

21^e Assemblée générale de CODATA, New Delhi, novembre 1998

Rapport sur les activités relatives aux données au Canada

On trouvera ci-dessous une liste, divisée en catégories, des activités relatives aux données qui ont lieu au Canada, telles qu'elles ont été signalées au Comité national canadien de CODATA (CNC/CODATA). On pourra obtenir de plus amples renseignements soit en obtenant l'URL ou l'adresse de courrier électronique qui figure dans la plupart des descriptions ou des rapporteurs cités avec leurs contributions respectives et présentés dans la section XI.

I. Sciences biologiques (M. Korab-Laskowska)

A. *Banques de données à accès public par Internet :*

1. **Organelle Genome Database (GOBASE)**

GOBASE est une base de données génomiques des organites à taxonomie étendue qui organise et intègre diverses données relatives aux organites. La version actuelle porte sur le sous-ensemble mitochondrial de données.

<http://megasun.bch.umontreal.ca/gobase/gobase.html>

2. **Protist Image Database (PID)**

PID fait partie du programme du génome de l'évolution moléculaire et des organites de l'Université de Montréal. Le PID offre, en direct, des images et de l'information sur la morphologie, la taxonomie et les relations phylogénétiques des protistes. La page Web de la PID offre aussi des liens à de nombreuses ressources en protistologie et dans des domaines connexes tels que la microbiologie, la mycologie, la phycologie et la protozoologie.

<http://megasun/protists/protists.html>

3. **Elegans Genetic Toolkit**

Le projet Genetic Toolkit est financé par une bourse du NIH National Center for Research Resources (NCR), accordée aux laboratoires de Ann Rose, de David Baillie et de Don Riddle (oeuvrant respectivement à l'University of British Columbia, à la Simon Fraser University et à l'University of Missouri). L'objectif du projet est d'offrir des «outils» génétiques visant à faciliter le clonage des gènes et l'analyse de leur fonction. La première étape a été la production et la caractérisation des réarrangements chromosomiques (appelés «balancers»), utilisés pour isoler et conserver les souches mutantes. On peut trouver sur le site Web l'information la plus récente qui soit sur les «balancers». Le projet aborde actuellement sa deuxième phase, qui vise à fournir des déficiences de chevauchement qui seront mises en parallèle avec les cartes génétiques et physiques.

<http://genekit.medgen.ubc.ca/gb.html>

4. Collection de cultures fongiques canadiennes

La collection de cultures fongiques canadiennes (CCFC) contient environ 10 500 souches de cultures fongiques représentant environ 2 500 espèces. La collection résulte de la consolidation graduelle de plusieurs collections de recherche individuelle. Elle est maintenant le dépôt principal de cultures fongiques pour la Direction générale de la recherche d'Agriculture et Agroalimentaire Canada et accepte en dépôt les lignées brevetées. Elle a aussi un mandat de banque de gènes pour les ressources génétiques microbiennes et fournit des cultures pures aux chercheurs en agriculture, foresterie, médecine, biotechnologie, tant du secteur privé que public. Plusieurs des échantillons de la collection sont des unicats et plusieurs sont nouveaux pour la science.

<http://res.agr.ca/brd/ccc/>

5. Annuaire des collections de cultures de micro-organismes au Canada

L'information recensée porte sur le nombre de collections, la diversité des exhibits, leur accessibilité, les modes de financement et les méthodes de préservation utilisées. On y distingue trois types de collections. Certaines contiennent un grand nombre de taxons et d'isolats tandis que d'autres renferment peu d'espèces mais constituent des collections nationales ou internationales de souches caractérisées importantes. La plupart de ces dernières ont bénéficié de soutien d'organismes en matière d'installation et d'exploitation. Les dernières, enfin, se caractérisent par le fait qu'elles appartenaient à des chercheurs et qu'elles subsistaient grâce à des budgets de programme ou des bourses de recherche universitaires.

<http://res.agr.ca/brd/ccc/ccfdir/ccfdire.html>

6. VectorDB

Cette base de données contient des annotations et de l'information séquentielle pour un grand nombre de vecteurs communément utilisés en biologie moléculaire. L'information sur plus de 2 600 vecteurs peut être recherchée à l'aide d'engins de recherche. Les vecteurs qui sont aussi dans la GenBank ont des liens directs à cette base de données par le biais du fureteur Entrez du NCBI.

<http://vectordb.atcg.com/>

7. Cystic Fibrosis Mutation Database

L'information contenue dans cette base de données est composée de l'information recueillie par le Cystic Fibrosis Genetic Analysis Consortium.

<http://www.genet.sickkids.on.ca/cftr/>

8. Le PAHdb

PAHdb contient des données sur le gène et les allèles de composés d'hydrocarbures aromatiques polycycliques. La base de données a été mise sur pied par une équipe de conservateurs du PAH Mutation Analysis Consortium.

<http://www.debelle.mcgill.ca/pahdb/>

9. Le serveur WWW de base de données de mutation de récepteurs androgènes

Cette ressource contient une base de données de mutation de gènes récepteurs androgènes, des cartes de mutation et des liens aux références et au site EMBL connexes.

<http://www.mcgill.ca/androgendb>

B. Organismes ou systèmes offrant l'accès à des bases de données internationales

10. Biologie moléculaire

Le Conseil national de recherches Canada (CNRC) a choisi Base4 pour fournir son service de base de données informatiques aux chercheurs en biologie. Le Système de base de données de biologie moléculaire (SBBM), établi en 1988 par l'Institut canadien de l'information scientifique et technique du CNRC, a été mis à niveau de façon exhaustive par Base4 et offert dans le cadre du service de la Mine du génome. En outre, Base4 offre un certain nombre de progiciels de service de premier ordre.

