From Insight to Action:
Analytics from Both Sides of the Brain

Michael O’Connell
Chief Analytics Officer
moconnell@tibco.com
@MichOConnell
about.me/moconnell
Both Sides of the Brain

- Fast & Slow

Insight to Action

1. Visual Analytics
2. Numerical Algorithms
3. Insight Execution

Insight to Action Case Studies

- Connected Equipment; IIoT
- Hi Tech Manufacturing

Come See the Demos

- Exhibition Hallway
Themes: Thinking Fast and Slow... from Both Sides of the Brain

System 1: Association Engine
System 2: Monitor & Control
Making Sense of the World

Insight  EVENTS  Action
Making Sense of the World – Some Key Steps

Insight • MODEL • Action
Smart Visual Analytics

Be first to insight, first to action

Analytics Apps

Build and broadcast smart analytics

Streaming Analytics

Continuous algorithmic awareness and automation

Visual analytics is like a bicycle for your business mind.
Smart Visual Analytics

Be first to insight, first to action

Visual analytics is like a bicycle for your business mind.
Visual Analytics – Interactive Spotfire visualization
Visual Analytics – Extending the Palette

Dials

Sankey

Gantt

Donut

Chord
Visual Analytics – Graph Configurations in Spotfire

Chip Contour Data Function
- Contour coloring
- Contour layers

Auto Wafer Data Function
- Auto-generate based on chip location data
- Wafer border
- Wafer reticle shot

Background Image
- Register

© Copyright 2000-2016 TIBCO Software Inc.
This document (including, without limitation, any product roadmap or statement of direction) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice. © Copyright 2000-2014 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
This document (including, without limitation, any product roadmap or statement of direction data) illustrates the planned testing, release and availability dates for TIBCO products and services. It is for informational purposes only and its contents are subject to change without notice. © Copyright 2000-2014 TIBCO Software Inc. All rights reserved. TIBCO Confidential & Proprietary Information.
Map Layers

Marker Layer

Feature Layer

Map Layer

WMS, TMS Layer

Image Layer
Map Elements

Marker Layer
- Color
- Shape
- Size
- Relative amounts
- Size

Feature Layer
- Color

Marker or Feature Layer
- Tooltip
- Labels
Examples of Spotfire Recommendations in Action
Easy dashboard setup for business users, dramatically faster creation of full-featured data analysis applications for analysts

The agile business intelligence market is growing rapidly, and as Gartner points out, the transition to business platforms that can be easily implemented and used by analysts and business users to find insights quickly—as well as by IT staff to quickly build analytics content to meet business requirements and deliver more timely business benefits. This drive for speed is about business value: accuracy and speed of interpretation for decision making, authoring, and development of data discovery applications, and task completion to enable developers to implement their ideas quickly and obtain accurate insights.

This paper describes a recommendation engine for the "TIBCO Spotfire" interactive graphical analysis system. Spotfire Recommendations makes data discovery fast and easy for both analysts and business users. The system uses metadata capping and built-in graphics taxonomy to produce a collection of interesting, sensible graphics choices applied to the data at hand. The user chooses from the suggestions and the software builds a dashboard of linked, brushable, configurative graphics with supporting data filters and graphics contexts that can be rapidly deployed to the canvas.

Dashboards in Spotfire

Do we have enough shelters for the homeless?

- Homeless Count: 610,042
- Bed Count: 730,376
- Bed Utilization: 84%
- CoC Count: 426
- Bed Deficit CoC Count: 108
- Bed Deficit CoC %: 25%

Trends in Homeless & Bed Totals

Top 15 States and Territories by Bed Utilization

- Greater than 1.0
- Between 0.9 and 1.0
- All other values

Search by County in Cross Tab:
Dashboards in Spotfire

Do we have enough shelters for the homeless?

- Homeless Count: 19,029
- Bed Count: 29,800
- Bed Utilization: 64%

- CoC Count: 25
- Bed Deficit: 7
- CoC %: 28%

Bed Utilization

Trends in Homeless & Bed Totals

Trends in Beds per Shelter Type

Bed Utilization by CoC

Top 15 States and Territories by Bed Utilization
Dashboards in Spotfire

Do we have enough shelters for the homeless?

- Homeless Count: 19,029
- Bed Count: 29,800
- Bed Utilization: 64%

CoC Count: 25
- Bed Deficit: 7
- Bed Deficit %: 28%

Bed Utilization

Trends in Homeless & Bed Totals

Top 15 States and Territories by Bed Utilization
Mobile – Responsive Design

Responsive Design
- Responsive to real estate
- Laptop, Tablet, Phone

Deployment Kit
- White label apps
#2. Numerical Models – What’s Needed

Analytics Apps
Build and broadcast smart analytics
Algorithms: Rules, Machine Learning & Optimization

- Declarative & Heuristic Rules
- SPC and Anomaly Detection
- Machine Learning
  - Supervised
  - Unsupervised
  - Gradient Boosting Machines
  - Random Forests
  - Deep Learning
- Optimization
  - Linear & Quadratic Programming
  - Genetic Algorithms
  - Process optimization
  - Capacity constraints
Algorithms: Machine Learning with MATLAB

Machine Learning finds predictive models in data without being told where to look

- **Supervised** – Solve known problems: \( y = f(X) \)
  - Build a model that predicts a condition (failure, success, ..)
  - What factors are driving failures?

- **Unsupervised** – Identify patterns, Detect anomalies \( X \) only
  - Are there new patterns or failure modes emerging?

