

STI

REVUE
N° 24

SCIENCE TECHNOLOGIE INDUSTRIE

Numéro spécial : «Le village mondial de la recherche »

Les technologies de l'information et l'activité scientifique :
tour d'horizon et principaux problèmes

Infrastructure de l'information pour le village mondial
de la recherche : la perspective européenne

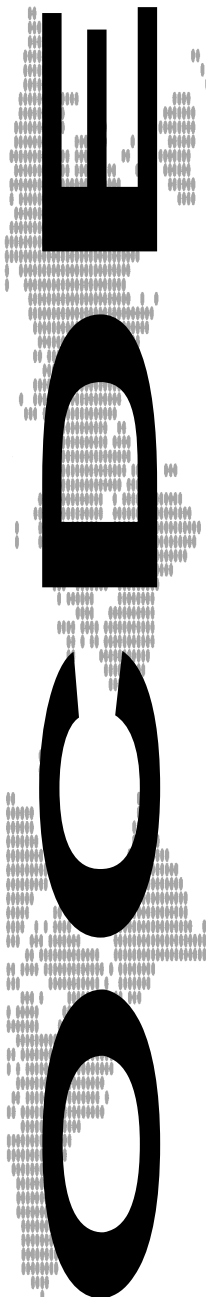
Les réseaux informatiques et l'université virtuelle

La publication électronique

La bibliothèque de recherche électronique du Danemark :
un outil du changement institutionnel

L'avenir des bases de données mathématiques

L'infrastructure de l'information en sciences sociales



© OCDE, 2000

© Logiciel, 1987-1996, Acrobat, marque déposée d'ADOBE.

Tous droits du producteur et du propriétaire de ce produit sont réservés. L'OCDE autorise la reproduction d'un seul exemplaire de ce programme pour usage personnel et non commercial uniquement. Sauf autorisation, la duplication, la location, le prêt, l'utilisation de ce produit pour exécution publique sont interdits. Ce programme, les données y afférentes et d'autres éléments doivent donc être traités comme toute autre documentation sur laquelle s'exerce la protection par le droit d'auteur.

Les demandes sont à adresser au :

Chef du Service des Publications,
Service des Publications de l'OCDE,
2, rue André-Pascal,
75775 Paris Cedex 16, France.

N° 24

STI
REVUE

Numéro spécial :
«Le village mondial
de la recherche»

ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES

En vertu de l'article 1^{er} de la Convention signée le 14 décembre 1960, à Paris, et entrée en vigueur le 30 septembre 1961, l'Organisation de Coopération et de Développement Économiques (OCDE) a pour objectif de promouvoir des politiques visant :

- à réaliser la plus forte expansion de l'économie et de l'emploi et une progression du niveau de vie dans les pays Membres, tout en maintenant la stabilité financière, et à contribuer ainsi au développement de l'économie mondiale ;
- à contribuer à une saine expansion économique dans les pays Membres, ainsi que les pays non membres, en voie de développement économique ;
- à contribuer à l'expansion du commerce mondial sur une base multilatérale et non discriminatoire conformément aux obligations internationales.

Les pays Membres originaires de l'OCDE sont : l'Allemagne, l'Autriche, la Belgique, le Canada, le Danemark, l'Espagne, les États-Unis, la France, la Grèce, l'Irlande, l'Islande, l'Italie, le Luxembourg, la Norvège, les Pays-Bas, le Portugal, le Royaume-Uni, la Suède, la Suisse et la Turquie. Les pays suivants sont ultérieurement devenus Membres par adhésion aux dates indiquées ci-après : le Japon (28 avril 1964), la Finlande (28 janvier 1969), l'Australie (7 juin 1971), la Nouvelle-Zélande (29 mai 1973), le Mexique (18 mai 1994), la République tchèque (21 décembre 1995), la Hongrie (7 mai 1996), la Pologne (22 novembre 1996) et la Corée (12 décembre 1996). La Commission des Communautés européennes participe aux travaux de l'OCDE (article 13 de la Convention de l'OCDE).

Également disponible sous le titre :

STI REVIEW
Special Issue on
"The Global Research Village"
No. 24

© OCDE 1999

Les permissions de reproduction partielle à usage non commercial ou destinée à une formation doivent être adressées au Centre français d'exploitation du droit de copie (CFC), 20, rue des Grands-Augustins, 75006 Paris, France, Tél. (33-1) 44 07 47 70, Fax (33-1) 46 34 67 19, pour tous les pays à l'exception des États-Unis. Aux États-Unis, l'autorisation doit être obtenue du Copyright Clearance Center, Service Client, (508)750-8400, 222 Rosewood Drive, Danvers, MA 01923 USA, or CCC Online: <http://www.copyright.com/>. Toute autre demande d'autorisation de reproduction ou de traduction totale ou partielle de cette publication doit être adressée aux Éditions de l'OCDE, 2, rue André-Pascal, 75775 Paris Cedex 16, France.

AVANT-PROPOS

Préparée par la Direction de la science, de la technologie et de l'industrie de l'OCDE, la *STI Revue* est publiée deux fois par an. Elle présente des études intéressantes pour les responsables politiques et les analystes concernés par les développements scientifiques, technologiques et industriels, en mettant l'accent sur les comparaisons entre pays, les analyses quantitatives des tendances nouvelles et les questions de politique récentes ou à venir. Suivant la nature des travaux de l'OCDE, la *STI Revue* examine les changements structurels et institutionnels qui se produisent tant au niveau global que régional, national et local. Certains numéros portent sur des thèmes spécifiques, comme les enquêtes sur le «comportement innovateur» des entreprises ou les problèmes d'emploi liés à la technologie.

Ce numéro de la *STI Revue* est plus particulièrement consacré aux incidences des technologies de l'information et des communications sur la recherche scientifique. Le Comité de la politique scientifique et technologique de l'OCDE examine l'évolution du rôle des pouvoirs publics dans le soutien du système scientifique, évolution déclenchée dans une large mesure par les vastes possibilités qu'offrent ces technologies. Les articles présentés dans ce numéro s'inspirent principalement des travaux de la 2^e Conférence sur le village mondial de la recherche intitulée «Maximiser les avantages de la technologie de l'information pour la science» ; celle-ci s'est tenue à Sintra (Portugal) et était organisée conjointement par le ministère portugais de la Science et de la Technologie et l'OCDE. Les principaux thèmes de ces articles sont le développement de l'infrastructure basée sur les TIC au service de la science dans les différents pays de l'OCDE et l'utilisation de cette infrastructure par les scientifiques pour communiquer, collaborer, accéder à l'information et publier les résultats de la recherche.

Les opinions exprimées dans cette publication ne reflètent pas nécessairement celles de l'Organisation ou de ses pays Membres. La *STI Revue* est publiée sous la responsabilité du Secrétaire général de l'OCDE.

TABLE DES MATIÈRES

Maximiser les technologies de l'information et l'activité scientifique : tour d'horizon et principaux problèmes <i>Jean-Éric Aubert et Viviane Bayar</i>	7
Infrastructure de l'information pour le village mondial de la recherche : la perspective européenne <i>Hans-Peter Axmann et Sabine Payr</i>	33
Les réseaux informatiques et l'université virtuelle <i>John P. Walsh et Ann Roselle</i>	55
La publication électronique <i>Karel Leeftang et Jim Gilmore</i>	91
La bibliothèque de recherche électronique du Danemark : un outil du changement institutionnel <i>Jens Thorhauge</i>	111
L'avenir des bases de données mathématiques <i>Jean-Pierre Bourguignon et la Société européenne des mathématiques</i>	123
L'infrastructure de l'information en sciences sociales <i>William Sims Bainbridge</i>	141

MAXIMISER LES TECHNOLOGIES DE L'INFORMATION ET L'ACTIVITÉ SCIENTIFIQUE : TOUR D'HORIZON ET PRINCIPAUX PROBLÈMES

Introduction

Les technologies de l'information et de la communication (TIC) sont devenues des outils essentiels de soutien des activités de recherche. Les progrès extraordinaires réalisés par ces technologies ont influencé pratiquement tous les aspects de l'activité scientifique, de la collecte et de l'analyse des données aux modes de diffusion des résultats de la recherche en passant par les modes de mise en relation des scientifiques de par le monde avec les nouvelles structures de collaboration. Les changements induits par l'utilisation des TIC se sont produits à une époque où les systèmes scientifiques étaient déjà confrontés à la nécessité de s'adapter à d'autres changements tels que la baisse du financement de la R-D par les pouvoirs publics, l'accent mis de plus en plus sur la pertinence économique et la transparence, les liens plus étroits avec la collectivité et le secteur des entreprises, l'attention accrue portée à la diffusion et à la coopération internationale en matière de recherche – et où simultanément ils prenaient une importance accrue dans des économies de plus en plus fondées sur le savoir.

En conséquence, de nombreux gouvernements s'intéressent de plus en plus au rôle des TIC dans la redéfinition des responsabilités des pouvoirs publics à l'égard du système scientifique. Dans le cadre de l'OCDE, une conférence sur le village mondial de la recherche, organisée par le Comité de la politique scientifique et technologique au Danemark en juin 1996, a engagé un dialogue entre responsables de l'action publique et la communauté scientifique afin de déterminer le rôle des gouvernements dans la mise à disposition et l'utilisation des ressources des TIC sur la scène scientifique mondiale.

Un rapport préparé par le Secrétariat de l'OCDE dans le courant de l'année 1997 a examiné ensuite comment l'utilisation des TIC a influencé la recherche dans les cinq domaines suivants : la communication entre scientifiques ; l'accès à l'information scientifique ; les instruments scientifiques ; l'édition électronique dans le domaine scientifique et l'enseignement et la formation scientifiques, en considérant à la fois les avantages et les problèmes de mise en œuvre. Sur la base

de ce rapport, une deuxième conférence sur le village mondial de la recherche s'est tenue à Sintra au Portugal en septembre 1998. Plus de 60 participants parmi lesquels figuraient des hauts responsables gouvernementaux des Ministères des sciences, de la technologie, de l'industrie et de l'éducation et des organisations intergouvernementales, ainsi que des représentants des instituts de recherche, du milieu universitaire et de l'industrie, ont discuté du rôle de l'État dans l'exploitation du potentiel que représentent pour la science les TIC.

Ce chapitre introductif souligne en premier lieu les différences de développement de l'infrastructure des TIC dans les différents pays de la zone OCDE. Il résume ensuite l'impact des TIC sur certains aspects du système scientifique : la communication entre scientifiques, les publications, les bibliothèques de recherche et les bases de données. Enfin, il présente les recommandations formulées lors de la Conférence de Sintra.

L'infrastructure des TIC

Toute une série de développements des TIC couvrant le matériel, les logiciels et les technologies de la gestion de réseau sont à la base des changements continus que connaît le système scientifique. L'infrastructure physique des TIC constitue le fondement des applications basées sur les TIC qui non seulement aident les scientifiques dans leurs tâches quotidiennes mais leur permettent d'entreprendre des projets de recherche qui auraient été impensables dans le passé. Il est nécessaire de disposer d'une infrastructure physique adéquate en matière de TIC pour appuyer des activités telles que l'acquisition, la transmission et le partage de données, la recherche informatique, l'accès aux instruments scientifiques, les communications informelles entre scientifiques et la diffusion formelle des résultats de la recherche. Le développement du réseau Internet et des applications interactives nouvelles utilisant bien souvent le multimédia, qui exigent la transmission de volumes considérables de données, a conduit à un accroissement rapide de la demande de large bande et de hautes performances.

Les États-Unis

L'infrastructure liée aux TIC est particulièrement développée aux États-Unis. Conçu initialement pour le secteur de la défense, le réseau Internet est passé par différentes phases de financement et de gestion jusqu'en 1995, date à laquelle les investissements sont passés aux mains d'intérêts commerciaux de plus en plus actifs. Le gouvernement américain continue néanmoins d'encourager très activement une recherche de pointe en vue de pousser la capacité des superordinateurs, des réseaux, des unités de stockage et des logiciels au-delà des limites actuelles, et ce dans l'intérêt des utilisateurs scientifiques (*Bainbridge*).

Dans le domaine de la gestion de réseau et des applications, la NSF a développé le service vBNS (*High Speed Backbone Network Service*) afin de porter à 622 mégabits/seconde (Mbps) dans un premier temps, puis à 2.4 gigabits/seconde (Gbps), la vitesse des réseaux de base existants d'Internet, qui sont les pivots de la transmission de données. Un projet « Internet2 » de coopération entre universités vise à étendre la capacité des réseaux de base perfectionnés aux campus des universités concernées (vitesse de transmission initiale de 2.4 Gbps entre les principaux nœuds d'ici janvier 1999 et 9.6 Gbps par la suite) et à élaborer un logiciel pour le nouveau réseau. Les deux projets (vBNS et Internet2) impliquent un partenariat avec l'industrie. En outre, le gouvernement soutient ces efforts en se concentrant sur les applications avancées avec l'initiative NGI d'Internet de nouvelle génération (à ce jour, les fonds ont servi à relier les universités au réseau vBNS et à améliorer les services sur les réseaux d'ordinateurs à grande vitesse).

Comme le gouvernement américain continue d'être le principal client pour le type de capacité de calcul qu'offrent les super ordinateurs (pour la modélisation et la simulation des phénomènes naturels complexes), l'investissement public demeure essentiel pour les avancées dans ce domaine en dépit du développement très rapide des TIC par l'industrie. Le gouvernement continue en outre d'investir des sommes considérables dans la recherche fondamentale de logiciels, alors que l'industrie a beaucoup moins investi dans ce domaine que dans celui du matériel informatique. Dans le même temps, les exigences liées à la collecte, au stockage, à l'analyse et au partage de données ont pris de l'importance et suscitent une plus grande attention.

L'Union européenne

La situation est bien différente en Europe où une infrastructure étendue de réseaux à hautes performances fait toujours défaut (*Axmann et Payr*). Chacun des États membres de l'Union européenne a financé le développement de réseaux nationaux de recherche et d'éducation (RNEN) qui fonctionnent à des vitesses de transmission variables et sont en général largement ouverts aux chercheurs, enseignants et étudiants dans un pays donné. Néanmoins, la construction de réseaux transeuropéens s'est avérée très difficile car jusqu'à ces dernières années l'organisation et le financement d'une infrastructure de base européenne ont été entravés par l'existence de monopoles nationaux dans les télécommunications. En 1998, par exemple, la largeur de bande entre les communautés universitaires française et belge n'était que de 6 Mbps.

Une réforme de la réglementation dans le secteur des télécommunications où des niveaux de coûts et de prix élevés ont freiné l'utilisation des TIC a été engagée en Europe et dans la plupart des pays la déréglementation est totale depuis janvier 1998. Il reste toutefois beaucoup à faire, en particulier dans certains pays où

l'introduction de nouvelles structures de marché s'opère lentement en dépit des efforts de la Commission européenne pour activer la déréglementation et la libéralisation du secteur. L'industrie est en cours de restructuration et ne s'est guère montrée active dans le fonctionnement en réseau de la recherche.

Jusqu'ici, les programmes de R-D de la Communauté européenne étaient la seule possibilité pour encourager la mise en réseau de la recherche au niveau européen. Le quatrième Programme-cadre, qui n'a pu que soutenir la recherche sur le fonctionnement en réseau et non pas l'établissement des réseaux de recherche, a permis néanmoins la création en 1997 de TEN-34, réseau dont la vitesse d'exploitation sera portée de 34 Mbps à 155 Mbps, ce qui permettra le raccordement des RNEN à des vitesses différentes. Ce n'est qu'au cours du cinquième Programme-cadre (1998-2002) que le financement direct des réseaux pour la recherche a pu être réalisé, grâce à un financement conjoint de la CE (40 %) et des États membres (60 %) pour un montant total de 172 millions de dollars.

Région Asie-Pacifique¹

Les efforts entrepris pour atteindre les établissements de recherche et d'enseignement de cette région sont relativement récents mais ils progressent à un rythme régulier. Le réseau APAN, qui relie l'Australie, la Chine, Hong-Kong (Chine), l'Indonésie, le Japon, la Corée, la Malaisie, les Philippines, Singapour, la Thaïlande et les États-Unis, a été lancé en 1997 pour établir des liaisons à grande vitesse (entre 45 et 155 Mbps) dans le domaine de la recherche et de l'enseignement entre les pays participants. Ce réseau est financé par la *National Science Foundation* (NSF) américaine et par l'Agence japonaise pour la science et la technologie. Son objectif est d'atténuer les disparités régionales de l'accès à l'Internet. Il est prévu d'étendre ce réseau à l'Europe et à l'Amérique latine en 1999.

Néanmoins, l'accès à l'Internet demeure insuffisant pour les scientifiques de certains pays de la région, tels que le Cambodge, le Laos et le Viêt-nam pour n'en citer que quelques-uns. Le *Asia-Pacific Networking Group* a été constitué pour coordonner les efforts ainsi que l'enseignement et de la formation des ingénieurs travaillant sur Internet. La langue constitue également un obstacle dans les pays non-anglophones de la région Asie-Pacifique, tant du point de vue de la communication entre les scientifiques que du temps que pourra prendre l'adaptation de nouveaux outils logiciels – deux ans dans le cas des outils Adobe PDF. Par conséquent, les outils de communication graphique se sont davantage développés dans ces régions que dans d'autres car les progrès de la technologie facilitent la transmission d'images fixes ou mobiles.

En août 1995, les agences et les ministères japonais concernés par l'infrastructure de l'information et de la recherche ont élaboré des directives. Les efforts entrepris visent à renforcer les infrastructures en reliant les organismes de recher-

¹⁰

che à un réseau qui transcende les frontières entre les instituts, les agences et ministères gouvernementaux ainsi qu'entre les pays. L'Agence japonaise pour la science et la technologie a un projet d'un coût de 400 millions de dollars sur cinq ans pour l'extension du réseau interministériel Imnet, réseau exploité par la Japan Science and Technology Corporation depuis 1997 et qui fait partie de l'APAN. L'objectif du projet est d'élaborer un réseau national ayant une capacité de transmission de 1 Gbps. Les efforts entrepris visent également une amélioration du contenu avec le développement et la maintenance d'une base de données contenant des informations sur les projets de recherche et de bases de données factuelles.

Le ministère de l'Éducation, des Sciences, des Sports et de la Culture a créé et perfectionné les systèmes de diffusion de l'information scientifique, les réseaux et bases de données. Le réseau japonais SINET qu'exploite le Centre national pour les systèmes d'informations scientifiques (NACSIS) du ministère de l'Éducation relie les universités et autres établissements au niveau national. Le ministère accélère le développement du réseau, qui opère actuellement à des vitesses comprises entre 6 et 150 Mbps et étend ses liaisons internationales qui atteignent actuellement la Thaïlande, les États-Unis et le Royaume-Uni. La liaison avec les États-Unis, qui connaît un trafic particulièrement important, a été accélérée en octobre 1997 ; la vitesse y a été portée à 45 Mbps. En mars 1997, le réseau reliait 613 établissements. Le NACSIS gère également des bases de données contenant les catalogues de toutes les universités japonaises, des listes de chercheurs japonais et des résumés de nombreux rapports scientifiques. Des versions en japonais et en anglais des catalogues sont disponibles sur le Web.

Stades de développement différents et problèmes connexes

Étant donné que l'infrastructure de recherche en est à des stades de développement très différents dans les différents pays de la zone OCDE et qu'actuellement elle est loin d'être mondiale, les problèmes qui se posent varient d'un pays à l'autre et d'une région à l'autre. Aux États-Unis, où le réseau pour la recherche fondamentale a été d'ores et déjà fortement développé, le gouvernement s'attache désormais à encourager le raccordement des universités au réseau de base et à stimuler la recherche avancée les infrastructures qui permettront aux scientifiques de repousser les frontières de la science. Des investissements en matériel, modules de logiciels, bases de données et établissements pour la formation des chercheurs devraient permettre de maximiser la rentabilité de l'infrastructure physique.

Au sein de l'Union européenne, l'existence de réseaux nationaux distincts pour la recherche reliés par des services d'interconnexions de réseaux assurés par les opérateurs publics nationaux de télécommunications, a conduit à une infrastructure inadéquate ayant des difficultés d'organisation et de financement. L'accent est mis aujourd'hui sur la libéralisation du secteur des télécommunications et sur

la création d'une infrastructure de base transeuropéenne. Le nouveau modèle de financement de l'infrastructure des TIC, adopté dans le cinquième Programme-cadre devrait être plus efficace que le précédent, bien qu'une incertitude budgétaire subsiste. En outre, certains appellent à une structure d'organisation différente, avec notamment un opérateur indépendant pour le réseau européen de recherche. Bien que les États-Unis soient fréquemment présentés en Europe comme le modèle à suivre en matière de connexions, la philosophie européenne a mis davantage l'accent sur l'élargissement de l'accès de tous les établissements de recherche et d'enseignement à un réseau de connaissances plutôt que sur une technologie de pointe et des applications avancées.

Au niveau mondial, de nouveaux efforts seront nécessaires si l'on a véritablement l'intention d'intégrer les communautés scientifiques des différents pays. Les pays participant au réseau APAN sont bien conscients que le fossé se creuse entre les scientifiques de la région Asie-Pacifique ayant accès à l'Internet et ceux ne l'ayant pas, et ils cherchent des sources de financement publiques et privées pour atténuer ces disparités. En Europe, la plupart des efforts se concentrent aujourd'hui sur les liaisons entre États membres de l'Union européenne et les connexions avec les autres pays ou entre ces pays sont négligées.

Par le passé, la demande de connexions avec les États-Unis a été particulièrement forte dans tous les pays, ce qui a eu pour conséquence que ceux-ci ont dû supporter l'essentiel des investissements nécessaires à l'établissement de liaisons avec les institutions et les réseaux américains. A présent que la demande et que les flux du trafic sont plus uniformément répartis à l'intérieur des États-Unis et en dehors, les pressions en vue d'un partage équitable des coûts et des tarifs entre les États-Unis et les autres pays se font plus grandes. Il faut désormais prendre de nouvelles dispositions concernant le financement des réseaux transcontinentaux, le partage des coûts et la fixation des prix.

Le soutien des pouvoirs publics continue d'être nécessaire pour encourager le développement de l'infrastructure technologique qui sous-tend le travail des scientifiques, et faire en sorte que ceux-ci soient formés à utiliser efficacement l'infrastructure des TIC. Les technologies de pointe qui sortent du champ de l'intérêt commercial, c'est-à-dire en gros les technologies qui permettent aux chercheurs de travailler dans des environnements de coopération répartie et en particulier les technologies de pilotage des instruments à distance, ne peuvent être développées sans l'aide des pouvoirs publics. La participation de l'État continuera d'être nécessaire même si des intérêts commerciaux venaient à effectuer les investissements requis dans la mesure où le financement de l'utilisation des réseaux par les scientifiques deviendrait alors un problème. Lorsqu'ils financent la science, les pouvoirs publics doivent prendre en compte de manière spécifique les besoins des TIC. Plus généralement, si l'on veut que les réseaux scientifiques se développent comme l'ont fait les réseaux téléphoniques – avec l'aide du marché – il faut éviter

d'isoler l'establishment scientifique et ses réseaux et encourager autant que faire se peut une expansion bénéficiant à un segment plus large de la société.

Impact des TIC sur les pratiques de la recherche et sur la productivité

L'une des applications les plus matures des TIC implique l'utilisation du réseau Internet pour les communications informelles entre scientifiques. Outre le fait qu'elle accroît considérablement la quantité, la qualité et la vitesse de communication entre les scientifiques, l'utilisation des TIC a eu également des effets divers sur l'organisation du travail dans les sciences (Walsh). Toutefois, les effets globaux de nouvelles pratiques de recherche sur les résultats scientifiques ne sont pas encore très clairs.

Évolution des pratiques de la recherche

Certains des effets des communications utilisant les TIC sont bien visibles. Une amélioration de la communication a contribué à une augmentation de la taille des réseaux professionnels et conduit à un accroissement des collaborations à distance, en particulier au niveau international. Elle a permis aux chercheurs de lever de nombreuses barrières dues à l'éloignement géographique comme le temps, le coût et la langue. Les TIC ont non seulement facilité les communications avec les collègues mais également permis l'accès aux équipements, aux logiciels et aux bases de données.

Le resserrement des liens entre des scientifiques géographiquement dispersés crée une communauté savante internationale plus dense. Pour un sujet de recherche donné, les TIC ont permis la création de groupes de travail plus complexes avec des structures plus fluides dans lesquelles c'est le sujet de la recherche plutôt que la proximité géographique qui détermine les décisions de collaboration.

La fréquence et la quantité des communications pendant un projet de recherche ont augmenté. Cela est particulièrement important dans le cas d'expériences à long terme, de travail en équipe et de zones de temps différentes car ces communications peuvent renforcer le sentiment d'attachement au groupe de recherche et d'engagement, la satisfaction apportée par le travail, et permettre une bonne coordination des différents éléments de projets de recherche s'étendant sur une longue période. Certaines indications montrent également que la communication par courrier électronique échappe dans une large mesure au jeu social qui généralement accompagne les face à face, voire même les conversations téléphoniques. Il peut en résulter une ambiance moins collégiale et un environnement de travail plus aliénant mais également une focalisation accrue sur la communication d'informations professionnelles.

Mais les indications concernant l'impact de certains autres changements induits par les TIC, notamment au niveau du statut et de la hiérarchie sont contradictoires. Les TIC peuvent faciliter l'accès des jeunes scientifiques et des chercheurs des établissements périphériques à des ressources cruciales mais éloignées, par exemple à des collègues et à des informations mais aussi à des installations de calcul, des logiciels ou des bases de données qui ont été de tout temps inégalement répartis. En revanche il n'est pas évident que ces chercheurs puissent prendre une part active aux projets ou apporter des contributions importantes. Une autre incertitude porte sur le point de savoir si les TIC conduisent à une plus grande collaboration interdisciplinaire ou au contraire à une fragmentation de la recherche, les chercheurs utilisant un temps de communication limité pour établir une interaction uniquement avec des personnes travaillant dans leur domaine de spécialité. Cela pourrait être bénéfique et permettre des économies d'échelle dans certains domaines scientifiques mais cela pourrait également réduire les fertilisations croisées d'idées entre disciplines.

L'impact net de ces nouvelles formes d'organisation n'est pas encore très clair. Globalement, l'utilisation des TIC risque de conduire davantage à un élargissement de la base scientifique qu'à une modification de la hiérarchie des établissements scientifiques. Outre les aspects sur lesquels les indications sont contradictoires, l'incertitude résulte de ce que les impacts sont atténués par le fait que l'utilisation mais aussi la possibilité de tirer profit de la technologie varient considérablement selon les domaines. Ceux dans lesquels l'interdépendance est grande et les interactions entre collaborateurs fréquentes, tels que les mathématiques, la physique et l'aérospatiale, dans lesquels les collaborateurs seront dispersés sont vraisemblablement les domaines qui en profiteront le plus. En revanche, dans des disciplines comme l'ornithologie et la botanique, les contraintes techniques liées à la transmission d'informations non textuelles et à un rythme de découverte relativement lent risquent de limiter les avantages.

Productivité

Une coopération et des communications accrues, un accès plus grand et plus rapide aux informations mais aussi une diffusion plus rapide et plus large des résultats de la recherche, des instruments de plus en plus puissants, efficaces et accessibles et des ressources numériques améliorées et plus accessibles pour l'enseignement et la formation sont autant de facteurs qui contribuent à la transformation continue du système scientifique.

Néanmoins, on ne sait pas très bien si ces développements contribuent à abaisser le coût de la recherche ou à améliorer la productivité du système scientifique. Les indications concernant l'utilisation des réseaux et la productivité des scientifiques ont une relation de corrélation qui n'est pas nécessairement causale.

La productivité des scientifiques, lorsqu'elle est mesurée par le nombre d'articles produits, ignore totalement l'aspect qualitatif qui est plus important et elle doit être pondérée par exemple par un examen de l'ampleur des citations. En outre, bien que les TIC semblent améliorer l'efficacité du travail des scientifiques, l'étendue de leur utilisation et leurs avantages potentiels varient selon la discipline. Qui plus est, l'utilisation des TIC peut impliquer des coûts d'apprentissage qui réduisent les gains potentiels de productivité.

Les TIC peuvent également conduire à des économies d'échelle et de champ mais les accords de collaboration sont de plus en plus spécialisés tandis que le nombre des percées scientifiques extra-disciplinaires augmente. La communication électronique est plus utile pour transmettre un savoir codifié ; elle l'est moins pour diffuser le savoir-faire ou les compétences des individus, ce que l'on considère de plus en plus pertinent. Globalement, les avancées scientifiques de pointe restent d'un coût élevé qui, dans certains cas, peut augmenter. En des temps de rigueur budgétaire, les gains de productivité peuvent, le cas échéant, contribuer au maintien du volume et de la qualité de la production scientifique (OCDE, 1998).

Édition électronique

Alors que les moyens informels de communication entre scientifiques sont devenus plus visibles avec l'utilisation des TIC, ces technologies ont également transformé la publication scientifique qui est le mode officiel de diffusion des résultats de la recherche (*Bourguignon ; Leeftang et Gilmore*).

Avant l'avènement de l'édition électronique, le chercheur qui voulait diffuser largement ses travaux à ses confrères devait les publier dans une revue périodique sur papier, tout comme n'importe quel autre auteur. Pour que son article soit publié, l'auteur cédait ses droits de copyright à un éditeur qui le protégeait contre le vol du contenu de son article et qui, grâce à ces droits, récupérait le coût de son investissement substantiel et dégagait un profit. Les utilisateurs potentiels avaient accès aux articles via les bibliothèques de recherche qui s'abonnaient à ces périodiques. Le prix des périodiques augmentant, le nombre des bibliothèques qui pouvaient se permettre de les acheter a diminué, limitant de plus en plus l'accès à ces articles. En conséquence, les travaux n'ont plus pu être diffusés aussi largement que les scientifiques le souhaitaient (*Harnad et Hemus, 1997*).

Impacts potentiels des TIC sur l'édition scientifique²

L'édition scientifique traverse actuellement une révolution qui affecte tous les aspects de la publication. Cette révolution se caractérise par :

- *De nouveaux outils de production des documents* : les auteurs eux-mêmes peuvent produire des manuscrits d'une grande qualité à l'aide de logiciels de traitement de textes sophistiqués tels que TeX.

- *Un accès élargi* : les documents stockés sont accessibles à partir de n'importe quel point du monde, ce qui modifie totalement la notion de distance.
- *Des contraintes économiques différentes* : jusqu'à présent, l'acceptation des manuscrits par les revues à l'issue d'un processus de sélection était directement liée aux coûts de production ; avec l'avènement de l'édition électronique, cette contrainte liée n'existe plus.
- *Stockage* : le stockage n'est plus une question de volume ; il est devenu plus sophistiqué, c'est-à-dire qu'il nécessite par exemple des outils appropriés pour transférer les fichiers d'un système d'exploitation sur un système de nouvelle génération.
- *La nécessité de disposer de nouveaux outils*, par exemple de moteurs de recherche appropriés et, du fait que la littérature primaire est disponible en ligne, d'un logiciel qui permette aux utilisateurs de passer d'un langage à un autre.
- *Types de document* : on pourra sous peu stocker des images animées.

L'impact des nouvelles technologies sur une science donnée dépend de la pondération des facteurs ci-dessus.

Le principal vecteur de l'édition électronique est le réseau Internet qui est fondé sur une philosophie du libre accès. Cela pose le problème du contrôle des documents disponibles sur Internet (copyright, authentification). De plus, les nouvelles technologies sont en train de changer les compétences professionnelles exigées des personnes qui interviennent dans la diffusion de l'information ; ces personnes doivent désormais être expertes en (télé)communications.

On observe dans ce contexte plusieurs tendances :

- Une vitesse accrue de diffusion des prépublications, due en partie au fait que les scientifiques eux-mêmes peuvent désormais élaborer des manuscrits d'un haut niveau de qualité.
- Des tentatives de lancement de formes de publications meilleur marché, essentiellement sous forme électronique.
- De nouvelles possibilités pour réunir rapidement et pratiquement sans effort un grand nombre de références grâce au format électronique des bases de données bibliographiques.
- Une moindre réticence à publier dans des revues spécialisées dans la mesure où les publications en ligne peuvent être plus aisément localisées.

De grandes incertitudes

L'avenir de la publication des revues est extrêmement incertain. L'une de ces incertitudes a trait aux économies de temps et de coûts que l'on peut effectivement réaliser avec l'édition électronique. L'économie de temps que permet l'édition élec-

tronique par rapport à l'impression n'est pas aussi avantageuse qu'on le pense parfois. Si le délai entre la soumission et la publication d'un texte peut être considérablement raccourci avec l'édition électronique, le résultat final dépend du caractère de la publication électronique et de la discipline scientifique (Kling et MacKim, 1997).

Par ailleurs, les estimations des économies de coûts que permet l'édition électronique diffèrent considérablement. Il est généralement admis que le supplément de coût qu'implique l'accès en ligne aux publications papier existantes est de l'ordre de 20 à 30 pour cent. Ce coût inclut le coût de l'accès et de la maintenance d'archives numériques mais ne prend pas en compte les coûts pour l'utilisateur (CIUS, 1998). Néanmoins, certaines études donnent à penser que ces économies dépendent très largement du nombre d'abonnés. Dans le cas d'une publication purement électronique, certaines indications donnent à penser qu'elle peut être produite pour environ le même coût qu'une version imprimée, mais l'expérience dans ce domaine est limitée (CIUS, 1998). Selon Stevan Harnad, rédacteur de longue date de revues papier et électroniques soumises à un comité de lecture et créateur d'archives électroniques sur les sciences cognitives, les économies réalisées par les éditeurs qui ont lancé une publication uniquement électronique peuvent représenter jusqu'à 75 pour cent du coût actuel d'une publication papier (Harnad et Hemus, 1997).

En dépit de ces nombreux avantages potentiels et réels, l'édition électronique de publications primaires et de bases de données secondaires a constitué jusqu'à une date récente un complément actualisé et rapidement accessible à une publication imprimée plutôt qu'un véritable substitut. Les recettes provenant des produits électroniques représentent toujours moins de 15 pour cent des recettes totales des éditeurs commerciaux traditionnels de revues scientifiques, technologiques et médicales (Leefflang et Gilmore).

Des applications ayant réussi dans certaines disciplines ne sont pas nécessairement réalisables ailleurs dans la mesure où les pratiques de collaboration, les utilisations de formats de publications pré-revues et les utilisations des revues de recherche diffèrent d'une discipline à l'autre (Kling et McKim, 1997). La Physics E-Print Archive, créée par Paul Ginsparg, est l'une des applications innovantes des TIC à la recherche fréquemment citée dans le domaine scientifique comme étant un modèle d'archivage électronique mondial de prétirages. Mais la physique des hautes énergies se caractérise par des pratiques particulières en matière de recherche. La recherche évolue autour d'un nombre limité d'instruments coûteux et implique d'importants groupes de recherche en collaboration travaillant ensemble sur de longues périodes. Lorsque les résultats de ces recherches sont soumis à une revue, ils ont en général été amplement révisés, le délai de publication est court et les prétirages peuvent être disponibles simultanément, ce qui diminue le risque de plagiat.

Cela est rarement le cas dans la recherche biologique, qui est une recherche très fragmentée impliquant un grand nombre de petits groupes de recherche et de chercheurs individuels. Les installations de recherche sont communes et relativement bon marché. La recherche biologique est également plus facile à vulgariser et à copier. Les chercheurs qui travaillent dans ce domaine sont donc réticents à partager les résultats de leurs recherches avant publication. Dans les secteurs de la biologie, tels que la recherche sur le cancer ou le sida, qui sont étroitement liés à des applications commerciales, les chercheurs travaillent bien souvent avec le secteur privé. Le fait que ces travaux puissent être lucratifs et qu'ils soient hautement concurrentiels explique que les chercheurs ne soient pas disposés à partager leurs méthodes de recherche, leurs matériaux et leurs résultats. En biologie, les publications sont donc axées sur les revues qui font examiner les articles par des pairs et les prétirages sont très rares.

Une large diffusion et un libre accès aux résultats de la recherche posent des problèmes de contrôle de la qualité (examen par les pairs) et de contrôle des documents diffusés sur le réseau Internet (validation de l'identité de l'auteur et date de publication, copyright). Il est absolument nécessaire tant pour les utilisateurs que pour l'évaluation des chercheurs et du système scientifique d'être bien conscients de l'inégale qualité du nombre grandissant de données et d'informations auxquelles les TIC donnent largement accès. Si éditeurs et universitaires s'accordent à penser qu'il est nécessaire que subsiste un contrôle par les pairs qui constitue un label de qualité pour les résultats de la recherche, l'édition électronique a rendu nécessaire la mise au point de nouvelles méthodes pour effectuer ce contrôle par les pairs qui, traditionnellement faisait partie du processus d'édition dans des revues de référence. Le Conseil international d'unions scientifiques (CIUS) a recommandé aux sociétés scientifiques et aux éditeurs d'établir des directives pour garantir la qualité et l'exactitude du processus d'examen par un comité de lecture dans un environnement numérique (1998a).

Les problèmes de copyright et les droits de propriété intellectuelle des auteurs sont en conflit avec les exigences de certains éditeurs commerciaux. Un fossé important subsiste entre les points de vue des éditeurs et des producteurs, d'une part et ceux des universitaires et des bibliothécaires d'autre part sur ce qui constitue une « utilisation loyale » de documents, d'images et de clips protégés par des droits de copyright sur les réseaux des campus et dans les cours en ligne (*Chronicle of Higher Education*, 1997). Les détenteurs de droits de copyright ont peur de perdre tout contrôle sur leur document une fois celui-ci sur le réseau alors que les universitaires veulent conserver leur droit de l'utiliser librement à des fins scolaires et éducatives. Ces divergences de points de vue ont conduit à suggérer que les chercheurs conservent leurs droits de copyright sur leurs travaux et qu'ils n'accordent aux revues qu'une licence limitée de publication de communications qu'elles ne possèdent pas, ce qui leur permettrait de diffuser sur le Web les résultats de leurs propres travaux (*Chronicle of Higher Education*, 1998).

Les éditeurs veulent protéger leurs investissements tout en jouant leur rôle qui est d'accroître les flux d'informations, et notamment de données, à destination de la communauté scientifique mondiale et de fixer les limites à une utilisation loyale. Généralement assemblées par les scientifiques dans le cadre de leurs recherches, les bases de données factuelles finissent fréquemment dans le domaine public. Celles qui continuent d'exister au-delà de quelques années sont généralement subventionnées par des établissements nationaux ou internationaux. Leur privatisation n'est guère probable étant donnée l'importance des subventions dans ce domaine dans différentes régions du monde et ce qu'il en coûterait pour l'utilisateur final. Les éditeurs appellent à des interactions entre acteurs publics et privés, en particulier en ce qui concerne le financement, afin de relever les nouveaux défis posés par les TIC, défis qui ont trait notamment aux conditions dans lesquelles des projets subventionnés et des projets non subventionnés de bases de données factuelles sont en concurrence sur les marchés internationaux (*Leeftang et Gilmore*).

Dans le domaine scientifique, l'édition électronique est dans une certaine mesure freinée par la réticence des chercheurs. Le fait que le réseau Internet aît la réputation de contenir des matériaux d'une qualité douteuse n'a pas aidé à persuader les scientifiques qu'il s'agit d'un vecteur approprié pour diffuser des travaux sérieux. En ce qui concerne la diffusion électronique de leurs travaux avant publication, ils craignent que les éditeurs voient dans cette initiative une publication préalable et qu'ils refusent de les publier. De plus, la reconnaissance et l'avancement des chercheurs reposent encore dans une très large mesure sur la publication de leurs travaux dans des revues reconnues, qui sont pratiquement toutes des revues sur support papier (*Harnad et Hemus, 1998*). Les pouvoirs publics pourraient apaiser ces craintes en accordant des subventions d'aide à la recherche pour couvrir les frais de publication des résultats de recherches dans des revues électroniques soumises à un comité de lecture, et en prenant en compte ces publications électroniques lorsqu'ils évaluent la productivité des universités en matière de publication de travaux de recherche en vue de déterminer l'aide à leur accorder (*Harnad et Hemus, 1998*).

Les mesures actuelles sont inadéquates pour assurer l'archivage et la disponibilité à long terme des publications et des communications de recherches numériques. Il est essentiel qu'éditeurs et scientifiques collaborent à l'établissement de principes de base couvrant l'archivage et la disponibilité à long terme. En outre, un financement public est essentiel pour créer et maintenir des archives des publications électroniques dans le domaine scientifique (*CIUS, 1998*). Dans les disciplines où une large diffusion des prétrages numériques est possible, les pouvoirs publics peuvent souhaiter accélérer le passage à l'électronique des revues de recherche soumises à un comité de lecture en subventionnant le développement d'archives de prétrages électroniques.

Bibliothèques numériques

Les progrès rapides des TIC ont permis de traiter de gros volumes d'informations et de données numériques à des vitesses de plus en plus grandes et pour des coûts moindres. Le stockage, le filtrage, le traitement, la compression et la recherche de données numériques pour analyse et retransmission ont été considérablement facilités. Les bases de données numériques constituent des outils de plus en plus puissants mis à la disposition d'une base de plus en plus large d'ingénieurs et de scientifiques, qui autorisent de nouvelles façons de travailler en offrant aux chercheurs un accès accru à l'information et améliorent considérablement l'efficacité des travaux reposant sur ces informations. Les outils Internet, en particulier, ont rendu l'information plus immédiatement accessible car les prestataires de services d'informations ont commencé à adopter des systèmes fonctionnant sur le Web. Les méthodes d'établissement du prix de l'information évoluent également et leur impact doit être pris en compte car il risque de limiter l'accès aux sources primaires d'information.

Les bibliothèques numériques stockent et manipulent des collections importantes de matériaux sous une forme de plus en plus électronique et elles permettent aux utilisateurs d'avoir rapidement accès à ces ressources quand ils le souhaitent et là où ils le souhaitent. Leur développement est étroitement lié à celui des systèmes d'information en réseau et à la capacité d'intégrer aux services traditionnels des services implantés sur la toile (le « net »). Ils se présentent de plus en plus comme des projets dans lesquels les bibliothèques établies coopèrent pour offrir un accès en réseau à des collections de ressources disparates sur des systèmes disparates (pouvant être, par exemple, à la fois des systèmes sur support papier et des systèmes numériques) par le biais d'une interface unique pour le profit des utilisateurs, et comme le moyen de capitaliser sur les investissements par la mise en commun de ressources. Cette évolution conduit à un besoin de formation et à des changements d'organisation qui sont essentiels pour bien tirer parti des investissements dans les TIC.

Les chercheurs ont souligné à maintes reprises la nécessité de maintenir le rôle central des bibliothèques lorsqu'elles se tournent vers les technologies numériques car elles sont les principaux outils qui leur permettent de projeter vers l'avenir leurs connaissances passées. Davantage de réflexions et d'efforts sont nécessaires pour réaliser ce lien efficacement dans un contexte de numérisation croissante des informations. Cela est particulièrement vrai lorsque tout le savoir passé ne peut être numérisé. A l'heure actuelle, le simple fait de rechercher des articles in extenso dans des bibliothèques électroniques peut être très laborieux. De même, les moyens appropriés de stockage et d'accès aux quantités énormes de données collectées font cruellement défaut dans certains domaines, notamment dans les sciences sociales.

Plusieurs gouvernements ont lancé des projets de bibliothèques numériques reposant sur des systèmes d'accès et des politiques de numérisation différents. Le Danemark (la bibliothèque danoise de recherche électronique), la France (Bibliothèque nationale de France), le Royaume-Uni (Initiatives en faveur de l'accès à la bibliothèque britannique) et les États-Unis (la bibliothèque américaine du Congrès) en sont quelques exemples. L'article de *Thorhaug* décrit le projet en coopération de trois ministères danois (Recherche, Culture et Éducation) qui vise à assurer aux chercheurs un accès facile, à partir de n'importe quel ordinateur, à l'intégralité du contenu des bibliothèques danoises fonctionnant en réseau. Une subvention d'un montant de 32 millions de dollars sur cinq ans assure le financement du projet. Le projet couvre l'infrastructure nationale, l'infrastructure des bibliothèques, les ressources numériques et les installations des utilisateurs.