<http://www.base4.com/>

11. Programme de l'ICRA en biologie évolutive (CIAR-PEB)

L'Institut canadien de recherche avancée (ICRA) prête main-forte à un réseau de chercheurs partout au Canada et de pays étrangers. L'objectif du programme en biologie évolutive (CIAR-PEB) est d'utiliser la base de données comparatives de séquences génomiques, auquel ce projet contribuera, pour formuler des concepts quant à l'évolution du génome, de la cellule et de la population et pour élaborer des algorithmes d'analyse de structure/fonction moléculaire qui pourront être ultérieurement appliqués à des problèmes en biotechnologie, en diversité microbologique et en technologie génétique/génomique. La page d'accueil du CIAR-PEB contient de l'information au sujet des programmes et activités réalisés dans le cadre du Projet ainsi que des liens et des ressources du monde entier en biologie évolutive et par modélisation numérique des molécules.

<http://megasun.bch.umontreal.ca/ciar/>

C. Grands projets de séquençage dont les données sont accessibles au public

12. Données génomiques de *Sulfolobus solfataricus*

Le projet de séquençage du génome de *Sulfolobus solfataricus* est une entreprise de collaboration de sept laboratoires du monde entier : trois laboratoires canadiens (W. Ford Doolittle, Université Dalhousie; Robert Charlebois, Université d'Ottawa; Mark Ragan, CRNC-IMB) et quatre laboratoires européens (Roger Garrett, Université de Copenhague; John van der Oost, Université de l'agriculture de Wageningen; Michel Duguet, Université Paris-Sud; Ib Groot Clausen, Novo Nordisk, Copenhague). Le projet a été lancé au milieu de 1993 grâce à l'appui du Programme canadien de technologie et d'analyse de génome (CTAG) et aux contributions de l'Institut canadien de recherche avancée (ICRA), du Conseil national de recherches Canada (CNRC) et du Conseil de recherches médicales du Canada (CRM). À la suite de la dévolution de CTAG, le projet a été poursuivi grâce au financement de l'Union européenne BIOTECH.

L'interaction entre les laboratoires est principalement réalisée par l'Internet. Les données sont transférées directement à un environnement UNIX à l'IBM pour le prétraitement et l'analyse par la base de données depuis les quatre laboratoires (Copenhague, Wageningen, Orsay et Halifax).

Le calcul des amorces, la recherche dans la base de données et les annotations sont automatisés à l'aide d'un logiciel conçu et programmé à IBM sous la supervision de Christoph W. Sensen (amorces), ou en collaboration entre MM. Sensen et Terry Gaasterland (Laboratoire national d'Argonne et l'University of Chicago; depuis 1998 à la Rockefeller University). Les installations informatiques sont celles de la Ressource de bioinformatique Canada (CBR-RBC) du CNRC. Les données traitées sont diffusées aux laboratoires par l'entremise d'un réseau protégé. Depuis le milieu de 1998 le projet est à plus de 80 % terminé.

(mark.ragan@nrc.ca)

13. Organelle Genome Megasequencing Program (OGMP)

L'OGMP repose sur une collaboration interdisciplinaire de sept groupes de recherche canadiens de l'Est du Canada, chacun de ceux-ci étant intéressés à l'évolution moléculaire, particulièrement aux mitochondries, aux plastes et aux bactéries. Ce projet concerté, qui reçoit l'appui du Programme canadien de technologie et d'analyse de génome (CTAG) gravite autour de la phylogénie des organites et prévoit l'établissement d'une installation de séquençage centralisée (Unité de mégaséquençage), qui constitue un centre névralgique de recherche. L'Unité de mégaséquençage est située à l'Université de Montréal. La division de la bioinformatique de l'OGMP est responsable du traitement et de l'analyse des données. Les séquences des génomes de mitochondrie de l'Unité de mégaséquençage seront diffusées à la communauté scientifique par l'entremise de GenBank et de GOBASE.

<http://megasun.bch.umontreal.ca/ogmproj.html>

14. Fungal Mitochondrial Genome Project (FMGP)

Le FMGP, un projet du groupe de B. F. Lang (Département de biochimie, Université de Montréal), reçoit l'appui du Conseil de recherches médicales du Canada (CRM). L'objectif du FMGP est d'établir la séquence génomique intégrale des mitochondries de l'ensemble des principales lignées fongiques, de conclure le défrichage de la branche fongique de l'«arbre de vie» et d'examiner l'expression des gènes, les introns et les éléments mobiles des mitochondries. Les pages Web du FMGP comprennent de l'information sur les généralités quant aux organismes, les cartes génétiques, le séquençage intégral, la phylogénie, etc.

<http://megasun.bch.umontreal.ca/People/lang/FMGP/FMGP.html>

II. Chimie (P. Mezey)

A. LOGKOW - Banque de données sur les coefficients de partage n-octanol/eau.

Monsieur James Sangster, des Laboratoires de recherche Sangster, à Montréal (Québec) a établi et coordonne une banque de données sur les coefficients de partage n-octanol/eau pour un grand nombre de molécules, données qui sont importantes pour de nombreux domaines de la chimie et de la biochimie. Ces données sont essentielles à l'établissement de comparaisons et à la formulation de prédictions de l'activité biochimique des molécules de médicaments théoriques ainsi que de contaminants de l'environnement.

(james.sangster@mail.polymtl.ca)

B. Canadian Domestic Substances & Non-Domestic Substances Databank.

TerraBase Inc. a lancé sur CD-ROM sa liste de substances intérieures canadiennes et étrangères (DSL&NDSL) sur CD-ROM. Cette banque de données recense plus de 66 400 substances visées par la Loi canadienne sur la protection de l'environnement (LCPE). Toxicity Data & QSAR, partie intégrale de la suite de progiciels TerraTox / TerraFit, contient des notices sur plus de 8 000 composés, dont plusieurs sont connus comme des carcinogènes, des mutagènes, des pesticides ou des dislocateurs endocriniens ou soupçonnés de l'être. Elle contient aussi actuellement plus de 5 000 références à des publications scientifiques récentes sur les rapports structure-activité quantitatifs (QSAR) et des sujets étroitement reliés (y compris les sources de données). (klaus.kaiser@cciw.ca)

