



Process mining: Business Intelligence (BI) and Business Process Management (BPM).

¹S.Sowjanya Chintalapati, ²Dr.Ch.G.V.N.Prasad, ³R.Vineela, ⁴J.Sowjanya

^{1,3,4} Assistant Professor, Dept of CSE, SRI INDU College of Engg and Technology, Ibrahimpatnam, Hyderabad, India

²Professor, Head of the Dept, Dept of CSE, SRI INDU College of Engg and Technology, Ibrahimpatnam, Hyderabad, India.

Abstract: *The increasing integration of information systems for the operation of business processes provide the basis for innovative data analysis approaches. Business processes become more and more complex and information systems support or even automate the execution of business transactions in modern companies. Business intelligence aims to support and improve decision making processes by providing methods and tools for analyzing data. BI solutions faced with the choice of purchasing a very expensive consolidated solution from a large vendor rather than acquiring best of breed tools. Business process management (BPM) enables organizations to align internal business functions with customer needs, and helps executive determine how to direct, monitor and measure company resources. When properly executed by qualified professionals, business process management has the capacity to reduce costs, enhance efficiency and productivity, and minimize errors and risk – thereby protecting and optimizing corporate resources.*

Keywords: *OLAP, Business Intelligence, performance analysis: event log, ERP, BPM, Process Mining..*

I. INTRODUCTION:

Companies use information systems to enhance the processing of their business transactions. Enterprise resource planning (ERP) and workflow management systems (WFMS) are the predominant information system types that are used to support and automate the execution of business processes. Business processes like procurement, operations, logistics, sales and human resources can hardly be imagined without the integration of information systems that support and monitor relevant activities in modern companies.

The increasing integration of information systems does not only provide the means to increase effectiveness and efficiency. It also opens up new possibilities of data access and analysis. When information systems are used for supporting and automating the processing of business transactions they generate data. This data can for example be used for improving business decisions. Technology has become an integral part of any organization. For example, current systems and installations are heavily controlled and monitored remotely by integrated internet technologies. Moreover, employing automated solutions in any line-of-business has become a trend.

The application of techniques and tools for generating information from digital data is called business intelligence (BI). That is, BI applications have been installed to support management in measuring company's performance and deriving appropriate decisions [3]. Among most important functions of BI are online analytical processing (OLAP), data mining, business performance management and predictive analytics.

II. BUSINESS INTELLIGENCE:

In order to evaluate relevant data at the click of a button, we need to create a customized business intelligence (BI) solution. The Block Diagram or simply the formation of a Business Intelligence is as shown in the figure i) [21]

i) Operational Data/Source Data:

Organizations typically store the extracted and transformed data in a separate database distinct and separate from the production systems - this insulates the production systems from any "run away" queries that could be generated by the users, and it allows the organizations to control updates to the data - since there is often a desire to ensure that everyone reports from the same time series!

ii) Extract-Transform-Load (ETL):

Extraction-Transformation-Loading is the process and tools whereby data is extracted from source or transactional systems. Typically, data is extracted and changed or transformed prior to loading into the reporting database (more commonly referred to as a Data Warehouse or Data Mart). The transformation is undertaken to accomplish three things: [25]a) Provide data in a consistent format. For example, the same data may be represented differently depending upon the source system. Gender may

be represented as Male/Female in one system, M/F in another, and even 1/2 in a third system! The ETL process supports the transformation of data to a consistent format.

b) Transform data to a format easily understood by the business. Transactional systems are designed to process information as quickly as possible. As a result, the Relational Data Models employed are often very complex and difficult to understand even for seasoned programmers. A Dimensional Data Model or Star Schema model is typically employed in order to represent data in a way that is easily understood by the business and that expedites database access.

c) Data may be enriched or corrected when transformed. For example, integration of multiple customer records to a single customer record may be undertaken, and rules may be applied to ensure that the data is correct. Customer records may be segregated by region or area based on the zip or postal code. This correction step more and more is being eschewed in favor of correcting data at the source level, as the corrective activities themselves can build in incorrect assumptions regarding the data.[2]

