature des articles

5.4. Quelques résultats.

MKD a permis de traiter les 2344 articles de congrès internationaux des tunnels, les 600 articles des JNGG et les congrès de l'AFTES et autres cours ou recommandations. L'ontologie en français comprend 22 sous-domaines et regroupe près de 6000 syntagmes (Faure et al., 2014b).

La fonction la plus usitée est la recherche d'articles semblables à un article donné. Au retour d'un congrès savoir qui travaille sur le même sujet que celui de l'article est très prisé.

Le prototype actuel, écrit en PHP et Perl avec une base MySQL, a permis de valider tous les points cités.

6. Conclusion

Mettre en œuvre MKD, pour préserver et diffuser (via un Intranet) l'expérience et le savoir d'un service c'est aussi :

- Rafraîchir les esprits, les mémoires, et mettre à plat la connaissance collective,
- S'approprier un métier rapidement, efficacement, le transfert de connaissances et de compétences est favorisé, la communication améliorée,
- Valoriser le domaine d'expertise de chacun et montrer le périmètre de compétence de chacun ce qui éveille l'intérêt des autres et permet l'enrichissement mutuel,
- Pérenniser le capital de savoir de l'entreprise,
- Gagner en réactivité, en productivité (on se comprend plus rapidement), on gagne en qualité (on se comprend mieux),
- Créer de la connaissance (par juxtaposition de connaissances) et optimiser l'organisation du service.

8MKD attend le service qui le mettra en œuvre pour plus de performance.

7. Références bibliographiques

- Aussenac-Gilles N. (2005). Modélisation et ontologies. Cours TCAN, Batz-sur-Mer.
- Faure R.M., Mascarelli D., Vaunat J., Leroueil S., Tavenas F. (1992). Present state of development of XPENT, expert system for slopes stability problems. *6th Int. Congress on Landslides Bell editor Christchurch*, pp. 1671-1678.
- Faure R-M. (1999). Data-bases and the management of landslides. *Int. Symp. on Landslides. Shikoku (Japan). IS Shikoku'99*, pp. 1317-1330.
- Faure R.M., Locat J., Thimus J.F., Picarelli L. (2001). Base de données internationale sur les glissements de terrain. *Revue Française de Géotechnique. Numéro 95 et 96 : Mouvements de terrain*, pp. 183-187.
- Faure R.M., Robert A.A., Thimus J.F., W. Schubert, Bourgeois E., Bois A.P. (2002). KBT- Base de connaissances internationales des tunnels. *Tunnels et Ouvrages Souterrains*, n°170, pp. 94-98.

