



What Big data can do in India: Issues in Gender & Higher Education

Abstract

The paper delves with the potential elements where big data can be an operative tool in removing gender inequalities in higher education in India. Big data term is used to denote the gigantic amount of information that want new technologies to assemble ,integrate ,store and finally evaluate the data. To expand the knowledge economy through gender mainstreaming, the formation of gender data base in education is necessary and here the significance on the use of existing data sources has to be scrutinised. The availability of big data enables our data scientists to locate the deliberations on gender concerns in the higher education area. Big data specifically, helps to identify the important aspects of gender inequality in higher education, it provides testimony on women's own interests about higher education and how women approach the gender driven policy deliberations in India. Big data is a complimentary tool to the prevailing sources of data .Big data can emphasise on information that is commonly ignored in usual surveys due to information bias. In specific terms the big data can make the voice of the agency participating in the process of higher education heard. The ability of big data to publish information is extensively acknowledged. The future of big data in Indian education system to remove gender inequality, lies in the creation of infrastructure to store, analyse and make policy discussions on the concerns of gender inequality in the sphere of higher education.

Key words : Big data, Higher Education, India, Gender, Policy issues

Introduction

India in conjoint with other developing countries faces numerous problems in data gathering, dissemination and analysis. Often there is duplication in data gathering process. For proper policy efforts in removing gender inequality particularly in the higher educational front proper suggestions are needed to overcome the data lacuna because without data there can be no developmental efforts. The basic objectives in data generation should be to avoid duplication, ensure efficiency in data collection and remove all clumsy data generation exercise. To drive a knowledge economy the creation of gender data base in education is necessary and here the importance and use of data sources has to be scrutinised. The availability of big data enables our data scientists to refocus the considerations on gender concerns in the higher education area. Big data in particular helps to identify the important aspects of gender inequality in higher education, it provides information on women's own concerns about higher education and how women access the gender driven policy considerations in India. Big data is a complimentary tool to the already existing sources of data generation and processing. Big data can focus on information that is usually overlooked in regular surveys due to information bias. In particular big data can make the voice of the agency participating in the process of higher education heard. The ability of big data to disseminate information is widely acknowledged, however the future of big data in Indian education system to remove gender inequality lies in the creation of infrastructure to store, analyse and create policy debates on the issues of gender inequality in the realm of higher education. This paper attempts to make a discussion on the existing data gaps in the higher educational sector and how big data may compliment the gender dimension to it. The paper further analyses how the feedback process of policy prescription on the gender equality aspect can be continually assessed with big data. The design of the paper is as follows Section II elaborates on the meaning of big data and the challenges associated with it. The importance of higher education for women in India and how big data can be effectively used as a gender equalising tool is elaborated upon in Section III. The nature of gaps in the collection and analysis of existing database is discussed in Section IV. Here the importance of big data to fill up the gaps is also outlined. Further through questionnaire design (applicable for big data collection) the feedback and perception of the stakeholders of the already existing gender sensitized policies in higher education is also discussed here. The paper is finally concluded in Section V

Big Data its meaning and major challenges

Big data term is used to refer to the huge amount of information that need new technologies to collect, assimilate, store and finally analyze the data. The data can be analysed in high speed and the system stores huge information content. Two popular definitions used include 3V definition- volume, variety and velocity and 4V definition apart from the categories referred to in 3V, 4V category adds value. A Big Database system allows the inclusion of wide variety of data, many different modes of analysis. The scope of predictive options is increased. Big data is often used to refer to large quantities of data, it can be structured and unstructured. The analysis with big data leads to better policy decisions.

The sources of big data are of three categories: Streaming data- this includes data that reaches the system of information through a web of connected devices. Social media data- this is the data on social interactions, it often unstructured so it imposes the task of structuring such data sets. Public sources- This includes data available from various governmental portals in the different countries. In today's globalised world big data often faces the need of accessing the data with high speed- this increases as the degree of granularity increases. Apart from maintaining the quality synchronizing the diverse sets of data sources to a common platform is a major prerequisite.