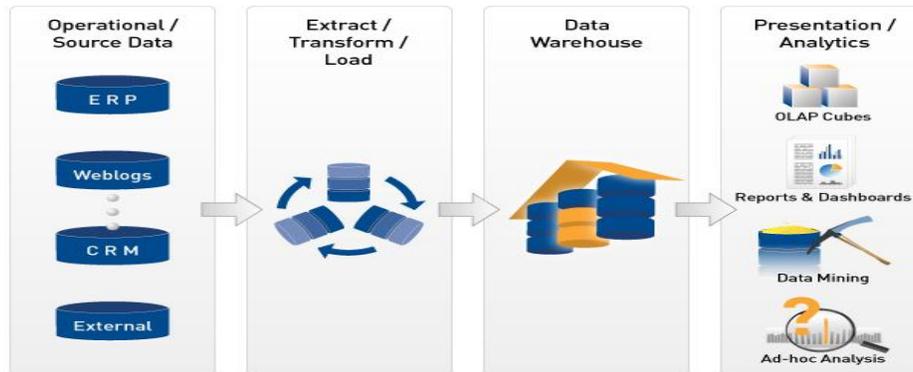


Fig i) Business Intelligence (BI)

iii) Reporting / Analysis:

There are a plethora of tools available for the surfacing and analysis of data. Most organizations will utilize multiple tools as there isn't a "one size fits all" BI tool. I often segregate their users into categories:

a) Consumers typically use delivered reports and/or dashboards (summarized information presented on a web page). If further analysis is required, they will have another member of the organization undertake this work. These are often senior executives within an organization who don't have the cycles to do the analysis and, more importantly, have individuals who can undertake the analysis.

b) Analysts will often utilize tools that allow a structured review of the data, such as configurable reports and queries or the use of data cubes, often referred to as OLAP or Online Analytical Processing, to slice and dice the data and to drill-through to details.

c) Explorers undertake detailed analysis of data. They will often use ad-hoc reporting tools or may program their own queries in SQL. Additionally, some users may utilize Data Mining tools to undertake advanced statistical analysis of data to drive out patterns and trends.

III. BUSINESS PROCESS MANAGEMENT (BPM):

At strategic level there is a need to deliver a technology platform for the business community to "Manage by Business Processes and not just by Business Transactions". To this effect, an innovative blend of Business Process Management (BPM) and Service-Oriented Architecture (SOA) needs to be employed to deliver a network of distributed business services (as engineered capabilities) that can be linked together dynamically to deploy an end-to-end visible business, thereby attracting early benefits out of emerging market opportunities.[22]

Unfortunately, these models are often completely disconnected from actual event data. Analysis results are unreliable because they are not based on observed facts, but on an idealized model of reality [23].

i) What really a Business Process Management: The BPM acronym is used loosely and its meaning often varies depending upon the context. Software companies often refer to BPM to describe the capabilities of their particular technology, while practitioners, management consultants and academics typically discuss a management discipline in defining BPM. Therefore BPM can be defined in any of the following ways.[24]

- A process of managing the business processes.
- A management discipline.
- A technology or set of technologies.
- A rapid application development framework.

ii) A process and management discipline:

BPM is a process and a management discipline. Enabling technology is meaningless without the processes and management disciplines for exploiting the technology - in this case the tools for managing an organization's business processes. BPM involves managing the end-to-end work that organizations perform to create value for their customers. The

performance of this work is how organizations fulfill their mission. An entire industry has emerged around BPM as organizations have recognized the importance of actively managing their business processes.

