
Notes de lecture

Rubrique préparée par Denis Maurel

Université François Rabelais Tours, LI (Laboratoire d'informatique)

Luc GRIVEL, La recherche d'information en contexte. Outils et usages applicatifs. Hermès-Lavoisier. 2011. 278 pages. ISBN 978 2746225817.

Lu par **Laurie ACENSIO**

Consultante R&D, société Altra Consulting-Pôle Formation.

Le développement d'applications basées sur des moteurs de recherche d'information implique un décloisonnement de l'exploitation des contenus au sein des systèmes d'information des entreprises. La mise en commun des contenus des documents, jusqu'alors autonome dans les applications documentaires, diminue l'hétérogénéité des sources d'information et amène à contextualiser le traitement de l'information. Par conséquent, l'accès contextuel à l'information modifie les pratiques de recherche supposant l'utilisation de nouveaux outils applicatifs. Cette nouvelle composante conduit le développement d'une ingénierie de la connaissance caractérisé par des modèles de représentation des connaissances (ontologie), méthodes de l'intelligence artificielle (raisonnement à partir de cas, traitement de langage naturel, fonction d'assistance...) et méthodes d'analyse des données (classification avec ou sans apprentissage). Cet ouvrage propose un éclairage sur les nouvelles problématiques sur la contextualisation de la recherche d'information, les solutions proposées et les retours d'expériences principalement issus du milieu industriel.

Le premier chapitre illustre un cas d'entreprise au sein d'EDF via un moteur de recherche d'information interne avec ses principales problématiques : indexation, recherche, restitution, gestion des droits et des profils utilisateurs. Pour améliorer l'accès à l'information, le moteur de recherche privilégie trois leviers : 1) devenir un composant essentiel de l'interaction entre les systèmes d'information et l'utilisateur en respectant les API des différents constituants rendant ainsi l'activité de recherche non intrusive ; 2) caractériser le contenu des documents textuels et multimédias via l'enrichissement de la description des documents (métadonnées) issus de ressources terminologiques et ontologiques ; 3) faire appel aux réseaux ou communautés pour faciliter le partage.

Le deuxième chapitre présente un état de l'art de la définition du contexte appliqué à la recherche d'information. Différentes approches dans la recherche d'information contextuelle dont la recherche interactive, la recherche adaptative et la recherche personnalisée sont abordées et illustrées par des exemples d'applications pour expliciter les distinctions.

Dans le troisième chapitre, l'apport du contexte des tâches est exploré dans des disciplines telles que l'intelligence artificielle et la recherche d'information. En effet, la recherche d'information se rapproche des agents assistants pour deux raisons principales : 1) la recherche d'information a besoin d'établir une relation avec ses usagers, cette relation devient socio-psychologique et s'appuie de plus en plus sur le traitement de la langue naturelle ; 2) les agents assistants (agents conversationnels) proposent des techniques issues du TAL (système de Q/R, système à base de clavardeurs, système de dialogues finalisés) pour éliciter le contexte de tâches.

Les travaux en recherche d'information contextuelle exploitent différentes formes de contextes (contexte de la requête, environnement de l'utilisateur, contexte des documents, contexte des interactions) afin d'adapter les réponses en fonction d'un contexte particulier. Le contexte est pris en compte 1) pour enrichir la requête (enrichissement par mots-clés ou par ontologie) ; 2) pour filtrer les réponses.

Les quatre chapitres suivants sont axés sur les apports de la sémantique dans la recherche d'information d'un point de vue utilisateur et concepteur.

Le quatrième chapitre décrit les principales fonctions sémantiques intégrées par des modules de traitement sémantique explicités par des techniques utilisées (lemmatisation, extraction d'entités nommées, analyse de sentiments, autosuggestion, correction orthographique, navigation à facettes...). Le projet « Infomagic » dédié à l'étude des interfaces homme-machine (IHM) a soulevé les problématiques des fusions des annotations notamment l'adaptation de l'architecture logicielle pour un gros volume de données et, d'un point de vue sémantique, les problèmes de normalisation des expressions numériques et temporelles.

