The Rise of Data Exchanges

Frictionless Integration of Third-Party Data

BY OLIVER VAGNER AND WAYNE ECKERSON

About the Author

Oliver Vagner is a strategy consultant, solution architect, trusted advisor, and industry analyst dedicated to helping enterprises get value from data and technology. He is a servant leader who believes the role of technology is to serve the business in delivering value to the enterprise. Oliver has built a comprehensive background in data analytics, with more than 25 years of consulting experience spanning multiple technology disciplines and industries. Oliver holds both a B.S. and an M.S. in industrial and systems engineering from the University of Florida.

Wayne W. Eckerson is an international thought leader in the data and analytics field since the early 1990s. He is a noted speaker, sought-after consultant, and widely read author. Eckerson has conducted many groundbreaking research studies, chaired numerous conferences, written two widely read books, and consulted with numerous organizations around the world. Eckerson is the president of Eckerson Group, a research and consulting firm that specializes in data analytics. He has degrees from Williams College and Wesleyan University.

About Eckerson Group

Eckerson Group helps organizations get more value from their data through research, consulting, and education. Our experts each have more than 25+ years of experience in the field, specializing in business intelligence, data architecture, data governance, analytics, and data management. We provide organizations with expert guidance during every step of their data and analytics journey. Get more value from your data. Put an expert on your side.

Learn what Eckerson Group can do for you!

Eckerson Group
RESEARCH · CONSULTING · ADVISORY
G E T · M O R E · V A L U E · F R O M · Y O U R · D A T A

About This Report

The report is sponsored by Harbr who has exclusive permission to syndicate its content.

Table of Contents

Executive Summary

Data exchanges are emerging as a key component of the data economy. By connecting data suppliers and consumers through a seamless experience, data exchanges have created a modern-day goldrush to help modern, digital organizations address their insatiable hunger for data. And, by eliminating the friction from finding, acquiring, and integrating data, these exchanges make it easy for any company to monetize its data assets and create new revenue streams.

To succeed, data exchanges need a robust data sharing platform that delivers transparency, trust, quality, security, and integration. The best data exchange platforms mimic the value proposition of leading consumer marketplaces, such as Amazon.com, Spotify, and Apple iTunes, by addressing the four Cs: convenience, cost, completeness, and community. These modern data exchanges reduce the traditional barriers that have made it nearly impossible for organizations to find, acquire, and integrate third-party data.

A host of companies are launching data marketplaces to facilitate data sharing among data suppliers and consumers. Some are global in nature, hosting a diverse range of data sets, suppliers, and consumers. Others focus on a single industry, functional area (e.g., sales and marketing), or type of data. Still others sell data exchange platforms to people or companies who want to run their own data marketplace. Cloud data platform providers have the upper hand since they've already captured the lion's share of data consumers who might be interested in sharing data.

Challenges in the Data Economy

Scenario #1 – Data science project

A data scientist in Boston needs additional features to enhance her customer churn model. She browses the Internet for data sets and, after several days, finds a data broker with suitable data. After several weeks working with procurement to purchase the data set, she downloads the data from a flaky FTP site, succeeding on the fifth attempt. She then waits several weeks for IT to map and load the data. Once she finally explores the data, she discovers it is littered with duplicates, incomprehensible labels, invalid numbers, and nulls. She can only use several thousand of the 200 million records. Frustrated, she starts the arduous process again.

Scenario #2 – Data monetization initiative

Meanwhile in Atlanta, a product manager at a point-of-sale service provider looks for opportunities to monetize restaurant transactions that his company has collected as part of its service business. He creates a web site to market and sell the data to interested parties, such as restaurant suppliers, government agencies, and market researchers. He soon realizes that he lacks the budget and time to ingest the data efficiently and integrate it with other sources to create marketable data products. Disappointed, the product manager shelves the effort for another year, forfeiting a revenue stream.

Data Sharing Issues

Sharing data between companies has never been easy—for either data suppliers or data consumers.

Pre-Internet. In the pre-Internet era, large data syndicators, such as IRI, Nielsen, and IMS Health, sent CDs of data to customers every month. As technology advanced, data transfers went digital, evolving from file transfer protocol (FTP) servers and email to cloud buckets (e.g., DropBox) and application programming interfaces (APIs). But these transfers are not always easy or quick to set up, and they take resources to troubleshoot, secure, and maintain.

Data Acquisition. Likewise, most individuals and companies struggle to find and acquire third-party data. They have to find a suitable source, which means browsing through dozens or even hundreds, of web sites run by list brokers and data syndicators. They need to evaluate the data, which may be impossible without a sample data set. Then they need to negotiate a price, especially if they want a custom cut of the data. Then they need to work with their purchasing department to set up a one-time purchase or monthly subscription.

Data integration. Finally, integrating third-party data with internal systems is fraught with challenges. List brokers and some data syndicators are notorious for delivering poor quality data, full of nulls, duplicates, errors, and schema inconsistencies. And data sets are not always at the right level of granularity. It must be cleaned and mapped before it can be queried. Data architects need to model and join the data with internal data, if common keys exist, and then design a method to update the tables when new data arrives. They also need to evaluate whether the external data exposes sensitive information or violates data regulations and compliance policies.

