Analytics for IBM Content Manager OnDemand

Or

Reports – a Love/Hate Relationship

Harvey C. Gross
C.T.O.
Datawatch

DATAWATCH

Agenda

1. Who is Datawatch
2. Reports – The Good, The Bad and The Ugly
3. History of Data Mining Methods
4. Governance, Data Lineage, Single Source of the Truth
5. Demonstration
6. Questions

DATAWATCH

About Datawatch

Founded in 1986, Public since 1992 (NASDAQ: DWCH)

Pioneer in real-time visualization and transformation of any data

Global operations and support

- US, UK, Germany, France, Australia, Singapore, Philippines

Extensive global customer base

- 99 of the Fortune 100

Resold and embedded by leading vendors

- SAP, IBM, Tibco, Thomson Reuters

Recognized leader and innovator

Gartner 2013 Cool Vendor

Ventana Research 2013 Technology Innovation Award WINNER

CRN Big Data 100

Buy-Side Technology Awards 2013 waterstechnology Winner

KM World Trend-Setting Products of 2013

DATAWATCH

Premier & Expansive Customer Base

Financial Services

citi
Bank of America
GMAC Mortgage
ING
HSBC
AIG
Credit Suisse
JPMorganChase
Barclays
Scottrade
Broadridge
John Hancock
Aetna

HealthCare

Good Samaritan Hospital
A Tradition of Caring
St. Anthony's Hospital
BayCare Health System
Baptist Health
UnitedHealthcare
The Chester County Hospital and Health System
LIFESPAN
WellPoint
Pfizer

Government

U.S. Army
IRS
United States Postal Service
Navy Federal Credit Union
U.S. Air Force
City of Atlanta
Public Service Credit Union
Together we're better

Other Industries

adidas
GE
logica
be brilliant together
Ford
Lowe's
Let's Build Something Together
Northrop
McKesson
Toyota
Sony

DATAWATCH

What's to Love

- An exact copy of customer facing documents
  - Invoice, statements, other customer communications
  - The official legal record of what the customer received
- Reports are a "snapshot in time" of the state of the business
- Authorized, verified, vetted information
  - The report is "always right"
  - Countless hours of development and testing

DATAWATCH

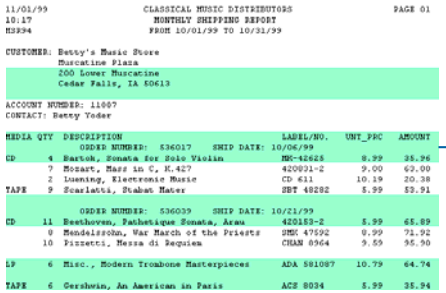
What I Hate

- Static Information
  - Sorted incorrectly for my needs
  - Contains TOO MUCH information
    - Not my department
    - Need to narrow down to anomalies
  - Cannot narrow focus
  - Takes forever to get to the right details
  - And … just as long to organize information into a meaningful summary
- Lots of information, but … It's not data!

DATAWATCH

Datawatch Report Mining Server

Resolve the relationship issue by turning information back into data



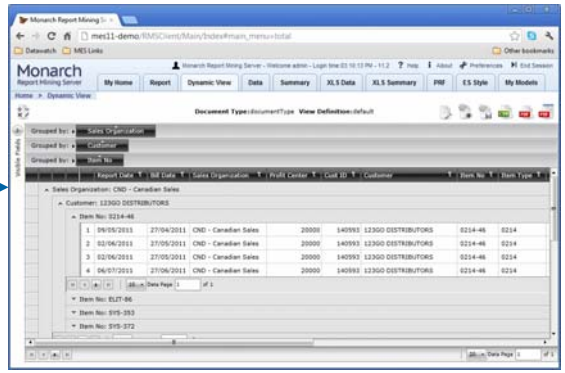
11/01/99 CLASSICAL MUSIC DISTRIBUTORS PAGE 01
10:17 MONTHLY SHIPPING REPORT
M3394 FROM 10/01/99 TO 10/31/99

CUSTOMER: Betty's Music Store
Muscatine Plaza
200 Lower Muscatine
Cedar Falls, IA 50613

