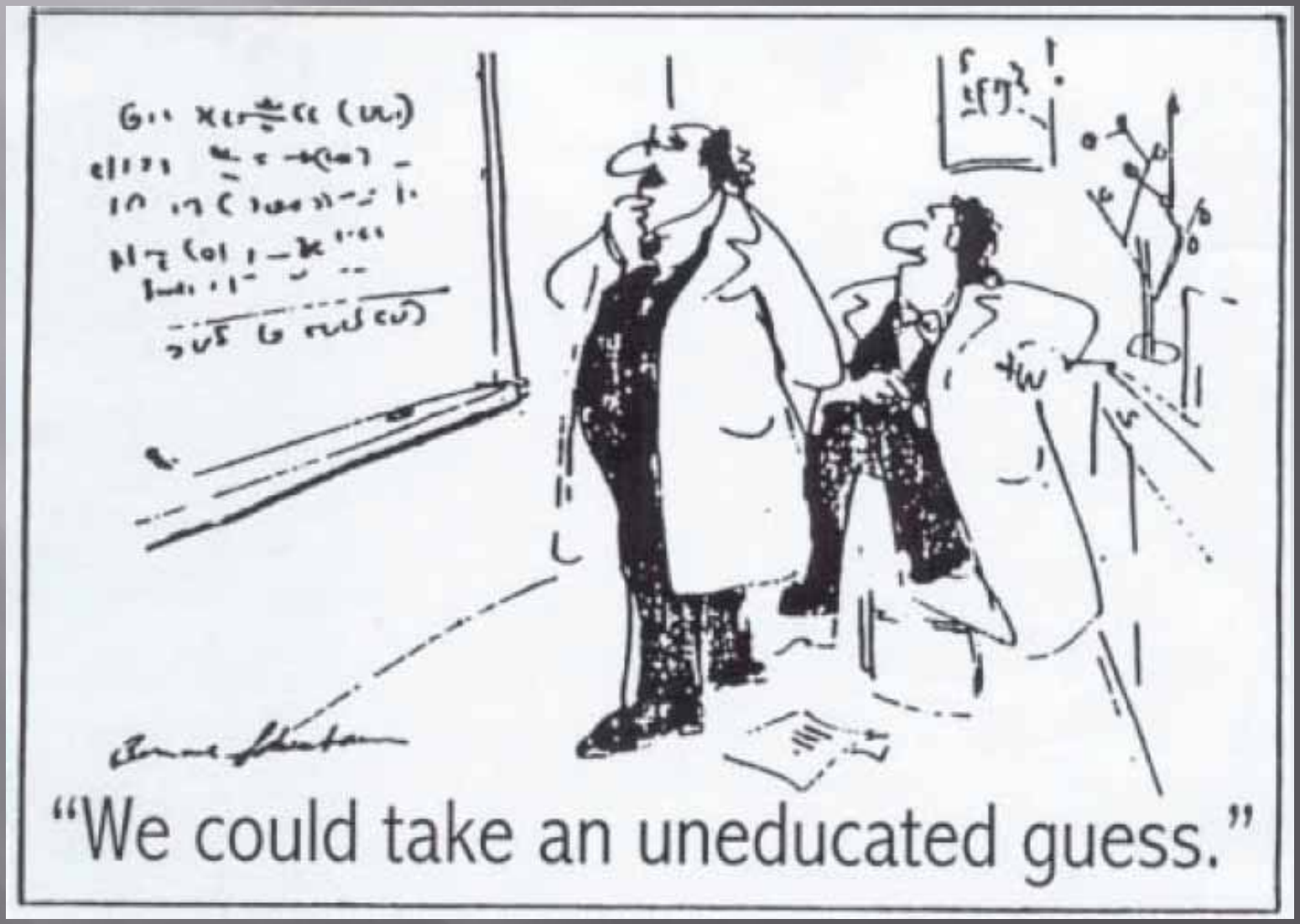
INTEGRATED DESCRIPTIVE/PREDICTIVE DATA MINING SOLUTION

Keith R Holdaway, SPE, SAS Institute Inc.

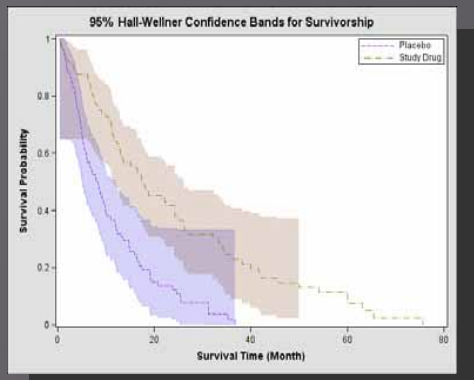
Agenda

- Introduction
- Data Cleansing, Transformation and Aggregation
- Multivariate and Exploratory Data Analysis
- Self-Organizing Maps and Artificial Neural Networks
- Genetic Algorithms , Fuzzy Logic and Predictive Models
- Conclusion

"More things in Heaven ..."



