Enhancing electronic collaboration in the South

Enrique Canessa, Fulvio Postogna and Sandro Radicella

The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.
Email: radionet@ictp.trieste.it

Virtual laboratories, based on the extensive use of electronic communication, could become a valuable device for combatting the brain drain from the South to the North

According to recent reports, the number of people globally who are connected electronically will double to 300 million by the year 2005, with the greatest growth occurring in Asia and South-America [1]. This expected growth will motivate further communication between people, and it is therefore worthwhile to imagine how electronic collaboration (e-collaboration) could help reverse the south-north brain drain.

To do this, it would first be useful to be able to quantify today's use of electronic media for the rapid dissemination of scientific research. This complex exercise would help to provide a picture of the current potential for e-collaboration in the scientific world in general, and in the research communities of developing countries in particular. (See [2] for a comprehensive picture of the state of technology and research in the region, with an updated list of 431 leading scientific research centres, institutes and universities.)

One potential application of such information is to the development of 'virtual laboratories'. These are projects that involve research collaboration carried out at a distance, the performance of remote (or distributed) computing, and the sharing of data between groups of scientists, each of whom remains in their home institutions.

Analysing the impact of e-mail for the dissemination of research could help assess whether further efforts should be put into creating future developments of Virtual Labs in the South. A limited but realistic picture can be obtained by looking at the Los Alamos e-preprint archives [3]. Of particular interest is knowledge of an author's geographical location, and how many of these preprints are posted electronically via e-mail from (one or more) less-developed country.

The Los Alamos archives incorporate 25,000 new electronic papers a year in topics ranging from non-linear dynamics to computer science [4]. The full condensed matter archive (cond-mat), for example, offers 14,769 e-print entries for the period Apr 1992 to Dec 1998, and the full High Energy Physics-Theory archive (hep-th) offers 16,653 e-prints

entries for Aug 1991 to Dec 1998. These are distributed as follows:

Years:	1991	1992	1993	1994	1995	1996	19971998
Number cond-mat e-prints:	-	219	687	1323	1825	2545	35334637
Number hep-th e-prints:	301	1386	2091	2442	2347	2626	26862774

From these numbers, it appears that the number of 'hep-th' entries has reached a plateau, while the number of 'cond-mat' entries continues to grow.

The 1998 (La)TeX-based sources - a popular computer language for composing formatted scientific text for high quality printing - include 4637 'cond-mat' e-prints and 2774 'hep-th' e-prints, without cross-listings, which means that they make up about 30 per cent of the complete Los Alamos archive for that year. A sample of 275 'cond-mat' e-prints, plus 201 'hep-th' e-prints, from January 1988 has been examined, representing 5.9 per cent and 7.3 pre cent of the respective entries for the whole of 1998.

From this subset, all countries mentioned in the e-prints are listed below. Also shown is the number of times that such countries were referred by the authors as their permanent and/or present address. These figures give a rough idea of the number of e-prints that are produced from scientists in the South.

	Country	Number of Cond-mat authors	Number of he-th authors
EUROPE:	Germany	87	37
	France	77	8
	Italy	55	47
	Spain	40	4
	UK	39	21
	Russia	32	35
	Switzerland	16	13
	Denmark	10	5
	Netherlands	9	20
	Poland	7	6
	Austria	6	2
	Belgium	5	20
	Hungary	4	-
	Ukraine	4	3
	Ireland	3	6
	Romania	3	1
	Slovenia	3	-
	Czech Rep.	2	1
	Finland	2	-
	Norway	2	2
	Sweden	2	1

	Estonia	1	-
	Greece	1	-
	Slovak Rep.	1	-
	Iceland	-	2
	Portugal	-	1
	USA	164	85
AMERICA:			
	Brazil	9	18
	Canada	7	9
	Argentina	7	2
	Cuba	2	-
	Mexico	2	1
	Colombia	-	4
	Costa Rica	-	2
	Japan	48	45
ASIA:			
	India	18	14
	Israel	15	6
	China	7	4
	Iran	5	2
	Armenia	4	-
	Korea	4	7
	Rep.Georgia	-	1
	Singapore	-	1
	Taiwan	-	1
	Australia	9	-
OCEANIA:			
	New Zealand	3	-
	Algeria	1	-
AFRICA:			

In this sample, there are 704 'cond-mat' authors who appeared in one or more e-prints that used e-media to disseminate their scientific work, and 416 'he-th' authors. There were 88 e-prints from groups of 'cond-mat' authors working in different countries (32 per cent of the total analysed), and 50 e-prints for 'he-th' authors (25 per cent of the total analysed). In both archives, only one e-print has resulted from a collaboration between South institutions alone, namely an India/Brazil collaboration for 'cond-mat', and Argentina/Brazil/Singapore for 'he-th'. In other words, only 1.14 per cent and 2 per cent of the considered e-prints for 'cond-mat' and 'he-th' respectively were the fruit of purely 'South' collaborative efforts.