Les projets de bibliothèques numériques visent à faire en sorte que les chercheurs puissent avoir accès à partir de leur ordinateur aux ressources publiées des différentes bibliothèques de recherche traditionnelles et au volume de plus en plus considérable d'informations électroniques diffusées en ligne. Le développement des bibliothèques numériques et des recherches afférentes prennent racine dans les initiatives d'infrastructure de l'information du début des années 90. On peut citer, par exemple, la *Digital Library Initiative* américaine centrée dans un premier temps sur les problèmes d'infrastructure des bibliothèques numériques de grande taille. Durant les quatre années qui ont constitué la première phase de cette initiative, la NSF, l'ARPA (*Advanced Research Projects Agency*) et la NASA ont financé des six projets chacune à hauteur d'environ 1 million de dollars par an.

Pour la deuxième phase, l'accent est mis non plus sur la technologie mais sur le contenu et sur les utilisateurs, en encourageant les partenariats de recherche entre l'université et l'industrie, en particulier avec les technologies interactives. Le but est non seulement de profiter des fonds supplémentaires que les entreprises peuvent être disposées à apporter et d'accélérer la commercialisation mais, ce qui est plus important, d'éviter d'être rattrapé par des évolutions commerciales rapides et imprévues.

Parmi les différentes composantes des bibliothèques numériques, c'est le contenu qui semble poser le plus de problèmes. En effet, bien qu'en principe les TIC puissent offrir aux utilisateurs, quel que soit l'endroit où ils se trouvent, l'accès à toutes les informations existant sur un sujet donné, ce potentiel n'a pas encore été pleinement réalisé. Des versions électroniques des publications traditionnelles et de nouveaux types de recherche automatique sont mis sur le marché. Les méthodes de conversion des articles *in extenso* et leur financement nécessitent un examen plus approfondi et, en particulier, une coopération internationale afin d'éviter que les mêmes tâches soient effectuées en parallèle dans de nombreux pays. Les matériels actuellement considérés comme faisant partie des bibliothèques numériques n'incluent pas les bases de données qui contiennent des données brutes

et il n'est pas certain qu'ils devraient en faire partie et s'il s'agit là d'un sujet de préoccupation pour les pouvoirs publics ou plutôt pour la profession des bibliothécaires. On se demande également dans quelle mesure les bibliothèques doivent être responsables de leur propre adaptation et de leur conversion au numérique sur leurs propres fonds ou grâce à des subventions publiques supplémentaires. Il en va de même pour le financement de la recherche, du développement et de la mise en œuvre d'outils d'indexage et de recherche. La construction de bibliothèques numériques est une tâche complexe et coûteuse qui nécessite un engagement à long terme.

Bases de données

Les bases de données bibliographiques (secondaires) fournissant des informations sur les travaux de recherche publiés sous la forme de citations, de résumés et de différents index permettent aux scientifiques et aux chercheurs d'identifier les publications correspondant à leurs besoins. Leur taille varie de plus de 20 millions d'enregistrements dans le *On-Line Computer Library Centre* à des millions d'entrées dans les bases de données en ligne pour disciplines spécifiques. Les TIC ont étendu les méthodes d'acheminement des données et offrent de meilleures possibilités de stockage, de recherche et d'extraction des données, comme le montre le cas des bases de données mathématiques. Dans de nombreux domaines, les scientifiques produisent désormais des ensembles de données factuelles auxquelles les scientifiques du monde entier peuvent avoir accès via Internet. Le réseau Internet fournit également à des scientifiques travaillant dans des pays différents de nouvelles possibilités de fusionner des ensembles locaux de données pour obtenir des ensembles globaux, en particulier dans le cadre de projets de recherche qui nécessitent des données en provenance du monde entier, comme cela est le cas dans le domaine des sciences biologiques et des sciences liées à la terre (Projet de génome humain, Programme international pour la géosphère et la biosphère). Des bases de données factuelles à grande échelle pourraient également transformer la recherche dans les sciences sociales mais à l'heure actuelle, elles font défaut dans ces disciplines.

Mathématiques : bases de données bibliographiques³

Les mathématiques sont un domaine particulièrement exigeant en matière d'informations bibliographiques car les documents écrits y jouent un rôle très spécifique. L'établissement d'une preuve rigoureuse, qui est l'objectif final de cette discipline, est tout aussi important aujourd'hui qu'il l'était dans le passé. Les travaux mathématiques publiés contiennent toutes les informations nécessaires au lecteur compétent pour procéder à un contrôle approfondi des données

22

(Bourguignon/Société européenne des mathématiques). Cela n'est évidemment pas le cas pour les travaux expérimentaux. De nouveaux concepts sont aussi développés dans le domaine des mathématiques mais ils ne remplacent pas automatiquement les concepts utilisés précédemment, comme cela est souvent le cas dans d'autres disciplines.

Bien que la nécessité de se référer à une littérature publiée antérieurement ne soit pas aussi aiguë dans toutes les disciplines, le cas des mathématiques met en lumière de nombreux problèmes et principes qu'il convient de considérer lorsqu'on tire parti des opportunités techniques offertes par les TIC dans l'élaboration de bases de données pour les diverses disciplines. Les TIC ont eu un impact particulier sur les méthodes d'acheminement des données et offrent des solutions plus performantes pour le stockage, la recherche et l'extraction des travaux de recherche de par le monde.

Les mathématiciens utilisent très largement deux bases de données biographiques comme outils de recherche. Les *Mathematical Reviews* nord-américaines et la *Zentralblatt-MATH* européenne, qui portent sur l'ensemble des publications mathématiques des années 1930 à ce jour. Ces bases de données sont de plus en plus accessibles par le biais d'un système international de doubles serveurs sur le réseau Internet (Berlin, Strasbourg, New York). Des serveurs locaux, soit régionaux soit desservant des institutions spécifiques, offrent un accès hautement performant, via le CD-ROM. Chaque notice comprend une analyse rédigée par un mathématicien spécialiste de la question. Un réseau de plus de 5 000 scientifiques apportent leur contribution à ces revues.

La possibilité pour les chercheurs d'accéder facilement à ces bases de données pour vérifier une affirmation, évaluer la production d'un chercheur ou se faire une idée des liens scientifiques existant entre deux sujets, pour ceux qui ne sont pas experts dans un domaine donné, est bien souvent à l'origine d'une découverte majeure. L'exploitation des deux bases de données coûte environ 1 million de dollars par an. Contrairement aux 60 000 notices ajoutées chaque année à chaque banque de données, la numérisation d'archives antérieures couvrant 200 000 notices entre 1868 et 1942 est actuellement financée sur les deniers de l'État car une implication de l'industrie aurait abouti à un produit d'un coût beaucoup trop élevé pour une utilisation répandue.

Plusieurs critères permettent d'évaluer la qualité d'une base de données. Sa globalité à l'intérieur des limites du domaine est très importante. La qualité de ses réviseurs est essentielle – les réviseurs ne sont généralement pas payés mais ils peuvent recevoir des maisons d'édition concernées quelques avantages annexes – et la participation d'experts du monde entier est un plus. La compétence professionnelle de l'équipe de rédacteurs qui distribue les articles aux réviseurs est également importante, de même que la vitesse de traitement des articles (dans les mathématiques, le délai standard est de six à huit mois à compter de leur publication

dans une revue). En outre, la classification et l'organisation des matériaux dans une base de données déterminent dans une large mesure leur exploitabilité tant pour les recherches, qui sont l'un des principaux avantages des nouveaux formats électroniques, que pour les consultations.

Le logiciel utilisé pour la recherche automatique est essentiel à une utilisation efficace de la base de données ; il doit satisfaire à un certain nombre d'exigences, notamment de facilité d'utilisation, d'efficacité, de souplesse et de robustesse. La base de données doit être rapide mais elle doit également ne comporter aucune ambiguïté quant à la manière dont les noms des auteurs, les titres des revues, etc., sont entrés dans la base. Cela n'est possible que si l'on utilise un thésaurus spécifique pour ces catégories. Celui-ci doit néanmoins être capable de reconnaître des noms entrés dans la base de données de manière différente ou une transcription à partir d'autres alphabets ou d'idéogrammes ayant varié au fil des années. Il doit être capable d'agir en interaction avec le logiciel standard existant et ne doit pas être excessivement sensible à la syntaxe.

Une autre question essentielle concerne le contrôle des bases de données. Dans la mesure où le contenu de ce type de base de données provient de la communauté scientifique elle-même, il pourrait être approprié que des sociétés savantes soient impliquées et détiennent les droits de copyright, comme dans le cas des mathématiques. Mais le danger pourrait être qu'un groupe cherche à imposer ses points de vue scientifiques sur un thème particulier. Les bases de données constituent un levier puissant et c'est là un bon argument en faveur d'un système concurrentiel à moins que des mesures appropriées ne soient prises pour préserver une gestion des bases de données indépendante des lobbies.

En ce qui concerne la fourniture de l'accès aux pays en développement, un accord de licence adopté pour des bases de données mathématiques mérite d'être noté. Un établissement appartenant à un pays mieux loti peut être jumelé à un établissement appartenant à un pays en développement, qui ne paie qu'un petit pourcentage du prix normal. Les deux établissements partagent certains services tels que la formation. Bien que le revenu supplémentaire pour le propriétaire de la base de données tende à être marginal, celui-ci peut toucher de nouveaux utilisateurs en leur rendant les produits accessibles.

Une étape critique de l'amélioration des services qu'offrent les bases de données bibliographiques a trait à l'interface avec la littérature primaire. L'établissement de liens développés (*hyperlinks*) entre les bases de données de revues et les articles *in extenso* augmentera considérablement l'efficacité de la recherche de références. Mais l'accès aux articles *in extenso* demeure difficile à ce stade.

Sciences sociales : bases de données à grande échelle

Un grand nombre de problèmes auxquels les gouvernements des pays de l'OCDE doivent s'attaquer aujourd'hui ont trait aux sciences sociales, c'est-à-dire à l'ensemble des disciplines qui traitent du comportement et des interactions des hommes et des institutions. Le développement durable, l'évolution du marché du travail, l'instabilité familiale et le vieillissement de la population ne sont que quelques exemples des problèmes majeurs auxquels sont confrontés les pays Membres et que ces sciences peuvent aider à résoudre.

Pourtant, les sciences sociales sont en général moins bien loties par rapport aux sciences naturelles en termes d'affectations budgétaires (en particulier en temps de crise), en termes d'impact politique et d'attrait auprès des étudiants. Dans une certaine mesure, les sciences sociales et les humanités reçoivent moins de financements que les autres disciplines scientifiques car elles n'exigent pas le même investissement en équipements (instruments et installations). Et pourtant, les sciences sociales ont leurs propres besoins en infrastructures de recherche ; elles peuvent notamment bénéficier de bases de données à grande échelle. Pour traiter de problèmes qui ont trait aux individus, à leur famille, à leur lieu de travail et à leurs entreprises, il faut effectuer des enquêtes auprès d'un grand nombre de personnes, sur de nombreuses dimensions et stocker les données tirées de ces enquêtes dans de très grandes bases de données (OCDE, 1999).

L'inadéquation des données et des possibilités de mise en commun de celles-ci constituent un obstacle majeur à l'intégration des sciences sociales dans l'élaboration de la politique. Des bases de données largement partagées font encore défaut. Dans la plupart des pays de l'OCDE, les bases de données existantes sont fragmentées et ne sont accessibles qu'à des cercles d'utilisateurs relativement étroits. En particulier, ces bases de données sont restées inaccessibles pour de nombreux types d'analyses secondaires que requiert l'élaboration d'options d'action à soumettre aux gouvernements. La plupart de ces ensembles de données sont tirés d'enquêtes nationales différentes, relativement limitées et coûteuses ayant généré des informations qui ne sont pas nécessairement compatibles les unes avec les autres (*Bainbridge*).

Les nouvelles TIC ont fourni la capacité technique de construire des bases de données électroniques complètes dans de nombreuses disciplines scientifiques. Des outils des TIC comme le réseau Internet permettent à pratiquement tous les chercheurs qualifiés d'avoir accès à des bases de données dispersées sur les sciences sociales. Ils peuvent également constituer un moyen moins coûteux de réaliser de nouvelles enquêtes. Les TIC peuvent en outre être utilisées pour appliquer les méthodologies expérimentales à l'étude de vastes réseaux d'interactions humaines et permettre de traiter un nombre de sujets beaucoup plus important qu'avant le développement et l'utilisation extensive de ces technologies et d'analyser leurs réponses.

Les études de marché, les systèmes d'élection et les réseaux sociaux sont autant de sujets qui devraient bénéficier de ces progrès.

L'utilité des bases de données pour les sciences sociales sera considérablement améliorée si leur portée est internationale, c'est-à-dire si des bases de données analogues existant dans différents pays sont reliées entre elles et si des chercheurs travaillant dans différents pays peuvent y avoir accès. Les organismes publics de certains pays ont entrepris récemment d'élargir leur soutien aux enquêtes pour couvrir la diffusion des résultats via le Web. Des chercheurs et des étudiants travaillant dans différentes régions du monde peuvent ainsi avoir accès à des données qu'ils peuvent soit analyser en ligne soit télécharger sur leurs ordinateurs, mais aussi à des résultats publiés, à des outils d'analyse et à des références.

Les défis que pose le développement de bases de données électroniques à grande échelle dans les sciences sociales sont nombreux. L'accent doit être mis sur la qualité des données, notamment sur leur fiabilité et leur normalisation en sus de leur stockage et de leur recherche automatique. Un problème spécifique aux sciences sociales (et aux sciences médicales), en particulier lorsqu'on cherche à maximiser l'accès des experts en sciences sociales et des responsables de l'action publique aux données sur les individus, est le respect de la vie privée. Quant aux enquêtes sur les entreprises, des questions de confidentialité des données peuvent se poser. Les questions relatives aux droits de propriété intellectuelle sur les bases de données doivent également être traitées. Enfin, et ce n'est pas le moindre des problèmes, il pourrait être nécessaire de développer un nouvel ensemble de valeurs culturelles vis-à-vis de la collecte et du partage des données (OCDE, 1999).

Des consultations entre gouvernements doivent être organisées pour traiter des problèmes techniques et politiques qu'impliquent une extension et une amélioration de l'efficacité de bases de données internationales largement accessibles, reliées électroniquement entre elles et contenant des données pertinentes pour la recherche en sciences sociales. A cet effet, un Atelier sur les infrastructures à grande échelle pour les sciences sociales, qui se tiendra à l'automne 1999 au Canada, examinera la relation existant entre le développement d'infrastructures dans les sciences sociales et son impact potentiel sur la productivité et sur les pratiques en matière de recherche.

Orientations de politique générale

Lors de la conférence de Sintra, les gouvernements⁴ ont souligné deux grandes opportunités qu'offre l'utilisation des TIC pour la science. L'une de ces opportunités est d'encourager la collaboration, non seulement à l'intérieur du système scientifique mais également entre le système scientifique et l'économie et la société

dans son ensemble. L'établissement de liens plus étroits entre le système scientifique et les entreprises, notamment l'établissement de partenariats entre l'université et l'entreprise pour le développement et l'utilisation des TIC, pourrait améliorer le processus d'innovation. Un resserrement des liens entre le système scientifique et la société réduirait l'isolement relatif de ce système, qui est perçue, l'une des raisons d'une compréhension et d'un soutien limités de ses activités par le public. Ces liens devraient également enrichir la culture scientifique et technologique, en se servant du fait que les ménages ont accès aux TIC et peuvent bénéficier de possibilités d'apprentissage tout au long de la vie. Les gouvernements reconnaissent néanmoins qu'il faut en premier donner la priorité aux besoins spécifiques de la communauté des chercheurs, qui est bien souvent le moteur du développement des TIC et de leurs applications connexes dans l'économie et dans la société.

Une autre grande opportunité est l'utilisation des TIC comme outil permettant de mieux intégrer les communautés scientifiques des pays peu développés dans le village mondial de la recherche. Pour les chercheurs de ces pays, qui n'ont guère la possibilité d'influencer leurs décideurs, le principal moyen de s'engager dans une coopération scientifique internationale est l'interaction directe avec les communautés scientifiques des pays développés, via le réseau Internet. Il est donc crucial de donner aux chercheurs des pays en développement les compétences et l'infrastructure en technologie de l'information requises pour participer aux projets internationaux de recherche.

Bien que les TIC offrent à l'entreprise scientifique des opportunités considérables, bon nombre de problèmes, dont certains avaient déjà été identifiés auparavant mais sur lesquels les progrès ont été lents, nécessitent une action des pouvoirs publics. Les participants à la Conférence de Sintra sont tombés d'accord sur les grandes lignes des tâches que les gouvernements devront accomplir⁵ :

- Mettre à la disposition des scientifiques une infrastructure hautement performante en matière d'information et de communication. A cet effet, les gouvernements devront fournir les financements nécessaires au développement de réseaux de recherche à grande vitesse, en particulier lorsque les réseaux commerciaux ne fournissent pas la capacité requise pour un prix raisonnable. Ils devront également assurer le leadership et le financement de la recherche et du développement des composantes de l'infrastructure pour lesquelles les chercheurs constituent la principale source de demande et que le secteur des entreprises est par conséquent peu susceptible de développer de sa propre initiative.
- Encourager la formation des scientifiques à l'utilisation des TIC, de sorte qu'ils soient capables et disposés à utiliser au mieux les outils des TIC. Bien que les scientifiques figurent parmi les travailleurs les plus qualifiés de la

zone OCDE et que l'enseignement universitaire porte une attention grandissante au développement des compétences en matière de TIC, les chercheurs peuvent n'être pas suffisamment conscients des potentialités qu'offrent ces technologies ou ne pas savoir les utiliser de la manière la plus efficace. Des travaux en collaboration exigent un investissement important dans l'apprentissage de l'utilisation des TIC. L'efficacité d'utilisation de ces technologies dépend également de plus en plus de l'existence d'un soutien technique. Les professionnels de l'information doivent donc être dûment formés pour apporter leur concours aux chercheurs, notamment pour l'organisation, la manipulation et l'accès rapide aux informations par le biais de nouveaux services fonctionnant en réseau.

- Stimuler l'offre de services et d'applications qui permettront d'améliorer les liens entre les sciences, les entreprises et la société en mettant en œuvre des cadres réglementaires appropriés. Les gouvernements de nombreux pays Membres de l'OCDE se sont déjà engagés dans une réforme de la réglementation affectant l'industrie des télécommunications, mais il reste encore beaucoup à faire.
- Faire en sorte que les cadres juridiques régissant les collaborations internationales en matière de recherche et de contrôle des informations en ligne, notamment les droits de propriété intellectuelle, les flux transfrontières de données, la protection de la vie privée, la sécurité de l'information et son authentification ne soient pas des entraves empêchant de tirer profit des avantages des TIC pour la science. Les gouvernements doivent continuer à rechercher une plus grande compatibilité entre les systèmes nationaux. Les développements ultérieurs de l'édition électronique, des bibliothèques numériques et des bases de données dépendent de la prise en compte des intérêts de la communauté scientifique dans le débat actuel concernant les droits respectifs des auteurs, des éditeurs, des bibliothécaires et des utilisateurs. Les gouvernements peuvent en outre, en coopération avec l'establishment scientifique et les entreprises commerciales, contribuer à lever les incertitudes relatives à l'accès aux infrastructures et aux bases de données, aux coûts d'utilisation, au contrôle de la qualité, à la reconnaissance et à la disponibilité à long terme, ainsi qu'à l'archivage des résultats des recherches diffusées en ligne.
- Entreprendre un effort international plus concerté pour réduire le fossé important qui existe entre le niveau de développement des infrastructures des pays développés et celles des pays en développement et réaliser une infrastructure globale et un village mondial de la recherche. Lorsqu'on considère l'ensemble des pays en développement, les disparités existant au niveau de l'infrastructure physique qui constitue la base de toutes les applications des TIC dans les sciences, sont frappantes. Un petit nombre de

projets, comme le réseau mondial du Projet 'Oxygène et le réseau APAN dans la région Asie-Pacifique, sont de bons exemples de ce que l'on peut accomplir dans ce domaine mais il reste beaucoup à faire et la réduction de ces disparités ne pourra être réalisée dans l'immédiat.

- Renforcer la coopération internationale en général pour éviter les duplications d'efforts et de financements dans des domaines tels que la conversion au numérique des résultats de recherche diffusés sur support papier afin de les incorporer aux bibliothèques numériques de recherche.
- Développer des indicateurs comparables au plan international pour une évaluation des développements et des performances, notamment du retour sur investissement dans ce domaine qui évolue rapidement. Il est nécessaire de pouvoir mesurer l'évolution de l'utilisation des TIC pour la recherche scientifique d'autant plus qu'elle est marquée par de nombreuses divergences de point de vues et d'organisation au niveau de l'infrastructure, du contenu et de l'accès entre les disciplines, les pays et les régions.

Nombre d'autres évolutions potentiellement bénéfiques pour la science risquent d'être sensiblement retardées si les pouvoirs publics n'accordent pas de financements supplémentaires. En règle générale, ces évolutions ont trait au développement et à la disponibilité à long terme du contenu numérique. Des finances publiques sont nécessaires pour la numérisation des matériaux sur support papier des bibliothèques numériques, l'archivage des publications électroniques, l'allongement de la durée de vie de bases de données importantes sinon uniques dans certaines disciplines, et la création de bases de données à grande échelle pour les sciences sociales.

Quelque soient les mesures que prendront les pouvoirs publics, il est difficile de profiter de tous les avantages qu'apportent les TIC aux systèmes scientifiques. Les technologies continuent d'évoluer à un rythme très rapide. Leurs potentiels et leurs limites sont encore incertaines. Les politiques à l'égard du système scientifique doivent prendre pleinement en compte ces incertitudes, et continuer à faire preuve de flexibilité.

Viviane Bayar
Jean-Éric Aubert

NOTES

1. Les informations concernant cette région ont été tirées de la présentation intitulée «Internet and Science in the Asia-Pacific Area: A Japanese Viewpoint», faite par M. H. Ishida de ASCII Future labs, Tama Art University, lors de la conférence de Sintra.
2. Cette section est extraite de la présentation intitulée «Maximiser les avantages de la technologie de l'information pour la science : le cas des bases de données mathématiques», faite par M. J.-P. Bourguignon, Président de la Société européenne des mathématiques, lors de la conférence de Sintra.
3. Cette section puise largement dans la présentation intitulée «Maximiser les avantages de la technologie de l'information pour la science : le cas des bases de données mathématiques», faite par M. J.-P. Bourguignon, Président de la Société européenne des mathématiques, lors de la conférence de Sintra.
4. M. José Mariano Gago, ministre portugais de la Science et de la Technologie, qui a présidé la conférence et M. Jan Trøjborg, ministre danois de la Recherche et des Technologies de l'Information.
5. Comme exprimé dans les conclusions adoptées à la fin de la conférence de Sintra.

BIBLIOGRAPHIE

- CHRONICLE OF HIGHER EDUCATION (1998),
«Scholars Seek New Copyright Rule to Ease Dissemination of Research on the Web»,
11 septembre.
- CIUS (1998),
«Executive Summary Report, ICSU Press Workshop on the Economics, Real Costs and
Benefits of Electronic Publishing in Science – A Technical Study», hors série du *Science
International Newsletter*, n° 68, Paris.
- CIUS (1998a),
«Meeting Reports», AAAS/ICSU Press/UNESCO Workshop on Electronic Publishing in
Science, n° 69, Paris.
- HARNAD, S. et M. HEMUS (1997),
«All Or None: No Stable Hybrid or Half-way Solutions for Launching the Learned Peri-
odical Literature into the Post Gutenberg Galaxy», [ftp://ftp.princeton.edu/pub/harnad/
HTML/harnad97.hybrid.pub.html](ftp://ftp.princeton.edu/pub/harnad/HTML/harnad97.hybrid.pub.html), 16 décembre.
- JAPAN SCIENCE AND TECHNOLOGY AGENCY (1997),
White Paper on Science and Technology 1997, Tokyo.
- KLING, R. et G. McKIM (1997),
«Electronic Media and Scientific Communication: Opportunities, Complexities and
Limitations», papier écrit pour l'OCDE, Indiana University School of Library and Infor-
mation Science, septembre, miméo.
- OCDE (1999),
Perspectives des communications 1999, Paris.
- OCDE (1998),
«The Global Research Village», http://www.oecd.org/dsti/sti/s_t/scs/prod/online.htm.
- OCDE (1998),
Perspectives de la science, de la technologie et de l'industrie 1998, Paris.
- OCDE (1999),
The Social Sciences at a Turning Point? Paris.

INFRASTRUCTURE DE L'INFORMATION POUR LE VILLAGE MONDIAL DE LA RECHERCHE : LA PERSPECTIVE EUROPÉENNE

Table des matières

I. Introduction	34
II. Le cinquième Programme-cadre.....	36
III. Les parties en présence.....	39
IV. Expériences et comparaisons.....	41
V. Le réseau d'interconnexion européen à haut débit.....	46
VI. Incidence pour l'Europe.....	48
VII. Conclusion.....	50
Notes	52
Bibliographie.....	53

Cet article a été rédigé par Hans-Peter Axmann et Sabine Payr. M. Axmann est Directeur de la division Informatique et Réseaux au sein du ministère fédéral autrichien des Sciences et des Transports et président de l'*European Group for Policy Co-ordination for Academic and Industrial Research Networking* (Groupe européen de coordination des réseaux de recherche universitaires et industriels – ENPG) (adresse électronique : hans-peter.axmann@bmf.gv.at). Mme Payr, consultante pour la télématique et les applications multimédia dans l'enseignement supérieur, est actuellement chargée de recherches auprès de la *Research Association for Computer Science* (Association de recherche en informatique). Elle a assumé les fonctions de secrétariat pour l'ENPG jusqu'en décembre 1998 (adresse électronique : s.payr@fgi.at).

I. INTRODUCTION

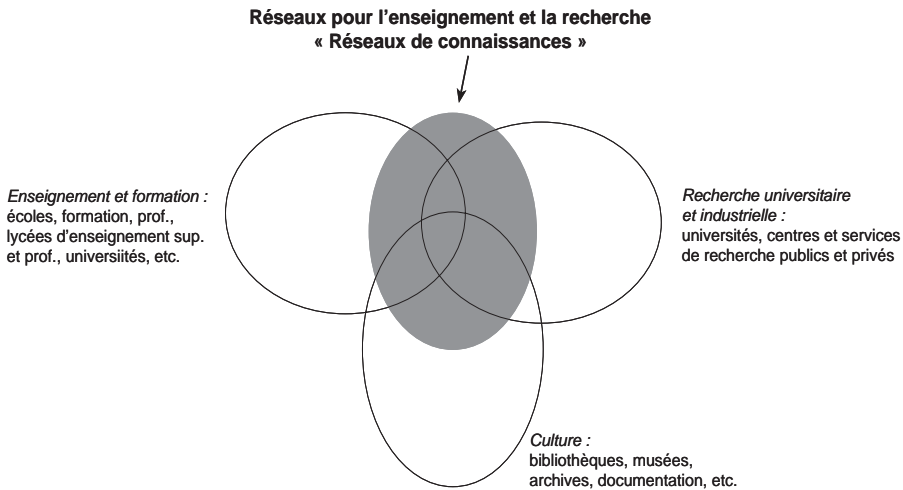
Les pays industrialisés sont fortement tributaires des innovations et des succès de la recherche et ils ont un intérêt vital à ce qu'elle livre une profusion d'excellents résultats. Dans la plupart des cas, ceux-ci se traduisent directement ou indirectement par de nouveaux produits industriels et contribuent ainsi nettement au développement économique. La recherche est donc également une condition préalable importante à la création de nouveaux emplois et à la préservation des emplois existants, ainsi qu'au maintien et à l'amélioration de la qualité de vie dans les pays de l'Union européenne. Elle est par ailleurs la réponse aux défis mondiaux soulevés par la concurrence technologique et économique.

Les gouvernements des pays industrialisés sont par conséquent tenus de créer un environnement favorable à la recherche et à l'innovation. La recherche ne peut en effet s'épanouir et être couronnée de succès que dans un environnement propice.

Pour instaurer un tel environnement, différentes méthodes et mesures sont envisageables. Elles vont des incitations fiscales en faveur des activités scientifiques jusqu'à l'implication indirecte et même directe des gouvernements dans les questions concernant la recherche (par exemple, la poursuite de politiques nationales pour la promotion de la recherche et le développement technologique, la stimulation des activités de recherche dans l'industrie, à la fois pour les grandes entreprises et les PME et enfin, la formation d'instituts de recherche investis de missions spécifiques).

Des réseaux à haut débit dédiés à la communication électronique sont devenus indispensables à cet environnement et presque toutes les recherches en sont largement tributaires. Aujourd'hui, les réseaux de recherche sont des outils importants, non seulement pour toutes les disciplines universitaires – Les sciences naturelles aux humanités en passant par la médecine – mais aussi pour l'interconnexion des bibliothèques, des musées et des institutions pédagogiques ou culturelles. Si l'on prend en compte les aspects liés à la « mission » de ces institutions, il semblerait plus approprié de parler de *réseaux de connaissances* (Figure 1). Cette terminologie n'étant pas encore communément acceptée, nous opterons dans cet article pour l'expression « réseaux de recherche ».

Figure 1. Réseaux de connaissances



Source : Auteurs.

Les fonds publics ont permis le développement rapide de plusieurs réseaux de recherche à l'échelon national. Cependant, faute d'une autorité nommément désignée, un développement similaire n'a pas été observé au niveau européen. D'où une structure de réseau hétérogène, pénalisant les pays dénués d'infrastructure de réseau nationale appropriée. Leur développement scientifique pâtit déjà d'une insuffisance de moyens de communication et l'on peut craindre que l'écart entre ces pays et ceux qui sont mieux équipés s'aggrave avec le temps.

La promotion des réseaux de recherche au niveau européen s'est limitée, jusqu'ici, aux divers programmes communautaires en faveur de la recherche et du développement technologique. En fait, le véritable premier réseau de recherche transeuropéen, IXI, a été constitué en 1990 dans le cadre du projet Eureka COSINE, commencé en 1987 et terminé en 1993. IXI a évolué pour devenir EuropaNET en 1992 et véhiculer le trafic à 2 Mbps suivant les protocoles X.25 et IP. EuropaNET, à son tour, a été remplacé par TEN-34 en 1997. Ce dernier, qui assure un débit nominal de 34 Mbps, a pu être créé uniquement grâce au soutien du 4^e PCRD (quatrième Programme-cadre de la Communauté économique européenne pour des actions communautaires de recherche, de développement technologique et de démonstration).

L'importance de l'établissement de réseaux pour la recherche et le développement technologique étant plus largement admise, cette question a été inscrite dans le 4^e PCRD. Les actions en faveur des réseaux de recherche sont devenues possibles sous la forme de projets spécifiques. Cependant, il faut distinguer deux aspects fondamentalement différents dans ce domaine :

- *Recherche sur l'établissement de réseaux* : développement et expérimentation de technologies, services et applications avancés dans le domaine des réseaux.
- *Réseaux pour la recherche* : mise en œuvre de technologies avancées d'établissement de réseaux dans toutes les disciplines et tous les domaines de la recherche.

Le quatrième Programme-cadre n'a ni considéré ces deux aspects, ni établi une distinction claire entre eux. Cependant, le soutien des réseaux de recherche (le second aspect) n'a été possible qu'indirectement, par la recherche sur l'établissement des réseaux (le premier aspect). C'est l'une des raisons pour laquelle le comité chargé de l'étude du Programme Applications Télématiques a qualifié le secteur de la recherche de « particulier » et conclu que « le travail d'infrastructure est si vaste qu'il pourrait constituer à lui seul une activité spécifique à part entière. » (Programme Applications Télématiques, 1996).

II. LE CINQUIÈME PROGRAMME-CADRE

La proposition de 5^e PCRD (Commission européenne, 1997) montre que ces commentaires et d'autres observations similaires sont maintenant pris en compte. Tous les programmes thématiques proposés jusqu'ici comportent un axe d'intervention spécifique en faveur de l'infrastructure de recherche. Le programme concernant la Société de l'Information, en particulier, cite les réseaux à large bande comme un élément indispensable de cette infrastructure.

Dans le processus de définition du 5^e PCRD, il est apparu clairement qu'il fallait adopter une approche différenciée concernant la méthode de mise en œuvre des actions sélectionnées et les instruments qui seront utilisés pour atteindre les objectifs fixés. Suivant la distinction établie plus haut, il est clair que :

- La recherche concernant l'établissement de réseaux doit demeurer un thème prioritaire. Elle se laisse assez facilement intégrer dans les dispositions des programmes-cadres, par le biais habituel du soutien par projet. Toutefois, la globalisation croissante exige que l'on se préoccupe plus de l'acceptabilité des résultats et des produits au niveau international, de manière à éviter des solutions isolées pour l'Europe.
- Les réseaux dédiés à la recherche doivent être considérés tout à la fois comme une technologie structurante et une infrastructure nécessaire. En

tant qu'outils essentiels pour la coopération entre les États membres et pour la recherche et le développement technologique, les réseaux de recherche doivent être considérés comme une activité transcendant l'ensemble des programmes et orientations thématiques, car il est évident que l'établissement de réseaux pour la recherche sert également tous les autres thèmes et actions figurant dans le 5^e PCRD.

Par ailleurs, une autre distinction s'impose. La longue expérience des communautés de la recherche et de l'enseignement, servies par un réseau fédérateur européen, permet de penser qu'il faut répondre à deux exigences de nature différente. Il est clair que la grande majorité des utilisateurs finals ont besoin d'un réseau de « production » stable et efficace pour leur travail quotidien. D'un autre côté, un groupe relativement réduit d'utilisateurs, notamment ceux qui travaillent dans la recherche sur l'établissement de réseaux et la mise au point de nouvelles applications en réseau, mais aussi ceux qui réalisent des essais et innovent dans la manière d'utiliser les applications existantes, ont besoin d'une infrastructure pour l'expérimentation et le test et donc, d'un réseau du type « banc d'essai ».

Par le passé, certains ont soutenu que ces modes d'utilisation différents justifiait l'établissement de réseaux distincts. Après examen, il apparaît toutefois clairement que les réseaux de production et les réseaux d'essai ont tant de caractéristiques en commun qu'il est judicieux de les installer sur la même infrastructure physique ; comme il est nécessaire de gérer ces réseaux dans les deux cas, il faut confier cette tâche à ceux qui se sont montrés capables d'administrer des réseaux à vocation scientifique, autrement dit les NREN (exploitants des réseaux desservant la recherche et l'enseignement à l'échelon national) et DANTE. Bien que la gestion du trafic Internet ne puisse actuellement faire l'objet que d'une obligation de moyens, les utilisateurs « ordinaires » ont eux aussi parfois besoin d'une certaine bande passante et d'une certaine qualité de service pour certaines applications. Les problèmes de gestion qui en résultent ressemblent à ceux que soulèvent l'établissement de liaisons spécialisées pour des applications de test isolées. En outre, une application testée avec satisfaction dans un milieu « sécurisé » est plus facile à transposer dans un milieu « réel », si les deux milieux utilisent la même infrastructure physique de base. Enfin, cette solution serait de loin plus économique sur le plan financier.

Compte tenu de leur portée limitée dans le temps, les instruments de soutien par projet ne peuvent refléter de manière appropriée les besoins des réseaux de recherche. Il faudrait adopter une approche plus directe et plus souple, du type « groupe de projet », et qui garantisse le soutien requis tout au long du 5^e PCRD. Qui plus est, il semble qu'il règne aujourd'hui un consensus pour que le soutien à l'infrastructure fasse l'objet de marchés publics.

Une condition préalable essentielle dans ce contexte est toutefois que le financement nécessaire à un réseau transeuropéen étendu à haut débit puisse être assuré sur au moins la durée du programme-cadre. Comme cela est dans l'intérêt aussi bien des autorités nationales que de la Commission européenne, un modèle de cofinancement a été suggéré (Axmann *et al.*, 1997). Si l'on suit l'exemple et si l'on tire les enseignements du TEN-34, un modèle de financement partagé semblerait le plus approprié. Suivant ce scénario, les États membres de l'UE contribueraient à hauteur de 60 % du total des coûts par leurs NREN et la Commission financerait les 40 % restants. Le même modèle devrait aussi s'appliquer à toutes les institutions de recherche européennes pour qu'elles puissent bénéficier de cette infrastructure.

A partir de données fournies par DANTE, on a été estimé qu'un réseau de transmission à 155 Mbps entre tous les États membres (avec deux lignes chacun pour l'insensibilité aux défaillances) coûterait chaque année quelque 125 millions d'ECU. En utilisant le modèle de partage des coûts 40-60 % décrit ci-dessus, la contribution de la CE s'élèverait à 50 millions d'ECU par an, soit au total 200 millions d'ECU sur les quatre ans du 5^ePCRD. Comparé au budget de 4 000 millions d'ECU initialement proposé par la CE pour l'ensemble du programme TSI (Technologies de la Société de l'Information), la contribution en faveur de l'infrastructure de réseau serait de 5 %, ce qui semblerait un niveau de frais très acceptable. Toutefois, les contraintes budgétaires dans la plupart des États membres ont réduit ce chiffre à celui de la position commune acceptée par le Conseil de l'Europe en février/mars 1998. De plus, les pourcentages utilisés dans la ventilation du budget donnée à titre indicatif le 8 juin 1998 montrent que le Conseil est prêt à affecter 4 % du budget TSI au soutien de l'infrastructure, alors que la CE n'a pour l'instant proposé que 3 %. En termes réels, les montants équivalents s'élèvent respectivement à 135 et 118 millions d'ECU. Cela est nettement insuffisant pour atteindre l'objectif souhaité.

Le récent processus de conciliation entre le Conseil et le Parlement européen s'est traduit par l'affectation de 3 600 millions d'ECU en faveur du programme TSI. L'application des pourcentages précédemment indiqués, à savoir 3 % (CE) et 4 % (suggestion de la présidence du Royaume-Uni), donnerait respectivement 108 et 144 millions d'ECU ; il est clair que la situation ne s'est pas améliorée de manière spectaculaire.

Le 22 décembre 1998, sous la présidence de l'Autriche, le Conseil des ministres de la Recherche a finalement décidé, au moment de statuer officiellement sur le 5^e PCRD, d'affecter 4.5 % de ce budget au programme TSI pour soutenir la recherche en matière d'établissement de réseaux, soit 161 millions d'ECU. Les services de la CE se sont montrés confiants quant à la possibilité de trouver des fonds supplémentaires par ailleurs. Cela semble assez prometteur, même si l'on ne sait pas

encore vraiment si la totalité de cette somme sera affectée à l'infrastructure de réseau ou quelle proportion sera dédiée à d'autres activités.

III. LES PARTIES EN PRÉSENCE

L'expérience a montré que la communauté de la recherche et de l'enseignement supérieur est celle qui avait besoin de l'infrastructure de réseau la plus avancée. Cependant, même un an après la libéralisation du marché des télécommunications en Europe, cette infrastructure n'est pas directement disponible comme un bien marchand. Les mécanismes habituels du marché ne s'appliquent donc pas. La CE et ses états membres ont reconnu la nécessité d'un (co-) financement public durable pour les réseaux de recherche. L'actuel PCRD, le 5^e en l'occurrence, devrait enfin fournir l'opportunité de financer cette infrastructure. La Commission sera responsable du cofinancement et des mécanismes de sa mise en oeuvre.

Il est essentiel que toutes les parties en Europe concernées par l'interconnexion des réseaux dans l'enseignement supérieur et la recherche conjuguent et harmonisent leurs efforts afin de planifier et de mettre en oeuvre cette infrastructure. Actuellement, les acteurs dans le domaine de l'établissement de réseaux de recherche sont les suivants :

- Les gouvernements nationaux et leurs organismes publics respectifs : comme indiqué plus haut, ce sont les responsables des réseaux de recherche à l'échelon national. Il est clair que les gouvernements nationaux doivent harmoniser leurs activités au plan européen, en particulier en ce qui concerne la connectivité intercontinentale et au sein même de l'Europe.
- L'ENPG (*European Group for Policy Co-ordination for Academic and Industrial Research Networking*) : groupe constitué par la plupart des pays européens comme un instrument de coordination des politiques nationales, il peut donc être considéré comme le représentant des gouvernements nationaux et des organismes publics¹.
- La Commission européenne (CE), qui peut intervenir dans le cadre du Traité (ou des Traités), pour promouvoir la recherche et soutenir l'infrastructure correspondante à l'échelon communautaire.
- Les NREN, qui sont les opérateurs des réseaux pour la recherche et l'enseignement supérieur au plan national.
- TERENA, qui est l'organisation réunissant les opérateurs de réseaux de recherche européens. Elle a essentiellement pour objectif de favoriser une meilleure interconnexion des réseaux de recherche sur le plan international,

pour le compte de ses membres et de ses utilisateurs. Sa mission consiste par conséquent à promouvoir et à participer au développement d'une infrastructure internationale d'information et de télécommunications de haute qualité au profit de la recherche et de l'enseignement.

- DANTE, qui est une société à but non lucratif, a été établie en 1993 par les réseaux de recherche nationaux en Europe dans le but de mettre en place des services internationaux sur réseaux pour les chercheurs européens.
- A ce jour, « l'industrie » n'a pas été un acteur très actif sur le plan des réseaux européens de recherche. On peut espérer, comme des exemples récents semblent l'indiquer, que l'importance d'une action et le potentiel pour le développement et les marchés industriels seront reconnus. L'industrie n'est pas seulement utile en tant que pourvoyeur de l'infrastructure physique, elle a aussi beaucoup à gagner comme partenaire. Pour qu'elle devienne un associé sérieux, il faudra toutefois que les conflits réels ou imaginaires entre les besoins des réseaux de recherche et les intérêts commerciaux soient aplanis. Sur le marché européen des télécommunications désormais en grande partie déréglementé, l'industrie est encore dans une phase de réorganisation. Lorsque les nouveaux opérateurs de télécommunications publiques (OTP) arriveront sur le marché, on peut supposer que ces derniers auront un intérêt vital à collaborer avec les institutions de recherche.

Les quatre acteurs aujourd'hui en mesure d'assumer le rôle de chefs de file et dont la collaboration est nécessaire sont par conséquent l'ENPG, la CE, TERENA et DANTE. Si chacun est prêt et résolu à assumer en coopération avec les autres le rôle qui lui revient, aucun conflit ne devrait opposer les parties en présence. On peut résumer leurs rôles respectifs comme suit.

L'ENPG coordonne et réunit les gouvernements nationaux qui, par leurs organismes respectifs de financement, définissent leurs politiques en matière d'établissement de réseaux de recherche. L'ENPG s'associe à la CE pour décider d'un plan d'action concret au niveau européen et notamment des décisions et engagements de contributions financières, ainsi que des procédures de mise en œuvre des actions décidées. Le rôle de TERENA est de faire en sorte que le plan d'action bénéficie de l'expérience et de l'engagement des réseaux nationaux. La société DANTE devrait se voir confier la tâche d'exécuter les travaux suivant les procédures commerciales habituelles et/ou d'éventuelles règles et dispositions de la CE. Ainsi, si les gouvernements nationaux sont responsables, au travers de l'ENPG et de la CE, de l'harmonisation des stratégies et de l'affectation des fonds, l'exécution effective des travaux sera confiée à TERENA, aux NREN, à la société DANTE et à leurs sous-traitants.

