Measuring and characterizing international collaboration patterns in Indian scientific research

Jyoti Dua1 · Vivek Kumar Singh1 · Hiran H. Lathabai2

Received: 23 September 2022 / Accepted: 10 July 2023 / Published online: 27 July 2023 © Akadémiai Kiadó, Budapest, Hungary 2023

Abstract
Scientific collaboration at international level has increased manifolds during last two decades. Collaboration is not only associated with higher research productivity but has also been found to be positively correlated with impact. Considering the benefits and advantages of international collaboration for the national science and technology systems of a country, policymakers in different countries have designed programs to promote international collaboration in science. This paper attempts to measure and characterize the international collaboration patterns in Indian scientific research for the last 20 years (2001–2020) using data from WoS. The existing methods from scientometrics and a new formalism of boost in productivity, citations and visibility, is used for the purpose. Analytical results show that India’s international collaboration has grown with a CAGR of 12.27% during this period, rising from 20.73% internationally collaborated papers in 2001 to 32.35% internationally collaborated papers in 2020. USA, Germany, England, and China remain the top collaborating partners in terms of volume of papers, however, the relative intensity of collaboration with South Korea and Saudi Arabia has increased significantly. Among the internationally collaborated papers of India, about 50% papers have an Indian researcher as lead (first) author, and further the internationally collaborated papers of India have more than 50% authors from India. The Indian indigenous scholarly ecosystem is found to be significantly productive, but slight and heavy dependence for impact and visibility, respectively, is found on international collaborations. The probable factors shaping the Indian international collaboration and the major policy implications for India are discussed.

Keywords Indian science · International collaboration · International research collaboration · Research collaboration · Scientific collaboration

Vivek Kumar Singh
vivekks12@gmail.com

1 Department of Computer Science, Banaras Hindu University, Varanasi 221005, India
2 Amrita-CREATE, Amrita Vishwa Vidhyapeetham, Amritapuri, Kollam 690525, India
**Introduction**

Collaborative research has become an integral part of any flourishing academic or industrial ecosystem. Collaboration and its benefits are well-studied from different perspectives. Mutual benefits in the form of resource sharing and knowledge transfer, complementary and/or common solutions for shared problems, etc., are the major drivers of collaborative research. Several studies have analyzed collaboration at the international level and have observed that it has risen linearly during the last two 2–3 decades, as measured in terms of number of internationally co-authored papers published (Adams, 2012; Glänzel, 2001; Lee & Bozeman, 2005; Leydesdorff & Wagner, 2008; Mattsson et al., 2008; Persson et al., 2004; Wagner & Leydesdorff, 2005). A strong affinity between collaboration and research productivity has also been observed along with the indications that collaboration may have a significant impact on citations (Abramo et al., 2011, 2017; Glänzel, 2001). The benefits of collaboration on research productivity were also explored by Abramo et al. (2009), Duc- tor (2015), Parish et al. (2018), etc. In an industrial context, inventor collaborations are found to have a profound effect on productivity (Favaro et al., 2012). It is to be noted that, after recognizing the benefits of international research collaborations, several national and international funding agencies have invested in policies that are intended to foster collaborative research (Boekholt et al., 2009; Jeong et al., 2014; Katz & Martin, 1997; Wagner & Leydesdorff, 2005; Wagner et al., 2001). The emphasis by different countries to engineer international collaborations indicates the vital role of collaborations in the progress of Science and Technology. Efforts to create and foster productive collaborations between nations are also a consequence of the increased awareness of the importance of collaborative research. The pursuit towards United Nation’s Sustainable Development Goals or SDGs calls for proactive research and development efforts at a global scale in different areas, which will require intense collaboration between countries.

In the knowledge-based economy, in terms of scientific, technological, and other types of knowledge, every nation is a prosumer (producer as well as consumer). However, the level of their production and consumption of knowledge varies in different degrees in different fields according to their overall capacities (indigenous capacities as well as the capacities built through external and/or extramural collaboration and various other schemes). India is steadily growing as a major knowledge producer, as is evident from its ranking in the top order by various study reports. For instance, two studies commissioned by DST, carried out by Elsevier’s Analytical Services (DST-Elsevier, 2020) and Clarivate Analytics (DST-Clarivate, 2019) placed India at the 5th and 9th positions, respectively, in research output volume. According to the National Science Foundation (NSF) report on Science and Engineering indicators (2020), India is ranked at 3rd position in terms of global research output (NSF Report, 2020). Though such a leap seems to be rapid, it is a cumulative consequence of a continuum of efforts through many decades. Establishment and management of fruitful collaborative ties with reputed international institutions or agencies had been one of the vital efforts that paid off well. The science and technology research in India post-independence has drawn a lot from different international collaborations. In fact, all the first-generation Indian Institutes of Technology involved collaborated approach at international level. For example, IIT Bombay (one among the institutions recognized by Govt. of India as institutions of eminence)

---

1 https://sdgs.un.org/goals.
was established in 1958 with funding from the then USSR and assistance from UNESCO (IIT Bombay-Institute history, 2014).

Over the last two decades, there have been major shifts in the global S&T landscape and India’s international interactions and knowledge networks have also evolved. Considering the fact that Reddy (2014) attributed the relatively low impact of Indian research (despite its 10th position in global share of research output) in his study over the period of 1996–2012 to the lesser share of collaborations on a scale comparable to top ranked countries, it is highly important to analyze whether collaborations have improved over time adequately. To some extent, the evidence or indications of improvement in collaboration can explain the increase in a country’s productivity and impact. There lies the importance of measuring the country’s international collaboration at this point in time. Even more important is the characterization or examination of the nature or characteristics of international collaborations. Specific concerns of such an examination are the identification degree/level of collaborations with countries and among different fields of research to determine top collaborating countries, top research areas/fields of collaboration, knowledge flows in such collaboration.

All the above factors increase the relevance of studies related to collaboration and address the concerns such as (i) assessment of the present status of collaborative research with respect to a country/region or any other entity, and (ii) characterization of collaboration patterns with respect to that entity. There lies the motivation and novelty of this research. In this work, we attempt to address these concerns for the country ‘India’. Thus, the objectives of this paper are to measure and characterize international collaboration patterns in Indian scientific research during 2001–2020 and understand how these collaborations have evolved recently, including discussing the relevant policy implications thereof. The international collaboration patterns for India are also analyzed through the new proposed formalism of Boost indicators, as discussed later. More precisely, the paper attempts to answer the following research questions:

RQ1: What proportion of research output of India involves international collaboration and whether the propensity to collaborate internationally has increased over time?

RQ2: What are the major collaborating research partner countries for India and how has the relative intensity of collaboration of India changed over time?

RQ3: Do the international research collaboration patterns vary across subject areas?

RQ4: What proportion of internationally collaborated papers from India have the first and/or majority of the author(s) from India?

RQ5: Do internationally collaborated papers of India get higher impact and visibility as compared to indigenous papers? If so, how much boost they provide to productivity, impact and visibility?

Achievement of these objectives can be vital for the extrication of suitable policy implications to improve the overall research output and impact of India. Further, answers to these questions present new insights about India’s international collaboration patterns. In all the above questions, if India is replaced by any other country name, and if analysis is carried out accordingly using the methodology used in this work, these questions can be answered in the context of that country. Thus, this work is not merely a descriptive analysis with respect to a country but provides a solid and diligent combination of existing methods as well as novel methods to answer these research questions in the context of any country.
Related work

Research collaboration involves a group of researchers working together as a team to achieve some common goals. Collaboration can involve cooperation at different levels—individuals, institutions, and countries (Katz & Martin, 1997). Several studies have analyzed collaboration at the international level and have observed that it has risen linearly during the last two 2–3 decades, as measured in terms of the number of internationally co-authored papers published (Adams, 2012; Glänzel, 2001; Lee & Bozeman, 2005; Leydesdorff & Wagner, 2008; Mattsson et al., 2008; Persson et al., 2004; Wagner & Leydesdorff, 2005). A strong affinity between collaboration and research productivity has also been observed, along with indications that collaboration may have a significant impact on citations (Abramo et al., 2011, 2017; Glänzel, 2001). The internationally collaborated papers are particularly found to attract higher citations as compared to domestically collaborated papers (Bordons et al., 1996; Bote et al., 2013; Glänzel & Schubert, 2001; Nguyen et al., 2017; Persson, 2010). It has also been pointed that, countries that are highly “open”, produce high impact research; where “open” is defined as having government policies that promote international collaboration and scientific mobility (Wagner & Jonkers, 2017).

Considering the benefits and advantages of international collaboration for the national science and technology systems of a country, policymakers in different countries have designed programs to promote different kinds of international collaboration (Boekholt et al., 2009; Katz & Martin, 1997; Wagner et al., 2001). There have also been efforts to identify the main driving factors of international collaboration in scientific research. Some of these major driving factors identified include science policies (Glänzel et al., 1999), need for use of complex equipment and technical infrastructure (Beaver, 2001; Birnholtz, 2007; D’Ippolito & Ruling, 2019), increasing specialization of science (Leahey & Reikowsky, 2008), need for knowledge sharing and training of scientists (Beaver, 2001; Katz & Martin, 1997) and globalization of science. The notion of proximities/distances is another important aspect analyzed in relation to international research collaboration (Balland et al., 2015; Boschma, 2005; Vieira et al., 2022).

In the broader context of the relevant factors and benefits and advantages of international collaboration, many previous studies focused their attention on measuring and understanding the international collaboration patterns of individual countries, such as for Brazil (Leta et al., 2002; McManus et al., 2020), Mexico (Marmolejo-Leyva et al., 2015), Korea (Kim, 2005), Vietnam (Nguyen et al., 2017), UK (Adams et al., 2007), and Russia (Pislyakov & Shukshina, 2014) etc. Several other previous studies focused on international collaboration in different regions, such as Africa (Adams et al., 2014; Mégñibêto, 2013), Asian region (Arunachalam & Doss, 2000; Arunachalam et al., 1994; Uddin & Singh, 2014), European region (Braun & Glaenzel, 1996; Glänzel et al., 1999), and the BRICS countries (Bouabid et al., 2016; Finardi, 2015; Finardi & Buratti, 2016; Leta et al., 2019; Singh & Hasan, 2015; Sokolov et al., 2019).

Some previous studies have analyzed the international collaboration in Indian scientific research (Anuradha & Urs, 2007; Arunachalam & Viswanathan, 2008; Basu & Aggarwal, 2001; Basu & Kumar, 2000; Gupta & Dhawan, 2003; Gupta et al., 2002). For example, Basu and Kumar (2000) analyzed the extent of international collaboration in Indian science for the 1990 to 1994 period from Science Citation Index (SCI) data and found an increase both in terms of output and the extent of the network. The bulk of Indian scientific co-operation was with developed Western nations and Japan. Basu and Aggarwal (2001) analyzed the data from major Indian institutions for 1997 from SCI and quantified the gain in impact.
through foreign collaboration. Their cluster analysis showed that different institutions had
different patterns of gain from international collaboration. Gupta et al. (2002) analyzed the
1992–1999 SCI data for India to find out how much did India collaborate with other South
Asian neighbors and found varying amount of collaboration with different South Asian
countries. Gupta and Dhawan (2003) analyzed India’s collaboration in Science & Technol-
ogy with China during 1994–1999 and found that the number of collaborated papers grew
from 21 in 1994 to 74 in 1999. However, it was also found that S&T collaboration between
India and China has been taking place mainly through multilateral channels and the output
through bilateral channels was very small (11.7%). Anuradha and Urs (2007) used corre-
spondence analysis to compute bibliometric indicators of India research collaboration pat-
terns during 1993 to 2000. Arunachalam and Viswanathan (2008) analyzed the Indian pub-
lication data for the 2000–2007 period to understand the collaboration between India and
China. The number of collaborated papers was found to have increased from 124 papers in
2000 to 361 in 2007, with multilateral collaboration being the main channel of collabora-
tion. Mahbuba and Rousseau (2010) also carried out a study on scientific research in the
selected countries in the Indian subcontinent and compared them with India.

