

The Data Warrior

Changing the world, one data model at a time. How can I help you?

AGILE METHODS AND DATA WAREHOUSING:

HOW TO DELIVER FASTER



KENT GRAZIANO

 @KentGraziano |  kentgraziano.com



Agenda

- My Bio
- Why Agile & DW
- 12 Agile Principles
- Applying Agile Concepts to DW
- Two week iterations?
- Real World metrics
- Becoming Agile
- Conclusion

My Bio



- › **Senior Technical Evangelist**, Snowflake Computing
- › Blogger: *The Data Warrior*
- › Certified Data Vault Master and DV 2.0 Practitioner
- › Oracle ACE Director (BI/DW)
- › Data Modeling, Data Architecture and Data Warehouse Specialist
 - › 30+ years in IT
 - › 25+ years of Oracle-related work
 - › 20+ years of data warehousing experience
- › Former-Member: Boulder BI Brain Trust (<http://www.boulderbibraintrust.org/>)
- › Author & Co-Author of a bunch of books
- › Past-President of Oracle Development Tools User Group and Rocky Mountain Oracle User Group

Why Agile & DW?

- Perceptions
 - DW too slow to produce results
 - DW projects “fail”
 - DW projects fail to adapt to business changes
- Goal
 - Change perceptions
 - Deliver value
 - Be more adaptable and flexible

Objectives

- Learn to apply some agile ideas to data warehouse and business intelligence efforts
- Answer the question – can we deliver results faster?

Manifesto for Agile Software Development

<http://agilemanifesto.org>

We are uncovering better ways of developing software by doing it and helping others do it.

Through this work we have come to value:

Individuals and interactions over processes and tools

Working software over comprehensive documentation

Customer collaboration over contract negotiation

Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more

Kent Beck
Mike Beedle
Arie van Bennekum
Alistair Cockburn
Ward Cunningham
Martin Fowler

James Grenning
Jim Highsmith
Andrew Hunt
Ron Jeffries
Jon Kern
Brian Marick

Robert C. Martin
Steve Mellor
Ken Schwaber
Jeff Sutherland
Dave Thomas



THE PRINCIPLES OF AGILE

Principle #1

- Our highest priority is to satisfy the customer through **early** and **continuous** delivery of valuable software.
 - Who is the customer?
 - What is “valuable software” in data warehousing?
 - BI reports
 - Dashboard interface
 - Working ETL code?
 - In the context of the customer!
 - Once we start – keep going!

Principle #2

- Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.
 - User Stories with Prioritized Backlog
 - Must be flexible and adaptable in thinking and design
 - Use code generators
 - Agile data engineering techniques
 - Flexible and performant data platform

Principle #3

- Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale.
 - Need good scope control!
 - Prioritized Backlog
 - One small chunk at a time
 - Agile Data Engineering – Data Vault

Principle #4

- Business people and developers must work together daily throughout the project.
 - DW MUST have the business involved
 - One of the Top 10 reasons for failure
 - This applies for BI reports
 - Daily interaction would be great!
 - But – politics and priorities may interfere!
 - At HP GBI/EDW – we used “war” room
 - At MSH – daily standups with Biz

Principle #5

- Build projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done.
 - Need people who WANT to be on the project
 - Lose the dead weight
 - Get training if needed
 - Keep units of work small to create an atmosphere of success
 - Don't try a Big Bang EDW

Principle #6

- The most efficient and effective method of conveying information to and within a development team is face-to-face conversation.
 - Daily team huddles
 - Co-located work space
 - While face-to-face is efficient, still need some documentation (or meta-data) for later
 - Use a tool like JIRA, VersionOne, Rally, LeanKit

Principle #7

- Working software is the primary measure of progress.
 - Applied to DW:
 - What is “working software?”
 - BI reports?
 - Tables definitions and working ETL code?
 - Think more broadly – it is not just a data entry screen

Principle #8

- Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely.
 - DW Programs last a long time – don't burn the team out with unreasonable deadlines
 - See P#5 – Motivated individuals
 - Good planning and scope control
 - No all nighters!
 - Smallest valuable unit of work possible (MVP)
 - Keep it moving like a production line
 - Pick (or develop) a standard, repeatable methodology
 - Study the Agile methods and adopt what works for your team
 - Data Vault Modeling Methodology