Many organizations have also recognized a gap in the skills necessary to be effective at BPM. Organizations such as the BPM Institute and the Association of Business Process Management Professionals (ABPMP) have been formed to help address this gap. [24] By identifying and sharing best practices while providing training and education opportunities, these and other organizations aid in the development of the necessary skills, knowledge and management practices for effective BPM. The ABPMP expects to launch a comprehensive body of knowledge and a professional certification for BPM Professionals in 2007. The ABPMP definition of Business Process Management is: "Business Process Management (BPM) is a disciplined approach to identify, design, execute, document, monitor, control, and measure both automated and non-automated business processes to achieve consistent, targeted results consistent with an organization's strategic goals. BPM involves the deliberate, collaborative and increasingly technology-aided definition, improvement, innovation, and management of end-to-end business processes that drive business results, create value, and enable an organization to meet its business objectives with more agility."

iii) Business Process Management workflows:

Many vendors have created application suites which enable organizations to better manage their business processes. These technologies typically involve tools to visually design and model business processes; simulate and test business processes; automate, control and measure business processes; and provide feedback and reporting on process performance. Some vendors have combined these functions into business process management suites that provide a complete integrated BPM platform, commonly referred to as a BPMS. Diagrammatical representation of Business Process Management Workflows is as shown in the figure below :(Fig ii)

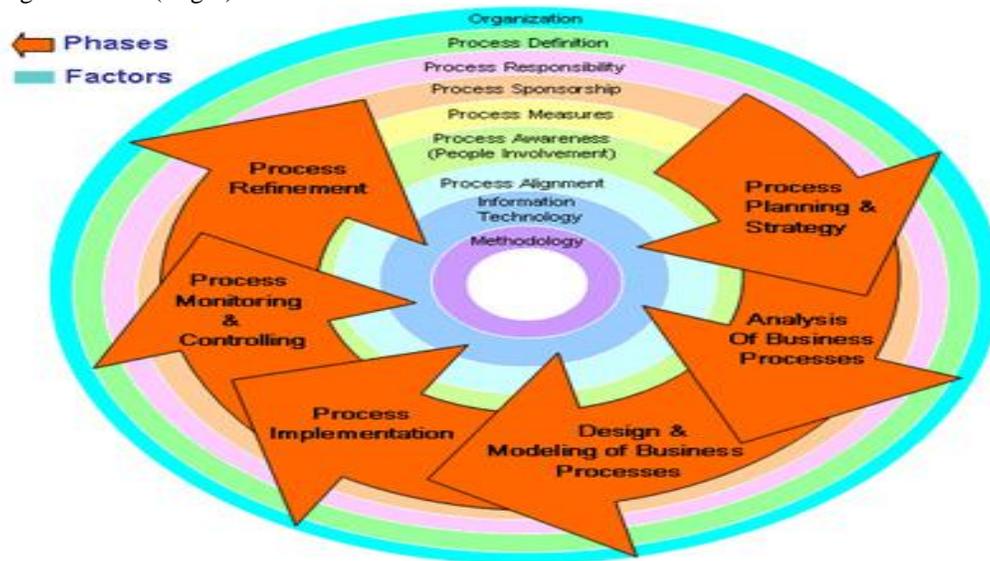


Fig ii) Business Process Management Workflows

Many organizations have a large number of legacy systems, typically designed to support specific functions such as manufacturing or sales. In order to manage the end-to-end work involved in business processes, a BPMS must be able to integrate with legacy systems across the organization in order to control work, get information or measure performance. A variety of new technologies have emerged to simplify integration efforts and the technology industry appears to be standardizing on a specific set of open technologies, commonly referred to as Web Services.

A common framework for how these technologies are deployed is also being adopted, most often referred to as a Service Oriented Architecture (SOA). By leveraging web services in a service oriented architecture construct, organizations can build and manage end-to-end business processes across organizational silos and their legacy systems. Many modern BPM technology solutions include the capability to interface to legacy systems through these standard interfaces, providing the tools to automate and orchestrate work across the entire organization.

iv) A rapid application development platform:

As BPM technologies were released into the market, many IT organizations have begun to recognize that the technology can be leveraged to develop applications faster and at a lower cost than traditional methods. The visual design capabilities and standards-based interfaces create an environment where skilled software engineers can rapidly define the behavior of software which traditionally required significant effort to develop in code.