Drilling Optimization Workflows



Analytical Datamart

Data QC

EDA

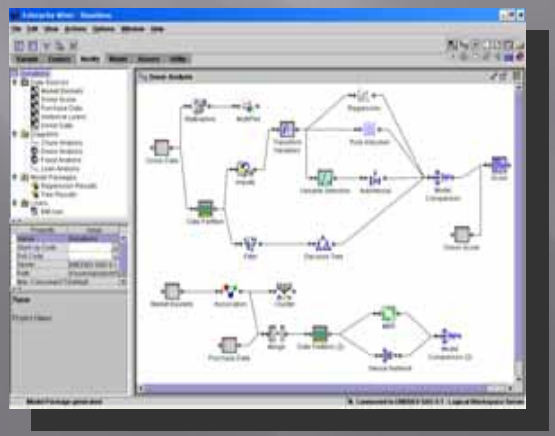
SOMs and ANNs

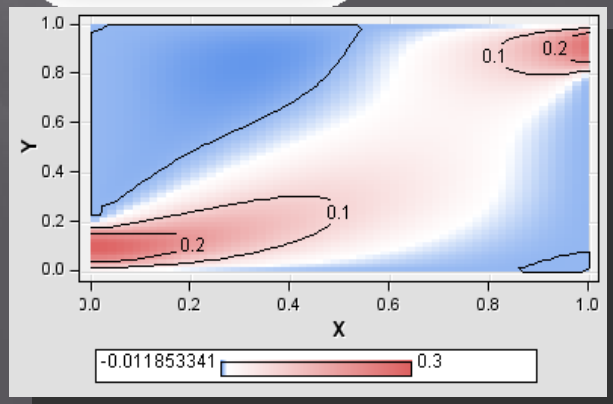
GA and Fuzzy Logic

Predictive Analytics

Analytical Datamart

Models

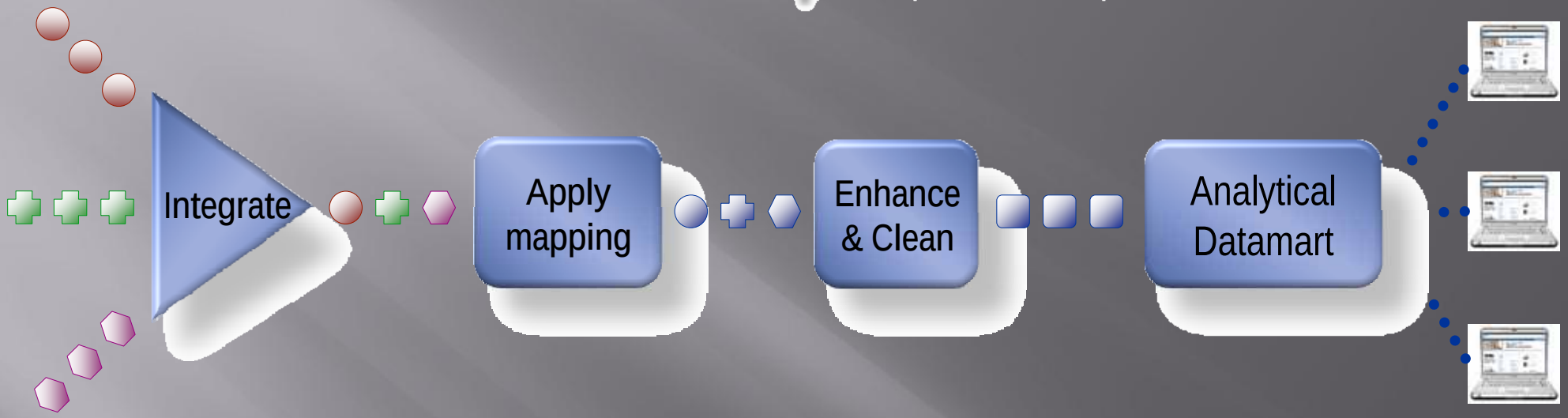




Data Cleansing, Transformation and Aggregation

Integrate, Enhance & Clean Data

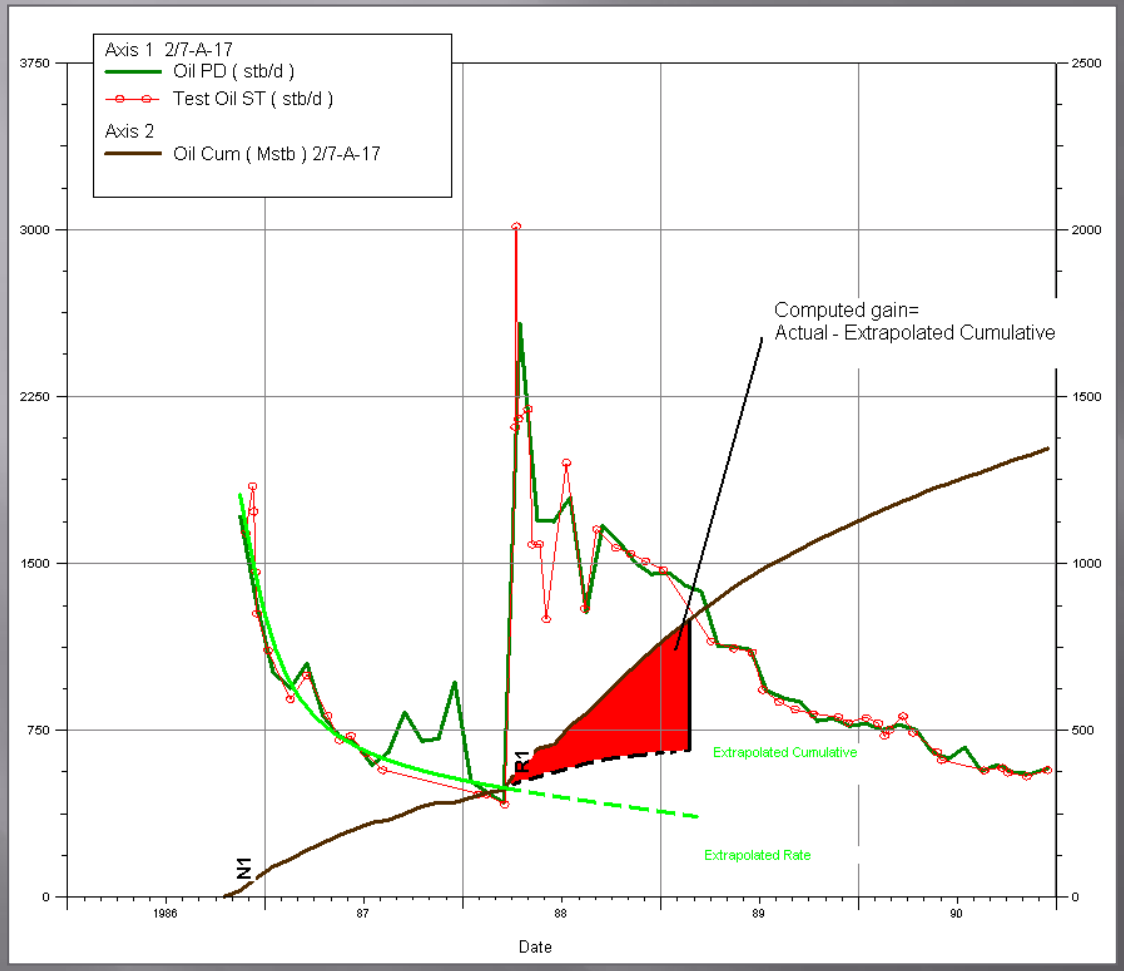
From multiple sources



Use anyone's data, from anywhere
Check data thoroughly
Format and map data correctly
Enhance data
Deliver quickly with no coding

Automation
Re-use
Consistency
Speed

Define Performance Criteria



Obtain Decline coefficient for the period before treatment:
12 months extrapolation to a “non-stimulated” scenario:
Gain computation for 3, 6 & 12 months case based on:

Cum. Oil $_{Post}$-Cum. Oil $_{Pre}$

Criteria that focus on Treatment Potential Deliverability

Segmentation using Data Mining

"Create groups of wells that have similar profiles in order to target wells as ideal candidates for stimulation."

