Research Report

Internet Presence as Knowledge Capacity: The Case of Research in Information and Communication Technology Infrastructure Reform

Sujata Gamage
gamage@lirne.net
Director of Knowledge Networks

Rohan Samarajiva
Executive Director

LIRNE*asia*
12 Balcombe Place
Colombo 00080
Sri Lanka
+94 11 267 1160
www.lirneasia.net

Abstract

Knowledge is an important driver of development. As the production and dissemination of knowledge become increasingly mediated by the Internet, the Internet presence and connectivity of researchers are becoming more valid than the conventionally used publication- and citation-based indicators. This article presents a methodology that includes the use of the Google Scholar search engine to locate knowledgeable individuals in Asia in a policy-relevant field, paying particular attention to locating researchers in developing countries or in nonacademic settings in Asia. Internet presence is not a guarantee of quality. Increasingly sophisticaticated search engines offer viable means of assessing research quality and enable us to measure the connectivity of researchers on the Internet. Although the focus of the research is information and communication technology infrastructure reform in East, Southeast, and South Asia, the method can be used to assess knowledge capacity and locate knowledgeable individuals in any field.

1. Introduction

Knowledge is an important driver of development. Universities, research institutes, and other knowledge organizations in developing countries are often called upon to contribute to development without an accurate understanding of the knowledge capabilities of these institutions.

The capacity of institutions, countries, or regions to acquire and use knowledge for development is typically expressed in terms of knowledge inputs (e.g., number of researchers) and outputs (e.g., publications, citations, and patents) (World Bank, 2005; Wagner et al., 2004; Archibugi & Cocco, 2004). The three citation indices, the Science Citation Index (SCI), the Social Sciences Citation Index (SSCI), and the Arts and Humanities Citation Index (A&HCI), all compiled by Thomson Scientific, have become authoritative sources for assessing the extent and the quality of publications at all levels of analysis (Hicks, 2004; Wagner et al., 2004). A country or a region is the unit of analysis in all three indices.

The inadequacy of these conventional measures came to light as a result of a knowledge networking initiative by LIRNE*asia,* a regional ICT policy and regulation research and capacity-building organization.[1] The central activity of the knowledge-networking initiative was the compila-

1. Its mission is "To improve the lives of the people of the emerging Asia–Pacific by facilitating their use of ICTs and re-

tion of a directory of Asia-based scholars in ICT reform.

Our work with SSCI publications from 2000–2005 yielded only 119 names from the 25 countries in South, Southeast, and East Asia that were included in the study—no scholars were identified from 15 countries.[2] Using Google Scholar, a beta version of a new search engine launched in late 2004, increased the number of names to 377 and increased the number of countries represented to 17. Our research showed that an Internet search engine can uncover more scholars from developing countries than the SSCI.

Internet presence by itself does not give validity to research outputs or to a researcher. In academia, research outputs are validated through peer review. What constitutes quality in academia has been subject to much debate (Boyer, 1990; Diamond & Adam, 2004). Some consensus now exists, because the set of attributes identified by Diamond and Adam (2004) appears in most U.S. university faculty-evaluation criteria and perhaps also in other countries.

Diamond and Adam (2004) identified seven criteria that characterize a scholarly work. Their criteria include both process and output attributes. If we focus on outputs, following are the five criteria that characterize a scholarly work:

1. Requires a high level of discipline-related expertise;
2. The work and its results are appropriately documented and disseminated;
3. Has significance beyond the individual context;
4. Breaks new ground or is innovative;
5. Is reviewed and judged to be meritorious and significant by a panel of peers.

Google Scholar does not divulge much information about its search algorithm except to state that its Web crawlers seek out "papers, theses, books, abstracts, and articles from academic publishers, professional societies, preprint repositories, universities, and other scholarly organizations." Not all documents available from these miscellaneous sources are peer reviewed and judged to be meritorious and significant, but Google Scholar and other search engines are able to give up-to-date information on the number of citations received by any document available on the Internet and lists other works that cite each document. Some peer-reviewed publications included in the SSCI hardly receive any citations after publication, and even if they did, such information is not freely available. A search engine result that includes connectivity information, or the number of citations to and from a document, can be a valid and convenient source of information on the quality and relevance of a research output.