Despite the above-mentioned previous studies having analyzed the international col-
laboration patterns in Indian scientific research, there is a need for a fresh and up to date
analysis due to various reasons. First, most of the existing studies on Indian international
research collaboration are quite old, with the most recent being done about 15 years ago,
whereas the Science and Technology research landscape of the world has changed signifi-
cantly since then. Second, the majority of the previous studies for India are for the pre-ICT
revolution era, and the collaboration patterns may have changed significantly since then as
ICT has diminished different kinds of distances in scientific collaboration. Third, the exist-
ing studies lack subject-specific collaboration pattern analysis. Fourth, the leadership role
and knowledge flows through international collaboration of India are not known. Fifth, to
the best of our knowledge, there is no existing analysis on the impact of international col-
laboration on citation and social media visibility of Indian research. This article, therefore,
emerges to bridge these research gaps by analyzing international collaboration patterns in
Indian scientific research during 2001–2020. In addition to measuring and characterizing
international collaboration patterns, the article also presents a discussion on the policy
implications of the observations.

Data and method

The research publication data of India for the period 2001–2020 has been used for the
bibliometric analysis. The data is downloaded from Web of Science database core col-
lection by using the search query: $CU = \text{India}$ and $PY = (2001–2020)$ and $LA = \text{English}$,
where CU field refers to country name, PY to publication year and LA to the language.
Publication data downloaded was restricted to document types ‘Article’ and ‘Review’, as these are considered the most important scholarly outputs. The data downloaded
was pre-processed to remove duplicate and null values. A total of 784,157 publication
records were retrieved in this way. There were 60 metadata fields in the publication
record data downloaded, such as author address (C1), document type (DT), publication
type (PT) etc. The analysis mainly involved processing of information in DI, C1, WC
and Z9 fields. The DI represent digital object identifier (DOI) and C1 denotes author’s
affiliation, including affiliation country, WC represents web of science categories, and
Z9 denotes citation count. The C1 field information was used to identify publication records that involved authors only from India (indigenous papers) and those that had authors from some other countries in addition to India (internationally collaborated papers). A total of 200,498 publication records were identified as internationally collaborated papers (ICP) and 583,659 publication records were indigenous papers.

The publication records of the two categories were then analyzed to compute different patterns. The year-wise ICP proportion for India was identified and plotted. The Compound Annual Growth Rate (CAGR) of the ICP instances as well as India’s total output is also computed. The CAGR is computed by using the following formula:

\[
\text{CAGR} = \left( \frac{V_{\text{final}}}{V_{\text{begin}}} \right)^{\frac{1}{t}} - 1 \times 100, 
\]

where, \(V_{\text{final}}\) is number of publication records in the year 2020, \(V_{\text{begin}}\) is the number of publication records in the year 2001, and \(t\) is the time period in years. The major collaborating countries for India and the number of collaborated papers with them was identified and plotted next.

Thereafter, the relative intensity of collaboration (RIC) (Fuchs et al., 2021) of India with respect to 20 major countries was computed for the given period. For this purpose, the number of publication records for the 20 countries was obtained from Web of Science. The RIC is formulated as the ratio of the share of the collaborations of actors X and Y within all collaborations of X to the share of collaborations of Y within all collaborations of the system excluding collaborations of X. It is expressed as:

\[
\text{RIC}(X, Y) = \frac{C_{xy} \times (T - C_x)}{C_x \times (C_y - C_{xy})}, 
\]

where, \(C_{xy}\) denotes the number of collaborations between two countries X and Y, \(C_x\) is the total number of collaborations of country X, \(C_y\) is the total number of collaborations of country Y and T represents the total number of pairwise collaborated publications of countries under study. The rationale behind RIC index is that it does not depend solely on \(C_{xy}\) rather it depends on other factor \(C_y\) also. In short, we can say, \(RIC(X, Y)\) is proportional to \((C_{xy}/C_y)\). Here, it may be noted that there are several other indices available for analysis of collaboration patterns, such as Jaccard and Salton-Ochiai indices and Probabilistic Partnership Index (Yamashita et al., 2006). There are also other indicators to evaluate the performance of a country in scientific research based on size-dependent and size-independent metrics (Mahbuba & Rousseau, 2010). In the present case, RIC computations are favored as it is asymmetric, that is, its value for collaborating units X and Y differs from its value between Y and X and it studies the relative importance of country X for country Y while taking other relations into account such as, total weight of the collaboration network. However, other indices can also be computed, and a more detailed comparison among these indicators could be made in future subject to availability of sufficient data.

Next the subject area-wise international collaboration patterns are analyzed. For this purpose, the publication records are categorized into 14 broad subject areas by using the scheme proposed in (Rupika et al., 2016). These 14 broad subject areas are: Agriculture (AGR), Arts & Humanities (AH), Biology (BIO), Chemistry (CHE), Engineering (ENG), Environment (ENV), Geology (GEO), Information (INF), Material (MAR), Mathematics (MAT), Medicine (MED), Multidisciplinary (MUL), Physics (PHY), Social Science (SS). The international collaboration patterns in these 14 subject areas
are identified, and the major collaborating countries in each of these subject areas are listed.

The next analysis involved identifying what proportion of India’s ICP instances have an Indian author as first/lead author. This was done by processing the author affiliation field data. Similarly, the affiliating country information for all authors in a paper was processed to find out how many of the ICP instances of India have more than 50% authors from India. The subject area differentiation for these patterns is also identified.

Finally, the citation impact and social media visibility of ICP and indigenous papers from India were computed and compared. The citation count was obtained from the Z9 field of the publication metadata. As, traditional scientometric indicators such as citations, show the impact of research on a specific domain; they do not capture the wider relevance of the research. Therefore, alternative metrics (Twitter, Facebook and Blogs) that may provide a broader range of impact of the research which add attention score to the publications, were also considered. These can be used to assess how far the research is disseminated on social platforms. The social media visibility on different platforms was obtained from Altmetric.com. In order to obtain altmetric data from https://www.altmetric.com/, a DOI look up was performed for all the DOIs in the WoS data. Out of the 784,157 publication records, a total of 187,566 records are found to be covered by Altmetric.com, which is about 23.91% of the total data. The altmetric data was obtained for a total of 74,035 publications out of 200,498 ICP instances, and 113,531 publications out of 583,659 indigenous publications. Altmetric.com has 46 fields in the data, including DOI, Title, Twitter mentions, Facebook mentions, News mentions, Altmetric Attention Score, OA Status, Subjects (FoR), Publication Date, URI, etc. Out of these, this study mainly used data for Twitter, Facebook, and Blog platforms. The data from Altmetric.com was downloaded in the month of August 2021.

RQ 5 deals with the effectiveness of collaboration in terms of increased productivity, impact, and visibility by virtue of publications, citations, and altmetric attention. First part of the question can be answered by comparing the mean values of collaborated and non-collaborated sets of publications, citations, and altmetric attentions through a standard statistical test namely the z-test. The second part can be answered by determining the net boost or increase in productivity, impact, and altmetric attention scores. This can be done with the help of an indicator that considers and expresses the indigenous productivity, impact, and visibility against the total productivity, impact, and visibility. As such an attempt is not found in the related literature, we devise some simple indicators that can be useful in this regard.

While existing collaboration studies mostly explored whether collaborations increased the productivity of a country and impact of a country, to what extent a particular country is benefited in its productivity and impact by collaboration is not well studied. For instance, Ni and An (2018) studied the relationship between international collaboration, number of papers, and number of citations from an economic perspective by analyzing international economic collaboration types, the number of countries at different economic levels, and the economic situations of the first or corresponding authors. That work has not determined the extent of boost/increase collaborations provided to the productivity and impact of such countries. Also, while collaboration can be beneficial for countries, excess collaboration is not healthy (this fact also serves as the rationale behind the usage of fractional count in some productivity assessment studies). When it comes to international collaborations, an excess of international collaboration can indicate the possibility of weakness in indigenous scholarly ecosystem of the country. This being the nature of collaborations, indicators that are capable of reflecting this
nature of collaboration is necessary. As the extent of increase or boost in productivity, impact, visibility etc., are supposed to be reflected by these indicators, we name these indicators as boost indicators. These are discussed next.

**Productivity boost (βₚ)** It is the ratio of total productivity (indicated by total number of publications or TP) to the total indigenous productivity (indicated by total number of indigenously produced publications or TIP) of a country, expressed in percentage.

\[
βₚ = \left[ \frac{TP}{TIP} - 1 \right] \times 100\%.
\]

If there are collaborations, \(βₚ > 0\). Higher the value of \(βₚ\), greater the boost in productivity due to collaborations. However, higher values of boost can also indicate greater dependence or reliance of a country on international collaborative research which can in turn indicate the prevalence of a weak scholarly ecosystem in a country. It is difficult to determine what is the ideal value of \(βₚ\). As a rule of thumb, if \(βₚ > 50\%\), a country is more reliant/dependent on foreign collaborations for productivity than indigenous scholarly ecosystem. If \(βₚ > 100\%\), country has gone too far in dependence and has to make extra proactive efforts to work towards self-sufficiency. If a country has an infinite \(βₚ\) (\(TIP=0\) and \(TP\) value of 1 or above), it signifies absolute dependence on collaboration. Countries which are in the beginning stage of establishing a research ecosystem are more likely to have such a \(βₚ\).

**Citations boost (βₖ)** It is the ratio of total citations (TC) to the citations received by total indigenous publications (TIC) of a country.

\[
βₖ = \left[ \frac{TC}{TIC} - 1 \right] \times 100\%.
\]

As a rule of thumb, if \(βₖ > 50\%\), a country is more reliant/dependent on foreign collaborations for impact than indigenous scholarly ecosystem. If \(βₖ > 100\%\), country is much dependent or indigenous productivity is attracting very low relative impact and reach. Country has to make extra proactive efforts to work towards choosing impactful sources to disseminate their research output or improving visibility and reach of their indigenous scholarly works or both.

Similarly, visibility related boost measures determined from various altmetric sources such as twitter, facebook and blogs can be computed in the same way as that of \(βₖ\) and are denoted as \(βₖ(t), βₖ(f), βₖ(b)\), respectively.

Boost ratio of impact per unit boost in productivity (\(γₖ\)): It is the net boost of impact per unit boost of productivity due to international collaborations.

\[
γₖ = \frac{βₖ}{βₚ}.
\]

If \(γₖ < 1\), collaborations are less rewarding and if \(γₖ > 1\), collaborations are rewarding. Greater the value of \(γₖ\), greater the benefit of collaboration. However, as high \(βₖ\) can indicate higher dependency or prevalence of low impact of indigenous productivity, desirable value of \(γₖ\) can be in the range \(1 < γₖ < 2\). Similarly boost ratio of visibility indicators are denoted as \(γₖ(t), γₖ(f)\) and \(γₖ(b)\), for twitter, facebook, and blogs, respectively.