C. Base de données de toxicologie moléculaire, Environnement Canada.

Pour obtenir plus d'information, vous pouvez communiquer avec M. Mark Lewis, Direction de l'évaluation des substances chimiques commerciales, Environnement Canada, Place Vincent Massey, 14^e étage, 351, boul. St. Joseph, Hull (Québec) K1A 0H3, Canada, tél. : 819-953-7199; télécopieur : 819-953-4936; courrier électronique : mark.lewis@ec.gc.ca

D. Base de données sur les hydrocarbures aromatiques polycycliques.

Les professeurs Bruce Greenberg et G. Dixon, de l'University of Waterloo (Ontario), conservent et enrichissent une banque de données sur les activités photochimiques et la toxicité aquatique des hydrocarbures aromatiques polycycliques ainsi que des produits de leur photo-oxydation. On s'attend à ce que les données sur la toxicité et les propriétés chimiques signalées dans cette base se révèlent utiles autant pour les laboratoires universitaires que pour l'industrie chimique, en tant qu'outils d'évaluation des risques toxicologiques.

(Professeurs Bruce Greenberg et G. Dixon, Département de biologie, University of Waterloo, Ontario.)

E. Base de données sur les activités de pesticides et d'herbicides en présence de contaminants métalliques.

Le professeur P. Ming Huang, du Département de la science des sols de l'University of Saskatchewan, à Saskatoon, conserve et enrichit une base de données sur les activités de pesticides et d'herbicides en présence de contaminants métalliques et ayant une incidence sur l'interface sol - système racinaire. Axée sur des techniques de recherche de similarités, la base de données devrait permettre d'améliorer la prédiction des effets néfastes des nouveaux pesticides et herbicides lancés sur le marché.

(Professeur P. Ming Huang, Département de la science des sols de l'University of Saskatchewan, Saskatoon, SK Canada)

F. Base de données d'affinité de fixation de l'hémoglobine.

Le professeur Kannan Krishnan, du Département de médecine du travail et d'hygiène du milieu de la Faculté de médecine de l'Université de Montréal (Québec), conserve et enrichit une base de données sur les constantes d'affinité de fixation de l'hémoglobine pour un grand nombre de

molécules organiques. On a déjà recours à cette base pour examiner les effets néfastes de substances toxiques.

(Professeur Kannan Krishnan, Dép. médecine du travail et d'hygiène du milieu, Faculté de médecine, Université de Montréal, Québec.)

G. Base de données sur la toxicité des métaux.

Les professeurs Beverly Hale, University of Guelph (Ontario) et Francine Denizeau, Département de chimie de l'Université du Québec à Montréal (Québec), conservent et mettent à jour une base de données sur la toxicité des métaux, y compris le cadmium et le zinc. On s'y intéresse en particulier à l'assimilation par les différents types de grains.

H. Base de données sur la densité du groupe d'électrons fonctionnel des composés carbonylés carcinogènes.

Monsieur Serge Lamy, Santé Canada et le P^r Mezey, de l'University of Saskatchewan sont à constituer une base de données de densité d'électrons du groupe fonctionnel de composés carbonylés carcinogènes qu'on retrouve dans l'échappement des véhicules.

(mezey@sask.usask.ca)

I. Base de données de densité des molécules d'électron organiques halogénées

Le professeur Mezey, de l'University of Saskatchewan conserve et enrichit une base de données de formes moléculaires d'une série de molécules organiques halogénées. La base de données de formes d'hydrocarbures aromatiques polycycliques plus ancienne est continuellement mise à jour. Ces bases de données de formes ont de nouvelles applications dans l'industrie pharmaceutique, dans la nouvelle recherche de pointe, dans l'évaluation du risque toxicologique dans le cadre de l'évaluation du risque quantitatif du RCCT (Réseau canadien des centres de toxicologie).

(mezey@sask.usask.ca)

III. Cristallographie (G. Wood)

A. Base de données cristallographiques des métaux du CNRC (CRYSTALMET)

Aux termes d'une licence exclusive du CNRC, Toth Information Systems non seulement conserve et met à jour la base de données mais a mis au point deux nouveaux produits sur CD-ROM appelés CRYSTALMET et CRYSTALMET-P. La version CD de CRYSTALMET est dotée d'un nouveau système de recherche et d'une série d'outils d'analyse; CRYSTALMET-P est semblable à la base de données CRYSTALMET de base mais contient aussi des modèles calculés de poudre. Les deux produits ont été lancés le 1^{er} octobre 1998.

(<http://www.nrc.ca/programs/toth/>)

B. Cambridge Structural Database (CSD)

La base de données CSD est distribuée au Canada par le P^r George Ferguson à l'University of Guelph. Dix-neuf sites universitaires canadiens ont maintenant leur propre système CSD. Les CD-ROM CSD sont distribués aux sites à la mi-avril et à la mi-octobre chaque année. L'accès à la

base de données CSD est ensuite possible pour les groupes compris dans la licence multi-usager pertinente à chaque université.

(george@angus.chembio.uoguelph.ca)

IV. Sciences de la terre (R. Berman)

A. Répertoire des communications géoscientifiques canadiennes de la CGC

Le répertoire des publications géoscientifiques canadiennes (<http://ntserv.gis.nrcan.gc.ca>) est constitué par la CGC afin de fournir un accès Internet «à guichet unique» à toutes les publications géoscientifiques au Canada. L'objectif est de permettre à l'utilisateur d'effectuer de la recherche de texte et des interrogations spatiales interactives. Actuellement le répertoire est composé des cartes du CGC et de publications géoscientifiques de Terre-Neuve. D'autres types de publications de la CGC, ainsi que les données d'autres provinces, seront bientôt offertes.

B. Base de données bibliographiques de Ressources naturelles Canada

GEOSCAN est une base de données bibliographiques renfermant les publications de la CGC de Ressources naturelles Canada. En plus de 40 000 fiches, elle contient des données bibliographiques, géographiques et de contrôle thématique de toutes les publications de la CGC et de plusieurs de ses contributions à des publications externes. Les clients peuvent accéder à GEOSCAN à (http://www.nrcan.gc.ca/ess/esic/esic_e.html (ang.) ou http://www.nrcan.gc.ca/ess/esic/esic_f.html (fr).