The terms *presence* and *connectivity* proposed here resonate well with the terms *link economy* and *hit economy* used in e-commerce to value an enterprise (Rogers, 2002):

> [A]n organization's Net presence derives from far more than site design and service delivery, and the maintenance of one's frame around the rest of the Web. One way to think through new notions of Net presence is to understand two types of Internet economies, the "hit economy" and the "link economy."
>
> Whether by portal or search engine placement, preferred sites are granted a larger audience (more hits). The organizational strategy thus revolves around establishing robust portal and search engine presence. In all, the combination of crawler-digestible coding, key word information design and favored placement is an organization's modus operandus in the hit economy. Robust Net presence is subsequently demonstrated on hit tables, which drive Web advertising, the seminal form of e-commerce.
>
> On the Web, "granting a link" (as a reference in science) and "receiving a link" (as a citation in science) are akin to positioning oneself and being positioned by another, respectively. Cognizance of such positioning may lead to consideration of presence strategy. (Rogers, 2002)

Drawing on the analogy between the terms proposed for research enterprises and those used in e-commerce, the following definitions are used in this article:

lated infrastructures; by catalyzing the reform of laws, policies, and regulations to enable those uses through the conduct of policy-relevant research, training, and advocacy with emphasis on building in-situ expertise."

2. This study focused on eight South Asian (Afghanistan, India, Pakistan, Bangladesh, Nepal, Bhutan, Sri Lanka, and the Maldives), 12 South East Asian (Brunei Darussalam, Cambodia, Hong Kong China, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Timor Leste, and Vietnam), and five East Asian (China, Taiwan ROC, Japan, DPR Korea, and RO Korea) countries.

- **Presence:** number of "hits" shown for a researcher on a scholarly search engine. Criterion is met by at least one validated hit.
- **Connectivity:**
  - Number and nature of links given
  - Number and nature of links received

These Web-based measures, or *Webometrics* (Björneborn & Ingwersen, 2004; Henzinger & Lawrence, 2004), represent the emerging modes of production, dissemination, and use of knowledge that are increasingly mediated by the Internet better than conventional measures that are limited to publications and citations in a selected set of journals. Webometrics are gaining acceptance, too. For example, a ranking of the Web presence of universities was published in July 2007.[3] Web presence is also used as an indicator to capture the complex processes that characterize the usability of research and innovation to research processes (Katz & Cothey, 2006). Web-centered metrics are not limited to knowledge-based interactions within communities; they are becoming central to the study of interactions within and between all sorts of communities (Adamic & Adar, 2001).

Internet search engines can be used by anyone with Internet access, even the most basic dial-up version. In contrast, proprietary databases such as SSCI are not easily available in most countries in Asia. For example, the SSCI is owned by Thomson Scientific. It indexes close to 14,000 international journals to provide a proprietary but authoritative basis for assessing knowledge products and producers worldwide. This database and the database of the U.S. Patent and Trademark Office (USPTO) are commonly used to gauge the extent and the quality of the knowledge base at the individual, organizational, national, or regional levels. What Webometrics may lose in rigor of measurement can be overcome by what they offer in terms of exposure to knowledge producers.

In this article, we present data that demonstrate the feasibility of using Internet presence and connectivity to identify and assess research capacity in ICT infrastructure reform. The discussion of connectivity is preliminary.

We begin with a short introduction to ICT infrastructure reform in section 2, followed by a discussion on current measures of knowledge capacity in section 3. connectivity attributes are discussed in section 4, focusing on the additional exposure received by scholars in developing countries as a result of their Internet presence. In section 5, we compare the ability of researchers in nonacademic settings with those in academic settings by using examples from India and Korea.

2. ICT Infrastructure Reforms in Asia

Infrastructure reforms are vital to economic development. ICT, energy, and transportation are three of the key infrastructures. Building infrastructure or providing infrastructure services, however, does not mean that the builders have to be inventors, too. For example, it is common nowadays for mobile network operators to completely outsource the design and even the operation of their network infrastructure to equipment manufacturers or others. Buying infrastructure inputs is often the more viable for infrastructure operators in developing countries than building them. When infrastructure operators buy inputs, further technical expertise is needed to adapt the existing technologies to local use (Gamage & Samarajiva, 2003).

