So the New Operational System Is Going to Destroy All Your Existing Reporting: Managing Requirements and Quality in an Agile Large-Scale EDW Project

> Richard Tench TDWI Vancouver Chapter March 4, 2016



building trust. driving confidence.

Agenda

- Background: about ICBC and INFO Release 2
- Managing requirements on a large scale agile project
- Quality assurance and test automation

About ICBC and its Transformation Program

- The Insurance Corporation of B.C. is a provincial crown established in 1973 to provide universal auto insurance to B.C. motorists
- ICBC's multi-year Transformation Program (TP) replaces aging technology and renews business processes to help provide better customer service, faster claims settlement, mitigation of future rate changes and lower operating costs
 - Most TP projects have been completed (e.g. new rating engine, modern integrated Claims Management system)







About INFO

- INFO is ICBC's enterprise business information project
- It is responsible for developing and implementing an Enterprise Information Strategy* to:
 - Enable timely fact-based decisions
 - Provide a single version of the truth
 - Deliver an enterprise view of all information assets
- The INFO project implements the core infrastructure, enabling technologies, business disciplines, and data governance to deliver the Enterprise Information Strategy.



*See appendix



About the Enterprise Data Warehouse

- The EDW integrates a broad array of legacy and modern source applications
 - INFO Release 1 (claims focus) integrated 13 source applications. Subject areas are:
 - Claim, Policy, Product, Vehicle, Person/Org, Driver, Broker, Human Resource, Supplier, Contravention, Reference Data, User Data
 - INFO Release 2 (insurance focus) integrates five additional applications and reuses/enhances eight of the R1 sources
 - New subject areas: Policy, Account, Payment, Legacy Claim,
 - Other projects integrate four additional source applications, with another four underway
- The EDW architecture:





INFO Release 2 statement of vision

Provide to the enterprise *essential* reporting & extracts using Policy and/or Vehicle data integrated with enterprise data

- *Essential* is defined as both:
 - The information is mandatory, as determined in prioritization, and
 - Already have the information today or intrinsic to new processes introduced in ISAS
- Operational reporting is included, i.e. not just performance / trending / strategic reports
- Solution types are:
 - Reports (Web Intelligence)
 - Data Exploration (Tableau)
 - Extracts (DB2 or flat file) are for advanced analysis, e.g. SAS and specialized actuarial tools
 - Ad hoc reporting capability for Business Insights team



Business representation on INFO R2



Business Prime advocates at executive level

Each Change Board member represents division, managing scope and priority

Product Owners provide requirements and represent the "Who" in their user stories

Business SMEs clarify requirements



Gathering enterprise user stories ain't easy

- So the vision was set, but how to go about gathering the report-level requirements?
- There are currently approximately 1000 existing reports based on legacy policy/premium data sources
- Almost every single report will break when the new COTS system goes live
- Current reports provide considerable operational and informational value
- We could...
 - Assess each of the 1000 reports: who uses it, why do they use it?
 - Turn them off and see who complains



- Start with an empy slate...zero-based requirements
- Find out what truly matters



About user stories

- So we
 - Requested and received divisional reps: the Change Board
 - Canvassed for Product Owners, and ended up with 19 (and 25 SMEs)
 - Conducted 25 sessions with owners and SMEs, asking the key questions:
 - Who are you? What do you need? Why do you need it?
 - Resulting in 178 user stories... Yikes!!! Examples:

010	a	Vane	Tayer,	
I	want	my Rock	tet to m	ove back
and	torth arrow	When ws	1 press	left and
rian				

Name	Who	What	Why
Terminated Prorate Plates	Supervisor, Prorate Licensing	Daily list of prorate plates that have been terminated by cancellation, replacement, or substitution	To determine whether an 'exit tax' needs to be assessed
Y-List	Risk Underwriting Manager, Supervisor, and Insurance Rating Analyst	All new and returning resident policies from out of province (OOP) that have provided OOP claim letters to justify a CRS discount (i.e. the "Y" or "N" list)	To determine their eligibility for a discount and track outcomes
Risk exposure & market penetration analysis	Product Manager	Written and earned exposures, and penetration rates by product, broker, vehicle and customer dimensions	To determine what products to sell or not sell

How did we arrive at scope?

- Initial cull: Product Owners slotted user stories into four Priorities as follows:
 - 1. Must be delivered at COTS go-live
 - 2. Essential to have, but may be delivered later
 - Product owners specified Timing Requirements (relative or absolute)
 - 3. Didn't make the cut ("nice to have")
 - 4. Not required
- Approximately 125 stories remained at Priority 1 (P1) or Priority 2 (P2)
- Significant challenge with the new application was that few understood the specific impacts on requirements
- This process was repeated a couple times in the project, reducing to ~100 stories
 - Change Board was critical for this
 - We had the assistance of business specialists from the operational project





How did we manage scope going forward?

- Scope and priority changed were reviewed by Change Board, with final approval by Business Prime and Executive Sponsor
- As a result of budget issues, overall scope was cut three further times
- How did we know what to cut?
 - The Change Board representatives placed a value on each user story, using Value Points
 - Each rep was given 20 value points per story to allocate however he or she wanted
 - All user stories and underlying developer stories were estimated in Story Points
 - Subsequent scope reductions were done using value-for-cost as the main metric
 - Stories were grouped into Scope Packages based on architectural requirements, e.g. new source, new ASA, etc.
 - Red line was drawn to match budget
 - Change Board evaluated stories below the red line, and slightly above the red line

Retain: higher value-for-cost

De-scope: lower value-for-cost

None of this would have worked well if the business areas hadn't been highly cooperative.



