# EXPLORING BIG DATA AND ANALYTICAL TOOLS FOR EDUCATIONAL ECOSYSTEM

**Prof. SANDHYA THAKKAR**
Assistant Professor,
Department of Information Technology,
S.I.A. College of Higher Education,
Mumbai

**Mr. JOEL JOSEPH**
Software Developer
Department of Information Technology,
S.I.A. College of Higher Education,
Mumbai

## ABSTRACT

*"Big data"- as popularly it is called, is extremely large data set that analyses patterns, trends, and an association [1].From human resources and product development to customer service and sales and marketing, companies are increasingly looking to Big Data to identify patterns and make predictions that can help for breakthrough innovations.Today it is a business priority, given its ability to profoundly affect commerce in the globally integrated economy. In addition to providing solutions to long-standing business challenges, big data inspires new ways to transform processes, organizations, entire industries and even society itself.*

*Big Data is progressively being used by prominent companies to outpace the competition. Be it established companies or start-ups; they are embracing data-focused strategies to outpace the competition[2]. Most of the organizations are using big data to target customer-centric outcomes, tap into internal data and build a better information ecosystem.There are considerable opportunities for using Big data in higher education as well. It would be highly beneficial to universities if they are using Big Data to continue to deliver the very best learning environments for the welfare of the society. There are also possibilities of using Big data to link research to education – both by making better use of latest research practices and outcomes to reform teaching and in enabling research activities to be undertaken as part of education[3]. The current paper aims at exploring the Big data Analytical tools for education ecosystem.*

***Keywords:*** *Big Data, Educational Data Mining, Data warehouse, Data discovery, resource sharing, Analytical tools*

## INTRODUCTION

Past two decades have seen rapid change in all aspect of business ecosystem right from business model, planning, execution along with technologies to support them. Companies have realized that technology can only benefit the business. So along with adopting technology came other duties like integration, maintenance and updating with the advancement. Early adopter of technology was banks. Manual paper entries were replaced with DOS programs and for the first time saw usage of computers and completely replaced manual work. Such adoption can be seen in all sphere of Business Ecosystem today. Today growth of data creation is on a vertical climb and is forecasted to be the same for the future. The main reason for adoption can thus can be summarized to make this large data chunks generated into humane analysable data called "Big data".

Business Intelligence and Analytics (BI&A) and the related field of big data analytics have become increasingly important in both the academic and the business communities over the past two decades[4]. Education Ecosystem is not any different from a Business Ecosystem from a point of Complexity, huge data generation and analysis needed to stream line even

regular Administrative work. For this paper we have chosen some big data application which can help to implement Big Data in Higher Education Ecosystem.

## BIG DATA AND ITS RELEVANCE IN HIGHER EDUCATION

One of the many arenas where officials could benefit from big data deployment is higher education. With their large volumes of student information, including enrollment, academic and disciplinary records, universities have the datasets needed to benefit from a targeted analytics project.Recent reports have suggested that administrators across the United States have become to realize the potential of analytics and have invested more resource into big data projects. Gartner's "Big Data, Bigger Opportunities: Investing in Information and Analytics" survey found that 42 % of IT professionals across all sectors reported investing in big data projects or planning to do so in the coming year, reported Campus Technology. The report also identified higher education as field with great promise for big data application, citing the availability of various sources of data.

## MORE DATA BRINGS GREATER INSIGHTS

Once universities begin investing more in Hadoop big data projects, they benefit from its many applications. SmartData Collective contributor Mark van Rijmenam argued that with the advent of massive open online courses, higher education administrators have an entirely new pool of data to access, which could provide even greater insights into a college's operations. With these tools, school officials could enhance various aspects of campus life, including student success and academic performance.