A BPMS typically comes with pre-packaged functionality, such as user interface and security frameworks, is vendor supported and is easier to support than custom code. While many systems do not address end-to-end business processes, developers have found that a BPMS can be leveraged to rapidly and inexpensively develop a wide variety of software

solutions for an organization's business users. While the BPMS market is consolidating onto a common platform capability, my observation is that BPMS vendors appear to be evolving from one of three areas:

- i) Vendors who have developed business process management capabilities which wrap around traditional content management technologies. These solutions are typically best deployed in document or content-centric processes, such as case management, claims management, licensing or permitting.
- ii) Vendors who have developed business process management capabilities which wrap around traditional middleware technologies, such as connectivity, messaging and transformation services. These solutions are typically best deployed for processes which will rely heavily on orchestration and integration of legacy systems and technologies.
- iii) So called "pure play" BPM vendors. These are typically new companies which have developed applications and application development frameworks specifically to focus on BPM capabilities. Many of these companies are being acquired by larger vendors, reducing the time to market for building a BPM capability by integrating these new technologies with legacy products.

Several of these vendors, however, appear to be surviving as standalone companies in part due to the fact that they are perceived to be "best of breed" in pure BPM capabilities. These solutions typically provide the most integrated and robust platform for rapid application development.

IV. AIM OF PROCESS MINING:

Data mining is the analysis of data for finding relationships and patterns. The patterns are an abstraction of the analyzed data. Abstraction reduces complexity and makes information available for the recipient. The aim of process mining is the extraction of information about business processes [4]

Process mining encompasses "techniques, tools and methods to discover, monitor and improve real processes by extracting knowledge from event logs [6]. The data that is generated during the execution of business processes in information systems is used for reconstructing process models. These models are useful for analyzing and optimizing processes.

Process mining is an innovative approach and builds a bridge between data mining and business process management. Process mining evolved in the context of analyzing software engineering processes by Cook and Wolf in the late 1990s (Cook and Wolf 1998). Agrawal and Gunopulos (Agrawal et al. 1998) and Herbst and Karagiannis (Herbst and Karagiannis 1998) [19] introduced process mining to the context of workflow management.

Major contributions to the field have been added during the last decade by van der Aalst and other research colleagues by developing mature mining algorithms and addressing a variety of topic related challenges [6]. This has led to a well developed set of methods and tools that are available for scientists and practitioners.

V. CHALLENGES IN PROCESS MINING:

In the context of today's rapidly changing environment, organizations are looking for new solutions to keep their businesses running efficiently. Slogans such as "Driving the Change" (Renault), "Changes for the Better" (Mitsubishi Semiconductor), "Empowering Change" (Credit Suisse First Boston), "New Thinking. New Possibilities" (Hyundai) are used more and more often. Furthermore, different areas of business research are trying to keep up with the change and process mining is not an exception. In 2011, the Process Mining Manifesto [7] was released to describe the state-of-the-art in process mining on one hand, and its current challenges, on the other hand. A year later, the project proposal "Mining Process Cubes from Event Data (PROCUBE)" in [4] suggested the so called process cube as a solution direction for some of these challenges. In the context of currently employed process mining solutions and using the Process Mining Manifesto as a reference, the PROCUBE project proposal presents several challenges that process mining is currently facing:

a) From small event data to big event data. Due to increased storage capacity and advanced technologies, the vast amount of available event data have become difficult to control and analyse. Most of the traditional process mining techniques operate with event logs whose size does not exceed several thousands cases and a couple hundred thousands events (for example, in BPI Challenges [2]).

However, nowadays corporations work on a different scale of event logs. Giants like Royal Dutch Shell, Walmart, IBM, would rather consider millions of events (a day or even a second) and this number will continue to grow. Ways to ensure that event data growth will not affect the importance of process mining techniques are constantly sought.

b) From homogeneous to heterogeneous processes: With the increasing complexity of an event log, chances are that the variability in its corresponding process increases as well. For example, events in an event log can present different levels of abstraction.

However many mining techniques assume that all events in an event log are logged at the same level of abstraction. In that sense, the diverse event log characteristics have to be properly considered.

c) From one to many processes: Many companies have their agencies spread across the globe. Let's take SAP AG as an example. Only its research and development units are located on four continents, but it has regional offices all around the world.

That is, SAP units are executing basically the same set of processes. Still, this does not exclude possible variations. For instance, there might be various in sequences due to the characteristics of a certain SAP distribution region (Germany, India, Brazil, Israel, Canada, China, and others).

