Assessment of India's research literature [☆]

Ronald N. Kostoff [a,*], Dustin Johnson [a,1], Christine A. Bowles [b], Sujit Bhattacharya [c],
Alan S. Icenhour [a], Kimberly Nikodym [a], Ryan B. Barth [b], Simha Dodbele [a]

[a] *Office of Naval Research, 875 N. Randolph St., Arlington, VA 22217, USA*
[b] *DDL-OMNI Engineering, LLC, 8260 Greensboro Drive, Suite 600, Mclean, VA 22102, USA*
[c] *National Institute of Science, Technology and Development Studies (NISTADS), Pusa Gate,
K.S. Krishnan Marg, New Delhi-110012, India*

Abstract

The structure and infrastructure of the Indian research literature were determined. A representative database of technical articles was extracted from the Science Citation Index/Social Science Citation Index (SCI/SSCI) [SCI. Certain data included herein are derived from the Science Citation Index/Social Science Citation Index prepared by the THOMSON SCIENTIFIC®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: ©Copyright THOMSON SCIENTIFIC® 2006. All rights reserved. [1]] for 2005, with each article containing at least one author with an India address. Document clustering was used to identify the main technical themes (core competencies) of Indian research. Aggregate India bibliometrics were also performed, emphasizing the value of collaborative research to India. A unique mapping approach was used to identify networks of organizations that published together, networks of organizations with common technical interests, and especially those organizations with common technical interests that did not co-publish extensively. Finally, trend analyses were performed using other year data from the SCI/SSCI to place the 2005 results in their proper historical context. Published by Elsevier Inc.

Keywords: India; Science and technology; Technology assessment; Core competencies; Research evaluation; Metrics; Bibliometrics; Text mining; Computational linguistics; Document clustering; CLUTO; Auto-correlation mapping; Cross-correlation mapping; Factor analysis; Factor matrix; Impact Factor

1. Introduction

South–East and East Asia have become dynamic growth areas, especially in science and technology (S&T) (see for example [2]). Our text mining studies of specific technologies over recent years have shown dramatic growth in research output production by China, South Korea, Taiwan, and Singapore (e.g., [3]), to name a few. As a result, we have started to adopt a national view of research output from some countries in the region, and are examining research products from individual countries. The preceding paper in this Special Issue was focused on an assessment of China's research enterprise. The present paper focuses on India's research enterprise, and the next paper in this Special Issue will compare the research outputs of the two countries.

The primary objective of the present study is to identify the S&T core competencies of India. In addition, temporal trends of significant research-related parameters will be presented. These trends will provide a context in which to interpret India's present research output status, and will provide support for the predictive conclusions that follow.

2. Background

The present study combines three concepts/approaches for the assessment of India's S&T literature: core competency determination, country technology assessments, and text mining assessments. The background for these three concepts, as well as a description of India's S&T enterprise, can be found in the Introduction of this Special Issue. India's S&T performance based on research output literature will now be summarized.

3. Approach and results

3.1. Overview

A taxonomy and detailed bibliometrics analyses are presented for 1 year (2005). Gross bibliometric trends are presented to place the detailed 2005 bibliometrics in perspective. The databases used for the bibliometrics and taxonomy analyses, the bibliometrics approaches, and the document clustering taxonomy approach, are described in the Introduction of this Special Issue.

3.2. Bibliometrics

Publication and citation bibliometrics were performed at the aggregate national level. In addition, bibliometrics of four core technologies were examined, and can be found in detail in [4].

3.2.1. Overall India bibliometrics
This section presents temporal publication trends, journals containing most articles, journals cited most frequently by Indian authors, most prolific institutions, and most collaborative countries for the aggregate India database.

3.2.1.1. Publication trends.   The first metric is number of articles as a function of time. All research articles in the SCI/SSCI having at least one author with an India address were retrieved for selected years