Implementing the necessary institutional reforms in the ICT sector requires specialized knowledge, especially for reforms that are not limited to transactions or "big-bang" reforms (Samarajiva & Gamage, 2007). This specialized knowledge includes topics such as economics, law, and public administration in addition to an understanding of information and communication technologies and the ability to continually formulate and implement policies that enable the least costly and most beneficial options for infrastructure development. Donor agencies such as the World Bank may provide technical assistance for specific institutional reforms, but even this must be complemented by local counterpart capacity.

There are several organizations in Asia devoted to aspects of research intended to contribute to ICT infrastructure reforms that focus on particular countries or subregions or that look at ICT reforms as part of infrastructure reforms in general. A more coordinated effort is needed to develop and nurture capacity in ICT infrastructure reform across developing Asia. With this work, we hope to identify some tools for assessing and using the capacity in Asia or

3. See www.webometrics.info

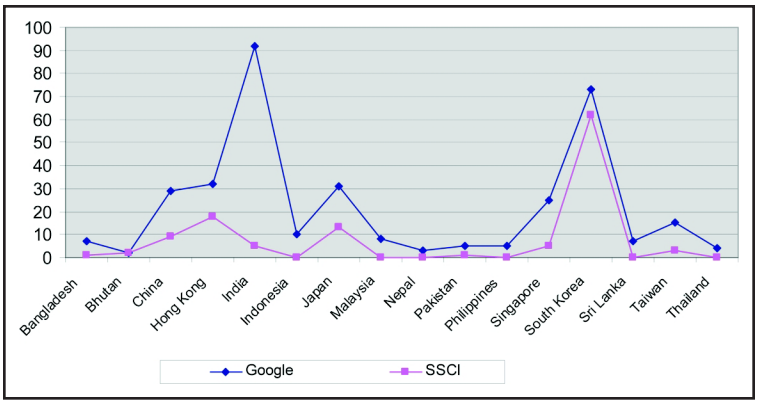

Figure 1. Distribution of scholarly documents on telecom reform by country and source, January 2000–August 2005.
Note: SSCI indicates Social Sciences Citation Index.

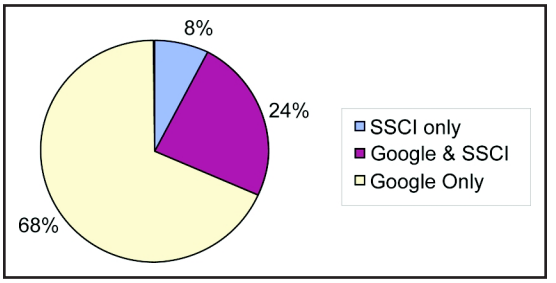

Figure 2. Distribution of scholarly documents on telecom reform by source, January 2000–August 2005.
Note: Sources were Google Scholar and SSCI, the Social Sciences Citation Index.

elsewhere to contribute to ICT infrastructure development through effective reform.

3. Presence

We searched the 2000–2005 SSCI for any records containing the keyword *telecom* or variants in the title or abstract, and any of the 22 country names in the address field (e.g., telecom* and India). The database was accessed through the Web of Science in August 2005. We manually filtered out entries that are not directly relevant to ICT infrastructure reforms to get 79 "Asian" records (or records with at least one Asian author) and 119 "Asian" authors, where "Asian" meant an author producing knowledge outputs from an address in Asia. We used the address of the author as given in the address field of a publication. If two articles yielded different locations, we conducted additional searches to determine the current location of an author. The search yielded names of scholars from 10 countries (Bangladesh, Bhutan, China, India, Japan, Hong Kong China, Pakistan, Singapore, Republic of Korea, and Taiwan [Republic of China]), but 16 other countries (Afghanistan, Brunei, Cambodia, Democratic People's Republic of Korea, Indonesia, Laos, Malaysia, Maldives, Myanmar, Nepal, Pakistan, Philippines, Sri Lanka, Thailand, Timor Leste, and Vietnam) were absent from the "map of knowledge" generated by SSCI.

Next, we used the term *telecom* and its variants and each country name to retrieve scholarly documents relevant to telecom infrastructure reform from Google Scholar. We retrieved only the first 200 hits for each country and manually filtered the records to get a total of 719 records, of which 226 had at least one Asia-based author, for a total of 348 authors. Google Scholar put six countries—Nepal, Indonesia, Malaysia, Philippines, Sri Lanka, and Thailand—back on the map.

