



Ask the Expert: Data Hub Q&A Session

Pooja Kachawaha, Community Events & User Groups Coordinator, Boomi

boomi

Upcoming Community Events

Stop Building an Integration for Every File Move

- **Date:** Oct 13, 2026
- **Time:** 10 AM AEDT
10 AM BST | 11 AM CEST

SFTP Client Connector creates scattered connections. Managed File Transfer consolidates them into one governed relationship.

Melbourne In-person User Group Meetup

- **Date:** Oct 27, 2026
- **Time:** 4-7PM AEDT

Join fellow Boomi users from across the Melbourne region for an evening of connection, conversation, and community.

Sydney In-person User Group Meetup

- **Date:** Oct 29, 2026
- **Time:** 4-7PM AEDT

Join fellow Boomi users from across the Sydney region for an evening of connection, conversation, and community.

Markus Schmidt

Boomi Professional Services (PSO) – Senior Consultant

- Germany, Munich / Frankfurt
- The last two years in a 'permanent' role
 - with a German leasing and finance company.
 - Boomi Integration and Datahub (Master-Data-Synchronization)
- With Boomi for more than six years (since April 2020),
 - working on various European projects.
- Prior to that: over 25 years as a freelance IT consultant,
 - including a "last" stint at Nestlé from 2013 to 2019.
- Background: Degree in Electrical Engineering.
- Hobbies: Software Architecture, Programming, Football, Hiking, Nature

markus.schmidt@boomi.com

*Anyone can programme (even AI can).
Developing software that will still work well in five years' time requires foresight.*

Questions



Question:

- What is Data Hub, what's it used for, and what makes it different/special compared to other tools?
- Can you share some reference/real-world use cases?

A data hub is a centralized location that collects data from many sources, processes it, and distributes it back out to systems that need it.

Traditional point-to-point architecture:

- Salesforce ↔ SAP
- Salesforce ↔ Accounting System
- Salesforce ↔ Data Warehouse
- SAP ↔ Accounting System
- SAP ↔ Data Warehouse
- (etc. – lots of direct connections)



Hub architecture:

```
Salesforce ↓  
SAP → [DataHub - Central Hub] → Accounting System  
Accounting System ↑                ↑  
Data Warehouse ←←←←←←←←←←←←←←←←
```

Instead of every system talking directly to every other system (chaotic), all systems connect through the hub. The Datahub becomes the **single source of truth that manages data quality, consistency, and synchronization.**

Benefits

Simpler and Scalable

-  Simpler — one central place to manage data instead of hundreds of point-to-point connections
-  Consistent — all systems pull from the same authoritative version
-  Scalable — adding a new system means connecting it to the hub, not to every other system

That's why it's called Boomi DataHub — it is

⇒ the hub through which all your critical master data *flows* and gets managed.

What is Data Hub

... and what it is not.

Data Hub is essentially "MDM built for integration," not isolated MDM System. If your organization needs to keep critical business data consistent across multiple systems and automate how that happens, DataHub's positioning as part of Boomi's broader platform is its unique advantage.

- Data Hub is this the central repository – the single point of truth – for the **synchronisation of company data**.
- Synchronisation processes can be built with **Boomi Integration**.
- **Data Governance** is part of Datahub as well as it can be designed with Boomi Integration.
- Datahub is **NOT a database!**
 - Neither it is it a central store where to keep (manage) your company global data (not a Master Data Repository).
- It is a HUB: A central point where multiple things connect, gather, or pass through.
- A master data repository is usually just another data sink that is fed by the Datahub.

Real-World use cases

- Whenever there is need to **synchronize data between systems**
- Without a need to change anything in a source system
 - Without even letting the source system know its data is synchronized (shared).
- Leaving all contributing systems untouched.

Use Case / Example

- A *Customer* is a business entity that is represented in various forms across different systems.
 - Sales Information, Billing, Delivery, CRM, ...
- While all systems have their unique information about a *Customer* (e.g. invoices) there is also *Customer* information that should be the same in all systems:
 - Name, Address, legal numbers (VAT ID) etc.
- And if you want to ensure some level of data quality
 - e.g. VAT is mandatory for all systems
- That's where Datahub comes into play!

