

Picking Winners and Losers in the Big Data Market

Michael E. Driscoll, CTO, Metamarkets
[@medriscoll](#)



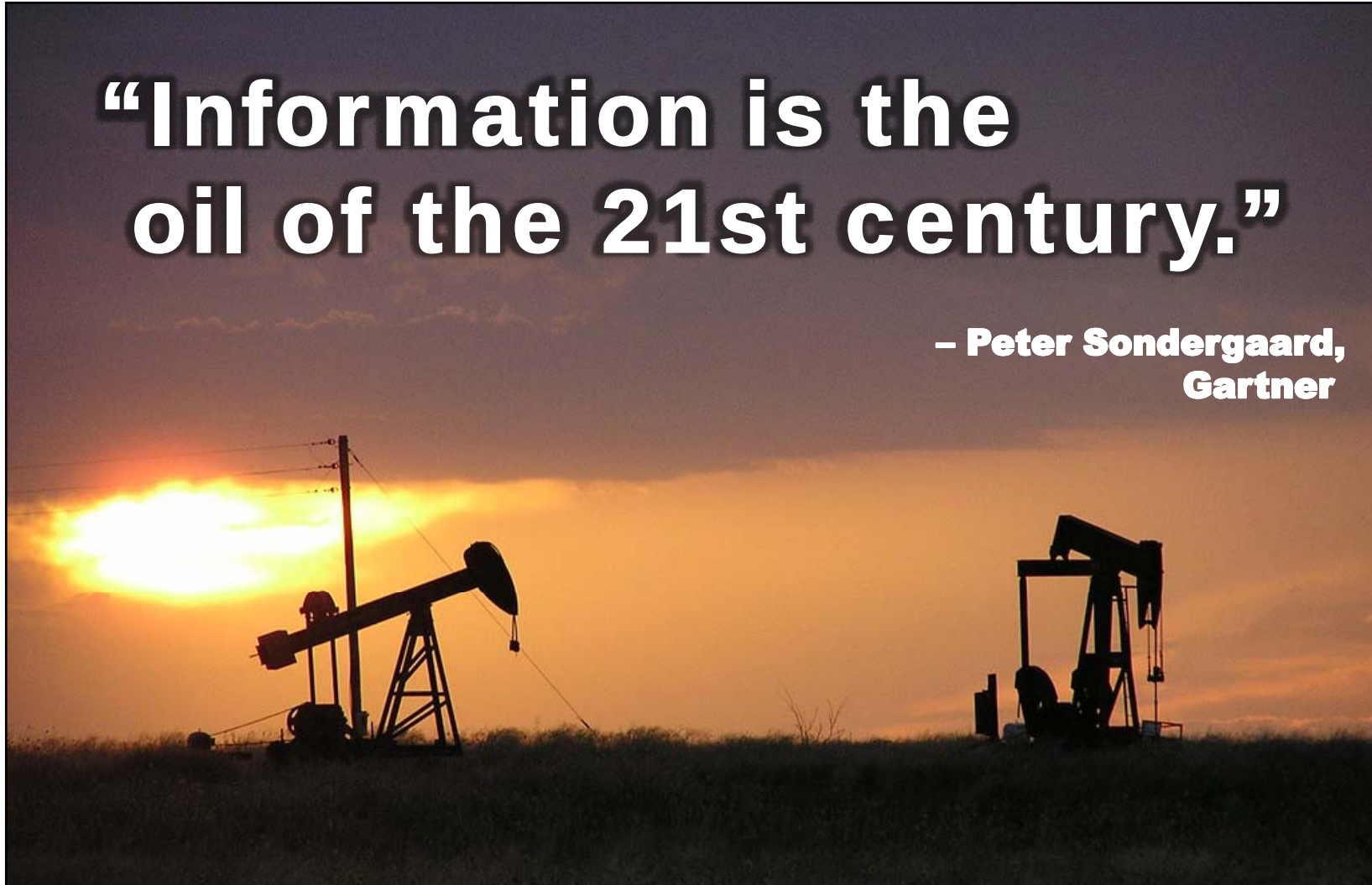
Open Source Business Conference | May 16, 2011

I. The Winning Opportunity

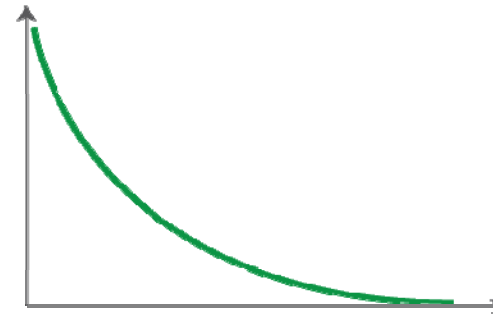
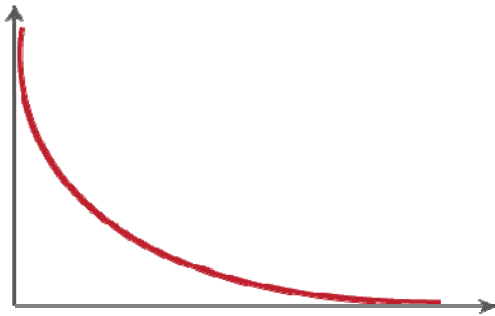
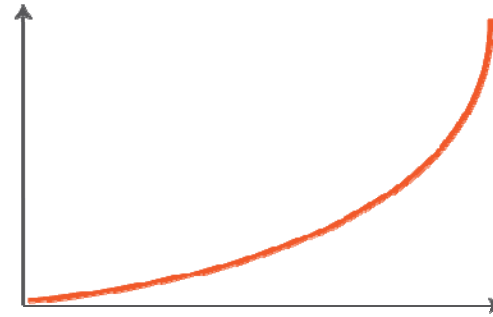
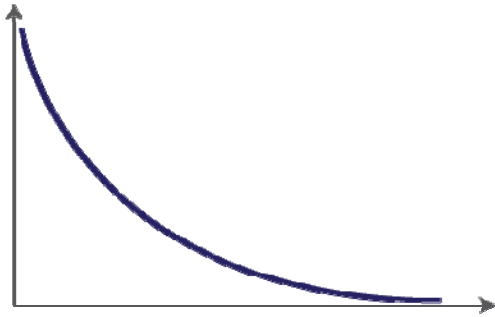