Principle #9

- Continuous attention to technical excellence and good design enhances agility.
 - Bad design + bad architecture = trouble
 - Symptom: can't build a requested data mart
 - Set standards and use them!
 - Frequent design reviews a must
 - Improves team skills – provides cross training
 - Over time – better designs, shorter review cycles
 - Faster delivery

Principle #10

- Simplicity--the art of maximizing the amount of work not done--is essential.
 - KISS – Keep it Simple Stupid
 - Write less code by hand
 - Use code generators! (No syntax errors – ever)
 - Oracle SDDM, ERWin, Vertabelo
 - Talend, SSIS
 - AnalytixDS, WhereScape RED
 - Modifications are easier – just regenerate the code
 - Virtualize initial reporting layers/data marts
 - Eliminate infrastructure admin & tuning
 - Agile DevOps!

Principle #11

- The best architectures, requirements, and designs emerge from self-organizing teams.
 - Team of smart, motivated people = success
 - We succeed (or fail) as a TEAM
 - Don't micro manage or pigeon-hole staff
 - Encourage team work and team thinking
 - Staff will gravitate to roles based on skills, interest, and personality
 - Then they have more buy-in to the process
 - Eliminates delays and bottlenecks by having shared responsibilities (no single point of failure)

Principle #12

- At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.
 - Retrospectives are a MUST!
 - Ask 3 questions:
 - What went right?
 - What went wrong?
 - What will we do different next sprint?
 - Related to self-organizing teams
 - Make finding the solution to a problem the team's problem
 - More buy-in to the solution

USING AGILE CONCEPTS

Agile Concepts for DW

- Team Huddles (Morning Scrum)
- Extreme Programming

Team Huddles

- Daily Standup Meeting
 - AKA Scrum
 - Morning Roll Call (FDD)
- Short meeting (< 15 minutes)
 - Every morning, mandatory attendance
 - Review assignments, accomplishments, backlogs, blockers
 - 3 Questions
- Immediate feedback and assistance
 - Keeps team motivated and on track (P #5)
 - Identifies constraints and bottlenecks early in the process
 - Eliminates backlogs more quickly via re-assignments
- Improves team work
- Supports self-organizing teams (P #11)

Extreme Programming (XP)

- Programmer works directly with the end user
 - At MSH used conference room or developer's cube
 - At HP used Virtual Classroom or NetMeeting
 - At DFA used WebEx
- In DW:
 - Best with developing BI reports,dashboards, visualizations
 - DW or data mart must already be populated
- Works for ETL as well!
 - Used war room with business to get near instant validation of ETL changes

Two week iterations?

- Goal is really a few weeks to a few months (see P#3)
- What is the deliverable?
 - A fact table for a star schema
 - A dimension table
 - A complete star (fact and all dimensions)
 - One piece of ETL code that populates a fact table
 - A function needed by the ETL code
 - A new report or query
- Who is the customer?
 - BI programmer?
 - Knowledge worker?
 - ETL programmer?

REAL WORLD METRICS

HP EDW Examples

- Business found missing report elements
- Solution: modify 3 tables to add 5 new columns in reporting model (star schema)
- Tasks:
 - Document requirements and ETL specs
 - Modify Logical & Physical model (w/peer review)
 - Rebuild tables in development
 - Develop and test ETL
 - MTI (Move To Integration) tables and code
 - Execute and test ETL
 - Modify report in UAT environment & test
- Result: Revised report ready in **18 hours, 44 minutes**
 - Less than 1 business day
- 2nd case: 6 tables, 16 new columns
 - Ready for UAT in 72 hours
- How? War room with Biz & IT Team



BECOMING AGILE IN THE REAL WORLD

Getting to Agile

- “better than average expertise”
 - Expert consulting and mentoring
 - Do the work (OTJ)
- At Denver Public Schools
 - Took two years before we could try being more “agile”
 - Needed experience in Data Warehousing, Oracle Designer, OWB, and the “process” of building, deploying, and maintaining an Oracle DW