IV. EXPÉRIENCES ET COMPARAISONS

La situation en Europe

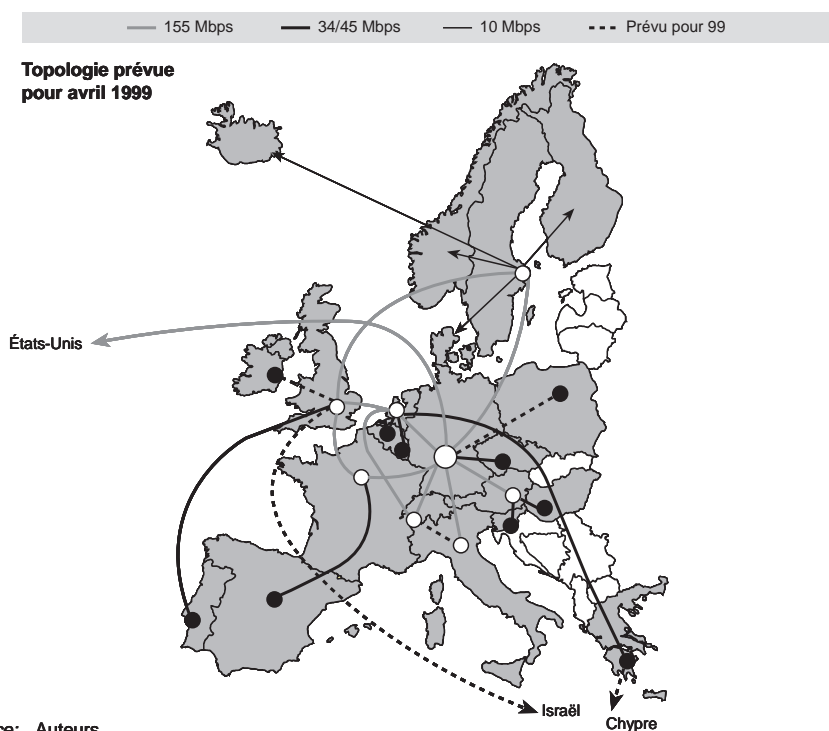
Voilà maintenant un peu plus de dix ans que les Européens s'efforcent de créer des réseaux électroniques numériques afin de permettre aux chercheurs et aux enseignants de mieux communiquer entre eux, non seulement à l'intérieur des frontières nationales mais aussi avec leurs homologues dans d'autres pays d'Europe et du monde. Ces activités ont conduit à l'établissement de réseaux nationaux de recherche dans tous les pays d'Europe et à l'instauration de divers mécanismes pour créer et assurer à l'échelon européen des services inter-réseaux. La manière dont cela a été réalisé a été conditionnée par la situation de monopole des exploitants de télécommunications qui prévalait alors dans les pays d'Europe et par les difficultés qui en ont résulté pour mettre en place et financer la nécessaire infrastructure de base européenne.

Comme indiqué plus haut, IXI a été le premier réseau de base européen. Lancé en 1990, c'est un réseau exclusivement X.25, fonctionnant suivant le protocole ISO à 64 Kbps. Il a été remplacé en 1992 par EuropaNET, lequel s'est transformé en un réseau à 2 Mbps fonctionnant suivant le protocole IP. En 1997, TEN-34 (débit nominal de 34 Mbps) a été lancé avec le soutien du 4^e PCRD de la Commission européenne. Vers la fin de cette année, il sera remplacé par TEN-155, lequel est constitué d'un réseau central fonctionnant à 155 Mbps auquel les NREN peuvent se connecter aux vitesses qu'ils estiment nécessaires (Figure 2)².

Le premier janvier 1998, le marché des télécommunications a été entièrement libéralisé dans la plupart des pays d'Europe. On espère par conséquent que la concurrence aidera à fournir l'infrastructure et les services indisponibles jusqu'ici et réduira aussi les coûts. De telles attentes sont motivées par l'adoption d'un mode de pensée axé sur le marché. On peut cependant prolonger ce raisonnement et parvenir à la conclusion qu'à l'avenir, la fourniture et la maintenance de l'infrastructure pourraient être entièrement confiées aux nouveaux exploitants, étant donné que le marché fournira aux clients ce qu'ils attendent et que ses règles leur garantiront des prix concurrentiels et des prestations intéressantes. L'expérience actuelle montre toutefois que cette situation, si elle se concrétise un jour, ne devrait pas s'instaurer avant un certain temps. Le secteur des télécommunications procédant tout juste à sa propre réorganisation, les considérations commerciales priment sur les besoins des communautés de la recherche et de l'enseignement.

Compte tenu de la nature du mandat des programmes-cadres, la recherche a pendant longtemps été considérée comme la seule raison légitime de soutenir les activités d'établissement de réseaux. Ce n'est que très récemment que l'on a reconnu que la formation et l'apprentissage – et en fait tous les domaines de

Figure 2. Ten-155



l'enseignement – sont essentiels au bien-être futur des citoyens européens. Cette constatation a contribué à changer la manière de voir des dirigeants, de sorte que l'établissement d'un réseau à des fins pédagogiques n'est plus considérée comme une activité « illégale » ou tout au plus tolérée sur un « réseau de recherche ».

La situation aux États-Unis

Le Gouvernement des États-Unis s'est clairement engagé en faveur de l'établissement de réseaux dès l'origine³, lorsque le Département de la défense, par la DARPA, a lancé le développement de l'ARPANET, lequel est resté en exploitation de 1969 à 1990. En 1986 est né NSFnet, réseau assurant un débit de 56 Kbps. Deux ans plus tard, celui-ci a été porté à 1.54 Mbps. En 1992, il a été relevé de nouveau et porté à 44 Mbps.

Le même raisonnement axé sur le marché a été appliqué à l'Internet pour les universitaires aux États-Unis, ce qui a entraîné l'arrêt du financement du NSFnet en

1995. Il faut admettre que, pendant de nombreuses années depuis 1987, les services offerts sur le réseau NSFnet pour la recherche et l'enseignement sont restés inégalés de par le monde, même si les utilisateurs ont toutefois remarqué que la privatisation de NSFnet et le fréquent engorgement du réseau commercial qui lui a succédé ont privé de nombreux universitaires des capacités de réseau indispensables à la poursuite de recherches d'un niveau mondial. Le fait que ce résultat involontaire ait eu un considérable impact négatif sur la communauté des chercheurs universitaires aux États-Unis a été reconnu.

Pour remédier à cette situation, les grandes universités ont réagi par le lancement de l'initiative Internet2⁴. Les principaux objectifs de ce projet consistent à permettre le développement d'applications de nouvelle génération, à restaurer des moyens de réseau de pointe pour la recherche et l'enseignement aux États-Unis et à favoriser le transfert des résultats obtenus sur le réseau Internet mondial. A cet effet, les membres de l'Internet2 ont constitué une société de capitaux. Dénommée *University Corporation for Advanced Internet Development* (UCAID), sa mission est de donner une impulsion et une orientation au développement de réseaux avancés au sein de la communauté universitaire. Actuellement, les universités à l'origine de l'Internet2 utilisent la dorsale vBNS, service sur réseau fédérateur à très grande vitesse, qui est interconnecté avec d'autres réseaux à financement fédéral. Lancé en 1995 à 155 Mbps dans le cadre d'un accord de collaboration sur cinq ans entre NSF et MCI, il fonctionne actuellement à 622 Mbps.

En 1998, l'UCAID a lancé le projet Abilene en association avec, entre autres, Qwest, Cisco et Nortel⁵. Opérationnel en janvier 1999, ce projet a fourni ses premiers services à haut débit en février 1999. Dans sa première phase, il offrira aux membres de l'UCAID des débits de 2.4 Gbps entre noeuds centraux, un chiffre qui devrait ensuite être porté à 9.6 Gbps. Il sera à n'en pas douter le réseau le plus performant au monde. Ce projet, ainsi que le programme de développement d'applications avancées de l'initiative Internet2, devraient avoir un extraordinaire impact sur l'établissement de réseaux.

Comparaison et enseignements à tirer

Comme indiqué plus haut, les pays d'Europe ont constitué des organisations nationales chargées de mettre en place et/ou d'exploiter leurs réseaux de recherche et d'enseignement respectifs. Ces réseaux étaient en fait de véritables «réseaux universitaires» qui ont dès le début fourni des services Internet aux membres de la faculté qui les utilisaient. Avec le temps, ces réseaux ont progressivement accueilli des professeurs, des chercheurs et des étudiants. Même si les politiques indiquant ceux qui sont autorisés à utiliser le réseau national varient d'un pays à l'autre en Europe (cela dépend en règle générale de ceux qui financent effectivement le réseau), on peut dire, tout à fait dans le droit fil de la

thèse défendue plus haut à propos des réseaux de connaissances, que le public des utilisateurs desservis est large.

Aux États-Unis, la situation diffère de manière vraiment frappante de celle de l'Europe. Aucune organisation nationale comparable n'est responsable de l'établissement des réseaux ; c'est peut-être ce qui fait que toutes les institutions de l'enseignement supérieur ne participent pas aux activités mentionnées plus haut. Particulièrement depuis la décision d'arrêter le financement du réseau NSFnet, la plupart des institutions ont dû faire appel à des services du commerce. La société UCAID a été mise en place pour servir en premier lieu (et exclusivement) ses 132 membres actuels, soit moins de 5 % de l'ensemble des institutions de l'enseignement supérieur aux États-Unis.

Compte tenu des efforts accomplis pour créer une Société Européenne de l'Information, il est fortement recommandé que la fourniture et l'utilisation d'un réseau d'interconnexion européen à haut débit ne soit pas réservées à une élite. C'est pourquoi la politique européenne devrait s'efforcer d'englober un public d'utilisateurs aussi large et diversifié que possible.

Le développement commercial de l'Internet a clairement montré que les prestataires commerciaux suivent un modèle qui ne prend pas (encore) en compte la nature particulière de l'utilisation de l'infrastructure propre à l'Internet et qui se base entièrement sur des principes industriels et commerciaux courants, autrement dit la génération de profits pour les actionnaires. La nécessité, pour la communauté universitaire, de disposer d'applications d'avant-garde lui permettant de répondre aux nouvelles exigences de la recherche, de l'enseignement et de l'apprentissage, figure bien loin dans la liste des priorités des exploitants de télécommunications publiques. Si l'on tire les enseignements de l'expérience de l'abandon du soutien du NSFnet aux États-Unis, il est clair qu'il faudrait éviter une évolution similaire en Europe. D'un autre côté, les exploitants de télécommunications européens devraient être encouragés à unir leurs forces avec les acteurs mentionnés plus haut afin d'accélérer le développement du projet Abilene en Europe.

Il a été dit à maintes reprises qu'un réseau d'interconnexion européen à haut débit était un préalable au développement d'applications nouvelles et innovantes. Nombre de chercheurs ont stigmatisé cette situation digne de la poule et de l'œuf : comme il n'existe pas d'applications « décisives », on ne voit pas la nécessité de réseaux hautes performances ; et comme il n'existe pas de réseaux hautes performances, les nouvelles technologies qui exigeraient de tels réseaux ne sont tout simplement pas développées.

Comme le montre l'exemple des États-Unis, un réseau d'interconnexion à haut débit est essentiel en Europe, si cette dernière souhaite jouer un rôle dans le développement de l'Internet de demain. L'absence d'une telle infrastructure en Europe serait désastreuse, les NREN tentant chacun de conclure leur propre accord

avec l'UCAID pour obtenir l'accès à Abilene et participer au développement global en qualité de « chercheurs invités des États-Unis » afin de ne pas rester à la traîne. En fait, plusieurs NREN sont déjà en négociations avec l'UCAID. On ne peut encore évaluer l'effet de cette situation sur la recherche en Europe et sur l'évolution de la coopération européenne au sein de la Société de l'Information. Il y a toutefois un risque que l'idée d'Espace européen de la recherche soit sévèrement freinée et ne prenne jamais corps.

Le tableau 1 offre un récapitulatif des différentes approches suivies en Europe et aux États-Unis.

Tableau 1. Comparaison entre les approches des États-Unis et de l'Europe pour l'établissement de réseau

Mise en réseau : approche européenne	Mise en réseau : approche des États-Unis
Infrastructure : IXI (X.25, protocole ISO, 64 Kbps) en 1990 EuropaNET (64 Kbps) lancé en 1992 a évolué en réseau IP à 2 Mbps TEN-34, débit nominal 34 Mbps (E3) en 1997 ; une seule connexion à 34 Mbps. TEN-155 central à 155 Mbps en 1998/99	Infrastructure : ARPANET commandé en 1969 par le Département américain de la Défense NSFNet (56 Kbps) créé en 1986 NSFNet à 1.54 Mbps (T1) en 1988 NSFNet à 44 Mbps (T3) en 1991 vBNS à 155 Mbps (OC3) en 1995 actuellement à 622Mbps (OC12) Abilene à 2.4Gbps (OC48) en 1999 évolution prévue vers 9.2Gbps (OC192)
Financement : – les fonds proviennent essentiellement des gouvernements nationaux – contribution de la CE	Financement : – financé à l'origine par le gouvernement – assuré maintenant par les utilisateurs et les sponsors
Objet : recherche (et enseignement)	Objet : recherche ET enseignement

De cette comparaison, on peut facilement conclure que l'Europe a commencé – assez tardivement – à concurrencer les États-Unis pour tenter de rattraper son retard au niveau de la bande passante. Ce constat soulève une question souvent posée : pourquoi la communauté de l'enseignement et de la recherche a-t-elle besoin d'autant de bande passante ? Interrogation qui est liée à une seconde question : existe-t-il une application « décisive » exigeant autant de bande passante ?

Ces questions appellent deux réponses. Premièrement, l'acceptation des services sur réseau a été si forte que la capacité des réseaux devrait être multipliée par deux tous les huit ou neuf mois en moyenne pour satisfaire à la demande. Cela entraîne un volume de trafic cumulé considérable sur les réseaux d'interconnexion. Deuxièmement, même si peu de gens l'ont remarqué, les applications sur le Web,

qui comprennent de plus en plus de graphiques et d'éléments animés, ainsi que d'autres applications novatrices, ont tout simplement besoin d'une bande passante plus élevée pour fonctionner correctement. Dans le tableau 2, nous avons essayé de calculer la bande passante nécessaire pour répondre aux attentes de la Société de l'Information.

Tableau 2. **Pourquoi une bande passante aussi importante est-elle nécessaire ?**

• Logiciel éducatif (« apprentissage » interactif plutôt qu'une simple « information »)	Gbps
• Modélisation scientifique	Centaines de Mbps à Gbps
• Entreposage intelligent de données	Gbps et temps de réaction rapide
• Télétravail/enseignement et formation à distance	6 – 10 Mbps
• Vidéo-conférence sur micro-ordinateur	6 – 10 Mbps
• Publications	Centaines de Mbps
• Voix sur protocole Internet	Pas encore vraiment connu
• Commerce électronique	Pas encore vraiment connu

V. LE RÉSEAU D'INTERCONNEXION EUROPÉEN A HAUT DÉBIT

L'idée d'un réseau d'interconnexion européen à haut débit a été défendue par grand nombre de personnes et d'équipes de travail, lors de nombreuses conférences et réunions (Williams, 1997; Axmann et Payr, 1997). Les questions de topologie et les problèmes techniques pouvant sans problème être confiées aux experts, le présent article met l'accent sur certaines réflexions générales. Elles devraient permettre à la fois de circonscrire les requêtes des utilisateurs et d'aider les responsables politiques à mieux comprendre les besoins et les exigences de l'établissement de réseaux et les avantages que l'on peut en retirer.

Les principes

Ce qui suit est une première esquisse du concept de réseau d'interconnexion européen à haut débit. Elle doit être affinée et faire l'objet d'un examen plus approfondi. Cependant, certains principes de base sont d'ores et déjà clairs :

- *Qualité et pas uniquement quantité* : les réseaux européens pour la recherche et de l'enseignement doivent disposer d'une importante bande passante et ce non plus seulement pour absorber le volume toujours croissant des services traditionnels. Pour les nouvelles demandes et applications, particulièrement dans le domaine de la communication en temps réel et multimédia,

comme la transmission de la voix ou de la vidéo, la qualité de service offerte par la nouvelle génération de réseaux de recherche et d'enseignement doit être améliorée. Le principe de l'obligation de moyen appliqué actuellement à l'Internet sera conservé, mais le réseau devra en plus satisfaire les exigences de bande passante à la demande, de temps de réaction rapide, de diffusion simultanée, etc.

- *Production et expérimentation* : bien qu'un réseau d'interconnexion européen à haut débit soit essentiellement requis pour la production, il constituerait un banc d'essai idéal pour de nouveaux protocoles et de nouvelles applications. Il ne semble pas irréaliste de supposer que de nouvelles technologies de réseau, comme la photonique, de même que les équipements appropriés, parviendront à maturité dans les toutes prochaines années ; quel meilleur banc d'essai qu'un réseau de production ! Le banc d'essai doit séduire tout à la fois les chercheurs des universités et de l'industrie, les exploitants de télécommunications publiques et les industriels des télécommunications et leur permettre de pratiquer des démonstrations et des évaluations des nouvelles technologies dans un environnement pré-concurrentiel « réel ». Par le passé, l'expérience a montré que si l'on prend les précautions appropriées, il est possible de combiner production et expérimentation. Les environnements de la recherche et de l'enseignement, friands de nouvelles applications et de nouveaux services, constituent un banc d'essai idéal pour les applications pilotes, car ces environnements sont prêts à tester et à utiliser des solutions « pas tout à fait au point ».
- *De nouvelles applications ne sont possibles que si le réseau est disponible* : le comité consultatif « Future of the Internet » déclarait il y a quelque temps que « ... de nombreuses applications intéressantes et précieuses seront développées en Europe, à la fois pour l'Internet de base et l'Internet hautes performances, dès que la bande passante augmentera et que la couverture transeuropéenne s'améliorera » (Williams, 1997). En concentrant les forces européennes sur l'amélioration de l'infrastructure de réseau à haut débit dans la recherche et l'enseignement, on pourrait sans doute créer un contexte propice à la création de débouchés et d'emplois grâce aux applications développées et à la migration vers l'Internet de base.
- *Assurer la cohésion, éviter tout écart technologique* : un réseau d'interconnexion européen à haut débit résultant d'une initiative commune en Europe doit absolument prendre en compte le processus de rapide mise à niveau que traversent certains membres actuels et futurs de l'UE. L'architecture de réseau doit par conséquent être suffisamment flexible pour accepter les différentes bandes passantes nécessaires aux liaisons entre les réseaux nationaux et l'infrastructure d'interconnexion et permettre de futures extensions.

- *Soutien européen pour une valeur ajoutée en Europe* : le bon fonctionnement d'un réseau d'interconnexion européen hautes performances dépend tout autant de l'efficacité de l'organisation générale de la gestion que de la participation active des NREN. Pour garantir la fourniture de services aux chercheurs et aux scientifiques européens du secteur universitaire ou industriel et pour le soutien des activités communautaires, il est légitime d'utiliser les fonds de la Communauté pour assumer une partie des coûts d'exploitation généraux du réseau d'interconnexion expressément liés à sa valeur ajoutée européenne.

Les avantages

Outre ses effets généraux à long terme, une infrastructure de réseau à hautes performances pour la recherche et l'enseignement en Europe peut apporter quelques avantages très concrets :

- *Coût des connexions intercontinentales* : un réseau commun pour la recherche et l'enseignement en Europe renforcerait la position de cette dernière dans les négociations sur le partage des coûts intercontinentaux (Karounos, 1997).
- *Avantages de la synergie* : si les pays d'Europe décidaient de construire un réseau d'interconnexion européen à hautes performances véritablement étendu, susceptible de garantir l'avenir de la recherche, du développement et de l'industrie, plutôt que d'ajouter au coup par coup des éléments isolés, il en résulterait certainement des coûts considérables à court terme. Toutefois, on peut sans se tromper supposer que le retour sur investissement serait proportionnel aux effets de synergie résultant d'un développement coordonné.
- *Partenariats université-industrie* : un réseau d'interconnexion européen à hautes performances devrait susciter l'intérêt des acteurs industriels qui comptent sur le marché des télécommunications en Europe. Les partenariats entre l'université et l'industrie profitent non seulement à la recherche sur l'établissement des réseaux, mais aussi à l'industrie, particulièrement par le fait qu'ils ouvrent la voie à la mise en place commerciale à grande échelle d'une infrastructure et d'une technologie de communications de type avancé et qu'ils contribuent à faire pleinement entrer l'Europe dans la Société de l'Information.

VI. INCIDENCES POUR L'EUROPE

Les universités et les établissements de l'enseignement supérieur (en fait, les établissements de tous types), les institutions de recherche et les organismes culturels

n'utilisent plus les réseaux seulement par commodité, ils sont désormais nombreux à se reposer sur eux comme un instrument vital de communication. Les réseaux donnent les moyens d'échanger des idées d'une manière rapide et efficace, de lancer et de poursuivre des travaux en collaboration, de diffuser des résultats de recherche, d'offrir un matériel d'enseignement et d'apprentissage à un public plus large, de solliciter la révision de travaux en cours, de créer des œuvres d'art avec de nouveaux « matériels » et de nouvelles méthodes de présentation et d'accomplir bien d'autres fonctions qui auraient été lourdes à gérer suivant la méthode « classique ».

La recherche en Europe bénéficiera d'un réseau d'interconnexion à haut débit à plusieurs égards : les applications sur réseaux de type avancé, comme la télécollaboration, les laboratoires virtuels ou les vidéo-conférences, permettront de suivre le rythme sans cesse plus rapide de production et de mise en application des résultats de la recherche. La mondialisation de la recherche impose une coopération et un partage des ressources au niveau international. La coopération s'est intensifiée, elle repose sur des intérêts communs et n'est plus limitée par la distance. Des équipements de recherche toujours plus perfectionnés et onéreux peuvent être partagés et sont par conséquent accessibles à tous les scientifiques européens.

Dans l'enseignement, l'établissement de réseaux à hautes performances aura un rôle important à jouer dans la réorganisation et la redéfinition du système éducatif imposés par les enjeux de la Société de l'Information : possibilité d'offrir un enseignement interactif mobilisant d'importantes ressources, axé sur la réalisation de projets et la solution de problèmes, sans contrainte de temps ou d'espace, tout en conservant la qualité et la force de l'interaction entre élèves et professeurs. Les cours et les programmes universitaires étant partagés entre différents sites et différents pays, les étudiants seront assurés de bénéficier du meilleur enseignement possible, quel que soit l'endroit où ils se trouvent.

Le rôle de l'établissement de réseaux dans la recherche et l'enseignement apparaît encore plus clairement lorsque l'on examine les objectifs et les priorités d'ensemble définies pour la recherche et le développement technologique en Europe (RDT) dans le 5^e PCRD (Commission européenne, 1996) et lorsque l'on tente d'imaginer si l'un quelconque de ces objectifs pourrait être atteint sans des réseaux de recherche efficaces et étendus :

- Confirmer et renforcer le rôle de la recherche européenne sur le plan international en encourageant la coopération des meilleures équipes de chercheurs en Europe.
- Créer un environnement propice à la recherche et instaurer dans ce domaine un contexte ouvert à de nouvelles idées et facilitant le lien, d'une part entre

recherche fondamentale et recherche appliquée et d'autre part, entre recherche universitaire et recherche industrielle.

- Création d'emplois : diverses études montrent que l'introduction rapide et à grande échelle de communications avancées dans toute l'Union européenne conduira directement ou indirectement à la création de six millions de nouveaux emplois d'ici l'an 2010.
- Mettre la recherche plus en adéquation avec le marché réel mais aussi, améliorer et accélérer l'exploitation des résultats de recherche, ce qui est impossible sans une coopération efficace entre les partenaires des communautés de la recherche, de l'industrie et des utilisateurs.
- Donner une dimension globale à la recherche et renforcer les partenariats à travers l'Europe de manière à répartir les risques et mettre en commun les ressources.

Les actions communautaires sortant du cadre des programmes RDT seront toujours plus tributaires et auront toujours plus besoin de réseaux de pointe. Parmi les actions et programmes actuels du 4^e PCRD dans le domaine de l'enseignement et de la formation, on peut par exemple citer :

- SOCRATES, dans le domaine de l'école et de l'enseignement supérieur.
- LEONARDO, dans le domaine de la formation professionnelle.
- Le programme FMC (programme de RDT dans le domaine de la formation et de la mobilité des chercheurs), qui est axé plus particulièrement sur l'échange de chercheurs et l'accès aux installations de recherche.

La coopération à l'échelle européenne entre ces institutions d'enseignement et de recherche occupe une place importante dans tous ces programmes. Et si les « échanges physiques » (programmes de bourses pour étudiants et apprentis chercheurs, conférences et ateliers internationaux, accès aux grandes installations de recherche pour tous les chercheurs des pays d'Europe) resteront certainement un instrument important pour favoriser l'intégration européenne et le potentiel scientifique et d'innovation en Europe, « les échanges virtuels » sous toutes leurs formes (salles de classe ou laboratoires virtuels, télé-conférences et télé-enseignement) auront bientôt à jouer un rôle tout aussi crucial et donneront à la fois une nouvelle dimension et une nouvelle qualité à la collaboration dans le domaine de la recherche et de l'enseignement en Europe.

VII. CONCLUSION

L'horizon dans lequel se situent les idées, les plans et les actions mentionnés précédemment n'est pas totalement indéterminé. On connaît déjà la date butoir à

laquelle les réseaux européens de recherche à haut débit devront être prêts à l'avenir. TEN-34 sera bientôt remplacé par TEN-155, grâce à des fonds de soutien du projet QUANTUM du quatrième Programme-cadre. TEN-155 aura également une durée de vie limitée. Toutefois, il garantira que l'aide à l'établissement de réseaux européens sera maintenue jusqu'à ce que les mesures du 5^e PCRD prennent effet.

Une action immédiate est par conséquent indispensable et les acteurs précédemment mentionnés se doivent de contribuer à l'élaboration et à la mise en œuvre d'un réseau d'interconnexion européen à haut débit. Aucun ne peut à lui seul donner corps à l'avenir des réseaux européens pour la recherche et l'enseignement. Beaucoup d'efforts et de négociations seront nécessaires et la coordination de tous les acteurs en présence est une nécessité absolue.

NOTES

6. Membres de l'ENPG (au mois de mai 1998) : Autriche, Belgique, Espagne, Finlande, Hongrie, Luxembourg, Norvège, Pays-Bas, Pologne, Portugal, Royaume-Uni, Slovaquie, Slovénie, Suède. La France, l'Irlande et l'Italie ont participé à des réunions en qualité d'observateurs. TERENA bénéficie du statut d'observateur permanent.
7. DANTE: <http://www.dante.org.uk/>.
8. Internet Society: <http://www.isoc.org/internet/history/>.
9. Projet Internet2: <http://www.internet2.edu/>.
10. University Corporation for Advanced Internet Development: <http://www.ucaid.edu/abilene/>.

BIBLIOGRAPHIE

- AXMANN, H.-P., P. LIZCANO et M. READ (1997),
« Report on Funding Research Infrastructure » remis au comité Programme Applications
Télématiques, le 18 juillet 1997.
- AXMANN, H.-P. et S. PAYR (1997),
« Towards a European High-performance Education and Research Network (HIPERnet) »,
JENC8, session 412, mai 1997.
- COMMISSION EUROPÉENNE (1996),
« Inventer demain : la recherche européenne au service du citoyen », Communication de
la Commission. COM(96)332 final, Bruxelles, juillet 1996.
- COMMISSION EUROPÉENNE (1997),
« Proposition de décision du Parlement européen et du Conseil relative au cinquième
programme-cadre de la Communauté européenne pour des actions de recherche, de
développement technologique et de démonstration (1998-2002) », COM(97)142, avril 1997.
- KAROUNOS, Theodoros (1997),
« Intercontinental Cost Sharing in the Internet between Europe and the USA », février 1997,
<http://www.dante.net/policy/co-fund.html>.
- PROGRAMME APPLICATIONS TÉLÉMATIQUES (1996),
Report of the Five-year Review Panel, octobre.
- WILLIAMS, David (1997),
« The Future of the Internet – What Role for Europe? », rapport provisoire d'un comité
consultatif, novembre 1997, <http://nicewww.cern.ch/~davidw/public/i2eureport.htm>

LES RÉSEAUX INFORMATIQUES ET L'UNIVERSITÉ VIRTUELLE

Table des matières

I. Introduction	56
II. Estimations de l'utilisation des réseaux informatiques.....	56
III. Mécanismes de collaboration	59
IV. Fréquence de la communication	65
V. Mondialisation de la science	69
VI. Effets de périphéricité	70
VII. Productivité.....	76
VIII. Contextes sociaux.....	78
IX. Implications pour la bibliothéconomie et les sciences de l'information	80
X. Conclusions	82
Notes	84
Bibliographie	85

Cet article a été rédigé par John P. Walsh, du Département de sociologie de l'Université de l'Illinois à Chicago (adresse électronique : JWalsh@uic.edu) et Ann Roselle, de l'Université d'Eastern Washington.

I. INTRODUCTION

Si de nombreuses publications ont examiné dans quelle mesure les réseaux informatiques étaient susceptibles de révolutionner le travail scientifique, seul un nombre assez limité d'études empiriques ont tenté de mesurer l'impact de cette technologie nouvelle. Le présent article présente une synthèse des publications existantes consacrées aux effets des réseaux informatiques sur le travail scientifique et examine l'incidence de ces réseaux scientifiques dans le domaine de la bibliothéconomie et des sciences de l'information. Les réseaux informatiques, qui créent une « université virtuelle » dans ce contexte, désignent la communication par ordinateur par l'intermédiaire d'un réseau (en l'occurrence, l'Internet). Le discours électronique (courrier électronique, services de messagerie électronique, services de conférence virtuelle, journaux électroniques, groupes de discussion en ligne) ainsi que l'accès aux bases de données informatisées et le transfert de fichiers électroniques sont inclus dans cette utilisation au sens large de l'expression « réseaux informatiques ».

L'article est structuré comme suit : la section II rend compte de l'utilisation de la communication par ordinateur dans certaines disciplines, tandis que la section III fournit certains détails concernant la collaboration à distance, la collaboration internationale et les modifications de l'organisation des groupes de travail qu'a entraînées la communication par ordinateur. La section IV explique les niveaux d'intensité de communication auxquels permet d'accéder la communication par ordinateur. La section V introduit l'idée de la coopération géographique par la mise en commun des bases de données, et la section VI se penche sur les changements dans les distinctions de statut et l'accès aux informations suite à la communication par ordinateur. La section VII présente les effets de la communication par ordinateur sur la productivité scientifique, la section VIII décrit les facteurs sociaux qui peuvent avoir une incidence sur le recours à la communication par ordinateur, et enfin la section IX fait la synthèse de l'utilisation de la communication par ordinateur par les professionnels des bibliothèques et de l'information, ainsi que de l'impact qu'a cette communication sur ces professionnels.

II. ESTIMATIONS DE L'UTILISATION DES RÉSEAUX INFORMATIQUES

Les recherches antérieures concernant la communication par ordinateur suggèrent que l'adoption de cette technique peut avoir toute une série d'effets

sur la structure sociale des organisations de travail (Sproull et Kiesler, 1991 ; Rice, 1994). Avant d'examiner ces effets en détail, il convient de souligner qu'ils sont conditionnés par l'utilisation, qui varie largement selon le domaine et la région.

Les analyses comparatives systématiques des taux d'utilisation de la communication par ordinateur dans les différents domaines et les différentes régions sont peu nombreuses, mais des estimations ont été sélectionnées dans diverses études afin de donner une idée de la manière dont l'utilisation peut varier. Par exemple, Walsh et Bayma (1996a), en se fondant sur des listes d'abonnés, ont conclu que les mathématiciens (dont 34 % signalent qu'ils possèdent une adresse de courrier électronique) et les physiciens (24 %) utilisaient largement le courrier électronique, et cela depuis 1991, alors que de nombreux biologistes expérimentaux (9 %) n'avaient pas encore adopté cette technique (tableau 1). De même, Mailman (1996) a constaté qu'en Amérique du Nord, le taux d'utilisation du courrier électronique était encore faible chez les ornithologues (15 %) en 1993.

Il n'empêche qu'en 1993, cette technique commençait à se répandre rapidement. Bishop (1994) signale, en se basant sur une enquête auprès des ingénieurs en aérospatiale dans l'industrie, dans le secteur public et dans les universités, qu'en 1993, 74 % d'entre eux utilisaient les réseaux, bien que 50 % seulement eussent accès aux réseaux extérieurs de recherche tels que l'Internet. Chez les universitaires (autres que les travailleurs de l'industrie ou des services publics) et chez les scientifiques (hors ingénieurs et gestionnaires), l'utilisation était quasi universelle (plus de 90 %). L'enquête de Cohen (1995) auprès des chimistes, des sociologues, des politicologues et des philosophes dans des collèges de jésuites et des universités aux États-Unis a relevé d'importants écarts d'utilisation selon les domaines, l'éventail allant de 82 % pour les chimistes à 55 % pour les philosophes (tableau 1). Des estimations récentes dues à Walsh *et al.* (à paraître) suggèrent que l'utilisation est maintenant largement répandue, mais qu'il subsiste d'importantes différences selon les disciplines, les utilisations estimées allant de 78 % pour les mathématiques à 61 % pour la sociologie (tableau 1). Il existe même une variation entre les sous-domaines de certaines disciplines (Walsh et Bayma, 1996a ; Hurd et Weller, 1997). Par exemple, Hurd et Weller (1997) comparent les chimistes de différents départements d'une université de recherche et concluent que l'utilisation du courrier électronique varie de 84 % pour ceux qui relèvent du département de la chimie (faculté des arts et des sciences) à 70 % pour ceux qui appartiennent au département de la chimie biologique (faculté de médecine)

Par conséquent, les estimations récentes semblent indiquer que la communication par ordinateur est devenue une pratique tout à fait courante chez les scientifiques, ce qui met en évidence l'importance de la question de son incidence sur le travail scientifique. Toutefois, même ces estimations récentes montrent qu'il existe des

Tableau I. **Utilisation des réseaux informatiques dans différentes disciplines, États-Unis**

Domaine	Année	Utilisation du courrier électronique (en pourcentage)	Année	Utilisation du courrier électronique (en pourcentage)	Source
Biologie expérimentale	1991	9	1997	70	1991 : Walsh et Bayma (1996a) 1997 : Walsh <i>et al.</i> (à paraître)
Mathématiques	1991	34	1997	78	
Physique	1991	24	1997	66	
Sociologie ¹			1997	61	
Tous domaines	1992	39			Schauder (1994)
Ornithologie	1993	15			Hailman (1996)
Génie aérospatial	1993	74			Bishop (1994)
Sciences biologiques	1993-94	89			Abels <i>et al.</i> (1996)
Ingénierie	1993-94	77			
Sciences médicales	1993-94	14			
Mathématiques/ informatique	1993-94	93			
Sciences physiques et géologiques	1993-94	80			
Sciences sociales	1993-94	55			
Chimie	1994	82			Cohen (1995)
Sociologie	1994	75			
Sciences politiques	1994	67			
Philosophie	1994	55			

Note : En raison de différences de stratégies de mesure et d'échantillonnage, les chiffres ne sont pas strictement comparables d'une étude à l'autre.

¹. Australie, États-Unis, Royaume-Uni.

Source : Walsh et Bayma (1996a) ; Schauder (1994) ; Hailman (1996) ; Bishop (1994) ; Abels *et al.* (1996) ; Cohen (1995) ; Walsh *et al.* (à paraître).

écarts importants d'un domaine à l'autre, ce qui semble indiquer que des facteurs sociaux ou techniques peuvent avoir une influence (Walsh et Bayma, 1996a). De plus, Hiltz et Johnson (1989) concluent que l'utilisation (et la satisfaction) ne sont pas nécessairement en relation avec des résultats bénéfiques. Nous examinerons, dans les sections qui suivent, les éléments qui indiquent que la communication par ordinateur a actuellement des effets mesurables sur le travail scientifique.

III. MÉCANISMES DE COLLABORATION

Collaboration à distance

Une des conclusions de la recherche concernant la communication par ordinateur est que celle-ci peut entraîner un accroissement de la taille du réseau social de celui qui y a recours. Dans une étude portant sur les océanographes, par exemple, Hesse *et al.* (1993) ont conclu que les grands utilisateurs du courrier électronique sont aussi ceux qui signalent les réseaux professionnels les plus étendus. L'étude concernant l'utilisation de la communication par ordinateur en biologie, chimie, mathématiques et physique, réalisée par Walsh et Bayma (1996b), a abouti à la conclusion que les collaborations se sont élargies et que les effets semblent associés à l'utilisation de cette technique. Un spécialiste de physique expérimentale des particules explique les choses comme suit :

« Les expériences sont en train de s'élargir. Nous tentons à présent de comprendre la nature de la matière, l'intérieur d'un proton. Il nous faut plus d'énergie, plus d'argent, du matériel de plus grande dimension. Les expériences sont aujourd'hui moins nombreuses parce que les ressources financières sont limitées. Par conséquent, chaque expérience acquiert une plus grande dimension. La taille des expériences croît. Ce n'est pas une conséquence des réseaux. Mais les réseaux sont un outil qui nous a facilité la vie. Cela n'a pas eu d'effet sur la taille de l'appareillage. Mais actuellement, une collaboration effective peut s'instaurer entre un plus grand nombre de personnes. » (Walsh et Bayma, 1996b).

Un des changements les plus significatifs dans l'organisation de travail de la science a été le renforcement de la collaboration à distance et, en particulier, de la collaboration internationale. Il a été démontré que les réseaux informatiques réduisaient la nécessité, pour ceux qui collaborent, de se trouver dans un même endroit (Bullen et Bennett, 1991 ; Finholt et Sproull, 1990). Carley et Wendt (1991) affirment que l'on voit apparaître une forme nouvelle de travail scientifique, ce qu'ils appellent un « groupe de recherche étendu ». Il s'agit de groupes de très grande taille, unifiés, cohérents et fortement coopératifs qui sont géographiquement dispersés mais dont la coordination se passe comme s'ils se trouvaient dans un même lieu et étaient dirigés par une seule personne. Ces groupes dépendent fortement de la communication par ordinateur pour coordonner leurs activités de travail. Carley et Wendt (1991) citent, comme exemple de cette manière nouvelle de travailler, le groupe Soar dans le domaine de la recherche en intelligence artificielle. D'autres ont utilisé le terme « collaboratoire » pour désigner un système de travail similaire (Lederberg et Uncapher, 1989 ; Finholt et Olson, 1996). Les collaboratoires offrent, indépendamment de la situation géographique, l'accès au matériel, aux collègues et aux bases de données qui fait traditionnellement partie de

l'organisation de la science en laboratoires. Finholt et Olson (1996) citent des laboratoires dans les sciences de l'atmosphère et les sciences spatiales, la biologie, la chimie, la médecine et la physique. Cette forme nouvelle de travail scientifique est illustrée par une des premières collaborations grâce à la communication par ordinateur, à savoir le projet de LIST commun. Orlikowski et Yates (1994) ont décrit le travail en groupe qui a donné naissance au LIST commun comme un réseau d'informaticiens qui communiquaient presque exclusivement par courrier électronique. Un exemple plus récent pris dans le domaine de l'informatique est la mise au point du système d'exploitation Linux, qui a évolué grâce à la collaboration en ligne de milliers de programmeurs du monde entier (Harmon, 1999).

Des travaux empiriques montrent qu'il existe une relation entre les réseaux et les mécanismes de collaboration. Sanderson (1996) décrit une collaboration dans le domaine de la physique atmosphérique qui comprenait des scientifiques de cinq sites au Canada, deux aux États-Unis et deux dans un autre pays. Ce projet de recherche a adopté le courrier électronique comme forme de communication de base. Tous les membres du groupe de recherche possèdent une adresse Internet et la plupart des membres du groupe ont signalé qu'ils envoyaient plusieurs messages électroniques par semaine aux autres membres du groupe de collaborateurs. Le courrier électronique a été préféré au téléphone pour les raisons suivantes : il peut être difficile de joindre des scientifiques qui se déplacent mais on peut les joindre à leur adresse virtuelle ; les messages écrits permettent un délai de formulation des réponses avant de réagir ; et, pour les collègues non anglophones, la communication orale est plus compliquée que la communication écrite. Sudweeks et Rafaeli (1996) relatent l'historique du « Projet H », une collaboration en ligne comprenant quelque 100 chercheurs dans 15 pays et dont le but était d'étudier le comportement face à la communication par ordinateur. Le groupe a planifié le projet, a prélevé l'échantillon (de messages électroniques sur les listes de BITNET, Usenet et Compuserve), et a créé la base de données commune sans jamais se rencontrer face à face. Le groupe a aussi créé une revue scientifique en ligne (*Journal of Computer-mediated Communication*), disponible sur des pages Web en Israël et aux États-Unis. Cohen (1995) a conclu qu'il existait une corrélation importante entre l'utilisation de la communication par ordinateur et la rédaction en collaboration. Si on ignore encore en partie si la communication par ordinateur entraîne une collaboration accrue, les résultats semblent indiquer que cette technique facilite à coup sûr ce type de collaborations (Walsh et Bayma, 1996b). Par exemple, l'article qui a annoncé la découverte du « quark t » citait 398 auteurs provenant de 34 institutions dans cinq pays (Walsh et Bayma, 1996b). La citation ci-dessus d'un spécialiste de physique des particules souligne l'importance de la communication par ordinateur en tant que moyen contribuant à faciliter les collaborations de très grande taille dans son domaine. De même, Bishop (1994) note que les ingénieurs en aérospa-

tiale affirment qu'un des effets des réseaux d'ordinateurs a été d'accroître la faisabilité et la taille de leurs efforts de collaboration. De plus, Walsh *et al.* (à paraître) signalent que 65 % d'un échantillon de scientifiques dans quatre domaines indiquent que l'utilisation du courrier électronique a renforcé leur collaboration (tableau 2). Ils signalent en outre qu'il existe une association positive entre l'utilisation du courrier électronique et la collaboration à distance.

Tableau 2. **Pourcentage de correspondants signalant un effet positif du courrier électronique sur les résultats scientifiques/professionnels, par domaine**

	Échantillon	Biologie	Math.	Physique	Sociologie	Écarts entre domaines
Contact avec les scientifiques d'autres institutions	85.2 %	86.8 %	87.8 %	77.5 %	88.3 %	**
Renseignements sur les conférences	78.7 %	70.0 %	89.8 %	81.9 %	72.8 %	***
Connaissance des scientifiques d'autres institutions	66.0 %	63.7 %	77.6 %	49.3 %	72.5 %	**
Ampleur de la collaboration	65.4 %	56.0 %	63.3 %	57.7 %	48.5 %	
Productivité de la recherche	58.5 %	58.2 %	69.4 %	64.8 %	41.7 %	***
Connaissance des appels à communications	58.0 %	38.9 %	71.4 %	68.1 %	52.9 %	***
Distraction par rapport à la recherche	50.4 %	40.7 %	44.9 %	59.2 %	57.3 %	**
Renseignements sur les bourses	49.1 %	51.1 %	47.9 %	45.8 %	51.5 %	
Ampleur des domaines de recherche	41.9 %	36.3 %	51.0 %	45.1 %	35.0 %	†
Prestige académique	20.3 %	16.5 %	28.6 %	13.9 %	21.8 %	
Participation des étudiants de troisième cycle à la recherche	15.1 %	17.0 %	6.4 %	18.6 %	18.8 %	
Capacité d'obtenir des bourses	9.2 %	6.7 %	12.5 %	9.7 %	7.8 %	
N =	314	91	49	72	103	

Note : Les calculs sur échantillons ont été pondérés de manière à contrôler les écarts de réponses par domaine.
*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$ † $p \leq 0.10$.

Source : Walsh *et al.* (à paraître).

Collaboration internationale

Bien que, de tout temps, la science ait été une activité internationale, ce qui a changé ces dernières années, c'est la fréquence des collaborations internationales. Le tableau 3 montre l'augmentation du nombre de collaborations internationales depuis 1981, par domaine. Ce tableau révèle que, pour l'ensemble des domaines, le pourcentage de communications publiées dont les auteurs appartiennent à plus

Tableau 3. **Pourcentage de publications résultant d'une collaboration internationale, à l'échelle mondiale et aux États-Unis, par domaine, 1981-1995**

Domaine	1981	1981-1985	1986	1991	1991-1995	1991-1995	1995
	Monde	États-Unis	Monde	Monde	Monde	États-Unis	Monde
Mathématiques	9	11	13	17	19	19	n.d.
Physique	9	14	11	16	19	25	n.d.
Biologie	5	7	7	10	11	13	n.d.
Chimie	5	9	6	9	11	14	n.d.
Tous domaines	6	9	8	11	13	16	15

n.d. = non disponible.