**“Information is the
oil of the 21st century.”**

**– Peter Sondergaard,
Gartner**

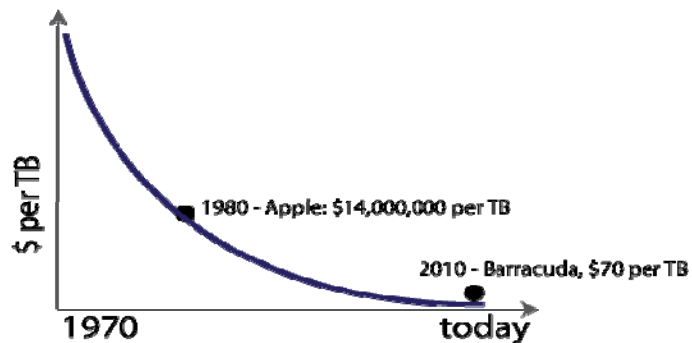


Big Data Force #1: The Attack of the Exponentials

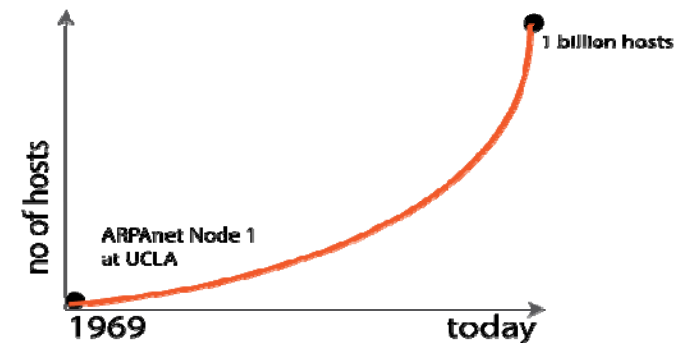


Big Data Force #1: The Attack of the Exponentials

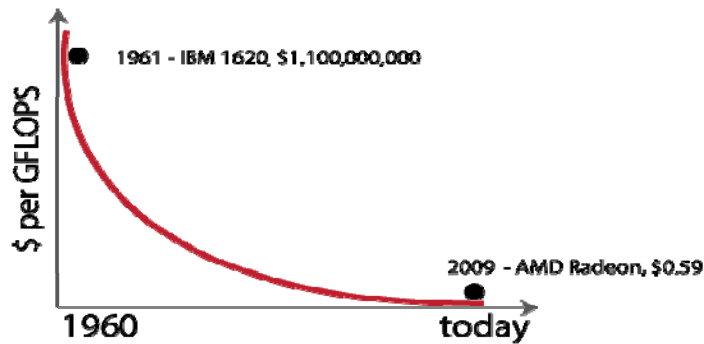
storage cost



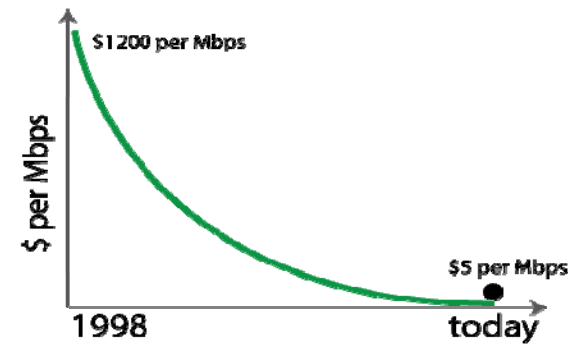
network access



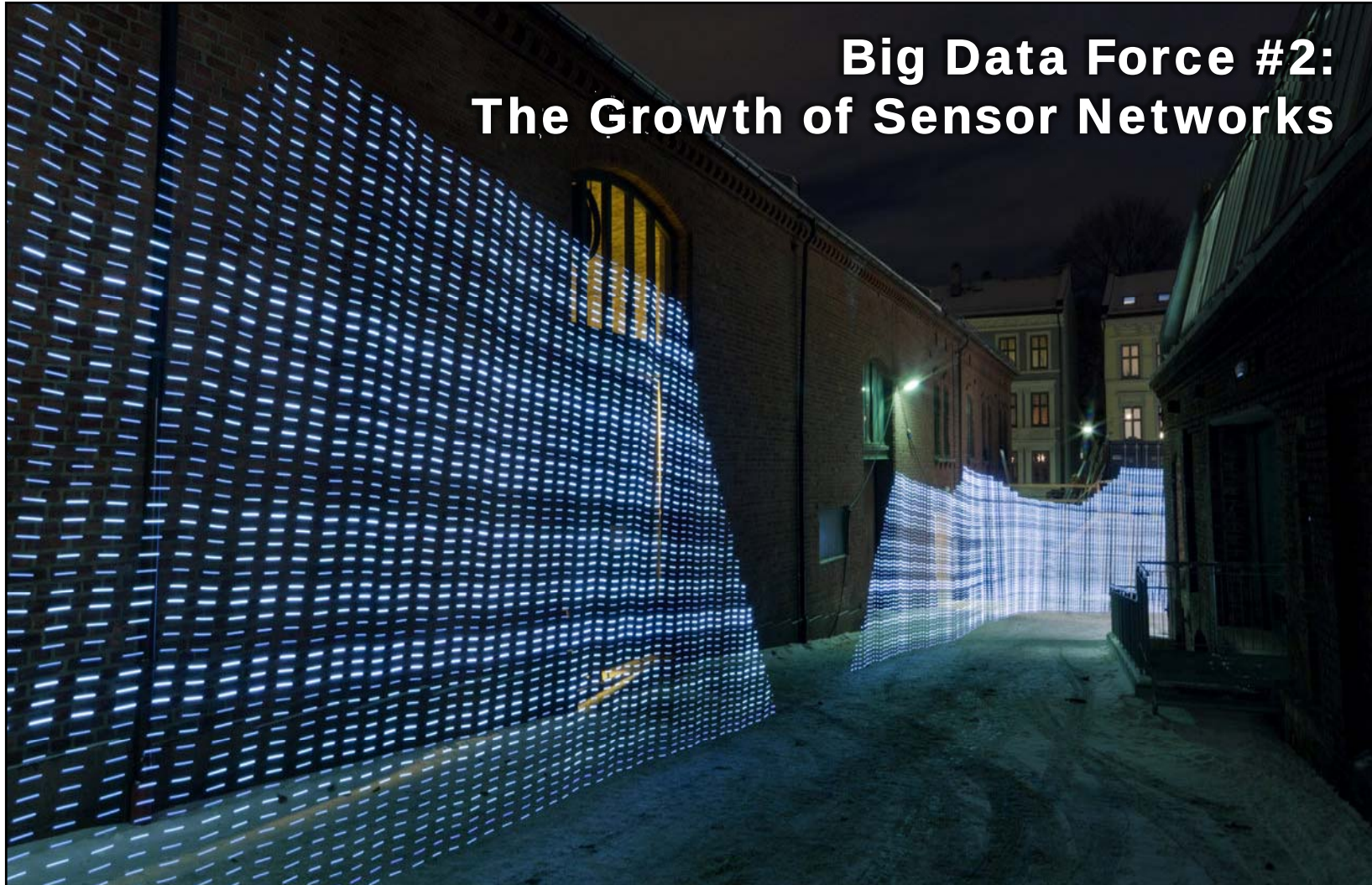
CPU cost



bandwidth cost



Big Data Force #2: The Growth of Sensor Networks

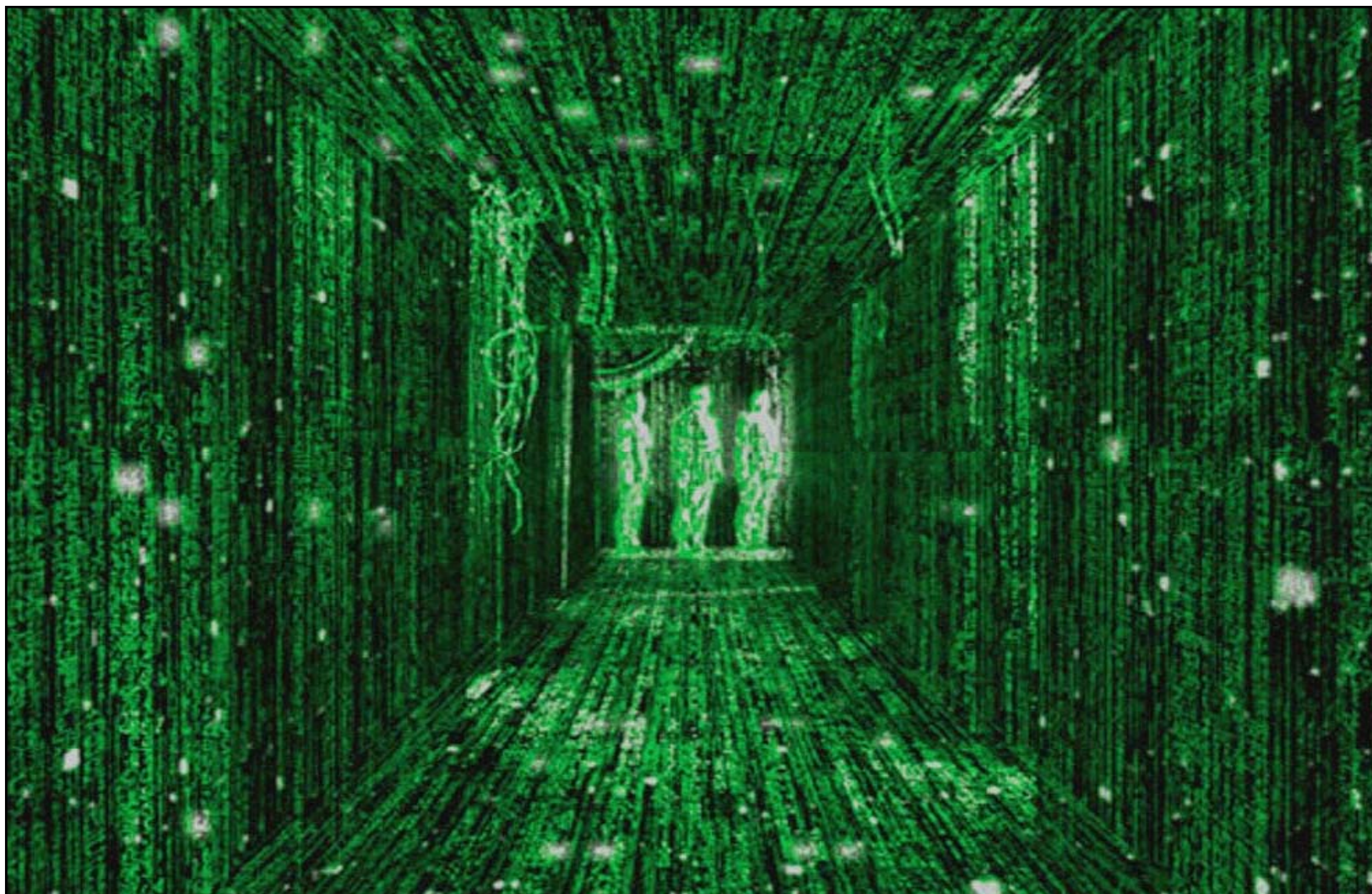


Big Data Force #3: The Rise of Cloud Computing



The Intersection of these Three Forces Yields An Explosion of Data

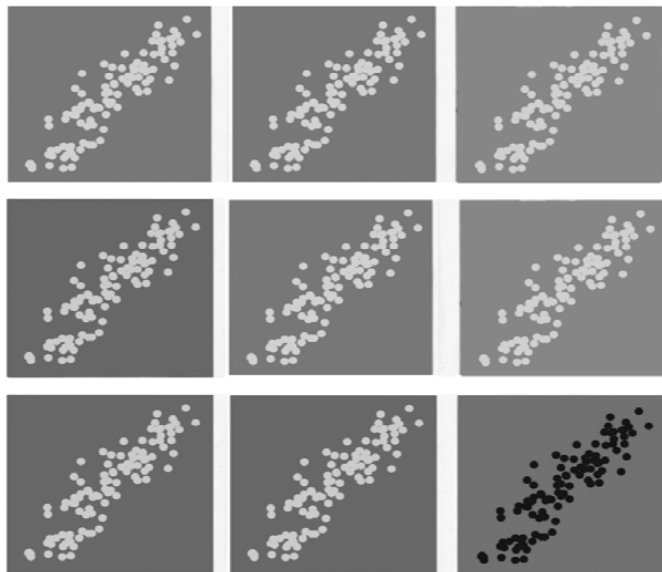




II. Winning Skills



“The sexy job in the next ten years will be statisticians...”
- Hal Varian

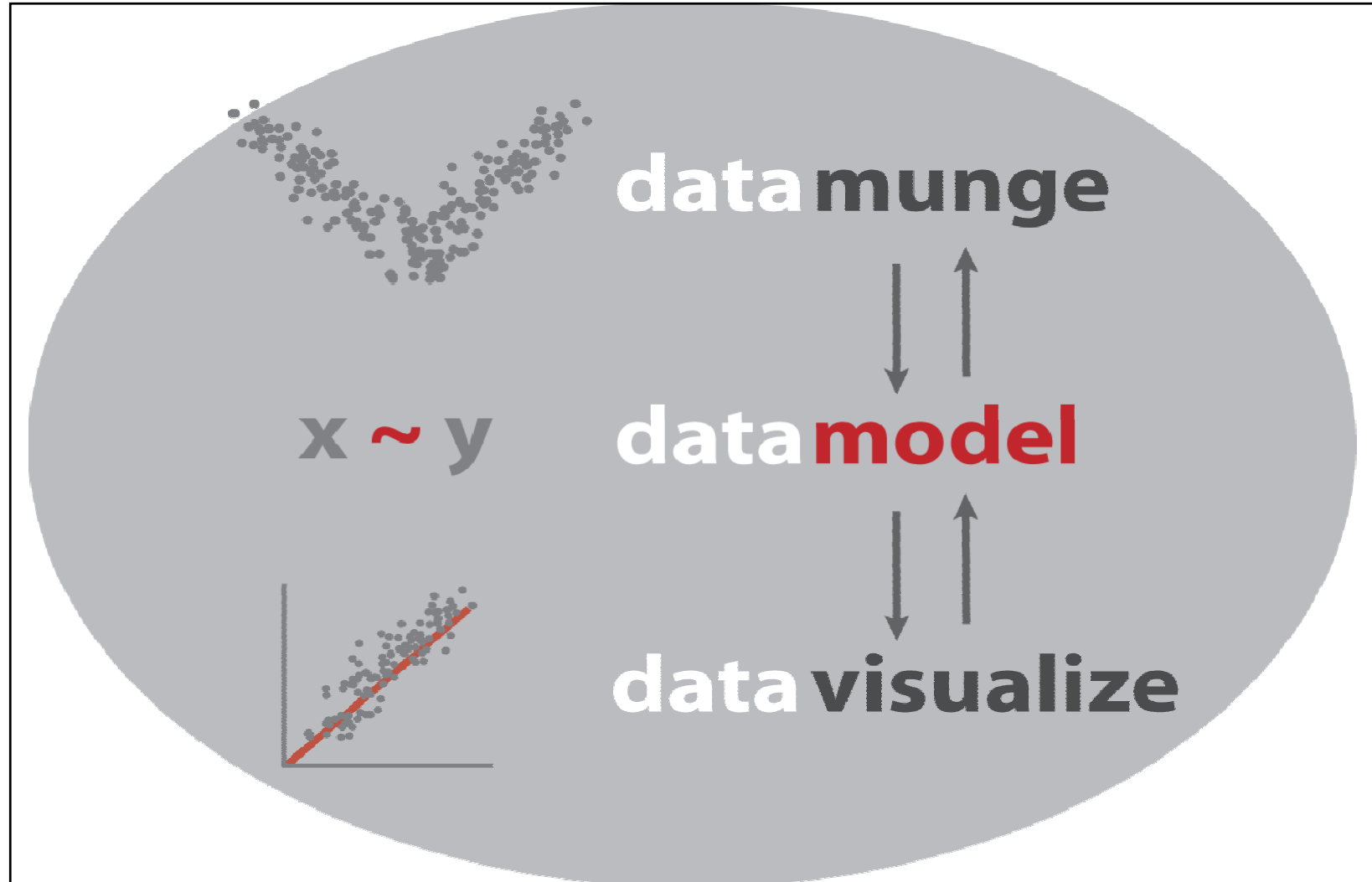


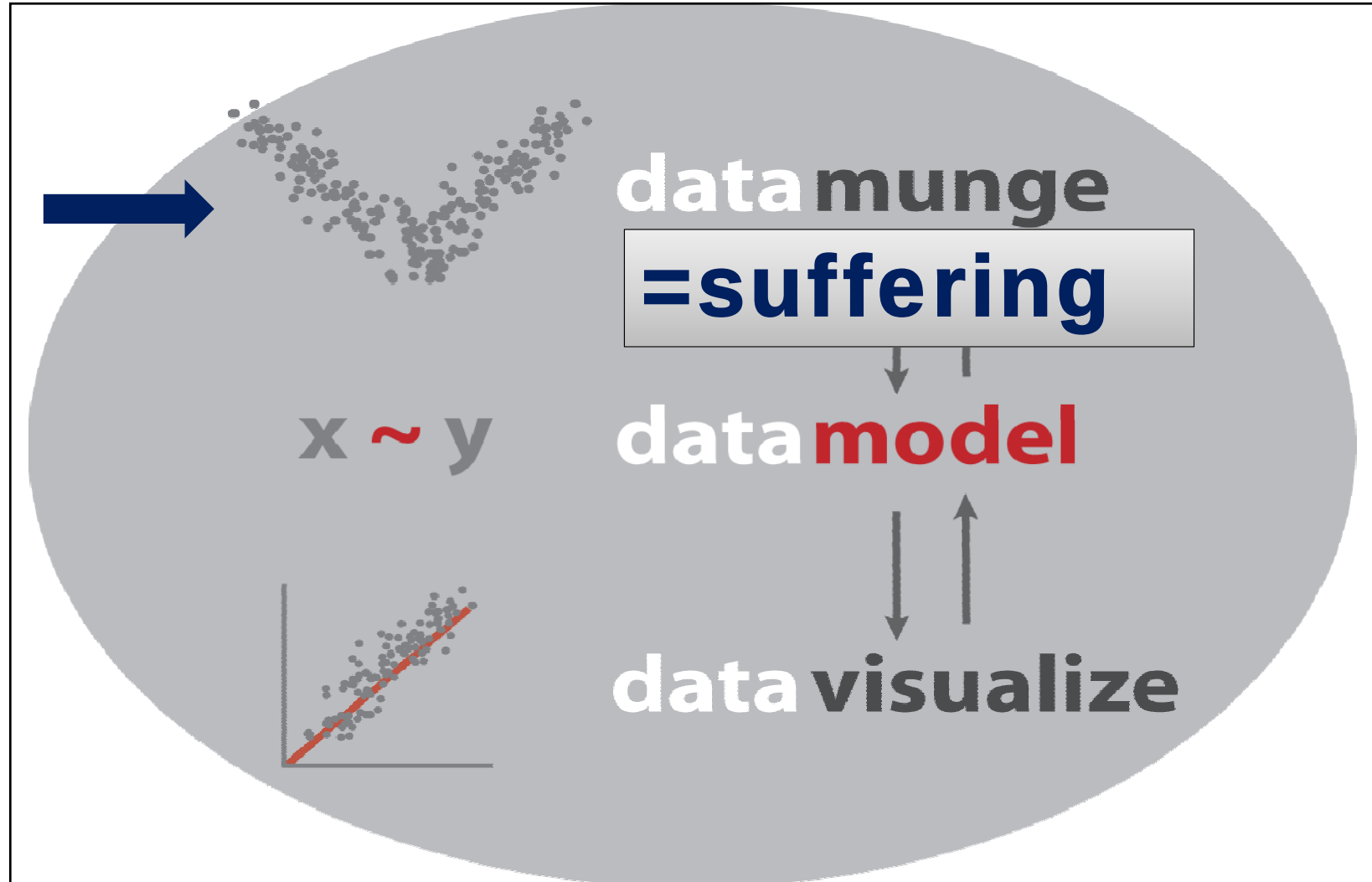
+



=



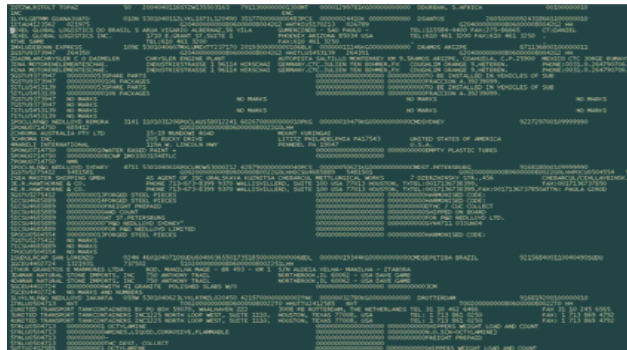




```

1STZWLRSTOLT TOPAZ 50 2004040116STZW135503163 7911300000001300MT 00001299781KG0000000000 DDURBAN, S.AFRICA 00100000010
INC EMC
1LYKLG8TMM GUANAJUATO 010N 5301040112LYKLTST120490 3517700000000483PCS 00000024410K 000000000000 DSANTOS 2605000000924318661100000010
1TITAU4123562 021975 G2040000000008060000080042G2 HHTRIU5170213 026789 G2040000000008060000080042G2 HH
EXEL GLOBAL LOGISTICS DO BRASIL S ARUA VIGARIO ALBERNAZ, 56 VILA GUMERCINDO - SAO PAULO - TEL:115584-4400 FAX:275-86665 CT:DANIEL
EXEL GLOBAL LOGISTICS INC. 1710 E. GRANT ST. SUITE 1 PHOENIX ARIZONA 85034 USA TEL:610 461 3200 FAX:610 461 3250
4THE SAME 109E 5301040607MXLUMEXTT2737170 2019300000000106BLK 00000011146KG000000000000 DRAMOS ARIZPE 871136800100000011
1MXLUDEBONN EXPRESS 109E 5301040607MXLUMEXTT2737170 2019300000000106BLK 00000011146KG000000000000 DRAMOS ARIZPE 871136800100000011
1GSTU9373947 264350 G2040000000008060000080042G2 HHITLU5453139 264351 G2040000000008060000080042G2 HH
2DAILMARCHRYSLER C O DAIMELER CHRYSLER ENGINE PLANT AUTOPISTA SALTILLO MONTERREY KM 9.5 RAMOS ARIZPE, COAHUILA, C.P.25900 MEXICO CTC JORGE RUMAYOR
3INA MOTORENELEMENTESCHAE. INDUSTRIESTRASSE 1 96114 HIRSCHAI GERMANY, CTC JULIEN TEN BOHMER, FX COUGHLIN ORANGE 9, HETEREN. PHONE:0031.0.264790706.
4INA MOTORENELEMENTESCHAE. INDUSTRIESTRASSE 1 96114 HIRSCHAI GERMANY, CTC JULIEN TEN BOHMER, FX COUGHLIN ORANGE 9, HETEREN. PHONE:0031.0.264790706.
5GSTU9373947 00000000053SPARE PARTS 000000000000000000 000000000000 BE INSTALLED IN VEHICLES OF SUB
5GSTU9373947 00000000000106 PACKAGES 000000000000000000 000000000000 FRACCION A.39239099.
5ITLU5453139 00000000053SPARE PARTS 000000000000000000 000000000000 BE INSTALLED IN VEHICLES OF SUB
5ITLU5453139 00000000000106 PACKAGES 000000000000000000 000000000000 FRACCION A.39239099.
7GSTU9373947 NO MARKS NO MARKS NO MARKS NO MARKS
7GSTU9373947 NO MARKS NO MARKS NO MARKS NO MARKS
7ITLU5453139 NO MARKS NO MARKS NO MARKS NO MARKS
7ITLU5453139 NO MARKS NO MARKS NO MARKS NO MARKS
1POCLLRP&O NEDLLOYD REMURA 3141 1101031206POCLAUS58012241 6026700000000010PKG 00000010479KG000000000000CMDSYDNEY 922729700109999990
1PONU0714750 685412 G0020000000008060000080022GOLHH MOUNT KURINGAI
2CHROMA AUSTRALIA PTY LTD 15-19 MUNDOWI ROAD LITITZ PHILADELPHIA PA17543 UNITED STATES OF AMERICA
3CHROMA INC. 205 BUCKY DRIVE PENNDEL PA 19047 U.S.A.
4MARELI INTERNATIONAL 119A W. LINCOLN HWY PENNDEL PA 19047 U.S.A.
5PONU0714750 00000000010WATER BASED PAINT + 000000000000000000 000000000000 EMPTY PLASTIC TUBES
