Journal of Industrial Microbiology (1996) 17, 498–504

Networking the microbial diversity information

VP Canhos[1],GP Manfio[2] and DAL Canhos[3]

[1]*Faculdade de Engenharia de Alimentos (FEA), Universidade Estadual de Campinas (UNICAMP), Campinas, SP;* [2]*Tropical Culture Collection (CCT);* [3]*Tropical Database (BDT), Fundação de Pesquisas e Tecnologia 'André Tosello', Campinas, SP, Brazil*

The Internet, the electronic global network of computers, has expanded significantly over the last few years, connecting segments from academia, industry and the private sector. The use of the Internet by researchers has diversified substantially from the primordial exchange of electronic messages to the current use of highly developed graphic interfaces capable of accessing images, sounds and retrieving information from databases in different computers around the world. The reduced costs for data storage and ease of connection to the net have nourished a steady exponential growth in the number of sites. Microbiologists now benefit from a wealth of databases and Web sites with microbiological information, ranging from culture collection data, information on microbial diversity, taxonomic and phylogenetic data, biotechnological applications and taxonomic expertise. Efforts are required to stimulate the integration of complex molecular and image databases with more traditional information resources, such as bibliographic, strain data, compendia of species descriptions and metabolic products databases. Coordinating and linking mechanisms are needed in order to promote the development of protocols, guidelines and minimal standards to ensure data quality. Future challenges to the scientific use of Internet resources include the setting up of clearing-house mechanisms and virtual libraries for the organization of the microbiological resources on the network.

Keywords: microbial diversity; information networks; Internet; electronic publishing; digital libraries

The growing need for networked information

A new biology is emerging from recent developments in the areas of automated sequencing and robotics, associated with data analysis and information technology [27]. Several laboratories are now using the Internet to help manage collaborative research projects, while major public nucleotide and protein sequence databases have switched their operations to the World Wide Web, the fastest growing and most innovative part of the Internet. The Web allows users to browse through pages containing text, graphics and multimedia applications such as sounds and images in motion. The information can be transparently accessed through hypertext links, by using the Hypertext Markup Language (HTML) and Universal Resources Locators (URLs) that tell the user on which Web server the requested information can be found on the Internet.

The genome projects (Table 1) and the rest of biology are merging at the genetic maps, the sequence databases and the peer-reviewed scientific literature, with sequences acting as specific retrieval keys cross-referenced with complementary phenotypic and physiological information [4].

Projects viewed with scepticism in the early 1980s, such as the determination of the complete 12.5-megabase genome sequence of the yeast *Saccharomyces cerevisiae*, are closing in on the complete yeast genome sequence [38]. The sequencing of the nematode *Caenorhabditis elegans* is projected to be completed by the end of 1998. The Human Genome Project final goal, to complete the 3 billion nucleotide sequence of the human genome, is also well on its way. The project has revealed more than 150 000 sequences and GenBank is receiving submissions at a rate of 1500 sequences per day, with the released data being heavily utilized by biomedical researchers [4].

Meanwhile, complete genome sequences are already available for some prokaryotes, such as *Haemophilus influenzae* [17] and *Mycoplasma genitalium* [18], while sequencing projects of larger microbial genomes are underway (Table 1). While more and more biologists are browsing through linked databases, clicking from gene sequences to protein structure to journal articles, the number of electronic hyperlinked publications is multiplying with traditional hardcopy journals reaching the cyberspace [6,12,15,20,23].

Increasingly networked information is not only providing speedy answers to scientific queries, but is also broadening the universe of possible questions. The hyperlinked and networked information will help to solve the huge uncertainties of microbial diversity. We still know very little about the viable but non-culturable microorganisms [30] and the value of microbial diversity [37].

