

LE MANIEMENT DE LA CONNAISSANCE EN GÉOTECHNIQUE, DE 1985 À CE JOUR.

Faure René-Michel

Centre d'Etude des Tunnels, 69500 BRON, France, rene-michel.faure@equipement.gouv.fr

Faure Nicolas,

Equipe Modème, Université de Lyon, France, nfaure@univ-lyon3.fr

RÉSUMÉ

L'article rappelle les tentatives des projets XPENT et WASSS pour collecter et mettre à disposition les connaissances en géotechnique et tire les enseignements de ces projets. Puis l'article présente le projet RAMCESH comme suite logique des projets précédents. L'usage des techniques actuelles de l'informatique permet d'obtenir un outil capable de synthétiser la mémoire d'un service et/ou la mémoire technique d'une communauté scientifique. Le projet RAMCESH est centré sur les espaces souterrains habités mais peut facilement être décliné vers d'autres domaines comme celui des glissements de terrain et aider à la recherche de solutions préventives. En présentant une des applications de RAMCESH sur la comparaison du contenu scientifique de textes de référence, les auteurs évoquent les potentialités de l'outil.

ABSTRACT

This paper reminds the XPENT and WASSS projects about data and knowledge storage for extracting the results of these projects. Then the paper presents the RAMCESH project as a logical continuation of the previous projects. The updated computer technologies allow the realization of a tool able to store all the memory of a team or the technical knowledge of a scientific community. RAMCESH project is actually targeted on underground works but it is easy to direct it to other domains such as landslides for hazards mitigation. When presenting a Ramcesh result about the comparison of the scientific knowledge included into two reference texts, the authors show the potentialities of the tool.

1. INTRODUCTION

L'ensemble des trois projets présentés peut être vu comme le fruit d'un long processus de réflexion puis de maturation, aiguillonné par l'évolution des technologies informatiques et répondant à un besoin d'accès rapide à la connaissance. De plus en plus, les questions posées à l'ingénieur doivent être résolues rapidement, très rapidement et sans faute. Il lui faut donc retrouver le savoir correspondant au problème posé dans les meilleurs délais. Cela pose le problème de l'acquisition du savoir. 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 la connaissance dont on a besoin pour traiter le problème posé par un contexte particulier? On a cru arriver à ce but avec l'aide des systèmes experts, des bases de données puis des moteurs de recherche. Aujourd'hui, ceux qui utilisent 'Google' en connaissent les limites, car la contextualisation des connaissances n'y est pas incluse et l'on se retrouve avec quelques milliers de références à explorer. C'est aussi le problème d'un service de documentation, qui, à une question posée, délivre bien trop de documents, car on assiste à une prolifération des textes avec beaucoup de redondance. Alors, comment choisir le bon document qui contient la réponse cherchée.

Les trois projets présentés sont une réponse à cette quête d'une information juste, pertinente et limitée en volume, au bon moment.

2. LE PROJET XPENT

XPENT (Faure *et al.*, 1988) (Mascarelli *et al.*, 1992) a été un projet de coopération entre l'école nationale des Travaux Publics de l'Etat (ENTPE) et l'université Laval à Québec pour la mise au point d'un système expert en stabilité de pentes. Financé en partie par France-Québec, ce projet a été aussi l'occasion de nombreux échanges d'étudiants. Du point de vue géotechnique, cela fut l'occasion de nombreuses discussions animées entre experts et de découvrir un monde ouvert (du point de vue informatique), ce qui signifie que la géotechnique est une science de la nature et que tout n'est pas connu. Un monde ouvert au sens informatique est un monde où toutes les règles ne sont pas connues, il peut toujours arriver quelque chose de non prévu. Dans le monde de la géotechnique c'est une réalité quotidienne, qui est aussi vraie pour tous les construits en liaison avec le sol. A l'opposé, un monde fermé est un monde parfaitement connu, où tous les éléments sont recensés. Le monde de l'industrie est souvent de ce type, dans un processus industriel, sur une chaîne de montage par exemple, seules les pièces prévues peuvent arriver. Ceci explique en partie pourquoi les systèmes experts des années 80 ont eu de belles réussites dans le monde industriel, mais jamais dans les sciences de la terre. La découverte d'un gisement par le système expert 'Prospector' semble bien n'avoir été qu'un heureux concours de circonstances. (voir annexe 1 : Tentative de définition de la géotechnique)