ACCOUNT NUMBER: 11067
CONTACT: Betty Yoder

MEDIA	QTY	DESCRIPTION	LABEL/NO.	UNT_PRC	AMOUNT
		ORDER NUMBER: 536017 SHIP DATE: 10/06/99			
CD	4	Bartok, Sonata for Solo Violin	MK-42625	8.99	35.96
	7	Mozart, Mass in C, K.427	420891-2	9.00	63.00
	2	Luening, Electronic Music	CD 611	10.19	20.38
TAPE	9	Scarlatti, Stabat Mater	SBT 48282	5.99	53.91
		ORDER NUMBER: 536039 SHIP DATE: 10/21/99			
CD	11	Beethoven, Pathetique Sonata, Arau	420153-2	5.99	65.89
	8	Mendelssohn, War March of the Priests	SMK 47592	9.99	71.92
	10	Pizzetti, Messa di Requiem	CHAN 8964	9.59	95.90
LP	6	Misc., Modern Trombone Masterpieces	ADA 581087	10.79	64.74
TAPE	6	Gershwin, An American in Paris	ACS 8034	5.99	35.94

DATAWATCH
Report mining server



DATAWATCH

Datawatch Report Mining Server

- **Transforms CMOD** into a search-to-action *Big Data* asset.
- **Increases productivity** of report users by providing one click transformation of reports into actionable formats
  - Interactive Web Analytic Framework
  - Excel spreadsheets
  - PDF Documents
- **Reduces IT burden** by eliminating need for ad-hoc reporting, sub-reports, or live database connectivity
- **Extends CMOD ROI** by eliminating the need for costly data marts or warehouses to meet key data and report mining needs

DATAWATCH

Datawatch Report Mining Server

- **Provides ad-hoc analysis,** investigation and auditing capabilities on core corporate data
- **Deploys quickly and easy** with browser client
- **Easily integrates** with IBM Content Manager on Demand with a plug-in for ICN

DATAWATCH

Report Mining Historical Methods

Data Entry

Slow
Labor Intensive
Prone to Errors

DATAWATCH

Report Mining Historical Methods

Cut and Paste

Slow
Labor Intensive
Prone to Errors

DATAWATCH

Report Mining Historical Methods

Parsing

A parser is a software component that takes input data (frequently text) and builds a data structure – often some kind of parse tree, abstract tree, syntax tree or other hierarchical structure – giving a structural representation of the input, checking for correct syntax in the process. The parsing may be preceded or followed by other steps, or these may be combined into a single step. The parser is often preceded by a separate lexical analysis , which creates tokens from the sequence of input characters. Parsers may be programmed by hand or may be automatically or semi-automatically generated by a parser generator. Parsing is complementary to templating , which produces formatted *output.*

Complicated
Expensive
Difficult to Change

DATAWATCH

Report Mining Historical Methods

Desktop Tools

- Create a template with drag-and-drop graphic tools
- Extract the data from the report using the previously created template
- Transform the report into Excel
- Use Excel for analysis
  - Perform calculations, add additional fields, make predictions
- Distribute at will

Datawatch Modeler

The Product Formerly Known as ***Monarch***

DATAWATCH

Report Mining Historical Methods

Desktop Tools

Datawatch Modeler

The Product Formerly Known as ***Monarch***

- *Multiple users may use different methods*
- *Multiple users may arrive at different conclusions*
- *Can result in no two instances of distributed data being the same*
  - *Loss of "single source of the truth"*
- *Data Lineage in question*
- *No Governance*

DATAWATCH

Server Based Report Mining

Current State of the Art

Server-based Report Mining that is tightly integrated with the underlying content system.

- *Web-based*
- *Security*
- *Single, enterprise-wide extraction template*
- *One set of instructions*
- *Single "source of the truth"*
- *Flexibility*
- *Data Lineage maintained*

DATAWATCH

Demonstration

Demonstration Scenarios

- Retail Store Analytics
  - Input – Cash Register Receipts
  - Output – Sales Analysis by Store, Category and Brand
  - Output – Sales Analysis by Cashier
- SAP Sales Report for Bicycle Parts Distributor
  - Input SAP Sales Report
  - Output – Customer Sales Analytics by Div. Low Margin Reporting

DATAWATCH

Demonstration

REGISTER REPORT PAGE: 3

STORE : COCO PARADISE
DATE : 03/01/2012

RECEIPT : 6628
CASHIER : SHAE HINES
TIME : 06:34:18

CATEGORY	ITEM	BRAND	DESCRIPTION	QTY	PRICE	AMOUNT
CONFECTIONARY	100110386	TREVIZE	RO T30X2 BIPACK	1	35.80	35.80
CONFECTIONARY	100171542	OLIVAW	CHOCO VALUE PACK	1	50.38	50.38
					SUBTOTAL :	86.18
					LESS DISCOUNT :(	0.00)
					ADD TAXES :	6.03
					TOTAL :	92.21