Challenges with big data

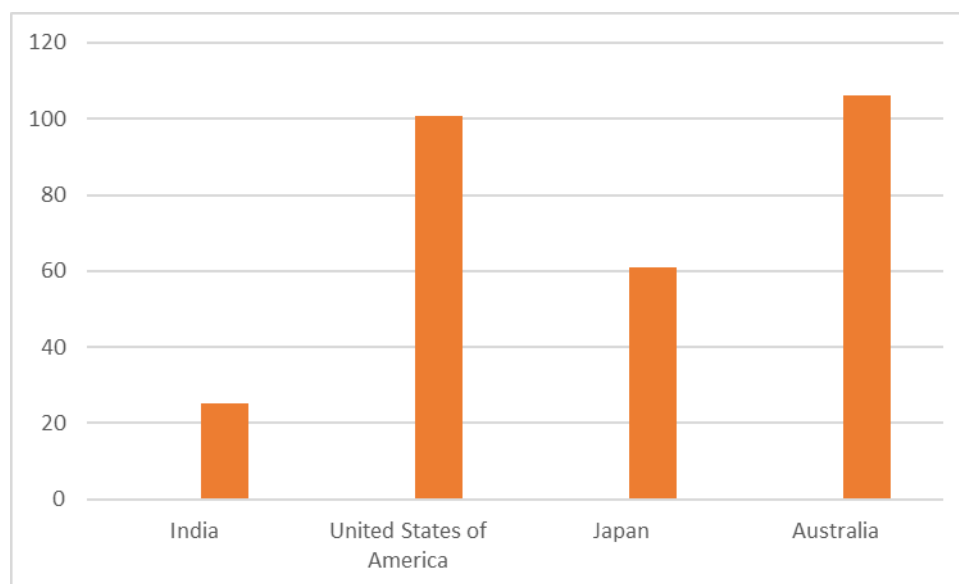
In today's era of information technology big data is becoming inseparable in major policy decisions. However there are a number of challenges with big data. Challenges with big data: are mainly of three types 1) ethical, 2) analytical and 3) technological. The analytical challenge includes interpretation and selection bias and lack of access to tools that interpret big data sets. Collection of big data to remove gender gaps need different initiatives both at institutional and state levels. The merit of big data is that it frees a data interpreter to make causal inferences. Big data generates causal relations between variables. The ethical challenge relates to security and privacy of the individual stake holder. This is of considerable importance in a developing country like India where security of girls may be affected by disclosure of information. The technological challenge includes- storing and analyzing data, disseminating information. Further aggregating the information at a centralized level is an utmost technological challenge because there is a need in India to translate diverse languages and types on a common platform.

III. Gender gaps in higher education in India and the relevance of Big data

Education is the precursor in creating the knowledge economy in India. The need of today's economy is to generate a new set of skills that will meet the new demand of the market. Higher education and further tertiary education is essential for furtherance of the knowledge economy. It is observed by rating organisations worldwide that despite progress Indian universities are still to put efforts to maintain high quality standards. Against this backdrop comes the need to expand gender based higher education in India to foster inclusive growth. The National Empowerment Policy (2001) for Women promised that equal access of all girls will be ensured in the realm of education. Further time targets had been focussed to ensure speedy enrollment of girls from socially backward communities. Gender equality and women's empowerment can be increasingly achieved by increasing the capability of the women. Education is an investment whereby gender equality and gender empowerment can be achieved. In the era of big data, education and investment in human capital can be successfully designed and can be a successful guide to policy decisions. UNESCO (1993) identified the following areas where significant databank needs to be developed to ensure higher enrollment of women in the higher educational level--'family attitudes; cultural stereotyping; propagation of the glass ceiling syndrome'(which promotes the discriminatory character of advancement). The development of legislation in today's inclusive era of growth to enable women to participate in higher education is being increasingly felt. Interestingly The National Assessment and Accreditation Council (NAAC) (2009/2010) study reveals that there is ghettoisation of women in higher education, this is not a healthy trend particularly when we are interested to ensure that girls reap the human capital benefits of higher education. As per Eleventh plan document (2007-2012) enrollment in higher education stand at 19% for males and 15.2% for females, so there is still gender disparity in enrollment in higher education in India. World Bank (2005) observes that in the era of knowledge economy India requires to put in more efforts to use knowledge for enhancing the productivity of the economy. The task of

propagating the knowledge across all sections of the community should be a development objective of utmost importance.

Figure I: Gross Enrolment Ratio in Higher Education among Women (2014), India and some country experiences



Source:Unesco (data files), compilation self

The figure (I) points out that as far international country comparisons are concerned the position of women in the realm of higher education is at low levels. The gap between developed nations of the world continue to broaden. At this point it is useful to consider the use of big data in helping to enhance the information on women's individual needs, abilities and the aggregative behaviour of the same on higher educational front. The technology of big data can help us to locate the areas of inequalities at granular levels, (implying how large and detailed data sources can be utilised at a disaggregated level). The prescription of policies for women to pursue higher education in India has been slow owing to lack of feedback and lack of adequate database. For effectiveness of policy implementation to remove gender inequality in higher education the application of big data is of utmost importance.