Ce projet XPENT a permis de saisir le comportement de l'expert par les interviews que la méthode KOD (Knowledge Oriented Design) imposait. L'expert ne développe pas son

expertise avec grande cohérence, c'est un parcours fait d'allers-retours, quasiment fractal, avec uniquement le but d'arriver à une solution à l'aide de nombreuses références, pas toujours judicieuses, mais servant de support à cette pensée tumultueuse. A la fin du projet XPENT, la nécessité d'une base de cas où les experts tireraient des parties de raisonnement a été mise en évidence (Mascarelli, 1994).

XPENT a donné lieu à de nombreux articles et trois thèses ont été soutenues sur le sujet. (Faure *et al.*, 1992) La géotechnique ne pouvant être une suite de déductions du fait de contradictions qui conduisaient à une analyse sans fin, une autre approche était nécessaire et pour réutiliser facilement actigrammes (décomposition d'un processus en processus élémentaires) et règles de production mises à plat, quelques années plus tard, le projet WASSS était lancé.

3. LE PROJET WASSS

Le projet WASSS (Faure *et al.*, 1999) (World Area Slope Stability Server) a suivi le projet XPENT dans le cadre du groupe technique pour les glissements de terrain (TC 11) de la société internationale de mécanique des sols et des fondations (ISSMFE). D'une communauté mono langue nous étions dans une communauté multi langues. Pour se comprendre, le premier travail a été la création d'un glossaire multi langues. Cette tentative a mis en évidence des exigences souvent contradictoires entre les pays membres qui ont été intégrées dans un serveur de cas. (Faure *et al.*, 2001)

Nous retiendrons de ce projet, la difficulté du dialogue entre communautés et la simplification extrême qu'il faut apporter aux managements informatiques et aux théories mises en oeuvre. Complicé à la rigueur mais surtout pas complexe. Les termes 'complicé' et 'complexe' sont ceux utilisés dans la systémique (étude des systèmes) où 'complicé' qualifie un processus simple à comprendre mais avec de nombreuses répétitions et branchements, si bien qu'une machine est une aide nécessaire, alors que l'utilisateur suit facilement les enchaînements. 'Complexe' correspond à un système dont l'abstraction nécessite un gros effort de compréhension, et la machine qui traduit le processus apparaît comme une boîte noire.

La disparition brutale du TC11, et la dispersion de ses acteurs, a arrêté le projet en 2000, cependant un goût d'inachevé, et l'apparition de nouvelles techniques informatiques, relancent dans un cadre plus restreint, une suite à cette quête de l'outil fabuleux d'aide au géotechnicien.

4. LE PROJET RAMCESH

Le projet Ramcesh (Recueil Assisté et Maniement des Connaissances des Espaces Souterrains Habités) est un projet conduit conjointement par la société Solem Expertises sa et le Centre d'Etude des Tunnels, avec l'aide

de l'université Lyon 3 (Laboratoire Modème) pour stocker sans redondance et retrouver instantanément la connaissance correspondant à un contexte donné. (Faure, 2006)

Le contexte de Ramcesh

Depuis quelques années la nécessité du recueil des connaissances est admise par tous dans le but principal de transmettre savoir et savoir-faire entre générations. Devant la difficulté du processus, les experts ne sont pas enthousiastes, beaucoup d'applications de 'management des connaissances' produisent une cartographie des connaissances de l'entreprise. Cette cartographie est alors utilisée par le Directeur des Ressources Humaines qui peut ainsi faire une adéquation entre les moyens humains et les domaines de l'entreprise. Le 'livre des connaissances' produit par la méthode MASK (Ermine, 2003) est un processus de mise à plat de la connaissance dont l'appropriation par un nouveau venu reste classique, par la lecture. Nous avons voulu aller plus loin et faire que la connaissance puisse 'réagir' à un contexte donné, pour par exemple, détecter une situation à risque demandant une action particulière. Cependant, le choix de cette action reste du fait de l'utilisateur et l'outil Ramcesh apparaît comme un bon compagnon qui avertit, mais ne décide pas. Cette transmission du savoir 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 de voir sur son écran tout ce qu'il faut savoir sur ce contexte. L'art de l'ingénieur n'est que facilité.