Trends and developments in networking

The Internet

The best known computer network is the Internet, a network of networks, which encompasses an estimated 50 000 networks worldwide. The Internet now is doubling in size every year [5] and is experiencing exponential growth in the number of networks, host computers, and volume of

Correspondence: Dr VP Canhos, Base de Dados Tropical, Rua Latino Coelho, 1301 CEP 13087-010, Campinas, SP, Brazil. E-mail: vcanhos@bdt.org.br
Received 15 May 1996; accepted 29 August 1996

Table 1 List of some Internet sites and resource centers related to genome mapping and whole-organism sequencing

Category/group	Description	Internet address
Microbial genome maps and gene products		
Bacillus subtilis	Non-Redundant Database for *Bacillus subtilis* (NRSub)	http://acnuc.univ-lyon1.fr/nrsub/nrsub.html
	The BSORF database: The International Project of Sequencing the *Bacillus subtilis* Genome	http://bacillus.genome.ad.jp/BSORF-DB.html
	SWISS-PROT Protein Sequence Data Bank: *Bacillus subtilis* gene products	http://expasy.hcuge.ch/cgi-bin/lists?subtilis.txt
Escherichia coli	The *E. coli* Genome Center	http://ecoliftp.genetics.wisc.edu/
	CGSC: *E. coli* Genetic Stock Center http://cgsc.biology.yale.edu/	
Filamentous fungi	Fungal Genetics Stock Center	http://kumchttp.mc.ukans.edu/research/fgsc/main.html
	PathoGenes: fungal pathogens of small-grain cereals	http://probe.nalusda.gov:8300/cgibin/browse/pathogenes
Haemophilus influenzae	The *Haemophilus influenzae* Rd Genome Database (HIDB)	http://www.tigr.org/tdb/mdb/hidb/hidb.html
Mycobacterium sp	The *Mycobacterium* DataBase (MycDB)	http://kiev.physchem.kth.se/MycDB.html
Mycoplasma genitalium	The *Mycoplasma genitalium* Genome Database (MGDB)	http://www.tigr.org/tdb/mdb/mgdb/mdgb.html
Saccharomyces cerevisiae	The *Saccharomyces* Genome Database	http://genome-www.stanford.edu/Saccharomyces/
Several species	Micado (formerly MadBase): Database for Microbial Genomes	http://locus.jouy.inra.fr/cgibin/genmic/madbase/progs/madbase.operl
	Virtual Genome Center: tools for DNA and protein analysis	http://alces.med.umn.edu/VGC.html
	Yeast virtual library	http://genome-www.stanford.edu/Saccharomyces/VL-yeast.html
Plant and animal genomes		
Plants	Plant Genome Databases: several species	http://probe.nalusda.gov:8300/
	Additional resources	http://probe.nalusda.gov:8000/elsewhere/indexbio.html#plant
Insects	Mosquito Genomics WWW Server	http://klab.agsci.colostate.edu/
	Additional resources	http://probe.nalusda.gov:8000/elsewhere/indexbio.html#insect
Human	Human Genome Project homepage	http://www.ornl.gov/TechResources/Human-Genome/home.html
Animals	Livestock Animal Genome Databases	http://tetra.gig.usda.gov:8400/
	Additional resources	http://probe.nalusda.gov:8000/elsewhere/indexbio.html#animal

traffic. The number of host computers has increased from 200 in 1983 to almost 10 million at the beginning of 1996 (Network Wizards, http://www.nw.com/zone/WWW/top.html). The non-USA connections are growing more rapidly, with the projected crossover estimated to occur in 1996. More than 160 countries currently have some form of connection to the Internet and in 1995 more than 30 developing countries were connected to the Internet for the first time. It is predicted that by the end of 1996 only five or six African countries will lack Internet connections [8]. Every 30 minutes another network is connected and the number of Internet users is expected to reach 200 million by the year 2000 [5]. The Regional Map to Internet Connectivity (http://info.isoc.org:80/images/mapv14.gif), regularly updated by the Internet Society, and the International E-mail Accessibility page (http://www.ee.ic.ac.uk/mis/country-codes.html) are good sources of information on the types of connectivity available and e-mail access in the different regions of the world.

The value of the Internet is increasing with the size and diversity of its user population, the power and the sophistication of its applications and the capability of its infrastructure. The continuous upgrading of the Internet backbone is enabling new applications such as teleconferencing, multimedia electronic mail, image analysis and on demand electronic publishing [13,16,25]

The US National Science Foundation (NSF) is planning to create a high speed computer network that will allow US scientists to carry out activities in virtual reality laboratories. NSF is providing the support to link the NSF's supercomputer centers through a 155-megabit very high speed backbone network service (vBNS), that in the future will link virtual reality laboratories [7,24,26]. The immediate impact of this programme will be to allow scientific interaction at high speeds, allowing the operation of instruments by remote control and the analysis of huge amounts of data by supercomputers.

