**REPORT ON THE SURVEY ON**
**DATA MANAGEMENT PRACTICES OF HEIS IN INDIA**

## *Introduction*

Today's Higher Education Institutions (HEIs) are functioning in a very competitive, dynamic and data-driven environment. It is becoming essential to understand and appreciate the significance of data management from an end-to-end perspective in both academic and administrative environments. The higher education administration is bound to adopt data management systems and practices to ensure accurate and timely reporting to both internal and external stakeholders. A well planned and executed Data management system would provide meaningful information to the leadership to make both short and long term strategies. Further, data management can be extended to bring out insights through analytics to grow and sustain competition.

In India, with the advent of various initiatives like NAAC's data-based accreditation framework, NIRF, AISHE survey etc., the need for robust data management practices within the HEIs' has taken precedence.
This survey is the second phase of the project 'EQUAM-BI' (Enhancing Quality Assurance Management & Benchmarking Strategies in Indian Universities) a prestigious Erasmus+ project. The project coordinators are University of Barcelona (UB) and ANECA, Spain and NAAC, India. The Survey is broadly structured to understand the type of data collected, persons responsible for data collection, the mode and purpose of collection, and the IT infrastructure that facilitates the process and reliability of data.

The objectives are:

- To assess the data collection, collation and management processes currently in place in the Indian HEIs.
- To understand the purpose and extent of IT enablement in different processes with regard to data collected, persons responsible, reliability and suitability of tools, support structures and level of data trustworthiness currently available in certain core processes like Student Life Cycle, Human Resources, Finance, Purchase and Maintenance, Research and Community Outreach.
- To understand the processes and systems, HEI's have developed to comply with the continuous data requirements of regulatory, statutory, accreditation and ranking agencies and the challenges faced.

## *Methodology*

Primary data was collected through a survey instrument and secondary data from the websites of regulatory bodies like the University Grants Commission and Ministry of Human Resource Development (All India Survey of Higher Education). In the subsequent paragraphs the survey instrument development and sample selection method are explained.

*Survey Instrument*

The study was conducted based on primary survey of HEIs and has also drawn information from secondary sources. This study has drawn inputs from several relevant information sources like websites of NAAC, NIRF, AISHE, HE Commission UK and IQAC Symbiosis International (Deemed University). The questionnaire sought both quantitative and qualitative information on:

1. Profile of the HEI
2. Extent of information technology enablement for the administration of different processes like Student Life Cycle Management, Human Resources, Finance, Purchase and Maintenance, Research, Community Outreach.
3. The reliability, usefulness, trustworthiness, suitability to measure the extent to which the outcomes (performance) have matched expectations.
4. The data security measures, existence of written policies, availability of support structures, expenditure incurred.

5. The systems in place for addressing the continuous data requirement for internal and external Quality Assurance purposes and the challenges that are being faced.

The draft questionnaire was shared with all partner universities from India and Europe, NAAC, ANECA for their suggestions and inputs. The revised questionnaire was then administered to the sample, selected as explained in the next section.

*Sample*

The number of HEIs in India (excluding affiliated colleges) was 949 as on June 15, 2018, which constituted the population of the study. The population was then classified based on the HEIs type of university (Central, State, Deemed-to-be, State Private and Institutes of National Importance) and their geographical location (States/Union Territories). The sample of HEIs was selected out of the population in three distinct steps.

- First, based on initial examination, the sample proportion was fixed at 10%-15% of the population (95 – 140 HEIs), based on the geographical spread and HEI type. Stratified random sampling method was followed to select the samples
- Second, it was decided to include the all 29 respondents which participated in the earlier study of the EQUAMBI conducted during July to November 2018. Further, emphasis was given to those types of HEIs which did not form part of the report of the previous study because they had not responded to the questionnaire of the study. So, concerted efforts were directed towards identifying and including institutes of national importance and State Private Universities.
- The 130 sample size was drawn from 29 states and two Union Territories of India.

The survey was administered online to the sample HEIs during the period March (3rd week) till April (2nd week). After a series of continuous follow ups valid and full responses were received from 29 HEI's only from 16 states of India. The maximum responses were received from Meghalaya (100%) and Tripura (100%), followed by Kerala (60%), Jharkhand (50%), Karnataka (46%) and Tamil Nadu (40%), Assam, Gujarat and Uttarakhand (33% each), Maharashtra (29%), Delhi, Telengana and West Bengal (25% each), Madhya Pradesh and Punjab (17% each) and Uttar Pradesh (10%).