Source : National Science Board (1993), annexe au tableau 5-24 ; National Science Board (1998), annexe au tableau 5-53.

d'un pays s'est considérablement accru. Ces collaborations sont les plus courantes dans les domaines plus susceptibles d'avoir recours à la communication informelle par ordinateur, notamment en mathématiques et en physique (Walsh et Bayma, 1996a).

Les scientifiques interrogés par Walsh et Bayma ont noté que leurs collaborations à distance seraient beaucoup plus difficiles, voire impossibles, sans le recours au courrier électronique, bon marché, rapide et asynchrone. Une des personnes interrogées (travaillant aux États-Unis dans le domaine de la physique de l'état solide) a relevé l'importance des réseaux pour faciliter la collaboration au niveau international :

« Globalement, cela [la constitution de réseaux] a été une bénédiction. C'est beaucoup plus facile pour correspondre et obtenir des données. C'est beaucoup plus efficace. C'est largement pratiqué. Je ne pourrais pas avoir ces collaborations s'il n'y avait pas de réseaux. Ce serait impossible. Nous n'avons eu que deux semaines de délai pour formuler cette proposition avec ce type en Autriche. Sans les réseaux, ce serait virtuellement impossible. » (Walsh et Bayma, 1996b).

Les personnes interrogées étaient largement d'accord pour dire qu'une des principales vertus de cette technique est qu'elle contribue à surmonter les obstacles résultant de la séparation géographique. Cela se remarque le plus dans les collaborations internationales, où le coût, les fuseaux horaires et la langue créent autant de barrières qui peuvent être surmontées par cet outil particulièrement bien adapté à cette tâche qu'est le courrier électronique par Internet. Cet aspect semble particulièrement important pour les chercheurs en Australie et en Nouvelle-Zélande (Ostbye et Welby, 1988 ; Bruce, 1994). Bruce (1994) note qu'un des grands avantages attribués à l'utilisation de l'Internet par son échantillon de chercheurs australiens est la possibilité de rester en contact avec les collaborateurs aux

Tableau 4. Nombre d'hôtes Internet pour 1 000 habitants, 1996

Rang	Pays	Nombre d'hôtes pour 1 000 habitants
1	Finlande	62
2	Islande	42
3	États-Unis	31
4	Norvège	30
5	Australie	24
6	Suède	23
7	Nouvelle-Zélande	23
8	Canada	23
9	Suisse	18
10	Singapour	18

Source : Ibrahim (1997).

États-Unis et en Europe. Les scientifiques en Australie et en Nouvelle-Zélande ont utilisé abondamment l'Internet pour obtenir un accès plus aisé aux communautés de recherche en Amérique du Nord et en Europe. Le tableau 4 présente les 10 premiers pays du point de vue du nombre d'hôtes Internet pour 1 000 habitants ; l'Australie et la Nouvelle-Zélande figurent toutes deux dans cette liste, avec des densités de réseau à peu près identiques à celles du Canada ou de la Suède (Jacobson, 1994).

Le courrier électronique, en tant que moyen de faciliter la collaboration, vient immédiatement après l'interaction face à face et permet aux collègues de continuer à collaborer à distance après avoir déblayé le terrain par des rencontres personnelles (Carley et Wendt, 1991 ; Sanderson, 1996). Les scientifiques ont souligné qu'il était important d'établir une compréhension commune du problème central de recherche par une interaction face à face intensive avant de passer aux collaborations par l'intermédiaire de l'ordinateur. Un physicien théoricien explique les choses comme suit :

« Cela [le courrier électronique] aide. Mais cela ne supprime pas la nécessité de se parler et de travailler ensemble. Une fois que le projet est lancé, une fois que le travail est terminé, on peut faire le reste en ligne. C'est difficile à faire par ordinateur. Il faut être proche de la conclusion ou avoir dépassé ce stade. Il est difficile de faire de la recherche par le biais de l'ordinateur. On peut communiquer les résultats. Mais ça ne remplace pas la discussion, le combat qu'il faut livrer. Mon collaborateur arrive du Brésil en septembre. Nous allons beaucoup travailler pendant qu'il sera ici. Nous rédigerons l'essentiel de la communication ensemble. Plus tard, nous ajouterons des choses [par courrier électronique]. » (Walsh et Bayma, 1996b).

Bien que la science ait toujours été une entreprise relativement internationale, cette mondialisation des domaines tisse aujourd'hui des liens plus étroits entre les scientifiques qui vivent dispersés, créant ainsi une communauté internationale de scientifiques formant un tissu plus serré. Ce n'est pas la communication par ordinateur qui est la cause de ces changements. Elle a toutefois fourni l'infrastructure qui permet à ces changements de se produire. Les possibilités de collaboration internationale ont toujours existé, mais celle-ci était assez rare avant que les réseaux ne deviennent d'usage courant. Le tableau 3 montre que, même aujourd'hui, la collaboration internationale, si elle est plus courante qu'il y a dix ans, reste relativement rare dans les domaines où la communication par ordinateur est moins répandue, notamment en chimie et en biologie expérimentale (Walsh et Bayma, 1996a).

Réorganisation des collaborations

Cette modification des structures de collaboration semble indiquer que celle-ci s'instaure de plus en plus par le biais de concordances sur le fond plutôt que par des liens géographiques ou personnels (Van Alstyne et Brynjolfsson, 1996a). Les recherches antérieures sur la communication par ordinateur (Finholt et Sproull, 1990 ; Feldman, 1987) ont conclu que celle-ci entraîne une réorganisation des groupes de travail qui met l'accent sur les concordances et néglige les coïncidences géographiques. Cette technique permet la constitution de groupes de travail de plus grande taille, plus complexes et dotés de structures plus fluides que ne l'autorise la communication face à face (Rice, 1994, chapitre 8). A l'instar de l'entreprise virtuelle, les réseaux contribuent à créer des équipes de recherche virtuelles qui relient toute une série de scientifiques, chacun apportant au projet ses compétences uniques. Les projets peuvent tirer parti des réseaux pour accéder aux compétences exactes qui sont nécessaires. Les chercheurs peuvent profiter des réseaux pour accéder à un plus large éventail de projets susceptibles d'avoir recours à leurs compétences. Les disciplines scientifiques ont, en tenant compte de ces effets, modifié leur organisation de travail, et les personnes interrogées attribuent en partie ces modifications à l'existence de la communication par ordinateur. Dans chaque situation, un thème commun est que l'adoption de la communication par ordinateur se fait de manière telle qu'elle reproduit les relations sociales et les pratiques de recherche locales entre les personnes qui collaborent à distance. Par conséquent, bien que la structure sociale soit quelque peu réorganisée suite au recours à la communication par ordinateur, il semble que cette réorganisation se limite en grande partie à la modification (l'élargissement) des personnes susceptibles de participer, le contenu de cette participation au groupe de recherche ne subissant que des modifications mineures. L'organisation de travail existante a été en grande partie reproduite sur une aire géographique plus vaste, la communication par ordinateur remplissant le rôle de lien que remplissait précédemment la communication face à face dans les collaborations locales (Allen, 1977).

Cette modification de l'organisation du groupe de travail pourrait devenir une des incidences les plus révolutionnaires de cette technique nouvelle. Les réseaux informatiques, en supprimant la composante géographique des structures de travail, sont en train de faciliter toute une série de nouvelles structures de travail, notamment les équipes multisites, le télétravail à domicile, les organisations décentralisées, les organisations virtuelles et autres formes d'équipes de travail géographiquement dispersées mais opérant en réseau. Toutefois, le point de savoir si ce changement est un avantage pour les personnes qu'il affecte ne fait pas l'unanimité. Par exemple, le télétravail à domicile permet aux travailleurs de rester chez eux et leur épargne les tracasseries des navettes et le coût et le temps nécessaires pour assurer, tous les matins, une présentation professionnelle de soi-même. En même temps, il coupe le travailleur de la stimulation sociale et professionnelle qui résulte de l'interaction informelle avec les collègues. De même, si les travailleurs se mettent à constituer des groupes de travail avec des collègues à distance, à leurs interactions face à face avec leurs collègues locaux pourraient se substituer des interactions, par l'intermédiaire de l'ordinateur, avec leurs collègues à distance (Van Alstyne et Brynjolfsson, 1996*b*). Bien que la communication par ordinateur puisse faciliter la collaboration interdisciplinaire, elle peut aussi entraîner une balkanisation de la science, les chercheurs utilisant leur temps limité de communication pour interagir uniquement avec les chercheurs de leur spécialité (n'importe où dans le monde). Il est possible que cette balkanisation des communautés scientifiques diminue demain la fécondation réciproque des idées d'une discipline à l'autre. On ne voit pas encore clairement quelle sera l'incidence de ces formes nouvelles d'organisation du travail scientifique sur les résultats.

Le courrier électronique est donc considéré comme un moyen d'étendre son réseau de collaborateurs potentiels et de faciliter les collaborations existantes en permettant aux contacts personnels de se poursuivre sous la forme de collaborations, même lorsque les collègues se sont quittés. Toutefois, la communication par ordinateur n'a pas été considérée comme pouvant se substituer entièrement à l'interaction face à face (impliquant souvent un déplacement) ou par téléphone.

IV. FRÉQUENCE DE LA COMMUNICATION

La technique de la communication par ordinateur, si on la compare à l'interaction face à face, impose aussi moins de contraintes sous l'angle de la quantité totale de communication produite. En règle générale, dans une rencontre face à face, une seule personne peut prendre la parole à la fois et le temps total de parole est d'habitude strictement limité. Certains types de communication par ordinateur, en revanche, permettent à tous les participants de s'exprimer simultanément. Par ailleurs, la discussion peut durer plus longtemps, puisque tous n'ont pas besoin

d'être réunis en même temps en excluant toute autre activité. (Siegel *et al.*, 1986 ; McGuire *et al.*, 1987 ; Dubrovsky *et al.*, 1991 ; Weisband, 1992).

Dans le domaine de la science universitaire, certains éléments indiquent que la communication par ordinateur peut contribuer à une augmentation globale de la quantité de communication réalisée au cours d'un projet de recherche (Walsh et Bayma, 1996*b*). Dans les expériences dans le domaine de la physique de la haute énergie, le courrier électronique a permis aux chercheurs de maintenir leur participation aux expériences de longue durée, même lorsqu'ils ne sont pas physiquement présents au laboratoire, en recourant aux listes de diffusion, aux messageries télématiques et au courrier électronique et en diffusant par voie électronique des publications préliminaires et d'autres informations (informelles) essentielles. Cette méthode de communication permet à tous les membres de la collaboration de « rester dans le coup ». Les mémos peuvent être diffusés simultanément à tous les membres, et ceux qui ne sont pas physiquement présents au laboratoire peuvent donc participer aux discussions en cours au sujet de l'expérience (Knorr Cetina, à paraître). La communication par ordinateur n'est toutefois pas suffisante par elle-même. La « communication par ordinateur » représente toute une série de types de communication (Orlikowski et Yates, 1994). Le choix des types qui sont utilisés affecte les résultats. Les modalités d'utilisation de la technique particulière sont, elles aussi, importantes.

Un obstacle supplémentaire à la communication dans le cadre du travail scientifique est la dispersion dans le temps. Les expériences en physique se font souvent en trois postes, et le courrier électronique offre donc un moyen commode de maintenir informés les membres de tous les postes. Des travaux antérieurs consacrés aux organisations de police ont montré que le travail par postes interfère avec les flux d'information (Maltz *et al.*, 1991). D'autres travaux antérieurs sur le courrier électronique ont prouvé que le courrier électronique peut contribuer à vaincre en partie l'isolement en matière d'information, et l'isolement social et psychologique qu'entraîne le travail par postes (Huff *et al.*, 1989). La communication par ordinateur peut jouer un rôle important d'intégration des équipes de travail qui sont dispersées tant géographiquement que chronologiquement. Un physicien de l'état solide a souligné l'importance de la communication par ordinateur dans ses collaborations, opposant la communication par ordinateur aux formes précédentes de coordination :

« Avant les réseaux, nous travaillions ensemble [en collaboration]. Nous utilisons différents modes. Nous nous rendions visite. Les réseaux offrent un moyen de collaboration très suivie. On peut communiquer comme si on était dans le local voisin. Si on le souhaite, on peut appeler des données, les examiner à l'écran. C'est comme si on était ensemble, tout en étant éloignés. Les fuseaux horaires peuvent même être un avantage. J'envoie un message à la fin

de la journée. Il travaille dessus pendant que je dors. J'ai la réponse le lendemain matin, heure locale.» (Walsh et Bayma, 1996b).

Ce qui précède semble indiquer que la communication par ordinateur permet une plus grande intensité de communication. Les collaborations fondées sur la communication par ordinateur permettent, plutôt que de longues périodes d'isolement ponctuées de rencontres face à face périodiques intenses, d'entretenir un niveau élevé de contact entre les collaborateurs distants. La communication par ordinateur contribue par ailleurs à surmonter le problème de la dispersion chronologique (résultant en l'espèce des fuseaux horaires plutôt que du travail posté). En outre, la dispersion chronologique est à présent considérée comme un avantage. En s'échangeant constamment les tâches de recherche, les collaborateurs peuvent créer un projet qui ne dort jamais (Sudweeks et Rafaeli, 1996).

De même, les chimistes et les biologistes ont relevé que les réseaux permettent une communication plus étroite avec des collègues éloignés. Cette possibilité est particulièrement importante pour les chimistes et les biologistes, dont les expériences couplées de manière lâche se trouveraient souvent désynchronisées suite à la fréquence faible des contacts dans les formes classiques de communication. Comme l'explique un chimiste :

« Nous menons des projets qui impliquent des lasers, la technique de l'ultra-vide, les faisceaux ioniques, les faisceaux moléculaires. Il faut des années pour réunir ces éléments, des années pour faire les expériences. Beaucoup de choses ont changé avec l'arrivée du courrier électronique et de son utilisation ouverte à tous. Avant, il y avait une expérience en cours en Californie ou n'importe où, on leur téléphonait tous les deux à trois mois, peut-être, pour coordonner les choses. A présent, avec le courrier électronique, ce n'est plus un problème. » (Walsh et Bayma, 1996b).

Un des effets de cette densité de communication accrue peut être un plus grand attachement au groupe de recherche et à la discipline. Le travail scientifique peut être extrêmement aliénant, en partie en raison de l'isolement résultant des heures de travail irrégulières et de la concentration sur un effort hautement spécialisé qui peut même n'intéresser guère les collègues locaux. Comme l'observe Huff *et al.* (1989), l'utilisation du courrier électronique peut contribuer à vaincre ce sentiment d'isolement et, dès lors, déboucher sur un engagement accru. La citation suivante des réflexions d'un physico-chimiste qui a recours aux réseaux pour la collaboration au niveau international souligne une des caractéristiques importantes du réseau, à savoir sa capacité à maintenir actifs des liens lâches :

« Les potins forment à coup sûr une composante importante des messages. Lorsqu'on ne voit les gens qu'une seule fois par an, on a besoin de leur faire savoir qu'on s'intéresse à eux en tant que personne, et pas seulement en tant

que voix décharnée à l'autre bout du fil. A défaut, je ne pense pas que l'on pourrait maintenir un projet de longue durée comme celui-ci.» (Hurd et Weller, 1997).

Si on s'imagine un réseau comme un ensemble de connexions qui se dégradent avec le temps si elles ne sont pas activées de temps à autre (un peu à l'instar de la conception que l'on a généralement des réseaux neuronaux), la communication électronique fournit un mécanisme important pour l'entretien des liens avec d'autres qui, probablement, quitteraient le réseau si la communication dépendait de moyens plus formels ou plus envahissants. Une utilisation courante des réseaux consiste à envoyer occasionnellement une note dont le message sous-jacent est «Je sais que vous êtes là et je souhaite toujours maintenir un lien avec vous». Le tableau 2 montre que l'effet le plus marqué du courrier électronique est de renforcer le contact avec ceux qui font partie d'autres institutions. Sproull et Kiesler (1991, p. 101) concluent que, bien qu'elle ne puisse remplacer complètement les interactions face à face, la communication par ordinateur peut offrir une voie d'interaction sociale à des personnes qui, sans cela, pourraient se trouver isolées. Cette interaction accrue peut déboucher sur une satisfaction et un engagement accrus à l'égard du travail.

On a relevé, en revanche, que la communication par courrier électronique supprime en grande partie les civilités et le bavardage qui accompagne généralement les conversations face à face, voire téléphoniques. Bien que certaines personnes interrogées aient considéré cela comme un avantage du courrier électronique (en ce sens que les coûts liés au temps se trouvent réduits), la collégialité ainsi créée est aussi plus aride et plus utilitaire¹. Si la communication par ordinateur permet certes une communication plus fréquente et plus interactive entre collègues éloignés, elle peut aussi créer une structure de communication entre collègues plus distante, plus utilitaire. Cet effet peut s'étendre à ceux qui travaillent dans la même institution. Les scientifiques ont relevé qu'ils utilisaient souvent le courrier électronique pour communiquer avec des collègues locaux et citent, comme raisons d'opter pour ce mode, les avantages du caractère asynchrone et du fait qu'ils perdent moins de temps en civilités. Bien que cela permette aux messages de circuler plus rapidement, cette façon de procéder pourrait aussi créer un milieu de travail plus isolé socialement et plus aliénant. En réalité, la communication par ordinateur peut avoir sur les personnes un effet à la fois d'intégration et d'isolement. Elle permet un contact plus fréquent et plus individualisé avec d'autres personnes éloignées (la distance pouvant être géographique ou sociale). En même temps, elle peut rendre ces interactions plus arides. Le résultat pourrait être un milieu ou une communauté de travail où chacun est relié à un plus grand nombre de personnes mais où ces liens deviennent plus utilitaires ou moins satisfaisants. Il est nécessaire de poursuivre la recherche en vue de préciser dans quelles conditions la commu-

nication par ordinateur devient un moyen pour intégrer dans des groupes virtuels des individus dispersés, et dans quelles autres elle est un médium qui isole et aliène les personnes (Walther, 1996).

V. MONDIALISATION DE LA SCIENCE

Outre ses effets sur la collaboration, il se peut aussi que la communication par ordinateur soit en train d'accroître une coopération géographique d'un caractère moins explicite (Sanderson, 1996). Dans une série de domaines, les scientifiques produisent actuellement des ensembles de données qui, par l'intermédiaire de l'Internet, sont accessibles aux scientifiques du monde entier. Les exemples les plus célèbres sont sans doute ceux qui appartiennent au domaine de la biologie. Il s'agit notamment du projet sur le génome humain et de la banque de données sur les protéines, qui offrent aux biologistes un ensemble croissant de données sur le matériel génétique et les protéines (Walsh et Bayma, 1996a). Ces ensembles de données (qui bénéficient de contributions en ligne et qui sont accessibles en ligne) fournissent aux scientifiques du monde entier des mécanismes permettant d'échanger les données et de combiner les résultats d'études coordonnées de manière lâche. En réalité, un grand nombre de revues scientifiques exigent aujourd'hui, comme condition préalable à leur publication, que les données soient déposées dans ces bases de données. Cette science mondialisée ne concerne pas uniquement la biologie. L'Internet offre aux scientifiques dans différents pays de nouvelles possibilités de combiner les séries de données locales et de créer des séries de données mondiales qui peuvent être utilisées pour répondre à des questions qui dépendent de la collecte de données provenant du monde entier. Par exemple, plusieurs projets combinent des données mondiales des géosciences (Bierly, 1988). Sanderson (1996) note la manière dont des groupes de physique atmosphérique ont étudié l'utilisation du protocole de transfert de fichiers et la connexion à distance à la fois pour maintenir l'intégrité du fichier très volumineux de données satellitaires nettoyées et pour permettre aux groupes de recherche dans trois pays d'accéder à ces données. Dans le domaine des sciences sociales, Rainwater et Smeeding (1988) décrivent l'étude luxembourgeoise sur les revenus qui, actuellement, rassemble des données concernant 20 pays. Ce projet permet aux chercheurs de procéder à une analyse des données au niveau des personnes en envoyant par courrier électronique vers le centre de données à Luxembourg des programmes SAS et SPSS. Le centre de données fait office de dépositaire des données et vérifie également les programmes afin de garantir la confidentialité des données. Les chercheurs peuvent ainsi accéder gratuitement à ces données du micro-niveau provenant de 20 pays² sans devoir se rendre à Luxembourg, tout en maintenant le caractère confidentiel nécessaire pour la collecte de ces données.

Dans les sciences humaines, Ruhleder (1994) et Everhart (1996) décrivent des utilisations similaires de l'Internet, dans lesquelles les séries de données sont des textes de la littérature classique ou médiévale. Il se peut donc que l'Internet soit en train de créer une science mondiale à la fois dans le sens de la collaboration internationale décrite ci-dessus et dans le sens de l'intégration des données provenant du monde entier et de leur mise à disposition aisée des chercheurs, quel que soit l'endroit où ils se trouvent.

VI. EFFETS DE PÉRIPHÉRICITÉ

Distinctions de statut et accès à l'information

D'une manière générale, les avis restent divergents quant à l'impact de la communication par ordinateur sur l'organisation de travail s'agissant, en particulier, des effets périphériques. Plusieurs études ont montré que la communication par ordinateur peut déboucher sur une décentralisation accrue ou sur un rapprochement des statuts au sein d'une organisation ou d'un groupe (Dubrovsky *et al.*, 1991 ; Sproull et Kiesler, 1986 ; Rice, 1980). Cette constatation est généralement attribuée au fait que l'interaction sur l'Internet fournit moins d'indications de statut que la communication face à face, ou même que l'interaction par courrier postal ou par téléphone. Les distinctions importantes qui n'apparaissent pas sont le rang hiérarchique et le sexe. Les décisions de groupe sont moins influencées par le statut de ceux qui proposent des solutions particulières (McGuire *et al.*, 1987). Toutefois, dans la mesure où les différences de statut sont reproduites dans ce genre nouveau de communication, les personnes de statut élevé continueront à peser plus fortement sur les décisions du groupe (Orlikowski et Yates, 1994). Weisband *et al.* (1995) ont constaté que les expériences visant à comparer la communication par ordinateur et les groupes qui communiquent face à face ne révèlent que peu de différences sur le plan de l'inégalité de participation au processus décisionnel. Ils suggèrent qu'une des explications possibles de l'écart entre leurs résultats et ceux de travaux précédents est que les membres du groupe étaient au fait des différences de statut entre les membres, même dans le cas de la communication par ordinateur. A mesure qu'elle se développe, la technique commence à introduire dans la communication davantage d'indications de statut, par exemple, des adresses électroniques plus facilement déchiffrables et des photos et des biographies sur les pages Web. Par ailleurs, on commence à voir apparaître d'autres mécanismes visant à réinstaurer les procédures de renforcement du statut qui caractérisent les anciennes techniques de communication (tels que le courrier postal et le téléphone). Par exemple, les scientifiques de statut élevé (et aussi, sans doute, les cadres de statut élevé) ont recours à des assistants chargés de filtrer leur courrier électronique au même titre que leurs appels téléphoniques et leur courrier postal.

De même, dans la mesure où cette technique viole certaines normes de travail ou certaines distinctions de travail existantes, il arrive que l'on n'y ait pas recours et, partant, qu'elle n'ait guère d'incidence (Sproull et Kiesler, 1991 ; Orlikowski, 1993).

En outre, la technique nouvelle peut transformer certaines des bases qui sous-tendent les distinctions de statut existantes. Barley (1990), par exemple, montre que les changements de technique au travail peuvent renforcer le statut des jeunes collègues, plus familiarisés avec les dernières techniques. Les réseaux d'ordinateurs peuvent aussi ouvrir plus largement l'accès à des ressources essentielles, notamment le matériel, les logiciels ou les bases de données informatiques (Dongarra et Grosse, 1987). L'accès aux ressources était auparavant réparti de manière inégale en science (Merton, 1968 ; Cole et Cole, 1967 ; Allison et Stewart, 1974). Dès lors que les réseaux sont en mesure de rendre ces ressources accessibles à un plus grand nombre de scientifiques, ils pourraient accroître la productivité des scientifiques périphériques (c'est-à-dire ceux de rang plus modeste, moins éminents ou qui ne font pas partie des grandes institutions). Cet accès renforcé pourrait réduire l'écart entre le noyau de scientifiques d'un domaine et leurs collègues moins éminents. Une étude récente sur l'utilisation des réseaux en océanographie a conclu que les jeunes océanographes qui utilisaient plus intensivement les réseaux avaient plus de chance de bénéficier d'une reconnaissance professionnelle que leurs collègues du même âge qui avaient moins recours au réseau (National Science Board, 1993). De même, les océanographes de l'intérieur qui utilisaient les réseaux publiaient plus que ceux qui étaient moins actifs sur les réseaux. Ces résultats semblent indiquer que la technique est peut-être en train de transformer la structure sociale de l'océanographie en offrant un meilleur accès aux scientifiques de statut inférieur. Toutefois, l'enquête de Cohen (1995) ne trouve pas d'indices de l'existence d'un processus de démocratisation. Par exemple, la communication par ordinateur n'a pas apporté d'avantages supplémentaires à ceux qui travaillaient dans des institutions moins prestigieuses ou aux jeunes scientifiques.

Un des résultats de la capacité du réseau à entretenir les liens avec les collaborateurs distants (comme relevé ci-dessus) est que les scientifiques qui ont été formés dans des centres de recherche peuvent maintenir leurs contacts de recherche lorsqu'ils vont travailler dans des institutions plus périphériques. Un mathématicien, interrogé au sujet de la nécessité d'être sur l'Internet pour être actif dans le domaine de la recherche, donne la réponse suivante :

« Il y a beaucoup de gens ici, beaucoup de gens qui passent ici, qui font le même type de travail que moi. Je peux me passer de l'Internet pour faire de la recherche. Mais j'ai un poste temporaire. L'année prochaine est ma dernière année. Là où j'irai après, il est possible que l'on soit obligé d'être sur le réseau. Par exemple, si je me retrouve au Nouveau-Mexique à Las Cruzas, le courrier électronique pourrait être vital si je veux maintenir mon activité de recherche. » (Walsh et Bayma, 1996b).

Ici aussi, le résultat semble indiquer que les réseaux électroniques ne sont pas forcément en train de se substituer au contact face à face ou de supprimer l'importance des centres de recherche dans ces disciplines. Ils sont plutôt utilisés pour éviter que ne s'atrophient les liens qui ont été créés au cours de ces contacts. Il se peut, vu sous cet angle, que le groupe des chercheurs actifs dans un certain domaine soit en train de s'élargir dans la mesure où ceux qui vont s'installer dans des institutions plus périphériques sont mieux en mesure de maintenir les liens indispensables pour rester actifs dans le domaine de la recherche. Dans le même ordre d'idées, un biologiste a soutenu que les conférences nationales et internationales ont gagné en importance à présent qu'il est possible de se familiariser avec le travail des participants vivant loin les uns des autres et qu'il est devenu intéressant de nouer des liens à distance.

Comme le relève un physicien expérimental, cet effet périphérique prend généralement la forme d'une participation accrue des institutions moins prestigieuses, plutôt que celle d'une restructuration de la hiérarchie des statuts :

« Je pourrais dire qu'elles [les institutions plus modestes/moins prestigieuses] savent mieux ce qui se passe dans les grandes institutions, mais je ne vois pas d'exemple d'une institution moins prestigieuse en Italie [son pays d'origine] qui ait vu son prestige accru grâce à l'Internet. » (Walsh et Bayma, 1996b).

De même, un mathématicien relève que les changements attribués au réseau résultent en grande partie de l'intégration d'un plus grand nombre de chercheurs de la base plutôt que d'une transformation du travail de ceux qui sont au sommet :

« Les mathématiques véritablement créatives sont un processus de création personnel. On dispose d'une grande quantité d'informations. Une bonne part de la recherche n'est pas sensationnelle. Pour une bonne partie, la meilleure parfois, elle consiste à s'écarter des grandes orientations. Faire du travail vraiment créatif. Les réseaux ne sont pas un facteur aussi important que cela. Si vous êtes en communication avec les géants dans un domaine, vous avez plus de chances d'être informé très tôt des idées nouvelles. Il faut faire partie de ce transfert d'informations. Certains des plus grands mathématiciens au monde n'ont aucun contact avec les ordinateurs. Pour un mathématicien de capacité moyenne, le courrier électronique améliore les chances de "donner le meilleur de soi-même", comme on dit à l'armée. Ici, au sommet, ils ne peuvent pas s'améliorer beaucoup, et personne n'oserait les harceler pour qu'ils changent. Le courrier électronique est un auxiliaire. Si on vous exile à Trifouilly-les-Oies, ce n'est plus, comme avant, une condamnation à mort. » (Walsh et Bayma, 1996b).

Comme le montrent les citations ci-dessus, le courrier électronique permet à ceux qui, précédemment, ont été exclus des informations les plus récentes d'avoir accès à celles-ci. Tous affirment que cet accès est important pour se tenir au courant

dans le domaine, et ceux qui travaillent dans des institutions moins prestigieuses sont à présent en mesure de participer aux communautés scientifiques. Bruce (1994) note que son échantillon de chercheurs australiens souligne que le réseau leur permet de rester au courant dans leurs domaines respectifs. La communication par ordinateur facilite par ailleurs, outre les contacts personnels potentiels, l'accès à d'autres types de ressources tels que les bases de données communes ou les équipements informatiques (Walsh et Bayma, 1996b ; Finholt *et al.*, 1995). Toutefois, ces personnes suggèrent aussi que le résultat n'est pas tant un nivellement de la science qu'un élargissement de celle-ci. Les personnes et les institutions qui sont au sommet de la hiérarchie restent les figures centrales. La différence, c'est qu'à présent, un plus large éventail de personnes et d'institutions apporte sa contribution à la masse des connaissances scientifiques ou, du moins, est en mesure de suivre de plus près les progrès dans leur domaine.

On trouve en physique théorique un exemple supplémentaire de cet effet d'élargissement plutôt que de nivellement. Les physiciens (en particulier les théoriciens de la physique des particules) ont commencé à institutionnaliser la diffusion électronique des publications préliminaires. Celles-ci constituent un canal important de communication informelle. Dans les années soixante-dix, Garvey et Griffith (1979) ont relevé que la diffusion des publications préliminaires avait augmenté, au point de devenir parfois une charge considérable pour les auteurs. Ils ont aussi relevé que «souvent, ceux qui ont le plus besoin de publications préliminaires – les jeunes scientifiques, les travailleurs dans les petites institutions, et les chercheurs dans les pays moins développés – ne sont pas ceux à qui elles parviennent» (Garvey et Griffith, 1979, p. 158). Au début des années 90, un physicien théoricien de Los Alamos a créé une messagerie télématique consacrée aux publications préliminaires (Taubes, 1993 ; Merz, à paraître). Cette messagerie fournit aux abonnés des résumés analytiques de toutes les nouvelles publications et leur permet d'envoyer une demande par courrier électronique et d'obtenir la publication complète, graphiques compris, par voie électronique. Le service a débuté en août 1991 et, en décembre 1992, il comptait déjà 8 000 abonnés. Très rapidement, la base de données s'enrichissait de 600 publications nouvelles par mois.

Cette base de données sur les publications préliminaires est aussi un exemple de la manière dont l'importation d'un type de communication d'un médium (publications préliminaires envoyées par la poste) dans un autre (messagerie électronique) représente un changement qui transforme et renforce à la fois la structure de communication existante au sein d'une communauté (Orlikowski et Yates, 1994). Cette base de données sur les publications préliminaires a contribué à intégrer, à l'échelle mondiale, un plus grand nombre de physiciens théoriciens dans le circuit des publications préliminaires. Toutefois, comme le souligne un physicien théoricien dans le témoignage ci-dessous, cet effet périphérique est tempéré par la

nécessité d'avoir accès à des informations locales pour traiter les documents qui se trouvent dans la base de données des publications préliminaires :

«Elles [les institutions périphériques] s'en sont bien trouvé. Si vous vous trouvez dans le tiers monde, dans les pays sous-développés ou dans une petite institution, en principe, l'accès est devenu facile. Mais il reste des obstacles. Avant, on était isolé parce qu'on ne pouvait obtenir les informations. Aujourd'hui, quelqu'un qui se trouve en Inde peut, en principe, obtenir les informations de la même manière que quelqu'un qui se trouve à Princeton. C'est un changement considérable. Toutefois, voir une publication et se trouver sur place sont deux choses différentes. Si vous êtes dans une grande institution, vous avez accès aux informations orales, aux colloques, vous pouvez parler à la personne. Ca, on ne l'a pas encore. Ils reçoivent plus de 100 publications par mois. Dans les grandes institutions, il y a quelqu'un qui sait ce qui est important et on peut se faire aider pour trier tout cela. Les réseaux sont une aide pour ceux qui ont les connaissances spécialisées requises. Eux peuvent tirer profit de cette nouvelle forme d'accès.» (Walsh et Bayma, 1996b).

Garvey (1979) a fait valoir que le filtrage et le tri qu'offre cette communication locale et informelle constituent une partie importante du processus de recherche des informations scientifiques. La citation ci-dessus semble indiquer que, si la communication par ordinateur est certes une aide, elle ne suffit pas en soi pour vaincre les inconvénients découlant de l'absence de contact direct avec les scientifiques les plus actifs dans le domaine que l'on pratique. Bien que la base de données des publications préliminaires permette en effet à un groupe plus vaste de personnes d'assister au dialogue entre les figures de proue de la recherche dans le domaine concerné, seul un petit sous-ensemble de personnes est en mesure de profiter de cet accès aux informations plus rapide et plus large qu'offre la base de données. La plupart de ceux qui constituent ce sous-ensemble sont les figures de proue de la recherche, qui s'échangent déjà des publications et qui communiquent par d'autres voies informelles.

Le courrier électronique facilite par ailleurs l'établissement de liens qui, sans lui, ne se créeraient peut-être jamais. La citation suivante d'un mathématicien montre comment la communication par ordinateur libère la communication de certaines des inhibitions sociales des autres formes de communication :

«Une fois que vous ouvrez la communication, vous pouvez avoir des interactions avec les gens. Peut-être n'oserais-je pas entrer à Harvard pour y rencontrer les gens face à face, mais je peux envoyer un courrier électronique à Harvard. Mon fils, lui, y allait carrément et prenait contact avec des gens prestigieux dans son domaine. Il ignorait qui ils étaient, mais je reconnaissais cer-

tains noms. On recule peut-être devant un contact face à face mais, dans le courrier électronique, on ignore à qui on parle.» (Walsh et Bayma, 1996b).

La citation suivante, due à un physicien théoricien, démontre la même chose, mais vu sous l'angle du destinataire de ces communications :

«Je n'ai pas entamé de travail avec des étrangers par le biais de l'Internet. Mais je me suis trouvé du côté du destinataire de ces demandes. D'autres ont écrit : "Cela a l'air intéressant. Nous travaillons sur un sujet X. Pouvez-vous nous faire des suggestions ?" Lorsqu'ils obtiennent un résultat, ils demandent à faire une publication conjointe. Jadis, on faisait cela par lettre. Mais le côté informel du courrier électronique apporte quelque chose. Dans le cas d'une lettre, on réfléchit plus longuement. [Le courrier électronique] est plus informel. On peut envoyer un bref message ; si c'était une lettre, on ne l'enverrait peut-être pas» (Walsh et Bayma, 1996b).

Comme l'observent Sproull et Kiesler (1986), les indications plus vagues sur le contexte social et le caractère informel du courrier électronique réduisent les contraintes qui pèsent sur les personnes de statut inférieur lorsqu'elles entrent en contact avec des personnes de statut supérieur au leur. La citation ci-dessus semble indiquer que les scientifiques qui pourraient reculer devant l'envoi d'une lettre destinée à créer un contact (par exemple, pour une question de différence de statut) seraient disposés à entamer un contact par courrier électronique. Il se peut dès lors que le courrier électronique facilite l'établissement de liens nouveaux entre collaborateurs distants et mette les scientifiques de statut inférieur en mesure d'interroger leurs collègues plus éminents. L'impact de ce changement n'est pas certain. La citation ci-dessus semble indiquer qu'un contact de ce type pourrait entraîner une disparité encore accrue entre les taux de publication, les scientifiques les plus éminents se trouvant rattachés à un réseau encore plus vaste de projets de recherche créés par le biais de contacts par courrier électronique. D'un autre côté, ces contacts pourraient permettre aux scientifiques qui, jusque là, ne disposaient pas de l'accès nécessaire pour rester au courant de devenir des participants actifs, voire des membres du noyau de base des scientifiques dans leurs domaines. Il faut toutefois relever que Cohen (1995) n'a constaté aucune interaction importante entre l'âge ou le prestige institutionnel et le recours à la communication par ordinateur en tant qu'indicateur prévisionnel de productivité.

Les réseaux peuvent aussi briser la forte corrélation entre la distance et les interrogations informelles (McGuire *et al.*, 1987 ; Kraut *et al.*, 1990). Chez les scientifiques, la capacité des réseaux à diffuser rapidement des questions à un grand nombre de personnes par le biais de groupes de discussion électroniques leur permet d'interagir avec des personnes disséminées sur une vaste zone comme si elles se trouvaient dans un local voisin (Sanderson, 1996). La déclaration suivante d'un mathématicien en est un exemple typique :

« J'ai eu recours une fois [au groupe de discussion sci.math]. J'ai écrit un article et l'ai envoyé à l'éditeur. Il m'a dit : "Ça a un rapport avec le résultat X de Untel." Mais il n'avait pas la référence. J'ai donc posé la question sur netnews. "Quelqu'un peut-il me donner la référence ?". Le lendemain, deux ou trois personnes ont envoyé la référence exacte. C'est une question facile, mais la réponse est difficile à trouver dans une bibliothèque. Il faudrait des heures pour la trouver. Je peux interroger plusieurs milliers de personnes à la fois. Il existe une chance que deux ou trois d'entre elles connaissent la réponse. Il n'existe pas d'autre façon de le faire. Il est impossible [en dehors du courrier électronique] de poser une question insignifiante à quelques milliers de personnes à la fois. La question est importante pour moi, mais pour personne d'autre. Ils ont été heureux de pouvoir m'aider. » (Walsh et Bayma, 1996a).

Constant *et al.* (1996) constatent que la communication par ordinateur facilite l'échange d'informations avec ceux qui se trouvent dans d'autres unités d'une grande société multinationale d'informatique. Finholt (1993) constate, dans la même société, que la communication par ordinateur permet l'accès à la mémoire organisationnelle par l'intermédiaire de sa base de données des questions diffusées précédemment. Le tableau 2 ci-dessus montre que le principal effet du courrier électronique consiste à augmenter les contacts avec ceux qui travaillent dans d'autres institutions. Van Alstyne et Brynjolfsson (1996b) affirment toutefois (en se basant sur un modèle économique formel) que cet accès accru aux collègues, quelle que soit la distance, conjugué avec les économies réalisées en matière d'échange d'informations et avec la maximalisation des avantages pour les partenaires des échanges, débouchera sur une inégalité accrue en science, puisque chaque personne limite ses interactions à ses pairs sur le plan de l'information, quel que soit l'endroit où ils se trouvent. Il est nécessaire de poursuivre les recherches empiriques si on veut y voir clair dans ces conclusions et ces prédictions contradictoires.

VII. PRODUCTIVITÉ

Il est possible, par ailleurs, que les réseaux informatiques renforcent la productivité des scientifiques, bien qu'ici, les données indiquent généralement une corrélation plutôt qu'une causalité. Plusieurs études antérieures ont conclu que ceux qui font des réseaux informatiques l'usage le plus intensif sont aussi ceux qui sont les plus productifs. Il se pourrait toutefois que ceux qui sont les plus productifs soient aussi ceux qui ont le plus besoin de profiter des avantages de la communication par ordinateur, ou qui y ont le plus intérêt. Une des premières études sur les systèmes de téléconférence informatisée (portant sur EIES aux États-Unis et COM en Suède) a conclu que les mesures objectives d'utilisation présentaient une asso-

ciation positive significative avec les mesures subjectives de productivité et de progrès de carrière pour des échantillons de scientifiques et de personnel de recherche-développement (Hiltz et Johnson, 1989 ; Hiltz, 1988). Hesse *et al.* (1993) a trouvé qu'il existait une corrélation positive entre l'utilisation des réseaux informatiques et la productivité, même après contrôle des autres variables liées à la productivité, notamment l'âge et le prestige de l'institution. Dans une étude portant sur des chimistes d'université au Royaume-Uni, Philip (1995) a constaté une corrélation entre la place du département dans la hiérarchie et l'utilisation des systèmes d'information chimiques en ligne. Cohen (1995) a également conclu à l'existence d'une corrélation significative entre l'utilisation de la communication par ordinateur, d'une part, et les publications et la reconnaissance professionnelle, d'autre part, ici aussi après contrôle des autres variables liées. Kaminer et Braunstien (1998) ont aussi constaté cette relation positive entre productivité et communication par ordinateur.

Certains éléments indiquent que les utilisateurs attribuent à la communication par ordinateur des effets sur la productivité. Les personnes interrogées par Cohen associaient un accroissement de productivité aux avantages offerts par la communication par ordinateur (en particulier, une communication plus facile avec les collègues sur d'autres campus). De même, des universitaires en Australie ont noté que les réseaux augmentaient leur efficacité au travail, bien que cette étude n'ait pas mesuré directement la productivité (Bruce, 1994). Bishop (1994) a constaté que, chez les ingénieurs aérospatiaux, les personnes interrogées avaient le sentiment que les réseaux augmentaient la quantité d'informations disponibles, renforçaient l'efficacité en matière d'établissement de contacts, et accroissait leur capacité à achever les projets dans les délais, ce qui semble indiquer que les réseaux améliorent effectivement l'efficacité des travailleurs. Son étude a toutefois conclu aussi que les réseaux pourraient accroître « le temps perdu en futilités ». Walsh *et al.* (à paraître) constatent que plus de la moitié des personnes interrogées estimaient que le courrier électronique renforçait la productivité et que les mathématiciens et les physiciens, en particulier, signalaient plus souvent un effet sur la productivité (tableau 2). Ils concluent également à l'existence d'une relation positive entre le recours à la communication par ordinateur et le nombre d'articles publiés. Comme Bishop, toutefois, ils constatent qu'un nombre important de personnes estiment que le courrier électronique multiplie aussi les distractions par rapport au travail (tableau 2).

S'agissant de l'impact de l'Internet sur la productivité de la science (et non plus de la productivité des scientifiques), les indications sont encore plus minces. On pourrait prétendre que l'accumulation plus rapide des articles et l'accès à une plus grande quantité de données devraient permettre une succession plus rapide des découvertes scientifiques. Merz (à paraître) suggère que l'archive des publications préliminaires en physique peut accroître la productivité scientifique non seulement

en rendant la recherche plus efficace et en accélérant la publication mais aussi en synchronisant l'activité de recherche et en permettant une répartition plus efficace du travail entre les équipes de recherche. Uzumeri et Snyder (1996) soutiennent qu'il existe une accélération de la science due à l'Internet s'agissant, en l'occurrence, de la découverte des défauts des produits. Ils comparent le cas du défaut du bouclier de Dalkon, qui est resté secret pendant six ans, et le défaut du processeur Pentium, qui a été mis à jour en six semaines suite à l'activité sur l'Internet. Il n'en demeure pas moins que la question de l'incidence de l'Internet sur le rythme des découvertes scientifiques reste sans réponse. Cole et Cole (1972) affirment que le rythme de la science est essentiellement déterminé par les figures de proue du domaine concerné et qu'une diminution de la participation de la majorité des scientifiques n'aurait guère d'incidence. Par extension, leur argument (l'hypothèse Ortega) semble indiquer qu'un élargissement de la participation de ceux qui travaillent dans les institutions périphériques ne devrait avoir que peu d'effet sur le rythme de la science. Bien qu'elle soit difficile à mesurer et à tester, cette question de l'impact de cette technique nouvelle sur le progrès scientifique constitue la question clé qui, pour l'heure, reste sans réponse.

