

CONTENTS

Chairman's Message

Principal's Message

HOD's Message

About The Department

The Memorandums

Quiz

About Big Data

Placement & Achievements

ABOUT OUR DEPARTMENT

We produce quality engineers with the knowledge of latest technologies and industrial trends and knowledge to meet the developing needs of the growing industries and society.



Nishant Tripathi
Chairman, SSIPMT, Raipur

SSIPMT is proud not only for being an inspiration to every individual who belongs to this institute, but also for being a motivation for students to follow their dream path. I'm proud to be in a position where at times I am a mentor for those who need it, a much needed positive authority figure to help guide the way. I am proud to be on this journey with my fellow teachers, students and every member of this institute who have made this institute what it is today. *As Paulo Coelho said, "Dreams are meant to be followed, life is meant to be lived".*



Dr. Alok Jain
Principal, SSIPMT, Raipur

I feel immensely proud to announce that department of computer science and engineering has manifested its introductory issue of the college newsletter "Cyber Trinity". The documentation would definitely outline the curriculum conducted and day-to-day development made by the department and incorporate more information about the latest technologies. At our institute, we aim to create quality engineers who are capable to contribute their best towards the technological advancement of our country. We aim to achieve the four pillars of India i.e. Human development, social development, economic development and environmental development.



Dr. J.P. Patra
HOD (CSE), SSIPMT, Raipur

I am happy to learn that the department of Computer Science and Engineering is bringing out the BIG DATA Newsletter of the department, which truly exhibits the talent of our budding computer science engineers. CSE department is continuously gifted with the best of the brains whom we believe that they are our future "Gamechangers". They have displayed their strengths through extra-ordinary achievements in academics, cultural, sports, etc. I am sure that the hard work and sincere efforts of the HOD and all staff of the department have kept their energies of our students properly channelled to deliver the best.

OUR FACULTY MEMBERS



Mr. Anand Tamrakar
Asst. Professor



Mr. Gurudatta Verma
Asst. Professor



Mr. Riju Bhattacharya
Asst. Professor



Mrs. Sumitra Samal
Asst. Professor



Mrs. Preeti Tuli
Asst. Professor



Mr. Gurpreet Singh Chhabra
Asst. Professor



Mr. Rajesh Deshmukh
Associate Professor



Mr. Tansen Patel
Asst. Professor



Mr. Siddharth Shukla
Asst. Professor



Mr. Teshu Gaurav Singh
Asst. Professor

MEMORANDUMS

EDITORIAL BOARD

Editor in Chief:
Asst. Prof. Taniya Jain

Designing & Layout:
Prerika Sanghvi
Pranjal Chawda
CSE 6th Sem.

Editorial Members:
Rashmi Patel
Shubham Khanna
Sagar Agarwal
Nikita Matlani
CSE 6th Sem.



Today, the data which is in petabytes will surely cross zetabytes in 2020. Hadoop is the technology that will efficiently collect, handle, store, transfer, process and analyse this data. And to handle such a big data for a traditional DBMS is difficult. That is why, HDFS, Map reduce, Yarn came into picture. This is beyond any doubt that in near future, big business houses will surely invest millions of bucks in hadoop and Big data technology. The reason is obvious - DIGITALIZATION.



I am very proud that our future engineers are learning more and more towards new technology which is Big Data. Big data helps the organizations to lead to better decisions and strategic business moves. Big Data and the way the organizations manage and derive insights from it is changing the way the world uses business information. Even the Indian government has always shown a keen interest in adopting new technologies, especially Big Data.

QUIZ

- 1** Distributed cache can cache files
a. Jar Files
b. read-only text files
c. archives
d. All of the above
- 2** Which of the following Hadoop config. files is used to define the heap size?
a. hdfs-site.xml
b. core-site.xml
c. hadoop-env.sh
d. mapred-site.xml
- 3** Which of the following method used to set the output directory
a. FileOutputStream.setOutputgetpath()
b. OutputFormat.setOutputpath()
c. FileOutputStream.setOutputpath()
d. OutputFormat.setOutputgetpath()
- 4** The total number of partitioner is equal to
a. The number of reducer
b. The number of mapper
c. The number of combiner
- 5** Which tool is used to distributes data evenly across datanode
a. Balancer
b. Disk Balancer

By Sagar Agarwal



BIG DATA

What is Big Data?