Définition de la connaissance utile au projet et développement théorique de l'outil

La "connaissance" n'existe pas en elle-même, elle ne se manifeste que par la réalisation de quelque chose. Pour réaliser et faire accepter un travail de conception il faut mettre en oeuvre la "connaissance", la plus exacte et la plus performante possible. Ceci induit une dimension téléologique et une dimension maïeutique dans la connaissance, la première permettant de qualifier la connaissance au vu de la qualité de la réalisation, la seconde de faire foisonner les connaissances afin d'en extraire d'autres pour aller plus loin.

C'est pourquoi, faire du recueil de connaissance dans l'absolu a peu de sens.

C'est aussi pour cela, que la contextualisation qui fonde une vérité au discours et reconnaît ou précise les buts poursuivis, devient nécessaire et omniprésente.

Pour respecter cet aspect dynamique de la connaissance, le projet Ramcesh aborde celle-ci d'un point de vue sémasiologique (c'est à dire à partir d'un signe lexical vers son sens) avec l'aide des trois formes suivantes:

- les catégories, définitions intuitives des signes, qui sont des ensembles indéfinis dont les limites sont floues.
- les occurrences, actualisations en contexte des catégories, qui fixent plus précisément le volume dénotatif associé à chaque signe.

- les instances, entités individuelles, décrites en fonction des occurrences.

Cette approche est centrale dans le projet Ramcesh. (Faure, 2007)

Pour traduire cette approche en outils informatiques, le modèle du signe et la théorie de la représentation du discours (DRT) ont été très utiles. Ces deux techniques permettent d'obtenir un modèle de représentation de la connaissance qui est mis en oeuvre à l'aide d'une ontologie. Le choix d'une ontologie simple (de poids léger) ne permet qu'une mise en oeuvre incomplète, ce qui a conduit à définir le granule de connaissance, pour apporter à l'ontologie choisie la souplesse d'emploi nécessaire. (Faure N, 2004 et 2006)

La mise en oeuvre de Ramcesh

Ramcesh est un projet basé sur les collaborations d'experts d'un domaine, impliqués dans un processus cumulatif de stockage de bribes de connaissance pour finaliser une base de connaissance capable de transmettre ainsi les connaissances d'un domaine et d'aider efficacement tout utilisateur dans son travail de conception ou de recherche. Ceux qui alimentent la base en seront les premiers utilisateurs et apprécieront les possibilités innovantes et puissantes de l'outil.

Le premier travail est de construire une base de connaissance suffisamment riche pour que son utilisation soit probante. Le choix des textes initiaux doit apporter ce premier niveau de complétude. A partir de ces textes l'ontologie s'enrichit et les granules correspondants sont écrits. L'ontologie contient les mots et syntagmes qui nomment les concepts de façon relativement floue, alors que les granules qui n'utilisent que ces concepts précisent le sens par une description du contexte d'emploi et de vérité de la proposition énoncée. Une interface très conviviale sur Internet permet une construction collaborative par des experts ayant une définition précise de leur rôle. (voir annexe 2)

Début 2008, l'ontologie comprend plus de 7000 syntagmes et quelques centaines de granules ont été écrits. Une thèse a été présentée en Octobre 2007. (N. Faure, 2007)

Premiers résultats

Un des résultats simples de cette application est la comparaison du contenu cognitif de deux textes. Nous avons choisi les recommandations sur l'étanchéité des tunnels du Centre d'Etude des Tunnels et de l'association française des travaux en souterrains. Trois familles de granules de connaissance apparaissent, celle des granules communs aux deux textes et celles des granules appartenant à un seul texte. Une harmonisation reste à faire.

5. CONCLUSION

De la géotechnique à l'acte de construire, la problématique reste sensiblement identique, aussi le projet Ramcesh peut

être appliqué à tout domaine d'investigation et de réalisation dans un domaine vaste et partiellement connu. Ramcesh oeuvrant dans un monde ouvert utilise une approche élargie par rapport à celle des Systèmes Experts, par l'ajout des apports des ontologies et des univers de modèles. Il repousse les limites d'usage des systèmes experts et des bases de données et devrait être une aide efficace aux systèmes documentaires. De plus, la spécificité de la géotechnique peut être transposée à bien d'autres domaines, ce qui devrait élargir son champ d'application.