Traditional process mining is oriented on stand-alone business processes. However, it is of great importance to be able to compare business processes of different organizations (units of an organization). For example, efficient and less efficient paths in different processes can be identified. Inefficient paths can be substituted and efficient paths can be applied to the rest of the processes to improve performance.

d) From steady-state to transient behavior: The change has a major impact not only on the size of event logs and on the necessity of dealing with many processes together, but also on the state of a business process. For example, companies should be able to quickly adjust to different business requirements.

As a result, their corresponding processes undergo different modifications. Current process mining techniques assume business processes to be in a steady-state [5]. However, it is important to understand the changing nature of a process and to react appropriately. The notion of concept drift was introduced in process mining [33] to capture this second-order dynamics.

Its target is to discover and analyze the dynamics of a process by detecting and adapting to change patterns in the ongoing work.

e) From online to online: As previously mentioned, systems produce an overwhelming amount of information. The idea of storing it as historical event data for later analysis, as it is currently done, may not seem as appealing any more. Instead, the emphasis should be more on the present and the future of an event. That is, an event should be analysed on-the-y and predictions on the contingency of its occurrence should be made based on existing historical data. As such, online analysis of event data is yet another process mining challenge.

Each of the issues discussed above, are extremely challenging. Analysing large scale event logs is difficult with the current process mining techniques. Solutions to mitigate some of the issues that appear when dealing with large scale event logs are proposed in [14], i.e., by event log simplification, by dealing with less-structured processes and others.

A framework for time-based operational support is described in [8]. In [16], an approach is offered to compare collections of process models corresponding to different Dutch municipalities. Nevertheless, there is still the need for more elaborated solutions and a unified way of approaching them.

VI. PERCEPTIVE PROCESS MINING:

The relationship between people, workflows, and the business applications ie already in use is dynamic landscape, and — in complex systems — locating the causes of process bottlenecks and breakdowns is often more a matter of luck than skill. All business applications generate large quantities of time and event data about your business processes. The record of actions people take while using such applications, and even the automated steps in your current systems, can produce potentially valuable insights when combined and viewed as a whole.

But trying to discover, monitor, and improve processes in a variety of application domains takes more than a clipboard and stopwatch. With business conditions constantly changing, you need a tool that can quickly adapt and comprehensively extract useful knowledge about the history — and future potential — of all your processes.

Perceptive Process Mining goes far beyond traditional process analysis methods like interviews, lead time measurement, and performance analysis outside the context of the process or organization.