**Citedness boost (βᵣₖ)** It is the ratio of total citedness (total cited ratio) to the citedness ratio of the indigenous publications.
Scientometrics (2023) 128:5081–5116

\[ \beta_{rc} = \left[ \frac{r_T}{r_{TI}} - 1 \right] \times 100\%, \]

where

\[ r_T = \frac{\text{total number of cited publications}}{\text{total number of publications}} = \frac{TP_{cited}}{TP}, \]

and

\[ r_{TI} = \frac{\text{total number of cited indigenous publications}}{\text{total number of indigenous publications}} = \frac{TIP_{cited}}{TIP}. \]

Citedness boost value greater than but close to 1 indicate that indigenous publications are also having considerably good citedness. \( \beta_{rc} \) and \( \beta_c \) can be together used to determine whether a country’s indigenous works are making enough impact. \( \beta_{rc} \) value closer to 1 (like < 1%), but considerably high \( \beta_c \) (like > 50%) can indicate that despite the potential of indigenous works to gain citations, a considerable amount of works is remaining under-cited or not getting enough citations. Visibility indicators can be also computed in a similar fashion using coverage instead of citedness and are denoted as \( \beta_{ra}(t), \beta_{ra}(f), \beta_{ra}(b) \) representing altmetric sources like twitter, facebook and blogs respectively.

**Boost ratio of impact per unit boost in citedness (\( \delta_c \))** It is the net boost of impact per unit boost of citedness due to international collaborations.

\[ \delta_c = \frac{\beta_c}{\beta_{rc}}. \]

Greater the value of \( \delta_c \), greater is the effectiveness of collaborations. If this ratio is very high, with not so reasonably high value of \( \beta_{rc} \) (i.e., \( \beta_{rc} < 1\% \)), it can indicate that a majority of collaborations are of good quality and is rewarding well. High value of \( \delta_c \) at a cost of reasonably high \( \beta_{rc} \) (i.e., \( \beta_{rc} > 1\% \)) can indicate that some relatively less-rewarding collaborations are there and these can be reviewed and decisions on whether to strengthen such collaborations or to minimize focus on such collaborations can be taken. There is a possibility that such collaborations can be a new tie and might take some time for reaping the fruits. Another possibility can be such collaborations are started long back but are mainly concentrated in very matured or obsolete themes. In such cases, decision on whether to stop further collaboration or start collaborating in new and hotter themes can be taken up.

Similarly boost ratio of visibility indicators per unit boost of coverage are denoted as \( \delta_a(t), \delta_a(f) \) and \( \delta_a(b) \), respectively for twitter, facebook and blogs.

**Results**

Now the results of the analysis are presented, organized according to the five research questions proposed.

**RQ1** What proportion of research output of India involves international collaboration and whether the propensity to collaborate internationally has increased over time?
For finding an answer to this question, the proportion of indigenous and ICP instances in Indian research output is identified, in a year-wise manner from 2001 to 2020. Table 1 presents the total research output, ICP and indigenous papers from India during the period 2001–2020. It can be observed that in the year 2001, a total of 2354 papers out of total 11,357 papers from India, involved international collaboration. This constitutes 20.73% of the total papers. By the year 2020, internationally collaborated papers have increased to 23,815, which is 32.35% of the total research output of 73,614 papers by India in the year. Thus, not only the volume of internationally collaborated papers has increased, but their proportionate share in Indian research output has also increased. For the whole 20-year period taken together, it was found that 25.57% of total research papers involved international collaboration. For a clearer visualization of the trend of international collaboration of India during the period and the rate of growth of international collaboration, the year-wise count of total research papers and the ICP instances are plotted together, and CAGR values for both are shown in Fig. 1. It can be observed that both, the total research papers from India as well as the internationally collaborated papers, have increased during the period. However, the rate of growth of ICP instances (12.27%) is higher as compared to rate of growth of total research output (9.8%). Thus, these results show that during the 2001–2020 period, India has about 1/4th of its research output involving international collaboration, and that India’s

| Year | Total No. of papers | No. and % of internationally collaborated papers | No. and % of indigenous papers |
|------|---------------------|-----------------------------------------------|-------------------------------|
| 2001 | 11357               | 2354 (20.73)                                  | 9003 (79.27)                  |
| 2002 | 13024               | 2707 (20.78)                                  | 10317 (79.22)                 |
| 2003 | 14237               | 2986 (20.97)                                  | 11251 (79.03)                 |
| 2004 | 15816               | 3409 (21.55)                                  | 12407 (78.45)                 |
| 2005 | 17945               | 3838 (21.39)                                  | 14107 (78.61)                 |
| 2006 | 21252               | 4623 (21.75)                                  | 16629 (78.25)                 |
| 2007 | 24441               | 5188 (21.23)                                  | 19253 (78.77)                 |
| 2008 | 22387               | 4810 (21.49)                                  | 17577 (78.51)                 |
| 2009 | 25624               | 5784 (22.57)                                  | 19840 (77.43)                 |
| 2010 | 34508               | 8089 (23.44)                                  | 26419 (76.56)                 |
| 2011 | 38219               | 9071 (23.73)                                  | 29148 (76.27)                 |
| 2012 | 42328               | 10018 (23.67)                                 | 32310 (76.33)                 |
| 2013 | 47290               | 11181 (23.64)                                 | 36109 (76.36)                 |
| 2014 | 53349               | 12807 (24.01)                                 | 40542 (75.99)                 |
| 2015 | 55542               | 13771 (24.79)                                 | 41771 (75.21)                 |
| 2016 | 61477               | 16412 (26.70)                                 | 45065 (73.30)                 |
| 2017 | 66016               | 17746 (26.88)                                 | 48270 (73.12)                 |
| 2018 | 70767               | 19871 (28.08)                                 | 50896 (71.92)                 |
| 2019 | 74964               | 22018 (29.37)                                 | 52946 (70.63)                 |
| 2020 | 73614               | 23815 (32.35)                                 | 49779 (67.65)                 |
| Total| 784157              | 200498 (25.57)                                | 583659 (74.43)                |
propensity to collaborate internationally has increased during the period, rising from 20.73% research papers in 2001 to 32.35% research papers in 2020. This is a significant rise in proportionate share, especially considering the fact that total research output from India during this period has also increased significantly.

The study also tried to measure what part of internationally collaborated papers from India involved bilateral and multilateral collaboration. Here, bilateral means all such papers where there is a collaboration between researcher(s) from India and one or more researchers from another country. Multilateral collaboration means research papers that involve researchers from three or more countries. Figure 2 shows a plot of the year-wise proportion of bilateral and multilateral collaborated research papers from among the total internationally collaborated research papers. Out of total ICP instances, 32.79% research papers are instances of multilateral collaboration, and 67.21% research papers are instances of bilateral collaboration. However, the proportion of multilateral collaborations is seen rising during the observation period.
Fig. 3 Graphical visualization of major collaborating countries with India (2001–2020)

RQ2 What are the major collaborating research partner countries for India and how has the relative intensity of collaboration of India changed over time?

The second research question involved identifying the major collaborating partner countries for India. For this purpose, the top countries with which India obtained highest collaborated output during the period, are identified. Figure 3 shows a graphical visualization illustrating the major collaborating partner countries of India during the 2001–2020 period, along with the intensity of each collaboration. The darker shade represents higher collaboration. It can be clearly observed that many of the major collaborating partners with India are located distantly from India (the USA being the farthest and yet having the highest collaborated output). Only, China is a neighboring country for India, and South Korea, Malaysia and Taiwan are Southeast Asian neighbors of India. India’s collaboration with its South Asian neighbors is found to be extremely small. Brazil and South Africa are also among the emerging international collaboration partners for India, though their proportionate share is low. The ‘Discussion’ section of the paper presents more discussion on these observations in the light of proximity theories of research collaboration.

In order to understand the levels of international collaboration further, Table 2 presents the number and percentage of ICP instances for the top 20 countries with India, grouped into two blocks of 10 year each, i.e., during 2001–2010 and 2011–2020. Further, the CAGR values of internationally collaborated papers with each country during the period are also computed. It can be observed that, USA, Germany, England, South Korea and China are the top 5 major collaborating partners. USA accounts for the largest share of more than 30% ICP instances of India, although its share declined marginally from 34.14% in 2001–2010 period to 31.22% in 2011–2020 period. Germany is the 2nd largest collaborator, though here too there is a slight decline from 14.60% in 2001–2010 to 10.76% in 2011–2020. The collaborated research output with England, South Korea and China, however, has increased from 2001–2010 to 2011–2020 period. Among some interesting patterns to observe is Saudi Arabia that records a rapid growth in the collaborated research
output with India, from 0.74% in 2001–2010 period to 7.29% in 2011–2020 period, with a CAGR value of 33.36%. The collaboration with South Korea, Australia, and South Africa is also found to be increasing at a higher rate. Germany shows the lowest CAGR value, indicating that the collaboration between India and Germany is not increasing significantly. Thus, the values in the table point to a rise in India’s research collaboration with newer partners, while collaboration with existing partners either remained intact or decreased a bit in terms of proportion (though volume has increased). For example, India’s ICP instances with USA have increased from 14,951 papers in 2001–2010 to 48,927 papers in 2011–2020, an increase of more than three times. However, at the same the time proportionate share of USA in India’s ICP instances has decreased, indicating the emergence of new collaboration partners.

In order to understand the changing patterns in a more systematic manner, the Relative Intensity of Collaboration (RIC) of India is computed with respect to these 20 major collaborating partners. It has been shown that comparing collaborating partners of similar sizes is simple but doing so will be a biased comparison if their sizes are drastically different. The relative intensity of collaboration is one such index that can address this (Fuchs et al., 2021). The RIC index is an asymmetric index that can be used to measure collaboration within a scientific network. The RIC curves for India are plotted in Fig. 4. It can be observed that the RIC with respect to USA, Germany and England has remained at almost the same level (with a marginal decrease), whereas RIC with respect to China has

| Countries          | Internationally collaborated papers | CAGR% (2001–2020) |
|--------------------|-------------------------------------|-------------------|
|                    | 2001–2010                          | 2011–2020         |
| USA                | 14951 (34.14%)                      | 48927 (31.22%)    | 10.59 |
| Germany            | 6395 (14.60%)                       | 16859 (10.76%)    | 9.51  |
| England            | 4520 (10.32%)                       | 18272 (11.66%)    | 13.10 |
| South Korea        | 3186 (7.28%)                        | 15698 (10.02%)    | 19.79 |
| China              | 2433 (5.56%)                        | 15856 (10.12%)    | 18.89 |
| France             | 3725 (8.51%)                        | 12220 (7.80%)     | 10.28 |
| Japan              | 4302 (9.82%)                        | 10807 (6.90%)     | 9.81  |
| Australia          | 2005 (4.58%)                        | 11996 (7.65%)     | 17.12 |
| Italy              | 2102 (4.80%)                        | 10393 (6.63%)     | 13.09 |
| Canada             | 2394 (5.47%)                        | 9386 (5.99%)      | 12.27 |
| Saudi Arabia       | 322 (0.74%)                         | 11425 (7.29%)     | 33.36 |
| Spain              | 1419 (3.24%)                        | 8395 (5.36%)      | 14.72 |
| Taiwan             | 1680 (3.84%)                        | 6731 (4.30%)      | 14.08 |
| Russia             | 1546 (3.53%)                        | 6722 (4.29%)      | 13.65 |
| Switzerland        | 1516 (3.46%)                        | 6718 (4.29%)      | 14.30 |
| Brazil             | 1142 (2.61%)                        | 6742 (4.30%)      | 15.87 |
| Malaysia           | 1179 (2.69%)                        | 6595 (4.21%)      | 18.46 |
| Netherlands        | 1443 (3.30%)                        | 6005 (3.83%)      | 11.72 |
| South Africa       | 526 (1.20%)                         | 6123 (3.91%)      | 19.09 |
| Poland             | 1106 (2.53%)                        | 5522 (3.52%)      | 14.87 |

Some internationally collaborated papers may involve multiple countries and hence the overall % value may be greater than 100.
decreased. The major growth in RIC is seen with respect to South Korea and Saudi Arabia. The RIC patterns in respect to countries like Canada, Japan, Spain, Switzerland, Netherlands, Poland, and Brazil remained largely at the same level during the 20-year period, suggesting relatively stable collaboration with these countries. The RIC with respect to countries like Italy, Australia, Saudi Arabia, Russia, and South Africa shows an upward trend. The trend for Taiwan and Malaysia shows a decline. Thus, the RIC plots confirm that India’s collaboration intensity with some of its old partners has decreased and that it has found some new collaboration partners, most notable being South Korea and Saudi Arabia.