VIII. CONTEXTES SOCIAUX

De nombreux effets de la communication par ordinateur peuvent dépendre de divers facteurs sociaux (Kling, 1995). Si cette technique peut certes être utilisée pour démocratiser l'accès à diverses ressources, elle peut aussi être utilisée de manière à garantir que seuls ceux qui travaillent dans les institutions d'élite aient accès à ces ressources (Rice, 1994). Orlikowski et Yates (1994) soutiennent que le registre utilisé pour la communication dans une communauté en ligne est fortement influencé par les normes d'interaction que les membres ont héritées des modes antérieurs d'interaction. Bien qu'il existe certaines indications de l'impact de la communication par ordinateur, nous suspectons que le contexte dans lequel cette communication est adoptée est un facteur de médiation important pour expliquer les effets de ce mode de communication (Sproull et Kiesler, 1991). Par exemple, Walther (1997) constate que la structuration de la collaboration et les attitudes des individus vis-à-vis de cette collaboration (durable ou de courte durée, de groupe ou entre personnes) interagissent avec les effets de la communication par ordinateur, en ce sens que certaines conditions produisent un effort plus grand, et d'autres un effort moins grand que dans les groupes qui ont un contact face à face.

Les effets périphériques dépendent de l'universalité de l'accès. Dans bon nombre de circonstances, ces conditions ne sont pas remplies. Par exemple, si l'accès à l'Internet est quasi universel chez les universitaires aux États-Unis (Jacobson,

1994 ; Gurbaxani, 1990), cet accès est nettement moins uniforme en dehors des États-Unis (Jacobson, 1994). En outre, la communication par ordinateur peut avoir une structure telle qu'elle renforce les distinctions de statut. Par exemple, les messageries télématiques telles que la hiérarchie sci.* peuvent être utilisées pour relier les membres travaillant dans le même domaine. Elles peuvent être utilisées pour annoncer de nouvelles découvertes, pour discuter de questions de fond et pour obtenir des réponses de collègues inconnus. La liste de diffusion propre à un domaine constitue un mécanisme similaire. Il existe toutefois une différence importante : l'existence d'une messagerie télématique du type «netnews» est communiquée à un grand nombre de serveurs informatiques («le monde»), tandis que l'existence de la liste de diffusion ne peut être annoncée que par contact direct par le biais de liens de recherche existants. Il peut même arriver que l'accès à la liste soit limité, de sorte que seul un groupe déterminé peut s'abonner. Si on adopte la première forme, toute personne ayant accès aux messageries télématiques peut être informée des discussions au sein du groupe électronique et y participer. Si on opte pour la seconde forme, seules les personnes retenues auront accès aux informations transmises à la liste. S'il s'agit bien, dans les deux cas, d'exemples de communication par ordinateur, les résultats peuvent être très différents, s'agissant de l'accès des personnes qui ne sont pas encore reliées directement au noyau des scientifiques dans un domaine donné.

De plus, certains indices montrent qu'il existe un effet d'ajustement. Il semble que certains domaines soient potentiellement mieux en mesure que d'autres de tirer profit de cette technique. Walsh et Bayma (1996a) et Bishop (1994) suggèrent que les domaines caractérisés par une forte interdépendance (c'est-à-dire dans lesquels les collaborateurs doivent réagir fréquemment aux activités des uns et des autres) et dans lesquels les collaborateurs risquent d'être dispersés (mathématiques, physique et ingénierie aérospatiale, par exemple) sont les domaines les plus susceptibles de profiter de la capacité de la communication par ordinateur d'assurer une communication fiable, rapide et asynchrone. Hailman (1996) estime que le faible niveau d'utilisation en ornithologie peut s'expliquer en partie par des limitations techniques (la difficulté de transmettre les informations non textuelles, courantes dans ce domaine) et le rythme relativement lent des découvertes (où les publications anciennes restent pertinentes et les longs délais de publication ne sont pas considérés comme un problème).

Une des incidences des réseaux informatiques est qu'ils élargissent l'accès à l'information. Bien que cette caractéristique présente plusieurs avantages, comme signalé ci-dessus, cet effet peut aussi être interprété négativement. Walsh et Bayma (1996a) notent que de nombreux scientifiques étaient peu enclins à utiliser les réseaux informatiques par crainte de perdre la propriété d'informations précieuses. Cette réaction est particulièrement marquée chez les scientifiques des domaines de la chimie et de la biologie expérimentale, où la présence de l'industrie

est très affirmée et où les droits de brevet représentent des sommes importantes. De même, Bishop (1994) a constaté que les ingénieurs aérospatiaux (dont beaucoup travaillent dans le domaine de la défense) étaient très préoccupés par les effets des réseaux informatiques sur la sécurité du système et par les fuites de renseignements couverts par le secret industriel. Ce problème de la propriété de l'information continuera à tarauder les décideurs qui s'efforcent d'élargir l'accès à l'Internet et son utilité tout en protégeant les droits de propriété sur l'information.

IX. IMPLICATIONS POUR LA BIBLIOTHÉCONOMIE ET LES SCIENCES DE L'INFORMATION

Recherche en bibliothéconomie et communication par ordinateur

Bien qu'il existe, dans le domaine de la bibliothéconomie et des sciences de l'information, un volume considérable de publications qui traitent de l'utilisation de la communication par ordinateur dans l'accomplissement des tâches et des responsabilités quotidiennes et pratiques, les publications qui se penchent sur l'effet de cette communication sur les activités de recherche dans ce domaine sont plus clairsemées. Ce fait pourrait peut-être être rapproché du volume moins important de recherche universitaire en bibliothéconomie par rapport aux disciplines scientifiques examinées ci-dessus.

Kovacs *et al.* (1995), dans une étude à l'échelle nationale portant sur l'utilisation des téléconférences informatisées par les professionnels de la bibliothéconomie et des sciences de l'information, ont toutefois révélé que ce moyen de communication électronique entre universitaires jouait un rôle dans la recherche. La majorité des personnes interrogées dans le cadre de cette étude ont indiqué que les téléconférences renforçaient d'autres sources d'information : revues professionnelles, conférences professionnelles, communications téléphoniques et courrier postal. Qui plus est, 38,4 % d'entre elles ont signalé que les téléconférences étaient en voie de se substituer à ces autres formes de communication classique. Enfin, 33 % ont indiqué qu'ils participaient à des téléconférences informatisées par intérêt professionnel et par intérêt pour la recherche. Tout comme plus de 130 forums de discussion pour bibliothécaires (Abbott, 1994), les revues scientifiques électroniques qui diffusent les résultats de la recherche commencent à apparaître, notamment LIBRES, *Library and Information Science Research Electronic Journal* ; *MC Journal* ; *The Journal of Academic Media* ; et *Public-Access Computer Systems Review*.

En 1991, l'*Association of College and Research Libraries* (ACRL) a mis en place un programme d'encadrement par courrier électronique dans le but d'encourager et d'améliorer la recherche en bibliothéconomie (Echavarria *et al.*, 1995). L'ACRL a pris conscience du fait que le groupe de chercheurs actifs était réduit et que la commu-

nication par ordinateur fournirait le mécanisme permettant d'atteindre un plus grand nombre de professionnels dans l'espoir d'encourager la recherche. Bien que le programme ait connu quelques échecs, il a aussi eu des incidences durables sur les compétences des participants en matière de recherche. Comme l'explique un des bénéficiaires :

«Je suis passé, l'été dernier, du service du catalogage au service de références, et cette décision de changer d'orientation a été motivée par ma participation au [programme d'encadrement par courrier électronique]. Je me suis rendu compte qu'au service du catalogage, l'environnement n'était pas propice à l'activité de recherche et j'ai cherché avidement à établir le contact avec les utilisateurs des bibliothèques. Ma participation au [programme d'encadrement par courrier électronique] a eu une forte influence sur ce que j'ai lu, ce que j'ai pensé et l'endroit où je me trouve aujourd'hui.» (Echavarria *et al.*, 1995, p. 358).

Une autre utilisation novatrice de la communication par ordinateur est la mise sur pied de programmes virtuels permettant d'obtenir le diplôme de maîtrise en bibliothéconomie et en sciences de l'information. Les programmes universitaires qui occupent l'avant-scène de l'enseignement à distance par Internet sont ceux de l'université de l'Arizona, de l'université de l'Illinois (Urbana-Champaign), et de l'université de Syracuse. Le programme virtuel de l'université de l'Arizona qui, en 1997, comptait quelque 70 étudiants de par le monde, a recours à une combinaison de diverses formes de communication par ordinateur. Les cours sont transmis sur l'Internet par le biais d'une page Web, les discussions de classe ont lieu par l'intermédiaire d'un forum de discussion électronique, les tâches et les tests sont diffusés par ftp, les projets de groupe des étudiants sont coordonnés par la voie du courrier électronique, les projets définitifs sont publiés sur une page Web, et les étudiants communiquent avec leurs professeurs par courrier électronique. Il faut aussi souligner que les étudiants doivent prendre 12 unités de valeur sur le campus de l'université de l'Arizona. Comme relevé ci-dessus, l'interaction face à face est importante dans un effort entrepris en collaboration. Il sera intéressant d'observer quels seront les effets durables que cette formation virtuelle aura sur ces nouveaux spécialistes de la bibliothéconomie et des sciences de l'information, et de voir si, lorsqu'ils s'engageront dans la recherche, ils seront plus portés à collaborer avec d'autres par le biais de la communication par ordinateur.

Le rôle du bibliothécaire dans l'enseignement virtuel

Les bibliothécaires et les spécialistes de l'information ont toujours été particulièrement attentifs à la façon dont la connaissance scientifique est créée et diffusée. Pour les bibliothécaires, suivre la trace de l'enseignement virtuel dans le cadre de ces processus constitue un défi supplémentaire. Dans le cadre de la section des

bibliothèques universitaires de la conférence annuelle de 1997 de l'American Library Association, des bibliothécaires et des scientifiques universitaires ont proposé ce thème de discussion (Garrison *et al.*, 1997). Un des bibliothécaires a estimé que, dès lors que les scientifiques universitaires modifient leurs méthodes, les bibliothécaires vont eux aussi devoir changer. Un scientifique universitaire a lui aussi indiqué qu'il est possible que les bibliothécaires, de gardiens de collections qu'ils étaient, sont peut-être en train de devenir des découvreurs et des organisateurs de l'information (Garrison *et al.*, p. 544).

Le travail et la communication scientifiques se déroulent de plus en plus dans une enceinte électronique, et l'on voit apparaître certains signes qui indiquent que les bibliothécaires sont en train d'assumer, en prévision de cette évolution, un rôle à part entière dans la recherche des ressources de communication par ordinateur et dans l'organisation de l'information électronique, de manière à permettre à la majorité d'accéder à ces ressources et de les appréhender. L'Association des bibliothèques de recherche publie un répertoire complet des revues, bulletins et groupes de discussion universitaires sous forme électronique. Divers bibliothécaires ont aussi créé des bibliographies et des répertoires de ressources de communication par ordinateur (Abbott, 1994). De plus, les bibliothécaires s'efforcent de rester au courant des informations scientifiques accessibles sur l'Internet en offrant de manière organisée des liens vers ces informations par le biais des pages d'accueil des bibliothèques. Ces liens ont pour but d'améliorer l'accès pour les personnes qui, auparavant, étaient exclues de ces réseaux d'information.

Enfin, à présent qu'il existe une université virtuelle, les bibliothécaires (et, en particulier, les bibliothécaires attachés à la recherche ou aux universités) doivent peut-être repenser leurs relations avec les chercheurs. Les bibliothécaires sont en train de mettre en place une communauté virtuelle propre, tandis qu'une communication virtuelle distincte a lieu pour les chercheurs de diverses disciplines. Il pourrait être intéressant de se demander s'il ne serait pas nécessaire d'adapter les mécanismes de communication des libraires avec les chercheurs. Un certain chevauchement à l'intérieur de ces communautés virtuelles pourrait présenter des avantages pour les deux parties. Par exemple, certains ont fait valoir que, par l'instruction et le contact étroit avec la faculté, les libraires peuvent contribuer à éviter la balkanisation du réseau de communication des chercheurs qui résulte souvent des écarts de niveaux de connaissance techniques au sein des facultés (Schwartz, 1994).

X. CONCLUSIONS

A mesure que la communication par ordinateur joue un rôle de plus en plus éminent dans le travail universitaire, on assiste à un développement significatif de

la recherche relative à l'incidence de cette technique nouvelle sur la science. Les travaux empiriques antérieurs concernant les effets de l'Internet sur la science semblent indiquer que le travail scientifique est en train de changer en profondeur. Le changement le plus notable est peut-être la transformation des mécanismes de collaboration. On a assisté au développement de la collaboration à distance et, en particulier, au développement de la collaboration internationale. Ce n'est pas uniquement la présence de cette technique nouvelle qui a provoqué ces changements. Celle-ci a, en réalité, facilité l'évolution des mécanismes de travail en permettant aux chercheurs d'accéder facilement à leurs collègues qui ont les mêmes intérêts, mais qui se trouvent dans une autre institution (parfois distantes de plusieurs fuseaux horaires). Pourtant, plusieurs études indiquent que ces collaborations sont souvent ponctuées de rencontres face à face destinées à convenir ensemble des questions et des procédures de recherche et à permettre les ajustements à mi-parcours. Les données longitudinales et les expériences sur le terrain restent nécessaires si on veut vérifier le lien entre la communication par ordinateur et les nouveaux mécanismes de collaboration en excluant les effets des autres changements dans les problèmes scientifiques et les conditions institutionnelles (comme les politiques de financement), afin de déterminer quelles ont été les incidences réelles.

La question de l'impact sur la hiérarchie des statuts en science est beaucoup plus controversée. Certaines études ont conclu que l'Internet ouvrait des opportunités nouvelles (en assurant l'accès aux données, aux collègues, aux résultats, etc.) pour ceux qui se trouvent au début de leur carrière ou dans des institutions assez périphériques (soit par la distance, soit par le statut). D'autres ont conclu qu'il n'y avait guère de preuve de l'existence d'avantages distincts pour les scientifiques en situation excentrique. Il semble que le principal effet ait été l'accroissement du nombre de chercheurs et d'institutions qui peuvent participer au dialogue dans une communauté scientifique. Il reste à déterminer si cette participation consiste essentiellement à assister passivement à la discussion, ou à apporter des contributions significatives qui, sinon, auraient pu être perdues.

Enfin, certains éléments indiquent que la communication par ordinateur améliore la productivité scientifique. Plusieurs études ont constaté l'existence de corrélations entre le recours à la communication par ordinateur et les articles publiés ou d'autres critères de productivité similaires. En outre, des études constatent que les personnes interrogées attribuent leurs augmentations de productivité aux gains d'efficacité que permet la communication par ordinateur. Il n'y a toutefois guère eu d'analyse longitudinale qui contribuerait à démêler le sens de la causalité en la matière. La question connexe, et plus fondamentale, de l'incidence de la communication par ordinateur sur le rythme des progrès de la science demeure en grande partie sans réponse. On ne dispose pas, pour l'instant, de mesures valables ni de données empiriques démontrant que l'extension prise par la communication par

ordinateur a entraîné une réduction du temps nécessaire pour faire de nouvelles découvertes scientifiques.

Il existe plusieurs domaines où de nouvelles recherches demeurent nécessaires. Outre les questions citées ci-dessus, il y a aussi celles des effets comparés de l'Internet entre pays et entre disciplines et de ses effets sur la science à l'échelle mondiale. On observe d'importants écarts, d'une discipline à l'autre, dans le recours à la communication par ordinateur. La recherche doit être poursuivie si on veut savoir dans quelle mesure ces écarts peuvent affecter la progression de la science dans différents domaines. Tout aussi importante est la nécessité de travaux empiriques de qualité concernant les utilisations et effets comparés de l'Internet entre pays. Bien que l'Internet permette en principe de surmonter de nombreux obstacles touchant les ressources et la géographie, on constatera peut-être que les avantages tirés de l'Internet dépendent de l'accès à des ressources locales qui ne sont pas disponibles de manière uniforme. Il n'existe pas encore de réponse claire à la question de savoir si l'Internet réduit les écarts de productivité scientifique et d'accès aux avantages de la science entre pays, ou s'il exacerbe ces différences.

NOTES

11. L'auteur souhaite remercier Michael Lynch pour avoir suggéré cet aspect de la communication par ordinateur.
12. Voir http://dpls.dacc.wisc.edu/apdu/lis_chart.html.

BIBLIOGRAPHIE

- ABBOTT, T. (éd.) (1994),
Internet World's on Internet, Mecklermedia, Westport.
- ABELS, E.G., P. LIEBSCHER et D.W. DENMAN (1996),
 «Factors that Influence the Use of Electronic Networks by Science and Engineering Faculty at Small Institutions. Part 1. Queries», *J. Amer. Soc. Inform. Sci.*, 47(2), pp. 146-158.
- ALLEN, T.J. (1977),
Managing the Flow of Technology, MIT Press, Cambridge, MA.
- ALLISON, P.D. et J.A. STEWART (1974),
 «Productivity Differences Among Scientists», *Am. Socio. Rev.*, 39, pp. 596-606.
- BARLEY, S.R. (1990),
 «The Alignment of Technology and Structure Through Roles and Networks», *Admin. Sci. Q.*, 35, pp. 61-103.
- BIERLY, E.W. (1988),
 «The World Climate Program: Collaboration and Communication on a Global Scale», *Ann. Amer. Acad. Polit. Soc. Sci.*, 495, pp. 106-116.
- BISHOP, A.P. (1994),
 «The Role of Computer Networks in Aerospace Engineering», *Libr. Trends*, 42(4), pp. 694-729.
- BRUCE, H. (1994),
 «Internet Services and Academic Work: An Australian Perspective», *Internet Res.*, 4(2), pp. 24-34.
- BULLEN, C. et J. BENNETT (1991),
 «Groupware in Practice», dans C. Dunlop et R. Kling (eds.). *Computerization and Controversy*, Academic, Boston, pp. 357-387.
- CARLEY, K. et K. WENDT (1991),
 «Electronic Mail and Scientific Communication: A Study of the Soar Extended Research Group», *Knowledge: Creation, Diffusion, Utilization*, 12, pp. 406-440.
- COHEN, J.A. (1995),
Computer Mediated Communication and Publication Productivity among Faculty in Association of Jesuit Colleges and Universities (AJCU) Institutions, thèse de doctorat, State University of New York, Buffalo.
- COLE, J.R. et S. COLE (1972),
 «The Ortega Hypothesis», *Science* 178, pp. 368-37.
- COLE, S. et J.R. COLE (1967),
 «Scientific Output and Recognition», *Am. Socio. Rev.*, 32, pp. 377-390.

- CONSTANT, D., L. SPROULL et S. KIESLER (1996),
«The Kindness of Strangers: On the Usefulness of Weak Ties for Technical Advice», *Organization Science*, 7(2), pp. 119-135.
- DONGARRA, J.J. et E. GROSSE (1987),
«Distribution of Mathematical Software via Electronic Mail», *Commun. ACM*, 30(5), pp. 403-407.
- DUBROVSKY, V.J., S. KIESLER et B. N. SETHNA (1991),
«The Equalization Phenomenon», *Human-Computer Interaction* 6, pp. 119-146.
- ECHAVARRIA, T., W.B. MITCHELL et K.L. NEWSOME (1995),
«Encouraging Research Through Electronic Mentoring: A Case Study», *Coll. & Res. Lib.*, 56, pp. 352-361.
- EVERHART, D. (1996),
«The Labyrinth: An Electronic Information Network for Medieval Studies», dans T.M. Harrison et T. Stephen (eds.), *Computer Networking and Scholarly Communication in the Twenty-First-Century University*, State University of New York Press, Albany, pp. 191-200.
- FELDMAN, M.S. (1987),
«Electronic Mail and Weak Ties in Organizations», *Office Tech. People*, 3, pp. 83-101.
- FINHOLT, T.A. (1993),
Outsiders on the Inside: Sharing Information through a Computer Archive (thèse non publiée), Department of Social and Decision Science, Carnegie-Mellon University.
- FINHOLT, T., S. LEWIS et W. MOTT (1995),
«Distance Learning in the Upper Atmospheric Research Collaboratory [UARC]», North Central Sociological Association Annual Meeting, Pittsburgh, PA.
- FINHOLT, T. et G. OLSON (1996),
From Laboratories to Collaboratories: A New Organizational Form for Scientific Collaboration, University of Michigan, Department of Psychology, document de travail (disponibles en s'adressant à finholt@umich.edu).
- FINHOLT, T. et L. SPROULL (1990),
«Electronic Groups at Work», *Org. Sci.*, 1, pp. 41-64.
- GARRISON, A., P. WALKER et L. TERHAAR (1997),
«ULS Asks Scholars: Do You Need Us?», *Coll. & Res. Lib. News*, 58, pp. 544-545.
- GARVEY, W.D. (1979),
Communication: The Essence of Science, Pergamon, New York.
- GARVEY, W.D. et B.C. GRIFFITH (1979),
«Scientific Communication as a Social System», dans W.D. Garvey (éd.), *Communication: The Essence of Science*, Pergamon, New York, pp. 148-164.
- GURBAXANI, V. (1990),
«Diffusion in Computing Networks: The Case of BITNET», *Commun. ACM*, 33(12), pp. 65-75.
- HAILMAN, J.P. (1996),
«Computer Networking in Ornithology», dans T.M. Harrison et T. Stephen (éds.), *Computer Networking and Scholarly Communication in the Twenty-First-Century University*, State University of New York Press, Albany, NY, pp. 167-175.
- HARMON, A. (1999),
«The Rebel Code», *NY Times*, le 21 février.

- HESSE, B.W., L.S. SPROULI, S. KIESLER et J. P. WALSH (1993),
«Returns to Science», *Commun. ACM*, 36(8), pp. 90-101.
- HILTZ, S.R. (1988),
«Productivity Enhancement from Computer-mediated Communication: A Systems Contingency Approach», *Commun. ACM*, 31(12), pp. 1438-1454.
- HILTZ, S.R. et K. JOHNSON (1989),
«Measuring Acceptance of Computer-mediated Communication Systems», *J. Amer. Soc. Inform. Sci.*, 40(6), pp. 386-397.
- HUFF, C., L. SPROULL et S. KIESLER (1989),
«Computer Communication and Organizational Commitment: Tracing the Relationship in a City Government», *J. App. Soc. Psychol.*, 19, pp. 1371-1391.
- HURD, J.M. et A.C. WELLER (1997),
«From Print to Electronic: The Adoption of Information Technology by Academic Chemists», *Science & Technology Libraries*, 16, pp. 147-170.
- IBRAHIM, Y.M. (1997),
«Finland: An Unlikely Home Base for Universal Use of Technology», *NY Times*, p. 1, le 20 janvier.
- JACOBSON, T.L. (1994),
«The Electronic Publishing Revolution is Not 'Global'», *J. Amer. Soc. Inform. Sci.*, 45(10), pp. 745-752.
- KAMINER, N. et Y.M. BRAUNSTIEN (1998),
«Bibliometric Analysis of the Impact of Internet Use on Scholarly Productivity», *J. Amer. Soc. Inform. Sci.*, 49, pp. 720-730.
- KLING, R. (1995),
«Electronic Publishing and Changes in the Systems of Scholarly Communication», *Connecting and Collaborating: Issues for the Sciences* (NSF Workshop), San Diego, CA.
- KNORR CETINA, K. (à paraître),
Epistemic Cultures: How Scientists Make Sense.
- KOVACS, D., K. ROBINSON et J. DIXON (1995),
«Scholarly E-Conferences on the Academic Networks: How Library and Information Science Professionals Use Them», *J. Amer. Soc. Inform. Sci.*, 46(4), pp. 244-253.
- KRAUT, R.E., C. EGIDO et J. GALEGHER (1990),
«Patterns of Contact and Communication in Scientific Research Collaboration», dans J. Galegher, R.E. Kraut et C. Egido (éds.), *Intellectual Teamwork*, Hillsdale, Lawrence Earlbaum Associates, NJ, pp. 149-171.
- LEDERBERG, L. et K. UNCAPHER (1989),
Towards a National Collaboratory: Report of an Invitational Workshop at the Rockefeller University, National Science Foundation, Directorate for Computer and Information Science, Washington, DC.
- MALTZ, M., A. GORDON et W. FRIEDMAN (1991),
Mapping Crime in Its Community Setting, Springer-Verlag, New York.
- McGUIRE, T., S. KIESLER et J. SIEGEL (1987),
«Group and Computer-mediated Discussion Effects in Risk Decision Making», *J. Pers. Soc. Psychol.*, 52, pp. 917-930.
- MERTON, R.K. (1968), «The Matthew Effect in Science», *Science* 159, pp. 56-63.

- MERZ, M. (à paraître),
«Formen der Internetnutzung in der Wissenschaft» («Forms of Internet Use in Science»), dans R. Werle et C. Lang (éds.), *Modell Internet? Entwicklungsperspektiven neuer Kommunikationsnetze*, Campus, Frankfurt/Maine.
- NATIONAL SCIENCE BOARD (1993),
Science and Engineering Indicators-1993, GPO, Washington, DC.
- NATIONAL SCIENCE BOARD (1998),
Science and Engineering Indicators-1993, GPO, Washington, DC.
- ORLIKOWSKI, W.J. (1993),
«Learning From Notes: Organizational Issues in Groupware Implementation», *Info. Soc.*, 9, pp. 237-250.
- ORLIKOWSKI, W.J. et J. YATES (1994),
«Genre Repertoire: The Structuring of Communicative Practices in Organizations», *Admin. Sci. Q.*, 39, pp. 541-574.
- OSTBYE, T. et T. WELBY (1988),
«Electronic Communication with Colleagues Overseas», *New Zeal. Med. J.*, 101, pp. 145-147.
- PHILIP, G. (1995),
««Use of 'Leading-Edge' Information Systems by Academic Chemists in the UK.» Part 1. The Results of a Preliminary Investigation», *J. Inf. Sci.*, 21, pp. 187-199.
- RAINWATER, L. et T.M. SMEEDING (1988),
«The Luxembourg Income Study: The Use of International Telecommunications in Comparative Social Research», *Ann. Amer. Acad. Polit. Soc. Sci.*, 495, pp. 95-105.
- RICE, R.E. (1980),
«Impact of Computer-mediated Organizational and Interpersonal Communication», *Annu. Rev. Inf. Sci. Technol.*, 15, pp. 221-249.
- RICE, R.E. (1994),
«Network Analysis and Computer-mediated Communication Systems», dans S. Wasserman et J. Galaskiewicz (eds.), *Advances in Social Network Analysis*, Sage, CA, pp. 167-203.
- RUHLEDER, K. (1994),
«Rich and Lean Presentations of Information for Knowledge Workers», *ACM Trans. Off. Inf. Syst.*, 12, pp. 208-230.
- SANDERSON, D. (1996),
«Co-operative and Collaborative Mediated Research», dans T.M. Harrison et T. Stephen (éds.), *Computer Networking and Scholarly Communication in the Twenty-First-Century University*, State University of New York Press, Albany, pp. 95-114.
- SCHAUDER, D. (1994),
«Electronic Publishing of Professional Articles: Attitudes of Academics and Implications for the Scholarly Communication Industry», *J. Amer. Soc. Inform. Sci.*, 45(2), pp. 73-100.
- SCHWARTZ, C.A. (1994),
«The Strength of Weak Ties in Electronic Development of the Scholarly Communication System», *Coll. & Res. Lib.*, 55, pp. 529-540.
- SIEGEL, J., V. DUBROVSKY, S. KIESLER et T. McGUIRE (1986),
«Group Processes in Computer-mediated Communication», *Organ. Behav. Hum. Decis. Process.*, 37, pp. 157-187.

- SPROULL, L. et S. KIESLER (1986),
 «Reducing Social Context Cues», *Manag. Sci.*, 32, pp. 1492-1512.
- SPROULL, L. et S. KIESLER (1991),
Connections, MIT, Cambridge, MA.
- SUDWEEKS, F. et S. RAFAELI (1996),
 «How Do You Get a Hundred Strangers to Agree? Computer-mediated Communication and Collaboration», dans T.M. Harrison et T. Stephen (éds.), *Computer Networking and Scholarly Communication in the Twenty-First-Century University*, State University of New York Press, Albany, pp. 115-136.
- TAUBES, G. (1993),
 «Publication by Electronic Mail Takes Physics by Storm», *Science*, 259, 26 février.
- UZUMERI, M.V. et C.A. SNYDER (1996),
 «Information Technology and Accelerated Science: The Case of the Pentium TM Flaw», *Calif. Manag. R.*, 38(2), pp. 44-63.
- VAN ALSTYNE, M. et E. BRYNJOLFSSON (1996a),
 «Communication Networks and the Rise of an Information Elite», MIT Sloan School of Management, <http://web.mit.edu/marshall/www/home.html>.
- VAN ALSTYNE, M. et E. BRYNJOLFSSON (1996b),
 «Could the Internet Balkanize Science?», *Science*, 274, pp. 1479-1480 (29 novembre).
- WALSH, J.P. et T. BAYMA (1996a),
 «Computer Networks and Scientific Work», *Soc. Stud. Sci.*, 26 (3), pp. 661-703.
- WALSH, J.P. et T. BAYMA (1996b),
 «The Virtual College: Computer-mediated Communication and Scientific Work», *Info. Soc.*, 12, pp. 343-363.
- WALSH, J.P., S. KUCKER, N. MALONEY et S. GABBAY (à paraître),
 «Connecting Minds: CMC and Scientific Work», *J. Amer. Soc. Inform. Sci.*
- WALTHER, J.B. (1996),
 «Computer-mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction», *Comm. Res.*, 23, pp. 3-43.
- WALTHER, J.B. (1997),
 «Group and Interpersonal Effects in International Computer-mediated Collaboration», *Hum. Commun. Res.*, 23, pp. 342-369.
- WEISBAND, S.P. (1992),
 «Group Discussion and 1st Advocacy Effects in Computer-mediated and Face-to-face Decision-making Groups», *Organ. Behav. Hum. Decis. Process.*, 53, pp. 352-380.
- WEISBAND, S.P., S.K. SCHNEIDER et T. CONNOLLY (1995),
 «Computer-mediated Communication and Social Information: Status Salience and Status Differences», *Acad. Manag. J.*, 38, pp. 1124-1151.

LA PUBLICATION ÉLECTRONIQUE

Table des matières

I. Introduction	92
II. Science et information	92
III. Tendances de la publication électronique	94
IV. Perspectives et défis	103
V. Conclusion	109

Cet article a été rédigé par Karel Leeftang, Managing Director, Elsevier Science, Irlande et Jim Gilmore, Director General, CAB INTERNATIONAL, Royaume-Uni.

I. INTRODUCTION

L'information et les processus et systèmes servant à la gérer et à la diffuser ont une importance inhérente essentielle pour la science. Le processus scientifique repose sur l'utilisation des informations connues et il livre à son tour de nouvelles informations à diverses collectivités partout dans le monde. L'infrastructure générale de la science doit donc fournir les systèmes et outils permettant d'acquérir, de produire et de diffuser l'information, de manière efficace et pérenne. De plus, ces systèmes et outils doivent garantir le respect de normes de qualité et la protection des droits de propriété intellectuelle.

Bien sûr, la publication électronique n'est pas un phénomène nouveau : on utilise les ordinateurs comme outil de production dans le secteur de l'édition depuis les années 60, et on offre des produits dérivés sur des systèmes en ligne depuis près de 30 ans. Les principaux changements résultent de l'accessibilité et de la puissance croissantes des ordinateurs, notamment des ordinateurs personnels, en particulier ces dix dernières années. Du côté de la production, cela touche les éditeurs en leur permettant de réduire les coûts de publication et de créer une large gamme de nouveaux produits électroniques. L'éventail et les capacités des médias de transfert ont aussi augmenté, offrant de nouvelles possibilités de livrer des ressources informationnelles directement sur le bureau de l'utilisateur. On a vu le passage de la disquette au CD-ROM et au transfert par l'Internet, et bientôt au puissant support nouveau que constitue le DVD (*Digital Versatile Disk*).

Ces technologies ont non seulement élargi les canaux de livraison à la disposition des éditeurs, mais elles encouragent aussi l'apparition de petits éditeurs concurrents et permettent aux utilisateurs d'exploiter et de maîtriser leurs propres processus de publication. Ces tendances et ces nouveautés ont un impact dans tous les domaines, et particulièrement pour la science.

II. SCIENCE ET INFORMATION

Les systèmes de transfert et de diffusion de l'information dans la science se sont créés et ont évolué au cours des siècles. Ils continuent de reposer sur une combinaison de systèmes « formels » et « informels » : les systèmes informels comprennent des types variés de communications directes ou indirectes entre les

scientifiques, alors que les systèmes formels reposent sur des mécanismes de publication bien établis, économiquement viables et de qualité assurée. Dans l'évolution actuelle liée à la croissance des ressources de technologies de l'information disponibles, les systèmes informels (et l'information qu'ils fournissent) peuvent proliférer et les systèmes formels doivent s'adapter à un nouvel ensemble de besoins et d'attentes des scientifiques.

Avant de voir en détail les changements engendrés par les technologies de l'information, et leurs implications, il convient de souligner l'importance intrinsèque de l'information pour la science et d'examiner les systèmes utilisés actuellement ou dans le passé pour gérer l'information qui sert d'ingrédient au processus de la R-D et l'information qui est un produit de la recherche.

Quels que soient les mécanismes de livraison et les systèmes utilisés, la qualité, la cohérence, la fiabilité et la viabilité de ces systèmes d'information sont des besoins essentiels. Les systèmes et médias de livraison de l'information ont parcouru une longue évolution pour aboutir à un environnement de stabilité. Il faudra rétablir un tel environnement dans le nouveau monde des technologies de l'information si l'on veut conserver ces qualités.

Comme on l'a noté, il existe deux grands types de systèmes. La première catégorie est celle des systèmes informels, qui comportent diverses méthodes de communication entre les scientifiques. Comme pour la R-D en général, ce type de processus de communication est principalement financé par les pouvoirs publics et, dans une grande mesure, les coûts, «cachés» dans les budgets, ne sont pas explicitement quantifiés.

Alors qu'ils représentent un important élément du processus de transfert de l'information, les systèmes informels sont sensibles aux réductions de financement et aux discontinuités dues au changement du personnel scientifique. Ils comportent souvent le risque d'un amoindrissement ou de la perte de droits de propriété intellectuelle et ils sont axés sur une information informelle et non soumise à un contrôle.

La deuxième catégorie de systèmes de transfert d'information est celle des systèmes que l'on peut appeler «formels», comprenant de nombreux types d'œuvres publiées. Ces œuvres sont produites, utilisées et entretenues à la fois par le secteur public et le secteur privé, et elles requièrent des dépenses directes substantielles dans les budgets de la recherche. Ces systèmes sont généralement viables, en particulier quand ils ont une base commerciale, bien qu'ils soient aussi sensibles aux variations du financement public de la science et aux environnements économiques. On peut remarquer, par exemple, que beaucoup d'éditeurs internationaux ont été fortement touchés par la crise économique en Asie du Sud-Est.

Un point important à noter est que les systèmes formels de livraison de l'information établis offrent une qualité vérifiable grâce à leurs processus d'examen par les collègues de la spécialité et qu'ils assurent une protection des droits de propriété intellectuelle par les conventions de publication auxquelles s'ajoute la législation nationale.

La conséquence majeure que les progrès des technologies de l'information peuvent avoir sur cette infrastructure est l'éventualité d'une croissance explosive des systèmes informels, peut-être au détriment des systèmes formels. Les défauts des systèmes informels verraient alors leurs effets s'amplifier et les problèmes de qualité, de viabilité et autres aspects importants pourraient se multiplier. Il faut prendre conscience de ce risque et le gérer, de manière à ce que les technologies de l'information produisent des effets véritablement bénéfiques et qu'une nouvelle infrastructure, durable et de valeur, s'établisse.

III. TENDANCES DE LA PUBLICATION ÉLECTRONIQUE

Le facteur clé est l'accessibilité générale des ordinateurs personnels et équipements connexes dans le monde. D'un point de vue planétaire, ce facteur contribue puissamment à réduire le fossé technologique entre le monde développé et le monde en développement. Les nouvelles technologies accroissent la capacité de stocker, traiter et distribuer l'information et les données. Elles recèlent un énorme potentiel d'effets bénéfiques mais elles peuvent conduire à une prolifération d'informations et de données invalides et elles posent de sérieux problèmes de viabilité.

Le deuxième grand facteur est l'émergence de l'Internet comme moyen omniprésent de communication et de livraison de l'information. Ce réseau recèle lui aussi un énorme potentiel bénéfique, avec de nouvelles capacités de traitement puissantes, l'accessibilité immédiate de l'information à l'échelle internationale, l'accès à des ressources et médias nouveaux, et des méthodes avancées de gestion et d'obtention de l'information. Cependant, les risques sont beaucoup plus grands de favoriser la distribution d'informations de mauvaise qualité, et l'Internet pose les mêmes problèmes de viabilité, avec également le risque d'une dépendance à l'égard de réseaux surchargés, l'inégalité de l'accessibilité internationale et certaines préoccupations sérieuses concernant l'économie sous-jacente de son utilisation.

Nous examinerons et illustrerons les tendances et l'étendue de la publication électronique moderne sous quatre aspects principaux :

- bases de données secondaires ;
- livraison électronique de publications primaires ;
- bases de données et banques de données factuelles ; et
- lier et intégrer les systèmes.

Bases de données secondaires

Les services secondaires, notamment les résumés et les index, constituent depuis plus de cent ans le premier moyen d'accès sélectif aux publications scientifiques. Ils comprennent de grands services comme Chemical Abstracts, INSPEC (physique et électronique), Excerpta Medica et MEDLINE (médecine), CAB ABSTRACTS (agriculture et sciences de la vie appliquées) et Biological Abstracts, ainsi que des services nombreux et variés dans d'autres domaines des sciences naturelles et sociales. Au cours de toute cette période, on a utilisé ces services sous la forme « traditionnelle » des publications imprimées et, plus récemment, sous forme électronique avec les services en ligne, les CD-ROM et maintenant la livraison par l'Internet.

Les scientifiques du monde entier ont l'habitude et sont dépendants de ces services qui donnent un moyen sélectif d'accès au vaste corpus de la science publiée, et ces services conservent un rôle important dans le processus de transfert de l'information, même si les publications primaires sont de plus en plus disponibles sous forme électronique. Les services secondaires sont offerts par un ensemble d'organisations privées, sans but lucratif ou commerciales, avec des systèmes et modèles de tarification bien établis. Ils ont commencé sous la forme de services par abonnement basés sur des revues imprimées, qui en fait continuent d'être utilisées, et ont progressivement évolué vers un modèle d'utilisation ponctuelle et donc une tarification à la transaction sur des systèmes en ligne, y compris l'Internet.

Pour la plupart, ces services sont tenus d'être au moins autosuffisants financièrement et, dans certains cas, de faire des bénéfices. Ils reposent donc nécessairement sur de bons modèles et principes économiques et ils réussissent (ou échouent) dans la mesure où ils fournissent des services à valeur ajoutée aux utilisateurs et sont ainsi viables. C'est généralement un secteur changeant, où les grandes bases de données jouissent d'une stabilité et d'une viabilité durables alors que des services plus petits et plus spécialisés ont tendance à naître et finalement disparaître faute d'utilisateurs ou du fait de la perte d'une subvention.

La question clé pour l'avenir, encore sans réponse, est de savoir si ces services conserveront un rôle face à la concurrence, que l'on entrevoit, des publications primaires directement livrées par voie électronique. Nous pensons qu'ils garderont un intérêt et seront en fait nécessaires comme « outil de recherche » des informations ou comme une partie d'un mécanisme d'accès intelligent vers la pléthore des différents médias et mécanismes de livraison servant à fournir l'information primaire.

La livraison électronique des publications primaires

Elsevier Science est le plus grand éditeur scientifique du monde, avec plus de 1 100 titres. Certains de ces titres paraissent presque quotidiennement, beaucoup

d'autres sont hebdomadaires. Elsevier Science fait partie de la société Reed Elsevier, un des dix plus grands éditeurs du monde avec une importante activité dans l'information scientifique, professionnelle et commerciale.

Reed Elsevier est propriétaire de quelques grandes bases de données, comme Nexis Lexis pour l'information d'affaires et l'information juridique et EMBASE qui fournit des informations bibliographiques dans le domaine biomédical. Reed Elsevier et ses compagnies ont toujours eu pour politique de se concentrer sur le contenu en laissant d'autres parties s'occuper de la distribution mais, pensant que les distributeurs tiers ne répondent pas à certains besoins du marché, cette entreprise s'est maintenant directement engagée dans la mise au point de ses propres moteurs de recherche. Cela concerne principalement Lexis et Butterworth pour l'information juridique et Science Direct pour l'information scientifique. La part importante du marché interentreprises de l'information que détient Reed Elsevier a l'avantage de lui permettre de créer des moteurs de recherche pour des catégories d'utilisateurs variées, scientifiques, professionnelles ou commerciales.

En même temps, Reed Elsevier travaille en liaison étroite avec des réalisateurs de logiciel ; par exemple, cette compagnie a conclu un contrat d'alliance stratégique de cinq ans avec Microsoft et divers consortiums de bibliothèques, établissements universitaires et autres fournisseurs de contenu et intermédiaires. En fin de compte, c'est toujours l'utilisateur qui décidera du véhicule approprié pour obtenir l'information la meilleure, la plus complète et la plus récente. Le fournisseur de contenu est le partenaire le mieux à même d'organiser et de mettre à disposition cette information.

Avec les progrès rapides des technologies de l'information et la possibilité technique et économique de livrer électroniquement l'information jusqu'au poste de travail du chercheur, les éditeurs ont commencé à fournir électroniquement l'information scientifique. Au moyen de la même technologie que pour les bases de données secondaires décrites ci-dessus, on fournit au marché, depuis déjà un certain temps, des informations primaires par la voie électronique. Bien que d'autres éditeurs, consortiums de bibliothèques ou distributeurs offrent sur le marché un certain nombre d'activités et produits, nous prendrons les activités de livraison électronique d'informations primaires d'Elsevier Science comme exemple pour illustrer l'évolution générale du marché de l'information scientifique, technique et médicale (STM).

Elsevier a commencé à expérimenter la livraison en ligne du texte complet de revues scientifiques au début de la décennie 80, en constituant une base de données de quelque 25 000 pages et en offrant l'accès à une trentaine de bibliothèques d'Europe et des États-Unis. Ces expériences ont réussi techniquement mais elles étaient beaucoup trop en avance sur leur temps. Les enquêtes auprès des

utilisateurs montrèrent un intérêt à l'égard de l'accès à ces informations mais il était évident qu'il n'y avait pas de marché et la livraison à 300 bauds sur les lignes téléphoniques ordinaires était inefficace. Cette tentative fut donc suspendue en attendant que l'infrastructure et le marché se développent.

Le projet expérimental suivant pour la livraison électronique de texte complet commença en 1990, avec l'association d'Elsevier Science et d'un groupe d'universités américaines renommées dans un programme intensif d'expérimentations empiriques d'une durée de trois ans appelé Tulip (*The University Licensing Program*), visant à étudier du point de vue technique et économique la livraison de versions électroniques des titres d'Elsevier Science vers un environnement de bibliothèque numérique. Ce service fut inauguré en 1991 avec la livraison d'information bibliographique et de résumés dans une des principales universités des Pays-Bas, suivis en 1994 par des fichiers électroniques de texte complet pour toutes les revues d'Elsevier auxquelles la bibliothèque était abonnée. C'était une avancée logique, l'infrastructure étant en place et l'accès au texte complet étant considéré comme un élément permettant d'améliorer notre service à l'utilisateur final. L'expérience Tulip apporta à toutes les parties une grande quantité de renseignements. Techniquement, elle était beaucoup plus compliquée aux deux bouts de la chaîne, du côté du fournisseur et du côté récepteur, et économiquement elle nécessitait un investissement notable à ces deux extrémités.