## *Analysis & Findings*

The questionnaire comprised both quantitative and qualitative questions segmented into two broad sections. Section A captured the Profile of the respondents, and section B captured the Data Management processes and practices including qualitative questions.

**Section A : Profile of the Respondents:**

The first section on profile sought details regarding; Name, Email address, Type of HEI, Total number of Registered Students, Full-time Faculty Members and Non-Teaching Staff / Employees.

1. Out of the 29 HEI, 15 were State Universities (constituting 52%), seven Deemed Universities (constituting 24%), three Central universities (constituting 10%), two State Private Universities (constituting 6.9%) and two Institutes of National Importance (constituting 6.9%) of the sample size.
2. The distribution of registered students, full-time faculty members and non-teaching staff among the 29 sample HEIs is presented in the table below.

| Type of HEI | Type of HEI % | Registered Students (%) | Full-Time Faculty Members (%) | Non-Teaching Staff (%) |
|---|---|---|---|---|
| Central University | 10.0 | 9.7 | 22.6 | 11.3 |
| State University | 52.0 | 66.5 | 45.9 | 51.4 |
| State Private University | 7.0 | 1.5 | 0.5 | 1.0 |
| Deemed-to-be University | 24.0 | 19.5 | 28.9 | 32.9 |
| Institutes of National Importance | 7.0 | 2.8 | 2.1 | 3.4 |
| **Total** | **100.0** | **100.0** | **100.0** | **100.0** |

It can be seen that State Universities which formed the major portion of the sample among the type of HEIs had 67% of the registered students, 46% of the full-time faculty members and 51% of the non-teaching staff. Deemed-to-be Universities which constituted 24% of the sample had 20% of the students. 29% of faculty members and 33% of non-teaching staff. The Central Universities which formed 10% of the sample had 10% of students, 23% faculty members and 11% non-teaching staff. The Institutes of National importance which formed 7% of the sample size had 3% registered students, 2% full-time faculty members and 3% non-teaching staff. The state private universities which constituted 7% of the sample size had just 1.5% of the registered students, less than one percent of the full-time faculty members and just 1% of the non-teaching staff.

3. The 29 HEI samples which participated in this study represented only 16 Indian States. The distribution of the samples stated that the state of Karnataka contributed six HEIs (representing 20.7% of the sample), followed by Tamil Nadu with four HEIs (13.8% of the sample) and Kerala with three HEIs (10.3% of the sample). Assam, Gujarat and Maharashtra had two HEIs each (constituting 6.9%) participating in the survey. Only one HEI each (constituting only 3.4% of the sample) participated from the states of Jharkhand, Madhya Pradesh, Meghalaya, New Delhi, Punjab, Telengana, Tripura, Uttar Pradesh, Uttarakhand and West Bengal.
4. The regional distribution of the 29 HEI samples revealed 14 HEIs (representing 48% of the sample size) belong to Southern India, four HEIs (representing 14%) each belonged to Northern, Western and North Eastern India, two HEIs (7%) belonged to Eastern India and only one HEI (constituting 3%) was from Northern India.

**Section B : Data Management**

1. Nineteen HEIs (constituting 66%) had *greater than 75% up to 100%* had IT enablement for overall administration, seven HEIs (constituting 24%) *had greater than 50% up to 75%*, two HEIs (constituting 6.9%) *had greater than 25% up to 50%* and only one HEI (constituting 3.4%) had up to 25% IT enablement.
2. The analysis of the extent of IT enablement for data collection in certain key processes of the HEI showed the following;
    a. Student Lifecycle Management: Majority of the respondents (above 70%) reported that they had ***High to full IT* enablement** in *application and admissions, processing and declaration of results, academic calendar, conduct of examination, students' feedback and academic timetable*, *Course / Curriculum Development, Co-curricular/Extra-curricular Activities of Students, Grievances-Handling,* and *Students Progression.* Comparatively, less proportion of the respondents reported High to full IT enablement (above 50%) in *students' campus placement, and alumni engagement and students' medical support and*

*insurance*. It is also found that seven HEIs have not implemented / enabled IT in the components namely students feedback (1), co-curricular/extra-curricular activities, grievances handling (1 each), alumni engagement (2) and students medical support and insurance (1).