RQ3 Do the international research collaboration patterns vary across subject areas?

The subject area-wise patterns of India’s international collaboration are analyzed to understand whether the different subject areas exhibit different levels of international collaboration. Or in other words, whether research in some subject areas involves higher international collaboration as compared to others. The fourteen broad subject area grouping, as described in the ‘Data & Method’ section, is used for the purpose.

Fig 5 shows the year-wise proportion of indigenous and ICP instances for the fourteen subject areas. Social Science (SS), Mathematics (MAT) and Multidisciplinary (MUL) are found to be the three subject areas with a relatively higher proportion of ICP instances. The percentage of ICP for these areas are 41.85% (for SS), 35.03% (for MAT), and 31.81% (for MUL). The other subject areas with relatively better proportion...
of papers as ICP are PHY (29.23%), GEO (27.54%), and ENV (26.23%). Out of all the subject areas, Agriculture (AGR) has the lowest proportion of ICP, with only 20.26% of the total research papers during 2001–2020 involving international collaboration. It may be noted that the CAGR\textsubscript{ICP} values for all subject areas are greater than CAGR\textsubscript{Non-ICP}, suggesting that the ICP proportion in each subject area has increased. The MUL subject area, with a CAGR\textsubscript{ICP} of 22.38% is the one to record the highest growth in internationally collaborated research output. This is followed by INF with CAGR\textsubscript{ICP} value of 17.41%. PHY and AGR are the two subject areas with lowest CAGR\textsubscript{ICP} value of 8.27%, 9.87%, respectively.

Fig. 5  ICP proportion and trends in different subject areas
Now the major collaborating countries in various subject areas are identified. For this purpose, the collaborated publication records in each subject area are analyzed and the major collaborating partners in each subject area are identified. Table 3 presents the top 10 collaborating countries (arranged in descending order of collaborated papers) for all the fourteen subject areas. While USA is found to have the highest number of collaborated papers across all subject areas, the proportion of collaborated papers varies from 47.93% in MED to 20.42% in MAR subject area. The other major collaborating countries in each subject area are found to be somewhat different in each case. In AGR, other than USA, the other major collaborators are Australia, China, Germany, and South Korea. In AH, the
major collaborators (in addition to USA) are England, China, and Australia. In BIO subject area, the other major collaborators are England, Germany, and South Korea; whereas in CHE, these are South Korea, Germany and Japan. South Korea, China, England and Germany are the other major collaborators in ENG subject area. In MED, the collaborating countries are England, Australia and Canada. In MAT subject area the other major collaborating countries are China, Saudi Arabia and Germany; whereas in PHY subject area, these are Germany, South Korea and France. Thus, the international collaboration in different subject areas shows slightly different preferences for collaborating partners. Countries with more developed research in a subject area are likely to be among the major collaborating partners in that subject area. The existence of bilateral and multilateral cooperation programs in specific thematic areas may also be related to international collaboration with a specific country. The different proximity dimensions may also have a role to play in shaping subject area specific collaboration patterns with a set of specific countries. This is discussed further in the ‘Discussion’ section.

RQ4 What proportion of internationally collaborated papers from India have first and/or the majority author(s) from India?

The next research question that the study attempted to answer is about the nature of international collaboration that Indian researchers are found to engage in. For this purpose, the following two results were computed—(a) how many internationally collaborated research papers have an Indian researcher as first/lead author, and (b) what proportion of the internationally collaborated papers have 50% or more authors from India. The objective behind computing both these results is to find whether Indian researchers are playing the leading/ major role in the internationally collaborated papers they engage in, or its otherwise.

Table 4 presents the year-wise number and percentage of internationally collaborated papers that have an Indian researcher as the first author, and also the proportion of papers that have at least 50% authors from India. It can be observed that the proportion of internationally collaborated papers with an Indian researcher as first author has increased marginally in the period, reaching to 52.15% in 2020 from 47.88% in 2001. From 2013 onwards, all the values are more than 50%. Thus, about half of the internationally collaborated papers from India during the period had an Indian researcher as
## Table 3  Top 10 collaborating countries in different subject areas

| Subject area | Countries and the number and percentage of internationally collaborated papers |
|--------------|--------------------------------------------------------------------------------|
| AGR          | USA (2996) Australia (1119) China (978) Germany (922) South Korea (908) England (731) Japan (689) Saudi Arabia (561) Canada (530) France (472) |
|              | 25.61% 9.56% 8.36% 7.88% 7.76% 6.25% 5.89% 4.79% 4.53% 4.03% |
| AH           | USA (642) England (185) China (154) Australia (128) Germany (98) Canada (95) France (87) Saudi Arabia (56) South Korea (54) Netherlands (52) |
|              | 41.23% 11.88% 9.89% 8.22% 6.29% 6.10% 5.59% 3.60% 3.47% 3.34% |
| BIO          | USA (17,286) England (4294) Germany (3652) South Korea (3414) China (3056) France (2954) Australia (2795) Japan (2713) Italy (2531) Saudis Arabia (2531) |
|              | 36.72% 9.12% 8.60% 7.76% 7.25% 6.49% 6.27% 5.94% 5.76% 5.38% |
| CHE          | USA (10,032) South Korea (5145) Germany (4280) Japan (3260) Saudi Arabia (2933) England (2538) France (2517) China (2000) Taiwan (1999) Spain (1999) |
|              | 21.66% 11.11% 9.24% 7.40% 7.26% 6.33% 5.48% 5.44% 4.32% 4.32% |
| ENG          | USA (6236) South Korea (2698) China (2257) England (1845) Australia (1542) Saudi Arabia (1451) Canada (1397) Germany (1328) Malaysia (1289) Japan (1037) |
|              | 24.31% 10.52% 8.80% 7.19% 6.01% 5.56% 5.45% 5.18% 5.02% 4.04% |
| ENV          | USA (4916) China (1821) Germany (1623) Australia (1588) England (1447) South Korea (1359) Japan (1286) France (1039) Canada (916) Saudis Arabia (901) |
|              | 27.87% 10.32% 9.20% 9.00% 8.20% 7.70% 7.21% 5.89% 5.19% 5.11% |
| GEO          | USA (4868) England (1749) Germany (1709) Australia (1509) China (1487) Japan (1323) France (1290) South Korea (1126) Canada (1064) Italy (688) |
|              | 29.69% 10.67% 10.42% 9.20% 9.07% 8.07% 7.87% 6.87% 6.49% 4.20% |
| INF          | USA (4920) China (2126) England (1242) South Korea (1146) Australia (1057) Canada (1050) Saudi Arabia (838) France (707) Germany (698) Malaysia (625) |
|              | 30.69% 13.26% 7.75% 7.15% 6.59% 6.55% 5.23% 4.41% 4.35% 3.90% |
Some research papers may be tagged in more than one subject area, due to overlaps in WC field. Similarly, an internationally collaborated paper may involve multiple countries.

Table 3 (continued)

| Subject area | Countries and the number and percentage of internationally collaborated papers |
|--------------|--------------------------------------------------------------------------------|
| MAR          | USA   | South Korea | Germany | Japan   | Saudi Arabia | England | China   | France | Malaysia | Taiwan |
|              | (8602) | (5761)      | (4644)  | (3271)  | (2456)       | (2405)  | (2379)  | (2104) | (1877)   | (1652) |
|              | 20.42% | 13.67%      | 11.02%  | 7.76%   | 5.83%        | 5.71%   | 5.65%   | 4.99%  | 4.46%    | 3.92%  |
| MAT          | USA   | China       | Saudi Arabia | Germany | South Korea | Canada | France | England | Turkey | Spain |
|              | (3877) | (1621)      | (1423)  | (1078)  | (953)       | (910)   | (839)   | (801)   | (596)   | (591)  |
|              | 23.68% | 9.90%       | 8.69%   | 6.58%   | 5.82%       | 5.56%   | 5.12%   | 4.89%   | 3.64%    | 3.61%  |
| MED          | USA   | England     | Australia | Canada | Germany | China | Italy | France | Japan | Netherlands |
|              | (21597) | (8055)      | (5161)  | (4450)  | (3996)      | (3951)  | (3279)  | (3215)  | (3125)   | (2776) |
|              | 47.93% | 17.88%      | 11.45%  | 9.88%   | 8.87%       | 8.77%   | 7.28%   | 7.14%   | 6.94%    | 6.16%  |
| MUL          | USA   | England     | Germany | China | Australia | Saudi Arabia | France | Canada | South Korea | Japan |
|              | (2897) | (1040)      | (729)   | (622)   | (573)       | (514)   | (501)   | (467)   | (439)    | (431)  |
|              | 40.60% | 14.57%      | 10.22%  | 8.72%   | 8.03%       | 7.20%   | 7.02%   | 6.54%   | 6.15%    | 6.04%  |
| PHY          | USA   | Germany     | South Korea | France | England | China | Italy | Japan | Russia | Spain |
|              | (23217) | (15557)     | (10849) | (10465) | (9757)     | (9514)  | (8722)  | (8586)  | (8580)   | (6821) |
|              | 31.75% | 21.27%      | 14.83%  | 14.31%  | 13.34%      | 13.01%  | 11.93%  | 11.74%  | 11.73%   | 9.33%  |
| SS           | USA   | England     | Australia | Canada | China | Germany | Netherlands | France | Switzerland | South Africa |
|              | (6928) | (3375)      | (1777)  | (1430)  | (1398)      | (1031)  | (990)   | (876)   | (841)    | (724)  |
|              | 46.93% | 19.98%      | 10.52%  | 8.47%   | 8.28%       | 6.10%   | 5.86%   | 5.19%   | 4.98%    | 4.29%  |
the first author. This indicates that more than half of the internationally collaborated research papers of India have a major role in research design and conceptualization played by the Indian researcher. Similarly, when we measure the quantum of Indian internationally collaborated papers with at least 50% of Indian authors, we observe that more than 50% of internationally collaborated papers have at least 50% of author contributions from India. Although, it is also found that this proportion has decreased from 66.86% in 2001 to 51.12% in 2020. There could be multiple reasons for the decrease, such as shift towards more team-oriented science, mobility of Indian scientists etc. However, at the same time, it is important to note that while the proportion of research papers with at least 50% authors from India may have decreased, the proportion of papers with an Indian researcher as the lead author has increased marginally during this period.