5PONU0714750 00000000000ECN# 1M033031548TLC 000000000000000000
7PONU0714750 NMR 000000000000000000
1POCLNLP&O NEDLLOYD SYDNEY 4711 5301040616POCLMOW53000212 42879000000000040PCS 00000050621KG000000000000CMDSYDNEY 916818000109999990
1GSTU5275412 5481581 G0020000000008060000080022GOLHHICSU4685889 5481501 G0020000000008060000080022GOLHHPOCU0504554 548
2SEA MASTER SHIPPING GMBH AS AGENT OF JSC URALS KAYA KUZNETSA CHEBARCUL METTURGICAL WORKS 7 DZERZHINSKY STR., 456 CHEBARCUL/CEHLAYBINSK, R
3E.R.HAWTHORNE & CO. PHONE 713-673-8399 9370 WALLISVILLERD, SUITE 100 USA 77013 HOUSTON, TX TEL:0017136738399, FAX:0017136737850
4E.R.HAWTHORNE & CO. PHONE 713-673-8399 9370 WALLISVILLERD, SUITE 100 USA 77013 HOUSTON, TX TEL:0017136738399, FAX:0017136737850ATTN: PAULA GIROD
5GSTU5275412 00000000013FORGED STEEL PIECES 000000000000000000 000000000000 HARMONISED CODE:
5ICSU4685889 00000000014FORGED STEEL PIECES 000000000000000000 000000000000 HARMONISED CODE:
5ICSU4685889 00000000000FREIGHT PREPAID 000000000000000000 000000000000 DTHC / CUC COLLECT
5ICSU4685889 00000000000AND COUNT 000000000000000000 000000000000 SHIPPED ON BOARD
5ICSU4685889 00000000000AT ST.PETERSBURG 000000000000000000 000000000000 FOR P&O NEDLLOYD LTD.
5ICSU4685889 00000000000"P&O NEDLLOYD SYDNEY" 000000000000000000 000000000000 SYN4711 03JUN04
5ICSU4685889 00000000000FOR P&O NEDLLOYD LIMITED 000000000000000000 000000000000
5POCU0504554 00000000013FORGED STEEL PIECES 000000000000000000 000000000000 HARMONISED CODE:
7GSTU5275412 NO MARKS NO MARKS NO MARKS NO MARKS
7ICSU4685889 NO MARKS NO MARKS NO MARKS NO MARKS
7POCU0504554 NO MARKS NO MARKS NO MARKS NO MARKS
1SDULRC&O SAN LORENZO 024N 46010407105UDU6040036550173518500000000006BDL 00000010344KG000000000000CMDSCTPITIBA BRAZIL 9215684001100404905UDU
1GCEU4402724 1321931 737502 5102000000000806000008002251LHH 000000000000000000
2THOR GRANITOS E MARMORES LTDA ROD. MANILHA MAGE - BR 493 - KM 1 S/N ALDEIA VELHA- MANILHA - ITABORA
3DAMAR NATURAL STONE IMPORTS, INC 750 ANTHONY TRAIL NORTHBROOK, IL 60062 - USA DAVE GAME
4DAMAR NATURAL STONE IMPORTS, INC 750 ANTHONY TRAIL NORTHBROOK, IL 60062 - USA DAVE GAME
5GCEU4402724 00000000006WITH 41 GRANITE POLISHED SLABS W/O 000000000000000000 000000000000 3CM
7GCEU4402724 NO MARKS AND NUMBERS NO MARKS AND NUMBERS NO MARKS AND NUMBERS NO MARKS AND NUMBERS
1LYKLNLP&O NEDLLOYD JAKARTA 059W 5301040623LYKLRMTM2020450 4215700000000002TNK 00000032780KG000000000000 DROTTERDAM 916819200100000010
1TRLU0504713 NVT T0020000000008060000080022T0 HHUTTU2412585 NVT T0020000000008060000080022T0 HH
2UNITED TRANSPORT TANKCONTAINERS BV PO BOX 59070, WAALHAVEN 222 3008 PB ROTTERDAM, THE NETHERLANDS TEL 31 10 462 6466 FAX 31 10 245 6565
3UNITED TRANSPORT TANKCONTAINERS INC1225 NORTH LOOP WEST, SUITE 1110, HOUSTON, TEXAS 77008, USA TEL: 1 713 861 0250 FAX: 1 713 869 4792
4UNITED TRANSPORT TANKCONTAINERS INC1225 NORTH LOOP WEST, SUITE 1110, HOUSTON, TEXAS 77008, USA TEL: 1 713 861 0250 FAX: 1 713 869 4792
5TRLU0504713 00000000001 OCTYLAMINE 000000000000000000 000000000000 SHIPPERS WEIGHT LOAD AND COUNT
5TRLU0504713 00000000000AMINES, LIQUID, CORROSIVE, FLAMMABLE 000000000000000000 000000000000 (S(N-OCTYLAMINE)
5TRLU0504713 00000000000 000000000000000000 000000000000 FREIGHT PREPAID
5TRLU0504713 00000000000THC DEST. COLLECT 000000000000000000 000000000000
5TTTU2412585 00000000001 OCTYLAMINE 000000000000000000 000000000000 SHIPPERS WEIGHT LOAD AND COUNT