All countries increased their presence with Google Scholar. India increased its presence significantly, from 5 to 92 records. South Korea increased its Internet presence only marginally compared with other countries (Figure 1).

While the SSCI search yielded only 119 names, the Google Scholar search yielded 348 names. Because 90 names were found in both datasets, the two sources combined yielded 377 names of which 258 names (or 68% of the total) were represented only in Google Scholar (Table 2).

The higher presence of researchers on Google Scholar relative to the SSCI reflects the fact that the criteria used by Google Scholar to identify a publication as scholarly are simple and mechanical. A Web crawler used by Google assesses metadata about a document published on the Web to determine the document's scholarly nature, but Google is not explicit about the specifics of metadata. SSCI, on the other hand, indexes only documents published by peer-reviewed journals. Panels of experts periodically

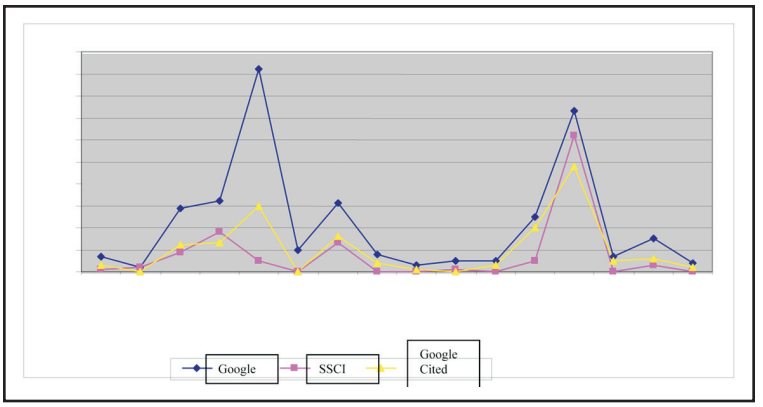

Figure 3. Number of scholarly documents on telecom reform by Asian authors by country and source, January 2000–August 2005.
Note: Sources were Google Scholar, the Social Sciences Citation Index (SSCI), and Google Scholar records with citations (Google Cited).

review which journals should be indexed and which journals should be removed.

4. Connectivity

Given the mechanical nature of the selection of documents by Google Scholar, Internet presence by itself should not be taken as an indicator of scholarly quality. Some of the records found on Google were preprints, conference proceedings, abstracts, and student papers that had not been validated through peer review. Google Scholar provides the number of citations for each record, and if there are several versions of the same article, they are presented as one record with the citations aggregated and the source of each version noted.

If we limit the scope to records with citations, all countries reduce their presence, and India's presence is reduced more than others. The "Google presence" of South Korea and Hong Kong is less than their "SSCI presence," and Bhutan, Indonesia, and Pakistan lose their presence altogether (Figure 3). The cases of South Korea and Hong Kong deserve special attention. Although their researchers have more validity in terms of SSCI publications, that validity is not commensurately reflected on the Internet.

Restricting the records to those with citations or those indexed in SSCI reduced coverage from 16 countries to 13 countries and researcher coverage to 224, but the overall coverage is still greater than the 10 countries and 119 researchers yielded by SSCI alone (Tables 1 and 2).

Table 1 *Number of Countries with Scholarly Documents by Asian Authors on Telecom Reform by Source, January 2000–August 2005*

Source	Number of countries covered
Google Scholar	16
Google Scholar with citations	13
SSCI only	10

Note. SSCI indicates Social Sciences Citation Index.

Using the number of citations as a validation tool may unnecessarily restrict some documents. Some of the publications may not have received citations because they were posted recently or because they did not provide full texts online. Therefore, it is advisable to maintain as many names as possible in a roster or a database and evaluate individuals as necessary, using complementary methods. From a usability point of view, it is better to have more names because citations only indicate academic relevance.

In fact, judging by the URLs of the citations, more than 61% of all citations on Google Scholar come from publications archived in university or academic publisher Web sites (Table 3).