- Faure N. (2004). Le granule de connaissances. *Actes Inforsid 2004, Biarritz*
- Faure R.M., Faure N., Hémond G. (2006). The use of knowledge management in the management of tunnels and tunnel projects. *ITA-AITES World Tunnel Congress. Séoul, 11p.*
- Faure R.M., Faure N., Hémond G.(2007). Recueil Assisté et Maniement des Connaissances des Espaces Souterrains Habités. *Tunnels et Ouvrages Souterrains, n° 202, pp. 351-363.*
- Faure N.(2007). Ontologies, thesauruses and information system for tunnels. *ITA-AITES World Tunnel Congress. Prague, Bartok ed. pp. 65-68.*
- Faure N., Faure R.M., Balasch M.A., Cottaz Y.(2008). Mise en forme de granules de connaissances à l'aide de l'outil RAMCESH. *Congrès International AFTES, Monaco, pp. 505-514.*
- Faure N., Hémond G., Gress J.C., Faure R.M. (2012). Course validation using an automatic system for sorting similar technical papers. *SFGE conference, Galway*
- Faure N., Thimus J.F., Faure R.M. (2014a). Analyse statistique lexicale pour la construction d'une base de connaissances. *Tunnels et Espaces Souterrains, n°244, pp. 317-328.*
- Faure N., Thimus J.F., Faure R.M.(2014b). Analyse statistique lexicale des articles des congrès JNGG pour des recherches personnalisées sur le contenu. *Journées Nationales de Géotechnique et de Géologie de l'Ingénieur, JNGG2014, Beauvais*
- Faure N., Faure R.M.(2015). Evolution des approches de la représentation des connaissances en géotechnique illustrée par les recherches du CETu. *Revue Travaux En Souterrain, n°251, pp. 361-375.*
- Favre J.L., Heritier B., Michalski R. (1992). MODELISOL : banque d'essais de référence pour l'ingénierie de demain. *Coll. Int. Géotechnique et Informatique, ENPC, Paris, France.*
- Greenwood J.R.(1988). Developments in computerised ground investigation data. *Ground Engineering, september*
- Gruber T. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition, 5(2), pp.199-220.*
- Guarino N. (1998). Formal ontology and information systems. *Amended version of a paper appeared in N. Guarino (ed.), Volume 46 Frontiers in Artificial Intelligence and Applications. IOS Press*
- Lacube J., Durville J.L.(1989). Un essai de fichier informatique sur les mouvements de terrain. *Bull. Liaison labo. P & Ch, pp. 86-89.*
- Laouini H., Deffayet M., Robert A., Gueniffey Y.(1992). Développement d'un système expert d'aide à la définition d'un programme de reconnaissance pour les tunnels. *Coll. Int. Géotechnique et Informatique, ENPC, Paris, France*
- Magnan JP.(1992). CESSOL: Bilan du Développement d'un Système-expert. *Geotechnique et Informatique, Proc. Int. Conf. on Geotechnics and Computers, Presses de l'École Nationale de Ponts et Chaussées, Paris*
- Magnan J.P. (2002). L'organisation du travail en géotechnique : développement, normalisation et artisanat. *Lettre de la Géotechnique 26 et 27, Société Internationale de la Mécanique des Sols et de la Géotechnique*
- Neches R., Fikes R., Finin T., Gruber T., Patil R., Senator T., Swartout W.R. (1991). Enabling Technology for Knowledge Sharing. *AI Magazine, Volume 12, No. 3*
- Soulié M., Lessard G., Vaucher J.(1992). Programme de dimensionnement des ouvrages en terre suivant l'approche orientée objet. *Coll. Int. Géotechnique et Informatique, ENPC, Paris, France*

Sowa J.F. (2001). Knowledge representation. Logical, philosophical and computational foundations. *Brooks/Cole Publishing Co, Conference on Conceptual Modeling: conceptual modeling*, pp. 383-396, November.

Site : www.pentes-tunnels.eu

8. Annexe : quelques précisions

En complément de la figure 1 qui montre la structure du granule il faut préciser que l'axiomatique qui traduit la phrase déductive d'un texte avec les seuls mots de l'ontologie est une succession de propositions (sujet = ce dont on parle, relation, ce qu'on en dit) à trois éléments pouvant être combinées avec des ET et des OU. Ces triplets sont regroupés en prémisse et conclusion reliées par une des cinq relations déductives impliquant (éventualité, obligation, négation, recommandation, risque). Lors de sa création le granule est comparé aux granules existants et il est introduit dans la base de connaissance que s'il apporte des connaissances nouvelles (nouveau granule ou granule existant modifié) afin que cette base ne contienne aucune redondance.

A titre d'exemple voici un granule.

Phrase	Relations	Univers	Modèle d'univers
SI			
Le coefficient de sécurité au glissement d'un mur en béton armé est inférieur à 1.5	Y est en Z X caractérise Y	Coefficient de sécurité au glissement, Mur, Béton armé	X = coefficient de sécurité au glissement Y = Mur Z = béton armé X < 1.5
ALORS (modalité : recommandation)			
Mettre une bêche à l'arrière de la semelle dont la hauteur soit au moins égale à l'épaisseur de la semelle	C caractérise A D caractérise B À à l'arrière de B	Bêche, Semelle, Hauteur, Épaisseur	À = Bêche B = Semelle C = Hauteur D = Épaisseur C > D

Tableau - Exemple simplifié de granule

C'est ce granule issu des théories de la représentation du discours qui permet la transcription simple et fidèle de toute forme d'information menant à la connaissance et ses possibilités de composition permettent d'exprimer de nouvelles connaissances.

La structure de l'ontologie est faite d'arbres plans (une relation principale pour tous les termes et quelques relations, dites de voisinage, pour lier deux concepts de nature différente mais qui sont souvent cités ensemble). Ces liens permettent le glissement sémantique qui élargit les questions posées et induit des réponses pertinentes. Les lignes pointillées d'un arbre à l'autre signifient que les deux concepts sont identiquement nommés, ce qui relie tous les graphes entre eux et permet de naviguer dans l'ensemble du vocabulaire, fait de syntagmes (groupe de mots) et non pas de mots simples.