The latter figures follow closely the trends that have been observed for scholarly publishing in refereed journals. At present, developing countries unfortunately account for only 2 per cent of the total scientific use and production (see, e.g., [5]). Scientific activity as measured by publications (appearing in the SCI-Science Citation Index database) for

different regions can be found in [6].

Our preliminary exercise does not provide any firm conclusions. To achieve this, the above figures would need to be increased to include the whole of the Los Alamos archives. It would also be useful to know the geographic distribution of scientists in each speciality, in order to search for possible correlations. Our goal here has merely been to stimulate some bold experiments in order to obtain more meaningful data that could be used to plan virtual laboratories in the region.

Virtual Offer

From the above results, it seems unlikely that virtual laboratories will be widely adopted by developing country scientists in the near future. One reason is the lack of adequate computing literacy in the academic community of countries that are new to the Internet. This is usually correlated with a lack of understanding of the potential offered by the network as a tool for spreading information (and contrast with its use for accessing information generated elsewhere to overcome scientific isolation). One implication is that *it is essential to familiarize the scientific communities with the use of Internet and of the many freeware tools that are currently available*.

Enlarging the demand for virtual laboratories means addressing questions such as: What does technology expects people to do? What are people prepared technically to digest? and how sensitive a society is to the change of mentality needed?. It is essential to work out how people will adapt to working in Virtual Labs.

Other important factors to understand are the human diversity, generational differences, and human relations based on the geographical location (and hence time difference) of partners in the virtual sessions. The electronic realm should meet the needs of the whole virtual community, old and young, avoiding an overload of information, asynchronicity and unnecessary usage complications [4,5]. Participation in interactive sessions from the South should make it easier to do science, not become another barrier to surpass. To increase the audience, organized information should be generated so that the volume and decentralization of available material does not limit the ability to access it. It should be set up to avoid redundancy of all distributed efforts, and should give potential benefits to most (if not all) participants. The administration of virtual laboratories should be simplified so that they can be run by scientists for the scientists in their own countries (see correspondence in [7]).

World-Wide Wait

Needless to say, connectivity in some Southern regions is in an early stage. In particular, sub-Saharan Africa has yet to make the kind of progress in both science and technology that has already been achieved in the more dynamic areas of growth in the South [2]. The limited bandwidth of the few available telecommunication lines in countries that are joining the Internet cause line congestion and make access exceedingly slow, often beyond the limit of usability. Furthermore, there remains a lack of computers in remote regions.

These problems reduce drastically the effectiveness of Internet as a working tool, and will delay the creation of South-South virtual laboratories.

To meet such problems, free support (or 'help desks') on information technology should be made available, offering 'pre-prepared kits' with programmes for enabling e-collaboration, as well as 'documentation' in conjunction with 'e-mail review newsletters' and 'on-site modular group training' by experts. All such activities would help to increase interest in virtual communities and effort at integration.

Therefore, *it is necessary to familiarize local system administrators with the most effective networked techniques for optimizing the use of bandwidth under continuous upgrades*.

For example, current Web technologies allow speed-based content negotiation to be performed. This means that different content can be transferred depending on the speed of the connection[8].

It is possible to navigate the whole Internet via e-mail (including dynamic content and preserving original pages layout), to transfer binary data (split into smaller blocks), to search scientific databases (such as the IAEA nuclear database), and to complete forms (such as Altavista, Yahoo). A standard Web browser and a MIME aware e-mail program is needed [9].

Grants or funding subsidies for paying fixed term Internet access should be made available to scientists from less-privileged countries. This would give them the opportunity to access the rich virtual environment, while the expected growth of Internet will dramatically lower the costs of access [2].

Multi-lingual facilities for non-Western character sets (ISO-8859-1), is another of the necessary requirements that would facilitate scientific communication throughout the developing world. For any system to become successful in the region, it should also offer full languages support [10]. Most Web browsers (besides AraMosaic, Sindbad) are not yet able to interpret bidirectional languages such as Arabic or Hebrew in simple terms, i.e., without the use of expensive auxiliary software, or plug-ins - a nightmare when connecting with limited bandwidth lines. This is being studied in the context of current HTML4.0 specifications and Mozilla user interfaces, which gathered the Netscape open-source initiative [11].