Many individuals who read and use records from the Internet may never cite them, because their purpose is not the creation of scholarly works, but the application of the knowledge. Currently, we are unable to give credit to an article that is never cited by other scholars, but is read widely by practitioners. It is quite possible that search engines of the future will be able to go beyond citation information offered by Google Scholar to tell us how often a scholarly work is accessed on the Internet and whether it was accessed by individuals associated with an educational site, a government site, or another site, and even provide a network map of connections for each researcher (Rogers, 1996). In the future, search engines are likely to give a range of metadata (e.g., how many viewed the site, how

Table 2 *Number of Asian Researchers in Telecom Reform Appearing in Google Scholar, by the Validity of Their Publications, January 2000–August 2005*

Validity of Publications	No.	Percentage (%)
At least one publication is indexed in SSCI	119	32
Publications not indexed in SSCI, but at least one is cited by others with records in Google Scholar	105	27
Publications are not indexed by SSCI or cited, according to Google Scholar	153	41
Total	**377**	**100**

Note. SSCI indicates Social Sciences Citation Index.

Table 3 *Number and Percentage of Citations of Documents on Telecom Reform by Asian Authors in Google Scholar by Citation Source, January 2000–August 2005*

Citation source	No. citations	Percentage (%)
University Web sites	526	30
Publisher Web sites	539	31
Other, unidentified	1,020	39
Total	**2,085**	**100**

Note. SSCI indicates Social Sciences Citation Index.

many downloaded the document) in addition to the citation information currently provided.

If such metadata were available, it would be possible to map the links from any one researcher to other individuals who had cited or downloaded the researcher's work. In the absence of a search engine that generates such metadata, a practical alternative would be to interview a representative sample of researchers in ICT infrastructure reform to identify how their work has been directly or indirectly communicated to other stakeholders in the sector.

Over all, Internet presence gives us a knowledge base that can be used as a starting point for further exploration of quality and relevance and gives exposure to researchers from developing countries who would not be known otherwise.

5. Academic Versus Nonacademic Researchers

The conventional methods of assessing knowledge capacity using publications in peer-reviewed journals give greater weight to knowledge produced in academic settings, because the SSCI-indexed journals are academic journals. With advances in the Internet and associated technologies, it is now possible for institutions and individuals outside academia to give and take knowledge through their presence on the Web. A change in the nature of the knowledge landscape was predicted by Gibbons and others as early as the 1990s in what they termed "the changing modes of knowledge production" (Gibbons et al., 1994). They termed knowledge produced in formal settings such as universities and research institutes as *mode 1* and knowledge produced in workplaces and other settings as *mode 2,* and predicted that mode 2 will gradually supersede mode 1. Although the concept is widely used in the literature, empirical work supporting it is lacking.

Internet search engines now make it possible to identify individuals and organizations who may not necessarily produce knowledge as part of their daily work or are not motivated or connected sufficiently to academic knowledge networks to receive exposure for their work. In this study, we attempt to make an empirical contribution to the mode 2 versus mode 1 debate.

In Table 4, we summarize the distribution of researchers by type of organization for a subset of authors from India and Korea who have received one

Table 4 *Distribution of Authors With One or More Citations From Telecom-Reform-Relevant Documents by Type of Organization in India and South Korea, January 2000–August 2005*

	Percentage of authors (%)	
Type of organization	**India (n = 29)**	**South Korea (n = 47)**
All universities	44	71
Research universities*	34	28
Other universities	10	43
Industry	17	11
Government	7	0
Nonprofit and other	31	19

Note. Sources were Google Scholar and the Social Sciences Citation Index.
**Research universities are those identified as such by* Asia Week *magazine in its 2000 "Survey of Best Universities in Asia"; all others are considered nonresearch universities.*

or more citations on Google Scholar. We identified the home institution of each researcher to derive the distribution shown below.

In India, the contribution of nonuniversity actors is 55% of the total. While it remains to be seen whether nonuniversity actors are equally prominent in other areas of research in India, the data for telecom reform researchers are in line with the postulate that knowledge production is increasingly being undertaken outside of the university (Gibbons et al., 1994; Delanty, 2001). As Delanty (2001, p. 6) states:

> Today knowledge has become more important and at the same time does not emanate from any one particular source. This restructuring in the mode of knowledge implies not the end of the university but its reconstitution. The great significance of the university today is that it can be the most important site of interconnectivity in what is now a knowledge society.

In the case of India, we would not have discovered 26 out of the 29 (or 90%) researchers, academic or nonacademic, if not for Internet search engines. Of the search engine discoveries, 50% are nonacademics.