Pour ce qui est de la stabilité des pentes et des situations à risque associées, l'outil issu de Ramcesh devrait permettre cette synthèse nationale et internationale, utilisant un langage commun (c'est le rôle que l'on attribue usuellement aux ontologies) et rassemblant les connaissances éparses et fragmentaires de toute une communauté qui sait que l'éventail des possibilités de la nature ne peut contenir dans la tête d'un expert à qui le temps de formation est maintenant compté.

Bibliographie

Ermine J.L., 2003, La gestion des connaissances, Hermès ed., 166p,

Faure N. (2004), Le granule de connaissances, Actes Inforsid 2004, Biarritz

Faure N., 2006, Représentation des connaissances géotechniques, Actes Inforsid 2006, Hammamet

Faure N., 2006, The use of knowledge management in the management of tunnels and tunnel projects. Proc. Int. Conf. ITA/AITES, Seoul, Korea

Faure N, 2007, Un système d'aide à la modélisation en géotechnique, Thèse Université Jean Moulin, Lyon, 151p

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. p1671-1678

Faure R-M., 1999, Data-bases and the management of landslides. Int. Symp. on Landslides. Shikoku (Japan). ISL Shikoku'99, p1317-1330

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.

Magnan JP., 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, 1992

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

Mascarelli D., Faure R.M., Kastner R., *Anatomie d'un projet à base de connaissances, XPENT, système de travail en ingénierie des pentes*, Actes du colloque Géotechnique et Informatique, ENPC, Paris, 1992

Mascarelli D., *Ingénierie des pentes instables : approche orientée modélisation de la connaissance*, Thèse de doctorat, 1994

Annexe 1 : Une tentative de définition de la géotechnique

La géotechnique est la science qui se préoccupe des interactions d'un sol et d'un construit, route, bâtiment ou ouvrage d'art. A ce terme de géotechnique est parfois substitué celui de géosciences, mais cette dernière appellation apparaît moins éclairante sur la nature empirique de l'activité géotechnique, c'est-à-dire essentiellement fondée sur la relation à un contexte donné. Toute démarche géotechnique ne peut se concevoir qu'en fonction d'une situation réelle ou supposée, d'une contextualisation d'un problème, et ce généralement par analogie avec une situation similaire déjà connue. Cela traduit, de fait, l'activité géotechnique comme un ensemble de savoir-faire et de techniques, comparables au tour de main d'un artisan (Magnan 2002). Le terme de géosciences est pourtant utilisé parce qu'il est aussi révélateur d'une des deux caractéristiques fondamentales de ce domaine, à savoir son hétérogénéité de pratiques. De fait, les géosciences constituent l'ensemble hétérogène des sciences, ou éléments scientifiques, mobilisées dans la pratique géotechnique.

Hétérogénéités techniques

La géotechnique est en effet le point de rencontre de spécialités distinctes : géologie, mécanique des milieux continus et discontinus, chimie, hydraulique, etc... Chacune de ces spécialités implique une conception a priori des éléments géotechniques, une terminologie et des pratiques particulières. Par exemple, le concept, aussi fondamental soit-il, de sol, ne sera pas forcément compris a priori de la même manière par un géologue ou un spécialiste de stabilité des pentes. Pour le premier, un sol est un ensemble de roches situé sous la couche meuble de la surface. Pour le second, en revanche, cette couche meuble, parfois appelée sol détritique et situé au-dessus du substratum rocheux rigide, est l'objet de son étude – « son » sol.

Hétérogénéités régionales

Un deuxième élément fonde l'hétérogénéité de la géotechnique, à savoir sa disparité régionale ; les types de roches et les configurations de sol sont liées à des régions, ce qui donne donc naissance à des terminologies et des concepts régionaux : en France, le granit du mont d'Ambin est un exemple et évoque une forme particulière de granit et certaines de ses propriétés ; le Toki Granite, autre exemple,

