SEMINAR TOPICS COVERED		
Definition of Big Data Environment	The Five V's of Big Data (V5)	BIG DATA Governance
Types of Big Data / Data Sources	Topical, Social, Temporal View	Data Streaming Applications
What is HADOOP/How Does it Work	Dremmel and Spanner	Business Drivers of High Volume
HDFS / Map Reduce / Latin / Pig Latin	Complex Event Processing (CEP)	Streaming Analytic Query (CQL)
Distributed Servers (Amazon EC2)	Semantic Processing (OWLIM)	Document Oriented Databases
Hybrid Architectures	Workflow Engines (OOZIE)	Workflow Balancing / Scalability
Distributed Processing (Map/Reduce)	Distributed Storage-Amazon S3)	Distributed File Systems (DFS)
SCRIBE and Logging Vehicles	Streaming Data Collection (Flume)	Columnar Databases (HBASE)
Linking Relational Database (e.g.Sqoop)	Building Unstructured Databases	Machine Learning Library(Mahout)
Clustering and Classification	NoSQL Database (e.g. Mongo)	CouchDB, VoltDB, ZooKeeper
Graph Databases and Visualization	Role of Data Modeling	Taxonomies and Ontologies
ETL: Unstructured/Structured	Metadata Sharing/Integration	Cassandra, Redis, HBase
Innovative Data Integration	Smart Sampling / Data Mining	Fortune 500 Typical Applications
Big Data Lifecycle	Best Practices / Best Technologies	Workload Driven Design
Practical Considerations	Future Trends & Technologies	Best of Breed Tools & Techniques

WHAT YOU WILL LEARN
Defining Big Data – What Is It and How Best To Utilize It
The Major Architectural Components of BIG Data – How They Fit Together
Big Data Technologies and How to Decide the Best for Your Organization
Linking Traditional Databases to the BIG DATA Environment
Data Management in Big Data Era: The Challenges and The Opportunities
Why so many Versions of HADOOP: Apache, IBM, Oracle, Google, Amazon, etc.
Implementing Moneyball Techniques to Leverage Big Data – The Role of Analytics
How to Architect Your Enterprise to Optimize Business Intelligence for BIG DATA
Why Globalization is Driving Integration and Massively Parallel Processing of Big Data
Best Practices For Managing Structured, Semi-Structured and Unstructured Big Data
What are the Leading Vendors in the BIG DATA Space: Databases, Reporting Tools, Query Languages
What are the Technology Skill Sets Needed for Training and Implementing a BIG DATA Project
How do you achieve Scalable Data Analysis and Visualization (SDAV)
Why does the role of Super Computing (e.g. IBM Watson) support the BIG DATA Environment
What Kinds of Applications Require GeoSpacial Analytics, Textual Disambiguation, Workload IsolatioN
How does the Usage of Mobile Communications and Social Media support your Big Data Projects
What is Crowd Sourcing, Longtail and Gamification Strategies
How Should you use the Forrester Wave Product Evaluation / Selection Criteria
White House Office of Science and Technology Policy (OSTP) - BIG DATA Research & Development Program (Co-Led by NIH and NSF): What Kinds of Applications and Overview of Technologies
Case Studies - Realizing Incredible Potential for Totally New Applications

SEMINAR ABSTRACT AND BIOGRAPHY

Since the dawn of time up until 2003 mankind has created 2 exabytes of data, but today at least that much data is created each and every day. This means that standard traditional databases and tools cannot process the huge volumes of data, nor can they discern and separate the “High Value” from the “Low Value” data in order to support analytical insights. David Menninger of EMC says "we used to ask whether we could afford to store information...today we ask whether we can afford to throw it away ". Five years ago, a scalable relational database cost $100k per terabyte and $20k per year maintenance, but today you can store the same amount of information for $1200 per year. It is no longer a technology decision but an economic one as the technical capabilities have progressed light years ahead in a relatively short amount of time.