```

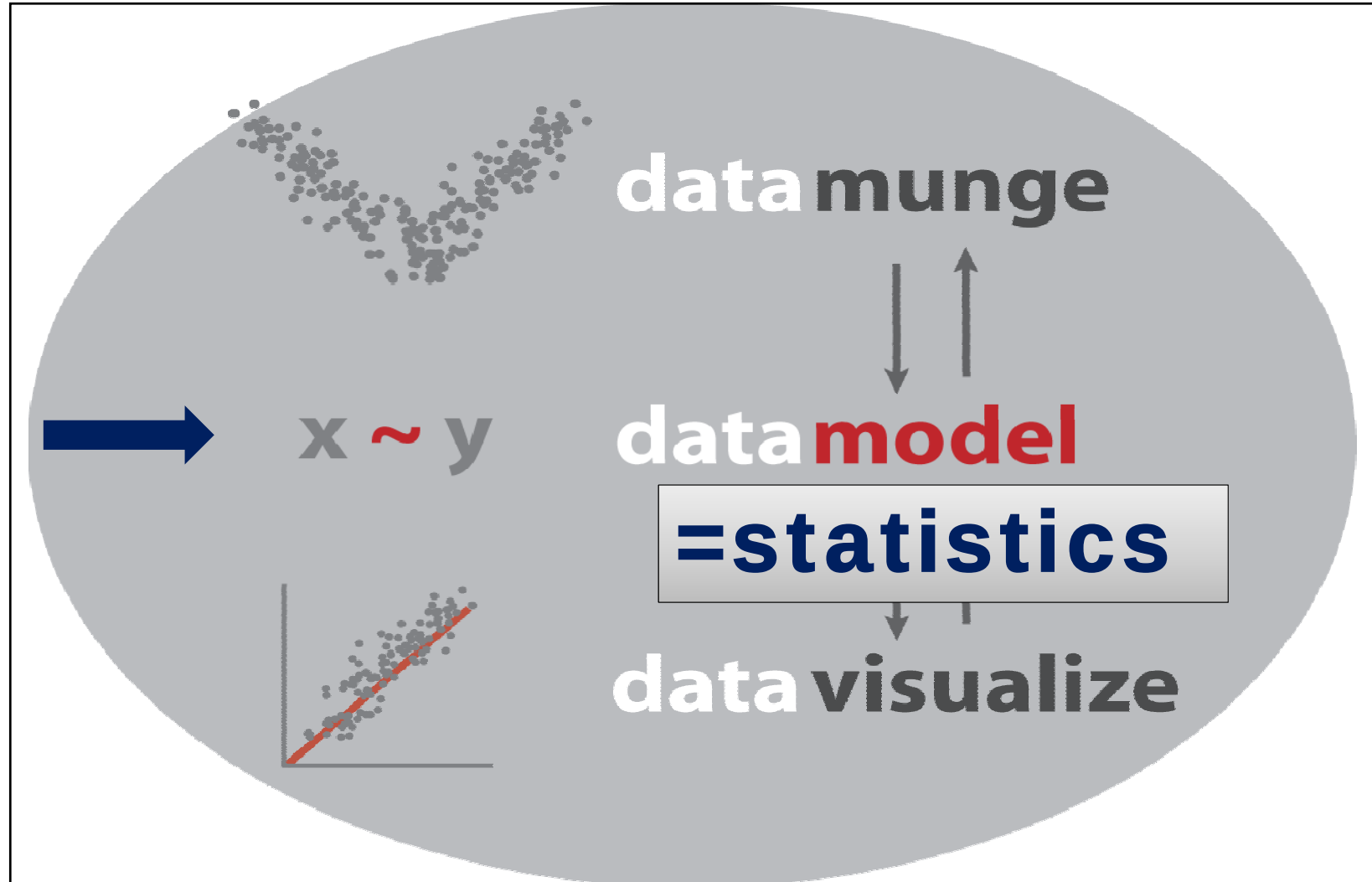


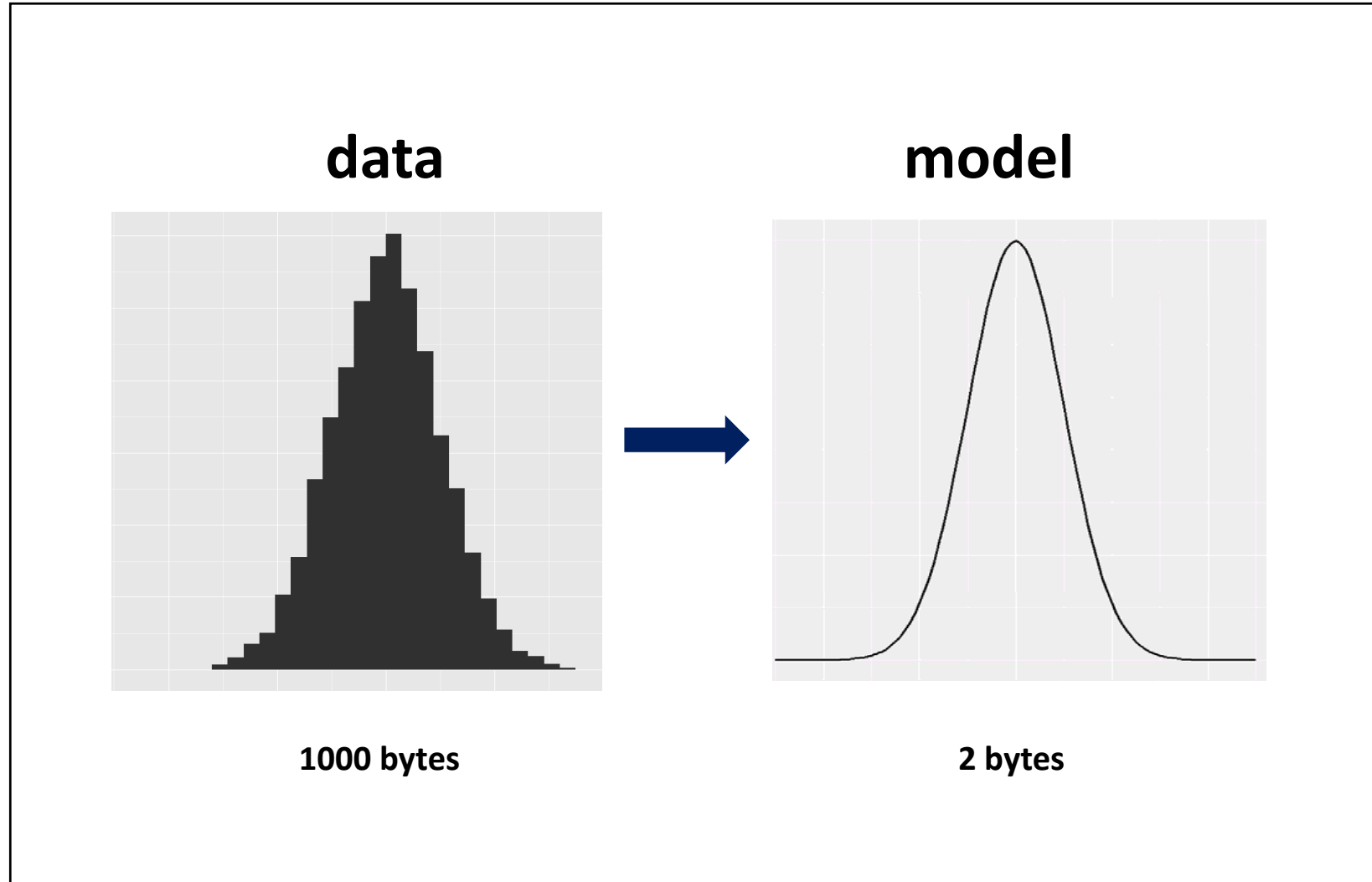
+



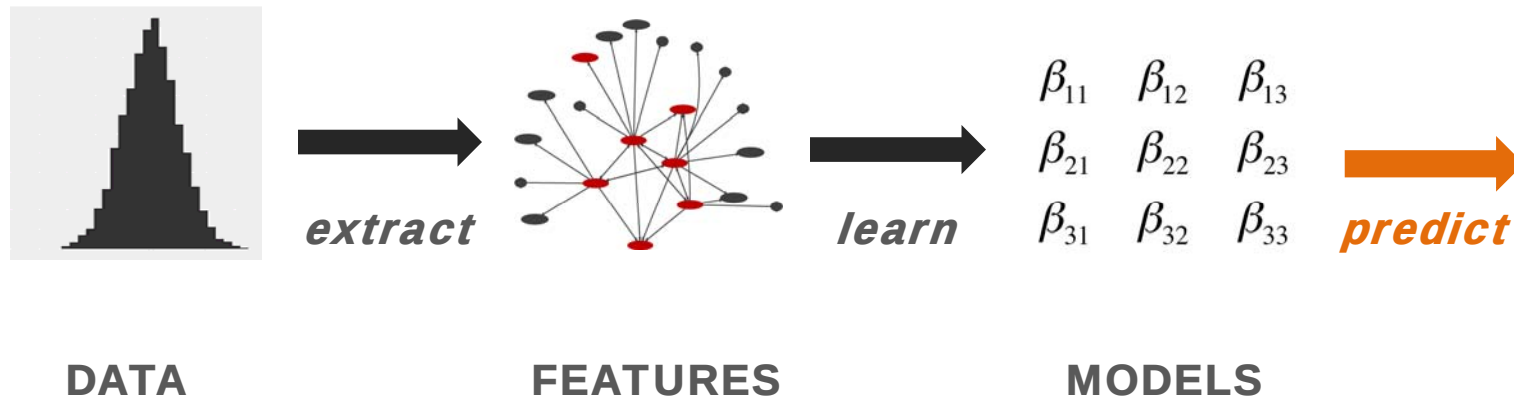
=

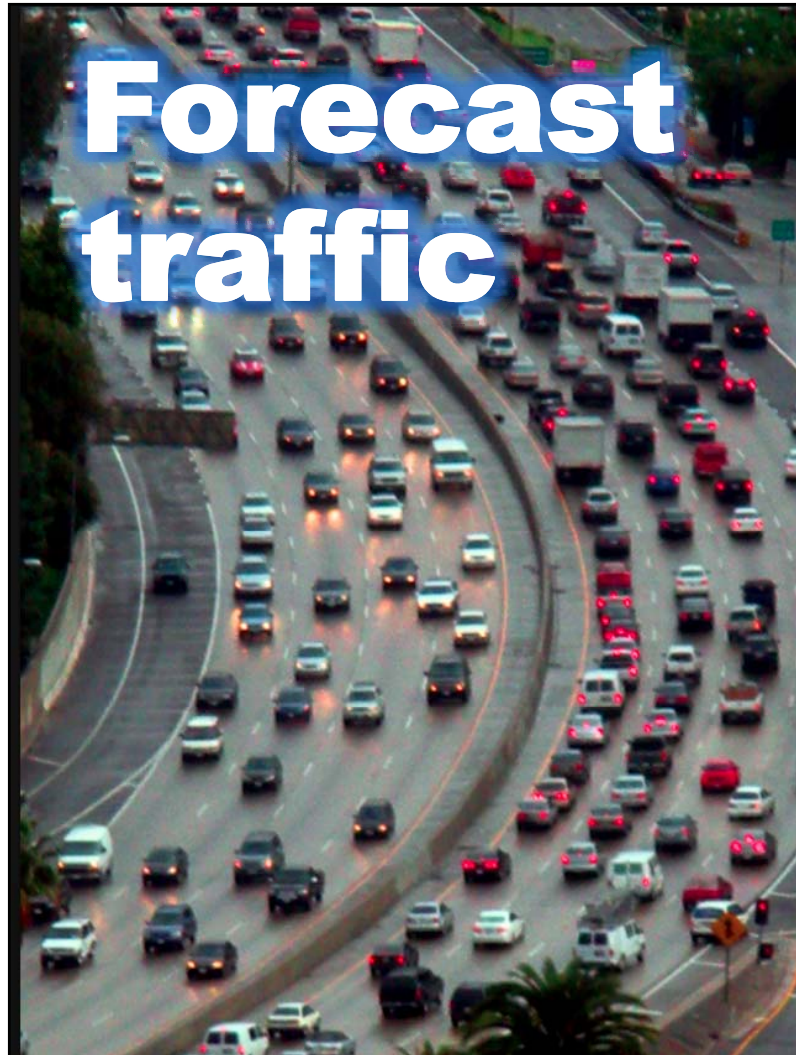
```
if ($foo =~
/\n {2,3}([A-Z]{5,7}) {2,5}/)
```