Le cinquième chapitre aborde les approches interactives de la navigation à facettes au sein des systèmes de recherche d'information. En s'appuyant sur des exemples, les auteurs décrivent plusieurs approches pour identifier les facettes, classifier les éléments de la collection dans les facettes (suggestions de termes-clés, regroupement des résultats). L'utilisateur peut élaborer sa requête progressivement favorisant ainsi une meilleure compréhension et une interaction avec le système. Les méthodes de construction de facettes sont souvent manuelles et par conséquent coûteuses. L'automatisation des méthodes de construction devient un enjeu important, de même que des techniques émergentes, comme la puissance des ontologies, s'adaptent à la navigation pour décrire et manipuler les données.

Le sixième chapitre aborde les apports d'une ontologie de domaine dans un système d'information de services d'accès au contenu (expansion sémantique des requêtes, recherche à facettes, contextualisation des contenus) qui peuvent être issus des différents types de sources (les ressources terminologiques, les taxonomies de publication, de bases de connaissances). Les méthodes de modélisation diffèrent selon les sources mais restent indépendantes du domaine métier.

Le septième chapitre est un exemple d'application de recherche d'information mettant en jeu des techniques d'assistance issues de l'intelligence artificielle. Le projet présenté vise à réaliser des outils d'assistance sur des appareils nomades destinés au grand public et aux professionnels de santé pour proposer un plan alimentaire adapté à la situation du patient en fonction de ses objectifs et de ses habitudes alimentaires. Le système fait appel au raisonnement à partir de cas, nommé *Case-Based Reasoning* (ou CBR), et qui constitue un bon exemple de rapprochement entre des techniques et des méthodes issues des sciences de l'information et de l'intelligence artificielle. Le cycle du CBR s'apparente au cycle de recherche documentaire. La représentation d'un cas simple doit être élaborée et facilitée par une ontologie de domaine. Celle-ci va être utilisée pour pondérer et compléter les descripteurs importants et l'ensemble des règles définies sur le domaine.

Le huitième chapitre présente une application des techniques d'analyse de données utilisant des techniques de classification incrémentale et un paradigme de communication entre les modèles de classification nommé MultiView Data Analysis (ou MVDA). Le mécanisme de communication entre les modèles est fondé sur l'utilisation des informations susceptibles d'être partagées par les différents modèles de classification, comme les données associées aux classes. Le raisonnement bayésien peut s'appuyer sur la distribution des données ou sur la distribution de leurs propriétés dans les classes. Cette technique permet d'identifier des relations thématiques entre deux classes de deux points de vue différents. Dans son état actuel, la méthode proposée permet donc de visualiser les correspondances principales se manifestant entre périodes de temps par l'intermédiaire de noyaux stables (mots-clés représentant la classe).

La question de la démocratisation de l'information et de la gouvernance des systèmes d'information est abordée dans le dernier chapitre. Elle est examinée à l'échelle du Web à travers une critique des moteurs de recherches (non-respect de la vie privée, opacité du classement des résultats de recherche, manque de transparence technique). L'auteur décrit deux visions de la démocratie à l'œuvre dans les moteurs de recherche actuels : une conception orientée communauté et une conception orientée société.

Cet ouvrage témoigne de l'amélioration progressive de la composante du contexte en recherche d'information au sein de la gouvernance des systèmes d'information des entreprises. La recherche d'information contextuelle est un domaine de recherche récent dont l'ouvrage propose un état de l'art et démontre l'intérêt des techniques d'assistance pour pallier la problématique de la volumétrie dans l'espace de recherche.

Nous pouvons noter que de nombreux chapitres insistent sur l'émergence du Web sémantique et plus particulièrement sur les ontologies pour la modélisation de l'accès à l'information.

De nombreux exemples d'applications sont mis en valeur par des illustrations permettant ainsi d'avoir une approche pragmatique des techniques actuellement utilisées au sein des entreprises françaises.

Pollet SAMVELIAN, Grammaire des prédicats complexes : les constructions nom-verbe. *Hermès-Lavoisier*. 2012. 330 pages. ISBN 978-2-7462-2365-3.