b. Human Resources: **High to full IT enablement** was reported by 80%-90% of the respondents in *payroll processing and staff recruitment*. Comparatively less respondents reported IT enablement was reported 70%-80% of the respondents in *faculty recruitment, faculty appraisal, and staff appraisal.* Only 60%-70% of the respondents in *faculty separation, staff separation and leave management*. And 50%-60% of the respondents in *faculty and staff grievance-handling, faculty and staff medical support and insurance.*

c. Finance: **High to full IT enablement** was reported in all four finance parameters namely, *accounting cycle, banking transactions, internal reporting / auditing, statutory reporting (GST, Income Tax, Budget, Financial Statements etc.)* by more than 90% of the respondents.

d. Purchase and Maintenance: **High to full IT enablement** was reported in all four purchase and maintenance parameters namely, *Assets / Project Purchase (CAPEX)* by 86% of the respondents, *Recurring Purchases (OPEX) by 79% of the respondents, Environment (Green/ Carbon Footprints)* by 72% of the respondents and *Infrastructure Maintenance* by 69% of the respondents. Two HEIs have not implemented / enabled IT in the components namely Environment (Green/ Carbon Footprints) and one in Infrastructure Maintenance.

*e.* Research: **High to full IT enablement** was reported in *Research Projects* by 93% of the respondents. *Research Publication (including Bibliometrics), Research Grants, Intellectual Property and Progression of Doctoral Students from Admission through Registration to Thesis Submission* by 86% of the respondents. *Doctoral Thesis Examination* by 82% of the respondents. *Major Research Equipment in possession and their maintenance* by 72% of the respondents. One HEI had *not implemented / enabled IT* in the components namely *Intellectual Property* and one HEI in *Major Research Equipment in possession and their maintenance.*

f. Community Outreach: *Institutional Social Responsibility (Short-term) by 72% of the respondents, Institutional Social Responsibility (Long-term) and Facilities/services available for use by outsiders* by 66% of the respondents. *Short term interventions (Financial Commitment) and Long term interventions (Financial Commitment)* by 59% of the respondents. Two respondents each have not enabled IT in *Institutional Social Responsibility (Short-term), Institutional Social Responsibility (Long-term), Short term interventions (Financial Commitment) and Long term interventions (Financial Commitment)* respectively. One respondent reported IT is not yet enabled in the *Facilities/services available for use by outsiders*.

3. The analysis of the 29 responses about the level of data management shows that 52% (15) *collect,* 24% (7) *store,* 14% (4) *retrieve* and 10% (3) respondents *share data centrally* at the University level only. 55% (16) *collect*, 28% (8) *store*, 3% (1) *retrieve* and 14% (4) respondents *share data de-centrally* at the Faculty/Department level only. 48% (14) *collect*, 24% (7) *store,* 10% (3) *retrieve* and 17% (5) respondents*, share data centrally* at the both the levels.
4. High to very high degree of reliability is reported with respect to data collection for all the parameters which include Teaching/Learning, Research, Internationalization, Facilities, Procedures and Outcomes. (Note: In the 10-point scale, 10-9 is highly reliable, 8-7 is reliable,6-5 Neutral, 4-3 Unreliable & 2-1 Highly Unreliable)
5. The respondents have reported that data collection for all the parameters which include Teaching/Learning, Research, Internationalization, Facilities, Procedures and Outcomes are useful to extremely useful for the Quality Assurance processes. (Note: In the 10-point scale, 10-9 is extremely useful, 8-7 is useful,6-5 Neutral, 4-3 Unuseful & 2-1 Not at all useful)
6. A vast majority of the respondents (90%) have reported that they agreed and strongly agreed that the data collection processes and practices at their HEI are completely trustworthy. (Note: In the 10-point scale, 10-9 is strongly agree, 8-7 agree, 6-5 neither agree nor disagree, 4-3 Disagree & 2-1 Strongly Disagree)
7. Around 80% of the respondents are satisfied and strongly satisfied that the data collection instruments are suitable to measure the extent to which outcomes matched the expectations.