est évoqué dans certains articles traitant de projets géotechniques au Japon et définit un autre type de granit, non caractérisé autrement que par sa provenance. Dans l'un et l'autre cas, les spécialistes peu familiers de la géologie alpine ou japonaise sont confrontés à des données incomplètes, car la dénomination particulière du granit implique l'importance de propriétés particulières. La généralisation des caractérisations régionales n'est pas nouvelle, et transparait assez facilement avec l'échelle des temps géologiques, le « calendrier géologique », où les étages et sous-étages géologiques portent les noms dérivés des régions typiquement représentatives des sols des époques correspondantes : Maestrichtien (Maestricht), Hettangien (Hettange), Coniacien (Cognac), etc... Ces travaux de dénomination et de calibrage correspondant sont souvent compliqués par le caractère évolutif de ces éléments qu'on pourrait croire figés : l'échelle des temps géologiques est ainsi remise à jour tous les quatre ans. Ces quelques exemples permettent de broser le tableau d'un environnement professionnel où les usages sont importants. Dans le cas de l'échelle des temps géologiques, cela correspond à l'enrichissement d'une colonne « autres dénominations », aux limites parfois mal identifiées, mais d'une manière générale cela correspond à l'enrichissement informel d'une polysémie déjà rendue redoutable par l'étendue du domaine géotechnique.

Afin d'illustrer la combinaison de ces éléments régionaux et terminologiques, citons enfin le cas typique de l'« argile de Mexico », qui, si elle est effectivement propre à Mexico, n'est pas de l'argile, mais un résidu détritique de roches volcaniques : synecdoque accolée à une régionalisation témoignant de la pérennisation d'une dénomination abusive mais couramment utilisée.

Un monde artisanal

Il y a comme une impossibilité de la connaissance pour l'abord d'un projet en géotechnique. On ne sait que bien peu de choses et tout est possible. Cet incertain est aussi source de contradictions, souvent apparentes et qui seront levées par une meilleure connaissance montrant que le processus de conception est itératif avec de nombreux retours sur le terrain.

Cette impossibilité de la connaissance n'est pas seulement le fait de la double hétérogénéité du domaine qui vient d'être décrite, mais aussi de la nature même des projets géotechniques. Ceux-ci commencent généralement par la conception d'un construit inexistant à implanter dans un environnement inconnu. Le sol, en effet, est un ensemble qui échappe aux méthodes d'observations usuelles, et notamment visuelles ; on ne peut en connaître les caractéristiques qu'au travers de tests et de sondages qui, pour aussi précis qu'ils soient, ne permettent qu'une approximation de la réalité. Par ailleurs, un sol donné n'est pas reproductible en laboratoire, et les sondages ont un coût élevé. De fait, un sol est un système complexe dont l'étendue physique elle-même est soumise à hypothèse, puisque ses caractéristiques peuvent être générées via des interactions à partir d'éléments dont il est a priori impossible de déterminer l'emplacement, la proximité ou même la nature. Quand à l'approche géostatistique elle est illusoire, le volume testé du sol, donc connu, correspondant dans les

meilleurs cas à une partie par milliard. (Magnan, Pouget, 1999)

On trouve là les raisons fondamentales de l'empirisme mentionné en premier lieu dans cet article, les causes de l'importance de l'expertise et la motivation principale pour le raisonnement par analogie, avec tous les risques qu'il comporte pour les projets de conception. C'est pourquoi la démarche pragmatique (au sens premier du terme, c'est-à-dire dont la validité est déterminée par les résultats) est très souvent utilisée en géotechnique.

Cet état de fait peut être résumé en une approche artisanale où chaque projet est une oeuvre unique.

Annexe 2 : Exemple de granules

Granule 1

Fragment de texte et référence

- Rec. AFTES GT9, collectif, 2001, TOS 168, pp134
- Le bituthène série 4000 peut être utilisé en étanchéité extradados si la température est supérieure à 5° C et si la protection est au moins équivalente à la classe 2 (voir article 7-4-2-3 du titre 3), sur dalle horizontale et sur paroi verticale sans charge d'eau.

Expression du granule, tous les syntagmes appartiennent à l'ontologie.

Si extradados et dalle ou paroi et température et charge d'eau et protection-(caractéristique classe de protection, référence article 7-4-2-3 du titre 3)

Alors (possibilité) bituthène série 4000

Modèle d'univers : extradados=dalle, extradados=paroi, (température>5°, unité-celsius), charge d'eau=0, classe de protection>=2

Les conditions exprimées dans le texte se retrouvent dans le modèle d'univers.