Total SPE 62880&78333

Well Group A
Low Porosity, High Permeability, Low Production, High Pressure, Flow, etc.

Well Group B
High Porosity, Low Permeability, Low Production, Low Pressure, Flow, etc.

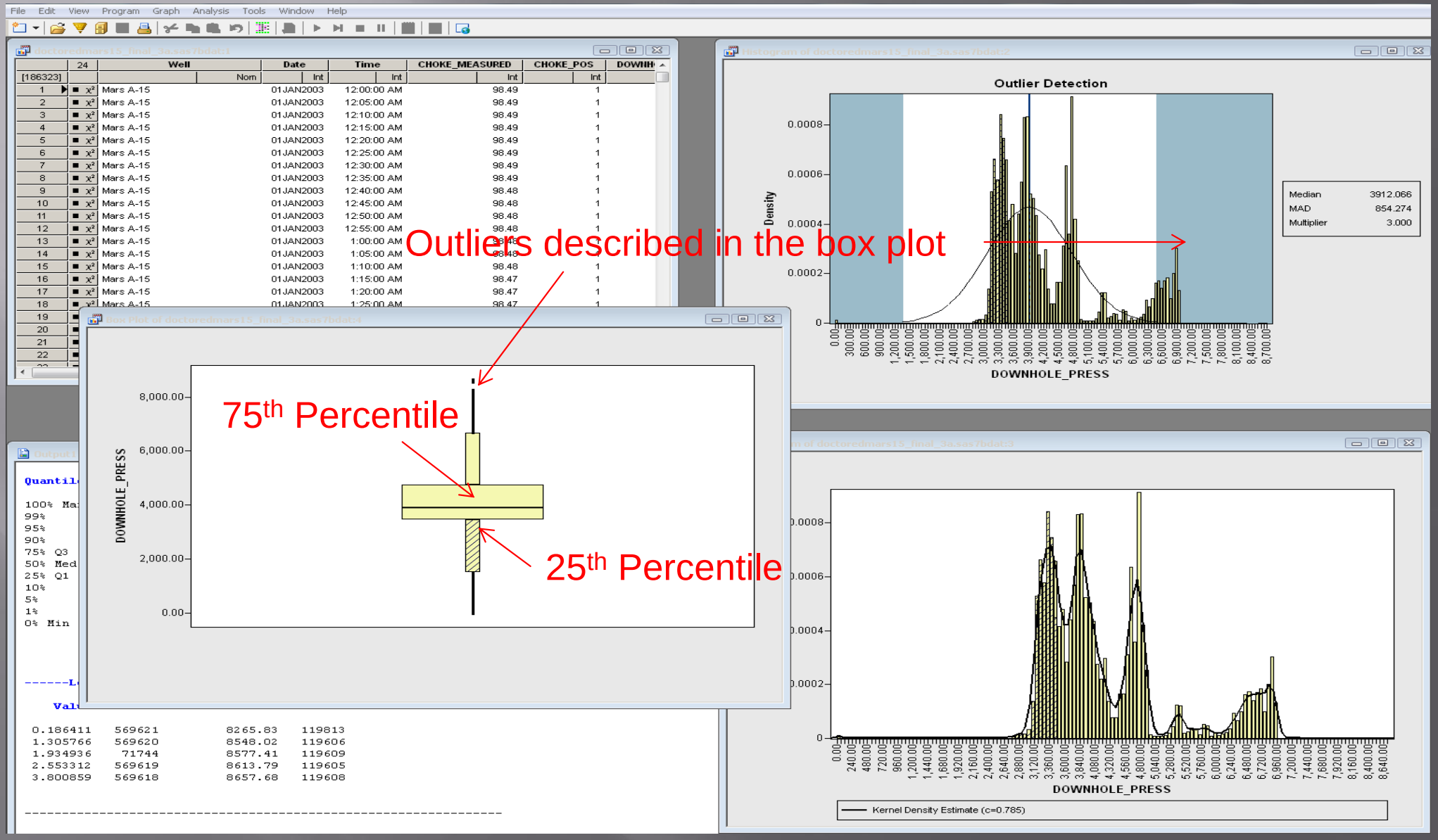
Well Group C
Low Porosity, Low Permeability, Low Production, High Pressure, Flow, etc.