Big data is a term used to refer to data sets that are too large or complex for traditional data-processing application software to adequately deal with. Data with many cases (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate. Big data challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy and data source. Big data was originally associated with three key concepts: volume, variety, and velocity. Other concepts later attributed with big data are veracity (i.e., how much noise is in the data) and value. Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time. Big data philosophy encompasses unstructured, semi-structured and structured data, however the main focus is on unstructured data. Big data "size" is a constantly moving target, as of 2012 ranging from a few dozen terabytes to many exabytes of data. Big data requires a set of techniques and technologies with new forms of integration to reveal insights from datasets that are diverse, complex, and of a massive scale. A 2016 definition states that "Big data represents the information assets characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value".

Characteristics of Big Data

Volume

The quantity of generated and stored data. The size of the data determines the value and potential insight, and whether it can be considered big data or not.

Variety

The type and nature of the data. This helps people who analyze it to effectively use the resulting insight. Big data draws from text, images, audio, video; plus it completes missing pieces through data fusion.

Velocity

In this context, the speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development. Big data is often available in real-time. Compared to small data, big data are produced more continually. Two kinds of velocity related to Big Data are the frequency of generation and the frequency of handling, recording, and publishing.

Veracity

It is the extended definition for big data, which refers to the data quality and the data value. The data quality of captured data can vary greatly, affecting the accurate analysis.



Technologies Used

Government

The use and adoption of big data within governmental processes allows efficiencies in terms of cost, productivity, and innovation. CRVS (Civil Registration and Vital Statistics) collects all certificates status from birth to death. CRVS is a source of big data for governments.

International development

Research on the effective usage of information and communication technologies for development (also known as ICT4D) suggests that big data technology can make important contributions but also present unique challenges to International development. Advancements in big data analysis offer cost-effective opportunities to improve decision-making in critical development areas such as health care, employment, economic productivity, crime, security, and natural disaster and resource management.

Manufacturing

Based on TCS 2013 Global Trend Study, improvements in supply planning and product quality provide the greatest benefit of big data for manufacturing. Big data provides an infrastructure for transparency in manufacturing industry, which is the ability to unravel uncertainties such as inconsistent component performance and availability. Predictive manufacturing as an applicable approach toward near-zero downtime and transparency requires vast amount of data and advanced prediction tools for a systematic process of data into useful information.

Healthcare

Big data analytics has helped healthcare improve by providing personalized medicine and prescriptive analytics, clinical risk intervention and predictive analytics, waste and care variability reduction, automated external and internal reporting of patient data, standardized medical terms and patient registries and fragmented point solutions.

Insurance

Health insurance providers are collecting data on social "determinants of health" such as food and TV consumption, marital status, clothing size and purchasing habits, from which they make predictions on health costs, in order to spot health issues in their clients.

BIG DATA



Big Data and Education

Big Data has become an integral part of almost every sector, and therefore the education sector can also no longer afford to remain insulated from the technology connect. A lot of education hubs have started using newer technological tools to improve pedagogy. Today, an Indian student has access to data and resources available anywhere in the world, which was impossible a couple of decades ago. Big data has been driving revolutionary changes in education. There hardly remains an area in education not impacted by big data. The changes are evident in the ways educational institutions are governed, course quality is managed and student participation and performance are managed. Compared to the education system in the past, the changes represent a radical shift in paradigm. Data is at the core of the revolution. Analytics is helping educational institutions to better manage course quality, student performance and behavior, and overall administration.