Ces expériences ont conduit au lancement d'*Elsevier Electronic Subscriptions* (EES) en 1995 qui, après des débuts assez lents, rencontre maintenant un succès croissant. Des bibliothèques universitaires, professionnelles, nationales et administratives sont abonnées à ce système par licences. EES est livré en quatre bases de données : fichier de table des matières, fichiers SGML limités pour la recherche par champ et les hyperliens, fichiers plats contenant le simple texte ASCII de toutes les pages et de tous les articles de chaque numéro, et fichiers d'articles contenant le texte complet et les graphiques de tous les articles publiés dans les revues d'Elsevier Science.

Depuis 1993, Elsevier Science a aussi mis sur le marché des revues électroniques séparées : *Nuclear Physics Electronic*, *Immunology Today Online*, *Crema-Combis* relié au Laboratoire européen de biologie moléculaire (EMBL), *New Astronomy* et beaucoup d'autres.

En 1997, Elsevier Science a lancé Science Direct, plate-forme de publication basée sur le Web qui offre aux bibliothèques et à leurs utilisateurs finals l'accès, sur leur table de bureau, au texte complet de toutes les revues d'Elsevier Science stocké à distance ainsi qu'au texte complet d'autres éditeurs participants. Des liens donnent facilement accès à des services de livraison de documents, à des bases de données commerciales, des ouvrages de référence, des bulletins d'information, des manuels et des passerelles vers d'autres types et sources d'information. Ce

produit ou programme répond pleinement aux besoins du monde de la recherche – guichet unique, accès à toute l'information pertinente correspondant aux besoins spécifiques du chercheur et résolution de problème – améliorant ainsi l'efficacité et, en conséquence, la situation économique du monde de la recherche.

Elsevier Science n'est pas le seul éditeur scientifique à offrir au monde de la science ces outils d'accès électronique et de recherche d'information. D'autres programmes ou produits similaires sont disponibles sur le marché, certains aussi élaborés mais d'autres moins. Un certain nombre de ces expériences n'ont pas survécu et, en l'absence d'une masse critique et d'un financement suffisants, il n'est pas possible de servir durablement le marché. En outre, comme les collectivités d'utilisateurs doivent aussi fortement investir dans l'infrastructure, le matériel et le logiciel, il est essentiel pour elles de choisir le bon partenariat avec leurs fournisseurs de contenu. Bien que les éditeurs commerciaux ne soient pas toujours populaires auprès des bibliothèques, il est juste de dire que les initiatives et l'évolution dues au secteur privé concernant la livraison électronique des informations primaires répondent effectivement aux besoins des collectivités d'utilisateurs. Il est évidemment difficile pour des fournisseurs de contenu sans but lucratif et non subventionnés d'assumer ces énormes investissements dans des moyens perfectionnés de production de manuscrits assistée par ordinateur, avec la conversion en fichiers structurés SGML ou HTML et des liens vers d'autres sources d'information. Il leur manque aussi la masse critique justifiant ces investissements et les recettes tirées des produits électroniques sont encore très faibles.

La plupart des prévisions, par exemple de l'Union européenne, sont très optimistes mais, en réalité, les recettes des produits électroniques pour la plupart des éditeurs STM classiques représentent moins de 15 % de leur chiffre d'affaires total. Ces recettes augmentent mais elles sont très insuffisantes pour financer les investissements nécessaires.

Il y a aussi le changement du modèle économique. Les bibliothèques ont l'habitude de payer des abonnements et sont chargées de mettre l'information à la disposition de leurs utilisateurs finals. Elles ont bâti leurs systèmes de budgétisation et de financement existants sur cette stratégie de prépaiement, comme l'ont fait aussi les éditeurs, mais la livraison de fichiers électroniques de texte complet, directement ou indirectement, au poste de travail de l'utilisateur final implique la mise au point de nouveaux systèmes de paiement et de budgétisation.

Cependant, la plupart des éditeurs offrent actuellement les contenus à leurs clients institutionnels sur la base d'un contrat de licence. En liaison avec les abonnements aux revues imprimées, des fichiers électroniques sont livrés contre un paiement additionnel, avec certaines conditions d'utilisation. De cette manière, la bibliothèque est capable d'établir son budget et peut mieux maîtriser ses coûts dans la livraison des revues et documents à l'intérieur de l'environnement institutionnel.

Des fichiers électroniques séparés, non liés à l'abonnement à une revue imprimée, peuvent être inclus dans la licence.

A l'avenir, les deux parties, bibliothèques et éditeurs, continueront de fonctionner sur la base de l'abonnement pour l'obtention d'informations liées à l'utilisation d'articles particuliers. Au moyen de logiciels perfectionnés, l'utilisateur aura accès à l'information verticale, à la revue et à l'information horizontale dans son domaine de recherche.

Bases de données factuelles et banques de données

L'avènement de l'Internet et les systèmes d'information électronique perfectionnés stimulent le développement d'une nouvelle gamme de bases de données. Il y a dix ans, si l'on mentionnait une base de données scientifique, il s'agissait probablement d'une base de données secondaire ou base de résumés qui permettait à l'utilisateur de parcourir des milliers de résumés de revue pour trouver l'article primaire recherché. Cependant, des domaines comme la biotechnologie ont connu une croissance presque sans précédent de « bases de données factuelles » telles que la *Nucleotide Sequence Database* de l'EMBL qui stocke des séquences de nucléotides annotées (base de données de séquences de gènes ou d'ADN).

Les bases de données sont des outils essentiels de la recherche moderne. On les manipule et on les analyse, et on en extrait des données pour leur ajouter une valeur. Ainsi, nous avons déjà des « bases de données construites à partir de bases de données » qui « traduisent les séquences de gènes en séquences d'acides aminés », et d'autres qui permettent à l'utilisateur de replier ces séquences d'acides aminés en modèles de protéine. Nous commençons maintenant à traduire la séquence brute d'information en protéines fonctionnelles.

Ces bases de données – petites ou grandes collections de données – sont habituellement assemblées par des scientifiques pour leurs propres besoins et ceux de leurs collègues. Des centaines de bases de données sont en construction, pour la plupart dans le domaine public, et les meilleures et les plus utilisées sont offertes au public par le biais de centres comme l'EMBL Outstation, l'*European Bioinformatics Institute* à Hinxton, Cambridge, ou les nombreux nœuds du réseau EMBnet (*European Molecular Biology Network*) dispersés dans toute l'Europe.

Les chiffres officiels de ces centres montrent que ces nouvelles banques de données prennent au moins trois à cinq ans pour atteindre une masse critique suffisante et nécessitent trois à quatre personnes à plein temps pour travailler sur la base et valider l'information, ainsi que des mécanismes de distribution permettant de rendre l'information accessible et utile pour le monde scientifique. Cependant, près de 50 % de ces banques de données disparaissent avant d'atteindre leurs cinq ans d'existence. Actuellement, celles qui survivent reçoivent des subventions

permanentes d'organismes nationaux ou supranationaux. Très peu (environ 1 %) comptent sur des ressources privées pour leur avenir.

Ce taux d'échec est inquiétant ; cela signifie, pour le moins, qu'on a gaspillé de l'argent et des efforts dans des projets sans avenir, mais cela peut aussi signifier que des données importantes ne sont plus disponibles parce que la base de données a été établie dans un cadre incertain. Résoudre ce problème est évidemment une tâche importante pour les autorités responsables du financement public, mais la croissance de ce secteur suscite des préoccupations. Y a-t-il assez de fonds publics pour entretenir ces bases de données en forte expansion ? Quels financements ces nouvelles bases de données peuvent-elles espérer ? Il paraît évident qu'un meilleur système d'information et de coordination est nécessaire autour de ce domaine, indiquant en détail les banques de données disponibles dans les différents domaines scientifiques. La propriété, le financement, l'accès aux données et le chevauchement entre les bases de données sont des points que l'on pourrait mieux étudier et décrire.

La crainte de la plupart de ces bases de données est de perdre leur financement public ou de ne plus pouvoir suivre les nouvelles technologies nécessaires pour rester compétitives. Actuellement, aux États-Unis, des bases de données comme MEDLINE, AGRICOLA, GENBANK et de nombreuses bases de données de biotechnologie disposent de fonds publics. L'EMBL, et dans une grande mesure l'Union européenne, financent les bases de données de l'EBI mais SWISS-PROT, une des bases de données primordiales dans le domaine de la biotechnologie, doit maintenant trouver d'autres sources de revenus parce que les fonds publics ne permettent plus de l'entretenir totalement. Si cela arrive à une base de données aussi prestigieuse, beaucoup d'autres seront bientôt touchées.

Peut-on les privatiser ? Pas vraiment, à moins que l'utilisateur n'accepte de payer le prix d'informations de valeur et à condition qu'il n'y ait pas de concurrence de la part de banques de données similaires entièrement subventionnées dans d'autres régions du monde. La science est, par définition, une activité mondiale mais l'accès à l'information scientifique est d'importance vitale pour chaque entité nationale ou supranationale. La dépendance à l'égard de banques de données d'accessibilité restreinte ou de banques de données non viables est une sérieuse menace pour le développement de la recherche. Nos hôtes d'aujourd'hui ont un rôle important à jouer dans la modélisation des banques de données actuelles et en même temps dans la création d'un environnement où les banques de données publiques et privées aient une existence économique durable.

Les subventions aux bases de données

Beaucoup de bases de données reçoivent des subventions publiques, notamment aux États-Unis où certains produits sont devenus synonymes « d'information

gratuite». Il est clair que cela fausse le marché mais cela touche aussi l'environnement scientifique qui dépend d'une information exacte et à large base.

Bien qu'il soit toujours difficile de faire face à la concurrence de produits subventionnés, la situation était plus supportable avant que l'Internet ait eu l'effet de supprimer les frontières géographiques. Aujourd'hui, des produits auparavant considérés comme « nationaux » sont distribués ou accessibles dans le monde entier, ce qui fausse gravement les autres marchés nationaux.

MEDLINE – produit très estimé que l'on peut utiliser plus ou moins gratuitement dans le monde entier – offre un excellent exemple de ce phénomène. Il met ainsi en péril non seulement les bases de données commerciales concurrentes mais aussi les bases de données sans but lucratif qui servent aussi ce marché. En outre, en raison de son accessibilité sur le World Wide Web, beaucoup d'hôtes et distributeurs de bases de données sont menacés d'une baisse de revenus et d'utilisation. En effet, les utilisateurs ont tendance à les quitter au profit des hôtes de bases de données financés par des fonds publics et à se désintéresser des hôtes commerciaux européens (et américains) qui offraient MEDLINE, ainsi que d'autres bases, à leurs clients.

Si certains utilisateurs se félicitent des mesures du gouvernement américain, d'autres se rendent compte qu'on ne peut pas compter sur une unique base de données pour satisfaire aux besoins de la recherche en matière d'information. MEDLINE fournit des résumés indexés pour 3 000 revues ou un peu plus. Cela ne représente qu'un pourcentage modéré du nombre total des revues de recherche biomédicale. En outre, bien naturellement, cette sélection a tendance à favoriser les titres américains et les utilisateurs sont ainsi essentiellement orientés vers des sources primaires américaines alors que des revues européennes ou d'autres pays sont passées sous silence. Le soutien du gouvernement américain confère aussi une certaine autorité qui fait naître l'impression que les meilleures revues sont incluses dans MEDLINE et donc que les meilleures revues sont américaines, même si ce n'est pas la vérité.

Tout aussi important est le fait que, si toute recherche dans MEDLINE peut permettre de localiser quelques articles, elle en manquera aussi beaucoup d'autres qui figurent dans des services concurrents (par exemple, le chevauchement entre MEDLINE et EMBASE – probablement son principal concurrent commercial – peut descendre à 40 % dans certains domaines). La fausse impression de sécurité qu'offre ce genre de résultat peut conduire à des erreurs d'orientation scientifique et être coûteuse, du fait que beaucoup de chercheurs risquent de suivre des voies de recherche déjà explorées par d'autres et publiées dans des revues tout aussi respectables mais dont les résumés ne figurent pas dans MEDLINE. (On notera que les chercheurs qui doivent être sûrs d'avoir examiné toutes les publications pertinentes, par exemple les chercheurs de compagnies pharmaceutiques,

utilisent toujours en plus des produits américains subventionnés une série d'autres bases de données)*.

En outre, parce que MEDLINE est peu coûteuse et librement accessible, beaucoup d'autres bases de données plus factuelles l'utilisent de plus en plus comme « outil de liaison » (c'est-à-dire qu'elles renvoient à MEDLINE à partir de leurs données). Cela accroît la réputation de MEDLINE elle-même mais cela favorise également beaucoup d'autres bases de données à financement américain qui sont aussi reliées aux fichiers de MEDLINE. En fait, on conduit l'utilisateur vers des revues et bases de données secondaires américaines au lieu de lui laisser chercher le meilleur produit dans son domaine.

Globalement, la distribution gratuite de cette base de données pourrait en fait être préjudiciable à l'utilisateur, et le gouvernement américain subventionne aussi d'autres bases de données dans d'autres spécialités comme l'agriculture, de telle sorte que cette situation se reproduit dans d'autres disciplines scientifiques. La science internationale aurait beaucoup à perdre si les bases de données complémentaires qui sont plus ou moins concurrentes de ces produits subventionnés devaient fermer boutique. L'utilisateur n'aurait alors accès qu'à une petite partie des publications reconnues nécessaires à une bonne R-D, ce qui ne profiterait à personne. L'OCDE est un lieu où l'on pourrait débattre de cette question compliquée en vue de parvenir à un meilleur équilibre.

Lier et intégrer les systèmes : le projet ADLIB

Dans l'avenir, c'est par la façon dont on pourra lier et intégrer les systèmes, services et bases de données décrits ci-dessus que l'utilisation des technologies modernes de l'information manifesterait vraiment sa puissance et ses possibilités, de manière à fournir aux scientifiques des méthodes d'accès à l'information encore plus utiles et commodes. L'association entre l'Internet, les ordinateurs personnels puissants et les outils logiciels modernes peut permettre au chercheur, à sa table de bureau, d'accéder à une large gamme de ressources d'information et de données, d'une manière qui était jusqu'ici impossible. Ces nouveaux types de services nécessiteront de nouveaux modèles économiques et tarifaires et, ce qui est un point important, une plus grande collaboration et des partenariats entre les fournisseurs d'information. Certains de ces services commencent déjà à voir le jour et on présente ci-dessous un exemple d'une des approches adoptées.

Le projet ADLIB (*Advanced Database Linkages in Biotechnology*) est une entreprise conjointe d'une douzaine d'éditeurs et de fournisseurs de données et de services européens dans le secteur européen de la bioinformatique. Il reçoit un soutien de

* Pour un avis indépendant à ce sujet, voir par exemple : Cecilia, M. Brown (1998), « The Benefits of Searching EMBASE versus MEDLINE for Pharmaceutical Information », *Online & CD-ROM Review*, vol. 22, n° 1.

l'Union européenne en tant que projet de démonstration visant à montrer comment on peut utiliser les nouvelles technologies pour améliorer l'accès des scientifiques aux informations et données européennes en matière de biotechnologie.

Ce concept repose sur une utilisation innovatrice des technologies pour relier à travers l'Internet diverses bases de données différentes au moyen d'une interface transparente. Au niveau le plus simple, par exemple, un chercheur peut interroger une ou plusieurs bases de données bibliographiques au moyen d'une interface intelligente, trouver les références d'une séquence de gènes et accéder ensuite automatiquement à la base de données qui contient la séquence détaillée. Nous avons donc pour la première fois un outil qui peut intégrer aux travaux publiés un large ensemble de références et permettre la liaison avec des données factuelles complètes et détaillées concernant, dans ce cas, les séquences de gènes.

La puissance et les possibilités de cette approche sont évidentes : elle laisse espérer que l'on pourra donner aux scientifiques du monde entier l'accès à des outils d'analyse et de recherche d'information beaucoup plus avancés que tout ce que l'on a connu dans le passé, et ce concept est applicable à pratiquement tous les domaines scientifiques.

IV. PERSPECTIVES ET DÉFIS

Dans le contexte de la publication, le défi à relever est d'utiliser au mieux la technologie et les ressources d'information pour servir efficacement le processus de la R-D, améliorer la qualité de la science et son rapport coût-efficacité, et améliorer la diffusion des résultats.

Cela doit nécessairement s'insérer dans le contexte plus large de l'infrastructure globale de la R-D. Il incombe aux gouvernements (en particulier) de veiller à ce que tous les systèmes de soutien soient sains, viables et pérennes. Les technologies de l'information touchent tous les systèmes formels ou informels permettant d'acquérir, utiliser, produire et diffuser l'information.

Ces changements ouvrent des possibilités incontestables mais ils comportent aussi des risques pour l'ordre existant et pour cet élément de l'infrastructure fondamentale de la science qu'est la publication. On ne peut pas laisser cette évolution trouver elle-même son équilibre et ses mécanismes ni se développer comme un phénomène de marché incontrôlé. Elle soulève des questions importantes pour l'action gouvernementale, des considérations sur la gestion de la recherche, des interrogations pour les milieux scientifiques sur les principes fondamentaux des processus de la recherche et, bien sûr, beaucoup de considérations économiques.

Les perspectives, et donc les effets bénéfiques, de l'utilisation des nouvelles technologies de l'information sont claires ; elles se manifesteront par l'amélioration

et l'accélération de la livraison de l'information, par le renforcement de la coopération internationale, par les possibilités de rationalisation des programmes en donnant plus largement accès à l'information sur les autres recherches internationales et, finalement, par une avancée plus rapide de la science.

On en observe certains aspects essentiels dans les tendances et les nouveautés décrites ci-dessus, avec l'apparition de services et de mécanismes de livraison nouveaux et plus puissants. On entrevoit en particulier la possibilité d'une beaucoup plus grande intégration des services et d'une collaboration accrue entre les fournisseurs de services. Ces nouveaux services et ressources, associés à une utilisation stratégique de la technologie, permettront aux scientifiques et aux gestionnaires de la science de développer des programmes de recherche sur une base d'information et de communication plus dynamique, en reconnaissant mieux l'importance (aussi bien scientifique qu'économique) des systèmes d'information servant le processus scientifique.

Dans une certaine mesure, ces nouveaux services réduiront le recours à l'entretien de collections de documents imprimés, avec les conséquences qui en résultent sur le plan des coûts, mais ils introduiront en même temps le besoin d'une nouvelle discipline dans la gestion (et la gestion des coûts) de l'utilisation des services électroniques. Ils permettront d'améliorer et d'accélérer la diffusion des résultats de la recherche dans le monde entier, et ils faciliteront la communication et l'interaction entre les scientifiques. Fait important, ils donneront accès à des informations auparavant inaccessibles, ou qui n'étaient accessibles qu'à grands frais et difficilement. Comme on l'a noté précédemment, cette nouvelle ère se manifestera par une intégration et une liaison accrues des ressources de données, permettant de trouver, d'analyser et de modéliser les données de manière plus puissante.

Cet énorme élargissement du champ de la publication électronique soulève un grand nombre de préoccupations, parmi lesquelles : les méthodes et mécanismes permettant de préserver et de vérifier la qualité de la science publiée, qui reposaient jusqu'à présent sur un processus très efficace d'examen par les collègues spécialistes ; la gestion et la sécurité des droits de propriété intellectuelle qui, là encore, sont soumis à des mécanismes formels de publication bien établis ; et la fonction essentielle d'archivage des résultats de la recherche scientifique, où il existe le réel danger d'un quasi-effondrement du système parce que l'information électronique prolifère plus rapidement que ne se créent les systèmes d'archivage nécessaires.

La fourniture d'information sur le Web est un grand plaisir pour beaucoup et un grand souci pour d'autres. Dans le milieu scientifique, les chercheurs peuvent facilement converser les uns avec les autres de leurs programmes de recherche et de leurs dernières découvertes. Nombreux sont ceux qui aiment «causer» de ces

nouveautés sur l'Internet et certains pensent que ces messageries collectives remplaceront un jour les revues. Il est incontestable que ces messageries remplissent une fonction, mais les scientifiques insistent tout autant sur le point que la science doit être validée.

Le service de pré-publication de Ginsparg est quelquefois présenté comme un exemple de la réussite de certains services sur l'Internet. Cependant, il faut prendre en compte deux points : premièrement, Ginsparg « se faisait porter » par le système d'examen des articles de la revue primaire, en ne plaçant un tirage préliminaire en ligne que quand l'article avait été accepté par une revue reconnue ; deuxièmement, le service était gratuit. Nous n'entrerons pas dans un débat sur le point de savoir s'il est moral d'utiliser pour les besoins de son propre service le système d'examen des articles d'une revue, mais il est évident que quelqu'un payait le contrôle des articles, considéré comme essentiel.

Les services d'information scientifique doivent faire preuve de qualité, de continuité et de fiabilité. De plus en plus, ils doivent être rapides et efficaces. Tant que le chercheur universitaire ne se soucie pas du temps qu'il faut pour trouver et sélectionner des informations pertinentes et de bonne qualité, il utilisera les informations qui circulent librement. Cependant, il n'y a pas de doute qu'avec l'accroissement des contraintes de temps, l'universitaire, comme le chercheur de l'industrie, voudra recourir à des informations validées et présélectionnées pour accélérer ses recherches et être efficace à moindre coût.

Des éditeurs fournissent ce service et ils utiliseront l'Internet pour faciliter la distribution d'informations scientifiques de valeur. Ils en feront une extension de leur activité présente : apporter à l'utilisateur des informations contrôlées par des collègues scientifiques. Les revues et les bases de données sur les travaux publiés seront de plus en plus intégrées aux bases de données factuelles mentionnées ci-dessus, de telle sorte que les utilisateurs bénéficieront d'un système de ressources d'information interconnectées offrant des contenus de qualité.

Il convient peut-être aussi de mentionner un autre aspect nouveau du besoin de validation. De plus en plus, on demande à l'examineur scientifique de porter un jugement sur des articles qui reposent sur une analyse de données qui ne peut se faire que par ordinateur. Jusqu'à présent, les chercheurs ne sont pas tenus de déposer ces données quand ils rédigent leur article et l'examineur doit leur faire confiance dans une certaine mesure. Dans l'avenir, il est clair qu'il faudra établir des « bases de données d'informations annexes ou auxiliaires » que la collectivité de la recherche pourra utiliser pour vérifier les conclusions ou mesures originales. Ce système existe déjà dans le domaine des séquences de nucléotides et il s'étendra bientôt à d'autres disciplines. Là encore, le financement sera un problème, bien que l'Union européenne s'intéresse aux moyens par lesquels on peut utiliser les fonds publics et privés pour assurer cet ensemble de tâches.

Les fournisseurs de contenu, les propriétaires de banques de données, les gouvernements et les milieux de la recherche débattent constamment de la question des droits de propriété intellectuelle dans un environnement numérique. Les termes d'*intellect* et de *propriété* ne semblent guère concorder. L'intellect est personnel et privé. La propriété concerne la frontière et l'interaction avec le monde. La propriété intellectuelle est donc liée au droit à la reconnaissance de l'auteur ou créateur individuel ainsi qu'au droit d'empêcher les atteintes.

Dans le monde scientifique, technique et médical (STM), il est bien admis que le droit d'auteur est transféré de l'auteur à l'éditeur. Les auteurs veulent être diffusés, ils attachent de l'importance aux idées exprimées dans leurs articles, à l'utilisation de leurs articles dans leurs établissements et à la mise en commun de leurs résultats avec les collaborateurs – ce qui ne s'oppose en rien aux intérêts commerciaux de l'éditeur STM.

Les éditeurs STM s'attachent à protéger l'intégrité de l'œuvre des auteurs STM ; ainsi, les droits intellectuels et les droits de propriété apparaissent comme complémentaires. Ce n'est pas tant la relation auteur-éditeur qui alimente un débat intense mais plutôt les limites à fixer pour l'usage loyal et la nécessité de protéger les droits économiques de l'éditeur.

Ainsi, l'éditeur a besoin du droit voisin de protéger son investissement dans l'organisation et la présentation de l'œuvre (déjà protégé par la loi britannique), ou il doit être en mesure d'obtenir par contrat ou licence le droit de poursuivre en justice les violations de ses droits économiques ou commerciaux. Cela doit aussi couvrir la protection des bases de données afin de protéger le savoir-faire et l'investissement apportés à l'organisation d'une base de données. Actuellement, la plupart des discussions sur le droit d'auteur mettent l'accent sur les aspects de « créativité » plutôt que sur la valeur en termes d'utilité et de complétude.

En Europe, le droit *sui generis* (ce qui veut dire simplement « nouveau » ou « différent ») de la Directive de l'Union européenne sur la protection juridique des bases de données est une approche – et une approche créative – visant à inciter les producteurs de bases de données à produire plus de bases de données en Europe. La Directive note que la protection est destinée au fabricant de la base de données afin de protéger son investissement.

Contrairement à la Directive de l'Union européenne, la protection des bases de données aux États-Unis qui a été adoptée par la Chambre des Représentants et qui est actuellement à l'étude au Congrès a pour base le concept de détournement (principes de la concurrence déloyale) et non la création d'un nouveau droit dans les activités commerciales. Cependant, l'approche de l'Union européenne et celle des États-Unis ont le même résultat. Toutes deux protègent le propriétaire contre une appropriation illicite.

Les éditeurs et leurs associations défendent très activement leurs droits de propriété intellectuelle tout en formulant et en proposant des droits voisins pour protéger leur activité commerciale, leur investissement et leur continuité. Bien sûr, l'opinion publique considère cela comme purement mercantile, mais il n'y a pas de différence entre les éditeurs commerciaux ou non commerciaux : tous considèrent qu'il leur incombe d'améliorer la circulation de l'information entre les scientifiques de manière économique, permanente et garantissant la qualité.

Cependant, les éditeurs ne fondent pas leurs activités, et en particulier leurs activités électroniques, sur une base essentiellement législative. Les expériences mentionnées ci-dessus ont apporté aux éditeurs et aux collectivités d'utilisateurs beaucoup d'informations sur les questions techniques et commerciales. Sur la base de ces constatations, des systèmes de licence ont été établis entre les deux parties, couvrant tous les aspects des conditions de livraison et d'utilisation. De cette manière, les deux parties savent à quoi s'en tenir et aucune n'a à craindre une utilisation incontrôlée de l'information dans ces collectivités.

L'information scientifique *n'est pas* à la croisée des chemins, malgré ce que disent certains critiques. Elle doit relever un certain nombre de défis qui devraient être à sa portée. De plus en plus, le besoin d'une interaction entre des initiatives de financement publiques et privées se fera sentir. La recherche de base découvre des faits de base, que les scientifiques utilisent pour des travaux appliqués ou pour la poursuite de leurs travaux intellectuels. La présente thèse est que ces données brutes de base devraient être dans le domaine public (par exemple, la Data Library de l'EMBL) mais que l'on peut assurément faire payer la valeur ajoutée à ces données. La valeur ajoutée implique des investissements et les investissements devraient être protégés.

Sans aucun doute, nous avons besoin d'un environnement où les projets subventionnés ne puissent être en concurrence avec les projets non subventionnés, et où les bases de données puissent évoluer à partir du niveau de la R-D de base auquel elles voient le jour pour devenir des produits exploitables (qu'ils soient financés par des fonds publics ou par l'utilisateur qui s'y abonne). Le chercheur dans le domaine des protéines qui, par exemple, recherche de nouvelles structures moléculaires pour créer un médicament explore aujourd'hui une trentaine de bases de données parmi la centaine disponible sur l'EMBnet. Il vérifie les données dans les bases de données secondaires et examine d'autres conclusions dans les revues originales. Tout cela doit se faire dans un environnement vérifié, stable, viable, en qui l'on ait toute confiance. L'éditeur joue ici un rôle clé qu'il conservera.

Cependant, ce n'est qu'une partie de la solution, parce que nous continuerons de rendre le contenu électronique facile à trouver, à obtenir et à négocier pour les utilisateurs tout en facilitant le suivi et la comptabilité pour les éditeurs et les bibliothèques. C'est ce qu'on peut appeler les transactions sur le contenu

électronique. La réponse doit se faire par des normes. Une norme doit s'étendre au-delà de toute collectivité particulière. Par exemple, les éditeurs tiennent à conserver leur identité de marque. Cependant, cela doit se faire à un niveau plus élevé que le moyen de communication ou l'interface appartenant en propre à un fournisseur. Il nous faut des normes dans les technologies habilitantes de base offrant une assise sûre sur laquelle on puisse créer et apporter une valeur ajoutée. Il nous faut des normes dans un certain nombre de domaines : nous pourrions attendre qu'elles apparaissent mais les éditeurs ont choisi de s'engager parce qu'elles sont essentielles à l'avenir de nos branches d'activité, avec tous les acteurs de la chaîne. Ces normes devraient couvrir : l'identification, les données descriptives, les données sur les droits, les formats de présentation et d'échange, l'identification de l'utilisateur, et l'autorisation et l'affranchissement des droits.

Il reste beaucoup de travail, mais on constate une grande coopération et beaucoup de bonne volonté. Comme exemple de cette collaboration, on peut mentionner la mise au point de numéros de série de document normalisés dans le cadre de la *Digital Object Identification* (identification d'objets numériques – DOI). A une réunion tenue à Francfort en octobre 1997, ce projet a été adopté par toute la collectivité des éditeurs STM ainsi que par les industries photographique et phonographique. C'est un point particulièrement important parce que toute information, dans quelque domaine ou sous quelque format que ce soit, pourrait être traitée comme un objet numérique. Peut-être heureusement d'un point de vue commercial, nous n'avons pas encore un monde de l'édition numérique homogène ; il subsiste des différences notables entre le texte, les images et le son, par exemple. Cependant, les technologies et normes sous-jacentes doivent converger : nous voulons, et c'est un besoin pour nos utilisateurs, présenter des images de structures chimiques avec un texte de chimie et peut-être des séquences phonétiques avec nos dictionnaires linguistiques. La technologie doit rendre cela possible mais, dans la pratique, cela implique que des normes sont nécessaires pour que l'on puisse pleinement tirer parti de l'Internet et du développement futur que l'on entrevoit pour ce réseau.

Ainsi, les transactions sur le contenu électronique nécessitent une infrastructure que nous puissions tous utiliser avec confiance, pour pouvoir consacrer nos efforts à bâtir de nouvelles solutions créatives et apporter une valeur ajoutée à nos produits au lieu de gaspiller notre temps à convertir des formats de fichiers ou à surmonter d'autres obstacles. L'établissement de ces normes ne se fera pas sans peine : cela exigera du temps et de la volonté, ainsi que la participation à des débats tels que celui-ci, le cas échéant à des commissions, et à des projets pilotes et prototypes. Les résultats de ce processus dépendront de l'ampleur de notre contribution.

Cet énorme défi à relever ouvre aussi de grandes perspectives. En lançant ce projet et en collaborant pour le bien mutuel de l'industrie de l'information

scientifique, et en consacrant des efforts à définir des normes efficaces, nous pouvons faire en sorte que l'infrastructure que nous bâtissons soit utilisée par la totalité du monde de la recherche pour le plus grand profit de la science.

V. CONCLUSION

Nous avons tendance à voir la science sous l'aspect de ses processus intellectuels et de sa contribution au développement de nouvelles techniques ou technologies. Dans une certaine mesure, nous pouvons oublier les vastes réseaux et systèmes de transfert de l'information qui soutiennent la science, et qui représente un coût très substantiel.

La communication scientifique obéissait jusqu'à présent à un ordre établi de longue date, reposant pour une large part sur des processus de publication traditionnels que l'on remet en question actuellement pour s'adapter à l'essor de la publication électronique, dû à la croissance spectaculaire de la puissance et de l'accessibilité des technologies de l'information. On a illustré dans le présent article les tendances de l'édition et on a décrit quelques-unes des questions économiques et stratégiques qui se posent actuellement à la communauté scientifique et à son infrastructure de publication.

Il s'ouvre sans aucun doute de larges et vastes perspectives, et le résultat final sera un nouvel environnement propre à assister et améliorer fortement le processus de la recherche. Toutefois, cela nécessitera un certain réexamen du rôle des éditeurs et, en fait, du rôle des scientifiques eux-mêmes dans la chaîne de transfert de l'information. Cela conduira aussi à revoir fondamentalement les coûts et l'économie de la communication scientifique, avec un impact croissant sur le financement de la recherche.

Les diverses questions soulevées ici, notamment des aspects comme les normes, les droits de propriété intellectuelle, et l'économie et la viabilité des services de publication électronique, ne seront plus le souci des seuls éditeurs mais deviendront de plus en plus pertinentes et importantes pour ceux qui gèrent ou financent la recherche. Toutes les milieux de la science devront collaborer pour établir une base viable et pérenne pour cette nouvelle ère de la publication électronique.

LA BIBLIOTHÈQUE DE RECHERCHE ÉLECTRONIQUE DU DANEMARK : UN OUTIL DU CHANGEMENT INSTITUTIONNEL

Table des matières

I Introduction	112
II. La bibliothèque de recherche électronique du Danemark	114

Cet article a été rédigé par Jens Thorhauge, membre de l'administration de la Bibliothèque nationale du Danemark.

I. INTRODUCTION

L'accès rapide à l'information – matière première du savoir – est un facteur stratégique essentiel dans de nombreux domaines, notamment ceux de la recherche, de l'enseignement et du commerce. Récemment, la Banque mondiale a affirmé que de tous les facteurs, le savoir était celui qui présentait le plus d'importance pour les états industrialisés. La Société de l'Information s'appuie sur la production et l'utilisation massive d'informations et elle se caractérise en outre par le nombre croissant de citoyens qui sont amenés à consulter des informations spécialisées dans leur travail quotidien.

L'importance de la documentation électronique réside dans le fait que, grâce aux réseaux, davantage de personnes peuvent accéder plus rapidement à un volume plus grand d'informations. Et, lorsque les bibliothèques combinent services traditionnels et services sur réseaux, les avantages potentiels pour les progrès de la connaissance sont encore plus importants. Cela s'applique bien sûr tout particulièrement aux chercheurs, mais l'impact sera également considérable sur le relèvement général des qualifications de la population active et sur la consultation et l'utilisation des informations par le public. Une interconnexion très poussée des systèmes d'information pour la recherche est essentielle si l'on veut favoriser une meilleure productivité dans le village mondial de la recherche.

Dans le monde entier, les bibliothèques évoluent de manière spectaculaire, en s'orientant toujours plus vers la numérisation au détriment du traditionnel support papier. Dans les bibliothèques de recherche d'Europe et des États-Unis plus particulièrement, d'énormes efforts sont déployés pour introduire de nouveaux services fondés sur les TI (technologies de l'information).

An cours des vingt dernières années, cette révolution de l'information s'est accélérée. La première étape a été celle de l'automatisation : les TI ont été utilisées pour des fonctions de gestion, telles que la circulation, l'acquisition, le catalogage et le contrôle des stocks. L'étape suivante a fait intervenir l'accès en ligne aux bases de données, généralement par des catalogues centraux offrant un point d'accès unique à plusieurs fonds bibliothécaires. Avec l'évolution des services sur réseau, les bibliothèques ont commencé à utiliser l'Internet. La dernière étape vers la bibliothèque numérique a vu les bibliothèques proposer des services Internet sur leur page d'accueil pour les utilisateurs distants, l'accès aux bases de données en texte intégral, les guides sur Internet, des fonctions de passation de commande,

la distribution électronique de documents et des services ciblés dans les domaines de la recherche, du commerce, de l'enseignement et de la formation aux TI. Cette évolution affecte l'organisation ainsi bien que les fonctions, les services et les objectifs spécifiques des bibliothèques. Le changement fondamental réside essentiellement dans le fait que les bibliothèques privilégient maintenant non plus le développement des collections mais plutôt l'établissement de liens et le service à l'utilisateur. Cette constatation révèle un changement profond dans la culture et les valeurs de la profession.

La manière dont on imaginait la bibliothèque numérique au début des années 1990 était simple : des fonds numérisés permettant d'accéder aux ressources de la bibliothèque depuis n'importe quel ordinateur disposant d'une connexion, l'objectif final étant l'accès direct à l'information ou la mise à disposition immédiate des informations à l'utilisateur. Avec l'essor d'Internet, de nouveaux termes sont apparus : bibliothèque électronique, bibliothèque virtuelle, bibliothèque numérique – le dernier étant probablement le plus exact. De récents travaux ont clairement montré que l'idée de bibliothèques totalement virtuelles, c'est-à-dire constituées d'un seul système entièrement numérique, n'était pas réaliste. Ce que nous recherchons à l'heure actuelle c'est un accès plus rapide et plus efficace à un volume plus important de documents de recherche au profit d'un public scientifique et universitaire plus vaste.

De récents articles présentent les bibliothèques numériques comme un ensemble de ressources et de systèmes disparates, qui desservent des communautés et des groupes d'utilisateurs spécifiques et répondent à des objectifs déterminés. Dans les sciences humaines en particulier, les bibliothèques conserveront sans doute indéfiniment un important fonds sur support papier.

Les bibliothèques numériques doivent relever d'importants défis : mettre en place une architecture technique appropriée, rassembler la masse critique de contenus numérisés susceptible d'attirer les utilisateurs, établir des normes et gérer les droits d'auteur, ce qui est encore plus difficile dans le monde numérique car il n'est pas possible de contrôler les copies. La protection des oeuvres numériques est une question essentielle encore sans réponse. En conclusion, la constitution de bibliothèques numériques s'avère difficile et coûteuse et demande des efforts durables et soutenus. L'enjeu au plan national consiste à établir un programme coordonné dans lequel les bibliothèques établies coopèrent et fournissent l'accès en réseau à leurs systèmes et ressources disparates par le biais d'une interface unique. Toutes les bibliothèques modernes permettent de consulter des documents numérisés ; elles ont maintenant la possibilité d'améliorer ce service en adhérant à un nouveau type de collaboration (qui pourrait s'avérer douloureux car il équivaut à renoncer en partie à leur liberté d'action). L'évolution technologique conduit inévitablement à une évolution organisationnelle, car une part croissante du fonds des bibliothèques scientifiques doit être partagée pour des raisons économiques

et les utilisateurs réclameront ce partage parce que la technologie offre justement de spectaculaires possibilités dans ce sens. On s'attend à ce que les institutions perdent de leur autonomie, mais l'on ne sait pas encore exactement jusqu'à quel point.

A court terme, la situation devrait évoluer vers un accès rapide et aisé aux articles et informations bibliographiques pour un public élargi. Les conditions sont donc réunies pour que l'on effectue un bond qualitatif vers un service qui permettrait aux bibliothèques d'atteindre plus efficacement leurs objectifs – au profit de l'ensemble de leurs utilisateurs.

Un autre aspect important du problème concerne les incidences sur les relations institutionnelles. La coopération en matière de ressources numériques devrait conduire à une collaboration dans d'autres domaines, tels que les systèmes de gestion, et devrait se traduire par une rationalisation accrue.

Le présent article s'intéresse particulièrement à la Bibliothèque de recherche électronique du Danemark qui fournit un exemple de stratégie pour changer les relations entre institutions, élaborer de nouveaux services et pérenniser un projet, sans financement spécifique.

II. LA BIBLIOTHÈQUE DE RECHERCHE ÉLECTRONIQUE DU DANEMARK

Au cours des cinq prochaines années (1998-2002), les bibliothèques de recherche du Danemark seront réorganisées de manière à former une seule bibliothèque numérique intégrée : la Bibliothèque de recherche électronique du Danemark. Les ministères de la Recherche et de l'Éducation ainsi que celui de la Culture investiront 200 millions de couronnes dans ce projet, dont l'objectif est de fournir aux chercheurs, étudiants, entreprises et autres professions un accès plus aisé, plus rapide et plus efficace aux découvertes les plus récentes de la recherche. Ce projet fait partie de la nouvelle initiative gouvernementale en faveur de la recherche et des TI, projet au profit duquel 126 millions de couronnes ont été débloqués au budget 1998. Les trois ministères concernés lui ont par ailleurs affecté 74 millions de couronnes supplémentaires, ce qui porte le total à 200 millions de couronnes.

Le concept

Ce réseau de bibliothèques de recherche va constituer un système virtuel, qui transcendera les frontières des bibliothèques régionales et locales, de manière simple et transparente, à l'intérieur du cadre juridique et économique existant. Dans tout le pays, les utilisateurs auront ainsi accès aux ressources documentaires traditionnelles ou numérisées de l'ensemble des fonds de ces bibliothèques.

La création de la Bibliothèque de recherche électronique du Danemark s'appuie sur la technologie existante. L'objectif est de proposer une offre d'information nationale efficiente, qui rende accessible les ressources développées dans les bibliothèques de recherche de l'ensemble du pays et permette à d'autres centres de s'intégrer dans ce système d'information virtuel. La Bibliothèque de recherche électronique apparaîtra comme un grand système d'information virtuel homogène et unique, dont l'effet global est obtenu par l'interconnexion des réseaux de bibliothèques de recherche et la mise en conformité aux normes de communication, notamment pour l'aide à la recherche, l'enregistrement et l'indexation, la description et la présentation des documents.

L'accès électronique aux informations du catalogue et à un nombre croissant de documents numérisés figurera parmi les fonctions offertes aux utilisateurs. Plus précisément, le projet doit assurer :

- Une recherche rapide et simultanée dans plusieurs bases de données, s'appuyant sur une combinaison de nombreux profils de recherche, indépendamment de la localisation physique des bases de données, nationales ou internationales.
- Un système de recherche convivial et harmonisé avec la même interface utilisateur, quelle que soit la base de données consultée.
- Une liste automatiquement actualisée des publications et périodiques, affichée par des robots qui utilisent des profils et critères de recherche individuels.

Le système sera accessible à tous : depuis le domicile, le bureau ou la bibliothèque.

Le contenu de la Bibliothèque de recherche électronique a été étudié en détail dans un rapport remis l'année dernière en février par un groupe de travail. La première préoccupation mentionnée dans ce rapport est la constitution d'une « masse critique » de documents archivés. A cet effet, il faut rendre disponible un grand volume de ressources dans des conditions qui permettent aux chercheurs d'y accéder aisément et rapidement. Dès que les avantages deviendront évidents, ils commenceront à utiliser les fonctions proposées. Le rapport identifie trois types de ressources documentaires importantes pour la mise en place de la Bibliothèque de recherche électronique et il recommande de leur accorder une attention hautement prioritaire dans la phase d'étude :

- Bases de données en texte intégral, principalement des revues, rendues accessibles par le biais de licences auprès de grands éditeurs.
- Bases de données de référence.
- Bibliographies nationales, principalement la base danoise DanBib.

Une priorité élevée doit par ailleurs être accordée aux outils qui font le lien entre la bibliothèque virtuelle et la bibliothèque physique, c'est-à-dire qui servent à la conversion des catalogues en sens inverse. Avec l'outil approprié, il est possible de numériser des parties de collections de documents. De toute évidence, la première étape dans la constitution d'une « masse critique » est bien de garantir l'accès au plus grand nombre possible de revues pertinentes.

Trois facteurs peuvent servir à décrire le concept de la Bibliothèque de recherche électronique du Danemark – le défi posé, l'objectif visé et les moyens employés :

- Le *défi* est de faire face à l'augmentation de la production d'informations scientifiques, et de permettre l'utilisation et la mise en œuvre de ces informations, qui ont un rôle essentiel dans le développement de la société, de l'enseignement et de l'activité économique.
- L'*objectif* est clair : Le Danemark souhaite bénéficier des avantages qualitatifs et économiques des bibliothèques numériques – et interconnectées – et offrir aux utilisateurs un accès direct aux informations scientifiques, quel que soit l'endroit où elles sont situées.
- Enfin, les *moyens* peuvent se résumer ainsi : Un réseau de bibliothèques scientifiques et de centres de documentation interconnectés en une structure cohérente permettant aux chercheurs et utilisateurs d'obtenir des résultats de recherche de manière plus aisée et plus rapide, et offrant un accès plus efficient aux informations scientifiques les plus récentes.