Question: Are there any initiatives to build APIs that fully support Quarantine Actions — specifically Submit, Edit, and Resubmit — for use within a Boomi Flow application?

Working with Quarantine

Apart from Flow..

Boomi Integration Connectors in general offer an easy to use and very convenient access to source systems. I recommend checking their functionality (and probably limitation) over their “native” REST / SOAP counterparts which are normally available, too. So, if Flow Datahub Connector does not satisfy our needs, check Datahub REST API:

- <https://developer.boomi.com/docs/datahub-api-reference>

Basically it is not a good idea to **edit quarantine entries**. The root cause why a record has entered quarantine is that the incoming source-entity did not meet the Datahub (business) requirements. The solution would be to fix the *source* record! Otherwise, you end up creating a ‘Golden Record’ that differs from the original data in the source system.

Using the Quarantine UI is the “last exit” to get it done – with all its consequences. I would recommend to not using Quarantine for daily operations.

Solution: Query Quarantine and inform the stakeholders (users) of the source system, that their system provided data that did not meet the Golden Record (company global) requirements, and let them fix and re-submit it. Ref. Data Ownership!

Question: What does real-time delivery look like with Data Hub? Are there specific use cases for it?



Real-Time Sync

Early-Access - [notes July 2026](#)

- Use the Data Hub Listener Shape to listen for Golden Record changes
 - And execute the required functionality right when the Golden Record is updated.
- Instead of polling of updates every n minutes/hours (fetching batches).
- Real-Time sync is based on Queues Event Streams (Queues) to trigger changes.
 - Licensing = ?
- I haven't seen functionality for quarantined real-time information.
- Datahub triggers the outbound delivery right when the Golden Record was updated.
- When I think of 'sync', I imagine something a bit different: end-to-end
 - However, this is out of Datahub's scope.
 - Sync must also take the source systems into account, their interfaces, as well, as the delivery to the target systems and everything that might happen in between.

Real-Time Sync

Step back ...

.. and think about the requirements and how you are going to solve them?

I do not remember any use-case where real-time synchronization was required. In addition, real-time sync means, the Datahub puts changes in the outgoing queue once they appear.

Real-Time is a solution for what? Where does the request come from?

- **What is your use-case?**

Think about *synchronization* – and data governance – and real-time: what if the record is incomplete and the record cannot be synchronized (not in real-time, not at all)? How are you going to handle real-time Quarantines?

If you want updates in real-time send it!

- Use APIs (with Boomi Integration) and build a [better] architecture.

Real World Example

Real-time “synchronization”

- *Customer* information entered in system A,
- to be enriched and approved in system B,
- to continue editing in system A.

Notes: Customer information entered in system A was not a “complete” customer record from the data governance perspective. It was more a partial customer or a “proposed customer” that needed approval. Actually this is not the use-case of synchronizing a customer!

It was a technical request for real-time to support the business process of customer creation.

- Do you really have the requirement to synchronize customer information in real-time over all systems?

The better way..

- Publish an API for /customer /approve
- Instead of real-time sync, let System A call that endpoint
 - To let System B approve and complete the “customer create request”.
- Once there is a complete, approved customer,
 - let System A create a Golden Record.

Question: Any plans for:

1. Allowing more complex queries?
2. Changes to record size limits/including BLOB like fields?
3. Improving speed of UI?

The better way..

- <https://boomi.com/blog/boomi-product-roadmap-q3-2026-everything-you-want-to-know/>

Golden Record Query and Profiling APIs: The Query End-Dated Golden Records endpoint now supports filtering and optional metadata fields, bringing it to parity with Query Golden Records. Golden record profiling statistics are also now retrievable via API, with up to 60 days of historical data for tracking quality trends over time.

- **Real-Time Delivery Mode (Early Access):** This capability continues to mature in early access – evolving as both an event-driven

I need your input: what are the real requirements and where do these questions come from?

Now .. or contact me directly: markus.schmidt@boomi.com



Thank you