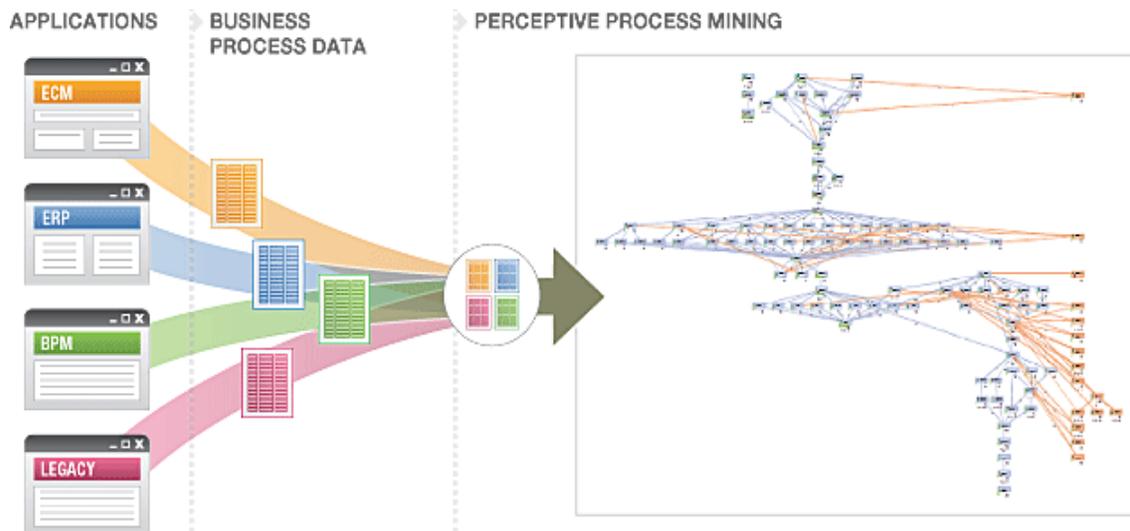


Fig iii) Perceptive Process Mining

By examining event data and applying process pattern recognition techniques to create process models in a fraction of the time and cost involved in traditional methods, process mining helps you view the interaction between the parts of your organization as a whole, so you can:

- i) Find problems:** Locate anomalies even before they become problems.
- ii) Understand root causes:** Identify the underlying problems instead of immediately visible symptoms.
- iii) Achieve compliance:** Uncover non-compliant situations and uncover incidental non-compliance that would otherwise be impossible to locate.
- iv) Compare performance:** Compare departments or similar organizations based on verifiable data, not hunches.
- v) Validate process changes:** Establish baselines for the current situation and use them to determine whether specific changes are effective or not.
- vi) Understand the entire process:** Departments often focus on completing their own subtask, unaware of the complete chain of actions that bring a product or service to the customer.
- vii) Gain transparency:** Since causes of problems can be identified and seen in the context of the organization's overall workflow, employees understand where and why change is needed.

The real impact of process mining is all about people. Continuous process improvement leads to more satisfied customers, providing a competitive advantage for your company. And by challenging assumptions about how employees work together, it can point the way to more effectively deploying your most valuable asset. Ultimately, Perceptive process mining can turn process improvement from an occasional, resource-intensive, disruptive effort into a routine part of your daily operation.[26]

VII. Conclusion

In this paper, we studied about the notion of process cubes. It gives end users the opportunity to analyze and explore processes interactively on the basis of a multidimensional view on event data. Process mining aims to bridge the gap between Business Intelligence (BI) and Business Process Management (BPM).

BI techniques are typically not process centric and provide rather simplistic reporting and dashboard functionalities. Sometimes BI tools offer more advanced data mining capabilities.

References:

- [1] PROCESS MINING Article by StB Prof. Dr. Nick Gehrke Nordakademie Chair for Information Systems Köllner Chaussee 11, D-25337 Elmshorn, nick.gehrke@nordakademie.de, Michael Werner, Dipl.-Wirt.-Inf., University of Hamburg, Chair for Information Systems Max-Brauer-Allee 60 D-22765 Hamburg michael.werner@wiso.uni-hamburg.de
- [2] T. H. Davenport. Putting the Enterprise into the Enterprise System. *Harvard Business Review*, 76(4):121-131, 1998.
- [3] R. Agrawal, A. Gupta, and S. Sarawagi. Modeling Multidimensional Databases. In *Proceedings of the Thirteenth International Conference on Data Engineering, ICDE '97*, pages 232-243, 1997.
- [4] W. M. P. Van der Aalst. Mining Process Cubes from Event Data (PROCUBE), project proposal (under review). 2012.
- [5] I.-M. Ailenei. Process Mining Tools: A Comparative Analysis. Master's thesis, Eindhoven University of Technology, 2011.
- [6] W. M. P. van der Aalst. Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In J. Liu M. Song, M. Wynn, editor, *Asia Pacific conference on Business Process Management (AP-BPM 2013)*, Lecture Notes in Business Information Processing, 2013.
- [7] Realizing a Process Cube allowing for the comparison of Event Data a master thesis by Tatiana MamalingaTU/e, Eindhoven, August 2013.
- [8] A. Berson and S. J. Smith. *Data Warehousing, Data Mining, and Olap*. 1997.
- [9] F. Melchert, R. Winter, and M. Klesse. Aligning Process Automation and Business Intelligence to Support Corporate Performance Management. In *AMCIS'04*, pages 507-507, 2004.
- [10] R.P. Jagadeesh Chandra Bose. *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*. PhD thesis, Eindhoven University of Technology, 2012.
- [11] J. C. A. M. Buijs. *Mapping Data Sources to XES in a Generic Way*. Master's thesis, Eindhoven University of Technology, 2010.
- [12] C. Li, M. Reichert, and A. Wombacher. The MINADEPT Clustering Approach for Discovering Reference Process Models Out of Process Variants. *International Journal of Cooperative Information Systems*, 19(3-4):159-203, 2010.
- [13] J. Munoz-Gama and J. Carmona. A Fresh Look at Precision in Process Conformance. In R. Hull, J. Mendling, and S. Tai, editors, *Business Process Management (BPM 2010)*, volume 6336 of *Lecture Notes in Computer Science*, pages 211-226. Springer-Verlag, Berlin, 2010.
- [14] Process mining evolved in the context of analyzing software engineering processes by Cook and Wolf in the late 1990s (Cook and Wolf 1998). Agrawal and Gunopulos (Agrawal et al. 1998) and Herbst and Karagiannis (Herbst and Karagiannis 1998)

