



NAVIGATING THE INSURANCE DATA MAZE

Abstract

This point-of-view offers the insurance industry a fresh perspective on the common pitfalls in dealing with the massive amounts of data available today. It also outlines a roadmap to successfully overcome these challenges and achieve the meaningful insights the industry needs.



A problem of plenty

Every insurance carrier talks about master data management, and the importance of having quality of data. However, the problem starts at the source. To borrow a cliched phrase, it's a problem of plenty!

The insurance industry adopted data analytics in the form of actuarial analysis long before analytics became mainstream across other industries. However, in the 90's, digital transformation caused an explosion of insurance data with inputs coming in from various sources, which the industry wasn't equipped to handle well. Now with multi-modal operations and sales processes with a heavy focus on insight-driven sales, managing and making sense of all that data is the key challenge for the insurance industry CXOs.

And a problem of quality

The problem of insurance data quality was highlighted by a recent BCG report¹ which noted the conundrum facing the industry: ***"AI can transform the insurance industry, but to do so, it needs quality data – which is exactly what insurers often lack"***.

Typically, a data journey starts with the onboarding of an intermediary or a customer. However, right at that stage, the process being manual and non-standardized makes it error-prone due to the human in the loop. Further, due to the loopholes in de-duplication checks or sometimes due to business urgency inducing the complete bypassing of any checks, duplicate intermediary and

customer records get created mostly with inadequate or twisted data. This opens up a Pandora's box of inaccurate data grouping, mapping, and aggregation, resulting in data which is inadequately cleansed or which requires enrichment, rectifying all of which involves monumental effort.

Another common factor for the insurance carriers' data quality woes is their huge daily data inflow. The human subjectivity due to manual key-in processes only allows a very small subset of this data to get into their system organically, to be used for meaningful insights. As a result, leveraging data for effective, advanced analytical models visualizing cross-sell, up-

sell, fraudulent claims or other scenarios becomes a distant dream. For example, the poor quality of risk and claims data hinders actuaries from coming up with accurate models for functional risk, claims, and pricing.

There are other less common factors underlying data quality issues as well. For example, mergers and acquisitions between insured entities or brokers also leads to data mapping and aggregation issues. However, that is a completely different problem to solve and probably an easier one.

¹<https://www.bcg.com/publications/2022/why-ignoring-ai-on-insurance-is-risky-for-insurance-ceos>



Another problem of siloes

The problem of data siloes is often misunderstood as just data being stored in separate silos. But it runs much deeper than that. In fact, disparate data sources in various operational data warehouses hosted in the vendor clouds may very well be the most optimized way of having a distributed data organization. The actual challenge lies in connecting all these siloes to derive meaningful business insights, and this is where most carriers fail miserably. Typically, the reasons behind this are similar to those underlying the problem of plenty and the problem of quality, as discussed earlier.



Navigating through the data maze

As The Oliver Wyman Risk Journal puts it, the *“hallmarks of data quality include accuracy at the point of entry, completeness of fields, congruency as it flows through the institution, consistency of interpretation and a stable approach to storage over time”*.²

Thus, insurance carriers looking to navigate through their maze of data need to undertake steps in three critical areas.

1. Data onboarding: To achieve data-driven business insights using any form of analytics, first the quality of the data needs to be improved through its cleansing and enrichment. Typically, data cleansing consists of deduplication checks and rectifying the groupings such as through parent-child mapping. Next, enrichment is all about improving the quality of the data, such as adding the missing details in existing records, and ensuring data uniformity.

For both data cleansing and data enrichment, it is advised to use global data standards. For example, integrating entity records with Dun & Bradstreet or Standard & Poor's will help to have a unique global identification number in addition to enrichment through correct addresses and industry codes, along with other details like revenue, number of employees and other such key information. This solves a major

standardization pain point without which underwriters would work with their own version of the nature and details of the client's business, and which would typically be faulty.

2. Data Maintenance: Once the data has been thoroughly cleansed and enriched, the next important step is to ensure it remains clean. This requires avoiding any manual addition of entity records, except in the case of exceptions. The best business practices to follow so as to avoid human errors, involve OCR/RPA tool-led data ingestion with additional validation checks or integration with standard third-party data sources.

3. Data hosting: Lastly, the hosting of data needs to be managed better through migration and modernization of on-premise storage solutions to cloud-native data stores.

Also, the conventional approach to handle data siloes has been to create a data lake

on premise and to dump all data from operations into it. However, this is not only complex, but also expensive and inefficient. A common concern with this approach often shared by the insurers include performance issues in running simultaneous reports, apart from the typical problems of faulty reports, and inability to derive meaningful insights.

Rightly so, the Gartner Hype Cycle for Data management 2021³ suggests that, data lake as a concept/solution is passing through the trough of disillusionment. However, multi-cloud and inter-cloud data management with a data fabric architecture is showing the promise of achieving the required business insights through connecting distributed data stores. Thus, all major business intelligence vendors, such as Tableau and Tibco, now support multi-cloud data storage and connectivity. Even Microsoft Power BI and IBM Cognos – BI platforms who own their own clouds – support multi-cloud analytics environments as well.

²<https://www.oliverwyman.com/our-expertise/insights/2011/may/the-oliver-wyman-risk-journal-vol-i-data-quality--the-truth-isnt-out-there.html>

³<https://www.gartner.com/en/documents/4004072>

Tying it all up together

While efficient data hosting is the most complex part of the puzzle to navigate through the maze of data challenges that the insurance industry faces, data onboarding and periodic maintenance are of equal importance. Unless all these three areas are given equal focus, data becomes stale and its insights are rendered ineffective.

Authors



Pinaki Ray

Principal Consultant, Digital Transformation Services, Infosys BPM

Pinaki is working as Principal Consultant with Infosys BPM's Digital Transformation Services, responsible for Insurance and Healthcare – Digital solution design and Service delivery. Prior to this he has worked for Wipro Technologies, TATA Consultancy Services and ICICI Lombard general Insurance Company Ltd.

Pinaki is a Fellow of Insurance Institute of India with over 18 years of experience in Insurance Business & Technology Consulting around P&C, Life, and Healthcare domain.



Sourav Ghosh

Senior Industry Principal, Infosys BPM

Sourav is a Senior Industry Principal with Infosys BPM's Digital Transformation Services, responsible for Industry Solutions – Global Digital solution design and Service delivery. An IBM-certified Design Thinking practitioner, he advises organizations on their operations strategy, assists them in improving profitability and efficiency of business processes, and helps in executing business transformation through calibration of operating model and technology.

Prior to Infosys, Sourav had been with IBM, Satyam, Tata Consultancy Services and Standard Chartered Bank across a variety of roles in India, the U.S., and the U.K..

For more information, contact infosysbpm@infosys.com



© 2022 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.

[Infosysbpm.com](https://www.infosysbpm.com)

Stay Connected

