

Big Data and Quality Management: A Glimpse Into the Future

*A Presentation for MNASQ
by Charles A. Liedtke, Ph.D.*

Strategic Improvement Systems

Strategy / Quality / Analytics / Improvement / Innovation

Excelsior, Minnesota USA

caliedtke@aol.com

www.strategicimprovementsystems.com

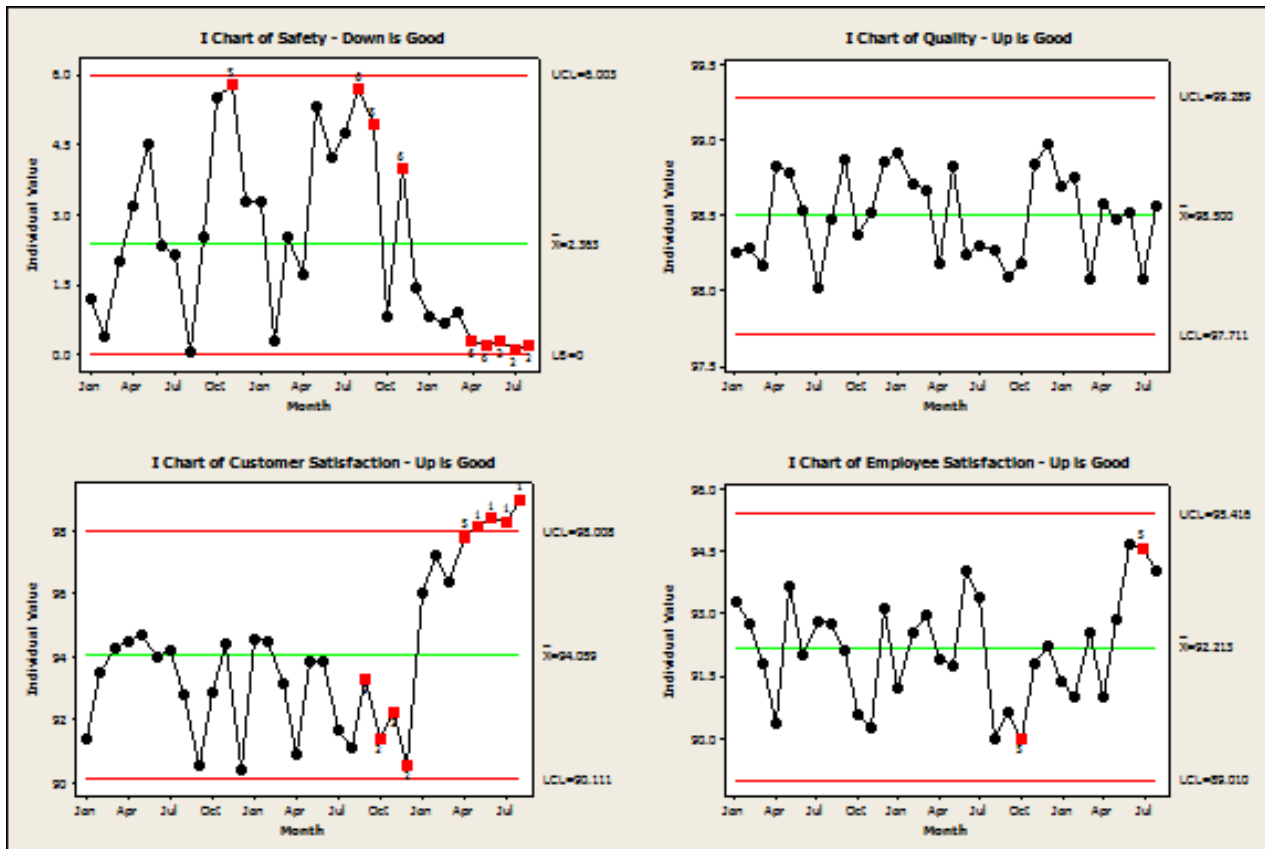
March 10, 2015



Context = “Performance”

There has been a significant increase in the use of performance scorecards/dashboards across most industries in the U.S.A.

Safety



Quality

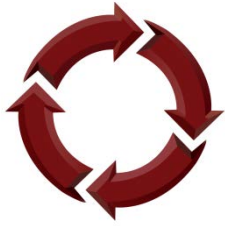


Customer Satisfaction



Employee Satisfaction





Some Improvement Approaches



Scientific
Management

Shewhart's
Control Chart

TQC
Dr. Feigenbaum

Japanese
TQM

Deming Prize

Process
Management

Deming's System
of Profound
Knowledge

QC Circles

Workout

Juran
Trilogy

BPR

Big Data

Balanced
Scorecard

ISO

Human Factors

DFSS

Lean (TPS)

Baldrige
Award

Reliability

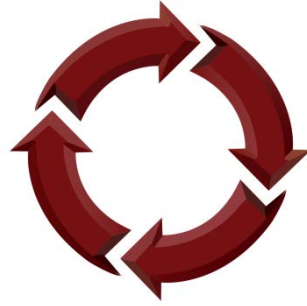
Theory of
Constraints

U.S.A.
TQM

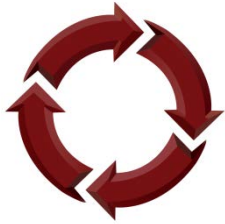
Six Sigma

Zero Defects

TPM



*Research on
Analytical Trends in the U.S.
(on-going)*



Selected References



2015-2015 Baldrige Excellence Framework by NIST

A Nation Transformed by Information by Chandler & Cortada (Eds.)

Big Data: A Revolution That Will Transform How We Live, Work, and Think by Mayer-Schonberger & Cukier.

Big Data @ Work: Dispelling the Myths, Uncovering the Opportunities by Davenport

Business Analytics: Data Analysis and Decision Making (5th Ed.) by Albright & Winston

Developing Analytic Talent: Becoming a Data Scientist by Granville

How Google Works by Schmidt & Rosenberg

Keeping Up with the Quants by Davenport & Kim

Nine Algorithms That Changed the Future by MacCormick & Bishop

Predictive Analytics: The Power To Predict Who Will Click, Buy, Lie, or Die by Siegel

Telematics and Big Data Analysis in Komatsu – Presentation by H. Asada of Komatsu at the 2014 International Conference on Quality

The Black Box Society by Pasquale

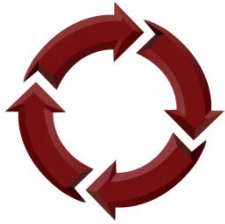
The Facebook Effect by Kirkpatrick

The Numbers Game: Baseball's Lifelong Fascination With Statistics by Schwarz

The Signal and the Noise by Silver

Turing's Cathedral: The Origins of the Digital Universe by Dyson

Uncharted: Big Data as a Lens on Human Culture by Aiden & Michel



Baldrige Award & Big Data



Baldrige Framework – Primary Questions

Is your organization doing as well as it could?

How do you know?

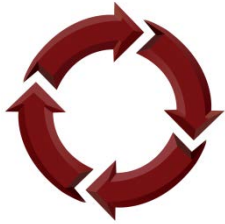
What and how should your organization improve or change?

Core Values and Concepts: Management by Fact

Category 4: Measurement, Analysis, and Knowledge Management

Changes from the 2013-2014 Criteria, Page 45

“Big data. For all organizations, turning data into knowledge and knowledge into useful strategic insights is the real challenge of big data. . . . In 2015, the Criteria incorporate an enhanced focus on data analytics, data integrity, and cybersecurity.”



American Statistician



American Statistician, Volume 68, Number 3, 2014

*A Coverage Probability Approach to Finding an Optimal Binomial
Confidence Procedure*

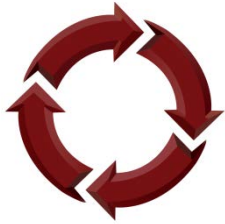
*New Statistical Tests for Detecting Disparate Impact Arising From
Two-Stage Selection Processes*

A Simple Density-Based Empirical Likelihood Ratio Test for Independence

*On a Simple Construction of a Bivariate Probability Function With a
Common Marginal Applying Bootstrap Methods to System Reliability*

Note: Recent journal article titles from the *American Statistician*.