How did we analyze the user stories?

- We used CA Agile (formerly known as Rally) to manage agile development
 - Cloud-based _
 - Earlier COTS projects used Rally
- CA Agile can be customized at a global level, or at a project level
 - We did both

—

- Example user story fields:
 - Standard: ID, Name, Owner, Plan Estimate (cost), Description
 - Custom: Change Board Rep, SMEs, Change Status, Change Type, Change Target, Acceptance Level
- We built extensive analytics using Excel pivot tables •





INFO R2 quality assurance approach

Sprint

- Each package of work is • pipelined.
- In Developers work ٠ stage ("In-Sprint"):
 - Test-led development
 - Unit testing
 - Integration testing, including regression
 - Demo (UAT level 1)
- In System Test work • stage:
 - Volume system testing, including regression
 - Demo (UAT level 2) _



*Demo



Test Criteria

In-Sprint testing is focused on user experience, as data is constructed

System testing is focused on user experience **and** correctness, as historical data is based on production data

	In Sp	print	Syst	em Test Sprint
Reports (and	User	Experience	User	Experience and Data Correctness
<u>other</u>	1.	Is the report look & feel easy and	1.	Is the report look & feel easy and intuitive?
<u>outputs)</u>	2. 3. 4. 5.	intuitive? Is the report consistent (per ICBC standards for report UI)? Are report filters and sorting relevant and meaningful? Is the data displayed in a structured and meaningful way that is usable by the end users Check on basic business rules to ensure the correctness of derived Data Are the reports rules and business labels as per DGO	2. 3. 4. 5. 6. 7.	Is the report consistent (as per any ICBC standards for report UI)? Are report filters and sorting relevant and meaningful? Is the data displayed in a structured and meaningful way that is usable by the end users Check on basic business rules to ensure the correctness of derived Data Are the reports rules and business labels as per DGO Are NFRs being met in full volume Data environment
Intermediate	Does	Data Make sense?	Is Da	ta Integrated?
<u>Data</u>	ata Are business rules being applied correctly throughout the transformation process? Integration Tests adequate? Do the test results presented by project QA give you confidence in the EDW quality of		Are ti all re	he transactions / records complete and with correct data from spective source systems of record?
			Syste	em Tests Adequate? Do the test results presented by Project QA team give you
	code?	>		 confidence in the quality of the reports operational set-up in following areas: a. execution sequence, b. performance, and c. dependency & recovery



INFO R2 quality assurance: test automation

- Data warehouses require extensive testing
 - R2 has 3,350 test cases
- To manage the workload and to provide quick feedback to developers, some kind of automation was required
- We reviewed the market and selected Zuzena
- Benefits the team found from test automation:
 - Reusability of test code across multiple teams and multiple schemas
 - Test repository is one-stop shop for test data, tests (SQL), and historical results
 - High volume regression testing took about half the time
 - Selective regression testing
 - Able to manage and automate loading of five separate test data sets
 - Automation of test results: able to track historical results of each test case
 - We use the tool for performance monitoring of the SQL scripts helped us tune for quicker test times



Diagram copyright © 2016 Elsevier Inc. All rights reserved

"The tool has been very effective for us in automating the integration and system testing streams. We have automated all of our ETL data load and validation across the three architectural layers."



INFO R2 Quality Assurance: challenges and how we dealt with them

Challenge	Response	Result
Finding data issues in reports is too late, resulting in considerable rework	Have business validate data well in advance	Lower defect rates on reports, less rework
Large production data volumes	 Use sub-setting scripts, and add/revise data to fit test cases and scenarios. Monitor performance for all runs, tuning as needed. Process one year at a time, if needed 	 Reduced handholding effort Improved performance for initial loads in particular High volume loads took less preparation t
Custom-masked test data broke RI and was unfriendly	Used InfoSphere Optim for masking	 Optim maintains the mapping of the real data to the masked data (no RI problems). Optim generates names that look like names.
Regression testing often took one to two weeks	Automate testing, running full regressions several times per sprint	 In-sprint regression testing typically took a few hours. System testing continued to take days or weeks due to data volumes
Untested Core columns (when not needed in reports)	Automate testing	Able to test all Core columns, even where not required for the in- scope reports reports
Required multiple test focus areas, e.g. person/org, premium, etc.	Automate testing	Without intervention, Zuzena loops through the different test data sets, recording the results of each.
Constant data changes from COTS source	At first: establish tighter connection with point-of-sale project. Manually track drops. Live with rework. Later on: focus on legacy data, setting new COTS messages aside	 Able to accomplish ~15% less per sprint than originally envisioned Focused on dealing with substantial data complexity in our legacy applications
Lack of test data provided from COTS project	Use QTP on PolicyCenter UI to generate test data	Basic test data from PolicyCenter



Appendix



Scope of INFO and the Enterprise Information Strategy

= 13 sub-strategies

INFO Release 0

Create 13 information management strategies including EDW (see diagram at right). Move to a data driven decision making organization and support the delivery of the Transformation Program

INFO Release 1

Develop a new core enterprise data warehouse and capability. Provide essential reporting & extracts using ClaimCenter data integrated with enterprise data.

INFO Release 2

Provide essential reporting & extracts using Insurance and Vehicle data integrated with enterprise data.