The world of Data Warehouse has been evolving rapidly over the last three years into an ecosystem that is architected with traditional data, Big Data, real time data, semantic interfaces, complex algorithms, new infrastructures and serves the analytical needs of users including mobile platforms. This seminar will be a two-day interactive session on how to understand the impact of Big Data on Data Warehousing. Does every enterprise need to adopt to the new evolutions of the Data Warehouse? Do the new components of the Data Warehouse provide real additional insights that can be harnessed to enable transformations in the enterprise? Does ETL have a future? How does one build new analytical platforms? Who is a data scientist and what role should that person play in your architecture?

According to James Kobielus, IBM's Big Data Evangelist, today's BIG DATA Developers must wrangle with a plethora of SQL-Like languages for big data access including HiveSQL, Cassandra QL, JAQL, SQOOP, Sparkll, Shark and DrQL. How do you navigate across the myriad of products and available tools?

How can the Forrester Wave Vendor Ranking and Evaluation Criteria be used to study and select what is best for your organization's needs? What types of applications are typically being implemented?

According to Ralph Kimball, "Data Warehousing has been demonstrating the value of data-driven insights for at least 20 years, ...but recently "many enterprises are looking seriously at unstructured data for the first time..." and it is going mainstream. How do you effectively link unstructured data, semi-structured data and classical relational data together to support your information architecture in order to meet the business needs and goals of your organization?

This seminar provides you a one of kind perspective on both the business and IT side of Big Data and the Data Warehouse. It includes live examples and real world case studies that will enable you to understand this subject. You will gain insights into the new world of Data Warehouse, how the new technologies are changing the way we design and build data warehouses, what will the new generation of reports and analytics look like and who are the companies that have forayed into this new world.

This two-day seminar will address these questions and many others related to BIG DATA. Practical approaches and techniques are contained throughout the presentation on the business, technical, and organizational aspects of Big Data. It will also address the skill sets required to support BIG DATA applications and BIG DATA Architectures. The seminar format is very interactive with attendee questions encouraged throughout the day, and attendee experiences shared as time permits.

ABOUT KRISH KRISHNAN

Mr. Krishnan is a recognized expert worldwide in the strategy, architecture and implementation of high performance data warehousing and database solutions. As Special Advisor to Bill Inmon (“The Father of Data Warehousing”) and TDWI Instructor and Author (The Data Warehouse Institute), he is a visionary data warehouse thought leader, ranked as one of the top 25 data warehouse consultants in the world, and an independent analyst, writing and speaking at industry leading conferences, user groups and trade publications. He has authored three eBooks, over 125 plus articles, viewpoints, whitepapers and case studies in Big Data, Business Intelligence, Data Warehousing and Data Warehouse Appliances and Architectures. He has co-authored a book with Bill Inmon entitled “Building The Unstructured Data Warehouse”, the ground-breaking publication upon which Bill Inmon’s “Unstructured Data” Seminar is based.

A recognized authority on Unstructured Data Integration, text mining and text analytics, along with Bill Inmon, he is promoting the next generation of data warehousing, primarily on DW 2.0 platforms with unstructured data integration and social intelligence as key areas in BIG Data and Analytics.

Specialties: Performance Tuning VLDBs (250TB+), Social Media Analytics, High Performance Data Warehouse Architecture and Design, Hadoop, BIG Data, NoSQL, Text Mining, Deep Network Analytics, Neural Networks, Artificial Intelligence.

Look for Krish Krishnan's articles on BIG DATA Analytics as Data Management Forum will be working with other publications during the 2014 timespan including IBM's Data Magazine, Information Management Magazine (Source Media) and DataBase Trends and Applications Magazine, as well as publishing many of these articles in the Data Management Forum eNewsletter.

WHO SHOULD ATTEND

This seminar is intended for business and IT Management and Executives and Technical Staff Interested in the Concepts and Usage of BIG DATA including: Project Directors, Data Warehousing, Data Architecture or Database Management executives and staff, Developers, Managers and Technicians responsible for planning, design and implementation, Data Stewards and Custodians, Architects, Metadata personnel, and anyone looking for a clear concise overview of data management and architecture as it relates to BIG DATA tools, technologies, techniques and applications. This is a "One-of-a-Kind" Intensive Tutorial Assuming no pre-requisite knowledge of BIG DATA.