DataPorts

Solving End-to-End Value Chain
Content Integration
Index

Introduction ........................................ 3

1. Concept ......................................... 4

2. Approach ...................................... 6

3. The Anatomy of a DataPort .................. 8

4. Implementation ................................. 9

5. Getting Started ................................. 10

6. Scale ............................................ 10

7. Implications .................................... 11

8. Conclusions ................................... 12
Introduction

Product content and content quality are key to commerce, and as supply chains accelerate so the impact of errors escalate. When the wrong product is shipped to a consumer because of errors in an online catalogue it costs the consumer, retailer, wholesaler, and manufacturer in satisfaction, shipping, restocking, lost sales, and reputation. When content errors occur upstream of the consumer, the impact worsens as multiple parties see losses on entire shipments of products. Bad demand and supply planning drives loss at scale. Content errors have the power to break not just individual transactions but longer-term trust and relationships, anywhere in the value chain.

The industry is asking for more efficient, accurate ways to move product content through the value chain - and the DataPorts project has emerged to address this need.

The fundamental concept of DataPorts is “lean content”. Just as we strive for “lean” in our physical value chains, we want the same for digital value chains. While many have already been addressing digital transformation within the enterprise, we must also do the same for digital transformation between enterprises.

This paper is a summary of the work done to date with the Consumer Goods Forum to answer a single initial question:

"Is there a durable way in which we can design and build an open-source technical framework for peer-to-peer content integration for value chain partners, in full support of current and future industry and open standards?"
1. Concept

Today, when a retailer needs to assemble a complete view of information about a product for an online product catalogue, that retailer will typically go to multiple web-based sources, search for information, transform the information obtained into a format fit for purpose, and assemble the information elements for publication to the catalogue.

Today that content integration process is difficult. It is difficult because the steps of the process and the technology required for search and retrieval of information vary sufficiently from source to source, and even from item to item such that it is difficult to fully automate. The content integration problem blocks operational efficiency, introduces product information errors, and slows down the speed to market - resulting in increased costs and lost sales.

One of the approaches traditionally proposed for solving this content integration problem is “data federation”. Data federation works by standardising on a common federated data model and mapping all data sources to that standardised model. Content might be mapped either real-time in response to requests, or it might be stored using the common data model in an intermediate data store ready for consumption.

The challenge though is that agreeing on a standard model is difficult and despite big efforts and successes in standardisation across the value chain there always seem to be exceptions and in most cases there remains data which does not fit the model. We’ve been asked to find a more general way to share information such that partners can choose when and where to apply standards, adapting dynamically to changing needs.
With DataPorts we aim to simplify and automate content integration tasks between value chain partners by virtualising value chain content such that it can be more simply discovered and utilised by value chain participants. Rather than hoping for a single common federated data model to be shared by all supply chain participants, DataPorts rely on role-specific data models connected peer-to-peer allowing for the consistent sharing of data between pairs or groups of supply chain partners. We do this through an integration approach known as Data Virtualisation.

**Data Federation**
Refers to the idea of creating an all encompassing data model which covers the data transfer needs for all participants and for all purposes. This can work for some situations, but in general it is a challenge, either because the resulting model is unreasonably large and complex, or because it’s not possible to come to a stable agreement on a model.

**Data Virtualisation**
A lean approach to a data integration in which data sources are decoupled or “virtualised” so that we can treat them alike, regardless of their location and implementation. This dramatically simplifies content integration tasks - especially across peer-to-peer networks of data sources.
The full report is available to members of The Consumer Goods Forum.

To read the full report, please login to your Members’ Dashboard or contact the CGF team about becoming a member.