Multivariate and Exploratory Data Analysis

EDA PHILOSOPHY

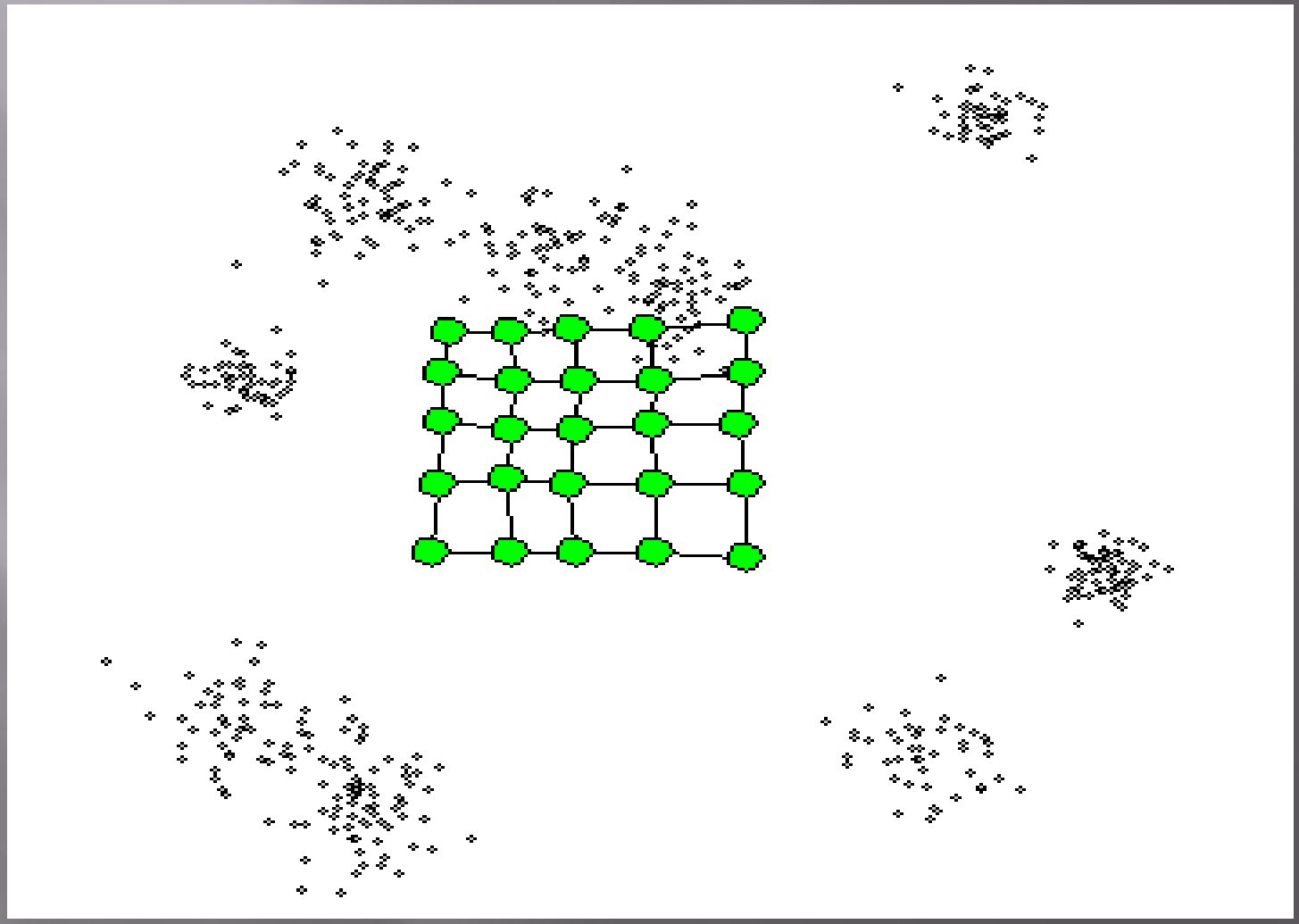
- Maximize insight into a data set
- Uncover underlying structure
- Extract important variables
- Detect outliers and anomalies
- Test underlying assumptions
- Determine optimal factor settings

