Strata: Data Wrangling for Insurance

Olivier de Garrigues, Solutions Lead EMEA
olivier@trifacta.com
Life as an actuary
“New technology is transforming the way we work, and it is allowing the competition to do better than what we can. The strange thing is we know the urgency, and yet there is inertia.”

Inga Beale, CEO of Lloyd’s of London (February 2017)
Conventional Approaches Inhibit User Empowerment

- Hand-Coding
- Technical Workflow Mapping
Self-service is now expected

How can I access the raw data?
What do you want to analyze?

Data Sources

I can’t tell you until I see the data – let me see the data first.
I can’t just point you to the raw data – you’ll need to tell me.

BUSINESS
IT
Over the past few years, **Self-Service Data Preparation** has emerged as a stand-alone category not simply a feature of Business Intelligence tools.
How to reconcile Insurance and Self-Service Data Preparation?

Or as we call it...

Data Wrangling
What is Data Wrangling?
Research Roots: Potter’s Wheel 2001

+ Real data, sampled on the fly
+ Menu-driven transforms
+ Immediate execution and feedback

[Raman & Hellerstein, VLDB01]
Research Roots: Open Source Data Wrangler, 2011

+ Predictive Transformation
+ Immediate feedback on multiple choices

[Kandel, Heer & Hellerstein, CHI 11]
Creating radical productivity for people who analyze data.
ANALYSIS & VISUALIZATION

Bridging the Gap Between Raw Data & Analysis

80% of the work in any data project is preparing the data for analysis.
Trifacta Approach: It’s All About The Experience
TRIFACTA WORKFLOW

1. Identify/Register Data
2. Predictive Interaction
3. Monitor and Adjust

Secure Access

Sample

Scale Up

Results

Schedulers

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Identify/Register Data

Sample

Results

Secure Access

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie

Visualization & Analysis

Consume

Predictive Interaction

Refine Sample

Schedule

Oozie
Insurance: Increase Wallet Share and Profitability of Existing Customers

Customer Data
Marketing Campaigns
Call Center, Webchat, Twitter
3rd Party Experian, Nielsen, FICO...
Web Logs
Insurance Agents/ Brokers

ANALYSIS & CONSUMPTION

TRIFACTA
Discovering Structuring  Cleaning  Enriching  Validating  Publishing

Ingestion  Processing  Storage
Insurances’ Big Data Projects
Driving Innovation and Business Efficiency

➔ Risk modeling
➔ Customer 360 for cross-sell and up-sell
➔ Claim processing and fraud detection
➔ Channel Attribution Modeling
  ➔ Market reach
  ➔ Influential drivers
➔ Quote Attribution
  ➔ Decision factors
  ➔ Prospect navigation
➔ Call center efficiency
➔ Telemetry
  ➔ New products and optimized pricing: i.e. pay per mile, teen control, drones controlling roof condition, etc.
  ➔ Customer reward
Challenges

• Increased pressure on underwriting profits due to volatile interest rates (*Interest rates have declined significantly since 2007-2008 global financial crisis)

• Competitive premiums require underwriters & actuaries to incorporate an increasing variety of diverse and complex data sources—telemetry, geospatial data, social media data, etc.

• UBI, or user-based insurance, relies upon IoT technologies and data to create personalized premium prices (*36% of insurers are projected to use UBI by 2020)
Streamline Policy Analysis Across Acquisitions

<table>
<thead>
<tr>
<th>BUSINESS CASE</th>
<th>TECHNICAL CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Must efficiently analyze claims for clients like Allianz, who transferred 1.1 billion in U.S. long-tailed liabilities to Enstar in 2016</td>
<td>• Lengthy Excel and Access macros procedures were hard to understand, modify, and update</td>
</tr>
<tr>
<td>• Inability to leverage variety of data across multi-billion dollar portfolio of insurance and reinsurance assets hinders growth</td>
<td>• Huge variety of data (Oracle, SQL, Excel files, etc.) from legacy solutions of insurance/reinsurance acquisitions</td>
</tr>
<tr>
<td>• Employ 1,300 people worldwide—needed self-service solution to leverage contextual knowledge of business users</td>
<td>• Business users couldn’t access raw data to drive deeper insights</td>
</tr>
<tr>
<td>• Must execute routine auditing to verify accurate premiums</td>
<td>• Lack of transparency from macros made it difficult to complete required audits</td>
</tr>
</tbody>
</table>
Benefits Using Trifacta

Benefits

• Automating preparation tasks in Trifacta improved policy analysis efficiency by 60%
• From more quickly identifying policy prices, Enstar has already seen profit gains of 10%
• Enstar has streamlined operations across its acquisitions to guarantee ROI for its portfolio
• Enstar has scaled technical resources and empowered people who know their data best
Trifacta Self-Service Data Preparation

**Editions**

- Hadoop Based
- Data Lake Initiatives
- Unlimited Volume & Scalability
- Enterprise Support
- Subscription Pricing

![Trifacta Wrangler Enterprise](image)

www.trifacta.com/start-wrangling

- Connected Desktop Application
- Up to 100mb Data Sets
- Community Support
- Free
We Wrote the Book on Data Wrangling (Literally)
THURSDAY, MAY 25th

10:45am - Book Signing
O’Reilly Bookstore Expo Booth #409
Get a free copy of the Principals of Data Wrangling Book!

12:05 to 12:45pm - Session
Capital Suite 15/16
“Driving the next wave of data lineage with automation, visualization & interaction”
Presented by: Sean Kandel, Co-Founder & CTO

Follow us: @Trifacta
Questions?

Download Trifacta Wrangler
trifacta.com/start-wrangling