In the case of Korea, we would not have discovered 15 out of the 47 (or 32%) researchers, academic or nonacademic, if not for Internet search engines. Nearly 50% of this group was constituted by nonacademics.

The Indian and Korean data on research on telecom reform clearly show that the Internet provides a unique means of locating nonacademic researchers; it also shows that many academic scholars, too, would not be discovered if not for Internet search engines. There is clearly a need to impress upon researchers in academia, and researchers or reflective practitioners outside of academia, the importance of maintaining a presence on the Internet.

In Korea, the nonresearch universities occupy an unexpectedly larger space on the knowledge map, suggesting a problem in the differentiation of universities as *research* and *nonresearch* institutions and pointing to the need for better differentiation of universities in the region.

6. Conclusions

The current focus on capacity as captured by instruments such as the SCI or SSCI may miss knowledge activities in smaller developing countries and in policy- or development-related fields. Publications in international journals do not come easily to those unconnected to international networks. There is also a finite amount of space in international journals to accommodate scholarly works from across the world.

As Internet search engines become ubiquitous and more sophisticated, an officer in the Telecommunications Authority of the Maldives, for example, is now able to type "telecom policy" in Google Scholar and begin a well-informed initiative to contact experts from any country. It is inconceivable that conventional citation indices will be used in this manner by individuals in policy or regulatory fields. If

universities and researchers wish to become better partners in development, they will have to better inform the world about their expertise—the Internet is the most logical means by which to do that.

Universities no longer have even a semblance of a monopoly on knowledge production, especially universities in the developing countries which, arguably, never had a monopoly on knowledge production. To locate knowledgeable people, the focus should be on the Internet presence and connectivity of individuals, irrespective of their affiliation or how they are represented in established citation indices. ■

References

Adamic, L. A., & Adar, E. (2001, May). You are what you link. Paper presented at the 10th International World Wide Web Conference, Hong Kong, China. Retrieved November 30, 2007, from http://www10.org/program/society/yawyl/YouAreWhatYouLink.htm

Archibugi, D., & Cocco, A. (2004). A new indicator of technological capabilities for developed and developing countries. *SPRU Electronic Working Paper Series* (Paper No. 111).

Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for Webometrics. *Journal of the American Society for Information Science and Technology, 55*(14), 1216–1227.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate.* Princeton, NJ: The Carnegie Foundation for Advancement of Teaching.

Delanty, G. (2001). *Challenging knowledge: The university in knowledge society.* Buckingham, UK: Open University Press.

Diamond, R. M., & Adam, B. E. (2004). Balancing institutional, disciplinary and faculty priorities with public and social needs: Defining scholarship for the 21st century. *Arts and Humanities in Higher Education, 3*(1), 29–40.

Gamage, S., & Samarajiva, R. (2003). Science and technology inputs for information and communication infrastructure development in Sri Lanka. In *Proceedings of the Second Biennial Conference on Science and Technology.* Colombo, Sri Lanka: National Science and Technology Commission.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). The new production of knowledge. In *The dynamics of science and research in contemporary societies.* Thousand Oaks, CA: Sage.

Henzinger, M., & Lawrence, S. (2004, April 6). Extracting knowledge from the World Wide Web. *Proceedings of the National Academy of Science, 101*(1), 5186–5191.

Hicks, D. (2004). Asian countries strengthen their research. *Issues in Science and Technology,* Summer, 75–78.

Katz, J. S., & Cothey, V. (2006). Free Content Web indicators for complex innovation systems. *Research Evaluation, 14*(2).

Rogers, R. (1996). The future of STS on the Web. *EASST Review, 15*(2).

Rogers, R. (2002). Operating issue networks on the Web. *Science as Culture, 11*(2).

Samarajiva, R., & Gamage, S. (2007). Bridging the digital divide: Building Asia–Pacific capacity for effective reforms. *The Information Society, 23*(2), 109–117.

Wagner, C. S., Brahmakulam, I., Jackson, B., Wong, A., & Yoda, T. (2004). *Science and technology collaboration: Building capacity in developing countries?* RAND Science and Technology.

World Bank. (2005). *Knowledge assessment methodology.* Retrieved September 30, 2005, from http://info.worldbank.org/etools/kam2005/index.htm