Lu par **Evelyne JACQUEY**

ATILF – CNRS – Nancy

L'objet de l'ouvrage de Pollet Samvelian est de contribuer à la description, la classification et la modélisation des prédicats complexes du persan. Pour les travaux présentés, l'auteure se concentre sur les constructions nom-verbe. Partant des particularités de ces constructions en persan et après avoir examiné leurs propriétés syntaxiques et sémantiques, elle fait émerger un certain nombre de régularités qui lui permettent de proposer une modélisation ancrée dans le cadre des grammaires de construction et qui intègre une compositionnalité particulière dite a posteriori.

Les prédicats nom-verbe que Pollet Samvelian définit comme des prédicats complexes recouvrent deux types de constructions : les constructions à verbe support (1) et les constructions à objet nu (2) (sans détermination, quantification ou modification) qui se rapprochent par certains côtés des constructions à objet incorporé.

(1) Bačče-hâ bâ tup **bâzi** **kard-and**
enfant-PL avec ballon jeu faire. R2-3PL

'Les enfants jouèrent avec le ballon'

(2) Bačče-hâ **šâm** **xord-and**
enfant-PL dîner manger. R2-3PL

'Les enfants dînèrent'

Ces deux types de constructions existent dans de nombreuses langues, de même que les interactions entre les mécanismes qui permettent de les produire. En persan cependant, le nombre de verbes lexicaux simples se réduit à une centaine à partir du moment où l'on écarte les verbes inusités ou archaïques, appartenant à un registre très littéraire ou très familier. Même en ne pratiquant pas cette sélection, le nombre total de verbes lexicaux simples en persan reste toujours en dessous de 300, variant de 252 à 27 pour la plus grande liste proposée à ce jour par Khanlari.

La rareté des verbes simples en persan a deux conséquences pour ce type de constructions à verbe support et à objet nu. Premièrement, les constructions à verbe support sont beaucoup plus nombreuses que dans d'autres langues, comme le

français par exemple. Deuxièmement, la polysémie des verbes simples en persan y est très importante. De ce fait, d'une part, certains verbes simples très usités ont plusieurs emplois lexicaux et plusieurs emplois supports, et, d'autre part, dans certains cas, la polysémie est tellement importante qu'il est impossible de décrire et d'analyser le sens des verbes concernés en dehors des multiples combinaisons où ils apparaissent. Autrement dit, dans certains cas, tout se passe comme si les propriétés de ces constructions étaient idiosyncrasiques et avaient plus à voir avec des locutions ou collocations.

Pollet Samvelian a pour objectif d'expliquer de manière complémentaire l'existence de ces deux propriétés qui peuvent paraître contradictoires *a priori* : la productivité des prédicats complexes d'une part et leur caractère plus ou moins collocationnel ou idiosyncrasique d'autre part.

Pour rendre compte et expliquer cette productivité, l'approche de Pollet Samvelian intègre une certaine part de compositionnalité dite *a posteriori*. La compositionnalité mise en œuvre s'appuie sur l'idée que les prédicats complexes seraient des expressions conventionnelles ou à paradigme restreint, aux composantes desquelles il est possible d'associer un sens plus ou moins consistant dans une combinaison particulière ou dans un ensemble de combinaisons. Autrement dit, sur le plan sémantique, les prédicats complexes étudiés se situent dans un continuum allant de l'opacité à la transparence. Corollairement, pour tenir compte de la variabilité des emplois des verbes simples entrant finalement, pour les plus usités d'entre eux, dans un continuum allant de verbe support à verbe lexical, et pour rendre compte des apports variés du verbe et du nom dans les prédicats complexes qu'elle veut représenter, Pollet Samvelian adopte une définition large de la notion de verbe support. Elle y inclut bien sûr des séquences relevant clairement des constructions à verbe support comme *Maryam faryâd zad* (Maryam cri frapper. R2, 'Maryam cria'), des constructions qui s'en rapprochent mais n'en sont pas et s'assimilent davantage à des collocations comme *Maryam otâq=râ jâru zad* (Maryam chambre = ODD balai frapper. R2, 'Maryam balaya la chambre') et d'autres constructions où le fonctionnement collocationnel s'étend à une classe d'objets comme la combinaison entre *piâno* ('jouer') et le verbe *zadan* qui s'interprète invariablement comme 'jouer' avec n'importe quel nom qui dénote un instrument de musique.