The patterns of leading and at least 50% Indian authors have also been computed for research publications in different subject areas. Table 5 presents the subject area-wise proportion of research papers with an Indian researcher as the first author and the proportion of research papers that have at least 50% authors from India. It is observed that, first author proportion lies between 36.15% (for SS subject area) and 60.24% (for CHE subject area). The subject areas with a good proportion of internationally collaborated research papers having an Indian lead author are: CHE (60.24%), MAR (56.84%), ENG (55.41%), INF (54.18%), BIO (51.03%), and PHY (50.00%). The subject areas with a lesser proportion of internationally collaborated papers having an Indian lead author are SS (36.15%), MED (40.47%), MUL (45.18%), and AGR (47.26%). In terms

| Year | No of internationally collaborated papers | Collaborated papers with Indian first author (%) | Collaborated papers with at least 50% Indian authors (%) |
|------|-----------------------------------------|-----------------------------------------------|--------------------------------------------------------|
| 2001 | 2354                                    | 1127 (47.88)                                  | 1574 (66.86)                                           |
| 2002 | 2707                                    | 1322 (48.84)                                  | 1780 (65.76)                                           |
| 2003 | 2986                                    | 1511 (50.60)                                  | 2018 (67.58)                                           |
| 2004 | 3409                                    | 1642 (48.17)                                  | 2184 (64.07)                                           |
| 2005 | 3838                                    | 1894 (49.35)                                  | 2426 (63.21)                                           |
| 2006 | 4623                                    | 2242 (48.50)                                  | 2975 (64.35)                                           |
| 2007 | 5188                                    | 2545 (49.06)                                  | 3351 (64.59)                                           |
| 2008 | 4810                                    | 2329 (48.42)                                  | 2961 (61.56)                                           |
| 2009 | 5784                                    | 2799 (48.39)                                  | 3554 (61.45)                                           |
| 2010 | 8089                                    | 3758 (46.46)                                  | 4913 (60.74)                                           |
| 2011 | 9071                                    | 4270 (47.07)                                  | 5483 (60.45)                                           |
| 2012 | 10018                                   | 4716 (47.08)                                  | 5785 (57.75)                                           |
| 2013 | 11181                                   | 5658 (50.60)                                  | 6274 (56.11)                                           |
| 2014 | 12807                                   | 6467 (50.50)                                  | 7018 (54.80)                                           |
| 2015 | 13771                                   | 6972 (50.63)                                  | 7386 (53.63)                                           |
| 2016 | 16412                                   | 8298 (50.56)                                  | 8832 (53.81)                                           |
| 2017 | 17746                                   | 9254 (52.15)                                  | 9508 (53.58)                                           |
| 2018 | 19871                                   | 10165 (51.15)                                 | 10478 (52.73)                                          |
| 2019 | 22018                                   | 11327 (51.44)                                 | 11601 (52.69)                                          |
| 2020 | 23815                                   | 12420 (52.15)                                 | 12175 (51.12)                                          |
of internationally collaborated papers with at least 50% Indian authors, it is observed that subject areas MAT (64%), CHE (63.96%), ENG (63.66%) and MAR (63.46%) have a higher proportion of internationally collaborated papers with at least 50% Indian authors. Many of these subject areas are also the one having higher proportion of research papers with Indian lead author. The subject areas of MED (44.65%), SS (47.54%). and MUL (47.69%) have less than 50% of research papers having at least 50% Indian authors. Thus, the patterns of lead as well as majority Indian authors in the internationally collaborated research papers are found to vary across different subject areas.

**RQ5** Do internationally collaborated papers of India get higher impact and visibility as compared to indigenous papers?

Many previous studies have highlighted advantages of international collaboration in terms of higher impact and visibility of research papers. Therefore, it has been attempted to find out whether this holds true in the case of Indian research papers involving international collaboration. For this purpose, the citation impact as well as social media visibility of internationally collaborated and indigenous research papers from India are computed.

### Table 5 Subject area-wise authorship patterns of internationally collaborated papers

| Subject area | No of papers in subject area | Collaborated papers with Indian first author (%) | Collaborated papers with 50% Indian authors (%) |
|--------------|-------------------------------|-----------------------------------------------|-----------------------------------------------|
| AGR          | 11700                         | 5530 (47.26)                                  | 6563 (56.09)                                  |
| AH           | 1557                          | 734 (47.14)                                   | 863 (55.43)                                   |
| BIO          | 47080                         | 24026 (51.03)                                 | 25738 (54.67)                                 |
| CHE          | 46306                         | 27896 (60.24)                                 | 29616 (63.96)                                 |
| ENG          | 25656                         | 14217 (55.41)                                 | 16333 (63.66)                                 |
| ENV          | 17641                         | 8554 (48.49)                                  | 9753 (55.29)                                  |
| GEO          | 16394                         | 8008 (48.85)                                  | 9047 (55.18)                                  |
| INF          | 16031                         | 8685 (54.18)                                  | 9439 (58.88)                                  |
| MAR          | 42131                         | 23948 (56.84)                                 | 26736 (63.46)                                 |
| MAT          | 16371                         | 7978 (48.73)                                  | 10478 (64.00)                                 |
| MED          | 45058                         | 18235 (40.47)                                 | 20117 (44.65)                                 |
| MUL          | 7136                          | 3224 (45.18)                                  | 3403 (47.69)                                  |
| PHY          | 73135                         | 36570 (50.00)                                 | 40182 (54.94)                                 |
| SS           | 16893                         | 6106 (36.15)                                  | 8031 (47.54)                                  |

aPapers may be tagged in more than one subject area, due to overlaps in WC field

### Table 6 Citation impact of ICP and indigenous papers

| Citation impact | ICP | Non-ICP |
|-----------------|-----|---------|
| No of collaborated papers | 200498 | 583659 |
| Cited %         | 87.68 | 85.55   |
| Citations per paper | 21.46 | 14.28   |
Table 6 presents the cited % and citations per paper for internationally collaborated and indigenous research papers. It can be observed that 87.68% of the total internationally collaborated papers get at least some citation as compared to 85.55% of the domestic papers. Thus, there is a slight edge in terms of citation potential seen for the internationally collaborated papers. Similarly, in terms of citations per paper, the value for internationally collaborated papers is 21.46 citations per paper as compared to 14.28 citations per paper for domestic papers, suggesting an edge in citations per paper for the internationally collaborated papers. The Z-test done in the case shows that the difference in average citations per paper between international collaborated publications and indigenous publications is significant (p < 0.001). Thus, apart from productivity, a significant boost in impact (citations) can be speculated. According to the indicator design related to various boosts discussed in ‘Data and method’ section, these boosts are computed next.

\[
\beta_p = \left[ \frac{TP}{TIP} - 1 \right] \times 100\% = \left[ \frac{784157}{583659} - 1 \right] \times 100\% = 34.4\%.
\]

**Productivity boost of India**

As India’s productivity boost is less than 50%, there is no over reliance of India on collaborations for productivity and this indicates sufficient strength of India’s indigenous scholarly ecosystem. However, the value is moderately high, indicating that some reduction, either through keeping only highly rewarding collaborative ties active or strengthening indigenous research through internal ties at the individual or organizational level or both is preferable.

\[
\beta_c = \left[ \frac{TC}{TIC} - 1 \right] \times 100\% = \left[ \frac{12640171}{8337273} - 1 \right] \times 100\% = 51.6\%.
\]

**Citations boost for India**

This value indicates that India is slightly more dependent on collaborations for citations than indigenous publications. Thus, though the scholarly ecosystem of India is sufficiently productive, it slightly fails somewhere to gather the necessary impact.

**Boost ratio of impact** On the other hand, how much rewarding collaborations are upon unit % increase in productivity can be computed using the boost ratio of impact.

\[
\gamma_c = \frac{\beta_c}{\beta_p} = \frac{51.6}{34.4} = 1.5.
\]

Thus, for India, for each 1% boost in productivity achieved through collaboration, 1.5% boost in citations is achieved. In case, if India wants to keep collaborative ties that bring more reward, the ones that increase \( \gamma_c \) without much improvement in \( \beta_p \) can be considered.

Further analysis may bring more clarity into this. Specially, how much boost in citedness is achieved with international collaborations needs to be determined.

Citedness boost of India can be computed in the following way:

\[
r_T = \frac{total\ number\ of\ cited\ publications}{total\ number\ of\ publications} = \frac{TP_{cited}}{TP} = \frac{675131}{784157} = 0.861,
\]

\[
r_{TI} = \frac{total\ number\ of\ cited\ indigenous\ publications}{total\ number\ of\ indigenous\ publications} = \frac{TIP_{cited}}{TIP} = \frac{499325}{583659} = 0.8555,
\]
Citedness boost achieved is found to be less than 1%, which may indicate that indigenous publications are not failing to attract citations but are not attracting enough citations. This points towards the possibility in scholarly ecosystem in India where productivity (volume of publications) is predominantly emphasized but there is not enough focus on the outlets (that might bring maximum visibility and attract citations) through which these publications are disseminated. More light on this can be shed with the indicator boost ratio of impact per unit boost in citedness.

Thus, in the case of India, per unit percentage boost in citedness achieved by collaboration, citations can improve by almost 81%. Thus, the major problem lies with the Indian indigenous scholarly ecosystem is not the insufficiency of citedness, but the failure to attract citations. Thus, Indian scholars should target not to merely improve the number of publications by publishing their research in ‘easy to publish’ kind of journals that may ensure citedness but cannot ensure enough reach and thereby impact, but to ensure the quality of their research work and target high end journals in respective fields that can ensure citedness as well as citations. Thus, $\beta_{c}$, $\beta_{rc}$ and $\delta_{c}$ together indicate a need to strive for ensuring maximum quality of indigenous research and publishing it through sources capable of providing maximum reach, thereby improving the chances of increasing the impact.

Various boost indicators related to visibility in various platforms that can be computed in similar fashion as that of the above-mentioned indicators are shown in Table 8. Before that let us examine Table 7.

Table 7 presents the social media visibility of internationally collaborated and indigenous papers, on different social media platforms. It is observed that ICP instances have in general higher visibility, as seen in terms of coverage by the Altmetric.com aggregator, with 36.93% coverage for ICP and 19.45% coverage for non-ICP papers. On Twitter platform, ICP instances get 23.28% coverage as against 12.23% for non-ICP instances. The average tweets per paper for ICP and non-ICP instances are 11.47 and 3.99, respectively. The Facebook and Blog platforms also show similar differentiation, with ICP instances

$$\beta_{rc} = \left( \frac{r_T}{r_{TI}} - 1 \right) \times 100\% = \left[ \frac{0.861}{0.855} - 1 \right] \times 100\% = 0.64\%.$$  

$$\delta_{c} = \frac{\beta_{c}}{\beta_{rc}} = \frac{51.6}{0.64} = 80.6.$$
getting higher coverage and visibility as compared to non-ICP instances. Thus, the internationally collaborated papers from India are found to attract higher social media visibility as compared to domestic papers, and upon Z-test these differences are found to be significant (p < 0.001).

From Table 8, it is clear that India is highly dependent on international collaborative research works (due to very high boosts in mentions and coverage) for altmetric attention in the form of mentions via twitter, facebook and blog platforms. High value of altmetric boost per productivity for all three platforms indicate that 1% increase in productivity achieved via international collaboration can yield 5.5%, 3.1%, and 6% boost in altmetric mentions for platforms twitter, facebook, and blogs, respectively. Also, it seems that coverage plays vital role in garnering altmetric attention than citedness play for citations. Currently, per 1% boost in coverage due to international collaboration, 8%, 3%, and 2.7% boost for mentions is ensured for twitter, facebook and blogs, respectively. Thus, unlike citedness, there is a serious issue with the coverage of indigenous publications on these platforms. Other than selecting top journals for disseminating their research (that ensure more altmetric coverage as these journals will be covered by altmetric.com), Indian researchers need to be active in social media to publicize and promote their new publications, especially for their indigenous works, to ensure visibility of such works. These strategies might improve citations and citedness as well other than altmetric attention. International and Indian journals managed by Indian publishers and editorial board should make sure that the journals are covered by altmetric.com and other altmetric players such as Plumx. Also, they should make it into their editorial policy to accept best quality research.

