

The following question and answer session with Forrester Principal Analyst Henry Peyret addresses the challenges organizations face as they explore the value in data catalogs.

### Q. “How would I get started with data catalogs?”

**Peyret:** The key thing to remember is that data catalogs are not a panacea for all data problems. I’m hearing customers express the perception that data catalogs are some sort of silver bullet for bringing lagging data practices into the future. I’m excited about data catalogs, but they should be seen as a part of a broader insights platform. That means having a strategy for both the catalog and the insights platform. The most important thing you can do to get started with all this is to engage with and understand the perspectives of the different roles, personas, or stakeholders that will have different – and potentially conflicting – data needs. The data manager, the data steward, and the data scientist all have different needs and your first step is to understand what they are. Then you can build a prioritized strategy that targets specific needs on a detailed timeline, complete with clearly stated objectives and KPIs.

### Q. “What are some examples of companies who are doing this successfully already?”

**Peyret:** The evolving vision for the insights-driven business is so new that it’s difficult to find great examples of firms that have nailed it. One of the best examples I can point to is a large financial services firm that I’m working with that clearly understands their past failures to advance their data capabilities and is working to build to the new vision. Cameron’s example of a global multi-billion dollar US-based firm that carved 2-3 specific objectives and 10 or 11 critical KPIs out of their enterprise-wide strategy to get some quick wins over a few months is very good. Nobody is an exemplar of the perfect future state at this point, but that company is looking at both the long term strategy and generating near-term value in a way that will get them to where they need to go.

### Q. “What is the difference between a data catalog and a business glossary?”

**Peyret:** A business glossary provides semantic information on the data elements available in an enterprise; these are often manually built. Catalogs contain semantics and more, including data sources, lineage, and even data usage information. Modern data catalogs are automated, built by software using machine learning to profile data and process usage logs; typically they can be annotated with manual updates as well. They should be seen as complementary to each other as the business glossary is describing more at conceptual level and the data catalog is more at physical level as the metadata is coming directly from data structures’ storage.

### Q. “What is changing in data architectures to make catalogs a necessary component?”

**Peyret:** Enterprise data architectures have evolved significantly over the last several years and are continuing to evolve. The placement of data in the cloud – including multi-cloud implementations – is key, as is the explosion of internal and external data sourcing. Usage scenarios for data are driving changes also: the need for self-service by a wide variety of user types – all with carefully controlled access – is a big change, and there are new use cases for the data including customer sentiment analysis, predictive analyses, and AI, as well as the need to drive new insights and value from data. These new data usages (rather than only transactional and reporting) are introducing new risks (bias is one of the examples we start to hear), new agile data management technologies (NoSQL, Hadoop, Time Sensitive Databases, Data Virtualization and more to come) and practices (DataOps) and adapted data governance. Catalogs can be a significant enabler in conjunction with data stewardship and data management technologies to the different roles that need to work with the data in these different scenarios.



### Henry Peyret

#### PRINCIPAL ANALYST SERVING ENTERPRISE ARCHITECTURE PROFESSIONALS

Henry serves enterprise architecture (EA) professionals by helping them understand the concepts, techniques, and tools they need to support the digital business. His current research focuses on integration — both internal to a business and business to business — and on data governance, stewardship, and discovery. Henry also has a separate research stream on how customers’ values affect companies and government.

Henry has more than 30 years of IT experience in a variety of consultative and leadership roles, and is the author or coauthor of many reference articles and books on IT architecture. His expertise spans the fields of master data management, enterprise application integration, middleware, adapters, business activity monitoring, enterprise architecture tools, and open source and risks issues.

Optimize your data management and governance strategy with Infogix. Visit [www.infogix.com](http://www.infogix.com) or call +1.630.505.1800 (US, Canada, and International), +44 1242 674 137 (UK and Europe).

