

Modata

Comparaison et complémentarité des approches par les données et par les modèles dans le cadre de la caractérisation d'un système technologique. Développement d'un cas illustratif.

Responsable : Bruno Pinaud – Labri

Co-responsable : Bruno Vallespir – IMS

Descriptif

L'approche par les données fait aujourd'hui florès dans le cadre de la caractérisation des systèmes, notamment technologiques. Cette approche consiste à profiter des nombreuses données récupérées sur le système en question et de mettre en œuvre des algorithmes qui permettent de mettre en évidence des corrélations entre ces données et, donc, entre divers éléments comportementaux de celui-ci. Mettant conjointement à profit le nombre important de données disponibles et les capacités de calcul des systèmes informatiques actuels, ces approches permettent d'obtenir des résultats intéressants dans bien des domaines.

D'un point de vue *métier*, ces approches ne manquent pas de poser question car l'ensemble des *connaissances* et des *modèles* accumulés sur ces systèmes, depuis des siècles pour certains, ne sont pas mobilisés dans ce type d'approches. La conclusion est-elle donc que les approches par les données rendent ces connaissances obsolètes ? Peu de scientifiques pensent comme ça et, comme partout ailleurs, la vérité doit être cherchée entre les deux.

Les connaissances permettent de comprendre les phénomènes observés. Elles permettent aussi d'aller beaucoup plus vite dans la résolution des problèmes. Par contre, les approches par les données permettent de combler les *trous dans la raquette scientifique* et, voire même dans certains cas, d'ouvrir de nouveaux domaines d'investigation scientifique.

L'avenir le plus raisonnable semble donc appartenir aux *approches hybrides*, qui combinent approches par les données et par les modèles et tireraient le meilleur des deux.

L'objectif de sujet proposé ici n'est, bien entendu, pas de répondre à la problématique qui vient d'être présentée. Le but est de proposer un exemple de comparaison entre les deux approches.

Pour cela, il sera d'abord proposé au stagiaire de choisir un système technique simple et de définir une problématique applicable à ce système. Par exemple, il est possible d'imaginer la détection de phénomènes conduisant à des risques de défaillance dans le cadre d'une politique de maintenance prédictive.

Il sera ensuite demandé au stagiaire de traiter la problématique identifiée par les deux approches : d'un côté, par l'usage des connaissances existantes sur ce type de systèmes et, d'un autre côté, par la mise en œuvre d'approches traitant les données captées sur le système.

Enfin, les deux modes de résolution seront comparés à travers un ensemble d'indicateurs à définir tels que la précision du résultat, la vitesse d'obtention de celui-ci, etc.

Le résultat constituera un *use case* illustratif de la thématique générale présentée en préambule.