WILLIAM (BILL) INMON

AGENDA (1-HOUR WEBINAR)

1. WELCOME & WEBINAR RULES (KEWAL DHARIWAL)
2. INTRODUCTIONS (STEPHEN KAISER, TERRI HOWALD)
3. SPONSOR – DATABRICKS (ROB SIGNORETTI, SAJITH APPUKUTTAN)
4. PRESENTATION (BILL INMON)
5. QUESTIONS (WADE WALKER)
6. THANK YOU (ERIC DRESSLER)
DAMA EDMONTON
Data Analytics and Business Intelligence for Data Professionals in Edmonton

http://dama-Edmonton.org
Hosted by
https://iccp.org

iccp
Institute for Certification of Computing Professionals

Get Certified, Hire Certified:
Big Data, Data Science, CyberSecurity Certifications

DAMA EDMONTON BOARD

INTRODUCTIONS/ORGANIZERS

STEPHEN KAISER
PRESIDENT, DAMA EDMONTON
DIRECTOR, MEMBER INSIGHTS SERVUS CREDIT UNION

TERRI HOWALD
VICE PRESIDENT, MEMBER SERVICES
DATA CONSULTANT, UNIVERSITY OF ALBERTA

ERIC DRESSLER
VICE PRESIDENT, DAMA EDMONTON CHAPTER
DATA SCIENCE INITIATIVES
DIRECTOR, DATA ANALYTICS, AMA

KEWAL DHARIWAL
FOUNDER, DAMA EDMONTON CHAPTER
EXECUTIVE DIRECTOR, ICCP

BOB HIEW
FOUNDER & PAST PRESIDENT, DAMA EDMONTON CHAPTER

SURINDER PADEM
VICE PRESIDENT, MARKETING & LOGISTICS
MANAGER, NAIT COMPUTER TRAINING CENTER
WILLIAM (BILL) INMON

AGENDA (1-HOUR WEBINAR)

1. WELCOME & WEBINAR RULES (KEWAL DHARIWAL)
2. INTRODUCTIONS (STEPHEN KAISER, TERRI HOWALD)
3. SPONSOR – DATABRICKS (ROB SIGNORETTI, SAJITH APPUKUTTAN)
4. PRESENTATION (BILL INMON)
5. QUESTIONS (WADE WALKER)
6. THANK YOU (ERIC DRESSLER)
WILLIAM (BILL) INMON

AGENDA (1-HOUR WEBINAR)

1. WELCOME & WEBINAR RULES (KEWAL DHARIWAL)
2. INTRODUCTIONS (STEPHEN KAISER, TERRI HOWALD)
3. SPONSOR – DATABRICKS (ROB SIGNORETTI, SAJITH APPUKUTTAN)
4. PRESENTATION (BILL INMON)
5. QUESTIONS (WADE WALKER)
6. THANK YOU (ERIC DRESSLER)
Click to add notes

Rob Signoretti, Enterprise Account Executive
Sajith Appukuttan, Solution Architect
Rob Signoretti, Enterprise Account Executive
Sajith Appukuttan, Solution Architect
A unified data analytics platform for accelerating innovation across data engineering, data science, and analytics

- Global company with over 5,000 customers and 450+ partners
- Original creators of popular yib data and machine learning open source projects

Logos: Apache Spark, Delta Lake, MLflow
Most organizations fail to unlock business value due to data, technology and people silos.
Unlocking business value: Four challenges

1. Data is messy, siloed, and slow
2. ML is hard, Production is harder
3. BI is limited to a fraction of data
4. Lack of enterprise readiness

- Data Engineers in IT
- Data Scientists in Business

- Fragmented security
- Poor reliability
- Disjointed governance
Databricks Unified Data Analytics Platform

Data science, ML, and analytics on one cloud platform

Access all business and big data in open data lake

Securely integrates with your cloud ecosystem

mlflow DATA SCIENCE WORKSPACE
Collaboration across the lifecycle

PyTorch
TensorFlow

UNIFIED DATA SERVICE
High quality data with great performance

Apache Spark

ENTERPRISE CLOUD SERVICE
A simple, scalable, and secure managed service

AWS

BI INTEGRATIONS
Access all your data

DELTA LAKE

BIG DATA & BUSINESS DATA
SPARK+AI SUMMIT
June 22-26 | Organized by databricks

THE VIRTUAL EVENT FOR DATA and ANALYTICS TEAMS

- Extended to 5 days with over 200 sessions
- 4x the pre-conference training
- Keynotes by visionaries and thought leaders

NOW FREE
DATA ARCHITECTURE

A presentation by
W H Inmon
All the data in the corporation
98% of corporate decisions made on 10% of the data.
Something is wrong here....