Recent Research



Journal of Quality Technology



Journal of Quality Technology, Volume 46, Number 4, October 2014

Fractional Brownian Fields for Response Surface Metamodeling

A Robust Estimator for Location in Phase I Based on an EWMA Chart

A Gaussian Process Control Chart for Monitoring Autocorrelated Process Data

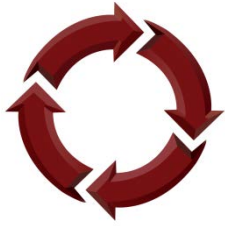
Parametric Yield Modeling Using Hidden Variable Logistic Regression

CUSUM Charts for Monitoring the Characteristic Life of Censored Weibull Lifetimes

Statistical Surface Monitoring by Spatial-Structure Modeling

Note: Recent journal article titles from the *Journal of Quality Technology*.

Recent Research



Quality Engineering

Quality Engineering, Volume 26, Number 4, October 2014

A Simplex Evolutionary Operation for Mixture Production Processes

The Best Location for Speed Bump Installation Using Taguchi and Classical Design of Experiments

Process Optimization of a Superfinishing Machine through Experimental Design and Mixed Response Surface Models

Using Statistical Methods to Assess a Surveillance Program

A Case Study to Select an Optimal Split-Plot Design for a Mixture-Process Experiment Based on Multiple Objectives

Analyzing Deficient Response Summaries from Designed Experiments

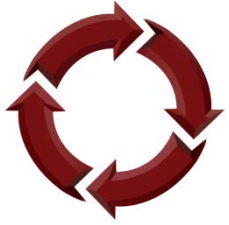
Mixed Acceptance Sampling Plans for Product Inspection Using Process Capability Index

A Large-Sample Confidence Interval for the Inverse Prediction of Quantile Differences in Logistic Regression for Two Independent Tests

A Graphical Method to Assist Quality Decisions throughout the Product Development Process

Recent Research

Note: Recent journal article titles from *Quality Engineering*.



Technometrics



Technometrics, Volume 56, Number 3, August 2014

Case-Deletion Diagnostics for Linear Mixed Models

Pointwise and Simultaneous Tolerance Limits Under Logistic Regression

A Test of Stationarity for Textured Images

The Inverse Gaussian Process as a Degradation Model

A Piecewise Single-Index Model for Dimension Reduction

Statistical Process Control Using a Dynamic Sampling Scheme

Time-Between-Event Control Charts for Sampling Inspection

Modeling Conditional Distributions for Functional Responses

Partition-Based Priors and Multiple Event Censoring

Surrogate Modeling of Computer Experiments with Different Mesh Densities

*Bayesian Uncertainty Quantification for Subsurface Inversion Using a
Multiscale Hierarchical Model*

A Bayesian Nonparametric Test for Minimal Repair

Recent Research

Note: Recent journal article titles from *Technometrics*.

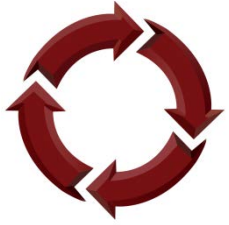


Recent Journal Summary



Very few new paths! Mainly advances in traditional *statistical methods* such as . . .

- Statistical Process Control
- Regression
- Reliability
- Design of Experiments
- Response Surface Methodology



Survey on *Data Analysis*

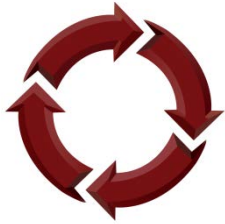


Phenomenon of Interest: “*Data Analysis Trends in the U.S.A.*”

Context: Performance Measurement & Improvement
(e.g., Quality, Safety, Speed, Productivity, etc.)

Sample Size: n = 16 organizations (8 with international operations)
8 Manufacturing; 4 Government; 3 Health Care; 1 Service

Findings: Presented at the *International Conference on Quality*,
Tokyo, Japan, October 19, 2014



Question 1 – “New”



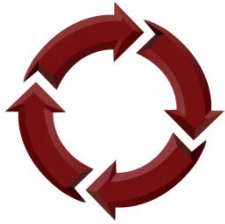
What is your organization doing that is new in terms of data analysis?

We are . . .

- establishing a data governance structure
- analyzing our overall cost structure
- implementing real-time dashboards
- analyzing customers in more depth
- being more systematic in analyzing new product and service ideas
- doing more with data visualization
- trying to be more effective at predicting performance outcomes
- studying population indicators

Governance
Real-Time
Visualization

Note: Partial summary of the responses.



Question 2 – “Changes”



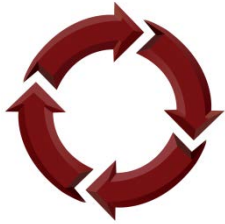
Data Quality
Prediction
Required Skill

How has data analysis changed in your organization?

We are . . .

- focusing more on the quality of our data
- holding more people accountable with our performance metrics
- using more automated real-time online dashboards
- using analytics more to predict and influence business outcomes
- tracking performance in more areas of our organization
- requiring a higher degree of statistical proof during decision-making
- integrating performance measures into the way we report and do our work
- now requiring that all leaders in the organization have quantitative skills

Note: Partial summary of the responses.



Question 3 – “Trends”



What data analysis trends are occurring in your industry?

- Everyone is moving to enterprise-wide information systems
- More use of social media metrics
- Trying to correlate positive social media reviews with performance
- Better end-to-end integration of data across customer lifecycles
- More uniformity in measuring performance (“apples to apples”)
- More public reporting and data transparency
- Higher degree of statistical proof required during decision-making
- Big Data (predictive analytics) is the new trend in health care

Note: Partial summary of the responses.

Social Media Metrics
Customer Lifecycles
Big Data



Question 4 – “Tools”



Which data analytic tools are used most in your organization?

- Statistical Process Control
- “Top Box” Scores
- Process Capability Analysis
- Logistic Regression
- Multivariate Analysis
- Linear Programming
- “What if” Scenario Modeling
- Time Series Plots
- Propensity Score Matching
- Structural Equation Modeling
- Excel, Access, R, Tableau, Minitab, etc.

Top Box
Scenario Modeling
Analytical Software

Note: Partial summary of the responses.



Question 5 – “Examples”

What is an example of how data analysis was used to improve performance?

“We utilized cluster analysis to determine which product offers were most meaningful to our customers allowing us to improve both product margin and customer experience.”

“We used a data driven method to reduce our emergency room ‘left without being seen’ metric from a high rate to an exceptionally low rate.”

“Warranty claims analysis helped identify defect frequency which led to a failure mode analysis which led to a process change which led to a reduction in warranty payments.”

“We used logistic regression to determine which attributes contributed most to the ability to resolve critical incidents within the targeted resolution time. As a result of the analysis, we were able to help a technology service provider significantly improve performance in incident resolution.”



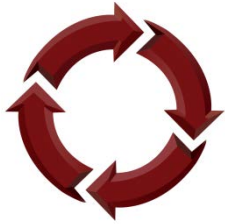
Question 6 – “Challenges”

What are the greatest challenges to effective data analysis?

- Getting accurate data
- Managing our data
- Turning our data into meaningful information
- Interpreting our data
- Telling a compelling story based on our data
- Getting people to “believe” the data—there is a lot of questioning of data
- Collecting more meaningful data and less meaningless data
- Making sure the appropriate people have access to the data
- Determining what to do about social media
- Deciding how to handle “instant feedback”
- Finding people with excellent quantitative skills
- Creating a data-oriented culture

Data Integrity
Finding Meaning
Creating a Data Culture

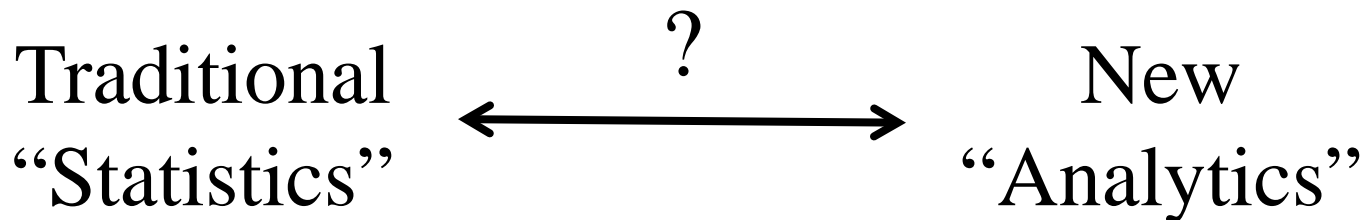
Note: Partial summary of the responses.



“Traditional” & “New”



Professional societies, such as the American Statistical Association and the American Society for Quality, are trying to determine what “Analytics” and “Big Data” mean for their professions (*existential moments*).