Exploratory Data Analysis

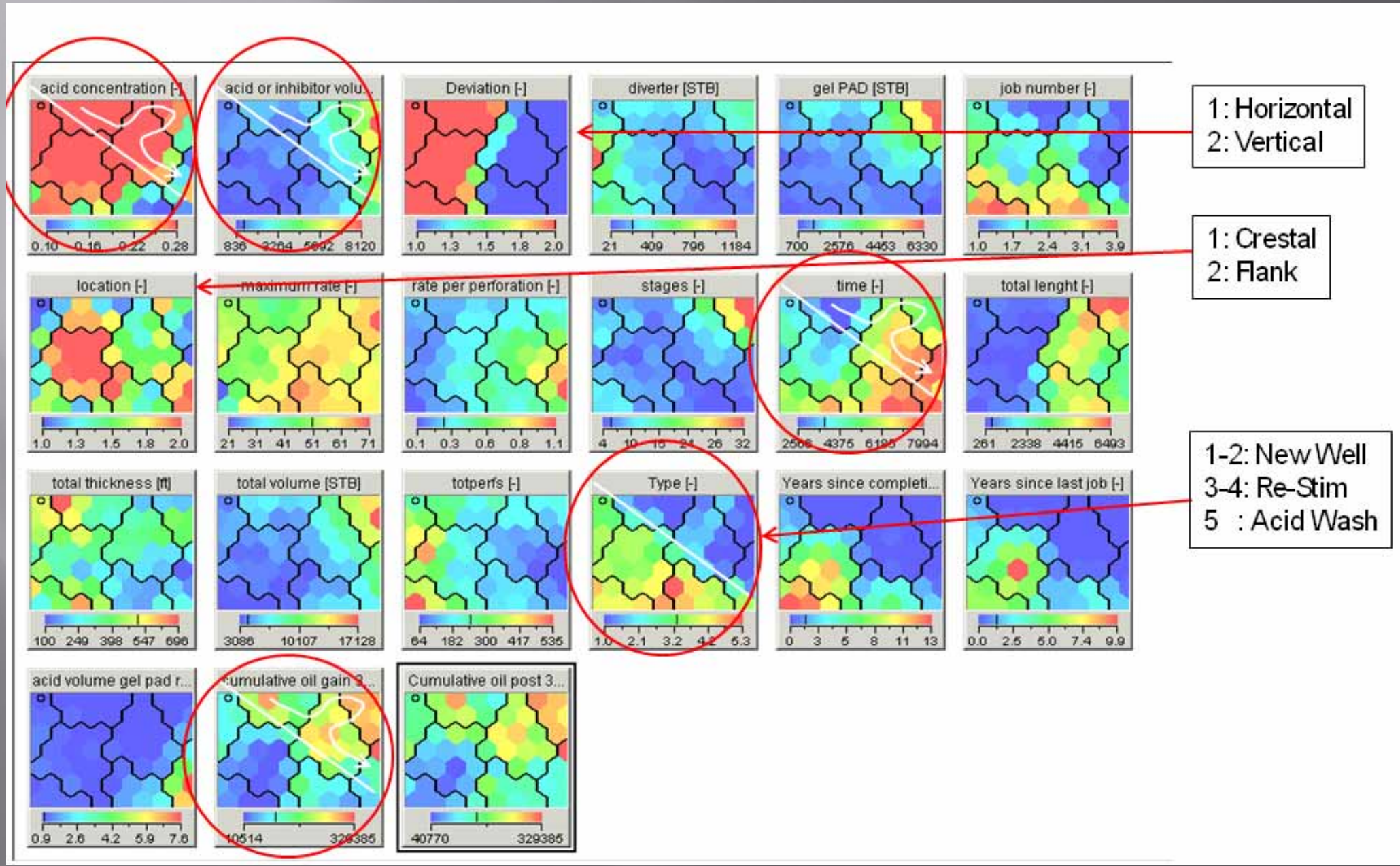


Self-Organizing Maps and Artificial Neural Networks

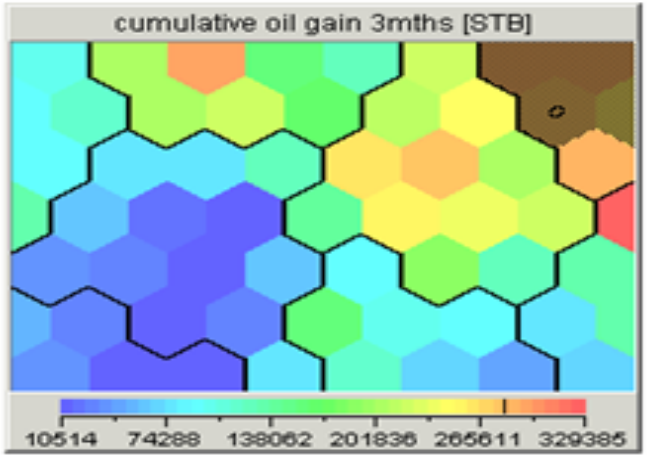
Self Organizing Maps

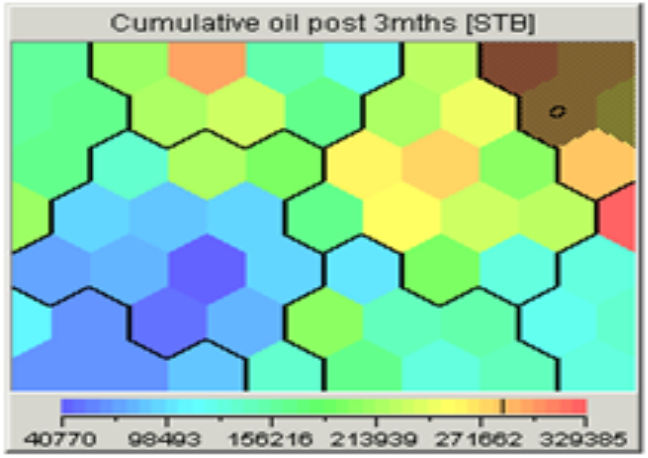


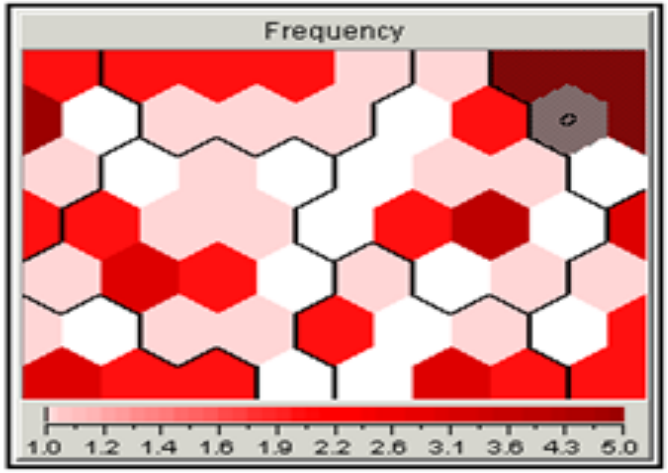
Acidization Stimulation Parameters



Critical Parameters in Horizontal Wells



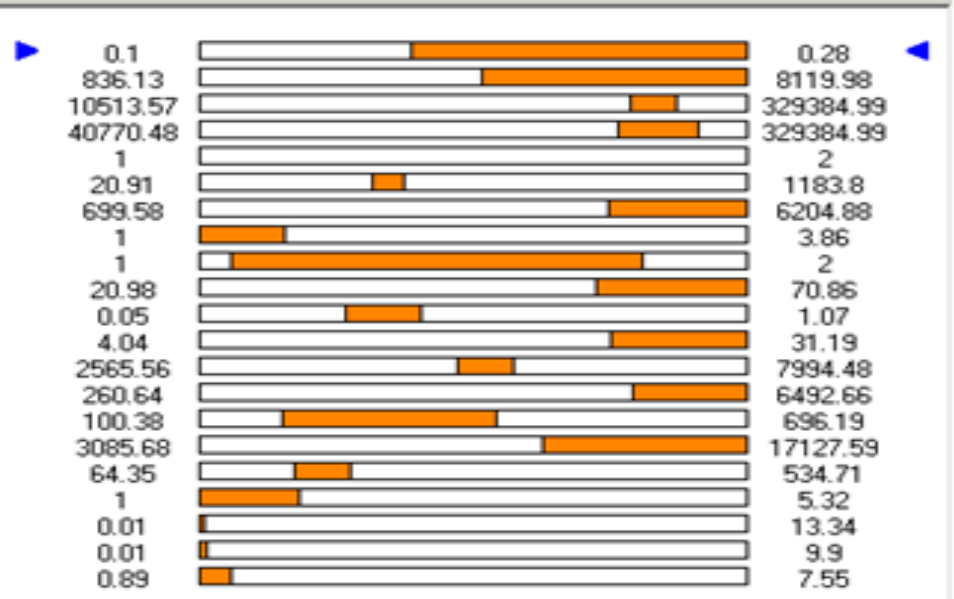




Range: Selection

Component	Value	Std. deviation	Minimum	Maximum
acid concentration [-]	0.243	0.049	0.172	0.280
acid or inhibitor volume [...	5916.424	1548.353	4559.438	8119.977
cumulative oil gain 3mths...	277831.790	12236.889	260715.648	289506.862
Cumulative oil post 3mth...	282966.097	17741.405	260715.648	304140.864
Deviation [-]	1.000	-0.000	1.000	1.000
diverter [STB]	438.921	26.453	388.417	460.298
gel PAD [STB]	5624.280	700.114	4802.060	6204.876
job number [-]	1.170	0.202	1.000	1.452
location [-]	1.426	0.312	1.056	1.807
maximum rate [-]	62.796	5.820	57.059	70.863
rate per perforation [-]	0.369	0.058	0.322	0.464
stages [-]	28.530	3.275	24.329	31.191
time [-]	5337.590	258.392	5101.667	5679.280
total lenght [-]	5990.779	483.670	5176.805	6492.659
total thickness [ft]	291.212	110.074	190.620	424.639
total volume [STB]	14156.426	2269.083	11864.371	17127.589
totperfs [-]	178.956	16.695	145.837	195.281
Type [-]	1.294	0.360	1.000	1.807
Years since completion [-]	0.084	0.082	0.018	0.202
Years since last job [-]	0.074	0.069	0.018	0.173
acid volume gel pad ratio...	1.047	0.184	0.886	1.307

