

MAD Analytics and
other
Interesting Trends

Keith Budge

Advisor CloudByte Inc.

VP and Managing Director Progress Software

Traditional Datawarehousing

Source, Extract, Normalize, Cleanse, De-Dupe,
Load, Aggregate, Analyze, Report

Expensive, Slow, Historic

Some Value

Datawarehousing Trends

Big data, tens, hundreds, thousands of terabytes

Very fast load and query speeds

Lower cost

Mostly still structured data or simple unstructured data (CDRs, IP logs, etc)

Google/MapReduce

MAD

Magnetic, Agile, Deep

- “MAD Skills: New Analysis Practices for Big Data”
 - Jeffrey Cohen, Greenplum
 - Brian Dolan, Fox Interactive Media
 - Mark Dunlap, Evergreen Technologies
 - Prof. Joseph M. Hellerstein, U.C. Berkeley
 - Caleb Welton, Greenplum

Data Scientists

HADOOP

Massive Volume

Extreme Low Cost (parallel architecture)

Complex Unstructured Data

Advanced Analytics

Open Source/Google/MapReduce

Lots of Start-Ups. Cloudera, Hortonworks, MapR

Platfora-Hadoop Analytics for the Enterprise

Convergence

Near Real Time/Real Time Analytics. “Rear View Mirror”

CEP Technologies and Approaches “Windscreen”
Forward View

Can We Predict the Future?

Demographics. Statistical Techniques. Etc.