Sampling Plans, Basic Seven Tools,
Statistical Process Control,
Hypothesis Testing, Regression,
Design of Experiments, Response
Surface Methodology, Reliability

Devices, Data Apps, Search,
Social Media, Consent, Privacy,
Personalized Data, Real-Time
Dashboards, Data Visualization,
Government Access, Data Breaches,
Computing Power, Cloud, Big Data



Need a Broader View of *Data*



Traditional Quality Management Data Types

- Verbatim
- Categorical
- Count
- Continuous

Some More Recent Additions (More Variety, More Sources)

- Audio Recordings, Pictures, & Videos
- “Big Data”
- Tweets (Positive, Neutral, Negative, Trending)
- Search Items, On-Line Comments, Data Tables
- Website Visits, Clicks, “Eyeballs”, *Likes*, Posts, etc.



Some Data Types & Sources



Focus Group
Data

Observation
Notes

Survey
Responses

Interview
Responses

Sensor Data

Public
Records

Messages

Clicks

Searches

Personal
Health Info

Transactions

Purchases

Calls

Friends

Posted
Pictures

Connections

Posted
Videos

Books

Songs

Movies/Shows

Likes

Blogs

Tweets

Website Visits

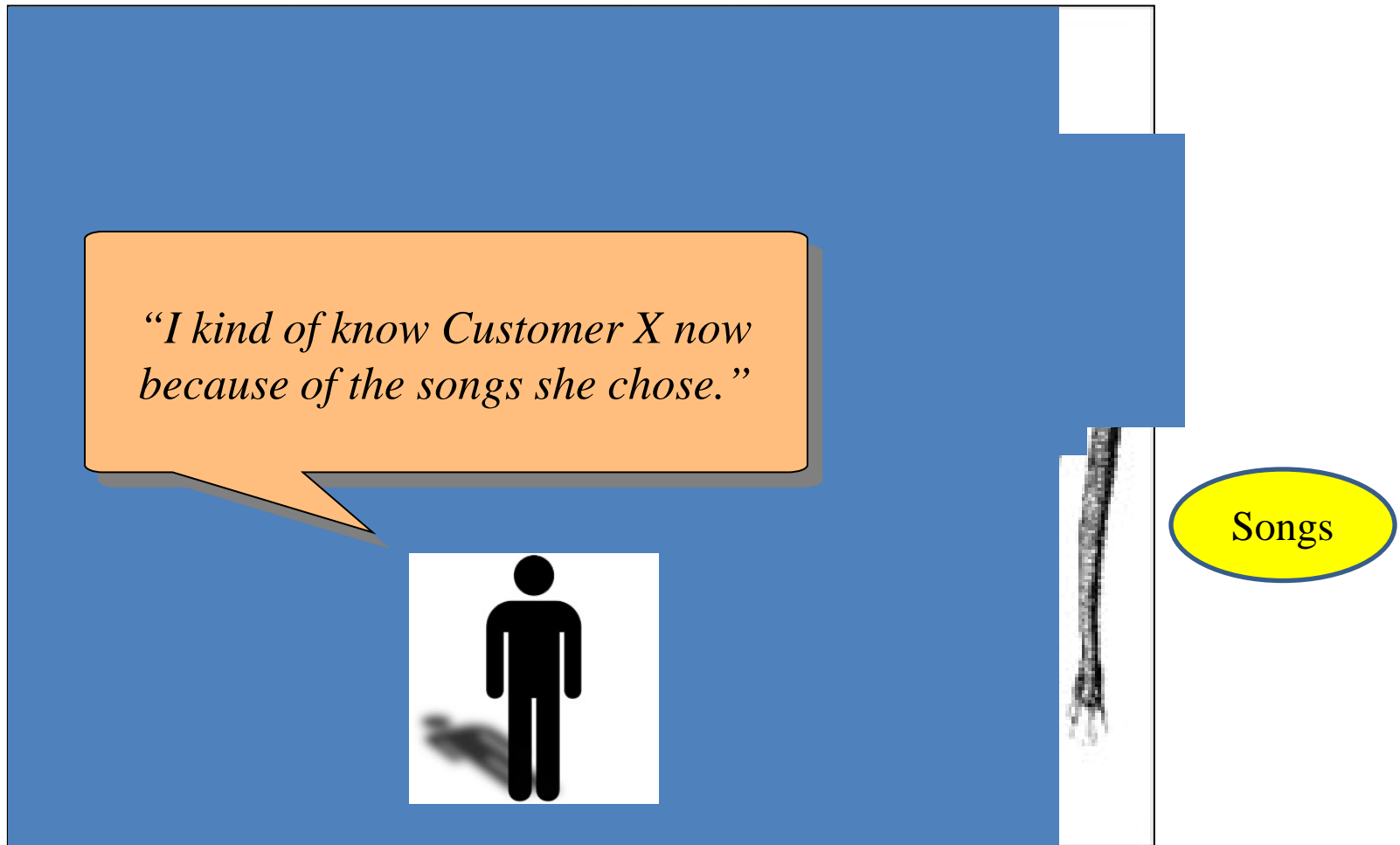
Biographical
Profiles

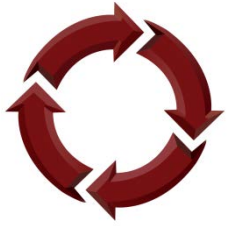
Demographic
Data

Travels/
Locations



Inferences from Song Choices





+ Inferences from Tweets



"I think I know Customer X even better now because of her Tweets."

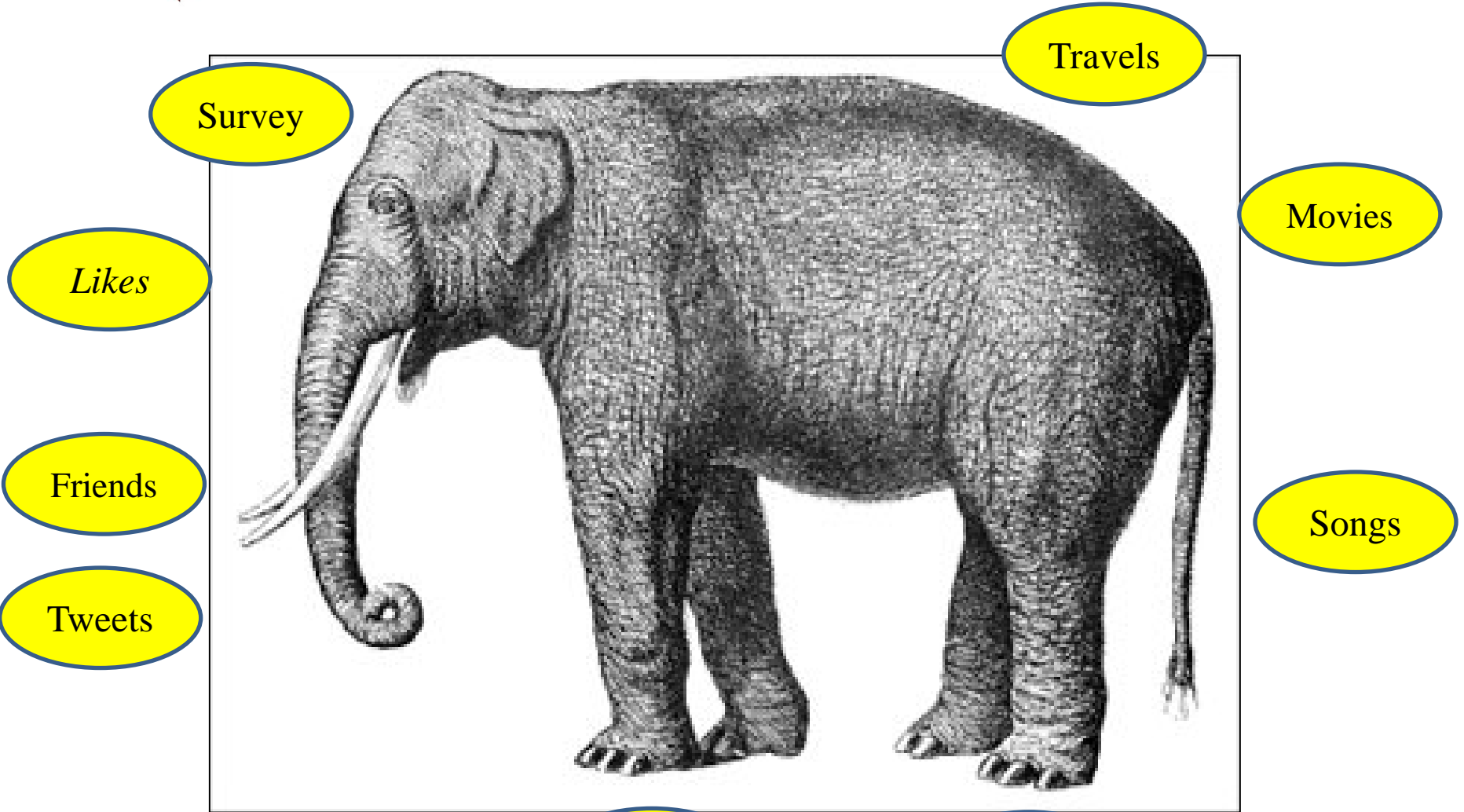


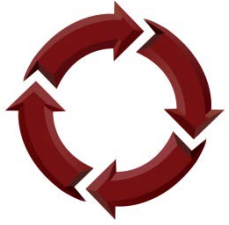
Tweets

Songs



“Oops—It is an elephant!”