- [15] J. C. A. M. Buijs, B. F. van Dongen, and W. M. P. van der Aalst. Towards Cross-Organizational Process Mining in Collections of Process Models and Their Executions. In Business Process Management Workshops (2), pages 2-13, 2011.
- [16] J. W. Buzydlowski, I.-Y. Song and L. Hassell. A **Framework for Object-Oriented On-Line** Analytic Processing. In Proceedings of the 1st ACM International workshop on Data ware-housing and OLAP, DOLAP '98, pages 10-15, 1998.
- [17] S. Chaudhuri and U. Dayal. An Overview of Data Warehousing and OLAP Technology. SIGMOD Record, 26(1):65-74, 1997.
- [18] S. Chaudhuri, U. Dayal, and V. Narasayya. An Overview of Business Intelligence Technology. Commun. ACM, 54(8):88-98, August 2011.
- [19] E. F. Codd, S. B. Codd, and C. T. Salley. Providing OLAP (On-Line Analytical Processing) to User-Analysis: An IT Mandate, 1993. White paper.
- [20] G. Colliat. OLAP, Relational, and Multidimensional Database Systems. SIGMOD Record, 25(3):64-69, 1996.]
- [21] <http://coresystems.ch/en/consulting/business-intelligence/>.
- [22]. <http://www.tcs.com/offering/business-intelligence-performance-management/business-process-management/Pages/default.aspx>.
- [23] <http://www.wis.win.tue.nl/~wvdaalst/publications/p651.pdf>.
- [24] <http://www.bpminstitute.org/resources/articles/what-bpm-anyway-business-process-management-explained>.
- [25] <http://www.sqlpower.ca/xbrlpower/page/biprimer>.
- [26] [.http://www.perceptivesoftware.com/products/perceptive-process/process-mining](http://www.perceptivesoftware.com/products/perceptive-process/process-mining).

Author’s Bibliography:



Sarada Sowjanya .C completed her Mtech from JNTU Kakinada with Distinction. She did her BE from RTM NAGPUR UNIVERSITY, Currently working as Asst. Professor in CSE dept, Sri Indu College of Engg & Technology Her research interests include: Data Mining, Computer Networks, and Data Base Management Systems.



Dr. Ch G.V.N. Prasad currently working as a Professor and HOD of CSE dept, Sri Indu College of Engg & Technology. He gained 12 years of experience in IT industry (8 years in National Informatics Centre, Govt. of India, as Scientist and Software Analyst in AT&T in US) and 11 years of experience in Teaching (As a Professor & HOD of CSE Dept).



R.Vineela Completed her Mtech from JNTUH with Distinction. She completed her MCA from St.John's PG College OU. Currently working as a Asst. Professor in CSE dept, Sri Indu College of Engg & Technology. Her areas of Interest are: Data Mining, Web Designing and Database Management System.



J.Sowjanya completed her Mtech from SIT, JNTUH with Distinction. She did her BE from Muffakham-Jah College of engineering and Technology, OU. Currently working as a Asst. Professor in CSE dept, Sri Indu College of Engg & Technology. Her areas of interest are: Data Mining, Web Technologies.