DATA DONE RIGHT

Applying Agile and XP Concepts to Enterprise Data Integration

#DataDoneRight
ABOUT THE SPEAKERS

JOE FRAZIER
- 12 years data integration experience
- Designer/Architect for enterprise standards
- Data Technologies Capability Lead
- frazij10@nationwide.com

RANDY LOUSHIN
- 6 years data integration experience
- Agile, Test Automation, Test Driven Development
- Technical lead
- loushir1@nationwide.com
IT/BUSINESS COLLABORATION

Give them exactly what they want
IT/BUSINESS COLLABORATION

Lack of business involvement often leads to…
IT/BUSINESS COLLABORATION

Lack of business involvement often leads to…

Where are the sales metrics I asked for?
IT/BUSINESS COLLABORATION

Lack of business involvement often leads to...

Missing Features

Where are the sales metrics I asked for?

Incorrect data

2 + 2 = 6
Lack of business involvement often leads to…

Where are the sales metrics I asked for?

- Missing Features
- Incorrect data
- Unneeded data

Name: John Doe
Address: 123 ABC St, Anytown, USA
Email: ABC@gmail.com
Phone: (555) 555-5555
Pets: 2 cats, 2 dogs, 1 fish
Lack of business involvement often leads to...

**Missing Features**

Where are the sales metrics I asked for?

**Incorrect data**

**Unneeded data**

**Rework**
Lack of business involvement often leads to...

**Missing Features**

Where are the sales metrics I asked for?

**Incorrect data**

**Unneeded data**

**Rework**

**Delayed Delivery**

Name: John Doe
Address: 123 ABC St, Anytown, USA
Email: ABC@gmail.com
Phone: (555) 555-5555
Pets: 2 cats, 2 dogs, 1 fish
IT/BUSINESS COLLABORATION

Eliminate these issues by…
IT/BUSINESS COLLABORATION

Eliminate these issues by...

Co-location
IT/BUSINESS COLLABORATION

Eliminate these issues by…

Co-location

3 Amigos
IT/BUSINESS COLLABORATION
Eliminate these issues by…

Co-location

3 Amigos

Iteration Planning
IT/BUSINESS COLLABORATION

Eliminate these issues by…

Co-location

Card sign-off

3 Amigos

Iteration Planning
IT/BUSINESS COLLABORATION

Eliminate these issues by...

Co-location

Show & Tell

Card sign-off

3 Amigos

Iteration Planning
DELIVERY MODEL CHANGE

Faster delivery, better results
DELIVERY MODEL CHANGE
DELIVERY MODEL CHANGE

1 Year Release

- Long Scoping – Analysis Paralysis
- Huge Development Effort – Risk!
- Testing At End – Probably Get Cut!
DELIVERY MODEL CHANGE

Smaller Scope
Release Faster
ITERATIVE, NOT INCREMENTAL!

UNTIL THE FINAL RELEASE...

- You cannot get feedback on overall solution
- Nothing is usable
ITERATIVE, NOT INCREMENTAL!

WITH THIS APPROACH...

• Something usable immediately
• Identify desired changes/updates early
BUILDING A CRAFTSMAN

Supercharge your associates
BUILDING A CRAFTSMAN
BUILDING A CRAFTSMAN

Dev/Test Disconnect
BUILDING A CRAFTSMAN

Dev/Test Disconnect

Arguments

No! Your code is bad!
Your tests are wrong!
BUILDING A CRAFTSMAN

Dev/Test Disconnect

Dev/Test/Repeat

Arguments

No! Your code is bad!

Your tests are wrong!
BUILDING A CRAFTSMAN

Dev/Test Disconnect

Dev/Test/Repeat

Personnel Dependency

Arguments

Your tests are wrong!

No! Your code is bad!
BUILDING A CRAFTSMAN
BUILDING A CRAFTSMAN

WHAT DO WE WANT?

• An employee that is skilled in both development and testing...a software craftsman
BUILDING A CRAFTSMAN

WHAT DO WE WANT?

• An employee that is skilled in both development and testing…a software craftsman

HOW TO ACHIEVE

• Paired programming
• Test Driven Development
• Mindset change
• Expand hiring reach
BUILDING A CRAFTSMAN

BENEFITS
BUILDING A CRAFTSMAN

BENEFITS

- Break down the wall between developers and testers
- Better understanding of the requirements
- Rigorous testing = higher quality code
- Anyone can test, anyone can develop
- Build a more skilled employee
BUILDING A CRAFTSMAN

BENEFITS

- Break down the wall between developers and testers
- Better understanding of the requirements
- Rigorous testing = higher quality code
- Anyone can test, anyone can develop
- Build a more skilled employee
BUILDING A CRAFTSMAN

BENEFITS

Break down the wall between developers and testers

Better understanding of the requirements

Rigorous testing = higher quality code

Anyone can test, anyone can develop

= Build a more skilled employee
BUILDING A CRAFTSMAN

**BENEFITS**

Break down the wall between developers and testers

Better understanding of the requirements

Rigorous testing = higher quality code

Anyone can test, anyone can develop

= Build a more skilled employee
BUILDING A CRAFTSMAN

**BENEFITS**

- Break down the wall between developers and testers
- Better understanding of the requirements
- Rigorous testing = higher quality code
- Anyone can test, anyone can develop
  
  =

  Build a more skilled employee
LIFE OF A STORY CARD

Demo
PROBLEM STATEMENT

Marketing needs to understand more about their customers.