C. Programme national de cartographie (NATMAP)

Il s'agit d'une initiative géoscientifique importante, créée en 1988 par la CGC, visant à accroître le niveau de cartographie géoscientifique au Canada par le biais de projets multi-organismes et multidisciplinaires. Les bases de données associées à chaque projet sont constituées de cartes géologiques du substratum, de cartes géologiques superficielles, de cartes de base topographiques, ainsi que de dépôts minéraux géophysiques, télédéfectés et des données géochronologiques. Certains ensembles de données NATMAP peuvent être obtenus pour la province des esclaves (<http://gis.nrcan.gc.ca/natmap/slave/slave.html>) et la marge du bouclier (<http://gis.nrcan.gc.ca/natmap/shield.html>).

D. Base de données géothermodynamiques

Une base de données uniforme des propriétés des formes extrêmes et des solutions minérales est conservée par la Division de la géologie du continent (DGC) de la CGC par le biais du progiciel de thermobarométrie TWQ (<http://www.gis.nrcan.gc.ca/twq.html>). La version la plus récente de cette base de données englobe de nouvelles mesures expérimentales sur la solubilité Ti et Al dans le biotite et Al dans l'orthopyroxène.

E. Logiciel de levés de terrain assistés par ordinateur

La DGC de la CGC a continué à développer un progiciel sur PC visant à aider aux levés de terrain assistés par ordinateur et à l'élaboration rapide de bases de données de sources ponctuelles. Le logiciel Fieldlog v3.0 (<http://gis.nrcan.gc.ca/fieldlog/Fieldlog.html>) comprend un certain nombre d'options très utiles, notamment : le traçage des symboles géologiques, les stéréonets, les graphes en étoile; des interrogations de base de données perfectionnées; des systèmes de coordonnées personnalisables; des projections; des transformations entre les systèmes de coordonnées; l'intégration avec le GPS et le logiciel d'enregistrement mobile.

F. Base de données du Nord de l'île de Baffin/presqu'île Melville

Une échelle numérique 1:500 000 sans coupure intégrée à une cartographie géoscientifique de la région du Nord de l'île de Baffin/presqu'île Melville est en voie d'être créée à la CGC. Cette base de connaissances géoscientifiques numériques (<http://gis.nrcan.gc.ca/baffin>) comprendra toute la cartographie géologique du substratum rocheux et superficiel à l'échelle régionale antérieure, les données géophysiques, les gîtes minéraux, la géochimie rocheuse et une base de données de géochronologie. Un fureteur de carte convivial permet d'effectuer des interrogations spatiales en tant qu'aide à la visualisation et à l'extraction d'information géoscientifique particulière.

G. La Section de levés géophysiques aériens de la Division des ressources minérales (DRM) de la CGC

La DRM tient à jour la base de données nationale de spectrométrie des rayons gamma aériens, constituée d'environ 1,5 million lignes-km de données spectrométriques de rayons gamma à multivariées (K, U, Th, exposition, Mag et VLF), surtout dans la région du Bouclier canadien. Les quadrillages de levés sont de 25 km, 5 km, 1 km et <250-500 m. Il est possible d'accéder aux données à natgam@gsc.nrcan.gc.ca ou <http://www.geophys.gsc.nrcan.gc.ca/airgs>

Base de données de reconnaissance géochimiques de la Division de la DRM de la CGC

La base de données de reconnaissance géochimique nationale (<http://gds.agg.gsc.nrcan.gc.ca>; courrier électronique : mmccurdy@nrcan.gc.ca) contient des données sur le terrain et des données analytiques de sédiments de cours d'eau et de lacs et des échantillons d'eau. Les données ont été recueillies de levés effectués depuis 1974 pour des régions choisies au Canada. Plus de 200 levés ont été menés à terme, représentant presque 200 000 échantillons sur environ 2 100 000 kilomètres carrés. Les levés ont été effectués à l'aide de méthodes d'échantillonnage et d'analyse uniformes.

H. Bases de données de la Division géoscientifique de l'Atlantique de la CGC

La Division de l'Atlantique de la CGC est le dépôt principal de données géophysiques marines (acoustiques, magnétiques, sismiques et gravitationnelles) d'échantillons instantanés et de carottes de sédiment et de collections rocheuses et paléontologiques obtenus dans le cadre d'une collaboration gouvernement/industrie dans les régions au large du Canada. Les bases de données (<http://agcwww.bio.ns.ca/pubprod/pubprod.html>) sont constituées de 1,3 million lignes-km de données de réflexion et de réfraction sismiques, de 8 000 km de données de réflexion sismique en profondeur, de 1,7 million d'observations gravitationnelles côtières, de 3,5 millions d'observations magnétiques embarquées au large, de 10 000 prélèvements obtenus lors de plus de

900 expéditions, de 250 000 échantillons de sédiments du fond marin et de roches, 800 fiches de forage maritime côtier, d'une base de données biostratigraphiques, géographiques et taxonomiques contenant 825 000 notices, et BASIN, une base de données d'information géologique et d'ingénierie sur plus de 300 puits de pétrole et d'exploration au large.

I. L'Inventaire des bases de données de ressources côtières de l'Atlantique

Accessible à (<http://www.ndi.nf.ca/ndi/aczisc/index.html>), l'inventaire dresse une liste descriptive de 608 bases de données pertinentes à la gestion intégrée et au développement viable des ressources côtières de l'Atlantique du Canada. Les descriptions dans la base de données comprennent des détails comme les contacts, la disponibilité, les formats, les géoréférences et les échelles, etc. Les bases de données sont tenues à jour par les gouvernements (fédéral, provinciaux et les administrations municipales-régionales), les établissements d'enseignement et de recherche, le secteur privé, les organismes non gouvernementaux et les organismes à but non lucratif.