Linking mechanisms and search engines

The potential of using networked hypertext and multimedia has prompted many users and publishers to create and explore countless innovative applications on the Internet [34]. Organizations distributed throughout the world are discovering both new opportunities provided by the information infrastructure and nurturing its development through the establishment of discussion lists, on-line databases, metadatabases, virtual libraries (VLs), special interest networks (SINs) and search engines.

Metadatabases: Metadatabases, which hold data about data, are analogous to library catalogues, which describe books yet are not the books themselves. Metadata are critical to understanding database contents by describing them and metadatabases are relevant to help users find

information about datasets without the need to access and investigate the data itself. Besides facilitating the discovery of existing datasets, metadata provide additional information about contents, quality and features of the datasets [14]. A good example of metadatabases is the National Spatial Data Infrastructure (NSDI), provided by the Federal Geographic Data Committee (FGDC, http://fgdc.er.usgs.gov/), which encompasses policies, standards, and procedures for organizations to produce cooperatively and share geo-spatial data.

Virtual libraries: A virtual library (VL) is an organized set of links to items (documents, software, images, databases) in the network. The purpose of a virtual library is to enable users of a site to find information that exists elsewhere in the network. Virtual libraries are a natural growth of the ability of modern client-server protocols (especially http and gopher) to provide seamless links to information anywhere in the Internet [19]. The World Wide Web Consortium (http://www.w3.org/pub/WWW/) holds the WWW Virtual Library (http://www.w3.org/pub/DataSources/bySubject/Overview.html). This is a distributed subject catalogue and includes a list of virtual libraries on a number of fields including those of interest to biodiversity. VLs related to microbiology are listed in Table 2.

Special interest networks: A special interest network (SIN) is a group of people and/or institutions who collaborate to provide information about a particular subject. SINs consist of a series of participating 'nodes' that contribute to the network's functions [19]. More specifically, the nodes are committed to the provision of public access to the unique information available at their sites and transparent links to other nodes in the network. A good example of a SIN is the European Molecular Biology Network (EMBNet, gopher://gopher.ebi.ac.uk:70/11/ and also at http://www.ebi.ac.uk/). EMBNet is a special interest network that serves the European molecular biology and biotechnology research community. It consists of nodes operated by biologically oriented centers in different European countries, featuring a number of services and activities, related mainly to genomic databases and sequence data. The Biodiversity Information Network (BIN21, http://www.bdt.org.br/bin21/bin21.html), another special interest network, was established to facilitate the access to all levels of biodiversity information, from molecular to biosphere [10]. BIN21 is developing affiliations with and network links to organizations that are already carrying out related activities [9,22].

Search engines: Due to the decentralized nature of the data on the Internet, tools for searching and reporting are fundamental for finding and retrieving information on relevant sites. This can be achieved by using powerful automated searching tools, which systematically retrieve and store descriptive information held in a large number of sites. The results are large databases, which hold information on resources available on thousands of Internet sites and web pages [31].

Searchers on such databases can be performed on several 'categories' of data, such as web pages, Usenet newsgroups, gopher and ftp sites and e-mail addresses. Examples of search engines in the Internet include the Alta Vista (http://www.altavista.digital.com/), InfoSeek (http://guide. infoseek.com/), Lycos (http://www.lycos.com/), and Yahoo (http://www.yahoo.com/), among several others (NetSearch, http://home.mcom.com/home/internet-search.html). A more recent addition to the list of search tools, the C|net (http://www.search.com/), is an integrated search engine which allows the user to search for information in over 250 search engines in the network.

Networking the scientific information

Before the 'Internet explosion' associated with major software and hardware developments, scientific publishing activities were characterized as top-down initiatives in the hands of traditional international publishers and major universities [34]. Meanwhile, the real revolution in the telecommunication and information fields has occurred from the bottom-up [29]. Personal computers and workstations, local area networks, cost-oriented leased lines, routers, network operating systems, the Internet, and other capabilities have empowered individuals and organizations to develop Web servers to make available pre-prints and other digital publications [33,36].