Selection [acid concentration [-]] - (.105, .28, .28, .28)



Conclusions

	Best Results	Worst Results
Acid strength	17 to 28%	14 to 26%
Acid Volume	4,500 to 8,120 bbls	2,450 to 5,680 bbls
Gel Pad	4,800 to 6,200 bbls	880 to 1,900 bbls
Acid / Pad Ratio	0.9 to 1.3	1.4 to 7.5
Location	Indifferent	Flank
Stages	24 to 31	5 to 7
Cum Oil Gain (3 Months)	260k to 290k bbls	124k to 134k bbls

Artificial Neural Network

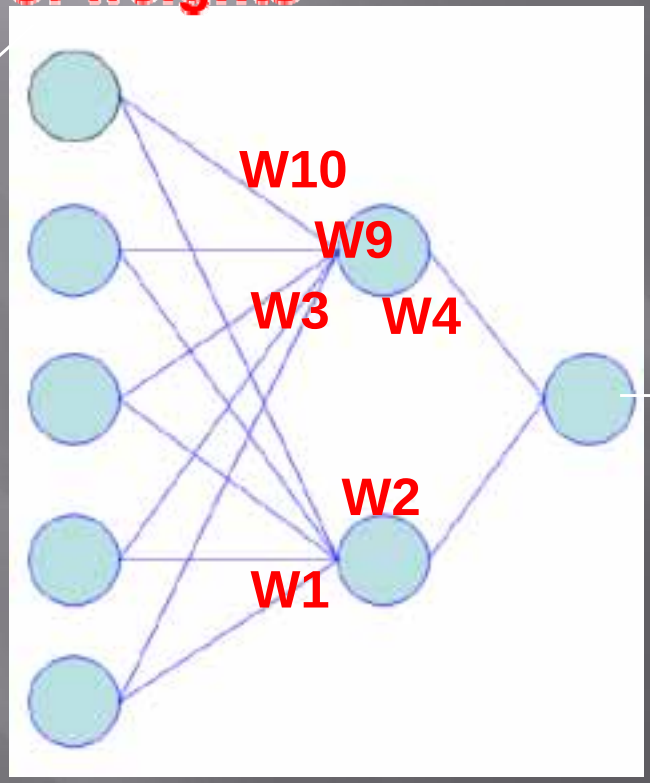
calculating error

Input Pattern

RMS $\Delta = Si - Oi$

$if\ \Delta \geq \varepsilon$: adjustment of weights

after calculating all patterns



Hidden Layer

Output Pattern

Recalculation of all patterns

Oil Production

Output Layer

Input Layer

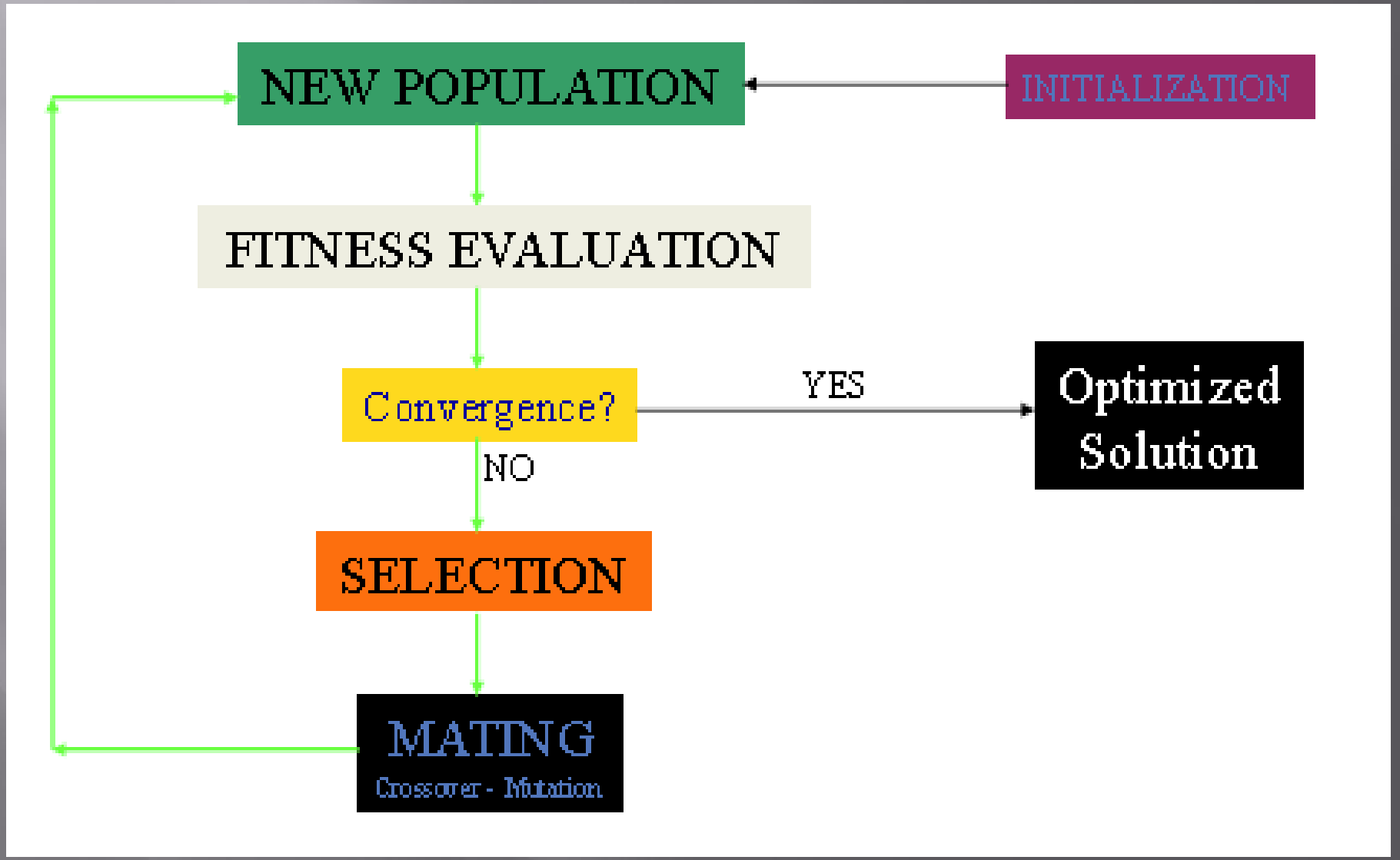
Shell SPE 135523

Tight gas well performance evaluation for hydraulic propped fracture treatment optimization in Pinedale Asset

Neural Network

Genetic Algorithms, Fuzzy Logic and Predictive Models

Genetic Algorithms



Fuzzy Logic

If the well shows a *high* potential for an increase in Cum03, Cum06 and / or Cum12 **And** has a *moderate* pressure, **And** has a *low* acid volume for the net pay completed, **Then** this well is a *good candidate* for (re)stimulation.