(Note: In the 10-point scale, 10-9 is Strongly Satisfied, 8-7 Satisfied, 6-5 Neither satisfied nor dissatisfied, 4-3 Dissatisfied & 2-1 Strongly Dissatisfied)

8. The analysis of the *authorities (Registrar, Information Officer (University level), Administrative Officer (at Institute level), IQAC Coordinator, Manager-IT, Chief Finance Officer and Third-party Vendor)* responsible for data management processes and practices for reporting purposes *(Accreditation & Ranking, Regulatory Bodies, Governance, Statutory Compliance (Audit & Taxation), Statutory Compliance (IT)* at the HEIs stated the following.
    a. Eleven respondents reported that the Registrar was responsible for all the five data management processes and practices for reporting purposes. Five respondents reported that Registrar is responsible for data reporting Regulatory Bodies and Governance only. All the other respondents responded with various combinations of data reporting purposes as the Registrar's responsibility.
    b. The role of the Information Officer and Administrative officer appeared to be distributed in all the different combinations of the data reporting purposes.
    c. Nearly 50% of the respondents (14 out of 29) stated that the IQAC coordinator is solely responsible for data relating to accreditation and ranking purposes. Five respondents reported that the IQAC coordinator is responsible for all the five data management processes and practices for reporting purposes.
    d. Twelve respondents stated that the Manager/Head of Information Technology is responsible for Information Technology related statutory compliances. Four respondents stated that Manager/Head of IT is also responsible for all the five data management processes and practices for reporting purposes.
    e. Nearly two-third of the respondents (19 out 29) stated that Statutory Compliance (Audit and Taxation) was solely the responsibility of Chief Finance Officer (CFO)/Head of Finance. Three respondents stated that CFO/Head of Finance is also responsible for Governance related data and another three respondents stated that Head of Finance is responsible for all the five data management processes and practices for reporting purposes.
    f. Eight respondents stated that they have outsourced the data management responsibility to third-party vendors, the break-up of which is; one for Governance and Statutory Compliance (Audit and Taxation), three each for Statutory Compliance (IT) and Governance, one for Statutory Compliance (Audit and Taxation).
9. 90% of the respondents had Firewall in place and Antivirus per terminal, 86% had Data Centre with precision AC and Fire alarms, and 72% had Centralized Antivirus Gateway.
10. Higher percentage of respondents reported the existence of written policies for information security (69% of respondents), data management (66%), and data collection procedures (62%) while 59% reported for policies for privacy and only 55% of the respondents reported the existence of written policies relating to data collection.
11. Almost all the respondents stated that they agree to strong agree (on a 5-point scale) that they have adequate data management support structures in the form of;
    a. Data Management processes (collect, store, retrieve and share), a dedicated center, people and hierarchies, access and authorization policies and process for Continuity Planning and Data Recovery (over 90% of the respondents). Periodical Data Backup processes (86% of the respondents), availability of sufficient Financial Resources (83% of the respondents) and hardware and software adequacy by 76% of the respondents.
12. The analysis of the responses relating to the ownership status of IT infrastructure revealed that above 90% of the respondents have fully-owned Computer hardware platforms and Operating system platforms, 83% fully-owned Networking, Telecommunication platforms and Databases, 79% of respondents fully-owned Internet platforms, 76% for System Integration platforms while only 62% of the respondents reported that they fully owned Enterprise Resource Planning (ERP) and other software applications. Though the proportion of fully-outsourced model of IT infrastructure is negligible among the respondents, it is interesting to find eight respondents had part-outsourced model their ERP systems, five respondents have part-outsourced their system integration and networking platforms and databases were partly outsourced by four respondents.
13. All the respondents have stated that they strongly agree or agree (on a 5-point scale) to the following parameters that satisfy the purpose of data management; *Creating a data repository,*

*Standardization of processes, Accessibility of data, Enhancing the productivity of staff and students, Data-Driven Decision Making, Paper-less or Less paper process environments, Real-time data processing, Timely reporting to internal and external stakeholders (University top management, Regulatory bodies and Accreditation and Ranking agencies), Reduce cost of operations, Information Transparency, Seamless integration of academic and administration processes, Timely conduct of examinations and declaration of results, Equity action, Benchmarking and Quality Improvement.*

14. Nine respondents have stated that they spend greater than 4 up to 6% of their total expenditure on IT infrastructure and processes. Six respondents each stated they spend in the range of greater than 2 up to 4% and greater than 8 up to 10% respectively. Three respondents each stated they spend in the range of greater than 6 up to 8% and less than 2% respectively. Only two respondents stated that they IT infrastructure spending ranged from greater than 10% (34% and 20% respectively).
15. 83% of the respondents have stated they are strongly satisfied / satisfied with the information technology infrastructure provided for data management, while 14% of the respondents gave neutral response and only one respondent expressed dissatisfaction.