From paper to bits

Information is becoming more digital and networked and less paper. Scientific networked information today includes not only conventional articles but also personal communications and live records, promoting collaborative work among scientists around the world [35]. Looking only from the perspective of who is connected and has access to today's information technology tools, this undoubtedly brings great advantages and possibilities such as:

- integrating and linking different information bases;
- electronic searching combined with analysis and modeling systems;
- establishing virtual libraries, metadatabases and special interest networks;
- having decentralized, integrated and multidisciplinary information sites;
- access to specialized information on specific subject matters, even if the researcher is not from the field;
- reduced costs of production and distribution;
- quick access to up-to-date on-line information, as opposed to traditional paper-publishing, where printing and distribution can be very slow.

Established publishers could face future market problems if they fail to react quickly [32]. Several publishers are slow in adapting to new technological developments, even though there are a growing number of software tools that allow the display of graphics, tables, complex equations and even offer analysis and modeling features. While big publishing companies are recognizing that the electronic journals are clearly the way to go, the wave of the future, reliable mechanisms for controlling and charging for access to an electronic publication, have not yet been established. Strategic planning is necessary to establish realistic cost

Table 2 List of some Internet sites and resource centers relevant to microbial diversity information

Category/group	Description	Internet address
Special interest networks and information resource centers		
BIN21	Biodiversity Information Network – Agenda 21	http://www.bdt.org.br/bin21/bin21.html
Biodiversity and Biological Collections Web Server	Dedicated to systematists and other biologists of the organismic kind	http://muse.bio.cornell.edu/
Center for Microbial Ecology	NSF Science and Technology Center for Microbial Ecology	http://www.cme.msu.edu/CME/
MSDN	Microbial Strain Data Network	http://www.bdt.org.br/bdt/msdn/
WDCM	World Data Center on Microorganisms	http://www.wdcm.riken.go.jp/
On-line culture collections		
ATCC	American Type Culture Collection, USA	http://www.atcc.org/
Brazilian National Catalogue of Strains	Holdings of 36 Brazilian culture collections: bacteria, fungi, yeasts, microalgae, cell culture, protozoa and viruses	http://www.bdt.org.br/bdt/index/biotecnologia/micropage/micromenu.catalogue/catbr
CCT	Tropical Culture Collection (Brazil)	http:///www.bdt.org.br/cct/
The China Directory of Microorganisms	Holdings of Chinese culture collections: actinomycetes, bacteria, fungi, yeasts and viruses	http://sun.im.ac.cn/database/catalogs.html
EFD	Entomopathogenic Fungi Databank, Brazil	gopher://asparagin.cenargen.embrapa.br/11/fungos
Japanese Catalogue of Microbial Strains	Holdings of the Japanese Federation for Culture Collections: bacteria, fungi, yeasts and viruses	http://www.wdcm.riken.go.jp/wdcm/JFCC.html
JCM	Japan Collection of Microorganisms	http://www.wdcm.riken.go.jp/JCM/
MGD	Microbial Germplasm Database: holdings of several 'research' collections	http://mgd.cordley.orst.edu/cgi-bin/mgd
MSDN	Microbial Strain Data Network	http://www.bdt.org.br/bdt/msdn/
• DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany	http://www.bdt.org.br/bdt/msdn/dsm/ and http://www.gbf-braunschweig.de/DSMZ/dsmzhome.html
• IBSO	Siberian Luminous Bacteria Database	http://www.bdt.org.br/bdt/msdn/ibso/
• IEGM	Institute of Ecology and Genetics of Microorganisms IEGM Regional Specialised Collection of Alkanotrophic Microorganisms, Russia	http://www.bdt.org.br/bdt/msdn/iegm/
• IMI	International Mycological Institute Culture Collection, UK: fungi	http://www.bdt.org.br/bdt/msdn/imi/
• IPPAS	KA Timiryazev Institute of Plant Physiology Collection, Russia: microalgae	http://www.bdt.org.br/bdt/msdn/ippas/
• Kemijski Institute	Slovenia Filamentous Fungi	http://www.bdt.org.br/bdt/msdn/slovenia/
• KMMGU	Russian Bacteria Database	http://www.bdt.org.br/bdt/msdn/kmmgu/
• LEBIN	Russian Basidiomycetes	http://www.bdt.org.br/bdt/msdn/lebin/
• MSU	Russian Yeasts Database	http:///www.bdt.org.br/bdt/msdn/msu/
• NCYC	UK National Collection of Yeast Cultures	http://www.bdt.org.br/bdt/msdn/ncyc/
• PETERHOF	Genetic Collection of Yeast Cultures	http://www.bdt.org.br/bdt/msdn/peterhof/
• RIAM	Russian Fungi	http://www.bdt.org.br/bdt/msdn/riam/
• VKM	Institute of Biochemistry and Physiology of Microorganisms, Russia: – bacteria, fungi and yeasts – genetically modified organisms	 http://www.bdt.org.br/bdt/msdn/vkm/strain/ http://www.bdt.org.br/bdt/msdn/vkm/modified/
• VKPM	National Collection of Industrial Microorganisms, Russia: bacteria, fungi, yeasts and actinomycetes	http://www.bdt.org.br/bdt/msdn/vkpm/
Virtual libraries		
Biosciences VL	WWW virtual library of biological sciences resources, including microbiology and virology	http://golgi.harvard.edu/biopages.html
Biotechnology VL	WWW virtual library of biotechnological resources: pharmaceutical development, genetic engineering, medical device development, and related fields such as pharmacology and toxicology	http://www.cato.com/interweb/cato/biotech/
Mycology VL	WWW virtual library of mycological resources in the Internet	http://muse.bio.cornell.edu/~fungi/fcollect.html
Biogeographical data and species inventories		
Fungi from Brazil	Checklist of fungi recorded by Batista and co-workers in Brazil	http://www.bdt.org.br/bdt/fungilistbr/
Species 2000	Indexing systems for all groups of organisms	http://sunrae.uel.ac.uk/species2000/
Systematic Agenda 2000	Taxonomic and species inventory effort	gopher://muse.bio.cornell.edu:70/11/standards/sa2000
WCMC	World Conservation and Monitoring Centre: worldwide conservation priorities	http://www.wcmc.org.uk/