The Promise of Predictive Analytics



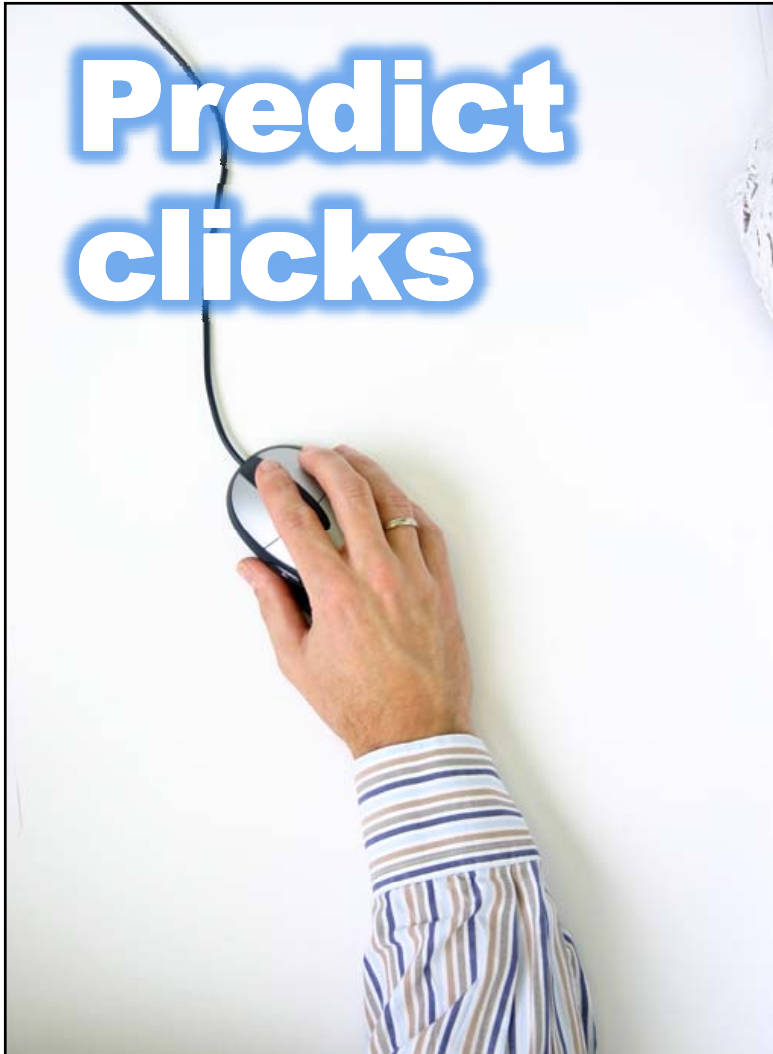


**Forecast
traffic**



**Recommend
a movie**

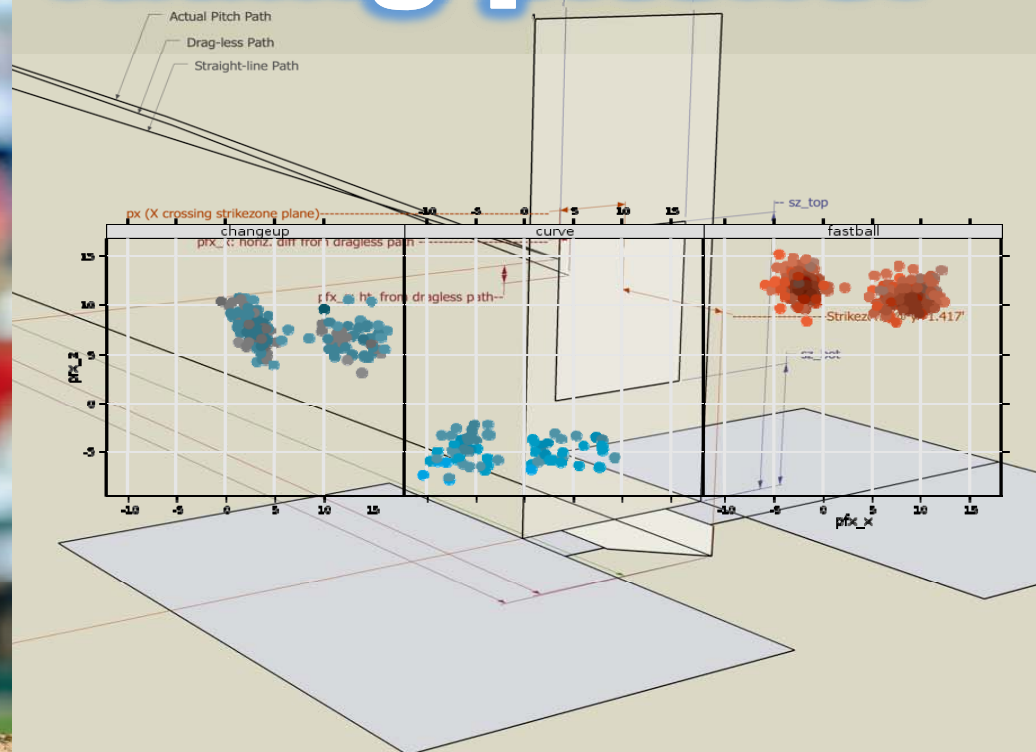
**Predict
clicks**

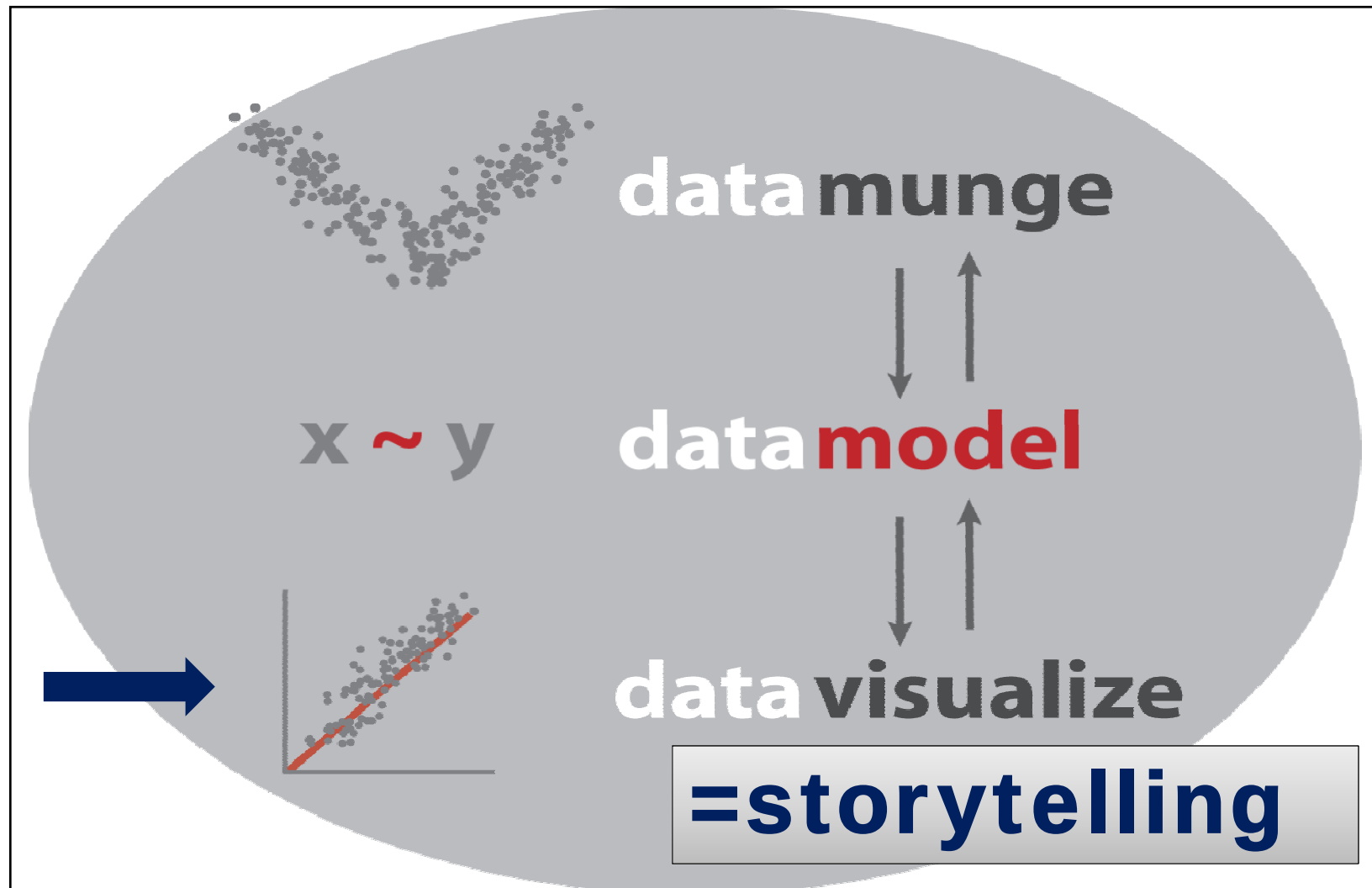


**Prevent
fraud**

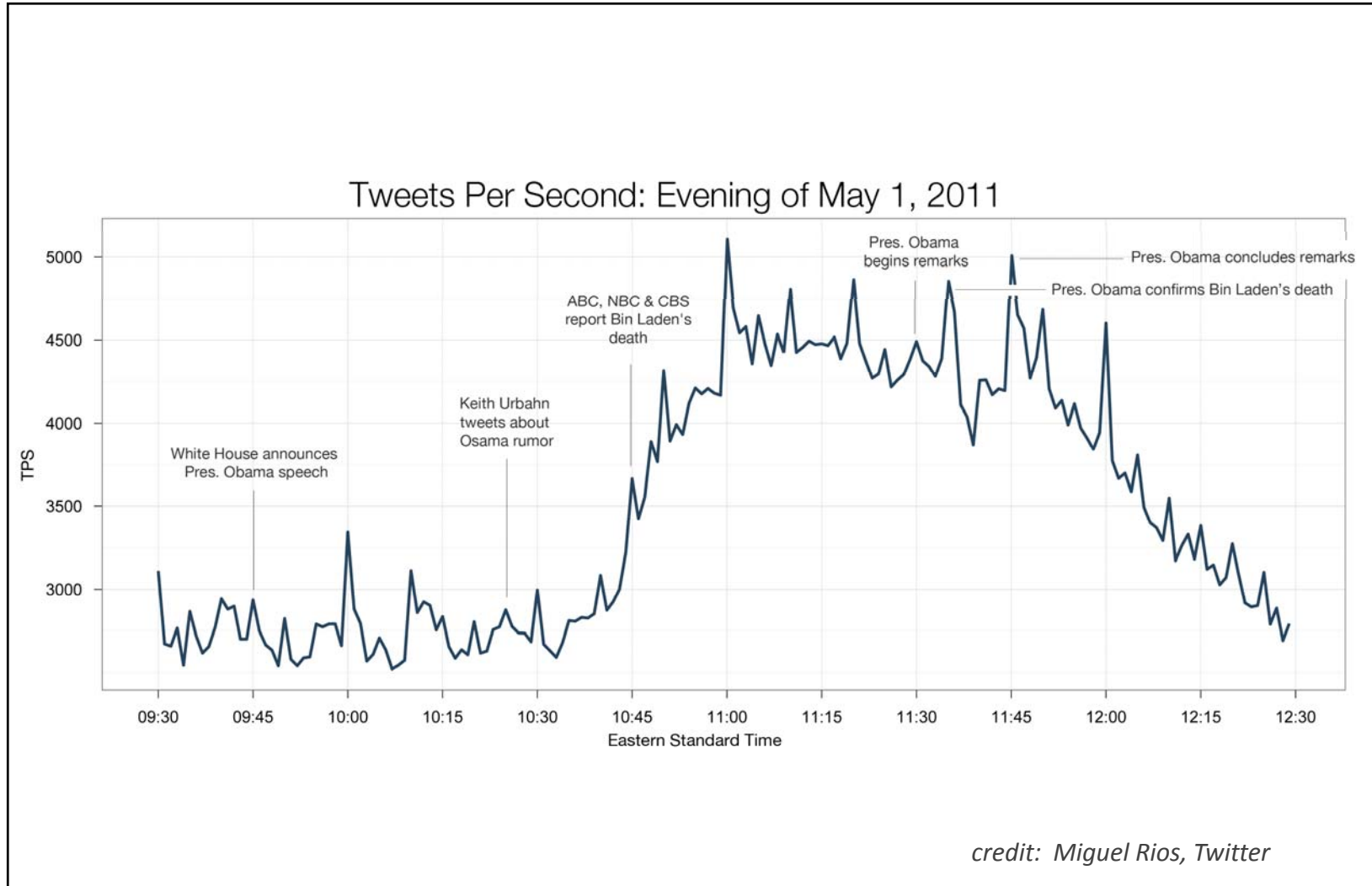


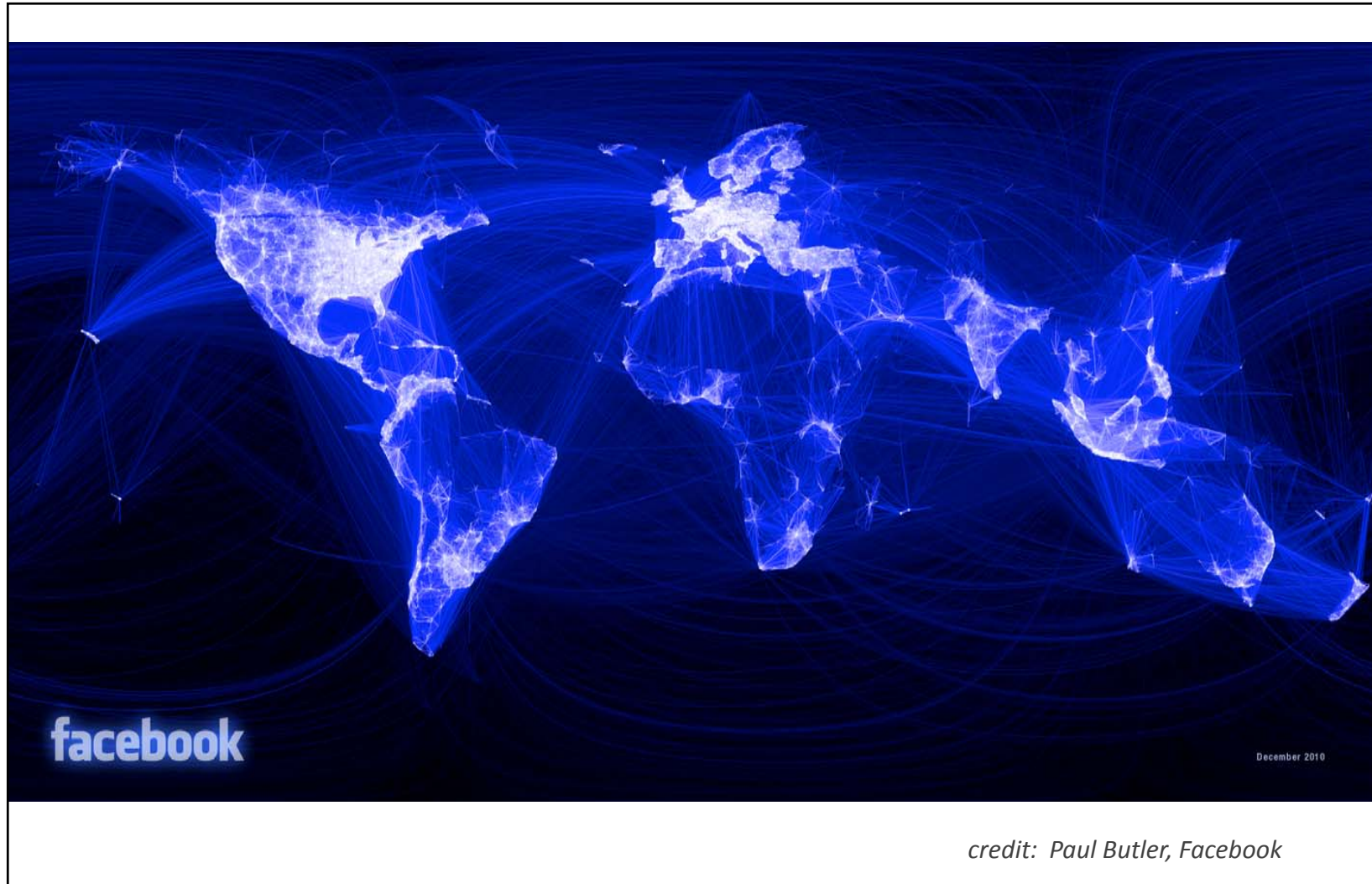
Classify a fading pitcher