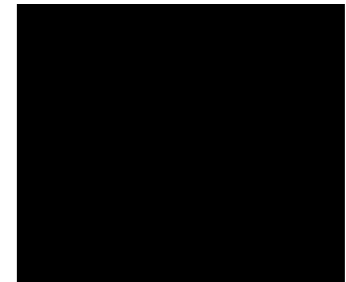


Black Box Algorithms



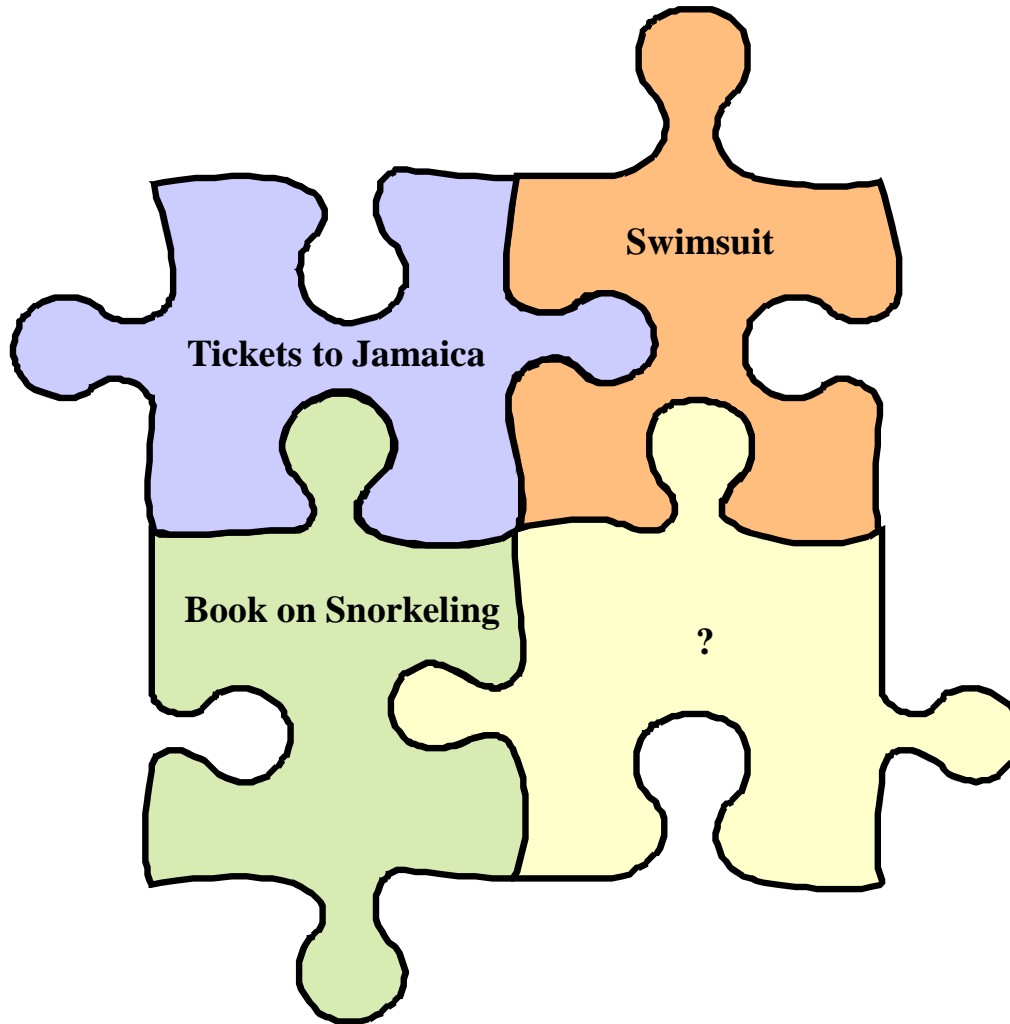
Some Black Box Algorithm Archetypes

- Preference Learning (Songs, Movies, Books, etc.)
- Search & Query
- Matching
- Discriminant Analysis
- Network Connections
- Geographic Spread
- Topic Trending
- Fill-In-The-Blank – Missing Puzzle Piece





Purchases - What Next?



**Snorkel
Mask
Fins
Sunscreen
Guide
Lessons
Etc.**



What is in common?



Born on the Bayou by CCR

Don't Fear the Reaper by Blue Oyster Cult

Fool for the City by Foghat

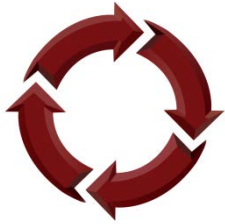
Go Your Own Way by Fleetwood Mac

Hair of the Dog by Nazareth

Never Been Any Reason by Head East

You Ain't Seen Nothing Yet by BTO

Working for the Weekend by Loverboy



Evaluate Algorithms

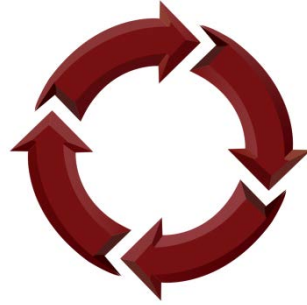


Algorithm Performance

Correct Predictions = 65%

Incorrect Predictions = 35%

		Actual Customer Behavior	
		Customer Renews Contract	Customer Does Not Renew Contract
Predicted Contract Renewal (Algorithm)	Customer Will Renew Contract	Correct 42%	Bad Surprise 21%
	Customer Will Not Renew Contract	Good Surprise 14%	Correct 23%

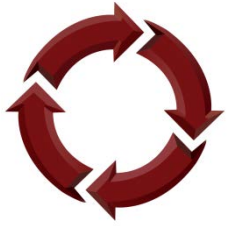


Definition of Big Data



“My company is skipping *Big Data*
and going right to *Gigantic Data*.”