Table 2 Continued

Category/group	Description	Internet address
Microbial taxonomy and phylogeny		
Bacterial systematics	Bacterial Nomenclature UP-TO-DATE Database (DSMZ)	http://www.bdt.org.br/bdt/bacterianame/ and http://www.gbf-braunschweig.de/DSMZ/bactnom/bactname.htm
	Bergey's Manual Trust: bacterial nomenclature list (under construction)	http://www.cme.msu.edu/bergeys/
	Approved Lists of Bacterial Names: WDCM	http://www.wdcm.riken.go.jp/htbin/bact_names.pl
Fungal systematics	Systematic Arrangement of Genera from 8th Edition of the Dictionary of the Fungi	http://www.cabi.org/institut/imi/taxon.htm
	WWW VL Mycology: Taxonomy	http://muse.bio.cornell.edu/~fungi/fgeneral.html#FBIO
Ribosomal sequences	Ribosomal Database Project (RDP)	http://rdp.life.uiuc.edu/
	RNA World	http://www.imb-jena.de/RNA.html
	BERLIN: 5S rRNA Database	http://www.caos.kun.nl/srs/wgetz?-info+BERLIN
	16S/23S rRNA Database	http://pundit.colorado.edu:8080/
Phylogenetic classification	The Ribosomal Database Project (RDP)	http://rdp.life.uiuc.edu/RDP/data/ssu.html
	The Tree of Life	http://phylogeny.arizona.edu/tree/phylogeny.html
	The Phylogeny of Life	http://ucmp1.berkeley.edu/allife/threedomains.html
Publications, publishers and communication forums		
BIOSCI	Set of electronic communication forums: include several discussion and news lists in biological sciences	http://www.bio.net/
Bioline Publications	Electronic publishing service that offers on-line version of several hardcopy microbiological journals	http://www.bdt.org.br/bioline/
BioMedNet	Forum for biological and medical on-line journals	http://BioMedNet.com/
Harvard Bio Labs Library	Searchable indexes and links to several on-line journals in Biology	http://golgi.harvard.edu/library/
NewJour website	Internet list for new journals and newsletters available on the Internet	http://gort.ucsd.edu/newjour/
Electronic Journals VL	WWW virtual library of electronic journals	http://www.edoc.com/ejournal/
Publishers VL	WWW virtual library of publishers	http://www.comlab.ox.ac.uk/archive/publishers.html

recovery programs and to adapt to the new technology. Important issues such as the guarantee of copyright and intellectual property on the Internet are still unclear and are being discussed on the network.