One of the ways educators can utilize big data tools is to analyze the performance and skill level of individual students and create a personalized learning experience that meets their specific needs. MOOCs could be especially beneficial in this instance, as the heaps of structured data contained within their records could be easily gathered and processed by data analytics tools. With more sophisticated analytics software, professors could monitor many different factors regarding student performance beyond simple right or wrong answers, including the amount of time needed to answer questions as well as any connections regarding the types of test questions that were skipped over. Education IT researchers have been developing adaptive learning software that can process this information during an assessment or lesson and recommend further exams or coursework based on a student's performance.

According to Mark van Rijmenam, big data tools could even analyze group dynamics, looking at the various strengths and weaknesses of individual students to determine the optimal arrangement. Group projects would have a more fairly distributed workload as the individuals involved would compliment each other's skill sets.

Ultimately, by enhancing the learning experience and improving student performance across the board, universities will be able to reduce dropout rates and increase their graduation numbers. In addition to the purely academic benefits, colleges could also expect a greater and potentially more loyal alumni base that would be more generous with its donations. Hadoop big data software can provide higher education IT teams with tools to build analytics programs customized for their specific needs and goals [4].

**NOW WHY HADOOP?**

While large Web 2.0 companies such as Google and Facebook use Hadoop to store and manage their huge data sets, Hadoop has also proven valuable for many other more traditional enterprises based on its five big advantages.

**1. Scalable**

Hadoop is a highly scalable storage platform, because it can store and distribute very large data sets across hundreds of inexpensive servers that operate in parallel. Unlike traditional relational database systems (RDBMS) that can't scale to process large amounts of data, Hadoop enables businesses to run applications on thousands of nodes involving thousands of terabytes of data.

**2. Cost effective**

Hadoop also offers a cost effective storage solution for businesses' exploding data sets. The problem with traditional relational database management systems is that it is extremely cost prohibitive to scale to such a degree in order to process such massive volumes of data. In an effort to reduce costs, many companies in the past would have had to down-sample data and classify it based on certain assumptions as to which data was the most valuable. The raw data would be deleted, as it would be too cost-prohibitive to keep. While this approach may have worked in the short term, this meant that when business priorities changed, the complete raw data set was not available, as it was too expensive to store. Hadoop, on the other hand, is designed as a scale-out architecture that can affordably store all of a company's data for later use. The cost savings are staggering: instead of costing thousands to tens of thousands of pounds per terabyte, Hadoop offers computing and storage capabilities for hundreds of pounds per terabyte.

**3. Flexible**

Hadoop enables businesses to easily access new data sources and tap into different types of data (both structured and unstructured) to generate value from that data. This means businesses can use Hadoop to derive valuable business insights from data sources such as social media, email conversations or clickstream data. In addition, Hadoop can be used for a wide variety of purposes, such as log processing, recommendation systems, data warehousing, market campaign analysis and fraud detection.

**4. Fast**

Hadoop's unique storage method is based on a distributed file system that basically 'maps' data wherever it is located on a cluster. The tools for data processing are often on the same servers where the data is located, resulting in much faster data processing. If you're dealing with large volumes of unstructured data, Hadoop is able to efficiently process terabytes of data in just minutes, and petabytes in hours.

**5. Resilient to failure**

A key advantage of using Hadoop is its fault tolerance. When data is sent to an individual node, that data is also replicated to other nodes in the cluster, which means that in the event of failure, there is another copy available for use.The MapR distribution goes beyond that by eliminating the NameNode and replacing it with a distributed No NameNode architecture that provides true high availability. Our architecture provides protection from both single and multiple failures.When it comes to handling large data sets in a safe and cost-effective manner, Hadoop has the advantage over relational database management systems, and its value for any size business will continue to increase as unstructured data continues to grow.