At a technical level, Linux O.S. can provide a cost-effective, alternative to help promoting distance e-collaboration. The open source software (OSS) is an added advantage to the Virtual Lab approach, though it may not be cheaper than proprietary software once the time spent in learning it is taken into account [12].

As for the telecommunication needs for a system of virtual scientific exchange, it would also be necessary to give wireless net access to scientists in remote or isolated areas, in order to overcome insufficient basic telecommunication services [13]. In so doing, mobile devices will play an important role in a world that is increasingly based on Internet-based global communications [14].

The use of VSAT technology is also likely to explode due to declining costs and the launch of new satellite series covering most continents on Ku-band, requiring smaller equipment. Costs can be sustained even further by using asymmetric satellite links (64 Kbytes rate or higher) to establish real-time audio/video, broadcasting and discussions with the

South. Other hubless options on a switched basis (e.g., TDMA-Time Division Multiple Access) could be tested depending on future needs and the growth of the system.

Within an international e-collaborative framework, the creation of large distributed SQL databases (for example for seismology research, such as 'Banco de datos CERESIS' [15]) is already under consideration in South-America. Within the same framework, it could be necessary to create a global database (or yellow pages) to enable a search for new virtual collaborators working in related topics.

Ruling, Rolling and Running

Planning, implementation and management are each valuable steps for creating virtual labroatories as an 'anti-brain drain' device. Motivated by new technological and human resources, a dynamic e-collaboration between Southern states could begin to flourish. The question therefore arises whether there is a need to establish rules in order to safeguard a sustainable development.

It is essential that fairly clear agreements will need to be reach between virtual working teams. In a practical sense, regulations agreed between two fixed-point users will differ from those working in a larger multi-user virtual infrastructure. In both cases, however, their overlapping interests will force negotiation. But they should be responsible for shaping the steps needed to enforce their virtual relations, and thus open new ways for interaction

For the settlement of, say, 'virtual' disputes - such as ownership of ideas - it would be necessary to provide assistance from ad-hoc advisory groups made up of experts on the issues in question. In the global village, such advisory teams could meet in virtual sessions and either provide recommendations, or merely ring alarm bells. It is soon to talk about the impact of common virtual sense, virtual friendship, virtual trust and ethics behind the scenes.

Science lives in the minds of men and women, not inside equipment and machines [16]. Science feeds on science. Men of science wish to work where other such men are, and where they have access to facilities [17]. The Virtual Lab approach could become a worthwhile service that helps to reduce scientific isolation while filling the need to transfer knowledge to the South in an unprecedented way.

References

[1] Web: www.news.com (March 1st, 1999).
[2] TWAS 1999 publication: *Profiles of Institutions for Scientific Exchange and Training in the South*.
[3] Web: xxx.lanl.gov (and site mirrors); see also discussion in *From Print to Electronic* (ASIS Monograph Series, 1996), S.Y. Crawford et al. Eds.; and also E. Canessa, *The Electronic Library* Vol.15 (1997) p.398.
[4] D. Butler, *Nature* Vol.397 (1999) p.195.
[5] G. Pastore and E. Canessa, *Computers in Physics* Vol.10 (1996) p.520.
[6] *UNESCO World Science Report*, H. Moore Ed. (Elsevier, 1998).
[7] F. Cabello, *Nature* Vol.397 (1999) p.557.
[8] Use Perl modules, such as *Apache::Throttle*, at www.modperl.org
[9] Freeware software at www.ictp.trieste.it/~www4mail
[10] Z. Khurshid, *The Electronic Library* Vol.16 (1998) p.171.
[11] Web: www.mozilla.org/projects.html
[12] A. Rubini, *Linux Journal* Feb. 1999, p.46.

[13] F. Postogna et al., *Linux Journal* Dec.1998, p.42.
[14] Web: w3c.org/Mobile
[15] Web: www.dgf.uchile.cl/ceresis.html
[16] M.J. Moravcsik, *On the Road to Word-wide Science* (World Scientific Publishing 1988).
[17] A. Salam, *Science in the Third World* (Alna Press, 1989).

Acknowledgements:
These Notes follow the IITAP initiative for *The Virtual Laboratory: Electronic Support for Cooperative Scientific Research*. Suggestions by Prof. J. Vary (Head IITAP, Iowa Univ. , USA) and Dr. G. Pastore (Trieste Univ., Italy) are acknowledged. Thanks are also due to Dr. A. Nobile (ICTP, Trieste, Italy) for valuable discussions.