Working Hypothesis

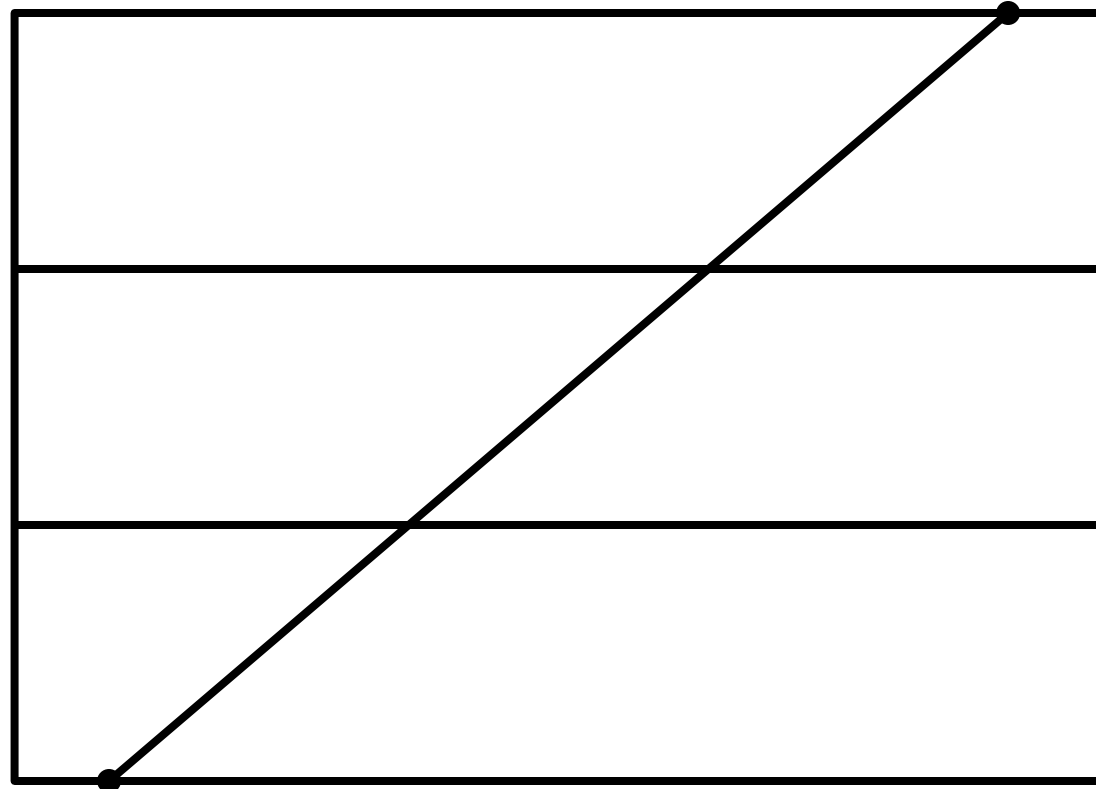


Media Attention

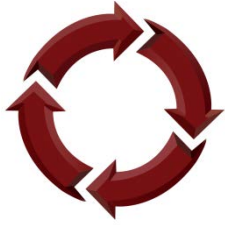
Big Data
(Exciting!)

Medium Data
(Don't Ignore!)

Small Data
(Don't Ignore!)



Application Opportunities



From *Zero Data* to *Big Data*



“*Zero Data*”

Why does our organization exist? What is our vision? What are our values?
How will we measure performance? What are our strategic priorities?
Keys to Success: Wisdom/Experience/Intuition/Passion/Insight/Judgment

Small
Data

How did our last customer feel about her experience?
What are the primary causes of the last few lost-time accidents?
Tools: Case Study Method, Key Issue Extraction, Root Cause Analysis, etc.

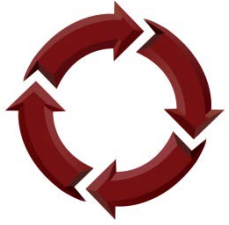
Medium
Data

What are the key drivers of *customer satisfaction* in our domestic market?
What has been our on-time delivery performance over the past three years?
Tools: Sampling, Control Chart, Regression, ANOVA, etc.

Big
Data
(Terabytes)

Who will become ill? Who will commit fraud? Which ads should we show?
Who will repurchase? What other books would people like?
Tools: Taxonomy Creation, Clustering, Cross-Correlation Analysis, etc.

Note: Make sure you understand *Zero Data* first.



Emergent Definition One

Definition 1: *Big Data* represents a situation where we have *significantly more data than usual* (e.g., a database containing 50,000,000 rows and 75 columns—terabytes).

Conclusion: This isn't a useful definition because *traditional statistical methods* can still be used and it doesn't adequately reflect what is happening in the *analytics movement*.



Emergent Definition Two

Definition 2: *Big Data* represents a situation involving a large amount of data consisting of multiple data types sometimes arriving real-time from multiple sources requiring exploratory data analysis and integrative analytical methods for problem-solving and problem-discovering.

Conclusion: This is a more useful definition consistent with what is happening in the *analytics movement* and it suggests the need for new analytical techniques and skills.

Implications: Synthesis, Hypothesis Creation, Modeling, Prediction, Security, Changing Preferences, Real-Time, Machine Learning, Data Automation, Mobile, Devices, etc.



Big Data @ Komatsu



International Conference on Quality

Hisashi Asada, Business Innovation Center

ICT Solution Division

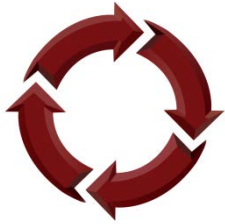
Komatsu Ltd.

Telematics and Big Data Analysis in Komatsu

October 19, 2014



- Major competitor of Caterpillar
- Continuous Enhancement of Corporate Value (the total sum of trust given to Komatsu by all stakeholders)
- Dantotsu Products, Dantotsu Service, Dantotsu Solution
- Study customer's earnings structure
- KOMTRAX System (products, communication satellite, cell phone, GPS, internet, etc.)



Under What Conditions?



“While there are plans also to build the store’s online jewelry business, some technological advances are a step too far for Ms von der Goltz.

She does not believe her buyers’ expertise can ever be replaced by that new retail obsession: big data. ‘Our buyers are editors as well as discoverers of new fashion.

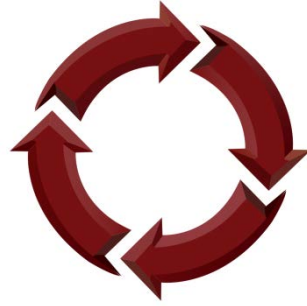
‘It’s different maybe if you’re buying for a supermarket, but when you’re talking about luxury, and especially jewelry, you have to know why a piece is going to retail at \$85,000 from its weight, its feel, to the quality of the stones used.’

Big data analysis also lacks that basic human element required in a luxury investment. ‘Jewelry is an emotional purchase,’ she says. ‘You have to get in pieces that you think will move people.’”

Financial Times, January 20, 2015

“Mastermind behind Bergdorf’s ambitious plan”