These indicators can be used for other countries also if data of publications, citation and social media visibility from other countries are downloaded and the procedures are followed as such.

For instance, from the data used in (Dua et al., 2022) for SAARC countries, we demonstrate the productivity and impact (citations) related boost indicators ($\beta_P$, $\beta_c$ and $\gamma_c$) for other seven countries in SAARC (i.e., Nepal, Bhutan, Afghanistan, Pakistan, Bangladesh, Sri Lanka, and Maldives).

From Table 9, other seven SAARC countries are found to be highly dependent on international collaborations for their productivity as well as impact, and for most of the

| Particulars                                      | Twitter | Facebook | Blogs |
|-------------------------------------------------|---------|----------|-------|
| Number of ICP covered in social media platforms | 46676   | 9925     | 6115  |
| Number of indigenous publications covered      | 71381   | 12490    | 4436  |
| Total number of publications covered            | 118057  | 22415    | 10551 |
| Number of mentions for ICP covered             | 535374  | 22133    | 11680 |
| Number of mentions for indigenous publications covered | 284810 | 20983    | 5634  |
| Total number of mentions for all publications covered | 820184 | 43116    | 17314 |
| $\beta_P$                                       | 188%    | 105.5%   | 207.3%|
| $\gamma_P$                                     | 5.5     | 3.1      | 6     |
| $\beta_c$                                      | 23.5%   | 35.5%    | 77.6% |
| $\delta_c$                                     | 8       | 3        | 2.7   |
countries, collaborations are found to be extremely rewarding. However, this also indicates an inherent weakness in their internal scholarly ecosystem, and most of these countries have to go a long way to strengthen their internal scholarly ecosystem. India, on the other hand, does not need that much effort to cut over reliance on international collaborations. These minimal results for other countries are given only as a demonstration that the indicators devised in this work can be used for assessing other countries also and we are not pursuing the whole analysis done in this work for all these countries, though such an endeavor is really useful.

Now we discuss the subject area-wise citation impact of international collaborations of India. Table 10 shows the subject area-wise overall impact of ICP and non-ICP instances of India for the period 2001–2020. Overall impact (consolidated for the last 20 years) is expressed as cited % and citations per paper (CPP). It can be seen that CHE (90.76%), AGR (89.90%) and BIO (89.54%) are three subject areas with relatively higher cited% in ICP, and subject areas such as CHE (90.73%), BIO (89.72%), and MAR (88.43%) have

| Table 9 | Citations related boost indicators for collaboration of other 7 SAARC countries |
|---------|---------------------------------------------------------------|
| Particulars | Afghanistan | Bangladesh | Bhutan | Maldives | Nepal | Pakistan | Sri Lanka |
| TIP | 78 | 7247 | 71 | 9 | 2114 | 56933 | 3721 |
| TP | 853 | 27760 | 701 | 227 | 9347 | 122689 | 11744 |
| TC | 26845 | 469137 | 12671 | 3327 | 175579 | 1551682 | 272157 |
| TIC | 445 | 67768 | 306 | 101 | 13541 | 509065 | 36322 |
| $\beta_P$ | 994% | 283% | 887% | 2422% | 342.1% | 115% | 216% |
| $\beta_c$ | 5933% | 692% | 4141% | 3294% | 1197% | 205% | 649% |
| $\gamma_c$ | 5.97 | 2.45 | 4.67 | 1.36 | 3.5 | 1.78 | 3 |

| Table 10 | Subject area wise citation impact |
|-----------|----------------------------------|
| Subjects | ICP | Non-ICP |
| | Cited % | Citations per paper | Cited % | Citations per paper | CPP ratio (CPP of Non-ICP/CPP of ICP) |
| AGR | 89.90 | 23.14 | 85.20 | 16.80 | 0.73 |
| AH | 75.47 | 9.66 | 61.99 | 4.83 | 0.5 |
| BIO | 89.54 | 22.69 | 89.72 | 17.57 | 0.77 |
| CHE | 90.76 | 20.86 | 90.73 | 18.08 | 0.87 |
| ENG | 86.29 | 17.91 | 84.30 | 14.54 | 0.81 |
| ENV | 88.91 | 22.57 | 87.09 | 16.94 | 0.75 |
| GEO | 89.77 | 24.88 | 87.34 | 17.20 | 0.69 |
| INF | 81.29 | 15.48 | 75.07 | 10.46 | 0.68 |
| MAR | 88.88 | 18.36 | 88.43 | 16.04 | 0.87 |
| MAT | 80.43 | 11.43 | 78.03 | 9.69 | 0.85 |
| MED | 88.37 | 25.78 | 86.29 | 13.99 | 0.54 |
| MUL | 83.16 | 25.89 | 73.24 | 8.75 | 0.34 |
| PHY | 88.98 | 21.67 | 86.05 | 12.20 | 0.56 |
| SS | 85.53 | 16.77 | 80.39 | 10.78 | 0.64 |
higher cited% in non-ICP instances. Thus, few subject areas have higher cited % for all types of papers. Out of all 14 subject areas, AH has the lowest cited% in ICP and non-ICP instances with values of 75.47% and 61.99%, respectively. MUL (25.89), MED (25.78), and GEO (24.88) are the top subject areas having high CPP in ICP, whereas CHE (18.08), BIO (17.57), and GEO (17.20) are the subject areas having higher CPP in non-ICP. Subject areas in which international collaborations made a significant impact in comparison to domestic collaborations can be determined by the CPP ratio, which is also shown in the table. While the international collaborations are found to have significantly more impact than domestic collaborations in the subject area MUL, impact of international collaborations is found to be exactly double as that of domestic collaborations in the case of AH. For areas MED and PHY, it is also close to double.

Discussion

The paper analyzed the research publication data of India during 2001 to 2020 to measure and characterize the international collaboration patterns. It is observed that the proportion of India’s internationally collaborated papers has grown significantly during the last 20 years, rising from 20.73% in 2001 to 32.35% in 2020. The CAGR value of internationally collaborated papers is 12.27%, which is higher than the growth rate of total research publications from India (9.80%). Given that international collaboration in research is known to have several advantages, it is advantageous for India to be able to engage in more internationally collaborated research. When compared to international collaboration levels with other developed countries, we find that India stands between the high ICP level of England and lower ICP level of China, among the major research output producing countries. While, England has a high proportion of 57.7% of its research papers as ICP instances, China has 20.3% of its research papers as ICP instances, as in 2016 (Singh et al., 2021). Among other major knowledge producing countries, USA has 36.8% of its research papers internationally collaborated, whereas Germany has 50.5% of its research papers as internationally collaborated. Thus, in terms of proportion of international collaboration in research, India is engaging in more internationally collaborated research as compared to China, is closer to USA, but much less than Germany and England. The growth in research output from India may also have a connection with rise in international collaboration, as has been observed earlier (Glanzel, 2001; Abramo, D’Angelo, & Solazzi, 2011; Abramo, D’Angelo & Murgia, 2017). Countries which have developed international collaboration in research have gained in terms of productivity as well as impact, and India may be no different.

International cooperation programs and their role The major international collaborating partners for India, during the 2001–2020 period, are USA, Germany, England, and China. Other important emerging collaborators are South Korea, Saudi Arabia, Australia, and South Africa. It may be noted that India is participant in many bilateral and multilateral international cooperation programs involving several countries, which could also have played some role in development of research collaboration with some of these countries.

2 https://dst.gov.in/international-st-cooperation.
For example, India has extensive collaboration programs in different areas with USA, UK, Germany, France etc. There are in fact dedicated organizations/centers working to promote cooperation and collaboration programs with these countries. These centres include Indo-US Science and Technology Forum,3 UK-India Education and Research Initiative,4 Indo-German Science and Technology Centre,5 Indo-French Centre for the Promotion of Advanced Research6 etc. Recently, under its look east policy, more cooperation programs have been instituted with Southeast Asian countries like South Korea, Philippines, Malaysia etc. However, unfortunately, the association of India’s neighboring South Asian countries in the SAARC intergovernmental organization, involving many kinds of cooperation has not been very fruitful in promoting research collaboration in these countries, as India’s international collaboration with its South Asian neighbors is found to be extremely small. The fact that most of the SAARC countries other than India are heavily dependent on foreign collaborations for productivity, impact, and visibility indicates a weakness in their indigenous scholarly ecosystem. Strengthening of the scholarly ecosystem can be achieved if there are long-term ties between these neighboring countries focused with specialized programs involving training and knowledge transfer for improving the indigenous ecosystem of the member countries. In that sense, bilateral and international cooperation programs have a definite role to play in promoting international research collaboration between countries. This is also evident from the fact that out of total internationally collaborated research papers from India, about 67.21% research papers are examples of bilateral collaboration. However, at the same time the proportion of multilateral collaboration is seen rising during this period. While India’s productivity and impact boost points to the benefits of these collaborations, a slight over dependence on collaborations for impact issues warning signals to the Indian researchers and policymakers. At the same time, India’s boost ratio of citation impact indicates the possibility of existence of well rewarding international collaborative ties. However, more benefits can be reaped by identifying more rewarding collaborating ties and strengthening those ties while maintaining other ties as such. However, present boost indicators are not capable of helping in determining rewarding and not-so-rewarding ties as more data about bilateral collaborations is required for that. Such a pursuit may be attempted as a future endeavor. The empirical observations in the paper can thus be used to assess the effectiveness of different cooperation programs of India to promote research collaboration. The lessons learned from the data and the past experiences can be used to modify the existing programs and to institute new international cooperation programs in science and technology.

**Role of proximity dimensions** One may also like to see the international collaboration patterns of India in the light of various proximity dimensions. Many previous studies have explored role of different proximities/distance dimensions as facilitators/inhibitors of international collaboration (Balland et al., 2015; Boschma, 2005; Vieira et al., 2022). The conceptualization of proximities and their role in scientific collaboration has been discussed in several previous studies (Caniëls et al., 2014; Fernández et al., 2016; Heringa et al., 2016; Kraut et al., 2002). The following distance/proximity dimensions are mainly analyzed in the previous studies: geographical (Bergé, 2017; Hoekman et al., 2010; Huang

---

3 https://iusstf.org/
4 http://ukieri.org/
5 https://www.igstc.org/
6 http://www.cefipra.org/
et al., 2013; Katz, 1994; Kuhn, 1996; Laursen et al., 2011; Nagpaul, 2003; Ponds et al., 2007; Wagner & Jonkers, 2017), socioeconomic (Fernández et al., 2016; Jiang et al., 2018; Sokolov-Mladenovic et al., 2016), political (Boschma, 2005; Cattaneo & Corbellini, 2011; Dawes et al., 2012; Whetsell et al., 2020), Cultural (Gui et al., 2019; Hwang, 2013; Knoben & Oerlemans, 2006; Narteh, 2008; Plotnikova & Rake, 2014; Suhay & Druckman, 2015; Taylor & Osland, 2015), intellectual (Cohen & Levinthal, 1990; Cummings & Teng, 2003; Hoekman et al., 2010; Powell et al., 1996), and excellence (Abramo et al., 2011, 2017; Jeong et al., 2014; Jones et al., 2008). In general, it has been observed that different kinds of proximities may facilitate international collaboration, however, the degree of impact of different proximity dimensions vary in different situations. Recently, Hou et al. (2021) reframed the major factors and explored the impact of scientific, economic, geopolitical and cultural factors on international research collaboration. They found that countries with large and equivalent scientific sizes, as well as economic sizes, are more likely to collaborate closely with each other. Co-memberships in intergovernmental organizations and cultural links (such as shared language and religion) also facilitate close collaboration between countries.

