

[Communiqués de presse](#)

Grâce à IBM, le National Institutes of Health accélère la découverte de médicaments et la recherche contre le cancer

Des entreprises spécialisées dans les biotechnologies collaborent avec IBM pour rendre publique une base de données consacrée à la chimie

Paris - 09 déc. 2011: IBM annonce aujourd'hui sa contribution à la mise en place d'une importante base de données consacrée aux données chimiques provenant des millions de brevets et des documents scientifiques du National Institutes of Health. Cette contribution permet aux scientifiques de visualiser plus facilement les liens importants entre les composants chimiques pour favoriser la découverte de médicaments et soutenir la recherche avancée contre le cancer.

En collaboration avec AstraZeneca, Bristol-Myers Squibb, DuPont et Pfizer, IBM fournit une base de données de plus de 2,4 milliards de composants chimiques issus de plus de 4,7 millions de brevets et 11 millions d'extraits de publications biomédicales de 1976 à 2000. L'annonce a été faite lors d'un forum IBM sur la compétitivité économique des Etats-Unis au 21ème siècle, explorant la façon dont les innovations et l'investissement du secteur privé peuvent être plus facilement partagés dans le domaine public.

L'utilisation de cette base de données publique apportera aux scientifiques du monde entier de nouveaux éclairages et de nouvelles pistes de recherche. Les scientifiques gagneront du temps en trouvant plus facilement l'information recherchée parmi les millions de pages consacrées aux brevets. Cela permettra également aux scientifiques d'analyser de plus grands nombres de documents qu'avec le processus manuel classique.

Les données ont été extraites grâce à Stratégic IP Insight Platform, une des solutions de BAO (Business Analytics and Optimization), mise en place via le SmartCloud d'IBM. Elle a été développée par la recherche d'IBM en collaboration avec les principales entreprises du secteur des biotechnologies. Cette nouvelle méthode pour classer et analyser l'énorme volume de brevets disponibles, les contenus scientifiques et les données moléculaires, fonctionne en mode cloud. Elle utilise l'analyse automatisée des images et une reconnaissance optique avancée des images chimiques et des symboles afin d'extraire les informations issues des brevets et des documents publiés. Une tâche réalisable aujourd'hui en seulement quelques secondes grâce à cette technologie, alors qu'il faudrait plusieurs semaines, voire des mois, pour la réaliser manuellement.

« L'infobésité continue d'être un défi pour la recherche de nouveaux médicaments et d'autres champs d'applications de la recherche scientifique », explique Steve Heller, directeur du projet pour la InChI Trust, une organisation à but non lucratif qui soutient le standard international InChI pour représenter les structures chimiques. « Les contenus riches sont souvent cachés dans les brevets, les dessins, les images ou les articles académiques. La contribution d'IBM et de ses collaborateurs consiste à faciliter l'usage de ces données par les scientifiques, et de les relier à d'autres données en utilisant la représentation de structure InChI pour en extraire de nouvelles connaissances. »

Sur les six dernières années, AstraZeneca, Bristol-Myers Squibb, DuPont et Pfizer ont travaillé sur ce projet avec la recherche d'IBM pour offrir un accès à une librairie chimique qui inclut les brevets de recherche sur les médicaments du monde entier. Les outils publics d'extraction de structure développés par les scientifiques de la National Institutes of Health furent également utilisés avec succès dans ce projet. En apportant leur expertise

dans les domaines de la pharmacie et de la chimie, ces sociétés vont soutenir cette base de connaissances en perpétuelle croissance.

« La communauté scientifique tirera un énorme avantage de cette avancée », explique Heller. « C'est un dispositif supplémentaire dans l'ensemble des données chimiques publiques. La compréhension des données et les nouvelles manières dont les scientifiques les appréhendent et les comparent avec d'autres données associées, devrait faciliter le développement de médicaments pour combattre les différentes formes de cancers et d'autres maladies humaines, et faciliter le développement d'autres composants médicaux. »

Les données seront à disposition de la National Center for Biotechnology Information (NCBI), une branche de la National Library of Medicine (NLM) et la Computer Aided Drug Design (CADD), du groupe de la National Cancer Institute (NCI) à la National Institutes of Health. Elles seront incorporées dans le NCBI PubChem, une ressource publique pour la communauté scientifique qui permet d'agrèger les résultats scientifiques, et dans les services du groupe NCI CADD tels que la Chemical Structure Lookup Service et la Chemical Identifier Resolver.

Le National Institutes of Health mettra le contenu à disposition sur le site PubChem

<http://pubchem.ncbi.nlm.nih.gov>

Plus d'informations sur IBM SIIP : www.ibm.com/gbs/bao/siip

A propos d'IBM <http://www.ibm.com/smarterplanet>

A propos d'IBM Life Sciences <http://www.ibm.com/lifesciences>