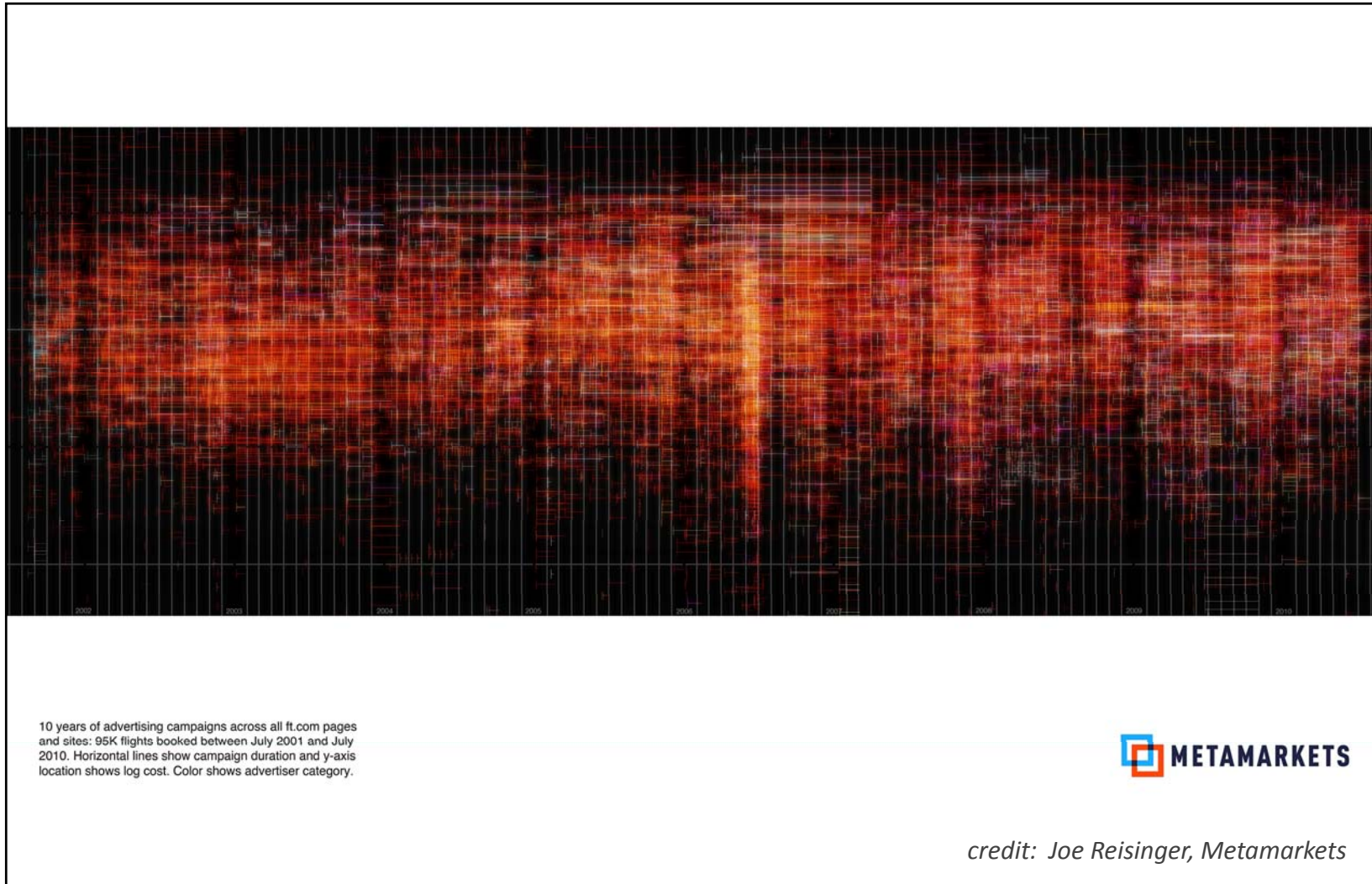




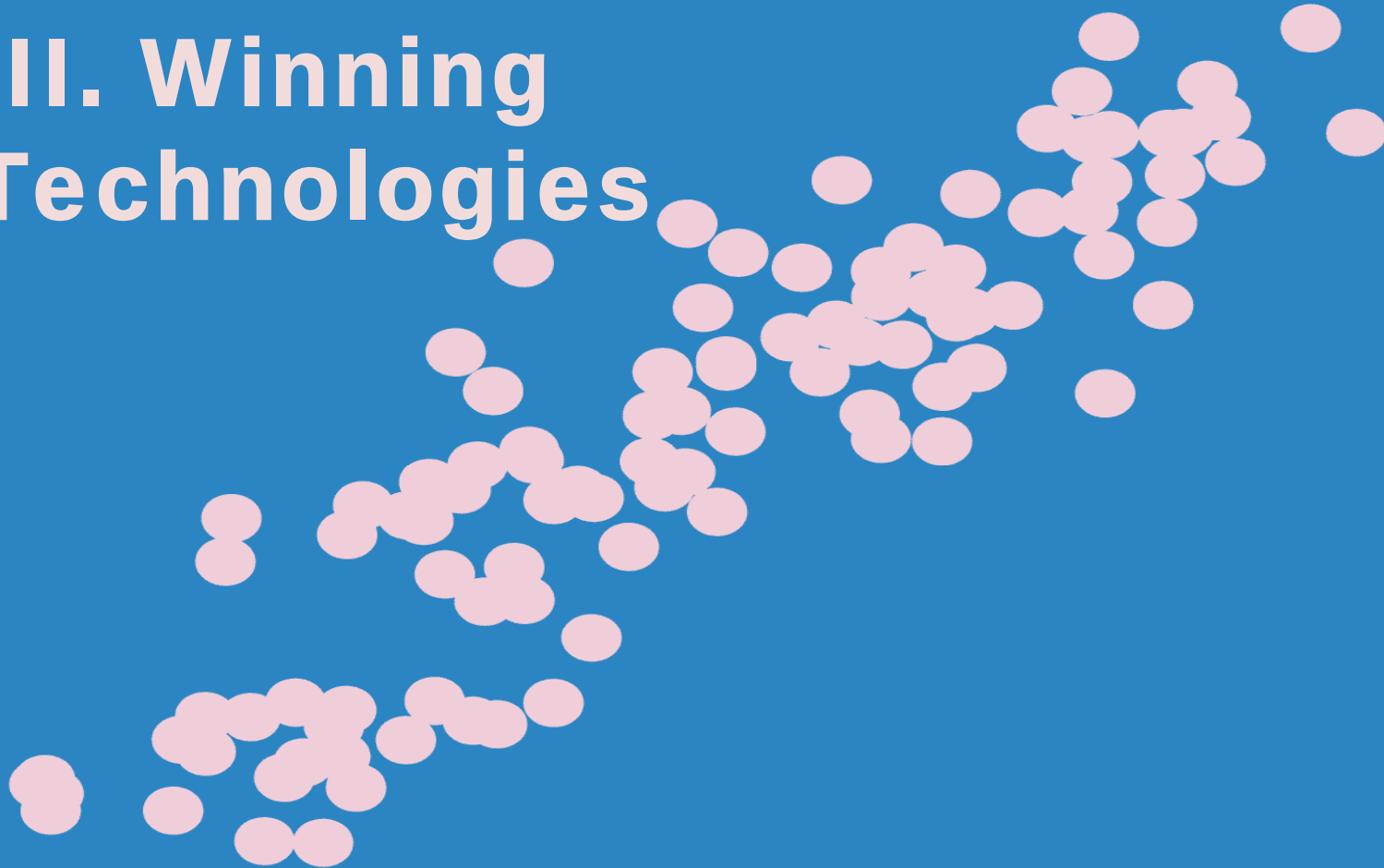




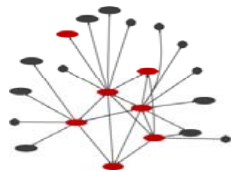




III. Winning Technologies



The Emerging Big Data Stack



Apps & Services
(Fraud Detection, Rec Engines)

$$\begin{matrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{matrix}$$

Analytics
(R, SPSS, SAS, SAP)



