

Unstructured Data:

Creating the Analytical Environment

Bill Inmon



4-5 December 2008, London

Fee £995 & VAT (£174.13) =
£1169.13
Group Booking & Series
Discounts Available

Overview

It is estimated that 80% of data in the corporation is textual. For years, managing textual data has meant placing documents in some form of ECM – Enterprise Content Management. But trying to do analysis on data found in ECM is a very different story than placing the documents there in the first place.

With DW 2.0 the idea arose that unstructured data is best placed in a data warehouse, where it can be analyzed along with other structured data found in the data warehouse. This seminar is about the work that needs to be done in order to take textual data out of the confines of documents and integrate the textual data into a data warehouse. This is a very down to earth seminar and workshop. The first day is a seminar based on the background material needed to understand the architecture surrounding the placement of text in an analytical, data warehouse environment. The second day is a workshop that shows – step by step – how text is converted into a data base that can then be placed into a data warehouse.

Learning Objectives

- Recognize the difference between search processing and textual analytics
- How to create a data warehouse that contains textual data
- What is required to turn textual data into data that is fit for a data warehouse
- What internal and external taxonomies are and why they are important
- What synonym concatenation is and why it is important
- What homographic resolution is and why it is important
- What semi structured data is and how it has to be handled
- How to integrate the textual integrated data to a data warehouse data model
- What stop words, stems, and standardized data elements are and how to handle them
- How to create a foundation that can be analyzed by standard BI tools
- How to scale the integration process
- How to read and interpret semi structured data
- How to create proximity variables (and why proximity variables are so important)
- How to integrate textual data with visualization
- Understand the difference between textual discovery and textual analytics

Seminar & Workshop Outline

Introduction

- Discovery
- Analysis

Search versus Analysis

- Transformation of text

Types of Unstructured Data

- Voice
- Image
- Text

Types of Textual Data

- Simple Unstructured Data
- Semi Structured Data
- Volumes of Data
- Textual versus Structured Data
- Scaling Volumes of Data

The Unstructured Data Base

- Iterative Development
- Imperfect Data

Integrating Simple Unstructured Data

- Stop Words
- Stems
- Synonyms
- Homographs
- Internal/External Taxonomies
- Alternate Spelling
- Multiple Languages
- Proximity Variables

- Date Standardization
- Text to Numeric Conversion
- Email (Screening)

Integrating Semi Structured Data

- Sub document Separation
- Looking at Hidden Characters
- Pattern Recognition
- Symbol Recognition
- Multiple Types of Indexing
- The Subject Oriented Index
- Index Trimming
- List processing

Linking Unstructured and Structured

- Dynamic links
- Static links
- By Name
- By variable
- By Communication Id
- By Business Id

Visualization

- SOMs
- Simple Unstructured SOM
- Semi Structured SOM
- The Discovery Process

Other Miscellaneous Topics

- Technology Infrastructure
- A Methodology for Unstructured Text

Data Management Series

Data Modelling Fundamentals

23-24 September 2008,
16-17 February 2009, London

Data Modelling Masterclass

25-26 September 2008,
18-19 February 2009, London

DW 2.0: Next Generation of Data Warehousing

2-3 December 2008, London

Unstructured Data

4-5 December 2008, London

Master Data Management

15-16 September 2008, London

Series Discounts

Attend more than one course in this series and you will be entitled to the following discounts:

2nd course	10%
3rd course	15%
4th course	20%
5th course	25%

Group Booking Discounts

If 5 delegates from the same organisation register at the same time for the same or various seminars, then the 5th delegate is free. We regret that this offer cannot be used in conjunction with the Series Discount.

Presenter



Bill Inmon, Inmon Consulting Services – known as the father of the data

warehouse – has written 47 books and over 1000 articles. Bill's books have been translated into 9 languages. Bill has a weekly newsletter with b-eye-network of over 55,000 recipients. Bill holds 8 software patents.

In-House Training

For an in-house presentation of this or any other IRM UK seminar, please contact us on:
T: +44 (0)20 8866 8366
E: customerservice@irmuk.co.uk

Endorsed by:



The Premier Organisation for data professionals in the UK

Special Features

● This seminar is based on two books which are commercially available. The first book is **Tapping into Unstructured Data**, Prentice Hall, 2007. The second book is **DW 2.0 – Architecture for the Next Generation of Data Warehousing**, El Sevier Press, 2008. Both of these books are available and the attendee is encouraged to read these books before attending the seminar.

● In addition there is a collection of white papers to which the attendee will be directed.