It is pertinent to mention here about an organisation which is painstakingly trying to deliver the need on big data to make it a toolkit to address gender equality issues-it is Data2X. The relevance of big data is reiterated in Data2X. We have lessons to learn from the platform of Data2X on our action plan on generating big data set on gender and higher education in India. Data2X is led by the UN foundation with support from from Bill & Melinda Gates Foundation and William & Flora Hewlett Foundation. The mission of Data2X is to provide a technical platform to improving the quality of gender data world wide. The action of both governmental and non governmental organisations in India becomes all the more challenging to initiate action on areas of data gap. It is acknowledged by Data2X that constraints continue in the lives of girls and women in developing countries because of limited data and biases relating to the importance of gender data collection. This exercise tries to generate motivation that to ensure higher education for women, it is necessary to develop statistics that reflect the inequality in educational access and quality. Data2X observes that at present in most developing countries measures on education quality is based on inputs which are not standardized. Again information on socially excluded girls should be better planned-because girls from socially excluded communities face double disadvantage as result not only the enrollment levels are poor but learning outcomes are also poor. The Data2X stresses on the importance of developing information base on transition rates of women from educational attainments to participation in workforce. Relating to higher educational area of big data sets Data2x opines that big data can generate interim data on observations on women's mobility conditions in pursuit of education, the conditions that are averse to their mobility and the suggestions from their view point to improve mobility. In sum the task of big data lies in supplementing the existing data sources on higher education in India and in redesigning questionnaire in the official statistical base. A commendable set of initiatives taken up by the University Grants Commission for making higher education for women a priority- include i) Day care centres in universities and college.

The purpose of this scheme is to promote enrolment of mothers willing to study, while their children are secured and protected. ii) Postgraduate scholarships for single girl child pursuing higher /technical education. The purpose of such a process is to support education for those who would be the only future prop of their elderly. The recognition is given to the families who value the norm for small families. iii) Construction of Women's Hostels for colleges. This is an initiative to provide infrastructural support to the colleges for their women students and teachers so as to enable the colleges to further the goal of expanding higher education for women. iv) Development of Women Studies in Universities and colleges. This is an effort in gender mainstreaming the process of higher education. v) Scheme of Capacity Building of Women Managers in Higher Education. This initiative is taken to put into vogue the recommendation of UNESCO (1993). The initiative would develop a gender friendly environment and remove the glass ceiling. vi) Post-Doctoral Fellowships for Women. This programme is a systematic effort to accelerate the research pursuits among the girls who have high research aspirations but are yet to get employed in the job market. So six formidable schemes are initiated to promote the higher educational pursuits of India's women. The accompanying task is to evaluate the efficacy of such schemes in promoting gender equality in higher learning. The first step towards evaluating the efficacy of such schemes lies in the generation of systematic and periodic data evaluation. Here the role of the big data is of utmost importance.

The broad issues that can be addressed with big data in the context of women in higher education in India are -1) the low enrollment of women in higher education ,(though it is increasing);2) on delivering a gender based higher education curricula; 3) evaluation of legislation procedure relating to women's higher educational accessibility and 4) development of gender based infrastructure and mobility. Kapur and Mehta (2004) argue that today Indian universities are not been used as institutions of signalling device rather centres of generating minimal qualification for preparation for job market needs. The outcome of such a process is growth of private coaching centres as parallel education system which have indeed become the space for deriving social mobility in education and also reaping the opportunity in job markets. However no statistical data base exists to examine the nature of mobility among socially excluded gender groups through parallel education system. This is a compelling area where data gaps needs to be filled. The big data can be useful solution here because of its technology.