*Qualitative Analysis*

16. The Questionnaire sought ~~to seek~~ inputs from the respondents regarding any practice or system that their HEI has developed to address the continuous data requirement for internal and external Quality. Most of the Universities responded that they have templates in Excel sheet, Custom made software, online portals developed in house for data collection, storage. Some have online learning management systems and system for financial management and one university has responded that they have installed ERP system. Dedicated IT portals through the pre-defined services indicators pre- embedded into the IT portals has been installed. This ensures quality check , relevance of the data to match the requirements of statutory, regulatory bodies and report generation. Few Universities have an IQAC department or a statistical cell for regular and timely reporting to regulatory and statutory bodies, one of the universities has established a NAAC Directorate and engaged faculty members to coordinate with departments for collection and compilation of data and are in the process of developing an android app. It is worth mentioning here that though the respondents have mentioned the importance of collecting data in an accurate and timely fashion as one of the priorities, the data management for majority of the Universities seems to be decentralized and collected through various levels of the University.

17. The challenges faced by universities for data collection required for internal and external reporting revealed a wide and varied response. Some of the challenges were relating to data collection and management whereas others were regarding the appropriateness of certain parameters identified by ranking/ accrediting agencies. They have been classified accordingly in the following paragraphs:

    (A) Data Collection & Management:

    a. Some Universities have responded that they face challenges regarding collection, monitoring and continuous updating of data, central data collection mechanism, lack of stakeholders' support, submitting timely, authentic and accurate data.
    b. Campuses are located at different locations, untrained staff leading to extensive follow ups and delays in data submission, high number of stakeholders, inadequate funding to support IT infrastructure, changes in policy regulations was also a hindrance.
    c. Universities found some data challenging to provide to authorities like NAAC and NIRF:
        i. community outreach where it is difficult to measure the impact and outcomes
        ii. student progression (progression to post-doctoral positions for those completing PhDs and number of students who qualify for NET/SET/GATE/UPSC/DST-INSPIRE and other fellowships/scholarships) and no institutional mechanisms exist for capturing this data in HEIs,

iii. Information from corporate/private sector is also difficult as on-campus recruitment can be provided, but there is unwillingness of corporate houses to divulge details of pay-package, cost to company (CTC), and so on necessarily makes the data incomplete and unreliable
iv. Data related to Research Publications, quantifying publication/citation data on STEM (science, technology, engineering, mathematics) disciplines with quality of research outputs in humanities and social science disciplines, especially when it comes to things like h-indices, listing in SCOPUS/Web of Science is a challenge. The problem is even more acute for research/scholarly publications in Indian languages.

(B) Appropriateness of Data

a. Majority of University have responded the challenge for catering to the heterogeneity of data needed by different statutory and regulatory, ranking and accreditation agencies.
b. Some universities felt marginalized regarding scores for "national character" and "regional diversity" as they are often prevented by their Acts, statutes, and regulations, from admitting students from other regions/states

(C) Some suggestions/recommendations:

i. include some quantification measure to evaluate the scholarly publications in the non-STEM areas.
ii. to collect such data directly from the bureau/body/organization (NET Bureau; SET Bureau, UPSC, DST; etc) which conducts such tests and/or awards such fellowships, however there would no way to ensure the accuracy of data collected for (b) above.
iii. Input-output analysis in terms of per capita expenditure and outcomes in terms of students or faculty numbers by ranking agencies should be considered.
iv. More awareness & training to be provided to the people at ground level (data generators at respective institutes) about the importance and need for having accurate & reliable data generation sources.

**Conclusion**

The study brought out the current status of Data Management in their processes and practices across the HEIs in India. The limitation of the study was the poor response rates. Overall, the study brings out that the understanding that HEIs have claimed to have adequate data management processes and practices in place and is quite aware of its purpose.

We conclude that the study:

- acknowledges the need for data management in the processes and practices of the HEIs toward ensuring information flow efficiency for better decision making.
- reiterates the need for establishing robust information technology systems to augment data flow throughout the HEI's organizational structure and to promote a data-driven culture across HEIs

This study can be extended to assess the maturity of information technology adoption from an 'input-process-output' towards insights driven decision making perspective to sustain growth and competitiveness.

*Dr. Bhama Venkataramani*
*Dr. S. Vijayakumar Bharathi*
*Ms. Swati Sahasrabuddhe*
*Symbiosis International (Deemed University), Pune INDIA*