J. Base de données de gîtes minéraux du ministère de l'Énergie et des Mines de Colombie-Britannique :

MINFILE (<http://www.ei.gov.bc.ca/geosmin/minfile/minfile.htm>) contient de l'information sur la géologie, les emplacements et les aspects économiques de plus de 11 900 dépôts et gîtes de métaux, de minéraux industriels et de mines de charbon en Colombie-Britannique. La base de données MINFILE se prête à la recherche, à la préparation de rapports et à la mise à jour à l'aide du progiciel MINFILE/pc. Le programme, les données et la documentation d'utilisateur peuvent être téléchargés à partir de ce site.

V. Géophysique (R. Coles)

A. Données géomagnétiques

Le Programme national de géomagnétisme de la Commission géologique du Canada (CGC) conserve l'archive des données numériques des observatoires magnétiques du Canada, décrivant les variations dans le temps du champ magnétique de la terre à certains points partout au Canada. Ces archives, d'environ 6 Gbits, contiennent des données à haute résolution recueillies par 13 observatoires au cours des 22 dernières années ainsi que des données historiques remontant aussi loin qu'à l'Année géophysique internationale et même plus loin encore. Les deux années de données les plus récentes sont présentées en direct. L'accès est offert aux chercheurs et aux intéressés du monde entier. Un courrier électronique géré par robot (DRM) est maintenant en fonction et les demandes personnalisées peuvent être traitées par le biais du ftp de l'Internet. La documentation sur l'acquisition des données, la disponibilité des données et les conditions d'accès se trouve sur le site Web à : www.geolab.nrcan.gc.ca. Les données plus âgées que huit jours peuvent être visualisées sous forme graphique sans restriction sur le Web. Il faut une autorisation spéciale pour visualiser les données plus récentes, et certaines conditions spéciales s'appliquent.

B. Données sismologiques

Dans le cadre du Programme de sismologie de la CGC les archives du Réseau sismographique national canadien sont tenues à jour. Les archives contiennent un très grand nombre de données sismographiques analogues datant d'aussi loin que le début des années 1900. Les archives de données récentes contiennent des données numériques recueillies depuis 1980, ce qui représente un téraoctet de séries chronologiques. On y retrouve aussi des données numériques datant de 1966 de l'ensemble sismographique de Yellowknife, utilisées pour les surveillances des essais nucléaires et des données dérivées de premier niveau sous forme d'épicentres de tremblement de terre pour le Canada. Un robot de diffusion de l'information par courrier électronique (DRM) a récemment été installé et celui-ci est beaucoup plus utilisé. Le système prévoit des liens directs avec le Centre international de données sismiques, situé à Washington DC. On trouve sur le site Web (www.seismo.nrcan.gc.ca) des données dérivées telles que les épicentres, de l'information à jour sur les récents tremblements de terre, un catalogue de disponibilité des données et des tracés de données de formes d'ondes pour les événements choisis.

C. Base de données aéromagnétiques

La section de la géophysique régionale de la CGC coordonne la Base de données aéromagnétiques nationale du Canada. La couverture commence en 1947 et les cartes analogues originales ont été converties en format numérique. La base de données décrit environ 80 % du Canada à l'échelle régionale, ce qui occupe environ 7 Gbits. Les données sont accessibles sous plusieurs formes : valeurs ponctuelles, valeurs réparties sur grille, tracés à n'importe quelle échelle, en n'importe quel format et par n'importe quel média, ou par le biais d'un transfert de fichier électronique. Un système de commande en direct est en fonction et les détails complets des données et des services peuvent être obtenus au site Web : (gdcinfo.agg.nrcan.gc.ca).

VI. Géospatiale (R. Tomlinson)

A. Accès aux données géospatiales

En collaboration avec d'autres administrations, Ressources naturelles Canada et le ministère de la Défense nationale sont à mettre au point les composants techniques clés de l'Infrastructure canadienne des données géospatiales (ICDG) qui permettra aux Canadiens d'accéder à une vaste quantité de données géospatiales par l'Internet. CEONet, un projet de l'ICDG, est à mettre au point un centre de répartition exhaustif à l'intention des fournisseurs et des utilisateurs des données et services géospatiaux. À ce jour plus de 4 900 fonds de données internationales sont inscrits dans CEONet dont environ 250 sont canadiens. Les fonds de données canadiens englobent bon nombre des fonds de données nationaux importants, y compris la Base nationale de données topographiques, les archives nationales de télédétection, la Base nationale de données de reconnaissance géochimique et plusieurs autres. CEONet offre des capacités de navigation et de recherche dans ces fonds de données et fournit un mécanisme pour rechercher directement les inventaires détaillés, p. ex., des feuilles topographiques uniques.

Bien que la plupart de ces données soient encore livrées en différé, bon nombre d'entre elles peuvent être téléchargées directement par le réseau. Un exemple important : l'Inventaire des

terres du Canada dont les données sont constituées de six thématiques différentes pour le tiers du Sud du Canada.

L'information sur CEONet se trouve à <http://cgdi.gc.ca/ceonet>. Les autres adresses de courrier électronique direct pour les fonds de données géospatiaux au Canada sont les suivantes :

Cartes et publications en aéronautique :	mél : aero@nrcan.gc.ca
Levés cadastraux :	mél : lst@NRCan.gc.ca
Levés géodésiques :	mél : information@geod.NRCan.gc.ca
Imagerie de télédétection :	mél : orderdesk@ccrs.nrcan.gc.ca
Atlas national :	mél : atlasinfo@ccrs.NRCan.gc.ca
Topographie numérique :	mél : bndt@CCG.NRCan.gc.ca
Topographiques :	mél : topo.maps@NRCan.gc.ca
Toponymie :	mél : geonames@NRCan.gc.ca
Photographies aériennes :	mél : NAPL@GeoCan.NRCan.gc.ca

B. Cadre de données géospatiales

Le cadre de données consiste en une assise de données géospatiales fiables qui permet d'intégrer des données à valeur ajoutée, le développement d'applications ou de collections de données détaillées. En tant qu'une des thématiques de l'ICDG, le cadre de données fournira des emplacements de référence précis en tant que base pour d'autres données de l'ICDG. De cette façon, les utilisateurs des données géospatiales pourront géocoder les données de diverses sources et à diverses échelles. Le cadre de l'assise sera principalement composé d'une Couche de cohérence des données, à laquelle on pourra accéder librement au milieu de 1998. Ressources naturelles Canada a été mandaté pour produire et mettre en oeuvre la Couche de cohérence des données de l'ICDG sur l'Internet conformément aux recommandations de l'ICDG.