- Easy to get started with MATLAB
  - Interactive, app-driven workflows
  - Work with business and engineering data (signal, images, financial, geospatial)
  - Deploy to IT systems or run on embedded systems
  - High quality libraries
Algorithms: Optimization with MATLAB

- Prescriptive Analytics – Support Decision-making
  - Find best solution when there are constraints on the process
  - What is the optimum allocation of resources for equipment maintenance? ...for energy production?

- Decision-making
  - Linear, Quadratic, Mixed-integer, Nonlinear

- Design
  - Nonlinear
  - Global: multistart, genetic algorithm, particle swarm, pattern search, simulated annealing

- Financial Applications
  - Portfolio Optimization, Risk Analytics, Econometrics

- Performance options
  - Multi-threaded, symbolic
  - On-demand Amazon EC2 with MATLAB Parallel Cloud
  - Compute cluster with MATLAB Distributed Computing Server

Predict and Optimize Energy production

Online Optimization of Building Energy Use
Machine Learning learns tasks using features extracted manually from data

Deep Learning learns both features and tasks directly from data

Deep learning – for image classification and computer vision

- Access to pre-trained models and datasets (eg ImageNet)
- Apps for data augmentation and labeling
- GPU for training acceleration
- High quality libraries: Autoencoders, CNNs
Modeling Yield and Quality

**Goal:** Predict Quality (e.g. Yield) as function of equipment and process attributes

- **Response:** Yield (continuous)
- **Predictors:** equipment and process attributes
  - Machines, assemblers, operators, date ranges,
  - Sensor data: pressure, temperature, …
  - Maintenance logs, control charts
  - Supplier data: electrical, chemical, physical characteristics
  - Defect inspection data
- **Big Data: many columns**
  - Wafer production: 1000 sensors * 1000 readings / sensor
  - Assembly: 1000-5000+ components in some assemblies
- **Models:** Gradient Boosting Machine works well
  - Root Cause / Fingerprints
Model: Gradient Boosting Machine

GBM Results

Predictor Importance - Effect on Yield

Predictor Interactions Summary Table

Heat Map Setup

Predictor Effect on Yield Detail

Predictor Interactions Detail
Demo
Reference Diagram

TIBCO Spotfire Server

MATLAB Production Server

HTTP(s)

MATLAB Analytics

MathWorks. MPSExtension

TIBCO Spotfire Web Player

MathWorks. MPSExtension

Spotfire Web

Spotfire Desktop

Mobile App

TIBCO Spotfire Web Player

MATLAB Analytics

HTTP(s)
Increasing Capacity and Redundancy

- TIBCO Spotfire Web Player
- MATLAB Production Server
- Load Balancer
- TIBCO Spotfire Server
- Spotfire Web
- Spotfire Desktop
- MathWorks. MPSExtension
- Mobile App
Overall Spotfire Architecture

CLIENT TIER

Spotfire Analyst
Spotfire Consumer and Business Author
Spotfire iPad App

Load Balancer

SECURITY AND ROUTING TIER

Spotfire Server

WORKER SERVICE TIER

Spotfire Web Player
Spotfire Automation Services
Spotfire Statistics Services

DATA TIER

Corporate LDAP
Spotfire Library and Audit Database
Data Sources

Schedulers and Automation Services
Web Service Clients

© Copyright 2000-2016 TIBCO Software Inc.
Streaming Analytics
Continuous algorithmic awareness and automation
The TIBCO Insight Platform

TIBCO Spotfire

MATLAB

TIBCO Streambase

```
Insight
```

MODEL

```
Action
```

```
TIBCO EMS, BW, ...
```
Streaming Analytics with Streambase

- Spotfire Connection
- Business Strategy
- Continuous Visualization
- Analytics designed by data scientists
- Automated Action

Load Data
- CSV File Reader
  - Load_ESP_Data
  - CleanData
  - Sequence
  - AddSlopeThresh
  - LiveViewPublish
- Build Time Window
- Exception Checking and Classification
  - Route Exceptions
  - Combine Output
  - Exception Reporting Output
  - Pressure Slope Exceeded
- Spotfire Automation Services
Streaming Analytics – with MATLAB injection

MATLAB Production Server
Example: Hard Drive Manufacturing

- Problem in week 17
  - Yield drops from 96% to 55%
  - Production reduced from 70K to 3K drives

- Machine Learning Model
- Parameter linked to head is primary culprit
- Publish Model to Event Server to monitor
Example: Hard Disk Manufacturing

- Data Refresh
- Thresholds
- GBM Model published from MATLAB to Streambase
- GBM Model Scores Data
- Notifications
- Interventions
- $10+ MM ROI generated

© Copyright 2000-2016 TIBCO Software Inc.
Community Wiki & Exchange

https://community.tibco.com/exchange

Extend Spotfire

IronPython scripts
C# Extensions
Automation Services

Data Access and Data Wrangling with TIBCO Spotfire®

Dashboard Design Best Practices Examples

Color Schemes

For instructions visit the following dedicated pages:

- Custom Design for customer specific Schemes
- C# Extensions
- JavaScript API
- IronPython Scripts
- Spotfire in web applications

Getting Started
Data Access and Wrangling
Visualizations
Maps
Analysis
Administration
Partners
Learn

HTML in Text Areas

Using HTML in Spotfire Text Areas is a great way to enhance the look and feel of your dashboard.
Contact

Thank you!

Michael O’Connell, PhD
Chief Analytics Officer
TIBCO Fellow
moconnell@tibco.com
@MichOConnell
+1-919-7401560