Data is everywhere – most of it in locations they cannot access.

Need a consolidated place for customer information.
PROBLEM STATEMENT

Marketing needs to understand more about their customers.

Data is everywhere – most of it in locations they cannot access

Need a consolidated place for customer information
PROBLEM STATEMENT

Marketing needs to understand more about their customers.

Data is everywhere – most of it in locations they cannot access.

Need a consolidated place for customer information.
<table>
<thead>
<tr>
<th>EPICS</th>
<th>IN ANALYSIS</th>
<th>READY FOR BUILD</th>
<th>ITERATION BACKLOG</th>
<th>IN BUILD</th>
<th>READY FOR SIGNOFF</th>
<th>DONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER INFO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
STORY BREAKDOWN
STORY BREAKDOWN

Customer Name and Location
As a marketing analyst, I want to know my customers name and location information, so that I can understand who my customers are and where they live.
As a marketing analyst, I want to know my customers name and location information, so that I can understand who my customers are and where they live.
As a marketing analyst, I want to know my customers name and location information, So that I can understand who my customers are and where they live.

As a marketing analyst, I want to know my customers preferred contact method, So that I can effectively communicate with them.
# REQUIREMENTS GATHERING

<table>
<thead>
<tr>
<th>Customer Project VMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPICS</strong></td>
</tr>
<tr>
<td>CUSTOMER INFO</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
REQUIREMENTS GATHERING

Customer Name and Location
Given new customer information is available
When the customer profile has been loaded
Then I will know my customer's name
And view my customer footprint by location
# Data Profiling

## Profile_AGILE2015_CUSTOMER_DETAIL - Column Profiling

<table>
<thead>
<tr>
<th>Name</th>
<th>Unique</th>
<th>% Unique</th>
<th>NULL</th>
<th>% Null</th>
<th>Datatype</th>
<th>% Infer</th>
<th>Documented Data</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER_ID</td>
<td>100</td>
<td>100.00</td>
<td>-</td>
<td>-</td>
<td>Integer(10)</td>
<td>100.00</td>
<td>number(15)</td>
<td>62564752</td>
<td>9736333...</td>
</tr>
<tr>
<td>NAME</td>
<td>100</td>
<td>100.00</td>
<td>-</td>
<td>-</td>
<td>String(18)</td>
<td>100.00</td>
<td>string(100)</td>
<td>Ankundi...</td>
<td>Zieme,Z...</td>
</tr>
<tr>
<td>ADDRESS</td>
<td>100</td>
<td>100.00</td>
<td>-</td>
<td>-</td>
<td>Date Time</td>
<td>100.00</td>
<td>string(500)</td>
<td>10222 F...</td>
<td>9851 Gu...</td>
</tr>
<tr>
<td>DOB</td>
<td>100</td>
<td>100.00</td>
<td>-</td>
<td>-</td>
<td>Date Time</td>
<td>100.00</td>
<td>string(38)</td>
<td>1915-11...</td>
<td>1996-10...</td>
</tr>
</tbody>
</table>

### Values

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Fre.</th>
<th>Perc.</th>
<th>Chart</th>
<th>Drill down</th>
</tr>
</thead>
<tbody>
<tr>
<td>X(7),X(7)</td>
<td>5</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(5),X(7)</td>
<td>5</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(6),X(8)</td>
<td>5</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(6),X(7)</td>
<td>7</td>
<td>7.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(5),X(5)</td>
<td>8</td>
<td>8.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(7),X(6)</td>
<td>9</td>
<td>9.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X(6),X(6)</td>
<td>10</td>
<td>10.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>51</td>
<td>51.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## MAPPING RULES

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>RULE</th>
<th>TARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer_Detail.Customer_Id</td>
<td>Straight move</td>
<td>Customer_Profile.Customer_Id</td>
</tr>
<tr>
<td>Customer_Detail.Name</td>
<td>Take string after comma &amp; uppercase (source format is LastName, FirstName)</td>
<td>Customer_Profile.F_Name</td>
</tr>
<tr>
<td>Customer_Detail.Name</td>
<td>Take string before comma &amp; uppercase (source format is LastName, FirstName)</td>
<td>Customer_Profile.L_Name</td>
</tr>
<tr>
<td>Customer_Detail.Address</td>
<td>Split on comma, extract street &amp; uppercase (source format is street,city,state,zip)</td>
<td>Customer_Profile.Street</td>
</tr>
<tr>
<td>Customer_Detail.Address</td>
<td>Split on comma, extract city &amp; uppercase (source format is street,city,state,zip)</td>
<td>Customer_Profile.City</td>
</tr>
<tr>
<td>Customer_Detail.Address</td>
<td>Split on comma, convert state to 2 chars (source format is street,city,state,zip)</td>
<td>Customer_Profile.State</td>
</tr>
<tr>
<td>Customer_Detail.Address</td>
<td>Split on comma, extract zip (source format is street,city,state,zip)</td>
<td>Customer_Profile.Zip</td>
</tr>
</tbody>
</table>
### READY FOR BUILD