By Prerika Sanghvi (CSE - 6th Sem.)
Rashmi Patel

ACTIVITIES AND PLACEMENTS

EVENTS

A number of events are organized by our society and the whole Computer Science & Engineering department comes together to make every activity a successful one. We have conducted various competitions such as Coding, Football, PC games, Mobile games, Athletics, etc. in which students from any branch can participate and workshops on IOT, Cloud Computing and BIGDATA. We have successful company campus drives and wherever our students go they hoist their flags high.



CONGRATULATIONS FOR SUCCESS - OUR PLACEMENTS 2018!



Alfiya Anjum
(Nucleus, GenPact, Techment Tech.)



Anubha Datey
(Capgemini, GenPact)



Astha Batra
(TCS, Capgemini, Yardi, HSBC)



Khushboo Krishnani
(Capgemini, GenPact, Yardi, ClickLab)



Ayushi Agrawal
(Capgemini, GenPact, Techment Tech.)



Saba Amrani
(Capgemini, GenPact, Techment Tech.)



Sonam Naidu
(Capgemini, GenPact, Collabera)



Arshiya Fatima
(Capgemini, Collabera)



Deepika Dawani
(GenPact, Collabera)



Rahul Gupta
(GenPact, Capgemini)



Rishabh Sonkar
(TCS, Nucleus)



Sanskar Sharma
(TCS, Capgemini)



Shweta Parmar
(Nucleus, Capgemini)



Snigdha Chowdhary
(Nucleus, Jain Softwares)



Aman Yerpude
(Jain Softwares)



Anurag Chandra
(Capgemini)



Astha Tiwari
(Capgemini)



Isha Gawhade
(Capgemini)



Linesh Gohil
(Capgemini)



Mayank Khakharia
(Capgemini)



Payal Agrawal
(Nucleus)



Prerna Sahu
(Capgemini)



Sidak Arora
(GenPact)



Yashashvi Vaidya
(GenPact)

ACHIEVEMENTS

NAME	BRANCH	EVENT	MEDAL/POSITION
Ayush Pandey	CSE-3 rd	Volleyball	National Level
Shubhanshu Markaam	CSE-3 rd	Football	Winner
Syed Tufail Ahmed	CSE-5 th	Football	Winner
Vaishali Pawar	CSE-3 rd	Rongoli	Winner
Darakshan Nazreen	CSE-5 th	Athletics-100M	Silver Medal (2 nd)
Darakshan Nazreen	CSE-5 th	Athletics-400M	Silver Medal (2 nd)
Ujjawala Katakwar	CSE-5 th	Table Tennis	Winner
Sanya Soni	CSE-5 th	Table Tennis	Winner
Varkha Shivhare	CSE-5 th	Basketball	Winner
Aashi Sharma	CSE-3 rd	Basketball	Winner
Ayush Pandey	CSE-3 rd	Volleyball	Winner
Akshat Parihar	CSE-3 rd	Volleyball	Winner

NAME	BRANCH	EVENT	MEDAL/POSITION
Darakshan Nazreen	CSE-5 th	Athletics-400m	Gold Medal (1 st)
Garima Chandrakar	CSE-5 th	Athletics-TripleJump	Gold Medal (1 st)
Priyanka Kashyap	CSE-5 th	Athletics-200m	Gold Medal (1 st)
Priyanka Kashyap	CSE-5 th	Athletics-LongJump	Gold Medal (1 st)
Darakshan Nazreen	CSE-5 th	Athletics-4X100m	Gold Medal (1 st)
Garima Chandrakar	CSE-5 th	Athletics-4X100m	Gold Medal (1 st)
Priyanka Kashyap	CSE-5 th	Athletics-4X100m	Gold Medal (1 st)
Priyanka Sahu	CSE-3 rd	Athletics-4X100m	Gold Medal (1 st)
Akshat Parihar	CSE-3 rd	Athletics-4X100m	Gold Medal (1 st)
Shubhanshu	CSE-3 rd	Athletics-4X100m	Gold Medal (1 st)
Varkha Shivhare	CSE-5 th	Handball	Runner
Nilima Thakur	CSE-5 th	Athletics-4X100m	Runner