RECEIPT : 6629
CASHIER : SHAE HINES
TIME : 06:37:48

CATEGORY	ITEM	BRAND	DESCRIPTION	QTY	PRICE	AMOUNT
CONFECTIONARY	100118213	OLIVAW	CHOCO COOKIENCRM DF	2	18.55	37.10
CONFECTIONARY	100118214	OLIVAW	CHOCO DF GIFTPOU	1	18.55	18.55
CONFECTIONARY	100118217	OLIVAW	CHOCO DF GIFTPOUCH	1	18.55	18.55
CONFECTIONARY	100118219	OLIVAW	DFGIFTP	1	18.23	18.23
					SUBTOTAL :	92.43
					LESS DISCOUNT :(	11.83)
					ADD TAXES :	5.64
					TOTAL :	86.24

RECEIPT : 6630
CASHIER : SHAE HINES
TIME : 06:40:11

CATEGORY	ITEM	BRAND	DESCRIPTION	QTY	PRICE	AMOUNT
CONFECTIONARY	34961	GUY GIRL	BOTE FRUITS	1	53.22	53.22
CONFECTIONARY	100109853	TODDLERONE	CHOCO DARK 6X100	2	23.09	46.18
					SUBTOTAL :	99.40
					LESS DISCOUNT :(	4.22)
					ADD TAXES :	6.67
					TOTAL :	101.85

RECEIPT : 6631
CASHIER : SHAE HINES
TIME : 06:49:31

CATEGORY	ITEM	BRAND	DESCRIPTION	QTY	PRICE	AMOUNT
CONFECTIONARY	62536	GLADIA	72 BAR 49G GODIVE	1	5.02	5.02
CONFECTIONARY	62538	GLADIA	CHOCO MILK BAR 49GC	1	5.02	5.02

DATAWATCH

Demonstration

Sales report (Copy of ZEBL0009) Output

02/03/2011
User ID: ILA
Date/Time: 02/03/2011 / 15:01:00
Report name: ZEBL0010
Page 1
SAP Client: 010

Sales Organization: CND - Canadian Sales

Material	Description	Sales Order #	Quantity	Net Value	Currency	Profit ctr	Cost of goods	Currency
Billing Date: 01/03/2011		Country: CA						
Customer: 120021	BUSINESS EQUIPMENT CENTER							
BA-8327	BEARING BALL	50013	17	340.00	USD	1000	102.00	USD
BE-2349	BB BALL BEARING	50023	18	720.00	USD	1000	144.00	USD
CR-7833	CHAINRING	50037	18	18.00	USD	1000	16.20	USD
Currency: USD		Total:	53	1,078.00			262.20	
Customer: 120021	BUSINESS EQUIPMENT CENTER	Total:	53	1,078.00			262.20	
Customer: 120022	VICTORY BIKES							
CR-7833	CHAINRING	50048	16	16.00	USD	1000	14.40	USD
DT-2377	DOWN TUBE	50059	6	24.00	USD	1000	9.60	USD
EC-R098	ROAD END CAPS	50068	6	30.00	USD	1000	6.00	USD
Currency: USD		Total:	28	70.00			30.00	
Customer: 120022	VICTORY BIKES	Total:	28	70.00			30.00	
Customer: 120023	NORSTAN BIKE HUT							
DT-2377	DOWN TUBE	50082	19	76.00	USD	1000	53.20	USD
EC-T209	TOURING END CAPS	50096	5	25.00	USD	1000	12.50	USD
BE-2908	HEADSET BALL BEARINGS	50110	14	700.00	USD	1000	630.00	USD
Currency: USD		Total:	38	801.00			695.70	
Customer: 120023	NORSTAN BIKE HUT	Total:	38	801.00			695.70	
Customer: 120024	FIRST RATE BICYCLES							
BE-2908	HEADSET BALL BEARINGS	50118	16	800.00	USD	1000	560.00	USD
BE-2349	BB BALL BEARING	50128	8	320.00	USD	1000	224.00	USD
Currency: USD		Total:	24	1,120.00			784.00	
Customer: 120024	FIRST RATE BICYCLES	Total:	24	1,120.00			784.00	
Billing Date: 01/03/2011		Total:	143	3,069.00			1,771.90	
Billing Date: 01/04/2011		Country: CA						
Customer: 120081	BLOOMINGTON MULTISPORT							
CA-5965	LL CRANKARM	50137	13	650.00	USD	1000	390.00	USD
Currency: USD		Total:	13	650.00			390.00	
Customer: 120081	BLOOMINGTON MULTISPORT	Total:	13	650.00			390.00	
Customer: 120082	ILLINOIS TREK & CLOTHING							
EC-T209	TOURING END CAPS	50147	13	65.00	USD	1000	52.00	USD

Thank you

DATAWATCH