An article on Elizabeth von der Goltz, Senior VP at Bergdorf Goodman



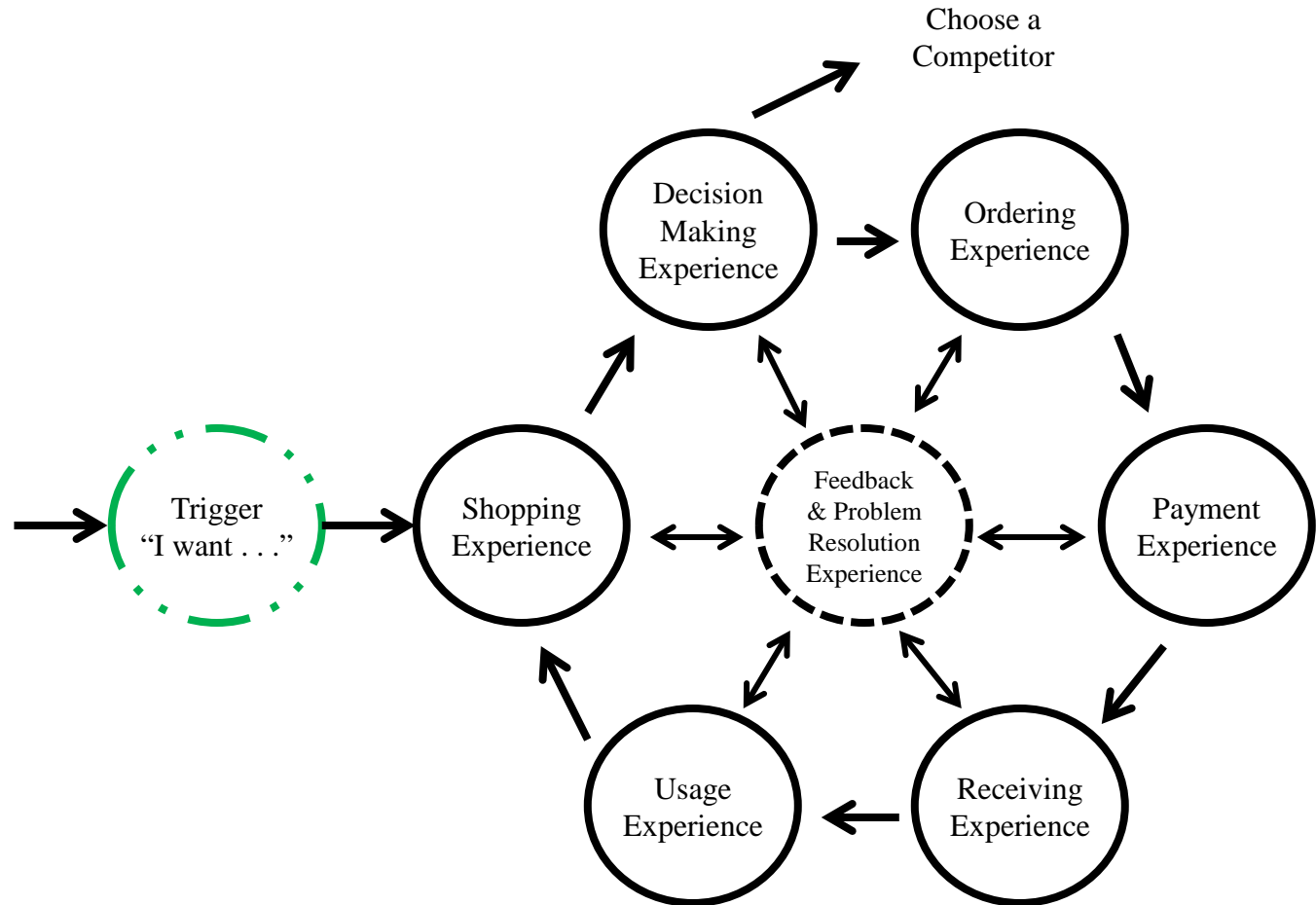
Analytical Frameworks

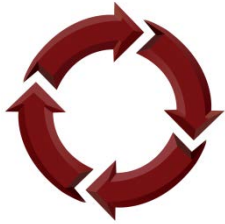


Analytics on Customer Experiences

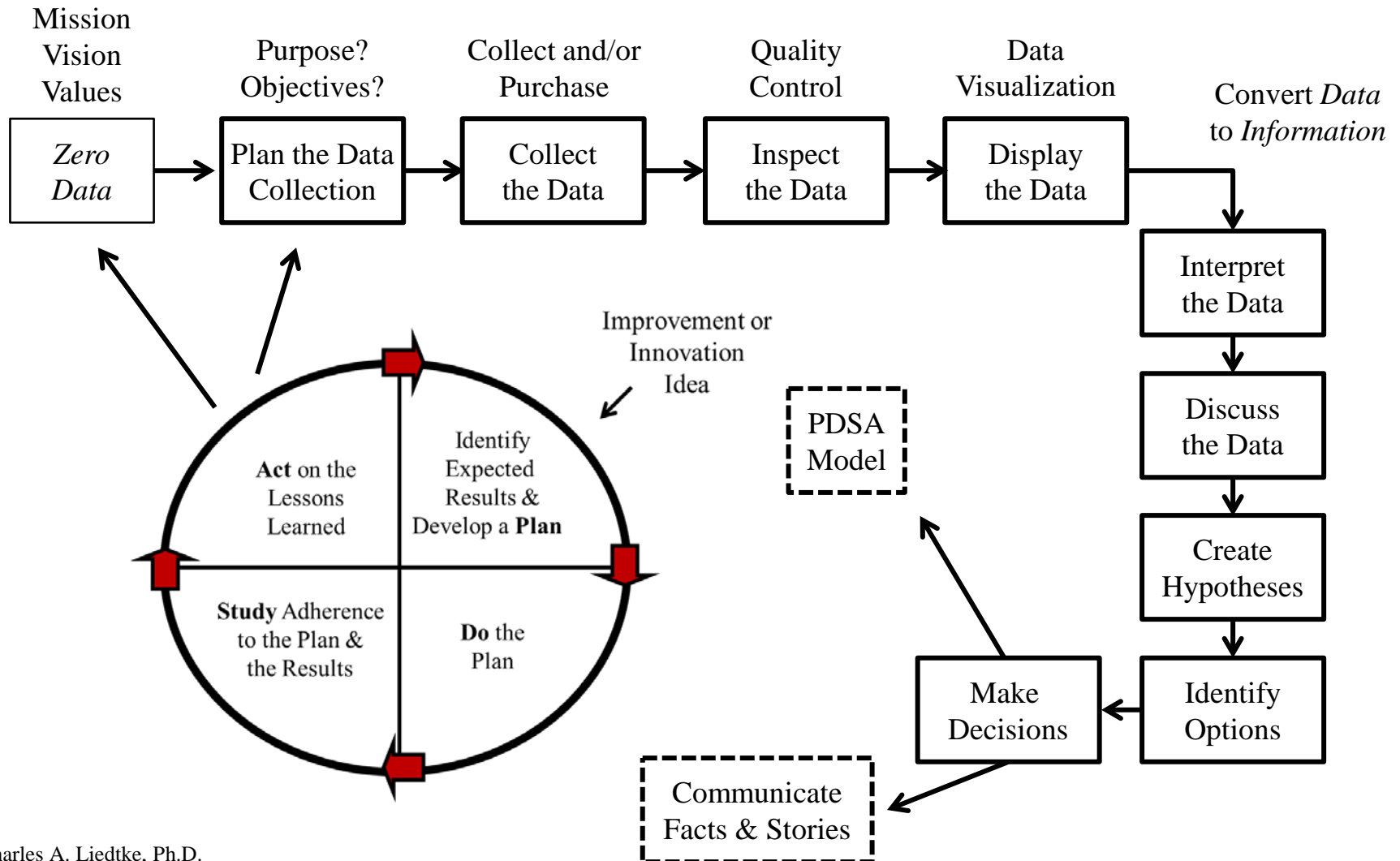


Current &
Potential
Customers



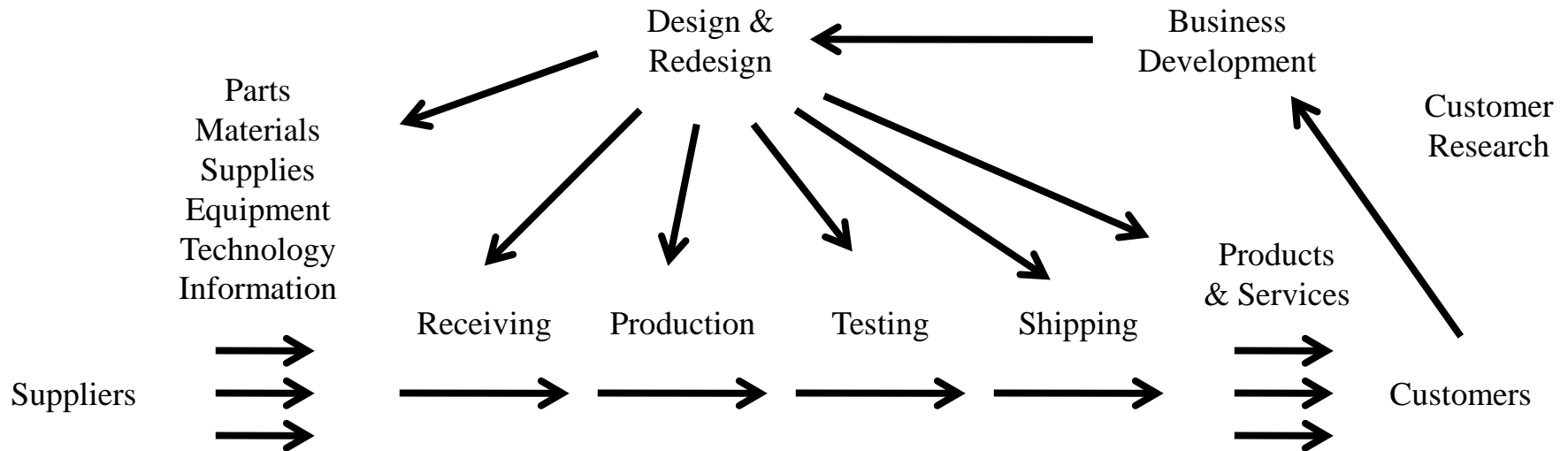


Emergent Analytical Model





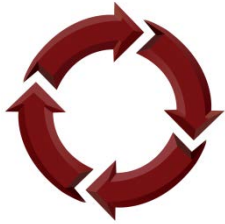