The existing gaps in data collection and the complementarities of Big data

The major sources of data on higher education in India is obtained from The All India Survey on Higher Education (AISHE) Ministry of Human Resource and Development, Government of India; National Sample Survey Organisation(NSSO), Ministry of Statistics and Programme Implementation, Government of India; Social and Cultural Tables,Census of India, Registrar General & Census Commissioner of India. The All India Survey on Higher Education (AISHE) provide data on Higher education. The survey covers data base on all recognised institutions of India which are associated in delivering higher education in India. The efficiency of this survey is it conducted through a paperless mode through the data collection portal <http://aishe.gov.in> thus it is a formidable initiation towards developing a possible big data base in future. The broad data parameters include teaching position, students enrolled in terms of discipline; level; across gender. Again social category wise distribution is also provided. The database captures representation of Minority students 'enrolment in various disciplines. There is also database on foreign student enrolment participation. A few indicators of educational development is also provided in the database- institution density; teacher -pupil ratio, Gender parity index. The survey is conducted on an annual basis. Though NSS data collection surveys focus on social and religious groups of households the data interpretation remains challenging because of overestimation possibilities. This is because of sample survey data and data becomes fewer as we move down the disaggregation scale. The Census of India, Registrar General Government of India, provides data on higher education. The Census definition of higher education is a broader compared to the definition provided by the AISHE. However district level aggregate data is available from the Census Tables.

The major data gaps identified with the existing data centres here may include---1) lack of cross country comparison standards; 2) lack of disaggregate data on exclusion and wealth quintiles and 3) the transition map from educational sector to workforce opportunities is essential –this is a major lacuna. These facts will help to disseminate information on the levels of education for the most disadvantaged girl students and their incorporation in the workforce. Further the existing measure of educational data on outcomes is in the domain of quantity of higher education in different courses thus it is not sufficient to judge quality outcomes on a standardized basis. In particular big data would be able to focus on understanding students potential and capabilities. To enable the institution to take corrective action to reduce dropout of students after evaluating the response of students from the data sets. Even the causes of failure and lack of interest can be assessed. Measures can be obtained from the analysis of big data on teacher effectiveness.

The central issues related with big data and gender equality framework could include:

Database on access, retention and completion

Location proximity of institutes of higher learning for girls in higher education

The data on percentage of continuation in higher education from high school education (across urban/rural/remote/hilly/location hazardous areas) (both boys and girls).

Security perceptions in public transport for girls travelling long distance. The information needs to be documented on the basis of class/caste religion/ wealth groups.

Reasons of discontinuance if any: household responsibilities, marriage/child bearing responsibilities/ cultural practices.

Tasks ahead

Big data can be effectively used to quantify the gender gap in relative and absolute terms in the realm of higher education in India. It can be used to assess the progress of women's learning abilities and assessing the changes over time. The data sets generated in the area of higher education can be utilised to study the impact of education in enhancing women's capabilities and their participation in economic activities. The data sets can complement the existing evaluation procedures on quality enhancement in learning. The data generated can eliminate response bias of direct questionnaire methods of surveys. The data can be generated from the following ways: Student assessment on learning outcomes, On Usage of internet and .Data generated from the social media of the required group of students. In the recent decade there is media reporting on gender based violence on students, continuance appraisal is necessary and proper research programmes need to be conducted to ensure that there is no inhibition from households to send their daughters to colleges. The big data can generate state /district specific surveys to focus on priority issues; can organise regular annual surveys to complement the existing surveys; can generate household based welfare /capability indicator based survey from where broad welfare index in relation to higher education can be calculated and can conduct institutional level surveys where the infrastructure mapping, research and teacher effectiveness can be assessed. The table (I) provides a comprehensive tabular format on the indicators which big data can generate on the well-being of women pursuing higher education. In addition to the format in Table (I), the Table (II) provides a schedule for specific assessment on the Scheme of Capacity Building of Women Managers in Higher Education, which is the most desirable policy option in India. The data generated has to be analysed with great speed depending on the degree of granularity. Again the quality assurance of the data generated must be ensured. Given the diversity of Indian demography, a large and diverse sets of data will be incorporated in the analytical platform, so proper synchronization across data sources is essential.

Conclusion

The paper explores the possible dimensions where big data can be an effective tool in removing gender inequalities in higher education in India. After exploring on the meaning and concept of big data the study specifically discusses the shortcomings of the existing data sets and where big data is required. In particular the paper identifies that big data can be used to quantify women's disadvantage in higher education particularly across the excluded population. In addition the relative position of women in comparison to the men can also be made. The study analyses that through big data the agents

themselves can discuss about the policy issues applicable to them, highlight on their perceptions and difficulties in the process of learning. Big data can assess the outcome of learning abilities of women in higher education and how their capabilities can be improved to contribute in the growth of knowledge economy. Big data can be unequivocally used to judge the relevance and use of gender data in higher education. Big data helps in mapping certain specific problems which in general may not affect the larger milieu of women. In sum big data has certain desirable features like : high speed , better coverage and comparability which make it relevant for use for reducing gender inequality