Enfin, son étude se veut à la fois réaliste et empiriquement adéquate. C'est pourquoi, les travaux présentés s'appuient sur l'examen d'un large ensemble de données provenant de différentes sources : exemples construits ou élicités auprès des locuteurs, données issues de la presse écrite, de la littérature, ou relevées sur Internet.

Dans le premier chapitre, Pollet Samvelian s'intéresse à la question du statut des prédicats complexes d'un point de vue morphologique et syntaxique. Elle y contrecarre les arguments plaçant en faveur d'un rapprochement des prédicats complexes du persan avec le statut de mot (V0). À l'encontre de ce rapprochement,

Pollet Samvelian rappelle, relativement au fait que les prédicats complexes puissent être des bases morphologiques, par exemple, qu'en français des séquences telles que *mettre en scène* peuvent être nominalisées, *mise en scène* ou *metteur en scène*, sans que celles-ci soient considérées comme des mots d'un point de vue morphologique ou syntaxique. De plus, l'auteure conteste, données à l'appui, l'impossibilité d'insérer des éléments ou de procéder aux transformations syntaxiques classiques comme l'extraction ou la passivation. Elle précise en fait que ces opérations sont plus fortement contraintes mais pas impossibles.

Le deuxième chapitre est consacré à un examen comparatif des propriétés syntaxiques des prédicats complexes et des séquences formées par un verbe et son complément, examen qui démontre leur similarité syntaxique du point de vue des propriétés admises sans contrainte particulière : séparabilité, construction passive, coordination, ellipse du verbe et topicalisation. Trois propriétés du nom, sans être grammaticalement exclues, sont plus rares et plus contraintes (en terme de registre, d'incidences sur la prosodie ou la structure informationnelle) : position postverbale, complément du verbe à l'infinitif ou ellipse. La différence entre les séquences à objet nu et les prédicats complexes est donc d'ordre sémantique. Syntactiquement, l'élément nominal d'un prédicat complexe est sélectionné par le verbe et apparaît dans la structure argumentale de celui-ci au même titre qu'une combinaison nom-verbe ordinaire. Sur le plan sémantique, cependant, le prédicat complexe renvoie à une seule relation sémantique et les autres compléments syntaxiques éventuels sont régis par le prédicat complexe dans son ensemble.

Ayant montré que les prédicats complexes sont à considérer comme des combinaisons syntaxiques ordinaires dans les chapitres précédents, le troisième chapitre aborde la question de la compositionnalité, question centrale pour les prédicats complexes, d'un point de vue théorique et relativement à leur productivité. L'auteure examine en particulier deux grands courants qui se sont intéressés à la question : les approches projectionnistes (ou lexicalistes) et les approches constructionnistes. Bien que ces approches ne soient pas satisfaisantes (analyse uniquement des séquences compositionnelles au détriment des séquences quasi idiomatiques et trop peu de souplesse dans les apports respectifs du verbe et du nom), l'auteure fait l'hypothèse qu'une approche compositionnelle est possible *via* la compositionnalité *a posteriori*. Avant de l'introduire, elle procède d'une part à un examen détaillé des apports variables du verbe et du nom et, d'autre part, à une description des différents cas de combinaisons sémantiques (des séquences compositionnelles aux séquences idiomatiques).

Le quatrième chapitre expose la modélisation proposée dans le cadre des grammaires de constructions. Celle-ci passe par un recensement, pour chaque prédicat en usage, de ses propriétés, par un regroupement des prédicats en différentes classes en fonction de leurs propriétés, et enfin par l'association d'une construction à chaque classe définie. Les constructions associées aux classes jouent le même rôle que les règles de construction en morphologie dans le courant lexématique : elles servent de modèles à la production et à l'interprétation de

nouveaux prédicats complexes. C'est donc bien par cette modélisation que Pollet Samvelian explique la productivité des prédicats complexes du persan.