Visit www.irmuk.co.uk for full course outline

Registration Information

Registration Fees:

Full payment or a purchase order is due prior to the event. Payment may be made in Sterling (£) or Euros. If paying in Euros the prevailing exchange rate of the country of the delegate or delegates' company is to be used. The total Euros remitted should be the amount required to purchase the sterling pound cost of the event on the day of payment. All delegates must add VAT (17.5%) to their total event fees. VAT may be reclaimed by delegates from the tax authorities after the event.

GROUP DISCOUNTS: For every 4 people from one company who register for the same event, at the same time, a 5th delegate may attend FREE! Please Note: We cannot offer any other published discounts / offers to those taking advantage of this group discount.

The registration fee includes the lectures, documentation, refreshment breaks and lunch on each day of the event. The cost of hotel accommodation is not included in the event fee.

UK Delegates: Expenses of travel, accommodation and subsistence incurred whilst attending any IRM UK event will be fully tax deductible by the employer company if attendance is undertaken to maintain professional skills of the employee attending.

Non-UK Delegates: Please check with your local tax authorities.

Cancellation Liability: In the event of cancellation of an event for any reason, IRM UK's liability is limited to the return of the registration fee only. IRM UK will not reimburse delegates for any travel or hotel cancellation fees or penalties.

Cancellation Policy: Cancellations must be received in writing at least two weeks before the commencement of the event and will be subject to a 10% administration fee. It is regretted that cancellations received within two weeks of the event date will be liable for the full fee. Substitutions can be made at any time.

Hotel Accommodation and Seminar Venue Details

IRM UK in association with JP Events, have arranged special discounted hotel rates at our various seminar venues and at other hotels nearby. Seminar venues will be confirmed upon registration. Alternatively, please visit our website.

Email: enquiries@jpeventsltd.com www.jpeventsltd.com

Tel: +44 (0)20 7428 9911 **Fax:** +44 (0)20 7428 9966

Alternatively, you may book directly at the relevant hotels

Seminar Timetables

08.30 – 09.00 Registration (first day only) 09.00 – 12.30 Event

12.30 – 13.30 Lunch 13.30 – 17.00 Event

Mailing Information

As we are using multiple mailing lists, there is the possibility that you may receive more than one brochure. If this is the case or if there is an error in your address details, please forward the incorrect mailing labels to us so we can update our database immediately.


Please tick here if you are unable to attend a seminar but would like your name added to our mailing list.


IRM UK Strategic IT Training Ltd


IRM UK are specialist in strategic IT training for IT & business professionals and managers. Our carefully selected presenters have superior technical knowledge, teaching skills and a wide range of practical business experience. They are some of the most influential technologists, methodologists and original thinkers in IT today. Noted for the participation of top level decision makers from both the corporate, user and vendor communities, and the lucid analysis of critical strategic and management issues, our events are condensed and rigorous combining technical explanations with management advice and discussions of future directions. Please visit our website for comprehensive information on our events and speakers www.irmuk.co.uk.

How to Register

 e-mail: customerservice@irmuk.co.uk

 website: www.irmuk.co.uk

 Fax: +44 (0)20 8866 7966

 Post: IRM UK Strategic IT Training Ltd,
Bishops Walk House, 19-23 High Street,
Pinner, Middlesex HA5 5PJ, UK

 Phone: +44 (0)20 8866 8366

Registration Form

Organisation

Address

City Postcode

Country

Telephone Fax

e-mail

Booking made by

Approving Manager Position

Please tick here if this completed form confirms your telephone registration.

Delegate Details

1. Surname Mr/Ms/Mrs/Dr

First Name Direct tel:

Job Title e-mail

Event Name & Date

2. Surname Mr/Ms/Mrs/Dr

First Name Direct tel:

Job Title e-mail

Event Name & Date

3. Surname Mr/Ms/Mrs/Dr

First Name Direct tel:

Job Title e-mail

Event Name & Date

Billing address if different from above

BOOKING CODE MUST BE QUOTED WHEN REGISTERING

BOOKING CODE = WBD

Payment Details: Please note that full payment or a purchase order is due prior to commencement of the event

Please tick appropriate method of payment:

Cheque, drawn on a UK bank, enclosed for £ made payable to IRM UK Strategic IT Training Ltd

Purchase Order Number

Bank Transfer (Please attach a copy of your bank transfer order to this booking) Contact IRM UK for our bank details on +44 (0)20 8866 8366. When paying by bank transfer, please ensure that the order includes the following details: the Event Title, the Delegate Name, and your Company Name. Please ensure the full amount is paid, including VAT.

Credit Card

American Express Visa Mastercard Eurocard Delta
Please charge £ to my account

Name on credit card (please print)

Cardholder's Signature Date

Account Number

Expiry Date Security Code

Credit card Billing Address if Different from above

Data Protection: The personal information that you provide will be held on a database by IRM UK. We may occasionally release your details to other reputable companies. If you prefer not to be included in this service please tick the box