La Couche de cohérence des données de l'ICDG (CCDI) contient des points faciles à identifier à diverses échelles cartographiques qui permettent de géoréférencer des données numériques (vecteur et trame) provenant de différentes sources. Les coordonnées géographiques utilisées pour définir les positions des points de la CCDI sont les degrés et les degrés décimaux de latitude et de longitude. L'unité de mesure utilisée pour sauvegarder les positions des points de la CCDI en coordonnées cartésiennes (X,Y) est le mètre. L'unité de mesure utilisée pour sauvegarder les valeurs angulaires est le degré (entier). La résolution des points CCDI est d'un mètre. Chaque point de CCDI compte au moins un jeu de caractéristiques qui définit son entrée dans la CCDI. Toutes les options de la CCDI sont des points de l'un des trois types suivants : intersection de ligne, centre de masse ou point. L'information sur la Couche de cohérence des données de l'ICDG se trouve à <http://www.ctis.nrcan.gc.ca/~cdal>.

C. Normes de données géospatiales

Le domaine des normes de données géospatiales a connu un niveau significatif d'activités au Canada au cours des deux dernières années. Le Canada a participé activement aux travaux de

normalisation du Comité TC211 d'information géographique/géomatique de l'Organisation internationale de normalisation (ISO) ainsi qu'aux autres travaux d'élaboration de normes géospatiales internationales ISO JTC1 SC32 WG4 SQL/MM – Spatial, ISO TC204 transport routier, de l'Organisation hydrographique internationale et du Digital Geographic Information Working Group (DGIWG) de l'OTAN. En outre, plusieurs entreprises canadiennes ont participé au consortium ouvert du SIG.

Les comités de l'OHI, du DGIWG et du TC204 transport routier ont élaboré des normes de produits qui décrivent les données utilisées dans des champs d'application particuliers. La norme S-57 de l'OHI est centrée sur la navigation maritime sécuritaire et est unique en ce qu'elle intègre une mise à jour dynamique des données pour maintenir la validité des cartes. Le comité DGIWG a une portée plus large, couvrant toute une gamme de données militaires. Ces normes ont été mises en oeuvre au Canada et les données sont disponibles. Le Canada a déployé des efforts significatifs pour harmoniser ces normes. La norme sur le transport routier est presque achevée et il y a un engagement commercial à la mettre en oeuvre. L'OGC représente un consortium industriel composé des principaux fournisseurs du SIG. Les travaux des comités ISO JTC1 SC32 WG4 SQL/MM – sur la partie 3 (Spatial) s'intéressent à la manière de sauvegarder les données spatiales et aux moyens d'accès aux bases de données plus grandes.

ISO TC211 veille à l'élaboration d'une gamme de normes à composants communs qui peuvent être utilisées comme pierre d'assise sur laquelle échafauder les autres normes. Les normes TC211 ISO 15046 seront mises en oeuvre à mesure que les profils et les normes comme la norme DIGEST, pourront reposer sur cette pierre d'assise. Les efforts d'harmonisation ont déjà produit des résultats significatifs : par exemple, l'OGC, SQL/MM – Spatial, DGIWG et l'OHI sont convenus ensemble de modèles de données communs pour la géométrie informatique sous le parapluie de TC211.

La participation canadienne à ces comités internationaux est prioritaire. L'activité se déroule sous les auspices de l'Office des normes générales du Canada – Comité sur la géomatique (ONGC-COG) par le biais du Conseil des normes du Canada (CNC). La structure de l'ONGC-COG reflète celle d'ISO TC211.

Le Comité technique COG – D. V. Hume, ministère des Affaires indiennes et du Nord canadien;
WG1 – Cadre de travail et modèle de référence – D^r K. Fadaie, Centre canadien de télédétection;
WG2 – Modèles géospatiaux et exploitation – (vacant);
WG3 – Administration des données géospatiales – P. Charlesworth, Ressources naturelles Canada;
WG4 – Service géospatial – P. Morin, ministère de la Défense nationale;
WG5 – Normes fonctionnelles – C. D. O'Brien, IDON Technologies Inc.;
CAC – Comité consultatif canadien – D. McKellar, ministère de la Défense nationale.

Il est prévu par le Comité interministériel pour la télédétection du gouvernement fédéral que les composants et les profils de produits de données enregistrés de TC211 constitueront les normes de l'Infrastructure canadienne de données géospatiales (ICDG).

VII. Environnement (B. Malone)

A. Bases de données pour l'analyse environnementale (# au cat. 16-506XCB)

Ce guide de référence unique sur CD-ROM aux données environnementales disponibles partout au Canada constitue un inventaire de quelque 1 200 bases de données conservées par 127 ministères des gouvernements fédéral, provinciaux et territoriaux. Recensant des sujets comme les cartes numériques en passant par les pluies acides jusqu'à la biodiversité, il s'agit d'un outil indispensable pour les chercheurs qui s'intéressent à la relation entre les activités humaines et les conditions et les tendances environnementales. Une vingtaine de variables décrivent chaque base de données dont une description sommaire, l'organisme responsable, les contacts, la liste des paramètres, le recensement géographique et la séquence de mise à jour.