Le dernier chapitre est consacré à la mise en œuvre exhaustive de son approche à l'ensemble des prédicats construits avec *zadan*. Ce verbe a des emplois supports et lexicaux, la classe des prédicats construits avec *zadan* est importante (500 membres) et s'enrichit régulièrement. Enfin, ces prédicats manifestent des ensembles de propriétés cohérents qui permettent de les classer, ce qui est adapté à l'approche de l'auteur.

Pierre LÉVY, La sphère sémantique 1 : computation, cognition, économie de l'information, Hermès-Lavoisier, 2011, 415 pages, ISBN 978-2-7462-2506-0.

Lu par Maryvonne ABRAHAM

École Nationale Supérieure des Télécommunications de Bretagne

En ces temps d'incertitudes concernant les divers bouleversements sociétaux qui se profilent dans le monde et les promesses avancées par la révolution numérique, le livre de Pierre Lévy ne peut qu'éveiller notre intérêt. L'auteur annonce en effet ce qui nous attend dans un monde où s'accroît l'informatisation de données de toute nature. Nous assistons tous à la numérisation de données de plus en plus précises nous concernant, inscrites et multipliées dans de nombreux entrepôts de données, sans que nous sachions exactement ce qui est filtré ou reconstruit de notre profil dans de multiples systèmes de représentation des connaissances dont les objectifs nous échappent. Une inquiétude légitime se porte alors sur les conclusions ayant un impact sur nos avenir et qui peuvent émerger de traitements que nous n'imaginons même pas. Dans ce contexte, La sphère sémantique peut apparaître comme une voie explicative, ou bien comme une menace supplémentaire, selon qu'on adhère ou non à la construction au développement très élaboré présentée par Pierre Lévy.

Le parcours de l'auteur, qu'il nous présente comme une aventure intellectuelle transdisciplinaire, correspond à la quête intellectuelle de la génération qui a vu croître les ordinateurs, et aussi aux ambitions dont ces machines constituaient le support, concernant l'esprit, l'intelligence et les pensées. L'influence des ordinateurs dans la société n'a cessé de laisser la place à l'avènement d'un pouvoir cybernétique qui se glisse partout : dans la finance, où nous l'avons accepté sans imaginer les folies découlant des calculs que cela pouvait entraîner, il ne s'agissait au départ que de calculer plus rapidement. L'informatique s'est insérée dans la médecine et la santé, aussi bien dans l'administration de ce domaine que dans des données plus médicales, présentées en entrées vers des aides aux raisonnements automatiques suivis de prises de décisions plus ou moins automatiques également. On se sent inclus malgré soi dans des champs de connaissances où nous avons l'impression de nous voir nous profiler dans une silhouette dans laquelle nous avons peine à nous reconnaître.

La réponse à ces doutes se trouve peut-être dans ce livre épais, de 415 pages, constitué de quinze chapitres regroupés en deux grandes parties intitulées « Philosophie de l'information » et « Modéliser la cognition », et qui sont précédées d'une introduction générale. Une abondante bibliographie et un index, très précieux dans un ouvrage d'une telle ampleur, complètent avec profit cette publication.

Pierre Lévy justifie nos inquiétudes sociétales à travers une réponse portée par le développement d'une longue quête qui apporte un résultat en termes d'étape dans l'évolution : rejoignant Vernadsky et Teilhard de Chardin, il introduit un nouveau degré de l'évolution, dominé par la convergence des intelligences, et qui se nomme la « noosphère ».

Il est difficile de résumer la partie philosophique, préliminaire à la justification du nouveau langage proposé, tant elle est argumentée, organisant des intuitions que le lecteur avait sans doute plus ou moins entrevues, sans pouvoir les relier avec autant de pertinence. Chaque phrase soutient une réflexion intéressante. Mais la grande originalité s'inscrit dans la proposition de modélisation de la cognition : *fonder raisonnablement une connaissance scientifique de l'esprit, une connaissance qui exploite autant que possible les nouvelles possibilités d'enregistrement et de calcul ubiquitaires dans le médium numérique [...] accroître l'intelligence collective d'un métalangage de description, IEML (Information Economy Meta Language) : une mémoire numérique participative commune à l'ensemble de l'humanité est en voie de constitution...*

C'est donc la deuxième grande partie que nous allons un peu mieux visiter. L'auteur déclare appréhender la communication, vue non pas comme un échange entre deux interlocuteurs mais comme un monde (p. 191), rejoignant l'objectif affiché de construire la noosphère.