Le projet imaginé au Danemark s'articule autour d'une plate-forme par laquelle, d'une part, les chercheurs dans les douze plus grandes bibliothèques scientifiques pourront accéder à un éventail élargi de revues électroniques mais aussi et surtout les petites bibliothèques scientifiques et les bibliothèques des institutions de recherche et d'enseignement de moindre importance bénéficieront d'un maximum d'avantages dans le cadre de licences nationales. Dans une phase ultérieure, les bibliothèques de comté disposeront d'un accès « à péage » et, à terme, toutes les bibliothèques du Danemark devraient être reliées à la Bibliothèque de recherche électronique.

Le projet prévoit de nouveaux services faisant appel à des profils utilisateur individuels et à la mise en place d'agents intelligents. Cela pourrait notamment présenter comme avantage particulier qu'en développant la Bibliothèque de recherche électronique, on puisse mettre à disposition non seulement le catalogue complet de toutes les publications scientifiques danoises, mais aussi leur présentation en texte intégral. Cette idée est étroitement liée à la politique de publication des universités et des chercheurs en liaison avec le code de l'édition électronique et institutionnelle. La voie est ouverte à l'établissement d'une nouvelle structure de publication – dans une perspective à plus long terme toutefois.

Pour concrétiser cette idée, le projet danois a identifié quatre composantes essentielles – chacune étant nécessaire pour atteindre l'objectif :

- L'infrastructure nationale.
- L'infrastructure des bibliothèques.
- Les ressources numériques.
- Les équipements pour les utilisateurs.

L'infrastructure nationale

L'infrastructure nationale comprend le réseau et les équipements de TI grâce auxquels les bibliothèques et les utilisateurs peuvent communiquer efficacement. Le réseau de TI choisi est le *Danish Research Net* (réseau danois pour la recherche). Ce réseau présente l'avantage qu'une partie appréciable des utilisateurs sont déjà connectés. Cette infrastructure nationale ne se limite cependant pas uniquement à la technologie. L'infrastructure globale exige l'adoption de principes communs, notamment pour les accords de licence nationale, l'échange d'informations, l'utilisation de normes internationales, l'uniformisation de l'accès des utilisateurs, etc. Les règles d'administration des utilisateurs doivent par ailleurs être homogènes et librement acceptées.

L'infrastructure des bibliothèques

Pour que toutes les bibliothèques s'intègrent dans cette bibliothèque virtuelle commune, une certaine modernisation est nécessaire. Jusqu'à présent, on admettait que chaque bibliothèque utilise ses propres systèmes de TI et procédures d'organisation. La création d'une bibliothèque de recherche virtuelle exige une standardisation sur le plan de la technologie et sur de nombreux points d'organisation.

Pour la mise en œuvre du concept, une nouvelle technologie sera nécessaire. Afin d'en tirer le maximum d'avantages, les bibliothèques doivent changer leurs habitudes et asseoir leur collaboration sur des normes et des orientations communes.

La Bibliothèque de recherche électronique du Danemark doit adopter un profil et un niveau de service qui fassent penser aux utilisateurs qu'ils sont les «clients» d'une entité virtuelle commune plutôt que de bibliothèques de recherche physiquement dispersées, régies par des conditions et des philosophies différentes. Un profil similaire doit être présenté à tous les fournisseurs. Il s'avère donc nécessaire d'élaborer un ensemble de règles communes pour formaliser cette coopération, et notamment :

- La possibilité de commande directe depuis n'importe quelle bibliothèque, ce qui affectera la collaboration entre bibliothèques et les appuis financiers par les institutions et les ministères concernés.

- L'acquisition d'un exemplaire numérique unique, ce qui exigera une certaine coordination, notamment en matière d'accords sur les licences nationales.

L'intégration de l'ensemble des utilisateurs des bibliothèques scientifiques exigera en outre que le niveau de service soit déterminé suivant un schéma directeur commun.

La collaboration accrue entre bibliothèques scientifiques rendra nécessaire une gestion et une coordination globales. Une coopération interministérielle devra s'instaurer, les participants locaux devant toutefois conserver leur indépendance afin que la dynamique du système soit préservée.

Les ressources numériques

Si pour les documents classiques «non numériques» la question la plus importante reste le catalogue, la numérisation de certaines parties des fonds est à l'étude. Pour les «nouveaux» documents, fournis sous forme numérique, plusieurs défis sont à relever et notamment : la nécessité d'une gestion efficace ; l'accès généralisé ; la protection contre les dommages et les abus ; et la migration vers les plates-formes technologiques du futur.

Au début, les serveurs de documents devront disposer de la capacité mémoire nécessaire pour des documents textuels au format HTML ou PDF. Ultérieurement, il faudra disposer de capacités mémoires supplémentaires pour un volume plus important de documents numérisés ou multimédia.

Les équipements pour les utilisateurs

Les équipements poseront un problème majeur aux utilisateurs de bibliothèques numériques, plus particulièrement sur le plan économique. L'essentiel est qu'ils soient en nombre suffisant : stations de travail, imprimantes et logiciels en libre accès pour la récupération et le traitement des informations.

Dans ce contexte, on envisage des ordinateurs personnels (PC) ou des ordinateurs de réseau (NC). Le choix est difficile à l'heure actuelle. Les bibliothèques et la majorité des utilisateurs sont familiarisés avec les PC et des logiciels de tous types sont disponibles pour ces appareils. D'un autre côté, les NC peuvent constituer une option intéressante – en particulier s'ils tiennent leurs promesses de sécurité, d'économie et plus particulièrement, de gain de temps pour l'administration du système.

Organisation et plans

L'étude

Le projet a été défini dans un descriptif établi en septembre 1996 par les trois ministères concernés. Une évaluation a été réalisée par UNI*C et Ernst and Young.

Le rapport résultant, disponible en anglais sous forme résumée, a été remis au printemps 1997.

Axée sur les aspects technologiques de l'information, l'évaluation ne traite que brièvement du développement des procédures et des conséquences organisationnelles. Elle isole les facteurs qui changeront les processus et l'organisation, sans proposer aucune recommandation ni conseil sur la manière de mettre en œuvre les changements.

Organisation du projet DEF

Le projet DEF réunira dans un premier temps les douze plus grandes bibliothèques de recherche, même si en principe toutes les bibliothèques seront invitées à participer.

L'organisation responsable du projet de mise en œuvre de la Bibliothèque de recherche électronique du Danemark comprend :

- Un *conseil de coordination* composé de membres des trois ministères.
- Un *comité directeur* de dix membres nommés par les ministères et les organisations.
- Un *secrétariat* intégré dans l'administration de la Bibliothèque nationale du Danemark.

Diverses compétences sont représentées au sein du comité directeur : gestion de bibliothèque, recherche et TI.

Le rôle de l'administration de la Bibliothèque nationale consistera de manière générale à exécuter les décisions du comité directeur ; en ce qui concerne les licences, c'est elle qui sera officiellement détentrice des licences nationales.

Les licences nationales seront définies, non pas par rapport aux utilisateurs, mais en relation avec l'aide économique de l'État. Conformément aux propositions du rapport, toute bibliothèque peut collaborer avec toute autre bibliothèque ou institution, former un consortium et négocier des licences. Le consortium peut demander une aide à l'administration de la Bibliothèque nationale du Danemark qui veillera à ce que toutes les institutions intéressées soient invitées à utiliser en commun la licence.

En liaison avec l'organisation de la Bibliothèque de recherche électronique, un groupe de travail sera constitué pour traiter de l'attribution des licences au plan national, qui pour l'heure semble le modèle privilégié pour la gestion de droits.

Projets pilotes

Six projets pilotes ont été lancés de 1997 à 1998, dotés d'un budget de 8.2 millions de couronnes. Ils couvrent la communication, la numérisation et l'archivage

numérique et touchent un large éventail d'utilisateurs de bibliothèques dans tout le pays.

Ces six projets sont les suivants :

- Bibliothèque universitaire virtuelle couvrant l'ensemble des informations numériques en sciences sociales (Bibliothèque nationale et universitaire d'Aarhus, 2 millions de couronnes).
- Accès rapide à la collection de CD-ROM de la Bibliothèque royale pour un large éventail d'utilisateurs dans les bibliothèques publiques et universitaires.
- Établissement d'une installation expérimentale pour la recherche d'informations numériques à grande échelle au profit d'un vaste éventail d'utilisateurs sur réseau local.
- Numérisation des publications scientifiques dans le domaine de la communication de manière à constituer une archive numérique pour la recherche et l'enseignement.
- Établissement d'un système d'information rapide et permanent pour la publication électronique d'ouvrages de recherche historique en collaboration entre une université, une bibliothèque universitaire et un éditeur universitaire.
- Constitution et promotion d'une base de données d'informations de référence combinant des articles scientifiques et une documentation scientifique en texte intégral.

État d'avancement du projet au début de 1999

- L'infrastructure nationale est opérationnelle.
- Un plan de mise à niveau des systèmes documentaires des 12 plus importantes bibliothèques scientifiques a été défini et sa mise en œuvre a commencé. Un autre plan visant 44 bibliothèques scientifiques a été approuvé et sa réalisation a démarré.
- Les ressources numériques augmentent rapidement. La délivrance de licences nationales a donné accès à des revues dans pratiquement tous les domaines.

La numérisation des anciens catalogues est une tâche énorme – une partie a été réalisée et un plan d'ensemble est prévu pour bientôt.

De manière similaire, un plan pour la numérisation des collections existantes n'a pas encore été défini et cette tâche est envisagée dans une perspective à long terme. Comme la numérisation des collections est coûteuse, il faut trouver des outils permettant de sélectionner les éléments appropriés dans les documents,

c'est-à-dire les parties présentant un intérêt particulier pour l'utilisateur et dont la numérisation offre des avantages évidents.

Les défis les plus ardues concernent la mise en place d'un système efficace de gestion des utilisateurs, la création d'une interface logique et intuitive et la gestion des droits. En l'occurrence, les utilisateurs devraient pouvoir s'identifier eux-mêmes pour accéder aux documents sous licence, bénéficier d'une orientation thématique claire et enfin, la gestion des droits ne devrait pas constituer un fardeau administratif insupportable.

Le Comité directeur étudie actuellement l'organisation future des bibliothèques scientifiques. La stratégie à la base du projet sur cinq ans consiste à accélérer l'inévitable changement de paradigme. Des fonds supplémentaires sont certainement utiles pour mettre à niveau les systèmes et permettre l'accès aux revues électroniques. Toutefois, si l'on en restait là, après cinq ans, il faudrait un nouveau projet pour répondre à d'incessants besoins de mise à niveau. Ce qu'il faut mettre en place, c'est un système susceptible de satisfaire d'autres priorités que celles du moment.

Deux possibilités sont offertes pour le changement d'organisation. La première consiste à se contenter d'employer les fonds pour moderniser les bibliothèques et acquérir des licences nationales pour accroître le volume d'informations numérisées. Cela reviendrait à améliorer le système que nous connaissons aujourd'hui, ce qui serait insuffisant pour le monde de demain.

L'autre extrême consisterait à établir un organe national qui gérerait de façon centralisée les bibliothèques scientifiques et déciderait de normes communes et négocierait les licences, mais qui laisserait un budget réduit et beaucoup de frustration aux branches locales de la bibliothèque numérique nationale. Un tel modèle tuerait l'initiative locale et bloquerait toute évolution ultérieure.

On peut tracer un parallèle entre le développement des bibliothèques numériques et celui de l'organisation des cartes de crédit. A l'origine, chaque banque disposait de sa propre carte, avec un nombre réduit de transactions et des dépenses de fonctionnement élevées. Le concept de la technologie des cartes de crédit était – comme chacun sait – de créer une carte universelle acceptée par toutes les banques. L'introduction de cette dernière était subordonnée au respect de normes sur le plan de la technique et de l'utilisation mais aussi de la déontologie. A partir de cette plate-forme, les cartes de crédit sont rapidement devenues universelles et leur rythme de progression et de développement est à son apogée, bien que la carte bancaire ait été et qu'elle soit encore administrée par une très petite organisation.

Dans le domaine documentaire, il suffit de règles très simples pour organiser une bibliothèque numérique nationale. L'adhésion n'est pas obligatoire mais dès qu'une bibliothèque en décide ainsi, elle est tenue d'accepter les règles, toutes les bibliothèques participantes peuvent et doivent innover en matière de services

communs et enfin, la bibliothèque numérique doit appartenir à ses adhérents et être administrée par ces derniers. Ce système s'appuie sur une informatique globale, sur des modèles de mise en réseau du type que l'on rencontre dans la Silicon Valley et sur la tradition danoise des sociétés coopératives.

L'objet essentiel du projet de Bibliothèque de recherche électronique du Danemark est d'améliorer nettement le service rendu aux utilisateurs, afin qu'à leur tour, ils encouragent fortement les bibliothèques à poursuivre ce service ainsi que leur collaboration.

L'AVENIR DES BASES DE DONNÉES MATHÉMATIQUES

Table des matières

I. Introduction.....	124
II. Généralités.....	125
III. Bases de données en mathématiques.....	134
<i>Annexe</i>	139

Cet article a été rédigé par Jean-Pierre Bourguignon et la Société européenne des mathématiques dans le cadre des travaux du Forum Mégascience de l'OCDE.

I. INTRODUCTION

Les scientifiques sont tous conscients du fait que la communication, la publication et l'archivage des travaux de recherche traversent actuellement une période de mutation sans précédent, grâce aux évolutions de la technologie électronique. Alors que le changement est tout autour de nous, force est d'admettre que nous ne pouvons encore voir où conduira cette révolution électronique dans le domaine de la publication de revues et de leur archivage dans les bibliothèques. La situation définitive qui se dégagera au début du ^{xxi}^e siècle est susceptible de varier énormément d'une discipline à l'autre. Cependant, alors que la revue scientifique traditionnelle est manifestement menacée si elle ne s'adapte pas au nouvel environnement, il apparaît déjà clairement que les bases de données scientifiques comptent parmi les grands bénéficiaires de la nouvelle technologie. Pour la première fois, il est techniquement possible, dans une discipline spécifique, de constituer des bases de données électroniques qui disposent de méthodes de recherche efficaces et flexibles, et qui peuvent devenir un outil de travail de base pour les chercheurs du domaine concerné dans le monde entier. Nous pensons que le moment est venu pour chaque discipline scientifique de saisir l'occasion magnifique qui lui est donnée de créer et de gérer intelligemment une structure qui soit accessible à tous les intervenants du domaine et qui devrait, à n'en pas douter, influencer insensiblement et profondément l'évolution à long terme de la discipline. Dans cet article, nous aborderons les questions complexes qui sous-tendent ce défi dans le domaine des mathématiques. La recherche de preuves rigoureuses est au cœur de tous les travaux dans ce domaine et comme celles-ci sont toujours contenues dans des structures mathématiques plus larges, il est vital pour le chercheur de cette spécialité de pouvoir citer les ouvrages précédemment publiés qui permettent de vérifier de manière précise la validité de ces structures. Dans certains domaines actuels des mathématiques, il n'est pas inhabituel de citer des travaux publiés il y a cent ans ou plus. De ce point de vue, les bases de données des publications passées en mathématiques diffèrent probablement quelque peu des bases de données nécessaires dans d'autres domaines de la recherche scientifique. Néanmoins, nous pensons que nombre de thèmes et de principes généraux présentés dans ce rapport sont communs à la plupart des disciplines.

II. GÉNÉRALITÉS

La révolution de l'édition scientifique

De par la nouvelle technologie électronique, la diffusion et la publication des recherches mathématiques connaissent actuellement une profonde révolution. Bien que l'on puisse supposer qu'il en est ainsi dans tous les principaux domaines scientifiques, nous nous concentrerons dans cet article essentiellement sur les mathématiques, eu égard au fait que c'est dans ce domaine que la Société mathématique européenne exerce ses compétences et que les mathématiciens se servent de leurs publications différemment des autres scientifiques. Toutefois, nombre de thèmes et de principes généraux sont de toute évidence communs à la plupart des grandes disciplines scientifiques.

Nous ne pensons pas qu'il soit exagéré d'établir un parallèle entre la révolution électronique actuelle et les évolutions spectaculaires dans la diffusion du savoir qu'ont entraînées l'invention de l'imprimerie au xv^e siècle en Europe. Dans la suite, nous étudierons quelques unes des évolutions spécifiques intervenant actuellement dans la publication d'ouvrages mathématiques.

L'évolutions de l'imprimerie

En typographie, la traditionnelle composition mécanique a presque été totalement remplacée par la composition automatique, dans laquelle l'imprimante est commandée par un programme informatique. Une évolution qui a conduit Donald Knuth à créer le programme TEX, lequel permet littéralement à l'auteur d'une publication mathématique de choisir sa propre composition suivant des normes élaborées en même temps qu'il rédige son travail. De plus en plus de mathématiciens utilisent TEX et les variantes qui lui ont succédé pour rédiger leurs textes et il est certain que cette méthode devrait avant longtemps devenir pratiquement universelle.

Cette évolution s'est avérée possible grâce à l'habitude que les scientifiques en général, et les mathématiciens en particulier, ont de l'informatique, ce qui leur permet de préparer eux-mêmes leurs textes. Cette nouvelle manière de rédiger des documents offre de nombreux avantages, les plus évidents étant la possibilité de procéder aisément à des modifications et de choisir librement les symboles et la présentation des formules mathématiques. L'éditeur reçoit maintenant une disquette ou un fichier électronique TEX et après seulement quelques modifications minimales destinées à garantir la cohérence de style dans l'ensemble de la revue concernée, il peut imprimer le document.

Ces quinze dernières années, le passage à la composition électronique a indubitablement entraîné une baisse considérable des coûts de fabrication des revues. Force est de constater que pratiquement rien ne prouve que des éditeurs du

commerce ou des institutions savantes aient répercuté les économies réalisées dans ce domaine auprès des acheteurs des revues. Avec cette nouvelle forme de production, les revues nouvelles paraissent par ailleurs pour la plupart automatiquement sous forme électronique. Les éditeurs diffusent de plus en plus cette version aux bibliothèques universitaires par le biais de licences de site, dans le cadre d'un forfait compris dans un abonnement à la revue. On peut légitimement se demander si, dans un futur pas trop éloigné, la production d'un exemplaire imprimé de la revue ne deviendra pas l'exception plutôt que la règle.

L'Internet

La création de l'Internet a révolutionné les communications entre les chercheurs en mathématiques ainsi que la diffusion des textes publiés à titre officieux ou officiel. Si les mathématiciens ont toujours beaucoup communiqué au niveau international (comme le prouve, par exemple, la création de l'Union mathématique internationale vers la fin des années 1890, bien longtemps avant que la plupart des autres disciplines scientifiques aient constitué des organes internationaux), l'Internet a fourni aux chercheurs en mathématiques du monde entier un nouveau support de choix pour se poser mutuellement des questions et s'échanger des versions préliminaires de textes qui souvent peuvent ne pas être destinés à la publication sous quelque forme que ce soit. Il est de plus en plus courant que les mathématiciens diffusent en libre lecture sur des sites Web les versions non encore imprimées de leurs travaux, bien avant qu'ils ne soient officiellement publiés.

Cette pratique soulève tout naturellement une foule de questions inédites et essentielles sur la gestion de base de la recherche scientifique, sur l'importance des maisons d'édition et sur toutes les formes plus officielles de publication et d'archivage des textes scientifiques. La simple mise à disposition d'un article sur un site Web doit-elle être considérée comme une forme de publication destinée à conférer la priorité scientifique (question rendue plus complexe encore par le fait que l'auteur peut à tout moment apporter des modifications à un tel article) ? Deviendra-t-il bientôt monnaie courante que la publication officielle de l'article soit sans rapport avec la communication de son contenu mathématique et qu'elle ait uniquement de l'importance pour la carrière de son auteur et pour l'archivage de l'article ? La publication sur le Web entraînera-t-elle la disparition des revues mathématiques classiques et de leurs éditeurs et conduira-t-elle à un « monde parfait » dans lequel les mathématiciens finiront en fait par publier eux-mêmes leurs ouvrages sur l'Internet ?

Il est impossible de prévoir actuellement les réponses à ce type de questions ou de prédire avec une quelconque certitude toutes les ramifications de la révolution électronique pour la publication mathématique. Néanmoins, ces questions

illustrent clairement l'intensité des mutations que connaît actuellement la publication de travaux dans la recherche mathématique.

Le rôle des bases de données

Dans cette partie, nous étudierons d'une part, pourquoi les chercheurs en mathématiques ont besoin de bases de données et d'autre part, pourquoi ce besoin est devenu plus aigu aujourd'hui, en grande partie grâce à la révolution électronique dans l'édition.

Principes de base de la rédaction en mathématiques

Il faut souligner que le besoin de bases de données dans la recherche mathématique ne date pas d'aujourd'hui. En effet, on peut dire premièrement que les mathématiques constituent la plus ancienne science exacte et que l'une de ses principales caractéristiques, qui se vérifie probablement moins dans les autres disciplines scientifiques modernes, réside dans le fait que les documents publiés il y a de nombreuses années peuvent s'avérer tout aussi pertinents pour un travail de recherche actuel que des mémoires appartenant à la bibliographie contemporaine. Deuxièmement, il ne faut pas oublier que la recherche de preuves rigoureuses est au centre de la recherche mathématique. Par conséquent, l'un des principes de base dans la rédaction d'un mémoire mathématique consiste depuis toujours pour l'auteur, pour tous les résultats majeurs dont il fait usage sans en donner lui-même une démonstration détaillée, à indiquer expressément ses sources dans la bibliographie existante. Troisièmement, les auteurs de mémoires mathématiques considèrent depuis toujours qu'ils se doivent absolument de citer et de prendre en compte l'ensemble des ouvrages précédemment publiés en relation étroite avec leur sujet. Cette obligation de base demeure l'un des piliers de la recherche mathématique aujourd'hui et sa disparition ferait un tort incalculable à une discipline qui s'enorgueillit de la rigueur et de la cohérence des nouvelles théories qu'elle crée.

Nouvelles tendances dans l'édition

Avec l'augmentation extraordinaire du nombre de revues et de toutes les formes de publications, résultat de la croissance du nombre de chercheurs en mathématiques dans le monde et de la révolution électronique dans l'édition, ce besoin historique de bases de données s'est accru. Le monde traditionnellement confortable de la recherche mathématique qui prévalait encore autour de 1980, dans lequel les mathématiciens pouvaient se tenir informés des ouvrages parus dans leur domaine simplement par l'étude régulière à la bibliothèque universitaire d'une liste relativement invariable de revues, solidement établies pour nombre d'entre elles, s'est maintenant profondément modifié sous de nombreux aspects.

Car il est un fait que le nombre de revues publiées ces vingt dernières années s'est considérablement accru.

La raison de cette augmentation est plus complexe que la simple croissance régulière du nombre de chercheurs en mathématiques. Elle s'explique d'abord certainement par le profit que les éditeurs du commerce et les sociétés savantes ont retiré de la publication, suite à l'introduction de la composition électronique. Elle s'explique ensuite par la spécialisation croissante de la recherche scientifique, qui a entraîné la création de nombreuses revues consacrées à des sous-domaines plutôt restreints des mathématiques. Elle s'explique enfin par la pression incessante exercée sur tous les universitaires pour que par des publications ils fassent concrètement la preuve de l'avancement de leurs recherches à toute une hiérarchie d'administrations chargées de les évaluer, depuis celle de leur propre université jusqu'à celle des organes délivrant les bourses de recherche et des organismes de financement de l'État. Il est inévitable que ces administrations en viennent à juger ce qu'elles considèrent être de bons résultats de recherche suivant des normes statistiques, des normes qui privilégient fortement le volume et la régularité des publications, plutôt que la qualité dans tous les sens absolus du terme. Cependant, plus que jamais, dans la recherche mathématique, comme dans tous les domaines de la science pure, il est un fait que ce sont les rares mémoires vraiment originaux qui font réellement avancer un sujet et qui ouvrent de nouveaux horizons essentiels pour la recherche.

Histoire des bases de données mathématiques

Avant de décrire brièvement la manière dont le chercheur en mathématiques utilise aujourd'hui une base de données électroniques, il nous semble important de retracer rapidement l'histoire de ces dernières dans ce domaine.

La première base de données mathématiques a été le *Jahrbuch über die Fortschritte der Mathematik*. Publiée de 1868 à 1942, elle était conçue de manière à ce que chaque volume contienne toutes les publications de la recherche mathématique d'une année donnée. Les documents publiés dans le *Jahrbuch* étaient d'une grande importance pour les mathématiciens et ils le sont encore aujourd'hui, comme semble le démontrer la préparation actuelle d'une version électronique (voir plus loin, pour plus de détails). Toutefois, le nombre sans cesse croissant de travaux mathématiques publiés chaque année a amené des retards dans la fabrication du *Jahrbuch* et en 1931, la maison d'édition Springer Verlag a lancé la publication de la revue de synthèse intitulée *Zentralblatt für Mathematik und ihre Grenzgebiete*.

A l'instigation de nombre de mathématiciens allemands émigrés, cette revue a été suivie en 1940 par la publication de la revue de synthèse *Mathematical Reviews* sous l'égide de la Société mathématique américaine. Ces deux revues de synthèse ont commencé à se transformer en de véritables bases de données électroniques

dans les années 1980 et constituent aujourd'hui les deux principales bases de données mathématiques dans le monde. Elles seront décrites plus en détail dans la suite.

Le passage à l'électronique

A cause des méthodes de recherche efficaces qu'elles offrent, la plupart des mathématiciens actuels utilisent déjà les bases de données dans leur version électronique. Habituellement, le mathématicien professionnel aura tendance à utiliser la base de données pour trouver ou vérifier la référence exacte d'un mémoire qu'il doit consulter ou citer. Dans ce cas, il peut généralement faire une recherche en fonction du nom de l'auteur et du titre du mémoire. Lorsqu'il existe une recension ou un résumé dans la base de données, cela peut entre autres s'avérer utile pour décider s'il est ou non nécessaire de consulter l'article dans sa totalité et de faire une visite à la bibliothèque ou de télécharger la version électronique du texte (disponible par le biais d'un abonnement à la revue ou à la carte). Parfois, il se peut que le chercheur connaisse seulement le nom de l'auteur ou qu'il ait confusément l'impression que tel ou tel auteur pourrait avoir rédigé un article en rapport avec son projet. Dans ce cas, il peut chercher dans la base de données tous les articles écrits par l'auteur concerné sur une période donnée. Nombre de perfectionnements sont déjà disponibles pour affiner la recherche. Par exemple, il est possible de préciser une catégorie dans une nomenclature de sujets ou de spécifier si la publication est un ouvrage, une revue ou un rapport. Il semble que ce ne soit plus qu'une question de temps avant que les mathématiciens n'utilisent les bases de données littéralement pour feuilletter la bibliographie concernant un domaine ou un problème particulier.

Critères de qualité d'une base de données

Il est clair que l'utilité d'une base de données dépend essentiellement de trois critères :

- Premièrement, elle doit être aussi complète que possible, c'est-à-dire contenir pratiquement toutes les publications mathématiques qui paraissent dans le monde sur un nombre d'années donné. Compte tenu du fait que la plupart des revues paraissent maintenant sous forme électronique, il devient assurément de plus en plus facile de satisfaire ce critère de base. Par conséquent, il est maintenant techniquement aisé pour un éditeur de transférer en partie ou en totalité le contenu d'une revue particulière vers une base de données (par exemple, l'éditeur peut commencer par transférer la table des matières, avec un résumé de l'auteur pour chaque article et ensuite le texte dans son entier si nécessaire pour des recensions ultérieures). S'il est vrai que les éditeurs se constituent de petites archives électroniques couvrant leurs

propres publications, il va de soi que des archives aussi limitées ne peuvent jamais remplacer une base de données réellement mondiale.

- Deuxièmement, elle doit disposer de méthodes de recherche efficaces et flexibles. En fait, on peut sans exagérer comparer le rôle des méthodes de recherche dans une base de données à la moelle épinière d'un être vivant. De nombreux travaux, et plus particulièrement ceux de la cellule MathDoc à Grenoble (voir plus loin), sont actuellement en cours pour développer un logiciel destiné à améliorer la métastructure des méthodes de recherche pour bases de données mathématiques.
- Troisièmement, les documents stockés dans la base de données sous une rubrique particulière doivent être sous une forme correcte et facilement exploitable par les méthodes de recherche et par le mathématicien qui consulte la base. Ce troisième critère rend absolument nécessaire l'intervention directe de l'homme dans la classification et l'organisation des documents dans cette dernière. Des débats ont eu lieu sur la possibilité de créer des bases de données scientifiques simplement par des machines réalisant des analyses statistiques de la totalité du texte de chaque article (les « recherches en texte libre »), mais il paraît inconcevable qu'un tel programme puisse fonctionner pour les mathématiques, notamment pour établir la classification du thème principal d'un texte ou pour choisir les articles qui seront recensés par une autre personne que l'auteur.

A qui appartiennent les bases de données ?

Il convient de souligner que les très grandes bases de données électroniques créées aujourd'hui dans une discipline scientifique particulière auront une durée de vie très longue et qu'il pourrait très bien ne plus jamais être très facile de créer de nouvelles bases fondamentalement différentes à un stade ultérieur d'évolution de cette discipline. Il est par conséquent très probable que ces bases de données finiront elles-mêmes par exercer insensiblement une influence subtile et notable sur l'évolution à long terme de la recherche dans le domaine concerné. Le fait de savoir à qui elles appartiennent s'avère donc être l'une des principales questions de principe à long terme pour tous les domaines essentiels de la recherche scientifique. Cela se vérifie particulièrement pour les mathématiques, où le recours aux références passées est la pierre angulaire de la discipline. Les scientifiques doivent par ailleurs être véritablement familiarisés avec la manière dont leurs bases de données sont construites, ce qui sera possible seulement s'ils ont leur mot à dire dans leur exploitation. On observe déjà sur l'Internet comment une requête rigoureusement identique peut donner des résultats tout à fait divergents selon la manière dont les différents moteurs de recherche structurent leurs données.

L'édification et la maintenance de grandes bases de données sont indéniablement coûteuses (voir plus loin l'analyse des coûts actuels de *Mathematical Reviews* et *Zentralblatt-MATH*). Il s'avère par ailleurs que les intérêts commerciaux jouent un rôle important dans l'exploitation des bases de données actuelles et qu'il continuera d'en être ainsi dans le futur. Toutefois, deux principes s'imposent presque automatiquement pour l'avenir des bases de données scientifiques à long terme.

Le premier est que la propriété de toutes les grandes bases de données devrait être clairement dévolue à la communauté scientifique responsable à l'origine d'une partie importante du matériel qu'elles renferment. Ce principe n'exclut bien sûr en aucune manière la participation active des partenaires commerciaux dans l'exploitation et la diffusion quotidiennes de la base de données. Deuxièmement, l'abonnement à une base de données devrait être fixé à un niveau raisonnable à l'issue d'une analyse transparente des coûts et les membres de la communauté scientifique qui l'utilisent devraient être préparés à supporter eux-mêmes ces coûts. Des fonds publics spéciaux pourront s'avérer de temps en temps nécessaires pour modifier ou actualiser la base. En dernière analyse, les bases de données connaîtront un plein essor uniquement avec la participation pleine et entière, autant sur le plan scientifique que financier, des scientifiques qui les utilisent.

Qui contrôle les bases de données mathématiques ?

La situation spécifique aux mathématiques est la suivante.

Mathematical Reviews est la propriété de la Société mathématique américaine et appartient donc en réalité à la communauté mathématique nord-américaine.

De nouveaux accords sont intervenus très récemment concernant la propriété de *Zentralblatt-MATH*. A l'avenir, quatre entités détiendront le droit d'auteur et le contrôle de cette base, à savoir Springer Verlag, FIZ Karlsruhe, l'Académie de Heidelberg et la Société mathématique européenne. Ces quatre organes ont convenu conjointement d'entreprendre tout ce qui était en leur pouvoir pour développer *Zentralblatt-MATH* de manière à en faire une base de données mathématiques internationale qui soit une émanation de la communauté mathématique européenne. Une étape importante a été franchie dans cette direction en France avec la création à Grenoble de la cellule MathDoc, sous la forme d'une coentreprise entre le CNRS, le ministère de l'Éducation et l'Université Joseph Fourier. Cette entité a réalisé d'importants travaux concernant le développement d'un logiciel pour bases de données (voir plus loin). Conformément à sa mission nationale en France, elle a par ailleurs amélioré le flux de publications françaises vers *Zentralblatt-MATH* et l'accès à cette dernière pour les institutions françaises (en particulier, elle a incité de petits groupes de chercheurs en mathématiques à former des consortiums pour l'achat d'abonnements). Ce n'est qu'une première étape et l'objectif

est de compter à terme des entités de production similaires dans de nombreux pays d'Europe.

On a bon espoir que d'importants domaines de collaboration se dégagent à l'avenir entre *Mathematical Reviews* et *Zentralblatt-MATH*. Toutefois, il s'avère par ailleurs que les mathématiciens et la recherche mathématique ne peuvent que bénéficier de deux bases de données exceptionnelles, qui si elles diffèrent par leur tradition et leur histoire, appartiennent aux deux plus importantes communautés mathématiques dans le monde et sont partout au service de tous les mathématiciens.

Le rôle des éditeurs et des bibliothèques

Incertitudes sur l'avenir de l'édition des revues

La révolution électronique dans l'édition s'est traduite par de nombreux avantages pour les éditeurs, mais elle pose aussi des défis majeurs, car il est difficile de prévoir toutes ses ramifications. Il est clair que, pour ce qui est de la publication de revues, l'avenir sera en grande partie déterminé par l'électronique. Ainsi, nous ne savons guère à l'heure actuelle combien de temps encore il sera économiquement possible pour les éditeurs d'imprimer et de diffuser des exemplaires papier de la plupart de leurs revues. D'un autre côté, la publication de revues exclusivement en ligne soulève une foule de questions fondamentales, tant pour les éditeurs que pour les utilisateurs. Il y a pour l'éditeur un danger évident de fortes réductions du nombre d'abonnements aux revues en ligne, particulièrement si les bibliothèques universitaires d'un pays ou d'une région commencent à regrouper leurs abonnements. Mais il existe également de sérieux inconvénients pour l'utilisateur. Par exemple, ce dernier est-il assuré par le règlement d'un abonnement à une revue en ligne pour une période donnée de pouvoir encore accéder à cette revue même après expiration de l'abonnement ? Compte tenu du besoin constant de consulter et de citer les anciens ouvrages comme expliqué précédemment, cette dernière question est primordiale pour les revues mathématiques.

Collaboration entre éditeurs et bases de données

Il semble que la meilleure réponse que les éditeurs puissent faire à long terme face à nombre de ces difficultés soit d'accepter le principe d'une étroite collaboration avec de grandes bases de données centrales appartenant à la communauté scientifique concernée et échappant par conséquent au cycle normal de la concurrence économique.

Le scénario idéal d'une telle collaboration pourrait se dérouler comme suit. A la date de publication de la revue, l'éditeur pourrait transférer une quantité restreinte du contenu (par exemple, une table des matières et un résumé rédigé par

l'auteur ou la première page de chaque article) vers la base de données centrale, laquelle mettrait immédiatement ces informations à la disposition de tous ses abonnés dans le cadre d'un service d'informations courantes. Ces informations seraient par ailleurs utilisées pour le fonctionnement normal de la base de données, notamment pour la classification des documents et le choix des recenseurs le cas échéant. Un tel transfert ne compromettrait aucunement les intérêts commerciaux vitaux de l'éditeur, car ceux qui souhaiteraient voir le texte de la revue dans son entier devraient alors s'acquitter d'un abonnement normal. On peut aussi estimer que le fait de placer la table des matières de la revue dans la section des informations courantes de la base de données constituerait pour celle-ci une bonne publicité.

Enfin, après que suffisamment de temps se soit écoulé pour être certain qu'il ne va pas perdre des abonnements (une période de plusieurs années, par exemple), l'éditeur pourrait transférer la totalité du texte de la revue vers la base de données centrale. Une fois cette opération effectuée, il incomberait alors à celle-ci de faire en sorte que la revue soit conservée indéfiniment sous forme électronique. *Zentralblatt-MATH* prévoit justement de lancer un tel programme en collaboration avec la Société mathématique européenne. On espère que ce programme, baptisé *Current Awareness Programme* (CAP), évolue rapidement de manière à rassembler de nombreux éditeurs.

Bibliothèques et archivage

La révolution électronique dans l'édition aura par ailleurs sans aucun doute de profondes répercussions sur les bibliothèques.

Premièrement, la nécessité d'archiver les ouvrages scientifiques électroniques s'avère encore plus cruciale lorsqu'il n'y a pas d'équivalent imprimé. Même si certaines maisons d'édition décident de conserver des archives détaillées des revues électroniques qu'elles produisent, les entreprises d'édition ne durent pas éternellement et elles seront soumises à des forces commerciales touchant en grande partie les activités présentes et futures de l'éditeur. Ainsi, il serait très dangereux de confier l'archivage des documents électroniques à ce dernier plutôt qu'aux bibliothèques, car une partie d'entre eux serait perdue à long terme.

Deuxièmement, il semble, grâce à l'aisance procurée dans la diffusion des données, que la révolution électronique devrait faciliter la mise en commun de ressources par les bibliothèques. On peut par conséquent imaginer qu'à l'avenir une bibliothèque nationale ou une grande bibliothèque régionale puisse s'abonner à une revue scientifique pour la diffuser à tout un ensemble d'utilisateurs et d'institutions. Elle ferait alors office d'archives où serait conservé le texte complet de la revue. Tout ceci donne fortement à penser qu'une centralisation des bibliothèques est pour bientôt. Par ailleurs, il convient de signaler que les catalogues

des principales bibliothèques se transformeront pratiquement à coup sûr en bases de données électroniques et qu'il pourrait à terme être difficile de les différencier d'un ensemble de bases de données se recoupant et couvrant tous les principaux sujets du catalogue. En conséquence, il est très probable que les bases de données scientifiques électroniques et les bibliothèques les plus importantes collaboreront beaucoup plus étroitement et qu'il faudra élaborer avec soin des modèles d'accords pour leur coopération.

Enfin, les bibliothèques devront également affronter les problèmes complexes de propriété intellectuelle soulevés par la révolution électronique et en particulier statuer sur l'institution d'un fondement juridique concernant la publication, le droit d'auteur et le dépôt légal.

III. BASES DE DONNÉES DANS LES MATHÉMATIQUES

La communauté mathématique internationale a la chance de posséder deux bases de données totalement indépendantes, florissantes et exhaustives, la première, *Zentralblatt-MATH*, établie en Europe, et l'autre, *Mathematical Reviews*, en Amérique du Nord. Par l'un des curieux hasards de l'histoire des mathématiques au xx^e siècle, ces bases ont été créées par la même personne, Otto Neugebauer, universitaire ayant fui l'Allemagne du régime nazi pour les États-Unis au milieu de sa carrière.

Aujourd'hui, force est de reconnaître que ces bases s'opposent et se complètent tout à la fois dans des domaines qui n'affectent pas leur souveraineté respective. Les paragraphes suivants donnent une description de ces bases dans leur état actuel et de certains projets connexes.

Zentralblatt-MATH* et *Mathematical Reviews

Comme indiqué précédemment, *Zentralblatt für Mathematik und ihre Grenzgebiete* a été fondée en 1931. Elle a récemment été rebaptisée *Zentralblatt-MATH* pour souligner sa transformation en base de données électronique à part entière. Publiée par les éditions Springer Verlag, elle est mise en forme par la Société mathématique européenne, FIZ Karlsruhe et l'Académie des sciences de Heidelberg. Fondée en 1940, *Mathematical Reviews* est publiée et mise en forme par la Société mathématique américaine. Ces deux revues ont commencé à se transformer en bases de données électroniques dans les années 80.

Service particulier du FachInformationZentrum (FIZ), dont le siège est à Karlsruhe, le centre d'opérations de *Zentralblatt-MATH* est établi à Berlin. *Mathematical Reviews* opère quant à elle depuis Ann Arbor, dans le Michigan. C'est à ces adresses qu'il est procédé à l'analyse et à la recension de l'ensemble des revues et des

ouvrages reçus. Ces bases s'efforcent d'être aussi exhaustives que possible et de couvrir tous les domaines des mathématiques pures et appliquées, mais aussi certaines matières mathématiques de disciplines afférentes telles que l'informatique, la physique mathématique, l'économie, l'ingénierie et les sciences biologiques. L'analyse des données brutes reçues est effectuée par des équipes de rédacteurs scientifiques qui travaillent respectivement à Berlin et à Ann Arbor. La première de leurs tâches consiste à affecter à chaque article reçu un code standard de classification.

Heureusement, depuis 1971, les deux bases ont collaboré à la mise au point d'un système commun de classification, connu sous le sigle MSC (Mathematics Subject Classification). Les éditeurs décident de la manière dont chaque article doit être traité dans la base de données, par exemple s'il doit faire l'objet d'une recension indépendante, d'une recension par l'auteur ou simplement d'une sorte de résumé standard. Les deux bases s'efforcent de disposer le plus possible de recensions indépendantes qui, il faut le souligner, sont rédigées par des recenseurs originaires de toutes les régions du monde. En annexe à ce document, on trouvera la répartition géographique par pays des quelque 5 000 recenseurs actuels de *Zentralblatt-MATH*. La contribution de tous les recenseurs à la constitution de ces bases de données est immense et c'est précisément ce qui les rend universelles et de la plus grande utilité pour les mathématiciens dans leurs recherches.

Pour chaque base, ces travaux se traduisent chaque année par la publication de plus de 60 000 notices, lesquelles constituent des sources d'informations vitales sur des articles tirés de plus de 2 000 revues, périodiques, ouvrages et comptes rendus de conférences.

Les deux bases sont accessibles sur trois supports et plus précisément, la traditionnelle version imprimée, l'accès en ligne et le CD-Rom. Actuellement, le prix de *Zentralblatt-MATH* (pour deux supports parmi les trois) est environ de 10 000 deutsche marks par an et celui de *Mathematical Reviews* (pour deux supports parmi les trois) avoisine 6 500 dollars des États-Unis.

Intégration des anciennes publications mathématiques dans les bases de données électroniques

A ce jour, le plus important projet d'intégration d'anciennes publications mathématiques dans une base de données électronique mais aussi d'archives électroniques de documents mathématiques classiques est le *Jahrbuch über die Fortschritte der Mathematik*. Financé par la Deutsche Forschungsgemeinschaft (DFG), il est supervisé par l'Université technique (TU) de Berlin (représentée par B. Wegner) et la Bibliothèque municipale et universitaire (SUB) de Göttingen

(représentée par E. Mittler), en étroite collaboration avec *Zentralblatt-MATH*. Pour simplifier la lecture, nous l'appellerons le Projet JFM.

Premières archives fiables et exhaustives des publications mathématiques, le *Jahrbuch über die Fortschritte der Mathematik* préfigure les bases de données modernes. De 1868 à 1942, il est paru un volume par an, chacun tentant de recenser tous les ouvrages de recherche mathématique parus dans l'année correspondante. Au total, le *Jahrbuch* couvre près de 200 000 articles. La mise en forme rédactionnelle a été réalisée par l'Académie des sciences de Prusse avec l'aide d'un bureau international de recenseurs et l'édition confiée aux éditions Walter de Gruyter. Jusqu'à l'arrivée de *Zentralblatt-MATH* en 1931, le *Jahrbuch* est demeuré la seule publication d'archives mathématiques. Le fait qu'il n'existe aucune publication d'archives avant 1868 rend évidemment très difficile l'intégration d'anciennes publications dans une base de données électronique.

Pour de nombreuses raisons, la version imprimée du *Jahrbuch* ne peut plus satisfaire les besoins actuels des chercheurs ainsi que des historiens en mathématiques. D'une part, il n'est pas aisé de trouver une édition imprimée complète du *Jahrbuch*, car la publication a cessé depuis longtemps. D'autre part, si chaque volume contient bien un index des auteurs, il n'existe pas de catalogue global et l'indexage par sujet est très incomplet. Enfin, aucune information n'est donnée sur l'importance d'un article pour l'évolution des recherches mathématiques.