Fuzzy set theory uses the rigorous precision of mathematics to manage imprecision of human expression and thoughts.

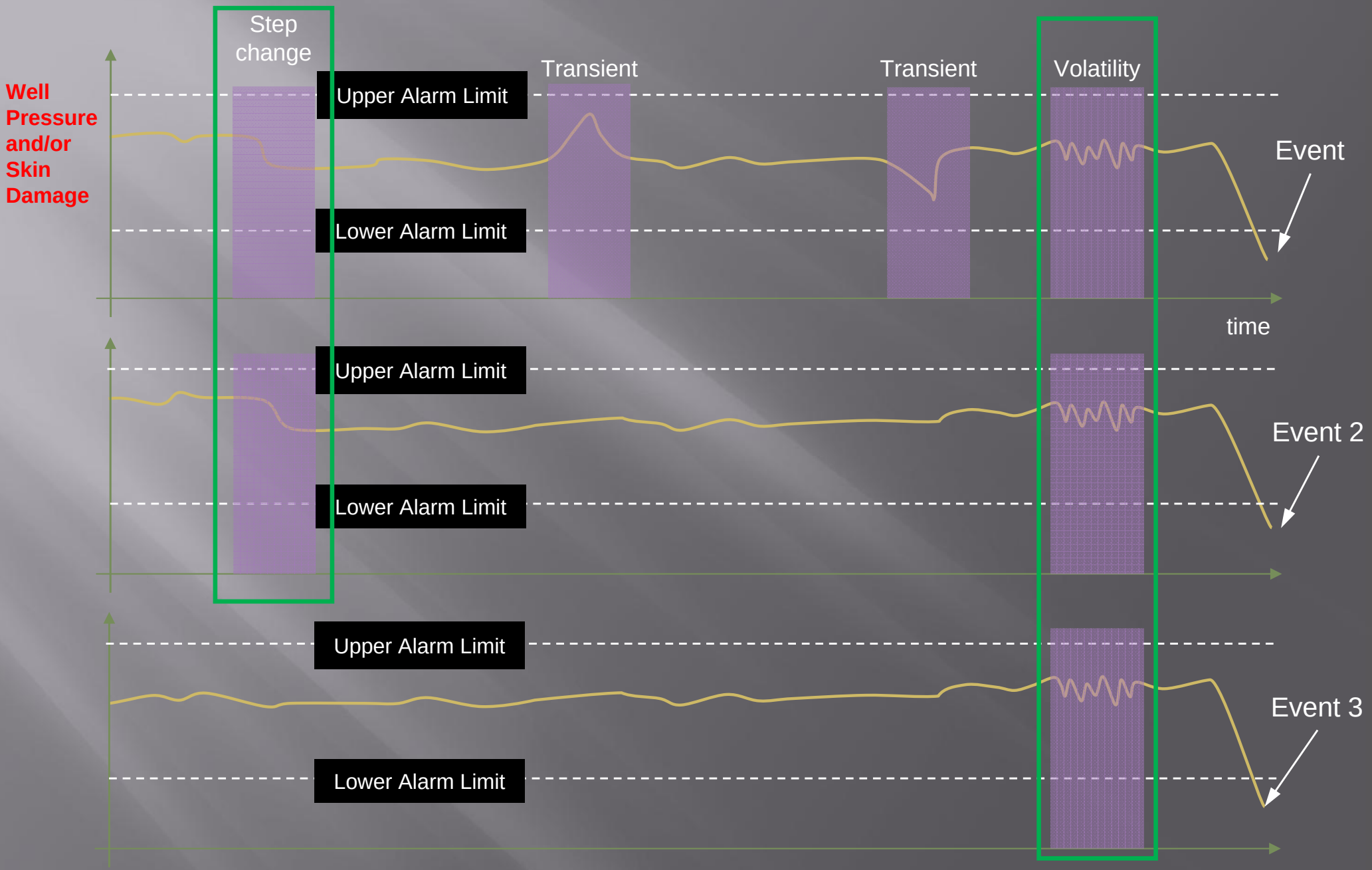
Definitions

Fuzzy logic is a tool to address uncertainty and deal with imprecision and lack of knowledge

"So far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain they do no refer to reality"