Table I: A Comprehensive Format on Indicators for Big data for Well-being and Higher Education among Women

The variable to be addressed	The expected feedback from such variable
Learning outcomes of Girls and Boys in Different levels of Higher education based on choice based credit system	The quality and teacher effectiveness can be assessed from this variable and the relative gender inequality can be decomposed. This should be conducted on annual basis
Girls from excluded population discontinuing education and the reasons of the discontinuance	The focus of this indicator is to study the scope of inclusion in higher education, the survey should be conducted repeatedly depending on the volume of data generated.
Girls from excluded population discontinuing education across various wealth /income quintile	The impact of wealth /income mobility on higher learning can be evaluated, on annual basis it should be conducted
Spending patterns in economically disadvantaged family in relation to college /university needs across boys and girls	Assessing gender disparity in spending patterns in households is the focus here
Frequency of Cash transfer to economically vulnerable girls' bank account from Government of India	The effectiveness of the scholarship programmes intended towards the girl child of India will be duly studied across budgetary allocation period.
Response of students on teacher student interactive learning in the classroom	The quality dimension of higher education will be monitored. The portal for this data source will be inaccessible to the teachers and administrator of the institution for anonymity and transparency, this should be conducted after every academic session.
Response of students on gender mainstreaming	This option will evaluate on how to expand the programme of Women studies in colleges
Reasons of lack of access for the excluded girls in higher education in educationally backward districts	This variable will generate feedback on the utility or otherwise of Day care centres in colleges or universities /the need for hostels for girls will also be assessed

Table (II): Schedule to evaluate the Scheme of Capacity Building of Women Managers in Higher Education, in India

To generate question on the opportunity of the female administrator to judge her performance in comparison to the male colleague.
To create a self-assessment format to evaluate whether the women and men administrators are able to handle turmoil situations, this aspect will be on the way towards gender sensitising that there is no

gender disparity across managerial traits.
--

To enable the administrators to mention how they handle gender stereotypes in workplace

References

- I. All India Survey on Higher Education (AISHE), Ministry of Human Resource Development. Government of India <http://aishe.nic.in/aishe/home> accessed on 2/1/2017 at 13:24 hours.
- II. Census Of India, Social Cultural Tables , Government of India, Ministry of Home affairs , Office of the Registrar General and Census Commissioner , India. http://censusindia.gov.in/Tables_Published/C-Series/c_series_tables_2001.aspx Accessed on 2/1/2016 at 13:32 hours
- III. Eleventh Five Year Plan (2007-2012) Planning Commission, Government of India. <http://www.planningcommission.nic.in/plans/planrel/11thf.htm> accessed on 3/1/2016 at 19:07 hours.
- IV. Gender data resources. <http://data2x.org/resources/> accessed on 1/1/2017 at 9:11 hours
- V. India - India and the knowledge economy - leveraging strengths and opportunities(2005). Report no 31267-IN World Bank. <http://documents.worldbank.org/curated/en/802981468253192956/> accessed on 3/1/2016, 19:15 hours
- VI. Kapur, D., & Mehta, P. B. (2004). Indian higher education reform: From half-baked socialism to half-baked capitalism. Center for international development working paper, 103.
- VII. National Policy for the Empowerment of Women (2001); Ministry of Women and Child Development, Government of India. <https://india.gov.in/national-policy-empowerment-women;>
- VIII. accessed on 3/01/2016;18:56 hours.
- IX. National Sample Survey. Ministry of Statistics and Programme Implementation, Government of India. <http://www.mospi.gov.in/national-sample-survey-office-nss> accessed on 2/1/2017 at 13:27 hours.
- X. Sixteenth Annual Report (2009/2010). National Assessment and Accreditation Council. <http://www.naac.gov.in/docs/Naac%20Annual%20Report.pdf> accessed on 5/1/2016,11:13 Hours.
- XI. UNESCO data files. http://data.uis.unesco.org/Index.aspx?DataSetCode=EDULIT_DS accessed on 3/1/2016,at 19:10hours
- XII. University Grants Commission, India http://www.ugc.ac.in/ugc_schemes/ accessed on 3/1/2016 at 19:22 hours

Dr. Sudeshna Ghosh
Associate Professor
Department of Economics
Scottish Church College
Kolkata

Copyright © 2012 - 2018 KCG. All Rights Reserved. | Powered By: Knowledge Consortium of Gujarat