Granule 2

Fragment de texte et référence

- Recommandations pour la conception et la protection parasismique des ouvrages souterrains, collectif, 2001, TOS 167, pp250
- si un sable, un sable vasard, un silt possède les caractéristiques suivantes, degré de saturation voisin de 100%, granulométrie assez uniforme de coefficient de Hazen inférieur à 15, de diamètre à 50% compris entre 0.05 et 1.5 mm, sous un état de contrainte effective verticale inférieur à 0.20 MPa en zone 1a et 1b, 0.25 MPa en zone 2, 0.30 MPa en zone 3, alors risque de liquéfaction.

Expression du granule, tous les syntagmes appartiennent à l'ontologie.

Si sable ou sable vasard ou silt et degré de saturation Sr et coefficient de hazen Cu et diamètre à 50% et contrainte effective verticale et zone sismique 1a et 1b

Alors risque liquéfaction du sol.

Modèle d'univers : Sr>90%, Cu<15, D50>0.05 unité mm, D50<1.5 unité mm, Sigma'v<0.20 unité MPa

Ce granule sera triplé pour exprimer les conditions zone 2 et zone 3.

Annexe 3 : Développement de la base de connaissance

Nous avons montré les raisons qui font que le monde de la géotechnique, au sens large, nécessite une approche originale avec un rôle important des experts. Important ne veut pas dire long et fastidieux. C'est d'ailleurs le point fort de Ramcesh de ne requérir des experts qu'une courte implication, qui de plus, peut être éparpillée dans le temps et l'espace.

Ce sont les experts du domaine qui contribuent au développement de la base de connaissance de Ramcesh qui comporte deux bases de données, celle qui contient l'ontologie et celle qui contient les granules. Ces deux bases forment la base de connaissance.

Les paragraphes suivant décrivent ce développement qui comprend le choix des textes fait par un collège d'experts, la recherche des fragments de textes porteur de connaissance et leur traitement pour augmenter la base de connaissance de Ramcesh.

La base sera-t-elle complète un jour? Sans doute jamais car le monde de la géotechnique est très productif et de nouvelles connaissances apparaissent tous les jours. Cependant, le processus cumulatif de Ramcesh permet d'approcher cette complétude grâce à une collaboration partagée du travail de traitement des textes. On peut imaginer, un jour, traiter les textes dès leur parution comme aujourd'hui les services documentations traitent l'existence des articles des revues.

Le choix des textes

Le choix des textes à traiter est fonction de l'avancement de la mise en forme de la connaissance du domaine. Au début les textes choisis seront des textes riches en connaissances comme les cours, les recommandations et les règlements. Pour ces textes un travail en profondeur de formalisation a déjà été fait. Lorsque la base de connaissance sera importante on traitera des textes plus spécifiques comme les articles de congrès ou de revues. La collaboration avec un service documentaire est alors naturelle.

La recherche des fragments de texte

Traiter un texte consiste à reconnaître les formes d'écriture contenant une relation de causalité. Si (ce contexte) alors (ces conséquences). La forme si alors n'est pas toujours

explicite, on trouve fréquemment : dans ce cas...il se passe cela, qui est aussi une forme causale. A la fin du traitement on obtient une suite de fragments de texte contenant cette connaissance déductive.

Dans la pratique cette mise en évidence des fragments se fait avec un simple traitement de texte, l'expert élimine toutes les parties du texte ne correspondant pas à une forme causale. Le stylo scanner est aussi un bon outil de sélection pour les textes dont on ne possède qu'une copie papier.

La mise en granule et l'augmentation de l'ontologie

La transformation de ces fragments de texte en granules, avec mise à jour de l'ontologie, est fortement assistée dans le processus Ramcesh. Ce processus comporte les étapes suivantes:

Pour tous les fragments de texte

Vérification et complétude si nécessaire de l'ontologie vis à vis des syntagmes contenus dans les fragments de texte.

Pour chaque fragment de texte

Mise en forme du granule correspondant, puis recherche des granules voisins (comparaison de signatures de granule), alors choix entre l'ajout du granule à la base de connaissance, son rejet avec ou non addition de sa référence dans le granule le plus voisin, ou modification d'un granule voisin pour compléter prémisse ou conclusion.

Cette partie peut être centralisée ou répartie et dans ce dernier cas des échanges de savoir faire entre "granuleurs" sont très souhaitables.