Production Viewed as a System



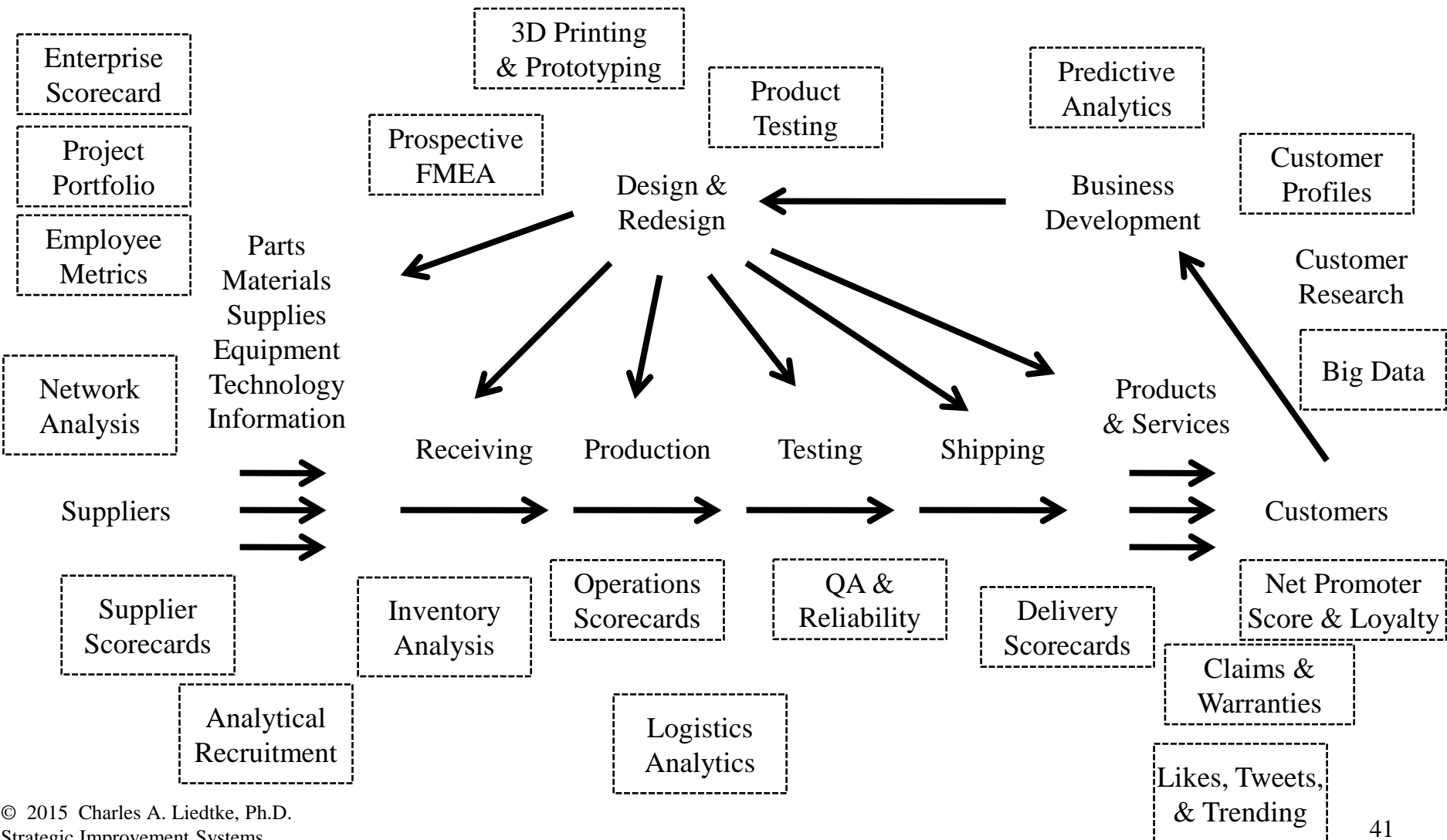
“This chart was first used in August 1950 at a conference with top management at the Hotel de Yama on Mount Hakone in Japan.”

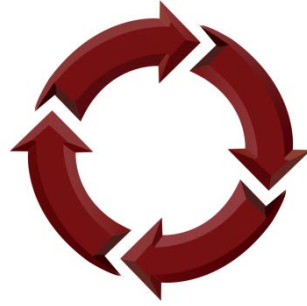
-- Excerpt from “*Out of the Crisis*” by Dr. W. Edwards Deming

Note: I have slightly modified the chart from Dr. Deming’s original form.