Decisions?
Very rich in business value
Process manually – a really bad idea
If we are serious about using text in decision making, we need to put text into a database.
Fitting a square peg in a round hole
The progression of textual analytics in technology

- Textual ETL
- NLP/taxonomies
- Tagging
- Sound-ex/stemming
- Blobs
- Comments
In order to fit text into a database, you need textual ETL.
Trying to sell a database
is really hard to do.
Nobody wants a database.
What people REALLY want are Visualizations
SOME EXAMPLES
Call center – what are people saying?
Hotel/restaurant feedback – listening to your customer
Product(s): Caesar Salad Parmesan Chicken Breast I have ordered a Caesar Salad with a Parmesan Chicken Breast twice now. Both times, the restaurant neglected to include the chicken in my to go order. Also, today I was given dressing that had separated. I called and was told that that’s just how it was going to be. I love your restaurant and was very disappointed. Not a good experience at all.

Typical feedback from a customer
Airline passenger feedback – how was your trip
Passengers talk on the Internet
Medical records
Dear Dr. Weber:

Thank you for referring your patient for consultation regarding a new area of microcalcifications on her most recent screening mammogram.

As you are aware, she is an otherwise healthy 43-year-old woman who underwent screening mammograms at the Women's Imaging Center in Hyannis, on June 19, 2002. This showed a tight cluster of microcalcifications in the mid to lower lateral aspect of the right breast. These are nodular, dense and heterogeneous in their character. Stereotactic core needle biopsy was recommended.

PAST MEDICAL HISTORY: Negative.

PAST SURGICAL HISTORY: Status-post appendectomy.

BREAST CANCER RISK FACTORS: Menarche was at age 12. Her last menstrual period was on July 2, 2002. She menstruates regularly, without signs or symptoms of menopause. She is gravida 1, para 0, TAB 1. She used oral contraceptives at age 15. She has no other history of hormonal therapy. She has no known exposure to diethylstilbestrol.

FAMILY HISTORY: A paternal aunt had fatal ovarian cancer diagnosed at age 70. A paternal first cousin had breast cancer diagnosed at age 36 and died at age 38. A paternal aunt had colon cancer diagnosed at age 58. Her maternal grandmother had bladder cancer at age 84 and died of a ruptured gallbladder.

MEDICATIONS:
(1) Aspirin p.r.n.
(7) Vitamin C

A typical medical record
Dear Dr. Weber:

Thank you for referring your patient, Ms. Johnson, to our vascular surgery clinic for a consultation regarding a new area of skin discoloration on her right arm.

As you know, she is an otherwise healthy 43-year-old woman who previously underwent screenings for vascular disease at the Women's Imaging Center in St. Louis, on June 19, 2002. This showed no signs of skin discoloration or other abnormalities.

PAST MEDICAL HISTORY: Ms. Johnson has no history of vascular disease or any other significant medical conditions. She has no history of smoking or other risk factors for vascular disease.

PAST SURGICAL HISTORY: Ms. Johnson has had no previous surgeries or procedures.

BREAST CANCER RISK FACTORS: Ms. Johnson has no family history of breast cancer. Her mother is alive and well, and her sister has not had breast cancer.

FAMILY HISTORY: Ms. Johnson's father died of a heart attack at age 60. Her mother is alive and well. Ms. Johnson has no other significant family history.

MEDICATIONS:
1. Aspirin p.r.n.
2. Calcium carbonate

A quick look inside the record...
The medical record recreated inside the data base
So who wants to turn their textual information into important decisions?

!!!!!!!!EVERYBODY!!!!!

Restaurants
Hotels
Banks
Telephone companies
Doctors
Hospitals
Pharmaceuticals
Airlines
Government agencies
Manufacturers
So who wants to turn their textual information into important decisions?

!!!!!!!EVERYBODY!!!!!
AGENDA (1-HOUR WEBINAR)

1. WELCOME & WEBINAR RULES (KEWAL DHARIWAL)
2. INTRODUCTIONS (STEPHEN KAISER, TERRI HOWARD)
3. SPONSOR – DATABRICKS (ROB SIGNORETTI, SAJITH APPLUKTIAN)
4. PRESENTATION (BILL INMON)
5. QUESTIONS (WADE WALKER)
6. THANK YOU (ERIC DRESSLER)
WILLIAM (BILL) İNMÓN

AGENDA (1-HOUR WEBINAR)

1. WELCOME & WEBINAR RULES (KEWAL DHARIWAL)
2. INTRODUCTIONS (STEPHEN KAISER, TERRI HOWALD)
3. SPONSOR – DATABRICKS (ROB SIGNORETTI, SAJITH APPUKUTTAN)
4. PRESENTATION (BILL İNMÓN)
5. QUESTIONS (WADE WALKER)
6. THANK YOU (ERIC DRESSLER)

DAMA Edmonton
Data Analytics and Business Intelligence for Data Professionals in Edmonton

http://dama-Edmonton.org
Hosted by https://iccp.org

iccp
Institute for Certification of Computing Professionals

Get Certified, Hire Certified: Big Data, Data Science, CyberSecurity Certifications