

L'avancement des travaux est affaire d'intégration

Dans le domaine du traitement informatique des langues, l'avancement des travaux est affaire d'intégration.

Sur le plan scientifique, afin de faire progresser les résultats de la science, il est nécessaire d'intégrer les connaissances développées dans des disciplines diverses comme la linguistique, l'informatique, les sciences cognitives et les sciences de l'information.

Sur le plan technologique, il faut accepter l'idée que les solutions aux problèmes ne sont pas seulement linguistiques, mais qu'elles peuvent provenir d'autres approches, statistiques, probabilistes, neuronales ou connexionnistes. Pour résoudre les difficultés éprouvées dans le traitement d'une langue donnée, il importe également de recourir aux approches multilingues, en plus d'employer les approches monolingues.

Enfin, sur le plan industriel, il convient, pour créer des produits performants et conviviaux, de combiner les habiletés techniques de spécialistes de la linguistique à celles d'experts en informatique, et de spécialistes en ergonomie et dans le domaine des interfaces.

Comme on le voit, l'avancement du secteur dépendra donc de l'intégration des efforts d'acteurs appartenant à différentes écoles ou disciplines, et de la complémentarité entre des technologies souvent perçues, à tort, comme concurrentes. Dans le domaine des industries de la langue, on a longtemps considéré que la linguistique était l'unique approche possible pour résoudre les problèmes inhérents au traitement automatique des langues naturelles ; il est temps de se rendre compte que, seule, l'hybridation des technologies donnera des résultats conformes aux attentes des usagers. On voit notamment les résultats de l'intégration des connaissances et des expertises dans le champ du traitement de l'information vers lequel convergent des spécialistes issus du monde de la bibliothéconomie, de la structuration du document électronique (SGML, XML, etc.) et de l'ingénierie linguistique.

L'intégration ou, au moins, la coordination des efforts des organismes publics sera aussi indispensable. À plusieurs reprises, on a parlé, pendant ce colloque, de projets visant à recenser, à valoriser et à rendre disponibles des ressources linguistiques, des corpus, des annuaires de compétence, des listes de projets, etc. Si on ne peut qu'admirer la multiplicité des projets en cours, en même temps, il faut

regretter le fait que lorsqu'un acteur francophone désire savoir ce qui est vraiment disponible, c'est quasiment mission impossible. De plus, s'il y a foisonnement d'initiatives nationales, régionales, multilatérales ou bilatérales, il y a aussi, dans de nombreux cas, des redoublements que l'on pourrait peut-être éviter.

Cela dit, différentes idées devraient guider les travaux des chercheurs du secteur du TIL dans les années à venir :

- premièrement, les interfaces de demain seront multimodales, c'est-à-dire qu'elles comporteront à la fois des composantes de traitement du texte, de la voix, de l'image, etc. Par conséquent, on ne peut plus travailler sur chaque question individuellement, en vase clos. Il faut plutôt penser en termes de complémentarité ;
- deuxièmement, les technologies linguistiques de demain devront être auto-apprenantes. Par exemple, le système permettant à l'acheteuse d'un téléphone cellulaire de dicter un numéro vocalement devrait être capable de s'habituer à elle. Il pourrait ainsi la reconnaître en toutes circonstances, par exemple même lorsqu'elle est enrhumée ;
- troisièmement, l'heure est au multilinguisme. Il faudra par exemple développer des outils permettant d'interroger une banque de données en une langue pour repérer un document écrit en une autre langue ;
- quatrièmement, il faudra continuer de travailler au traitement sémantique des connaissances. En effet, on passe progressivement d'une informatique de *contenant* à une informatique de *contenu*. Autrement dit, on ne peut plus traiter seulement le signal, par exemple les caractères composant un mot ; on doit aussi tenir compte du sens de ce dernier ;
- cinquièmement, les travaux devraient être menés en tenant compte des usages. Trop souvent, les développements techniques réalisés dans le domaine du traitement de la parole ou de l'écrit tiennent peu compte de la manière dont les gens utiliseront les outils mis au point. Il faudrait avoir une attitude plus modeste à l'égard de la technologie, admettre que les gens ne l'utilisent pas toujours comme il a été prévu.

*Stéphane Chaudiron,
chargé de mission Information scientifique et technique - Ingénierie linguistique, ministère de la Recherche, France.*