**6. Java**

Why not to use something that is something free, used worldwide and reliable with tag line "Code once, Run Anywhere" Java technologies like Hibernate, JSP can be used to access

front end input from multiple clusters. A recommended method to implement in our University or such large ecosystem [5]

## BIG DATA ANALYTICAL APPLICATIONSFOR HIGHER EDUCATION

### CourseSmart

CourseSmart Analytics is an innovative new solution that measures student engagement with digital course materials in order to provide administrators and faculty with actionable insights to help drive retention and outcomes. A centrepiece of CourseSmart Analytics is the CourseSmart Engagement Index, a proprietary algorithm that evaluates student usage data points and assimilates them into an overall assessment of students' engagement with digital material. Used in conjunction with other key indicators such as class participation, homework completion, and exam and other assessment results, CourseSmart Analytics can help improve student retention and outcomes, and ultimately, their success in your course or program.

### LabStats

LabStats is the easiest and most economical software solution for tracking computer usage in colleges and universities. Designed specifically for networked computer environments, LabStats provides a simple and customizable interface to manage every aspect of your computer labs. Save money, back up funding requests with real data and take control of your lab with LabStats. LabStats gathers usage information for every lab machine and stores it indefinitely in a high-efficiency database. LabStats lets lab managers view real-time usage charts in a web browser, rather than relying on lab assistants, manual headcounts and other homegrown solutions.

### Civitas Learning

Our platform enables us to ingest data from learning management systems, student information systems, and other sources, then build school-specific predictive models that discover hidden connections and identify key decision points that affect student success. Our products deliver current, personalized, data-informed advice to the front lines of higher education. Educational Ecosystem can interact with these clean, user-friendly apps, which distill deep analytical insights into straightforward, actionable guidance. By helping to improve a million small decisions across a college or university, they enable big improvements across the institution.

### Declara

At the core of the Declara platform is The CognitiveGraph a dynamic map of how people learn based on years of research across hundreds of subjects. The CognitiveGraphtakes every interaction an individual has over the network — be it access to a blog post, a connection to an expert, or a posting of "like"to deliver the right content and the right learning partners to the right people at just the right time to radically improve learning outcomes. Using advanced algorithms and machine learning, the Declara CognitiveGraph learns as the user learns.

### Knewton

Knewton is an education technology company that personalizes digital courses so every student is engaged and no one slips through the cracks. As students progress through a Knewton-powered course, Knewton figures out what each student knows and how that student learns best, then recommends what to study next. Teachers use Knewton-powered predictive analytics to detect gaps in knowledge and differentiate instruction for each student.

### Noel-Levitz

ForecastPlus for Recruitment is a highly advanced predictive modeling tool for enrollment management. Powered by customized, multi-variable analyses, it uses your past enrollment data to predict your future enrollment results. Proven effective on more than 100 campuses, ForecastPlus reduces uncertainty and duplication from the admissions process.

**Blackboard Analytics**

Blackboard Analytics for Learn combines the extensive data from Blackboard Learnwith student and course attributes from your SIS/ERP to create comprehensive reports and dashboards for your students, instructors, staff and leadership. With Blackboard Analytics for Learn, you have easy, self-service access to data that can help give you an enterprise-level perspective. Gain insight into user activity, course design and student grade and learning outcome data across departments and colleges, enabling you to improve your use of the Blackboard Learn platform in support of teaching, learning, and student success.

**Desire2Learn Insights**

Desire2Learn Insightsoffers a powerful suite of high-performance reports, rich data visualizations and predictive analytic capabilities that enable institutions to discover rich and meaningful insight about the integrity of their learning environment and deliver a perceptive, next generation experience. With the ability to harness the big data available in your learning environment, Insights' dynamic, outcomes-based reports enable key stakeholders to establish, track, measure and assess achievement across course, program or institution-level. While your focus for improvement might be accreditation support, program design and delivery, or mentoring at-risk learners – whatever the need – Desire2Learn Insights expertly transforms your learning environment data into actionable insights that guide new opportunities for learner and institutional success.

**IBM**

A complex world demands skilled and knowledgeable citizens, yet educational institutions at every level struggle to graduate students who can meet these requirements. As costs rise and funding shrinks, educators and administrators need deeper insight into which approaches work and which do not. Analytics provides them with the tools to measure performance and ensure students acquire the skills to succeed.