In case of Indian international collaboration, out of geographical, socio-economic, political, cultural, intellectual, and excellence proximities; the intellectual and excellence dimensions appear to be more effective. Most of the major collaborating countries for India are located quite far geographically and have different cultural settings. In case of England, political and cultural dimensions may be effective since India and England share similar political systems and common language used by academicians and researchers. Intellectual and Excellence dimensions may explain India’s higher international collaboration with USA, Germany, South Korea, Saudi Arabia, Australia etc. India’s international collaboration with its South Asian neighbors is extremely small to be noticed, therefore the geographical, socio-economic and cultural proximity dimensions do not appear to be playing a major role in this respect. These observations are, however, only indicative in nature and a more systematic procedure and an appropriate experimental design will be required to confirm the impact of various proximity dimensions on India’s international collaboration patterns.

**Knowledge flow in Indian international research collaboration** The authorship structure of internationally collaborated papers from India indicate that Indian authors play major role in such collaborated papers. The number of papers with Indian lead (first) author has increased during 2001–2020. Indian researchers are the lead authors in more than 50% of the internationally collaborated research papers. Similarly, more than 50% internationally collaborated research papers have at least 50% authors from India. Both these patterns together demonstrate that Indian researchers have played major role in the internationally collaborated research papers. This in turn may be an indication of the knowledge flows and important role played by Indian researchers in such collaborated research. The capability of Indian researchers in fostering international collaborations seem to have improved over the time. The more noticeable observation here is that this improvement is not at the cost of importance of their contribution, as Indian authors are found to be in leading position in the rising international collaboration. The role and contribution of Indian researchers in such collaborated output, however, is also found to varying in different subject areas, with some subject areas having more than 50% research papers with Indian lead author. The subject areas of CHE (60.24%), MAR (56.84%), ENG (55.41%), and INF (54.18%) have good proportion of the internationally collaborated papers with Indian lead author. This indicates that the international collaboration in the areas of Chemistry, Materials,
Engineering and Information Technology have a more impactful role of Indian researchers in the collaboration. The subject areas like SS (36.15%), MED (40.47%), MUL (45.18%) and AGR (47.26%) have lesser proportion of such papers, indicating inward knowledge flow in these areas. Interestingly, SS, MED and MUL are also the subject areas with less proportion of internationally collaborated papers having at least 50% authors from India. It appears that the subject areas where India may have a strong research publication base may have a more impactful participation of Indian researchers in international collaboration. This information may be used to institute subject/ theme specific collaboration programs on a complementing basis.

**Citation and altmetric impact of internationally collaborated research** The international collaboration advantage in research impact and visibility is also seen in Indian research output as measured in terms of citations and social media visibility. Several previous studies have demonstrated such advantages (Bordons et al., 1996; Bote et al, 2013; Glänzel & Schubert, 2001; Nguyen et al., 2017; Persson, 2010). In case of India, a marginal advantage in cited% is seen for internationally collaborated papers as compared to indigenous papers. The citations per paper value for internationally collaborated papers is also found to be higher than indigenous papers. There are subject area-wise variations observed too. The advantage is more profound in case of social media visibility. Both, the social media coverage and average mentions per paper is significantly higher for internationally collaborated papers as compared to indigenous papers. Several reasons could be responsible for this. An internationally collaborated paper gets more attention across research groups in different countries and hence is likely to be more impactful in terms of citations and social media visibility. Further, internationally collaborated papers may be addressing research problems of wider interest across national boundaries and hence they may be able to attract more attention.

**Some policy implications** As international collaborations have turned out well for India in terms of both social media visibility and citation impact in comparison to domestic collaborations (on exploration based on RQ5), we highly recommend Indian institutions to strengthen the efforts to engage in more international collaborations and maintain existing collaborations. Based on the results of investigation carried out for RQ3, such effort can be directed by giving more emphasis to subject areas such as Agriculture Science (AGR) and Engineering (ENG) where international collaboration is relatively weak. Though international collaborations are highly rewarding (as indicated by various indices devised in this work), considering the cost of establishing and maintaining such collaborations, barring some top-level institutions, most of the Indian institutions might be finding it hard to break the ice for that matter. Also, though the Indian indigenous scholarly ecosystem is found to be productive, the prevalence of inability to attract enough visibility and citations is revealed by the indicators we devised. So, it is high time Indian institutions made deliberate efforts to make domestic collaborations as successful in terms of intensity, citation impact and visibility on par with the international collaborations. As determination of right partner for collaboration is very much vital for successful collaborations, one of the major reasons for finding domestic collaborations to be less rewarding than international collaborations can be the choice of partners of collaboration. At the level of institutions, to engineer collaborations with compatible institutions that can complement their expertise, institutions should have an effective mechanism to identify their research strengths, i.e., the areas in which they are really strong (core expertise or core competency areas).
and the areas that have the potential to develop as a core competency area later (potential core competency areas). For this, they can adopt any sound assessment methods that help to reveal the research portfolio. Internal audit or assessment can be useful to an extent for this. But in order to identify suitable partners for an institution, it is important to know the strengths of other institutions too and such information is difficult to obtain. For policymakers at national level too, this exercise can be vital. This approach can be used for performance-based funding/resource allocation decisions in ‘thrust-areas’ or areas of national importance.

Apart from this, the same approach can be useful for capacity building for systematic institutionalization for nurturing institutional clusters of excellences in these thrust-areas to make the country globally competitive in these areas. Lathabai et al. (2021) developed a framework to determine research portfolio (that includes core competency areas and potential core competency areas) of academic institutions of a country and Lathabai et al. (2022) developed a recommendation system framework that can toss collaboration recommendations based on the core competency and potential core competency areas for institutions with respect to a field. Application of these frameworks can be extended beyond the limit of fields to explore suitable partner institutions for collaboration in multiple thematic areas. For that matter, one contribution of this research can be significant. As mentioned earlier, agriculture science (AGR) and engineering (ENG) are found to be the subjects with more domestic collaborations and from CPP ratio of these subject areas (0.73 and 0.81, respectively) domestic collaborations are found to be fairly rewarding in comparison to international. Thus, as there are signs of existence of impactful domestic collaborations, further explorations for establishment of more impactful domestic collaborations should be conducted. Considering this and the strategic importance of these fields for the nation, along with attempting to improve the level of international collaborations, systematic approach for bringing out impactful domestic collaborations with respect to important fields within these subjects by application of the above discussed recommendation system framework will be really valuable at institutional level as well as to the country as a whole. Apart from this, as discussed in the previous section, Indian researchers should take proactive effort in ensuring their research to be disseminated through high end journals in respective fields and devote time and energy in promoting the same in different media for outreach, especially blogs where Indian indigenous research is heavily ‘under-mentioned’. Journal administrators and senior researchers at editorial board positions should ensure the coverage of journal in various altmetric platforms. National policymakers can filter out less rewarding collaborations and review and make decision on how to continue with these collaborative ties.

As areas MUL, AH, MED and PHY are the areas where international collaborations are found to be more rewarding, efforts for maintaining such collaborations can be encouraged. However, does this point towards the weakness of domestic collaborations in these fields? Though an ideal value of the CPP ratio to indicate the alarming lack of self-sufficiency in a subject-area is neither given nor any rule of thumb is defined to do so in the available literature, values of CPP close to 0.5 and less can be taken as a warning signal. Special care should be taken in these subject areas to make domestic collaborations more impactful. Thus, this study can serve as a basis for prioritization of subjects (from which important fields can be picked up). However, accurate computations of the same can be done by adopting the indicators devised in this work to the field level. This is reserved as a future endeavor. Therefore, another recommendation to national level policymakers is that while considering key fields for fostering collaborations and to build institutional clusters of excellence, give special emphasis to the fields within MUL, AH, MED, PHY, AGR and
ENG. Since the major collaborating countries for different subject areas are found to be a little different, subject-specific strategies for international collaboration can be worked out. For example, England is the major collaborator in subject areas- SS, MUL, MED, BIO and GEO. China is a major collaborator in ENV and INF subject areas. Similarly, Germany is major collaborator in PHY subject area, whereas South Korea is in CHE and Australia is in AGR subject area. Thus, the subject specific differentiations in international collaborations can be observed and utilized for instituting appropriate collaboration programs.

Conclusion

The article measures and characterizes the international collaboration patterns in Indian scientific research during the last two decades (2001–2020). It is found that international collaboration in Indian scientific research has risen from 20.73% 2001 to 32.35% in 2020. Among internationally collaborated papers, a large proportion (67.21%) constitutes bilateral collaboration. USA, Germany, England remain major collaborators (in terms of volume) whereas South Korea and Saudi Arabia are emerging as major collaborators (in terms of relative intensity). Variations in international collaboration patterns are seen across different subject areas with Multidisciplinary, Arts & Humanities, Medical Science and Physics gaining most through international collaboration. Out of the internationally collaborated papers from India, more than 50% have an Indian researcher as lead author, and more than half of the papers have at least 50% authors from India. Upon usage of novel indicators, though India is found to have a productive internal scholarly ecosystem, its slight over reliance on external collaborations for impact, and heavy dependence on external collaborations for visibility in altmetric platforms is revealed. However, considering the fact that Indians had a commendable role in making external collaborative research successful, some proactive measures suggested in this work along with usage of diligent frameworks for determining the strength of institutions and collaboration recommendation system at organizational level and national level respectively can make the Indian indigenous scholarly ecosystem more self-sufficient in terms of visibility and impact.

The article not only answers the research questions proposed but also presents a comprehensive and updated analysis and characterization of international collaboration patterns in Indian scientific research, which can be useful for various purposes. The empirical observations in the paper can be used to assess the effectiveness of different cooperation programs of India to promote research collaboration. The lessons learned from the data and the past experiences can be used to modify the existing programs and to institute new international cooperation programs in science and technology. Some recommendations are made to policymakers at institutional and national level along with directions for further research that can be useful for the improvement of the overall academic performance of the country. This research not only presents the Indian case, but also presents a diligently designed sequence of methods, which is a combination of existing as well as novel methods that can be used to measure and characterize the collaboration patterns and other aspects of a country. Novel indicators designed to determine the effectiveness of external collaborations and level of dependency of a country on external collaborations for productivity and impact are demonstrated in the case of seven SAARC countries also.

As a future work, the exploration of devised indicators to determine the more rewarding fields of international collaboration and vital collaborating countries that contributed
most to these fields can be pursued, because such an endeavor may help to shape up future optimal collaboration pattern for a country and may aid policymakers of a country to design and engineer collaboration facilitation programs for a country with the countries with which collaborative ties are found to be most rewarding. An analysis of the devised indicator of boost by using data from other countries can also help in exploring the wider applicability of the proposed indicator, including a comparative analysis with other collaboration indices used in the literature.

Acknowledgements A preprint of this paper is available at arXiv. See: Dua et al. (2022).

Funding The authors declare that no funding was received for this work.

Declarations Conflict of interest The manuscript complies with ethical standards of the journal and there is no conflict of interests whatsoever.

References

Abramo, G., D’Angelo, A. C., & Murgia, G. (2017). The relationship among research productivity, research collaboration, and their determinants. *Journal of Informetrics, 11*(4), 1016–1030.

Abramo, G., D’Angelo, A. C., & Solazzi, M. (2011). The relationship between scientists’ research performance and the degree of internationalization of their research. *Scientometrics, 86*(3), 629–643.

Abramo, G., D’Angelo, C. A., & Di Costa, F. (2009). Research collaboration and productivity: Is there correlation? *Higher Education, 57*(2), 155–171.