(trephele@statcan.ca)

B. Registre national automatisé des résidus de produits chimiques toxiques

Les banques de spécimens, une collection de tissus biologiques archivés ou de matériaux environnementaux et de données connexes, représentent un genre particulier de base de données. Depuis le début des années soixante la banque de spécimens du Registre national automatisé des résidus de produits chimiques toxiques est passée d'une collection spécialisée conservée dans un congélateur bahut à l'une des plus grandes et plus importantes banques de spécimens biologiques au monde. Elle conserve actuellement 55 000 échantillons et plus de 400 000 sous-échantillons dans quatre congélateurs-chambres (-40 °), trois congélateurs bahuts à température ultra basse (-80 °) et deux congélateurs à azote liquide (-150 °). (La banque de spécimens du Service canadien de la faune, du Centre national de la recherche faunique, Service canadien de la faune, 100, rue Gamelin, Hull (Qc) J8Y 1V9).

La surveillance de l'environnement demeure l'un des outils fondamentaux pour la gestion efficace des stress anthropogéniques sur la biosphère au niveau local, national et international. Afin d'être efficace, la surveillance doit être suffisamment souple pour s'adapter aux changements dans la nature des problèmes qui se posent, aux technologies utilisées et à l'état de la connaissance. Il faut être en mesure de décrire l'état de l'environnement, d'évaluer les menaces sur l'environnement, d'analyser les effets anthropogéniques sur l'environnement et de fournir la base pour établir et surveiller nos réponses à ces stress. La banque de spécimens est l'une des façons les plus efficaces en termes de coût de réaliser ces buts.

Les banques de spécimens collectionnent et conservent le matériel de façon à permettre :
?? la surveillance en temps réel des paramètres environnementaux d'intérêt courant,

- ?? la surveillance en rétrospective pour étudier des paramètres d'intérêt nouvellement découverts, en vue de valider les méthodes analytiques nouvelles ou modifiées et de vérifier les résultats antérieurs,
- ?? la recherche écotoxicologique.

VIII. Données sur les propriétés de matériaux (G. Wood)

A. Vieillissement des structures de béton en environnement nucléaire

Avant de passer à la mise sur pied d'une Base de données internationale sur le vieillissement du béton, la question a été soumise par l'AIEA en 1995 pour une évaluation à l'OCDE/AEN à Paris tel que décrit dans le *Rapport sur les activités relatives aux données* au Canada en 1996.

De nombreuses réunions ont été convoquées par l'AEN et la base de données a reçu la priorité. À la réunion de l'AEN de mai 1997 le Canada (C. Seni) a présenté un rapport sur la façon de mettre en oeuvre la base de données. Cependant, les États membres des pays ont fait valoir leurs préoccupations du fait que la collaboration requise des installations nucléaires dans divers pays pourrait ne pas être obtenue et que l'AIEA serait mieux placée pour obtenir cette collaboration. Pendant cette période et avec la participation de C. Seni, l'AEN a aussi organisé deux ateliers internationaux portant sur le vieillissement du béton, qui deviendra plus tard une source d'enrichissement de la base de données sur le vieillissement du béton.

Au cours de la même période et à l'instar des activités de l'AIEA sur les questions de vieillissement du béton, C. Seni et D. Naus (É.-U.) ont mis sur pied, dans le cadre de la RILEM (International Reunion of Laboratories for Testing of Structures and Materials), un comité technique qui doit s'occuper des questions du vieillissement du béton en environnement nucléaire y compris de la base de données sur le vieillissement du béton. Les activités ont été coordonnées avec celles de l'AEN, par C. Seni. Le Comité technique de la RILEM nouvellement créé TC-160 MLN (Methodology for Life prediction of concrete structures in Nuclear power plants) a défini une structure pour la Base de données sur le vieillissement du béton qui sera adoptée par l'AIEA lorsqu'elle lancera le programme. C. Seni, B. Oland (É.-U.) et M. Johnston (R.-U.) ont pris une part active aux travaux.

Comme il est prévu par l'AIEA, le lancement de la Concrete Containment Database se fera en l'an 2000 lors d'une réunion d'experts qui sera convoquée pour décider de la façon de procéder.
(czeni@globalserve.net)

B. Computer Integrated Material Database (CIMDATA)

L'Institut des matériaux industriels du Conseil national de recherches Canada a produit un système intégré de bases de données de connaissances et d'information précises portant sur les plastiques, les alliages métalliques non ferreux, les alliages de fonte et les matériaux d'outillage.

Plus de 100 propriétés physiques et techniques et plus de 600 matériaux de moulage y sont recensés.

La composante «connaissances» du système portant sur le contrôle des procédés, par exemple, offre des réponses ou des suggestions aux problèmes pouvant surgir lors de la production avec certains procédés. Ainsi, un ingénieur qui se retrouve avec un moule à injection qui ne fonctionne pas pourrait interroger le système et être orienté vers une solution grâce aux compétences recensées dans la base.

Cette composante est étoffée par les bases de données d'information précises qui aident l'utilisateur en matière de sélection et de comparaison des propriétés des matériaux. Dans cette composante, l'utilisateur peut choisir parmi de nombreuses propriétés et de nombreux paramètres pour déterminer le matériau qui se prêtera le mieux à son application. (georges.salloum@nrc.ca)

IX. Physique - astrophysique (D. Durand)

Le Centre canadien de données astronomiques (CCDA) a continué à être un centre dynamique d'activités en matière de données astronomiques. En effet, le CCDA a élaboré une approche novatrice de la diffusion des données d'archive du télescope spatial Hubble, approche qui prévoit le stockage de données brutes sur CD-ROM dans une bibliothèque de disques optiques à 500 logements ainsi que l'étalonnage automatique des données lorsque celles-ci sont sollicitées des archives. Le système mise sur les plus récents logiciels d'étalonnage et les plus récents fichiers de référence. Les archives peuvent être consultés par l'interface World Wide Web, qui permet l'accès par toutes les plates-formes informatiques.

Le CCDA réalisera le transfert des archives du télescope Canada-France-Hawaii sur CD-ROM et sera aussi responsable d'offrir l'accès en direct aux données. Cette année, le CCDA commencera aussi à archiver les données de deux importantes installations astronomiques canadiennes, nommément le télescope James Clerk Maxwell à Hawaii et l'Observatoire fédéral de radio astrophysique à Penticton, en C.-B.