**PAR Framework by WCET**

The Predictive Analytics Reporting (PAR) Framework is a multi-institutional data mining collaborative that brings together 2 year, 4 year, public, proprietary, traditional, and progressive institutions to collaborate on identifying points of student loss and to find effective practices that improve student retention in U.S. higher education. With sixteen member institutions, over 1,700,000 anonymized student records and 8,100,000 institutionally de-identified course level records, the PAR Framework offers educational stakeholders a unique multi-institutional lens for examining dimensions of student success from both unified and contextual perspectives.

**SAP**

Graduate to the next level of performancewith best-in-class software to help your higher education or research institution run better. No matter the size of your organization, we can help boost your international ranking and reputation through superior research, improved student performance, and smarter financial management.

**iDashboards**

iDashboards Higher Education dashboard software is providing colleges and universities around the country better access to their data and greater insight into their key performance indicators (KPIs). Dashboards are being utilized to monitor Higher Education KPIs such as enrollment, accreditation, effectiveness, institutional research, financials and departmental metrics. Education Dashboards are customized for each institution taking into account their data sources and the goals of the institution.

**McGraw-Hill Connect**

The first and only analytics tool of its kind, Connect Insight is a series of visual data displays—each framed by an intuitive question—to provide at-a-glance information

regarding how your class is doing. Initially Connect Insightwill provide analysis on five key insights, available at a moment's notice from your tablet device.

**eBureau**

In the increasingly competitive market for online higher education, schools must cut through the noise to find and engage with qualified prospects. Higher education marketers must drive profitable new student acquisition through cost-effective investments in online advertising, including cost-per-lead and display advertising. Using eBureau's advanced predictive analytics at the moment of initial inquiry, marketers can: Lower cost-per-start. Meet and surpass enrollment goals. Boost conversion rates and improve customer service. Understand which leads are likely to apply, enroll, and stay enrolled. After enrollment, eBureau's big data and predictive analytics tools can help schools: Identify retention risk. Locate past students and alumni.

**Tableau Software**

Educational institutions have data everywhere: from student and donor data to classes, faculty and grants. Yale University, Cornell University, the University of Oxford and hundreds of other institutions use Tableau for their institutional research initiatives: Analyze student data: enrollment statistics, achievement, and demographics. Target areas of the country for alumni activities and development. Prepare Integrated Postsecondary Education Data System (IPEDS) and other reporting.

**HCONN**

Enrollment Analytics for Colleges and Universities is an analytical platform for your institution to increase educational performance. With access to relevant data from multiple sources, your organization can make better decisions by tracking progress against key performance indicators (KPIs).The Higher Education Analytics tools enables your organization to analyze and identify higher performance factors on test scores, analyze and discover attendance patterns, insight on behaviour patterns in relation to student performance, and gain a 360-degree view of teacher and student demographics. Colleges and Universities need to gain comprehensive and in-depth insight on enrolled students including experience, high school performance, academic programs, courses, locations, demographic insights, and external factors on enrolments.[6].

## CONCLUSION

In world there are of course many Big Data Tools available but most of them are suitable for a research oriented ecosystem where cost effectiveness play a vital role to make it more meaningful. The above mentioned some of the open source software can be much helpful. Even by using Apache's Hadoop, R-Language, Java technologies cost can be brought down greatly so that it can be used at Education Ecosystem and more over students can work on it themselves without relying on proprietary tools like SAP.

## REFERENCES

[1] http://www.uazone.org
[2] How Big Data is changing the business landscape By IANS | 2 Jan, 2016, 09.54AM IST
[3] http://ictactjournals.in
[4] http://www.bigdatalandscape.com/blog/big-data-higher-education-apps-change-everything
[5] http://hortonworks.com
[6] http://hortonworks.com