C'est dans le chapitre 7 que nous trouvons la définition de la sphère sémantique IEML, « système de coordonnées mathématiques spécialisé dans le balisage et la simulation d'écosystèmes d'idées »... le cœur en est la machine IEML : « il s'agit d'une société d'automates virtuels capables de produire, de transformer et de mesurer les circuits de la sphère sémantique ». Dans le métalangage IEML, « les concepts forment les nœuds et les relations syntagmatiques et paradigmatiques entre concepts forment les liens de ce gigantesque réseau. Comme chaque concept est traduit en langue naturelle¹ le métalangage IEML fonctionne comme langage pivot entre les langues et systèmes symboliques naturels ». Tout cela nous mène à de nouvelles perspectives en informatique et sciences humaines : « il faut considérer l'hypercortex comme le grand projet d'observatoire des humanités et des sciences sociales, comparable aux cyclotrons des physiciens ou aux agences spatiales de la conquête de l'espace ». L'auteur a fait le choix, pour IEML d'un langage régulier au sens de Chomsky (p. 210), et d'« un système de notation de type idéographique »

1. C'est nous qui soulignons, car c'est ici que notre communauté se trouve concernée.

puisque sa finalité est de coder et de manipuler le sens, contrairement aux notations de type phonétique dont la finalité est de coder le son.

Le chapitre 9 est une présentation générale de la sphère sémantique, organisée en un réseau de concepts qui constituent un métalangage, calculable, *pouvant servir de système d'adressage aux idées*. L'auteur décrit l'*idée* comme la conjonction d'un concept (une catégorie générale), d'un percept (une image complexe de type sensorimoteur) et d'un affect (une énergie d'attraction ou de répulsion). Les idées se présentent à l'esprit interconnectées en réseaux de relations. L'identification de chaque concept repose alors sur *le nœud de relations qu'il entretient avec les autres*, assurant le passage à un système calculatoire.

Le chapitre 10 nous présente le métalangage IEML, avec la question essentielle du codage des concepts, et de leur instanciation en langue naturelle. Ce chapitre 10, qui concerne le métalangage IEML, langage pivot, doit nous éclairer un peu plus sur les liens qu'entretiennent le langage et la sphère conceptuelle. On peut discuter le choix d'une grammaire formelle régulière qui est présentée par l'auteur comme une condition nécessaire de calculabilité. « Mais cette condition nécessaire n'est pas suffisante. Il faut coder les concepts de manière à utiliser utilement leur manipulation » (p. 270). Pour cela, « il faut en outre un type de codage ou de notation des concepts suffisamment compact et surtout suffisamment isomorphe à la structure des langues pour être automatiquement interprétable en langue naturelle » (p. 271) ...c'est-à-dire « un codage des concepts qui nous permette d'automatiser un calcul sémantique et non seulement un calcul arithmétique ou logique [...] Il existe une isomorphie entre les objets textuels et la topologie sémantique en IEML », concrétisée par une relation entre les objets textuels et leur représentation, qui conserve « le plus possible de traits pertinents de l'objet et notamment l'allure de ses transformations et de ses relations avec les objets de même espèce », afin de (p. 272) « se prêter de façon structurelle aux manifestations cognitives que nous effectuons sur les textes et leur sens », d'où la stratégie : « repérer d'abord la structure générale des opérations cognitives sur les objets linguistiques, puis intégrer cette structure dans un mécanisme opérant sur un langage régulier. [...] le métalangage IEML intègre les universaux de la structure des langues ».

Après avoir décrit très finement les multiples composantes observables de la langue, l'auteur aborde la question indispensable : entre texte et circuit, qu'est-ce que le sens ? Il fait alors appel au modèle de Melcu'k pour qui une langue naturelle peut se ramener à un ensemble de correspondances entre les sens et les textes de cette langue. Pour la machine sémantique naturelle, un texte est un arrangement de signifiants, il s'agit donc d'établir une correspondance entre chaînes de signifiants et circuits de signifiés ; une langue se présente ainsi comme un ensemble de règles qui font correspondre à un texte tous les sens possibles et à un sens tous les textes possibles. Face à cette profusion de correspondances, « Nous avons un besoin aigu d'un SUPERCERVEAU, c'est-à-dire de machines capables de penser à une échelle que l'homme seul ne saurait atteindre ».