<table>
<thead>
<tr>
<th>EPICS</th>
<th>IN ANALYSIS</th>
<th>READY FOR BUILD</th>
<th>ITERATION BACKLOG</th>
<th>IN BUILD</th>
<th>READY FOR SIGNOFF</th>
<th>DONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER INFO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Name and Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Contact Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GET READY FOR THE SPRINT
GET READY FOR THE SPRINT

Design Session
GET READY FOR THE SPRINT

Design Session

Sizing
GET READY FOR THE SPRINT
HIGH LEVEL DESIGN
# DESIGN, TEST, DEVELOP

## Customer Project VMS

<table>
<thead>
<tr>
<th>EPICS</th>
<th>IN ANALYSIS</th>
<th>READY FOR BUILD</th>
<th>ITERATION BACKLOG</th>
<th>IN BUILD</th>
<th>READY FOR SIGNOFF</th>
<th>DONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER INFO</td>
<td>Customer Contact Information</td>
<td>Customer Name and Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HIGH LEVEL DESIGN
HIGH LEVEL DESIGN

ACCEPTANCE TEST
### Design, Test, Develop

<table>
<thead>
<tr>
<th>Customer Project VMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPICS</strong></td>
</tr>
<tr>
<td>CUSTOMER INFO</td>
</tr>
</tbody>
</table>

**Notes:**
- Customer Contact information

LET’S START CODING!

Red, Green, Refactor

http://youtu.be/Yw utZRePChk
<table>
<thead>
<tr>
<th>EPICS</th>
<th>IN ANALYSIS</th>
<th>READY FOR BUILD</th>
<th>ITERATION BACKLOG</th>
<th>IN BUILD</th>
<th>READY FOR SIGNOFF</th>
<th>DONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER INFO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Contact Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer Name and Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CONTINUOUS INTEGRATION

Test, Develop, Commit!
CONTINUOUS INTEGRATION
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors.
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors

- Full, automated regression testing
- Find defects quicker
- Complete picture of your system – what’s working, what’s not
- Reduced time in system testing
- System upgrades are painless
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors

- Full, automated regression testing
- Find defects quicker
- Complete picture of your system – what’s working, what’s not
- Reduced time in system testing
- System upgrades are painless
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors.

- Full, automated regression testing
- Find defects quicker
- Complete picture of your system – what’s working, what’s not
- Reduced time in system testing
- System upgrades are painless
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors.

- Full, automated regression testing
- Find defects quicker
- Complete picture of your system – what’s working, what’s not
- Reduced time in system testing
- System upgrades are painless
CONTINUOUS INTEGRATION

The process of committing new code into a shared repository, running an automated build, and reporting on any errors.

- Full, automated regression testing
- Find defects quicker
- Complete picture of your system – what’s working, what’s not
- Reduced time in system testing
- System upgrades are painless
## Continuous Integration

### Jenkins

<table>
<thead>
<tr>
<th>S</th>
<th>W</th>
<th>Name</th>
<th>Last Success</th>
<th>Last Failure</th>
<th>Last Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><code>app_acquire_eparty</code></td>
<td>4 days 2 hr - #5</td>
<td>16 hr - #10</td>
<td>16 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3311</code></td>
<td>15 hr - #14</td>
<td>N/A</td>
<td>2 min 16 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3312</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>2 min 30 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3313</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>2 min 28 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3314</code></td>
<td>15 hr - #11</td>
<td>N/A</td>
<td>2 min 26 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3315</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>2 min 3 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3316</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>2 min 30 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3317</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>3 min 12 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3318</code></td>
<td>15 hr - #10</td>
<td>N/A</td>
<td>5 min 45 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3319</code></td>
<td>15 hr - #9</td>
<td>N/A</td>
<td>4 min 30 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>app_integrate_eparty_3320</code></td>
<td>15 hr - #9</td>
<td>N/A</td>
<td>2 min 55 sec</td>
</tr>
</tbody>
</table>
CONTINUOUS INTEGRATION

Jenkins

**Project app_acquire_eparty**

- **Recent Changes**
- **Latest Test Result** (7 failures / +4)

**Upstream Projects**

- Perl-EParty

**Downstream Projects**

![Test Result Trend](just show failures) enlarge
QUESTIONS??

Joe Frazier – frazij10@nationwide.com
Randy Loushin – loushir1@nationwide.com