Les astronomes ont aussi obtenu l'accès à deux versions du Digital Sky Survey par le biais du CCDA. À partir des plus de 300 CD disponibles actuellement, les usagers du Web peuvent repérer n'importe quelle partie du ciel sur leur propre ordinateur par l'Internet.

Le CCDA veille actuellement à la conception et à la mise en oeuvre de TERAPIX, une pipeline de traitement des données consacrée à la nouvelle caméra qui sera attachée au CFHT en 2001. Cette caméra, capable de présenter un angle de vision d'un degré, produira des images d'environ 8 Mo chacune. En raison du grand volume de données, 100 images par nuit pour environ 100 nuits par année, TERAPIX doit être en mesure de traiter ces images automatiquement.

(<http://cadcwww.hia.nrc.ca/>)

X Thermodynamique (J. Sangster)

A. Formulation analytique interactive en thermodynamique (F*A*I*T)

F*A*I*T est un système intégré canadien comprenant une base de données thermochimiques, un groupe de programmes ainsi que des données thermodynamiques cohérentes évaluées par les chercheurs. Le système contient actuellement des données sur plus de 5 000 composés chimiques ainsi que des bases de données de solutions représentant plus de 70 solutions non idéales (alliages métalliques, scories, mattes et sels fondus). F*A*I*T est accessible en direct auprès de l'Université McGill et est aussi offert en version PC.

(<http://www.crct.polymtl.ca>)

B. Programmes de recherche universitaire

Les professeurs C. B. Alcock et V. Itkin (University of Toronto) évaluent les données thermodynamiques des éléments (température de Debye, $C_p(T)$, enthalpie, entropie de troisième loi et propriétés de fusion). Les données de C_p sont décrites par différentes équations et des données sont recommandées.

Le professeur A. E. Mather (University of Alberta) mesure les liquides vapeur-liquide et les enthalpies pour la réaction et la solution des gaz acides en solution aqueuse de solvants organiques polaires (application en purification gazeuse). Il a contribué aux séries de données sur la solubilité de l'IUPAC en compilation et en évaluation de données sur le CO_2 pour l'eau et pour les systèmes non aqueux, ainsi que pour des solides et des liquides dans du CO_2 surcritique.

Le professeur J. Liemezs (University of British Columbia) compile et évalue les données pour l'élaboration d'équations d'état. Ces données servent aussi aux méthodes prédictives corrélatives pour les chaleurs de vaporisation, la tension superficielle et les propriétés de transport ainsi que pour les propriétés de thermodynamique des gaz parfaits.

Le professeur P. Englezos (University of British Columbia) mesure les équilibres de phase hydrate de gaz auxquels participent le méthane, le CO_2 , les hydrocarbures et l'azote. Les mesures couvrent aussi la solubilité du carbonate de calcium en présence de substances adsorbées.

XI. Comité national canadien de CODATA

Le Comité a continué de se réunir annuellement pendant le mandat biennal de coordination de l'Institut canadien de l'information scientifique et technique (ICIST). Monsieur Daniel Durand, remplaçant M. Denis Crabtree, et M. George Needler sont venus siéger au comité à titre de nouveaux membres; le mandat de M. Sangster a été renouvelé pour trois ans. La composition actuelle du comité, ainsi que les responsabilités de rapporteur pour le présent rapport, sont indiquées dans le tableau ci-dessous.

À sa réunion de mai 1997, le Comité a lancé un projet pilote pour promouvoir la sensibilité au Canada du besoin pour la qualité et l'uniformité des données. L'objectif du projet pilote était

d'assurer par entrevue l'assurance de la qualité de la méthodologie utilisée par les producteurs d'un certain nombre limité d'ensembles de données canadiens. Il est prévu qu'il faudra encore au moins six mois à un an avant de terminer l'étude.

Président	Rapporteur - section	Courrier électronique
Prof. Paul Mezey	Chimie	mezey@sask.usask.ca
Membres		
M. Daniel Durand	Physique-astrophysique	daniel.durand@hia.nrc.ca
M. Maria Korab-Laskowska	Biologie	korablam@bch.umontreal.ca
M. George Needler	Océanographie	G_Needler@bionet.bio.dfo.ca
M. James Sangster	Thermodynamique	jsangster@mail.polymtl.ca
M. Roger F. Tomlinson	Géospatiale	talgeo@magma.ca
Observateurs		
M. Robert Berman	Géoscience	berman@gsc.emr.ca
M. Andrée Bichon		andree.bichon@nrc.ca
M. Richard L. Coles	Géophysique	coles@geolab.emr.ca
M. Brian Malone	Environnement	bmalone@synapse.net
M. John R. Rodgers	Cristallographie	rodgers@snd.cisti.nrc.ca
Prof. Michel Sabourin		sabourm@ere.umontreal.ca
M. Larry Speers		speersli@em.agr.ca
M. Barry M. Wood		barry.wood@nrc.ca
Secrétariat		
M ^{me} Roxanne Rochon (secrétaire)		roxanne.rochon@nrc.ca
M. Gordon H. Wood (secrétaire exécutif)	Matériaux	gordon.wood@nrc.ca

Le Comité s'est acquitté de sa responsabilité de distribuer le bulletin CODATA à plus de 400 adresses au Canada. L'ICIST, à titre de secrétariat du Comité, a la distinction d'être l'hôte du site Web principal pour CODATA qui fait le lien avec toutes les autres activités CODATA dans le monde entier et offre les versions électroniques du bulletin, du manuel, de divers rapports, etc. (<http://www.cisti.nrc.ca/programs/codata/welcome.html>)

En outre, l'Institut Herzberg d'astrophysique du CNRC à Victoria a continué d'être l'hôte du site Web pour le CNC/CODATA. M. Durand en est le webmestre. (<http://cadwww.dao.nrc.ca/cnc-codata/>)