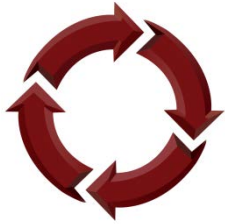


Current Analytics Activities

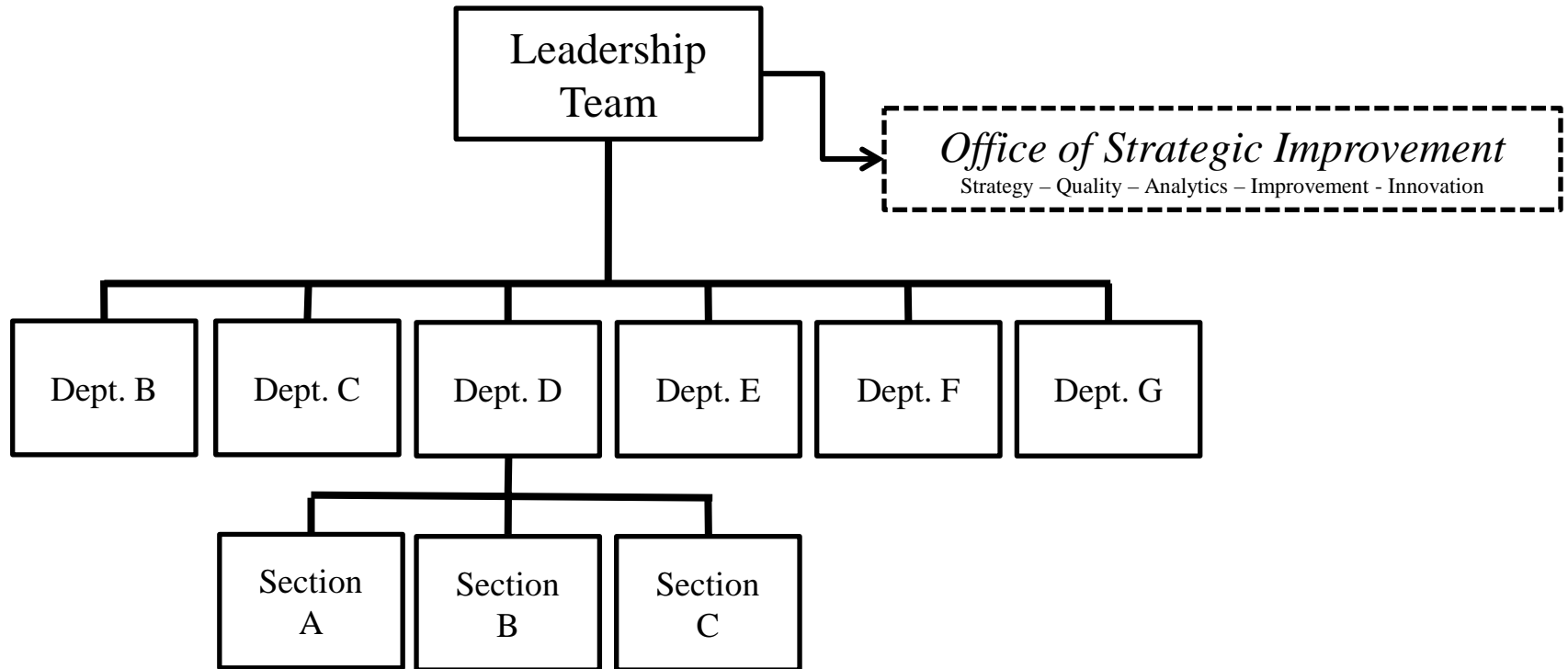




The Future



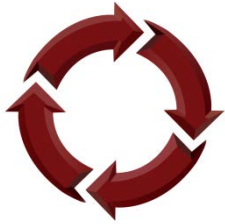
Potential Organization Chart



Alternative Titles

Office of Strategic Quality Management

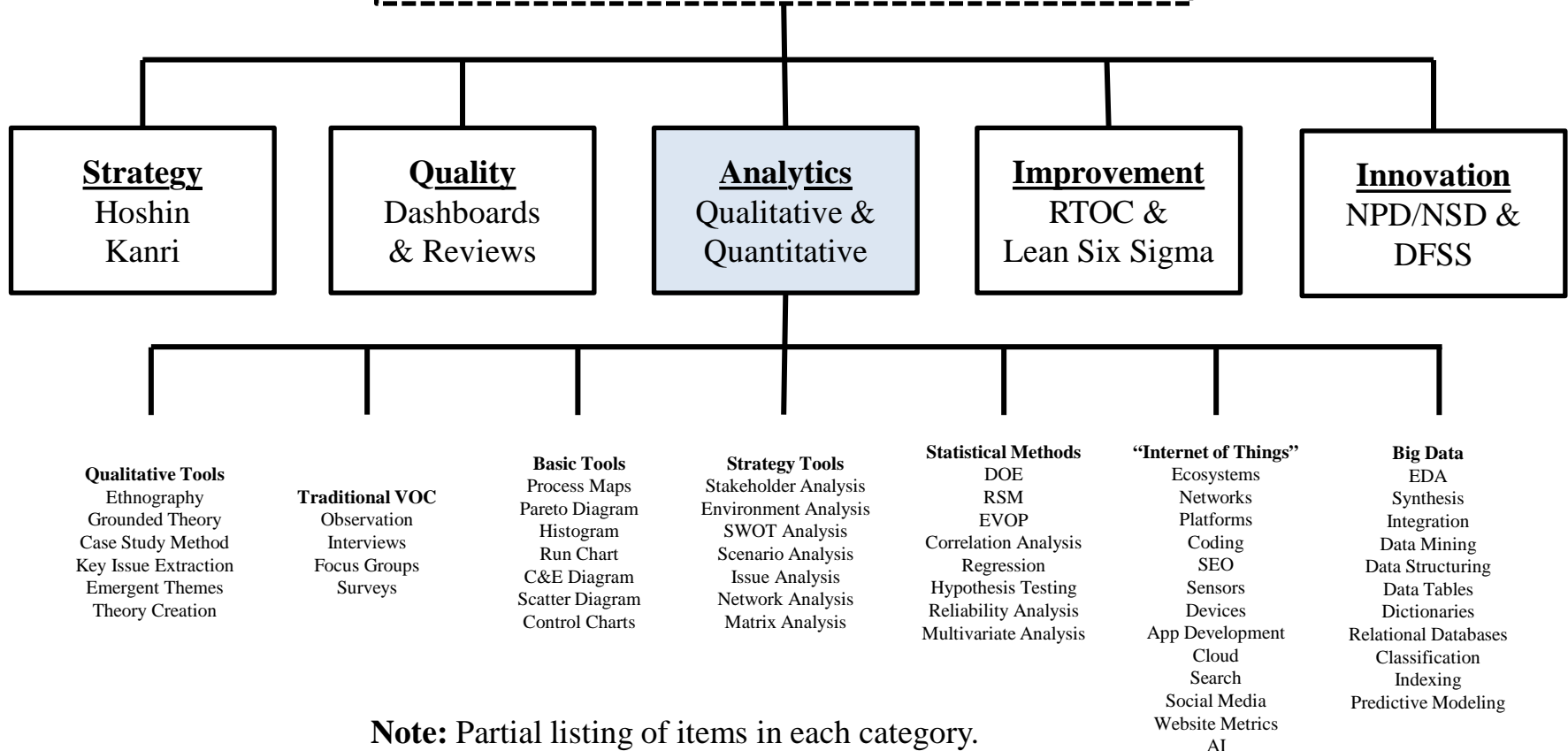
Office of Performance Excellence



Office of Strategic Improvement



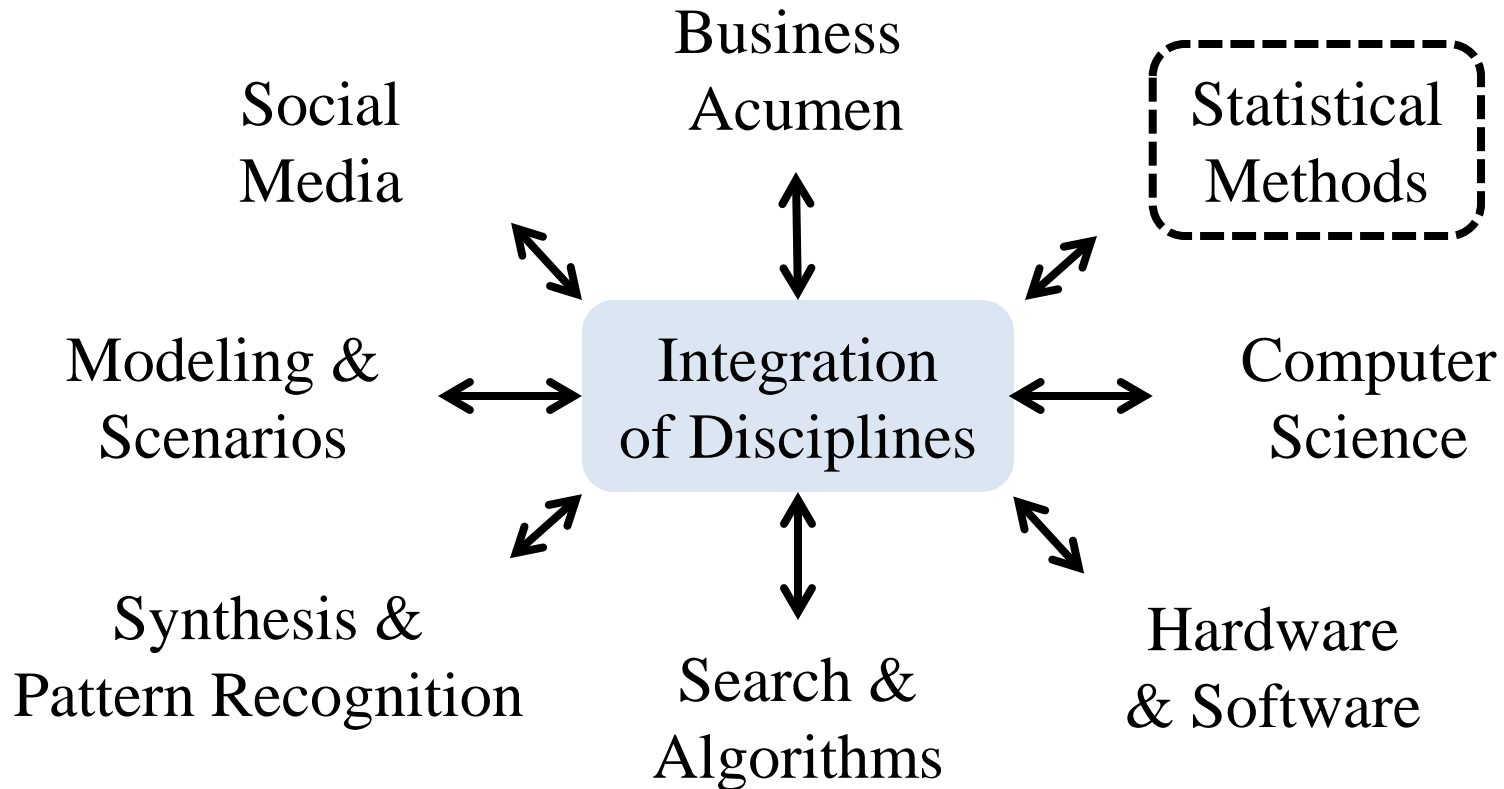
Office of Strategic Improvement
Strategy – Quality – Analytics – Improvement - Innovation



Note: Partial listing of items in each category.



Some Emerging Skill Areas



Note: *Analytics* and *Quality* professionals need at least some knowledge and skill in several areas like those depicted above.

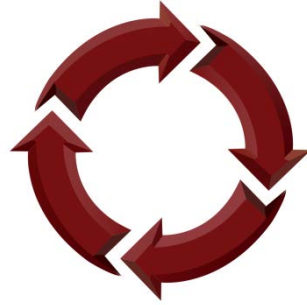


In Search of Eternal Truths



"The twentieth century has produced a world of conflicting visions, intense emotions, and unpredictable events, and the opportunities for grasping the substance of life have faded as the pace of activity has increased. Electronic media shuffle us through a myriad of experiences which would have baffled earlier generations and seem to produce in us a strange isolation from the reality of human history. Reflection is the most difficult of all our activities because we are no longer able to establish relative priorities from the multitude of sensations that engulf us. Times such as these seem to illuminate the classic expressions of eternal truths and great wisdom comes to stand out in the crowd of ordinary maxims."

- Vine Deloria, Jr.



Thank You!

Charles A. Liedtke, Ph.D.

Strategic Improvement Systems

Strategy / Quality / Analytics / Improvement / Innovation

Excelsior, Minnesota USA

caliedtke@aol.com

www.strategicimprovementsystems.com

LinkedIn Member