Data
RDBMS, Hadoop, CEP

Actions



Insights

Data

The Emerging Big Data Stack



Apps & Services

$$\begin{matrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{matrix}$$

Analytics

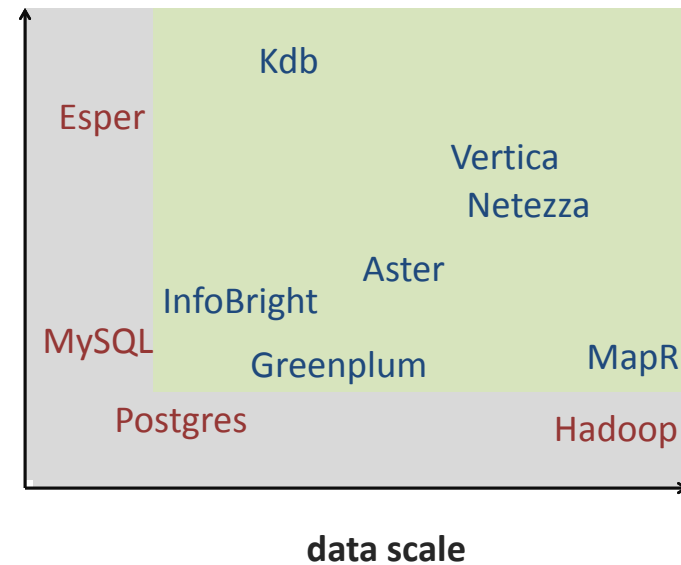


Data

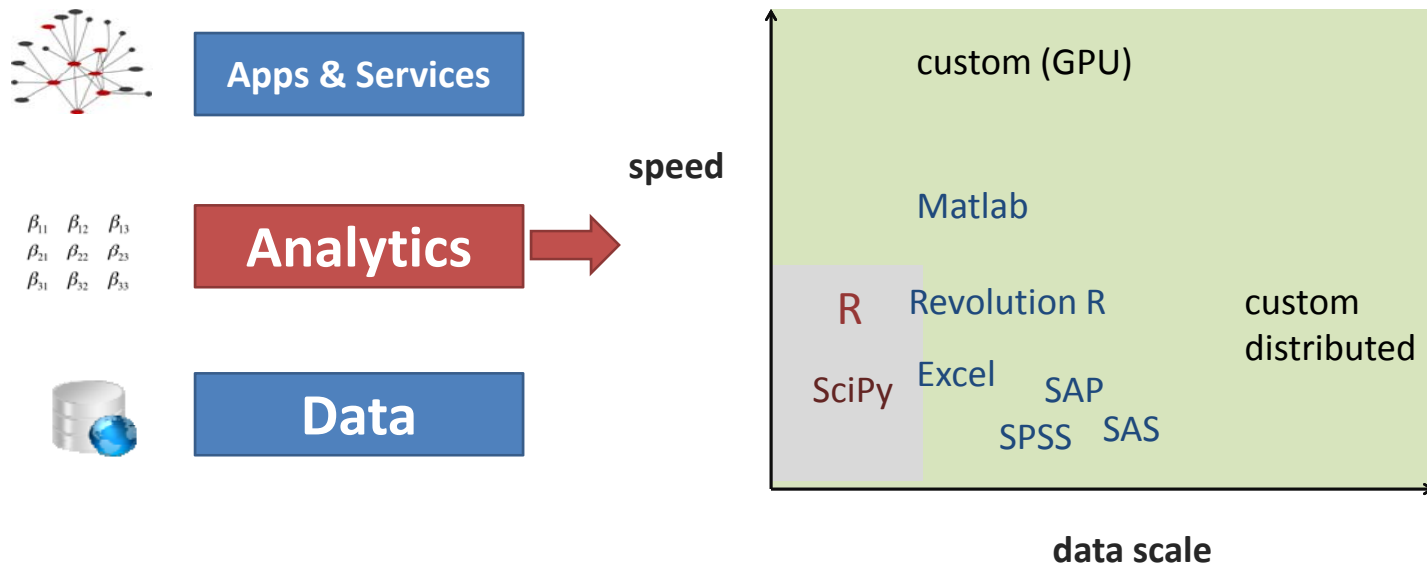


speed

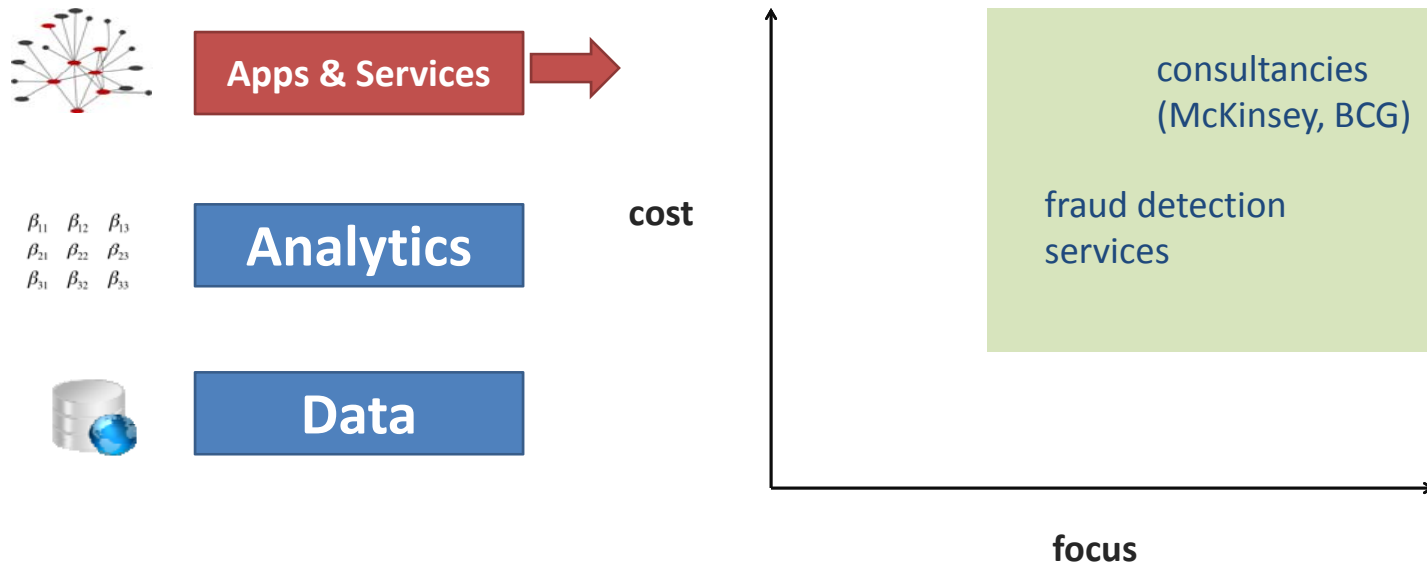
Competitive Matrix



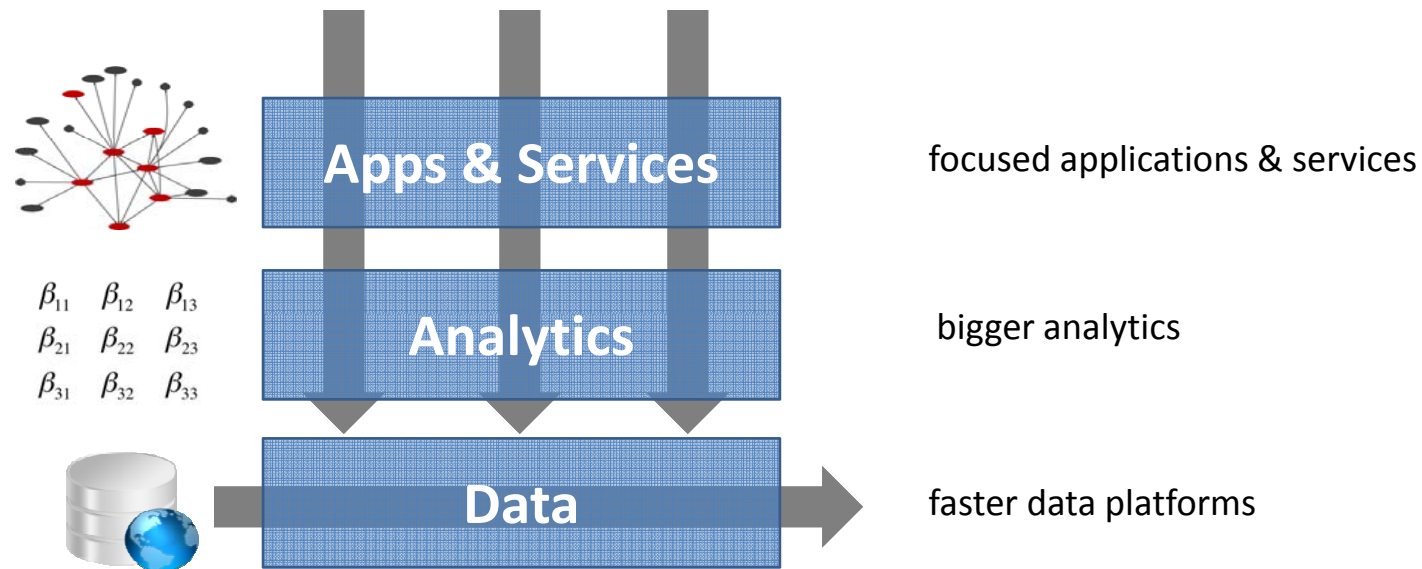
The Emerging Big Data Stack



The Emerging Big Data Stack



Winners: Focused, Big, and Fast

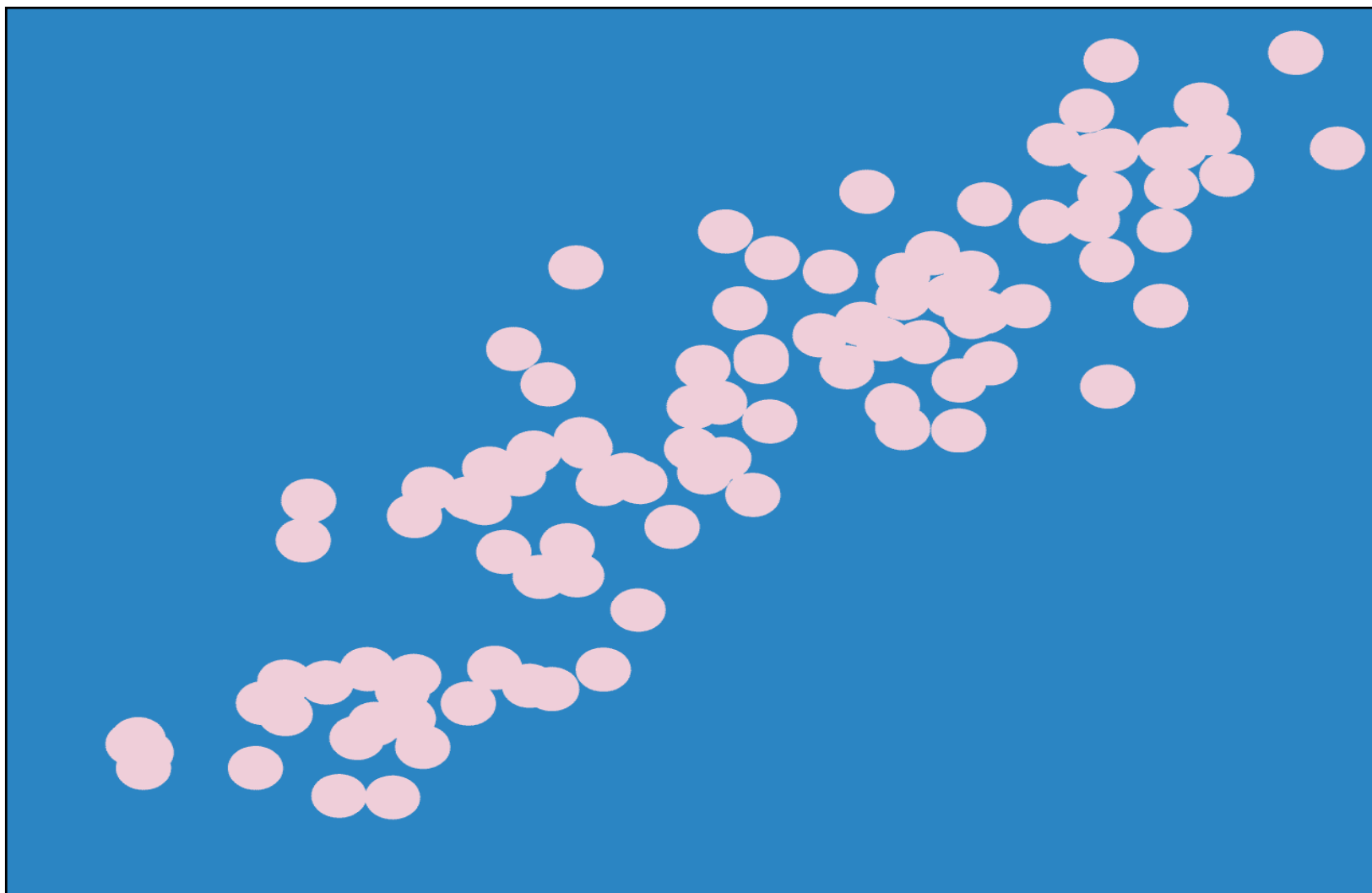


Thank You. Questions?