Network publishing

Network publications range from electronic versions of familiar paper items, such as books, journals and reports to the newly available on-line-only peer-reviewed journals (see Bioline Publications, http://www.bdt.org.br/bioline/). Novel productions, such as image databases or thematic compilations of pointers to many different sites are also relevant sources of information in the field of Biology (see The Tree of Life, http://phylogeny.arizona.edu/tree/ phylogeny.html and The Phylogeny of Life, http://ucmp1.berkeley.edu/alllife/threedomains.html).

One advantage of network publishing is that the material can be continuously updated in a decentralized fashion, and be instantly available to users in a transparent manner. The worldwide distributed sites can ensure that users have access to the current version of a publication by creating live links to the repositories of on-line information.

A Conference of Experts on Electronic Publishing in Science convened by ICSU Press (International Council of Scientific Unions) and UNESCO (United Nations Educational, Scientific and Cultural Organization) in February 1996 discussed the problems and opportunities presented by the new information/communication technologies. The conference recommended that strict peer review should be applied to scientific electronic journals and that codes of practice and guidelines for electronic archiving should be developed (http://www.Imcp.jussieu.fr/icsu/Information/Proc_0296/).

Publishers and scientists are considering using the interactive powers of the Internet to turn on-line journals into perpetual electronic conferences, where articles take the place of lectures and sprout on-line discussion groups and commentaries [33].

The ability to use the available Internet tools to develop innovative ways of carrying out traditional tasks, such as writing taxonomic descriptions in microbial systematics, could speed up these tasks. It is now possible to publish all new microbial species descriptions on the Internet, and have them immediately available to other researchers by using the resources of the World Wide Web and the Internet.

The digital documents, besides having embedded figures and tables, have the added power of multimedia and hypertext links to items anywhere in the world. The information can be text, images, animation, sound and scientific data. Documents can undergo similar editing procedures as they do now, including peer reviewing, and, once 'accepted', they become validly published.

Such a process could mean that all microbial species names would have their own 'Internet address' where they could be easily found and linked using hypertext to associated information such as phylogenetic data and type strains primary data, including images, culture collection information and strain-specific technological applications.

By the end of 1996, the number of electronic peer-

reviewed journals may increase by another order of magnitude, as publishers announce on-line offerings [33]. The full text and graphics of prestigious traditional journals in Biological Sciences, such as Science (http://science-mag.aaas.org/science), Nature (http://www.nature.com/) and Nucleic Acids Research (http://www.oup.co.uk/nar/) are already available on-line. Addresses for electronic journals and newsletters available in the Internet (nearly 2000 entries) can be found at the Electronic Journals VL (http://www.edoc.com/ejournal/) and NewJour websites (http://gort.ucsd.edu/newjour/).

Already a mouse click can take the microbiologist from one article to related articles in the same or other relevant journals and to information resources such as DNA sequence and protein databases. By offering authors' raw data or the software used in the analysis, some journals will even allow readers to double-check an authors' work.

The changes and re-organization of the existing efforts are moving towards a greater interactivity. The centralization era is over, and is being replaced by distributed and cooperative enterprises. New browsers that allow web users to create and modify documents on screen are under development. Groups of authorized users will be able to work simultaneously on the design of scientific experiments, update experimental protocols, interpret data as the results come along and edit the final manuscripts.

Digital libraries

Not only scientists and publishers, but also libraries are adapting to the innovative ways of producing, collecting and disseminating scientific information in the digital world. Several efforts are being carried out to open the way to weaving electronic journals and scientific libraries into a single interconnected database. The Digital Library Initiative (DLI), a project jointly sponsored by the US National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA) and the Advanced Research Projects Agency (ARPA), congregates six University Libraries in an effort to '... dramatically advance the means to collect, store and organize information in digital forms, and make it available for searching, retrieval, and processing via communication networks' (http://walrus.stanford.edu/diglib/pub/nsf.announce.html).