Le chapitre 11 décrit la machine sémantique qui assure *la construction automatique du mégaréseau de la sphère sémantique et de la traduction automatique en langues naturelles de ses nœuds et de ses liens*. « La machine sémantique IEML ouvre la voie à l'émergence d'un hypercortex capable de réfléchir l'intelligence collective humaine en utilisant la puissance de stockage et de calcul du médium numérique. » La solution, l'hypercortex, est décrite au chapitre 12, par une structure en couches de réseaux de plus en plus larges et organisés en métadonnées sémantiques symboliques reliées par des liens conceptuels en un « mille-feuilles de transcodings », encore loin de son achèvement réaliste, mais qui suggère une dimension au-dessus des langues naturelles et la coopération des spécialistes des sciences de l'homme et de la société.

Le chapitre 15 clôt ce travail gigantesque, revenant sur l'opposition entre la pauvreté de nos moyens humains particuliers et le bouillonnement continu d'innombrables constructions intellectuelles : « si nous voulons produire des images scientifiques de l'esprit, il nous faut sans doute construire un instrument d'observation capables de canaliser cette complexité inépuisable. »

Notre présentation, émaillée de citations, s'est efforcée de rendre compte d'une réflexion menée depuis plusieurs années par l'auteur, à qui on peut reprocher un très grand optimisme eu égard au monde de conflits de pouvoirs intellectuels qui nous entoure, ainsi qu'un survol un peu rapide de ce qu'est la diversité des langues naturelles dans le monde. Les langues sont décrites, non pas comme des systèmes organisés, mais comme des données observées de couches de différentes catégories, que l'hypercortex devra organiser, mais il faut en attendre la réalisation.

J'avoue que je n'ai pas eu assez de temps pour lire et relire lentement, comme il conviendrait de le faire, cet ouvrage qui me paraît essentiel dans le domaine de l'analyse du monde cybernétique que nous voyons se construire et auquel nous participons, de gré ou de force. Les outils informatiques classiques sont en place, les concepts se représentent bien par les réseaux sémantiques, la théorie des graphes supporte la circulation sur ces réseaux. Des modèles de représentation de la langue écrite et orale sont choisis. En somme, les deux grandes parties, philosophique et informatique, doivent trouver leur articulation dans la description des concepts, réalisée par un métalangage, concepts obtenus à partir de données langagières. C'est ce passage des données langagières au métalangage pivot qui nous semble encore assez mystérieux dans le projet. Une fois ce passage réalisé, on peut aussi se poser la question de notre compréhension de la noosphère intelligente ainsi constituée, voir quelles idées auront été retenues, suivant quels critères, et surtout quelles sont les conséquences pour nous-mêmes de cette intelligence dont nous ne saurons pas juger les effets sur la direction de nos vies. « La machine sémantique IEML qui simule les jeux d'interprétation collective comme des flux de courant dans des circuits » aboutirait à une « démocratie organisée selon un strict principe de séparation des pouvoirs » (p. 362). On peut se poser les questions d'un avenir de l'humanité sous la chape d'une intelligence collective à laquelle nous aurons participé de différentes manières, pas toujours en étant conscient de l'impact des idées que nous aurons

apportées, peut-être malencontreusement, sur le résultat final. Que peut-il en émerger ? Et comment savoir que cela nous convient ? La présentation optimiste d'un niveau qui nous dépasse auquel nous contribuerons mais sans connaître vraiment l'emprise que nous pouvons avoir sur lui ni surtout celle que ce niveau aura sur nous grâce à une machine, peut inquiéter le lecteur qui répugne à se laisser entraîner dans une argumentation si brillante, c'est aussi pour cette raison que je recommande ce livre : il faut y trouver la place des individus et de leur emprise personnelle sur leur vie et, pour une description plus concrète, attendons le deuxième volume de cette recherche.