Adams, J. (2012). The rise of research networks. *Nature, 490*(7420), 3350336.

Adams, J., Gurney, K., Hook, D., & Leydesdorff, L. (2014). International collaboration clusters in Africa. *Scientometrics, 98*(1), 547–556.

Adams, J., Gurney, K., & Marshall, S. (2007). Patterns of international collaboration for the UK and leading partners. *Science Focus, 2*(6), 1–19.

Anuradha, K., & Urs, S. (2007). Bibliometric indicators of Indian research collaboration patterns: A correspondence analysis. *Scientometrics, 71*(2), 179–189.

Arunachalam, S., & Doss, M. J. (2000). Mapping international collaboration in science in Asia through coauthorship analysis. *Current Science, 79*, 621–628.

Arunachalam, S., Srinivasan, R., & Raman, V. (1994). International collaboration in science: Participation by the Asian giants. *Scientometrics, 30*(1), 7–22.

Arunachalam, S., & Viswanathan, B. (2008). South-South cooperation: The case of Indo-Chinese collaboration in scientific research. *Current Science, 95*(3), 311–313.

Ballard, P. A., Boschma, R. A., & Frenken, K. (2015). Proximity and innovation: From statics to dynamics. *Regional Studies, 49*(6), 907–920.

Basu, A., & Aggarwal, R. (2001). International collaboration in science in India and its impact on institutional performance. *Scientometrics, 52*(3), 379–394.

Basu, A., & Kumar, B. V. (2000). International collaboration in Indian scientific papers. *Scientometrics, 48*(3), 381–402.

Beaver, D. D. (2001). Reflections on scientific collaboration, (and its study): Past, present, and future. *Scientometrics, 52*(3), 365–377. https://doi.org/10.1023/A:1014254214337

Bergé, L. R. (2017). Network proximity in the geography of research collaboration. *Papers in Regional Science, 96*(4), 785–815.

Birnholtz, J. P. (2007). When do researchers collaborate? Toward a model of collaboration propensity. *Journal of the American Society for Information Science and Technology, 58*(14), 2226–2239. https://doi.org/10.1002/asi.20684

Boekholt, P., Edler, J., Cunningham, P., & Flanagan, K. (2009). Drivers of international collaboration in research. *European Commission. https://doi.org/10.2777/81914

Bordons, M., Gomez, I., Fernandez, M. T., Zulueta, M. A., & Mendez, A. (1996). Local, domestic and international scientific collaboration in biomedical research. *Scientometrics, 37*(2), 279–295.

Boschma, R. A. (2005). Proximity and Innovation: A critical assessment. *Regional Studies, 39*(1), 61–74.
Bote, V. P. G., Olmeda-Gomez, C., & de Moya-Angeon, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of American Society for Information Science and Technology, 64*(2), 392–404.

Bouabid, H., Paul-Hus, A., & Larivière, V. (2016). Scientific collaboration and high-technology exchanges among BRICS and G-7 countries. *Scientometrics, 106*(3), 873–899.

Braun, T., & Glaenzel, W. (1996). International collaboration: Will it be keeping alive East European research? *Scientometrics, 36*(2), 247–254.

Caniëls, M. C., Kronenberg, K., & Werker, C. (2014). Conceptualizing proximity in research collaborations. In R. Rutten, P. Benneworth, D. Irawati, & F. Boekema (Eds.), *The social dynamics of innovation networks* (pp. 247–264). Routledge.

Canto, I., & Hannah, J. (2001). A partnership of equals? Academic collaboration between the United Kingdom and Brazil. *Journal of Studies in International Education, 5*(1), 26–41.

Cattaneo, E., & Corbellini, G. (2011). Science under politics: An Italian nightmare. *Embo Reports, 12*(1), 19–22. https://doi.org/10.1038/embor.2010.198

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive-capacity—A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152. https://doi.org/10.2307/2393553

Cummings, J. L., & Teng, B. S. (2003). Transferring R&D knowledge: The key factors affecting knowledge transfer success. *Journal of Engineering and Technology Management, 20*(1–2), 39–68. https://doi.org/10.1016/s0923-4748(03)00004-3

Dawes, S. S., Gharawi, M. A., & Burke, G. B. (2012). Transnational public sector knowledge networks: Knowledge and information sharing in a multi-dimensional context. *Government Information Quarterly, 29*(1), S112–S120. https://doi.org/10.1016/j.giq.2011.08.002

D'Ippolito, B., & Ruling, C. C. (2019). Research collaboration in large scale research infrastructures: Collaboration types and policy implications. *Research Policy, 48*(5), 1282–1296. https://doi.org/10.1016/j.respol.2019.01.011

DST-Clarivate Report (2019). Bibliometric Study of India’s Research Output and Collaboration (2011–2016), Vol. 01. A study commissioned by the Department of Science and Technology, Government of India, Feb. 2019.

DST-Elsevier Report (2020). International comparative performance of India’s research base: Scholarly output, impact, collaboration, knowledge transfer. A study commissioned by the department of science and technology, Government of India, 2011–2016.

Dua, J., Lathabai, H. H., & Singh, V. K. (2022). Measuring and characterizing research collaboration in SAARC countries. *Scientometrics*. https://doi.org/10.1007/s11192-022-04606-0

Ductor, L. (2015). Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics, 77*(3), 385–407.

Dustri, D., Ninka, E., & Turvani, M. (2012). Productivity in innovation: the role of inventor connections and mobility. MPRA Paper No. 38950. Retrieved May 23, 2012, from https://mpra.ub.uni-muenchen.de/38950/.

Fernández, A., Ferrándiz, E., & León, M. D. (2016). Proximity dimensions and scientific collaboration among academic institutions in Europe: The closer, the better? *Scientometrics, 106*(3), 1073–1092.

Finardi, U. (2015). Scientific collaboration between BRICS countries. *Scientometrics, 102*(2), 1139–1166.

Finardi, U., & Buratti, A. (2016). Scientific collaboration framework of BRICS countries: An analysis of international coauthorship. *Scientometrics, 109*(1), 433–446.

Fuchs, J. E., Sivertsen, G., & Rousseau, R. (2021). Measuring the relative intensity of collaboration within a network. *Scientometrics, 126*(10), 8673–8682.

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics, 51*(1), 69–115.

Glänzel, W., & Schubert, A. (2001). Double effort=double impact? A critical view at international co-authorship in chemistry. *Scientometrics, 50*(2), 199–214.

Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). A bibliometric analysis of international scientific cooperation of the European Union (1985–1995). *Scientometrics, 45*(2), 185–202.

Gui, Q. C., Liu, C., & Du, D. B. (2019). Globalization of science and international scientific collaboration: A network perspective. *Geoforum, 105*, 1–12. https://doi.org/10.1016/j.geoforum.2019.06.017

Gupta, B., & Dhawan, S. (2003). India’s collaboration with People’s Republic of China in science and technology: A scientometric analysis of coauthored papers during 1994–1999. *Scientometrics, 57*(1), 59–74.

Gupta, B. M., Munshi, U. M., & Mishra, P. K. (2002). S&T collaboration of India with other South Asian countries. *Current Science, 83*(10), 1201–1209.
Heringa, P. W., Hessels, L. K., & van der Zouwen, M. (2016). The influence of proximity dimensions on international research collaboration: An analysis of European water projects. *Industry and Innovation*, 23(8), 753–772.

Hoekman, J., Frenken, K., & Tijsse, R. J. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662–673.

Hou, L., Pan, Y., & Zhu, J. J. (2021). Impact of scientific, economic, geopolitical, and cultural factors on international research collaboration. *Journal of Informetrics*, 15(3), 101194.

Huang, Y., Shen, C., & Contractor, N. S. (2013). Distance matters: Exploring proximity and homophily in virtual world networks. *Decision Support Systems*, 55(4), 969–977.

Hwang, K. (2013). Effects of the language barrier on processes and performance of international scientific collaboration, collaborators’ participation, organizational integrity, and interorganizational relationships. *Science Communication*, 35(1), 3–31. https://doi.org/10.1177/1075547012437442

Institute history. (2014). IIT Bombay. Retrieved December 25, 2021, from https://www.iitb.ac.in/en/about-iit-bombay/institute-history.

Jeong, S., Choi, J. Y., & Kim, J. Y. (2014). On the drivers of international collaboration: The impact of informal communication, motivation, and research resources. *Science and Public Policy*, 41(4), 520–531.

Jiang, L., Zhu, N. B., Yang, Z. L., Xu, S., & Jun, M. (2018). The relationships between distance factors and international collaborative research outcomes: A bibliometric examination. *Journal of Informetrics*, 12(3), 618–630. https://doi.org/10.1016/j.joi.2018.04.004

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-University research teams: Shifting impact, geography, and stratification in science. *Science*, 322(5905), 1259–1262. https://doi.org/10.1126/science.1158357

Katz, J. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31(1), 31–43.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18.

Kim, M. J. (2005). Korean science and international collaboration, 1995–2000. *Scientometrics*, 63(2), 321–339.

Knoben, J., & Oerlemans, L. A. (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8(2), 71–89.

Kraut, R. E., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 137–162). American Psychological Association.

Kuhn, T. S. (1996). The structure of scientific revolutions: The University of Chicago Press.

Lathaibai, H. H., Nandy, A., & Singh, V. K. (2021). x-index: Identifying core competency and thematic research strengths of institutions using an NLP and network based ranking framework. *Scientometrics*, 126(12), 9557–9583.

Lathaibai, H. H., Nandy, A., & Singh, V. K. (2022). Institutional collaboration recommendation: An expertise-based framework using NLP and network analysis. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2022.118317

Laursen, K., Reichstein, T., & Salter, A. (2011). Exploring the effect of geographical proximity and university quality on university–industry collaboration in the United Kingdom. *Regional Studies*, 45(4), 507–523.

Leahy, E., & Reikowsky, R. C. (2008). Research specialization and collaboration patterns in sociology. *Social Studies of Science*, 38(3), 425–440. https://doi.org/10.1177/0306312707086190

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673–702.

Leta, J., & Chaimovich, H. (2002). Recognition and international collaboration: The Brazilian case. *Scientometrics*, 53(3), 325–335.

Leta, J., das Neves Machado, R., & Canchumani, R. M. L. (2019). Scientific collaboration among BRICS: Trends and priority areas. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 485–504). Cham: Springer.

Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2(4), 317–325.

Mahbuba, D., & Rousseau, R. (2010). Scientific research in the Indian subcontinent: Selected trends and indicators 1973–2007 comparing Bangladesh, Pakistan and Sri Lanka with India, the local giant. *Scientometrics*, 84(2), 403–420.

Marmolejo-Leyva, R., Perez-Angon, M. A., & Russell, J. M. (2015). Mobility and international collaboration: Case of the Mexican scientific diaspora. *PLoS ONE*, 10(6), e0126720.

Mattsson, P., Laget, P., Nilsson, A., & Sundberg, C. J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, 75(3), 555–574.

McManus, C., Neves, A. A. B., Maranhão, A. Q., Souza Filho, A. G., & Santana, J. M. (2020). International collaboration in Brazilian science: Financing and impact. *Scientometrics*, 125(3), 2745–2772.
Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy, 34*(10), 1608–1618.

Whetsell, T. A., Jonkers, K., Dimand, A.-M., Baas, J., & Wagner, C. S. (2020). Democracy, Complexity, and Science. *arXiv:1909.04468v3*.

Yamashita, Y., & Okubo, Y. (2006). Patterns of scientific collaboration between Japan and France: Intersectoral analysis using Probabilistic Partnership Index (PPI). *Scientometrics, 68*(2), 303–324.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.