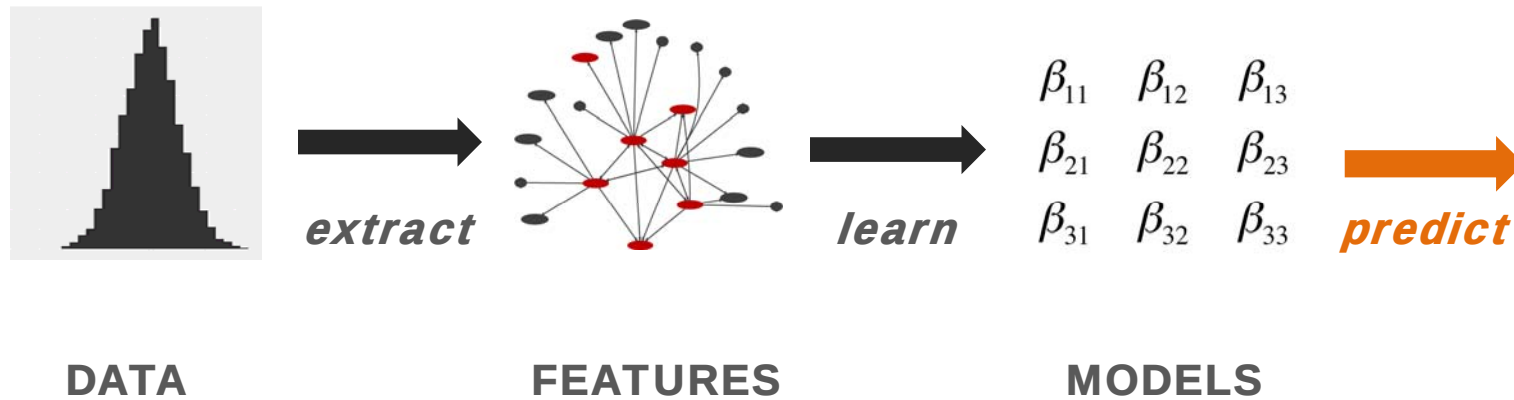
**Michael E. Driscoll, CTO, Metamarkets
@medriscoll**

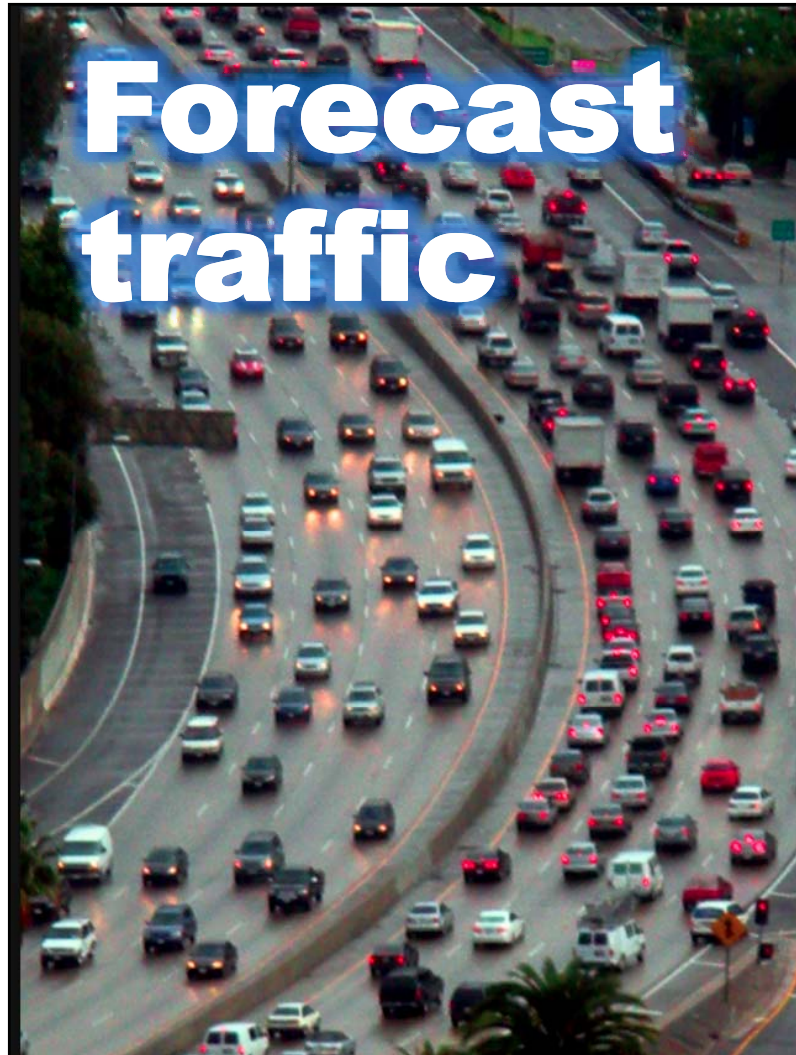


Open Source Business Conference | May 16, 2011



The Promise of Predictive Analytics



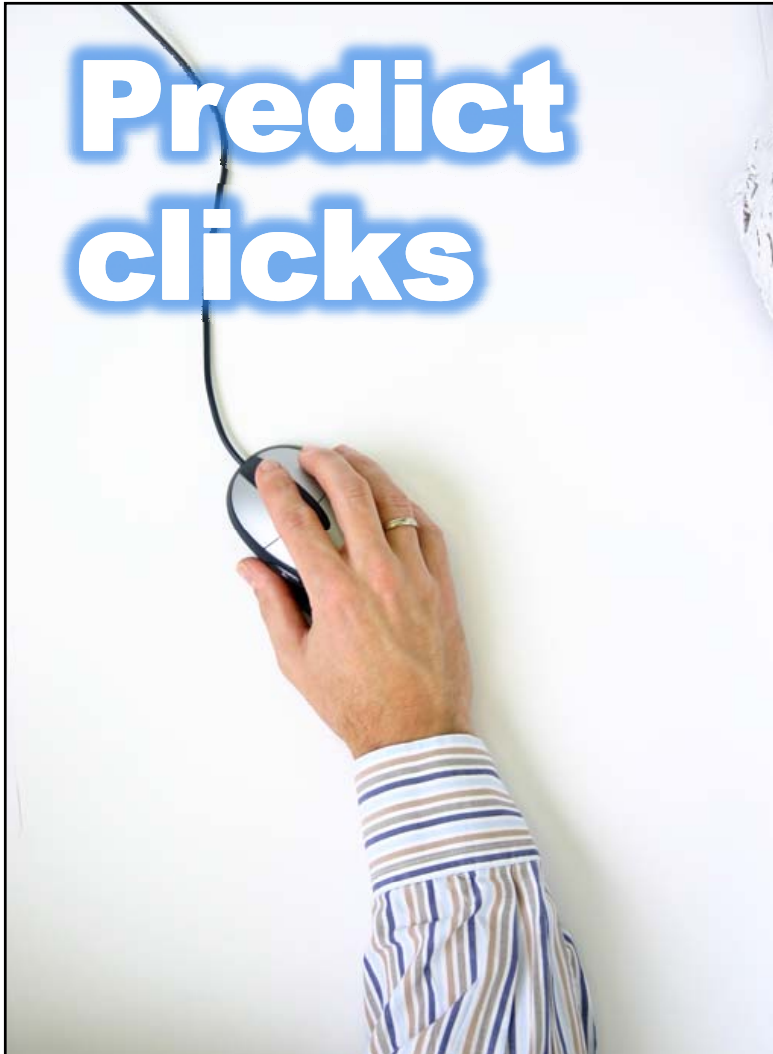


**Forecast
traffic**



**Recommend
a movie**

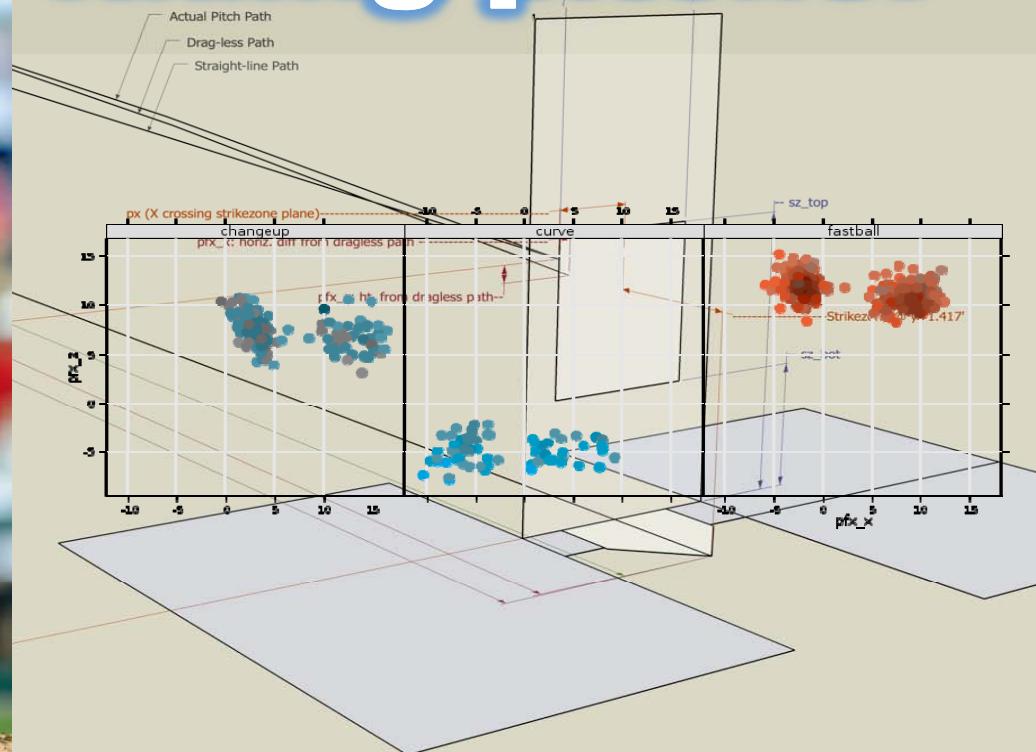
**Predict
clicks**



**Prevent
fraud**



Classify a fading pitcher



I. Data Privacy & Ownership



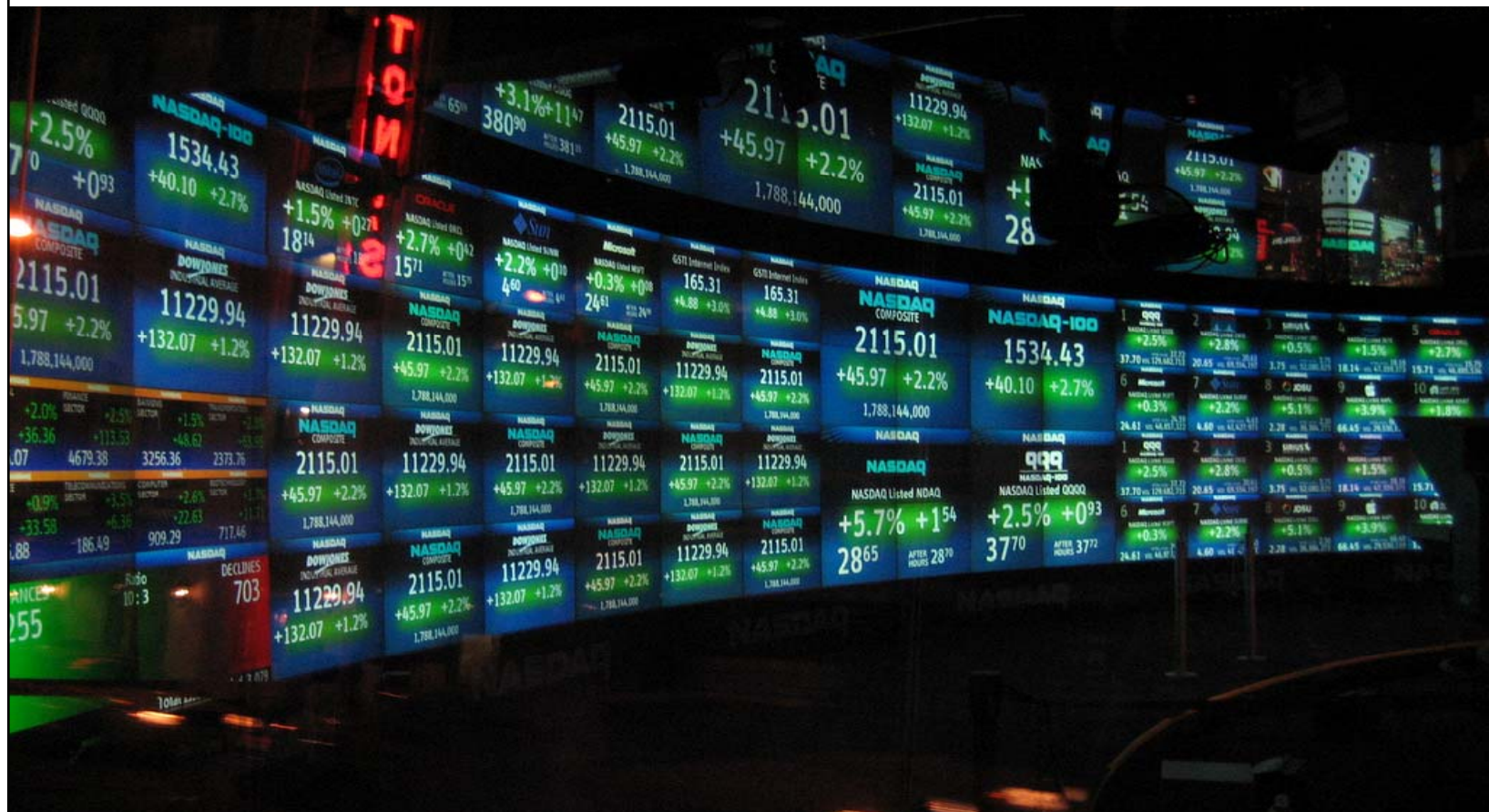
II. Scalability of Algorithms



III. User Data & Analytics



IV. Data Exchanges & Ecosystems



Predictive Analytics: The Consumer Crystal Ball

Omar Tawakol, Chief Executive Officer, Bluekai

Matthew J. Barkoff, Vice President, Badgeville

Scott Burke, Senior Vice President, Yahoo!

Theresa Gouw Ranzetta, Partner, Accel Partners

Moderator: Michael E. Driscoll, Metamarkets

MIT/Stanford Venture Lab | February 15, 2011