Le projet JFM comporte trois étapes. La première étape vise la constitution d'une base de données électronique à partir des documents contenus dans le *Jahrbuch*. Cette base aura la même structure que la version électronique de *Zentralblatt-MATH* et les utilisateurs pourront accéder simultanément aux deux services. Les directeurs de rédaction de cette base sont K. Dennis (Université de Cornell) et B. Wegner (Université technique de Berlin). La deuxième étape consiste, avec l'assistance d'un groupe d'experts internationaux, à faciliter l'utilisation de la base du *Jahrbuch*, notamment par l'établissement de liens internes entre les données, l'affectation de mots clés en anglais et d'un code de classification MSC à chaque notice, la normalisation des bibliographies, l'indication de l'emplacement des textes complets des notices et, le cas échéant, l'ajout de commentaires sur la valeur d'une notice particulière pour l'évolution des mathématiques. On espère qu'il sera possible à l'avenir de continuer d'actualiser et d'améliorer tous ces éléments, car les bases de données doivent évoluer comme des organismes vivants, si l'on veut qu'elles restent d'une utilité maximale pour les recherches du moment. La troisième étape consistera à identifier les notices recensées dans le *Jahrbuch* qui se sont avérées particulièrement importantes pour la poursuite de la recherche mathématique. On espère en outre identifier les documents difficiles à trouver sous forme imprimée (par exemple, plusieurs revues publiées sur la période allant de 1868 à 1942 utilisaient du papier de qualité tellement médiocre que l'on doute qu'il se soit conservé à long terme). Il est prévu de numériser le texte intégral de ces documents, soit environ

dix pour cent de toutes les notices recensées par le *Jahrbuch*. Les archives électroniques de ces documents sont conservées à la SUB (bibliothèque municipale et universitaire) de Göttingen et seront reliées à la base de données du *Jahrbuch*.

Le projet JFM illustre parfaitement la manière dont les fonds publics peuvent contribuer à la constitution d'une base de données scientifiques. Les éditeurs du commerce n'étaient pas intéressés par un projet de ce type pour la simple raison qu'il aurait conduit avec une quasi-certitude à un produit trop onéreux pour une large diffusion au sein de la communauté mathématique. Par le recours aux fonds publics, on espère faire bénéficier tous les mathématiciens du projet JFM à un coût très modéré, à la fois sous la forme d'un service en ligne et d'une base de données autonome stockée sur CD-Rom.

Logiciel et développement de méthodes de recherche efficaces

Nous avons vu plus haut que le logiciel qui sert non seulement à lancer des procédures de recherche efficaces et flexibles, mais aussi à intégrer de nouvelles données dans une base électronique, joue un rôle similaire à la moelle épinière chez un être vivant. Ce logiciel doit être adapté aux besoins des chercheurs en mathématiques et doit être conçu de manière à pouvoir évoluer avec les changements de modes et d'habitudes dans ce domaine. Comme les spécifications d'un tel logiciel doivent être induites par ses utilisateurs, il doit demeurer dans le domaine public et ne pas appartenir à des entités commerciales ni être sous le contrôle d'un quelconque monopole.

Spécifications du logiciel de consultation

Les logiciels existants peuvent aisément traiter des bases de données de la taille de *Zentralblatt* et de *Mathematical Reviews*. Pour le logiciel de consultation, la taille n'est pas véritablement un problème. Mais outre l'aspect général de convivialité, plusieurs caractéristiques sont d'une importance cruciale. Premièrement, il doit pouvoir fonctionner avec les logiciels standard (par exemple, «HTML», «postscript» et «PDF» pour les utilisateurs de l'Internet et «TEX», «BIBTEX» et, dans un proche avenir, «mathml», pour l'écriture mathématique). Deuxièmement, il doit pouvoir fonctionner sous différents systèmes d'exploitation, notamment «Unix» et «Windows NT». Troisièmement, il doit être suffisamment flexible pour que les utilisateurs puissent développer des aspects liés à leurs propres besoins et systèmes d'exploitation et alimenter, grâce à ces développements, les recherches de l'équipe responsable du programme principal. L'évolution de TEX ces vingt dernières années est un excellent exemple d'interactions de ce type entre utilisateurs et créateurs. Enfin, il doit être possible de l'exploiter sur un large éventail de matériels informatiques.

Enrichissement permanent de la base de données

Le logiciel doit par ailleurs être conçu de manière à s'adapter aux diverses extensions que subissent au cours du temps les données sur lesquelles il travaille. Ainsi, la base de données sera enrichie en permanence de nouveaux documents issus des revues mathématiques. Quelquefois, ce seront même des bases de données entières, comme dans le cas de la version électronique du *Jahrbuch* actuellement en préparation. Le logiciel devra également accepter de nouveaux champs de recherche et notamment distinguer entre recensions et résumés et pouvoir insérer l'intégralité du texte original ou de la traduction d'une notice. Le logiciel devrait par ailleurs prendre en compte les nouvelles possibilités de création de liens hypertexte entre notices au sein d'une même base de données ou avec d'autres bases de données, notamment avec les versions électroniques des revues.

Travaux de la cellule MathDoc

Les travaux de la cellule MathDoc à Grenoble sont axés sur la conception d'un logiciel correspondant aux exigences décrites ci-dessus. Jusqu'ici, elle s'est concentrée sur le développement des méthodes de recherche et des types de requêtes susceptibles d'être intégrées dans la base de données. La version actuelle du logiciel peut être utilisée sur tout un éventail de matériels et de systèmes d'exploitation, comme le montre clairement son installation sur un bon nombre de sites dans le monde entier. La prochaine étape doit consister à améliorer son adaptabilité au plan local, afin de permettre les interactions avec les administrateurs des serveurs locaux et les bases de données locales. On espère que cette deuxième étape sera franchie d'ici fin 1998 et que le nouveau logiciel sera rapidement placé dans le domaine public. De nombreux efforts sont par ailleurs déployés pour créer un logiciel qui simplifie l'entrée des documents dans la base de données. Pour l'essentiel, on espère qu'il permettra entre autres aux éditeurs d'entrer automatiquement la table des matières et les résumés des articles de leurs revues dans la base de données. Différentes agences de rédaction pourront ainsi travailler simultanément sur le contenu de cette dernière sans nuire à sa cohérence et à son homogénéité sur le plan scientifique. Enfin, il devra pouvoir aussi s'adapter à des pratiques de publication mathématique sans cesse en évolution et faire ainsi en sorte que toutes les formes de publication se retrouvent finalement dans la base de données aussi vite que possible.

Annexe

Recenseurs de Zentralblatt-MATH par pays

Dans le monde, on compte plus de 5 000 recenseurs (qui traitent environ 60 000 notices par an).

	Nombre de recenseurs		Nombre de recenseurs
Allemagne	1 302	Taipei chinois	22
États-Unis	1 227	Mexique	21
Fédération de Russie	353	Singapour	18
Roumanie	333	Cuba	15
France	313	Hong-Kong, Chine	15
Pologne	302	Turquie	15
Royaume-Uni	255	Latvija	14
Japon	251	Portugal	14
Italie	246	Viêt-nam	14
Inde	223	Croatie	12
Chine	215	Irlande	12
Canada	208	Moldova	12
République tchèque	197	Égypte	11
Ukraine	118	Corée	11
Pays-Bas	116	Arabie Saoudite	10
Autriche	113	Nigeria	9
Espagne	108	Ouzbékistan	9
Hongrie	105	Arménie	8
Bulgarie	101	Azerbaïdjan	8
Australie	93	Koweït	8
République slovaque	74	Kazakhstan	7
Suisse	69	Venezuela	7
Grèce	68	Algérie	5
Belgique	64	Slovénie	5
Bésil	59	État de Brunei Darussalam	4
Géorgie	52	Chili	4
Israël	49	République islamique d'Iran	4
Suède	46	Tunisie	4
Bélarus	44	Chypre	3
États de l'ex-Yougoslavie	34	Islande	3
Danemark	33	Tadjikistan	3
Nouvelle-Zélande	31	Colombie	2
Estonie	30	Jordanie	2
Norvège	27	Luxembourg	2
Lituanie	26	Mongolie	2
Afrique du Sud	24	Maroc	2
Argentine	22		

Ainsi que de nombreux pays avec un seul recenseur.

L'INFRASTRUCTURE DE L'INFORMATION EN SCIENCES SOCIALES

Table des matières

I. Problématique de l'infrastructure de l'information	142
II. L'infrastructure matérielle	143
III. L'infrastructure pour les sciences sociales	146
IV. Réseau international de sciences sociales intégrées	150
V. Niveaux d'investissements	152
VI. Conclusion	154

Cet article a été rédigé par William Sims Bainbridge, Ph.D, de la National Science Foundation (e-mail : wbainbri@nsf.gov). Les idées qu'il exprime ne reflètent pas nécessairement les points de vue de la National Science Foundation ou des États-Unis.

I. PROBLÉMATIQUE DE L'INFRASTRUCTURE DE L'INFORMATION

Il n'est pas exagéré de dire que l'infrastructure de l'information aux États-Unis aujourd'hui connaît à la fois une croissance explosive et une transformation révolutionnaire. La vaste mutation en cours dans le domaine des communications, de l'informatique et des sciences de l'information aborde une nouvelle phase, où les avancées vont de plus en plus dépendre de l'apport des sciences sociales et comportementales, tandis que les technologies de pointe vont transformer les méthodes des sciences humaines elles-mêmes. Dans cet environnement en pleine évolution, la définition même de l'infrastructure change.

Traditionnellement, le terme d'*infrastructure* connaît deux acceptions. En génie civil aussi bien que militaire, l'*infrastructure* désigne les installations physiques permanentes nécessaires aux activités essentielles. Ainsi, l'infrastructure d'une ville comprend les réseaux de voirie, d'assainissement, d'électricité et de téléphone. En sciences sociales, l'infrastructure désigne plus généralement les structures ou les ressources fondamentales partagées par une collectivité ou une organisation, et qui englobent des biens publics physiques et non physiques. C'est ainsi que l'infrastructure éducative ne comprend pas seulement les bâtiments scolaires, mais aussi les manuels, l'expertise des enseignants, l'organisation des établissements d'enseignement et les valeurs culturelles de la collectivité sur lesquelles s'appuie l'éducation.

Appliqué à la technologie de l'information, le terme d'*infrastructure* a parfois le sens étroit de grandes installations matérielles communes, telles que câbles et routeurs de réseaux, ordinateurs centraux ou superordinateurs. Cependant, l'informatique nous l'a appris, le logiciel est tout aussi important que le matériel. La question qui se pose aujourd'hui aux gouvernements est de savoir sur quoi axer leurs gros investissements pour créer les environnements informatiques et de communication dont on va avoir besoin dans l'avenir. Dans ce cas, l'infrastructure doit être considérée au sens large, comme englobant le matériel, les logiciels, les bases de données polyvalentes destinées à des clientèles multiples, et les organisations sociales qui les créent et les appliquent.

La révolution de l'information aujourd'hui est tout aussi commerciale que scientifique, et la privatisation de l'Internet montre bien qu'une grande partie de l'infrastructure matérielle de l'information de l'avenir relèvera non plus des États mais du secteur privé. Aux États-Unis, plusieurs sociétés offrent maintenant des

services intégrés – téléphone, télévision par câble et Internet – et la demande de fréquences va exploser avec la télévision haute définition et la musique et les films payables à la séance. Le rôle de l'État consiste à financer la recherche fondamentale qui permettra de développer les infrastructures de l'information du futur et à se servir de la communication scientifique pour expérimenter et mettre au point les technologies nécessaires.

Du point de vue des sciences sociales, il est certes extrêmement intéressant de pouvoir assurer des communications rapides par Internet entre grandes universités et établissements de recherche, mais, au-delà, le principal intérêt est de faire évoluer le monde vers le moment où toutes les écoles primaires, toutes les entreprises et tous les particuliers bénéficieront de ce même accès. Le laboratoire de recherche des sciences sociales étant la société en général, le champ de collecte des données est particulièrement vaste ; ces données n'intéressent pas seulement les étudiants des grandes universités, mais aussi les responsables politiques, les mass médias et le public en général. C'est pourquoi l'ère qui s'ouvre dans le développement de l'infrastructure de l'information dans le domaine des sciences sociales appelle une réflexion nouvelle et une recherche ambitieuse.

II. L'INFRASTRUCTURE MATÉRIELLE

L'infrastructure matérielle de l'informatique d'aujourd'hui est importante aux États-Unis, mais il convient de se demander avec quelle rapidité les capacités de ses différents éléments peuvent s'améliorer. L'État doit par exemple décider combien il va investir pour augmenter la rapidité des superordinateurs. L'industrie utilise largement l'informatique de pointe ; Boeing, par exemple, utilise des méthodes entièrement informatiques pour concevoir et mettre au point le prototype de ses nouveaux avions. Mais l'État en étant pratiquement le seul client, de nouvelles avancées dans ce domaine supposent qu'il consente d'importants investissements supplémentaires. Les applications scientifiques des superordinateurs sont notamment la modélisation de la formation des galaxies, l'analyse des processus complexes qui déterminent le climat et les conditions météorologiques, la conception de moteurs à haut rendement par la simulation précise de la combustion, et l'utilisation de techniques hautement itératives pour déterminer la structure conceptuelle de vastes ensembles de textes écrits. Avec le temps, les technologies mises au point pour les superordinateurs, telles que le parallélisme massif, s'étendront aux applications de l'enseignement et de l'industrie et aux produits de consommation mais si l'État n'en prenait pas l'initiative, les industriels ne mettraient peut-être jamais au point de tels systèmes.

Depuis peu, les ordinateurs les plus rapides du monde ont atteint la gamme dite des *teraflops*, avec des vitesses utiles soutenues de plus de mille milliards (10^{12})

d'opérations arithmétiques à virgule flottante par seconde. Le pionnier a été le *Department of Energy* avec son *Accelerated Strategic Computing Initiative*, mais la NASA, les *National Institutes of Health* et la *National Science Foundation* y coopèrent aussi. On approche d'une vitesse de quatre teraflops et l'on envisage sérieusement d'atteindre 40 teraflops d'ici cinq ans. En fait, l'objectif des études actuelles est de passer des teraflops à la *petascale*, ce qui implique non seulement un objectif à long terme de 10^{15} opérations à virgule flottante par seconde mais aussi un stockage de données à accès rapide de l'ordre de un trillion d'octets.

Les spécialistes des sciences sociales s'intéressent particulièrement à l'organisation sociale nécessaire pour utiliser efficacement ces machines. Récemment, la *National Science Foundation* a révolutionné l'organisation de ses centres de supercalcul pour créer des partenariats en réseau dont les centres nodaux sont dotés de superordinateurs, mais qui intègrent également toute une série d'équipements et d'équipes scientifiques réparties entre de nombreux points. Il existe aujourd'hui deux « *Partnerships of Advanced Computation Infrastructure (PACI)* ». L'un, la *National Computational Science Alliance*, est implanté dans l'ancien centre de supercalcul de l'Illinois bien connu pour avoir mis au point le premier système de navigation sur le Web, Mosaic, qui a précédé Netscape. L'autre, le *National Partnership for Advanced Computational Infrastructure*, se situe au vieux centre de supercalcul de San Diego. Chaque PACI est un réseau d'environ 50 universités et entreprises. Un chercheur situé dans l'un peut avoir besoin d'un logiciel établi dans un autre et exécuté sur un superordinateur dans un troisième pour analyser des données stockées dans un quatrième. Le passage en machine peut également se faire simultanément en plusieurs points.

Le *National Center for Atmospheric Research* de la *National Science Foundation* possède une infrastructure qui est un bon exemple de la diversité des grands équipements utilisés par la recherche scientifique, car elle comprend non seulement un superordinateur mais aussi les avions nécessaires pour recueillir les données. La recherche atmosphérique du XXI^e siècle nécessitera non seulement des superordinateurs pour la modélisation, mais aussi le recueil essentiellement en temps réel de plus d'un teraoctet de données par jour en provenance d'avions, de navires, de satellites et de stations terrestres et qui seront fournies aux scientifiques de par le monde. Comme le vent et la pluie, la science ne s'arrête pas aux frontières nationales.

On investit aussi beaucoup dans le développement de réseaux plus rapides. L'ARPANET d'origine utilisait des lignes téléphoniques ordinaires avec, pour le routage, des ordinateurs de conception existante mais spécialement adaptés, et l'explosion récente d'Internet et du World Wide Web a largement exploité les réseaux de communication existants. Cela dit, la création de l'Internet moderne a dès le début nécessité des efforts importants de recherche et développement, et l'explosion actuelle d'intérêt commercial n'a pas éliminé la nécessité de l'intervention du secteur public dans le développement des technologies de réseau. La NSF

et le gouvernement des États-Unis en général, sont déterminés à faire progresser les capacités des réseaux, entre autres les bandes passantes, sous des rubriques comme l'Internet de la prochaine génération. En outre, la NSF et d'autres agences vont encourager activement la recherche fondamentale dans des domaines aussi divers que la technologie des réseaux 100 % optiques et l'optimisation des barèmes de prix. Même si un jour l'industrie crée des réseaux qui répondent aux besoins commerciaux, il restera à déterminer qui paiera pour que les scientifiques puissent les utiliser.

La NSF a lourdement investi dans des efforts de développement liés au Net comme le banc d'essai gigabit, le matériel à mode de transfert asynchrone et les réseaux sans fil. La NSF a notamment constitué un partenariat avec MCI Corporation pour construire le *Backbone Network Service* à très haute performance (vBNS) destiné à relier les deux PACI à une centaine de grands établissements de recherche. Actuellement, le vBNS est en train de passer de 622 mégabits par seconde à 2.4 gigabits par seconde. En novembre 1997, Singapour y accédait officiellement, et assurer à tous les établissements scientifiques de par le monde des connexions fiables à haute bande passante doit être un objectif hautement prioritaire.

Sous la bannière d'Internet 2, l'*University Corporation for Advanced Internet Development* travaille avec plusieurs partenaires industriels – notamment Qwest, Nortel et Cisco – à la construction du réseau Abilene. L'objectif est de relier 60 universités en 1999, puis rapidement plus de 130, à 2.4 gigabits par seconde, avec des tronçons à 9.6 gigabits par seconde. Etant donné la taille géographique des États-Unis, il faudra plus de 13 000 miles de câble en fibre optique.

Naturellement, cela implique une coopération entre toutes sortes d'institutions appartenant à des secteurs aussi divers que l'éducation, l'industrie, le secteur public civil et militaire et les associations scientifiques. La création de l'Internet a été une retombée d'un projet militaire de la *Defense Advanced Projects Agency*, mis au point avec l'aide de la *National Science Foundation*, organisme civil. Ce genre d'application dérivée et de coopération se poursuit aujourd'hui. Ainsi, le *Naval Research Laboratory*, qui prend une part importante au développement de réseaux militaires de transmission de données sécurisées, travaille avec la *Shoah Visual History Foundation* de Steven Spielberg (VHF). L'objectif de cette dernière est de réaliser des enregistrements vidéo de récits de témoins directs de l'Holocauste et de constituer une archive informatisée pour les conserver et les rendre accessibles de façon analytiquement utile aux chercheurs, scientifiques, étudiants et simples citoyens. Plus de 46 000 entretiens ont été réalisés, et un système d'hypertexte multimédia permet à l'utilisateur de passer de fragments donnés d'un entretien à des documents écrits, cartes et fragments d'autres entretiens qui traitent du même sujet. Le *Naval Research Laboratory* a constaté que ce système se prêtait extraordinairement bien au test des capacités de ses réseaux, du fait qu'il exige une transmission souple de structures complexes de blocs de données très importants.

Comme on l'a déjà dit, certains types de logiciels doivent être considérés comme des éléments de l'infrastructure des technologies de l'information tout aussi importants que les grands équipements partagés. D'une façon générale, l'industrie a beaucoup moins investi dans la recherche fondamentale sur les logiciels que dans celle qui doit permettre la mise au point des équipements. Certes, le développement par Sun Microsystems du langage Java et la récente création d'un institut de recherche sur les logiciels par Microsoft indiquent que l'industrie est intéressée, mais l'État reste obligé d'investir beaucoup dans la recherche fondamentale et le développement de logiciels en tant que bien public. Le récent rapport d'étape de l'*Information Technology Advisory Committee* auprès du Président dit ceci : « Le logiciel est la nouvelle infrastructure physique de l'ère de l'information. Il est fondamental pour la réussite économique, la recherche scientifique et technique, et la sécurité nationale, et il est de plus en plus important pour le commerce, la communication, l'accès à l'information et l'infrastructure physique du pays. »

Une autre question, tout à fait différente, qui se pose au sujet de l'infrastructure matérielle, est celle de savoir dans quelle mesure les diverses sciences tirent un avantage de la mise en réseau des ressources partagées dans des dispositifs multi-utilisateurs et multi-usages virtuels, comme les « collaboratoires » en ligne, la télémanipulation d'instruments scientifiques et les liaisons hypermédias d'information entre divers fichiers de données. On peut aussi se demander dans quels types d'infrastructures non matérielles il faut investir en parallèle pour maximiser le rendement scientifique de l'infrastructure matérielle : fichiers de données massifs, bibliothèques de modules de logiciels, établissements de formation de chercheurs par exemple. Pour se faire une meilleure idée de ces questions, nous examinerons comment elles se posent dans le domaine des sciences sociales.

III. L'INFRASTRUCTURE POUR LES SCIENCES SOCIALES

Les problèmes d'infrastructure qui se posent aux sciences sociales donnent un bon aperçu des choix que toutes les sciences vont avoir à faire dans les années qui viennent. Elles manquent en effet cruellement de bases de données largement partagées et des collaboratoires basés sur le Web qui en permettraient la création et l'utilisation. Les chercheurs ont commencé à appliquer des méthodes expérimentales à l'étude de réseaux d'interactions humaines à grande échelle comme les marchés, les systèmes électoraux et les systèmes sociaux. La convergence de nouvelles puissances de calcul, de nouveaux outils logiciels et de l'extension des réseaux à large bande va désormais permettre de procéder sur Internet à des expériences à beaucoup plus grande échelle, portant sur des centaines, voire des milliers de sujets et s'étendant à travers les frontières et sur de longues durées, et ainsi de modéliser les relations du monde réel d'une façon beaucoup plus réaliste

qu'on ne pouvait le faire jusque-là. De même, la technologie de réalisation d'enquêtes sur Internet auprès de très grands nombres de personnes va révolutionner la recherche par enquête, si l'on peut mettre au point des techniques appropriées pour compenser le caractère non aléatoire des échantillons.

Une illustration impressionnante en est Survey 2000, questionnaire en anglais extrêmement complexe diffusé sur le réseau en novembre 1998 à plus de 50 000 personnes à travers le monde. Cette enquête, qui portait sur les migrations géographiques et les cultures régionales, questionnait les personnes interrogées sur leurs mouvements à intervalles de sept ans ; sur leurs goûts en matière de cuisine régionale, de musique et de littérature ; sur leur utilisation d'Internet et leur participation à la collectivité locale et leur demandait même de porter un jugement sur le style et la qualité d'une série de clips musicaux enregistrés. L'échantillon n'était pas aléatoire, mais un certain nombre d'éléments démographiques ont été pris en compte afin d'évaluer le caractère non aléatoire et de le compenser dans une certaine mesure. Les spécialistes de sciences sociales qui ont mis au point cette enquête et analysent actuellement le riche ensemble de données recueillies observent que, pour des raisons de coût et de complexité, les grandes enquêtes nationales traditionnelles n'auraient jamais permis de couvrir les mêmes sujets.

L'analyse des données est en cours, mais une tabulation provisoire indique que 40 747 résidents des États-Unis y ont répondu, plus 5 339 résidents canadiens. L'enquête avait fait l'objet d'une publicité considérable dans ces deux pays, mais un minimum de 100 réponses ont également été reçues de chacun des 33 pays suivants : Argentine (279), Australie (1 480), Belgique (159), Brésil (203), Chine (236), Chili (162), Colombie (174), Danemark (100), Finlande (128), France (368), Allemagne (417), Grèce (358), Hongrie (604), Inde (332), Irlande (203), Israël (299), Italie (430), Japon (275), Malaisie (283), Mexique (706), Pays-Bas (292), Nouvelle-Zélande (467), Norvège (146), Philippines (122), Portugal (111), Singapour (344), Afrique du Sud (382), Espagne (462), Suède (232), Suisse (184), Turquie (133), Royaume-Uni (1 431) et Venezuela (168).

Il y a quatre ans, les directions du calcul et des sciences sociales de la NSF ont conclu avec la *Defence Advanced Project Agency* et la NASA un partenariat pour le développement de la technologie nécessaire aux bibliothèques numériques. On imagine généralement les bibliothèques comme des lieux où l'on conserve livres et périodiques, mais en fait n'importe quel type de collection de données numériques utiles pour les scientifiques peut en constituer une. L'un des projets en cours de bibliothèques électroniques concerne essentiellement les sciences éducatives et les programmes d'informations télévisées, par exemple, tandis qu'un autre est centré sur la cartographie et un troisième archive et indexe le contenu des revues scientifiques. On pourrait imaginer une bibliothèque numérique anthropologique ou archéologique consacrée aux représentations tridimensionnelles d'artefacts. En

paléontologie, tous les spécimens types de toutes les espèces provenant de toutes les collections, grandes et petites, devraient être représentés sur le Web sous forme multidimensionnelle. Il est clair qu'à cet égard, des organisations non gouvernementales comme le *Getty Information Institute*, spécialisé dans l'histoire de l'art, et l'*American Association of Museums* joueront un rôle clé.

Une nouvelle phase de l'Initiative « bibliothèques électroniques » vient d'être lancée par la *National Science Foundation*, la *Defence Advanced Research Projects Agency*, la *National Library of Medicine*, la *Library of Congress*, la NASA et le *National Endowment for the Humanities*. D'autre part, à la NSF, la Division de l'information et des systèmes intelligents et la Division des programmes internationaux coopèrent à la mise au point d'une nouvelle Initiative « bibliothèques électroniques internationales » qui mettrait en place des partenariats binationaux pour financer la recherche sur des systèmes d'information capables de fonctionner dans de multiples langues, formats, médias et contextes sociaux et organisationnels.

Dans le système public partagé des États-Unis, le gouvernement fédéral n'est pas le seul en pointe. L'État de Californie vient d'entamer un programme extrêmement ambitieux avec le lancement le 20 janvier 1999 de la *California Digital Library*. Dans l'idée de Richard Atkinson, Président de l'Université de Californie, ce vaste projet va ouvrir une ère où nos bibliothèques viendront à nous, sur simple pression d'un bouton, où que nous soyons, à n'importe quel moment.

En sciences sociales, c'est peut-être l'infrastructure de données qui est la ressource scientifique répartie la plus importante. De nombreux ensembles de données d'enquête sont largement exploités par une multitude de chercheurs différents qui en font une analyse secondaire, et un certain nombre de grands questionnaires ont été mis au point essentiellement pour servir d'infrastructure de données. Cela fait des décennies que la NSF finance la *General Social Survey*, le *Panel Study of Income Dynamics* et l'*American National Election Study*. Les données de ces trois enquêtes sont aujourd'hui librement accessibles à n'importe qui sur le réseau mondial. Il y a trois ans, l'*Inter-University Consortium for Political and Social Research* a ouvert le site *Web General Social Survey*, financé par la NSF, auquel étudiants et chercheurs du monde entier ont accès gratuitement. Des répertoires de codes bien organisés permettent d'accéder aux 3 000 questions incluses dans l'enquête depuis ses débuts en 1972 et une hyperliaison renvoie de chacune d'entre elles à des tables de données et à des résumés de l'une quelconque des 3 000 publications relatives à cette enquête et qui traitent de cette question. On peut effectuer en ligne une analyse statistique suffisante pour la plupart des usages, et les chercheurs qui ont besoin de faire des analyses plus poussées à l'aide de leur propre logiciel peuvent télécharger les données. Les travaux en cours permettront d'introduire les données des dernières enquêtes et de perfectionner les systèmes de recherche et d'analyse.

De nombreux organismes gouvernementaux proposent maintenant sur le Web des sources de données intéressantes pour la recherche et l'enseignement en sciences sociales. C'est ainsi que la *Division of Science Resources Studies* (SRS) de la NSF propose sur son site un grand nombre de ses rapports et tables de données. Le personnel de la SRS a participé à la création de FedStats, site qui donne un accès facile aux informations statistiques de 70 agences du gouvernement des États-Unis. Le programme *Digital Government* de la NSF va même plus loin en développant des techniques qui amélioreront la gestion interne des agences et les encourageront à fournir électroniquement des données aux chercheurs et au public en général.

Un problème difficile aujourd'hui pour les États est de savoir comment maximiser l'accès des spécialistes en sciences sociales et du gouvernement à de vastes ensembles de données concernant les humains, tout en préservant l'anonymat des personnes. Les données brutes des grandes enquêtes nationales sont disponibles sur le Web, mais sans les codes de localisation géographique des personnes enquêtées, de sorte que l'utilisateur ne peut pas en déduire l'identité des individus. Le GSS et le PSID donnent les géocodes à certains chercheurs sélectionnés, dans des conditions contrôlées, afin qu'ils puissent, à des fins scientifiques, accéder à d'autres données à base géographique. L'étude de Luxembourg sur les revenus permet à des chercheurs dûment inscrits appartenant aux 25 pays Membres de présenter des demandes d'analyse statistique en ligne et en donne les résultats sous une forme qui préserve l'intégrité des données personnelles.

On espère pouvoir mettre au point des systèmes automatiques permettant aux usagers de procéder à des analyses sophistiquées en ligne, tout en les empêchant d'effectuer la séquence de calculs dont ils pourraient déduire les identités des personnes ou organismes inclus dans l'ensemble de données. Si l'on n'y parvient pas, l'infrastructure scientifique devra obligatoirement comprendre des organisations sociales chargées de la protection de la confidentialité. La NSF a récemment entrepris d'en explorer les possibilités, en coopération avec l'*US Census Bureau* et d'autres organismes. Le *Boston Research Data Center*, notamment, offre un environnement protégé où les chercheurs ont accès à des fichiers confidentiels sur des micropanels, et deux nouveaux centres sont en cours d'ouverture en Californie.

Cela dit, quel que soit le soin avec lequel on met en place des systèmes de sécurité technologique et organisationnelle, l'élargissement de l'accès à un plus vaste public de chercheurs va inévitablement accroître le risque de violation des engagements de confidentialité. Une infraction spectaculaire mettant en cause des données confidentielles pourrait jeter le discrédit sur l'ensemble du milieu de la recherche en sciences sociales. Un récent rapport adressé à la Commission sur la protection des infrastructures critiques attachée à la présidence estime que les sociétés modernes vont devenir de plus en plus vulnérables aux attaques par Internet, à la fois parce qu'Internet devient vital pour le commerce et les communications et parce qu'il n'a jamais été conçu pour résister à des attaquants déterminés et

compétents. Dans le même temps, les défenseurs des libertés civiles craignent que celles-ci ne soient mises à mal par une tentative inconsidérée de renforcement de la sécurité. Aussi les questions de sécurité méritent-elles une vigilance particulière à l'heure où se développent les ressources scientifiques de toutes sortes basées sur Internet. Les recherches en sciences sociales relatives aux conditions sociales, organisationnelles et culturelles qui assurent ou menacent la sécurité peuvent contribuer indirectement à préparer l'infrastructure future de l'information pour l'ensemble des sciences.

IV. RÉSEAU INTERNATIONAL DE SCIENCES SOCIALES INTÉGRÉES

Un nouveau type d'infrastructure pourrait être créé pour les sciences sociales dans la décennie qui vient : l'*International Network for Integrated Social Science*. Il s'agirait d'un collaboratoire transdisciplinaire, basé sur Internet, qui fournirait à tous les chercheurs en sciences sociales et comportementales de par le monde les bases de données, logiciels et matériels et autres ressources nécessaires pour la conduite de recherches au niveau mondial qui intègrent des méthodologies d'expérimentation et d'enquête, géographiques et économiques, à beaucoup plus grande échelle qu'on ne pouvait le faire précédemment. Cette infrastructure permettra d'assurer une recherche et un enseignement de pointe en économie, sociologie, psychologie, sciences politiques, géographie et autres domaines connexes.

Un modèle d'infrastructure de ce genre distribuerait les principales fonctions entre huit à dix centres basés sur le Web pour la collecte, l'archivage, l'extraction et l'intégration des données sociales, comportementales et économiques. Ces centres nécessiteraient des moyens de pointe pour le traitement et la mise en réseau d'une capacité de mémoire massive, de nouveaux outils d'analyse, et la mise au point d'outils logiciels et de visualisation innovants. Y seraient en outre reliés, en tant que sites extérieurs, des milliers d'établissements d'enseignement, de musées, d'archives gouvernementales et d'entreprises, avec des protocoles communs pour le transfert de données et l'imputation des coûts.

Aujourd'hui, les données des sciences sociales proviennent de diverses sources – études financées par la NSF, enquêtes du gouvernement des États-Unis, fichiers internationaux, recherches menées par des entreprises – et elles ne sont pas aisément accessibles pour l'analyse secondaire et la méta-analyse qui permettraient d'en tirer le maximum sur le plan scientifique. La liaison et la distribution de ces bases de données par le Web exigent que soient résolus de nombreux problèmes difficiles concernant l'indexage, l'archivage et la recherche de données. Il faudra créer des réseaux de chercheurs appartenant à divers centres pour résoudre les problèmes de fiabilité, de confidentialité, de stockage et d'extraction des données. Il faudra également prendre en compte les problèmes de différence de formats de

données et la nécessité de mettre au point des nouvelles techniques d'analyse spécifiquement adaptées aux données basées sur le Web.

De nombreux sujets d'études complexes des sciences sociales nécessitent des interactions entre un beaucoup plus grand nombre d'intervenants que ne peuvent en contenir les laboratoires classiques ; c'est le cas notamment des marchés, des systèmes électoraux et des réseaux sociaux. Avec un laboratoire en réseau, les expériences peuvent être menées sur des milliers de sujets dans des centaines de lieux éloignés. Ce « netlab », qui transcendera les frontières régionales et nationales, pourra travailler sur de nouveaux échantillons de population et mener des recherches expérimentales sur l'impact de la culture et de la géographie. La durée des expériences pourra être considérablement allongée et couvrir des processus évolutifs jamais étudiés auparavant. Alors que l'intérêt des laboratoires classiques de sciences sociales pour l'enseignement était limité, peu d'universités en étant dotées, l'expérimentation en laboratoire va désormais pouvoir faire partie de l'enseignement normal du premier cycle en sciences sociales et comportementales.

La présence sur le Web d'une base de données complète et efficiente en sciences sociales et comportementales, permettrait d'appuyer plus solidement les recherches sur les études antérieures, et de rapprocher les données nouvelles des données existantes pour les comparer et les étalonner. Les expériences en sciences sociales et comportementales basées sur Internet seront combinées aux enquêtes par questionnaire et les deux pourront être reliées grâce à l'analyse géographique, à d'autres sources de données, telles que données de recensement, statistiques économiques, et données d'autres expérimentations et enquêtes. La diffusion sur Internet de questionnaires à de très grands nombres de personnes va révolutionner le domaine des enquêtes. Les enquêtes nationales aujourd'hui sont si coûteuses que le nombre de personnes interrogées est généralement trop faible ou limité à trop peu de lieux pour permettre une intégration à d'autres sources de données au moyen des systèmes d'information géographique. Ce n'est qu'en passant sur le Web pour la conduite des expérimentations et des enquêtes que les chercheurs en sciences sociales et comportementales pourront étudier des quantités suffisantes de cas pour intégrer les différentes disciplines et méthodes d'analyse.

Le Réseau international de sciences sociales intégrées ouvrira la voie en développant de nouvelles techniques informatiques et statistiques pour conjuguer les différentes méthodes de recherche classiques et transcender leurs limitations traditionnelles. Beaucoup d'études qui seront effectuées sur ce réseau vont représenter un défi pour la prochaine génération d'Internet, car elles nécessiteront des communications très rapides et très fiables entre de nombreux sites ; d'autre part, l'envergure internationale du projet remet en question les structures et modèles actuels de la recherche. Les méthodes existantes de recherche et de distribution des données ne permettront pas de faire face au très grand nombre d'individus et d'organisations des secteurs de la science, de l'enseignement, de l'administration et des

affaires qui auront accès aux bases de données. Le réseau procédera à des enquêtes de calibration en utilisant des méthodes standards pour compenser le caractère non aléatoire des échantillons obtenus en menant les enquêtes et les expériences sur Internet. Il créera des logiciels d'intégration des expérimentations aux questionnaires et des deux à des systèmes d'information géographique. Tel un grand télescope, ce réseau constituera l'infrastructure physique et le cadre organisationnel de la recherche pour les scientifiques de nombreuses universités. Son archive distribuée fournira données et outils d'analyse statistique aux chercheurs et étudiants du monde entier.

Dans la société de l'information du ^{xxi}^e siècle, les données que contiendra ce réseau international d'archives seront extrêmement utiles pour l'enseignement, les affaires et les citoyens en général. D'où la nécessité de développer des moyens techniques et institutionnels qui permettront de fournir les données à autant d'utilisateurs que possible, tout en préservant la confidentialité et les droits de propriété intellectuelle qui peuvent s'attacher à certains sous-ensembles de données. Les découvertes scientifiques étant un bien public universel, l'accessibilité, pour les besoins de l'analyse scientifique, des données d'origine publique est un impératif. Chaque grande base de données du réseau doit être dotée de procédures permettant l'accès des scientifiques accrédités à des données par ailleurs confidentielles, soit dans le cadre d'une protection légale stricte contre les utilisations abusives, soit par des systèmes techniques limitant l'accès à des données de caractère général qui ne portent pas atteinte à la confidentialité.

Le fait de pouvoir se référer aux résultats, aux mesures et à l'expérience pratique des études qui ont été archivées permettra de perfectionner les collectes de données ultérieures. Les études longitudinales pourront établir des comparaisons chronologiques de divers ensembles de données pour déterminer les tendances sociales et économiques. Chaque nouvelle étude sera conçue de façon que les données en soient archivées automatiquement et instantanément, et les publications scientifiques seront reliées aux ensembles de données sur lesquels elles se basent, de façon que le réseau constitue un système de connaissances universel.

V. NIVEAUX D'INVESTISSEMENT

Il est très difficile d'estimer le montant investi chaque année aux États-Unis dans l'infrastructure de l'information, à la fois parce qu'on n'a pas de définition claire du phénomène et parce qu'un très grand nombre d'organismes, publics et privés, réalisent des investissements de ce type. On peut toutefois en dresser un tableau précis pour certains secteurs, ce qui donnera une idée des ordres de grandeur, aussi bien absolus que relatifs.

Chaque année, les agences fédérales déclarent les montants investis en recherche et développement dans le domaine de l'informatique, de l'information et des communications. Pour l'exercice 1998, le total s'est élevé à environ 1 070 000 000 millions de dollars. Deux agences comptent pour plus de la moitié de ce chiffre, la *Defense Advanced Projects Agency* (321 millions de dollars) et la *National Science Foundation* (284 millions de dollars), qui sont responsables de la création d'Internet. Le reste se répartit entre dix autres agences : le *Department of Energy* (129 millions de dollars), la NASA (128 millions de dollars), les *National Institutes of Health* (92 millions de dollars), la *National Security Agency* (36 millions de dollars), le *National Institute of Standards and Technology* (27 millions de dollars), le *Department of Veterans Affairs* (22 millions de dollars), le *Department of Education* (12 millions de dollars), la *National Oceanic and Atmospheric Administration* (8 millions de dollars), l'*Agency for Health Care Policy and Research* (6 millions de dollars) et l'*Environmental Protection Agency* (5 millions de dollars).

Ces importantes sommes ne comprennent pas la recherche et le développement dans les domaines des communications et de la gestion de l'information, qui n'impliquent pas d'informatique de pointe, et elles ne couvrent qu'une très faible part des coûts réels de déploiement des réseaux informatiques opérationnels-matériels et logiciels. Les cinq grands domaines entre lesquels se répartit le montant de 1 070 000 000 millions de dollars sont : le calcul haut de gamme (462 millions de dollars), les réseaux à grande échelle (255 millions de dollars), les systèmes de haute sécurité (33 millions de dollars), les systèmes centrés sur l'homme (280 millions de dollars) et l'enseignement, la formation et les ressources humaines en informatique (39 millions de dollars).

Le vBNS est budgété à 50 millions de dollars et les universités qui participent 50 millions de dollars par an dans Abilene et les travaux qui s'y rapportent. De même, la phase actuelle de l'Initiative fédérale « bibliothèques électroniques » représente un investissement de plus de 50 millions de dollars. Celui consacré à la 20 millions de dollars par an fournis par l'État de Californie, et autant par les universités participantes, peut-être pris sur des chapitres de dépenses plus classiques. Si, comme beaucoup s'y attendent, la publication des revues scientifiques devient entièrement électronique d'ici quelques années, une grande partie du contenu des bibliothèques deviendra progressivement numérique.

Pour les sciences sociales, les investissements ont toujours été plus limités, et il faudrait les augmenter de façon considérable pour atteindre les objectifs que nous avons évoqués. Les trois principales enquêtes financées par la NSF ne coûte 17 milliards de dollars sur les quatre années couvertes par les dotations actuelles : la *Panel Study of Income Dynamics* (9 392 050 dollars), les *National Election Studies* (3 954 259 dollars) et la *General Social Survey* (3 835 000 dollars). Selon les meilleures estimations actuelles, le coût des projets de recherche, développement et

démonstration de la partie américaine de l'*International Network for Integrated Social Science* est beaucoup plus élevé : environ 100 millions de dollars.

Tous ces montants ne sont rien par rapport à l'intérêt que présente la nouvelle infrastructure de l'information pour la science et pour la société en général. Dès 1996, 64 % des écoles primaires et secondaires américaines avaient accès à Internet et la moitié des entreprises avaient des sites sur le Web. Internet est d'ailleurs maintenant le lieu de vastes transactions : c'est ainsi que Netscape a été récemment acheté par AOL pour 4 200 millions de dollars et Excite Corporation, qui au début était un simple moteur de recherche sur le Web, s'est vendue 6 700 millions de dollars. Malgré le ralentissement actuel de l'économie de nombreux pays, il y a de bonnes raisons de croire aux prédictions les plus optimistes quant à la valeur du futur système d'information mondial. La participation des sciences sociales et comportementales à la création de cet ambitieux système de calcul et de communication en maximisera les avantages pour la société en même temps qu'elle permettra des découvertes dans les sciences humaines.

VI. CONCLUSION

La science étant un partenariat international, cela impose, lorsqu'on considère ces questions d'infrastructure, de raisonner au niveau mondial. Nombreux sont les pays qui fabriquent des matériels informatiques et de communication, et il faudrait que leurs scientifiques s'élèvent au-dessus des considérations de concurrence nationaliste pour créer l'infrastructure physique nécessaire à la science du *xxi*^e siècle. Le logiciel devra dans la mesure du possible être modulaire et compatible ; aussi la création de logiciels scientifiques exigera-t-elle une coopération internationale. Les États devront s'engager à mettre les données qui sont leur propriété à la disposition des scientifiques et instaurer des collaborations internationales pour la collecte de données nouvelles. La technologie d'Internet et du World Wide Web permettra de créer de nouveaux types de centres de recherche virtuels et un système de partage des données au niveau mondial qui donneront une nouvelle dimension aux travaux des scientifiques et des enseignants partout dans le monde. C'est par la plus large coopération internationale possible, et en considérant que cette infrastructure est constituée à la fois d'équipements, de logiciels, de bases de données et d'organisations sociales, que l'on pourra le mieux faire avancer la construction de l'infrastructure d'information scientifique de l'avenir.

PUBLICATIONS DE L'OCDE, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(90 1999 24 2 P) ISBN 92-64-26178-8 – n° 50782 1999