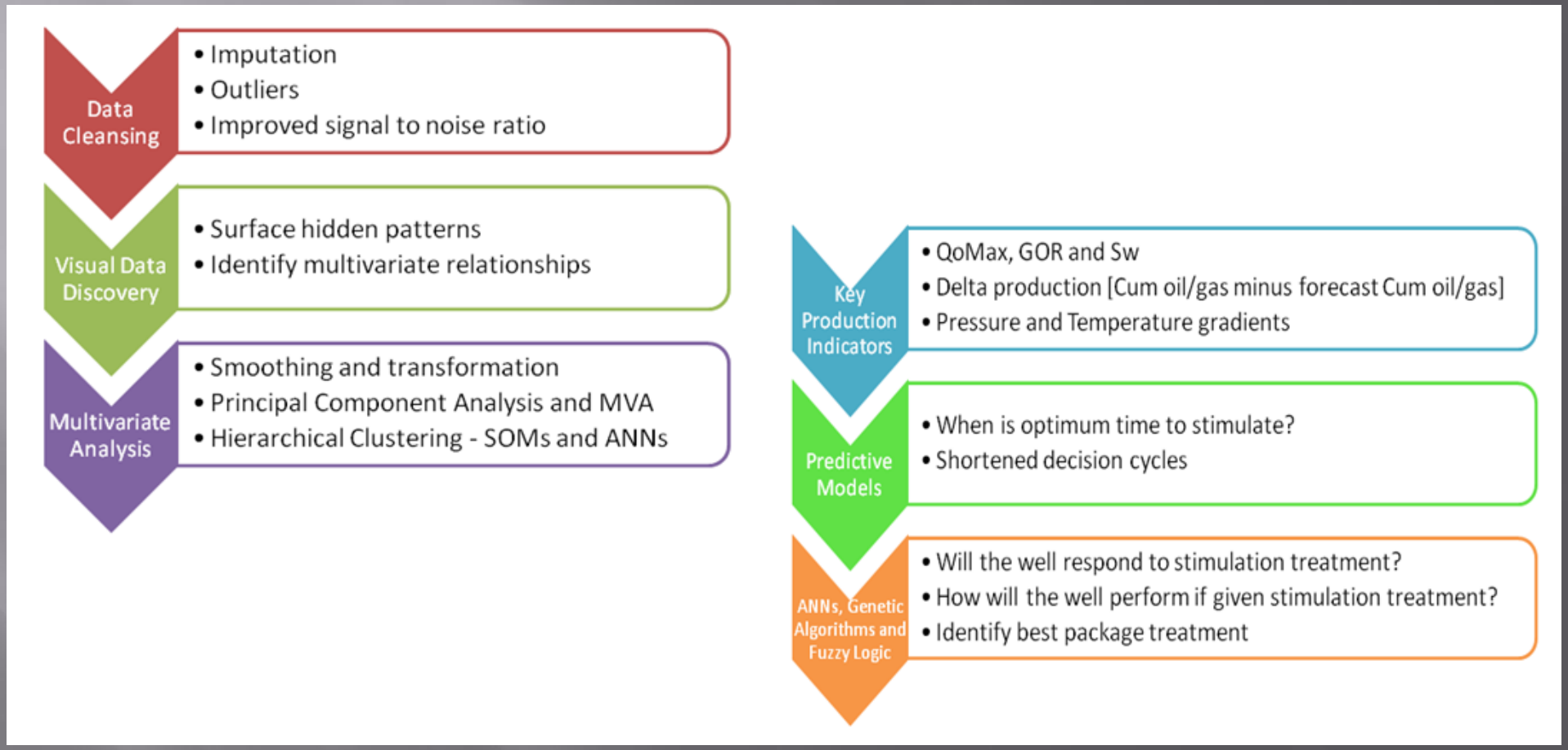
Albert Einstein

Predictive Models

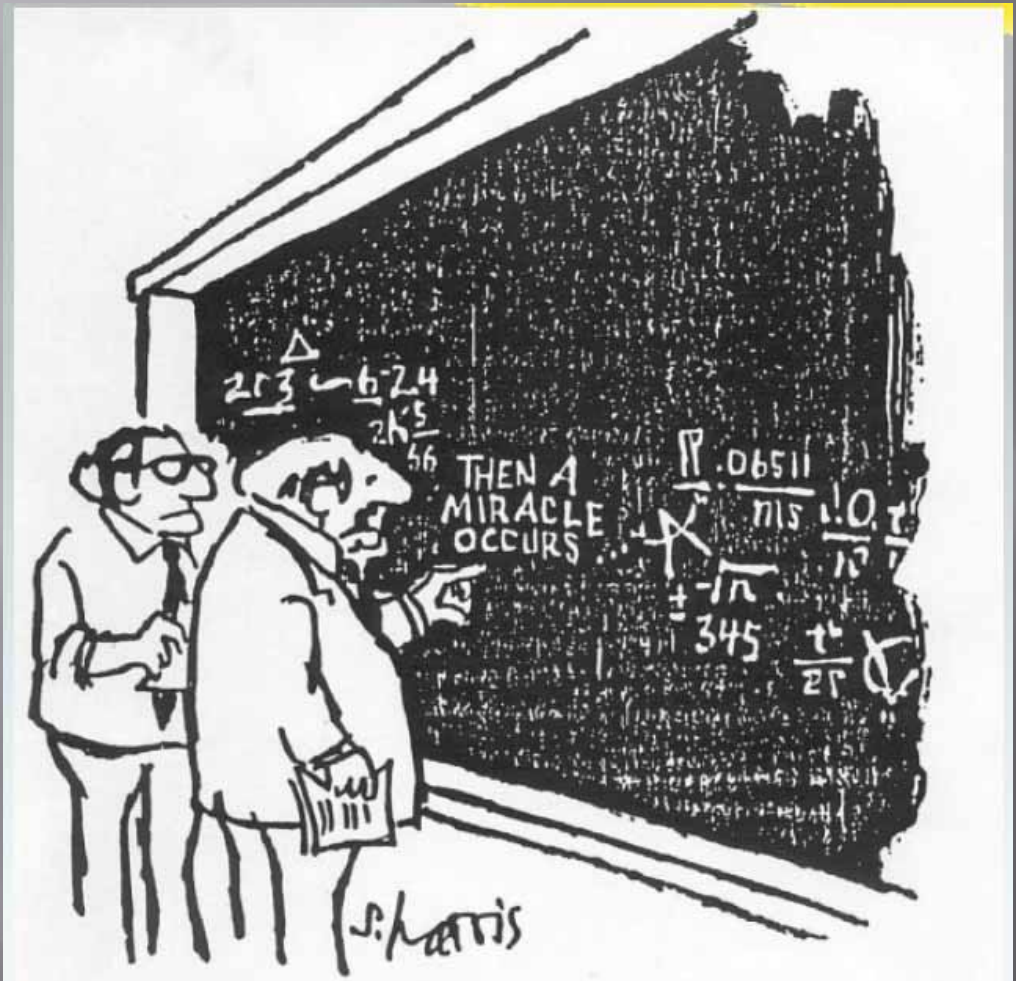


Conclusion

Analytical Workflow



Knowledge Evolves from Details



"I think you should be more explicit here in step two."

SPE142509:Tight Gas
SPE125368:Reservoir Characterization
SPE143701:DCA
SPE127269:Digital Oilfields
SPE62880:Total Field Re-engineering
SPE78333:Total Fast track Mature field
SPE135523:Shell Hydraulic Fracture

QUESTIONS

"We are overwhelmed by information, not because there is too much, but because we don't know how to tame it. Information lies stagnant in rapidly expanding pools as our ability to collect and warehouse it increases, but our ability to make sense of and communicate it remains inert, largely without notice."

Stephen Few, "Now You See It"