Looking towards the future

Considering the interdisciplinary nature of modern research in microbiology (from molecular biology to polyphasic systematics and industrial applications) and the convergence of computing, information, and communication, it becomes clear that the use of networks in the microbiology community is only just beginning. Associated with the continuous improvements and periodic breakthroughs that open entirely new possibilities, we have this truly cultural revolution that, among many other issues is helping to promote interdisciplinary and intercultural collaboration.

The challenge of 'wiring up' developing countries will depend on the ability of local governments and international organizations to ensure that the critical technical issues are considered and that the appropriate enabling physical infrastructure is put in place [28].

New strategies and mechanisms for information development and dissemination are required to complement the top-down approaches to more responsive bottom-up initiatives involving scientists, policy makers and citizens in developing and industrialized countries. To develop such mechanisms on a broader basis will not be trivial and will require substantial changes in the mindset of policy makers. Steps to encourage the paradigm shift necessary to achieve this goal will require key interventions including international efforts to overcome the:

- inadequate infrastructure, technical and political problems related to the Internet topology and connectivity in some regions of the world;
- problems associated with the establishment and operation of information and service resources, including infrastructure and expertise;
- digital illiteracy and the inadequate awareness of the potential of telematics.

The implementation of international programmes to document the importance and role of systematic biology in sustainable development is contributing to the enhancement of the microbial diversity knowledge base. The 'Sustainable Biosphere Initiative' [1], the 'Systematics Agenda 2000' [2]and the 'Microbial Diversity 21' [21] have stimulated debate on the importance of systematics in human affairs. Within the context of Diversitas, the International Union of Biological Sciences (IUBS) Biodiversity Programme, the international Biodiversity Information Network (BIN-21, http://www.bdt.org.br/bin21/bin21.html) was established in order to facilitate the dissemination of biodiversity information worldwide [9–11].

Other developments under the sponsorship of IUBS, such as the Species 2000 Project (http://www.uel.ac.uk/species2000/), will be fundamental to help clarify microbial diversity uncertainties. Species 2000, to be operated by a federation of database organizations, aims at indexing names of all known species of microorganisms, plants and animals [3], in cooperation with CODATA (Committee on Data for Science and Technology; http://www.cisti.nrc.ca/programs/codata/), IUMS (International Union of Microbiological Societies: http://www.Imcp.jussieu.fr/~fabrice/icsu/Membership/SUM/iums.html) and UNEP (The United Nations Environmental Programme; http://unep.unep.no/).

In order to enable capacity-building developments, it is important that international organizations work in close collaboration with the Secretariat to the Convention on Biological Diversity (CBD) through the Clearing-House Mechanism (CHM). Devised to be an instrument to facilitate scientific and technological cooperation between industrialized and developing countries, the CHM (http://www.istar.ca/biodiv/) will be developed on an incremental basis. Through the pilot phase implementation (1996–1997) the CHM existing information resources on the Internet are being interconnected, navigation tools are being developed and the CBD relevant gaps/links are being identified. Future developments will rely on the strengthening of networking strategies between the active partners,

and on the interconnection of decentralized information sources.

Organizations such as the International Union of Biological Sciences (IUBS), the International Union of Microbiological Societies (IUMS) and the World Federation for Culture Collections (WFCC) have major roles to play in assisting the process towards greater equality of access to the scientific information on the microbial diversity.