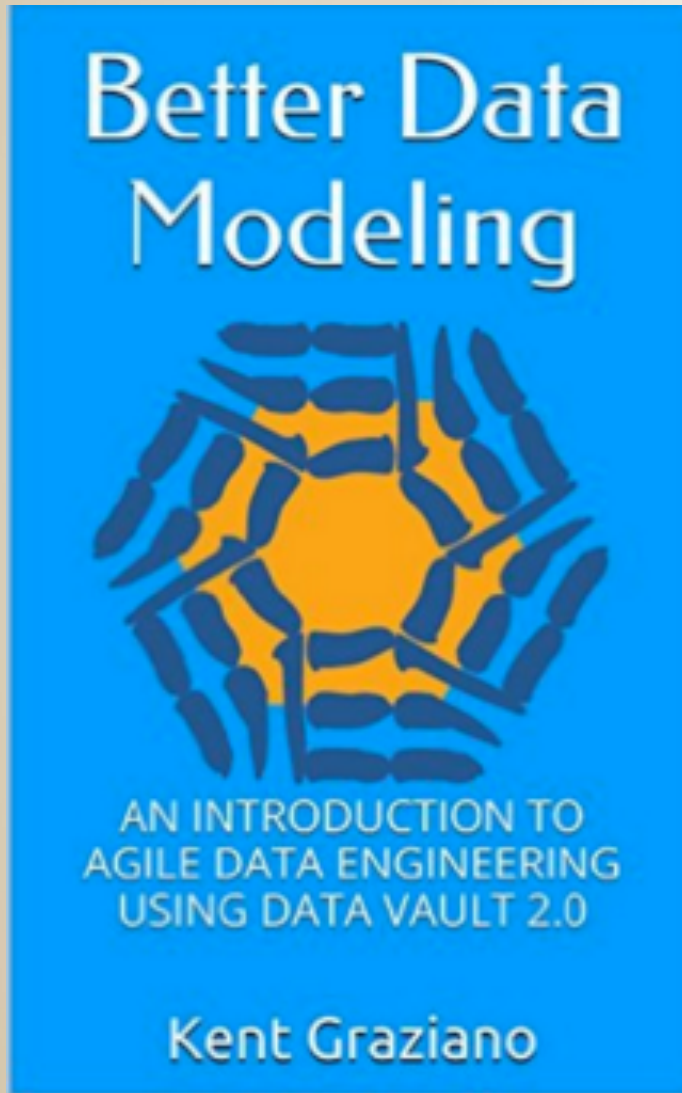
Getting to Agile

- At HP GBI/EDW it took about a year
 - Needed the right team and the right management support
 - Also the right project with willing business users
- At McKesson over a year setting standards and training staff
 - Built templates in JIRA for repeatable tasks
 - Then 3 week sprints
- At DFA – Day 1!
 - Experienced team with agile BI/DW coach
 - Used modified KanBan

Getting to Agile

- Cannot instantly become “agile”
- Must develop an agile mindset & culture
- Implement agile tools, processes, and methods
- Get an Agile Coach if necessary
 - Specifically an Agile BI/DW or Data Coach

SHAMELESS PLUG: My Intro Book



Available on
Amazon.com

<http://www.amazon.com/Better-Data-Modeling-Introduction-Engineering-ebook/dp/B018BREV1C/>

Other References

- Agile Management for Software Engineering: Applying the Theory of Constraints for Business Results by David J. Anderson
- CASE Method Fast-track: A RAD Approach by Richard Barker & Dai Clegg
- The Goal by Eliyahu M. Goldratt
- The Business of Data Vault Modeling by Dan Linstedt, Kent Graziano, & Hans Hultgren
- Super Charge Your Data Warehouse by Dan Linstedt
- Test Automation by Lynn Winterboer
<http://www.slideshare.net/AgileDenver/lynn-winterboer-test-automation>

Conclusion

- Agile concepts can be applied to data warehouse and BI projects
 - Not a purist definition!
 - Try to apply the principles – be creative
- Suggested approaches
 - Use team huddles
 - Use pair programming to increase quality and cross training
 - Use agile data platform - Snowflake Elastic DW!
 - Agile data modeling - Data Vault 2.0!
 - Use code generators
 - Read about Agile Methods
- Be flexible and give it a try

CONTACT INFORMATION

KENT GRAZIANO

Snowflake Computing
Snowflake.net



kent.graziano@snowflake.net



@KentGraziano



<http://kentgraziano.com>



The Data Warrior

Changing the world, one data model at a time. How can I help you?