References

1 Anonymous. 1991. Sustainable biosphere initiative. Ecology 72: 371–412.
2 Anonymous. 1991. Systematics Agenda 2000: integrating biological diversity and societal needs. System Zool 40: 520–523.
3 Bisby FA. 1994. Global master species databases and biodiversity. Biol Int 29: 33–38.
4 Bogusky MS. 1995. The turning point in genome research. Trends Biochem Sci 20: 295–296.
5 Bournellis C. 1995. Internet 95: the Internet's phenomenal growth is mirrored in starting statistics. Internet World November: 47–52.
6 Bungay HR. 1995. Why electronic hypertext journals are inevitable. Binary 7: 89–90.
7 Buttler B. 1996. US plans for virtual laboratories by Internet. Nature 380: 93.
8 Buttler B, S Barker and C Levitin. 1996. Research and the Internet connection. Nature 380: 377–381.
9 Canhos DAL, VP Canhos and BE Kirsop (eds). 1994. Linking mechanisms for biodiversity information (UNEP sponsored workshop). Tropical Foundation, Campinas (SP), Brazil.
10 Canhos VP, D Lange, BE Kirsop, S Nandi and E Ross (eds). 1992. Needs and specifications for a biodiversity information network. United Nations Environment Programme, Nairobi.
11 Canhos VP, GP Manfio and LD Blaine. 1993. Software tools and databases for bacterial systematics and their dissemination via global networks. Antonie van Leeuwenhoek 64: 205–229.
12 Carling B. 1996. Crawling around the Web: biology and the Internet. Biologist 43: 9–12.
13 Cerf VG. 1995. Computer networking: global infrastructure for the 21st century. [URL = http://www.cs.washington.edu/homes/lazowska/cra/networks.html].
14 Crossley D. 1995. Wais though the web—discovering environmental information. [URL = http://atlas.erin.gov.au/gis/develop/WWW_fall94/paper.html].
15 Edwards A, B Kirsop and D McKenney. 1995. The Internet and the microbiologist. Binary 7: 91–94.
16 European Commission. 1995. The G-7 Information Society Conference. Conference Programme. [URL = http://www.ispo.cec.be/g7/g7main.html].
17 Fleischmann RD *et al*. 1995. Whole genome random sequencing assembly of *Haemophilus influenzae* Rd. Science 269: 496–512.
18 Fraser CM *et al*. 1995. The minimal gene complement of *Mycoplasma genitalium*. Science 270: 397–403.
19 Green DG and JR Croft. 1994. Proposal for implementing a Biodiversity Information Network. In: Linking Mechanisms for Biodiversity Information. Proceedings of a Workshop for the Biodiversity Information Network. Base de Dados Tropical, Campinas, São Paulo, Brazil.
20 Harper R. 1995. World Wide Web resources for the biologist. Trends Genet 11: 223.
21 Hawksworth DL and RR Colwell. 1992. Microbial Diversity 21: biodiversity amongst microorganisms and its relevance. Biodiv Conser 1: 221–226.
22 Kirsop B and VP Canhos. 1995. Biodiversity information transfer: some existing initiatives and how to link them. In: Microbial Diversity and Ecosystem Function (Allsopp D, RR Colwell and DL Hawksworth, eds), pp 439–445, CAB International, UK.
23 Knight P. 1995. Microbiology resources on the Internet. ASM News 61: 397–400.
24 National Research Council. 1993. National Collaboratories: Applying Information Technology for Scientific Research. National Academy Press, Washington, DC.
25 National Research Council. 1994. Realizing the Information Future: the Internet and Beyond. National Academy Press, Washington, DC.
26 Office of Science and Technology Policy. 1994. High performance computing and communications: toward a national information infrastructure. USA.
27 Olson MV. 1995. A time to sequence. Science 270: 394.
28 Pelsinsky N. 1996. Connectivity and developing countries. Personal communication (available from nmpelsin@welchlink.welch.jhu.edu).
29 Rutkowski A-M. 1994. The present and the future of the Internet: five faces. Keynote address at Networld:Interop 94, Tokyo. [URL = http://info.isoc.org:80/speeches/interop-tokyo.html].
30 Stackebrandt E. 1994. The uncertainties of microbiological diversity. In: The Biodiversity of Microorganisms and the Role of Microbial Resources Centers (Kirsop B and DL Hawksworth, eds), World Federation for Culture Collections.
31 Steinberg S. 1996. Seek and ye shall find (maybe). Wired May: 108.
32 Stix G. 1994. The speed of write. Scient Amer December: 72–77.
33 Taubes G. 1995. Indexing the Internet. Science 269: 1356.
34 Taubes G. 1996. Science journals go wired. Science 271: 764.
35 Taubes G. 1996. Speed of publication: stuck in first gear. Science 271: 765.
36 Taubes G. 1996. Electronic preprints point the way to 'author empowerment'. Science 271: 767.
37 Tiedje JM. 1994. Microbial diversity: of value to whom? ASM News 60: 524–525.
38 Williams N. 1995. Closing in on the complete yeast genome sequence. Science 268: 1560.