The whole process is time consuming, and for many data consumers, not worth the effort. It prevents data suppliers from achieving the economies of scale needed to run a sustainable business. It makes it harder for value-added providers, who can enrich, map, and validate third-party data to get traction. Finally, it discourages potential suppliers, especially companies who generate data as a by-product of doing business, from selling their data to companies that could truly benefit from it, thereby preventing those suppliers from generating new revenue streams for themselves

Rise of the Data Exchange

The Promise. Modern, cloud-based data exchanges deliver a better data sharing model. They create a digital marketplace of data assets where data suppliers and consumers can exchange data in a frictionless and seamless way. They make it easy for data suppliers to package, publish, sell, and distribute data assets and for data consumers to discover relevant data sets and, in some cases, integrate them into their internal systems.

Why Now? Data exchanges address a growing thirst for third-party data, ranging from weather, geographic, and streaming sensor data to traffic and economic data. Data scientists want external data to enrich their statistical and machine learning models; data analysts need third-party data to enhance analyses, plans, and forecasts; product managers want demographic, psychographic, and social media data to create 360-degree views of customers; many companies want clickstream data to personalize user experiences and create effective recommendation engines, among other things.

A key driver of data exchanges are popular consumer marketplaces, such as Amazon and Apple iTunes, which have revolutionized how we shop. Amazon's consumer marketplace started with books and now sells every imaginable product. Apple iTunes started with music and now offers any digital media product, including videos, movies, and podcasts. Both have disrupted multiple industries by digitizing interactions between suppliers and consumers.

With a relentless focus on convenience, cost, completeness, and community, these early consumer marketplaces provide data entrepreneurs a ready roadmap for how to create a modern data marketplace. In the last five years, dozens of data companies have launched data marketplaces serving every industry, functional area, and role with a vast diversity of data sets in every conceivable format. Meanwhile, traditional data syndicators are creating marketplaces of their own or leveraging emerging data exchanges as new distribution channels.

The Data Goldrush. Most significantly, data exchanges herald a new era in data management. In the new data economy, organizations are not just consumers of data; they are suppliers, too. Automobile makers, for instance, now recognize that the data about their cars and how customers drive them is nearly as valuable as the vehicles themselves. As these organizations digitize their products, they produce a goldmine of information. A data exchange gives them an opportunity to create a new and potentially lucrative revenue stream by monetizing that data.

What Is a Data Exchange?

A data exchange facilitates the transfer of data between a data supplier and data consumer. Traditionally, a *data supplier* is a data broker, data syndicator, or open data provider. But increasingly, it is any company that wants to share or sell its own data. A data consumer is any individual or company that wants to collect or purchase data. Note that we don't call them "buyers" and "sellers" because many share data freely or execute payments (i.e., subscriptions) outside of a data exchange.

Other key players are data exchange providers and operators. A *data exchange provider* is a company that sells data exchange technology, which we call a data exchange platform (DXP). *A data exchange operator* is a company that uses a DXP to create and manage a data sharing site, such as a private data exchange or a public data marketplace. In most cases, a data exchange provider is also a data exchange operator. For instance, **Dawex** and **Narrative** are primarily data exchange providers, but they also run public data marketplaces. In contrast, **Harbr** only provides data exchange technology for exchange operators.

Types of Data Exchanges

To facilitate the transfer of data, data exchanges primarily use FTP servers, APIs, and various cloud sharing mechanisms. But what's more interesting than the physical data transfer is the relationship between data suppliers and consumers and how they engage with each other to share data. There are three types of data exchanges that define this relationship:

1. **Peer-to-Peer Data Exchange.** A peer-to-peer exchange facilitates direct data sharing between two companies that want to share data on an ad hoc basis. But it could also be used to facilitate data sharing between two or more divisions in the same company. For instance, a large hospital with two data warehouses, one for business operations and one for research, might use a peer-to-peer exchange to share subsets of data.

2. **Private Data Exchange.** A private data exchange takes many forms. It could be an industry consortium, such as Verisk, which collects and standardizes insurance data and distributes it to participating members and regulators. Also, a large company such as Walmart might use a private data exchange to share inventory data with its suppliers and collect shipment data in return. Or a service provider, such as an ad tech firm, might use a private data exchange to distribute account activity to individual customers. Some providers, such as **Harbr**, provide platforms explicitly to support private data exchanges.

3. **Data Marketplace.** A data marketplace is a public data exchange open to any company that wants to supply or consume data. Some marketplaces are global in scope, spanning every industry, function, and type of data. For example, AWS, **Snowflake**, **Dawex**, and **Narrative** have global data marketplaces. But there are also niche data marketplaces that focus on a single industry. For instance, **Crux Informatics** is a data exchange for financial firms that provides a network and services to connect suppliers and consumers. Another example is **Terbine**, a marketplace that sells only machine-generated data.

Ideally, a data exchange supports all three modes, but only Dawex and Snowflake do so today. Amazon offers a data marketplace but not private exchanges. Microsoft supports private data sharing but not a global data marketplace.

Figure 1. Data Exchange Framework
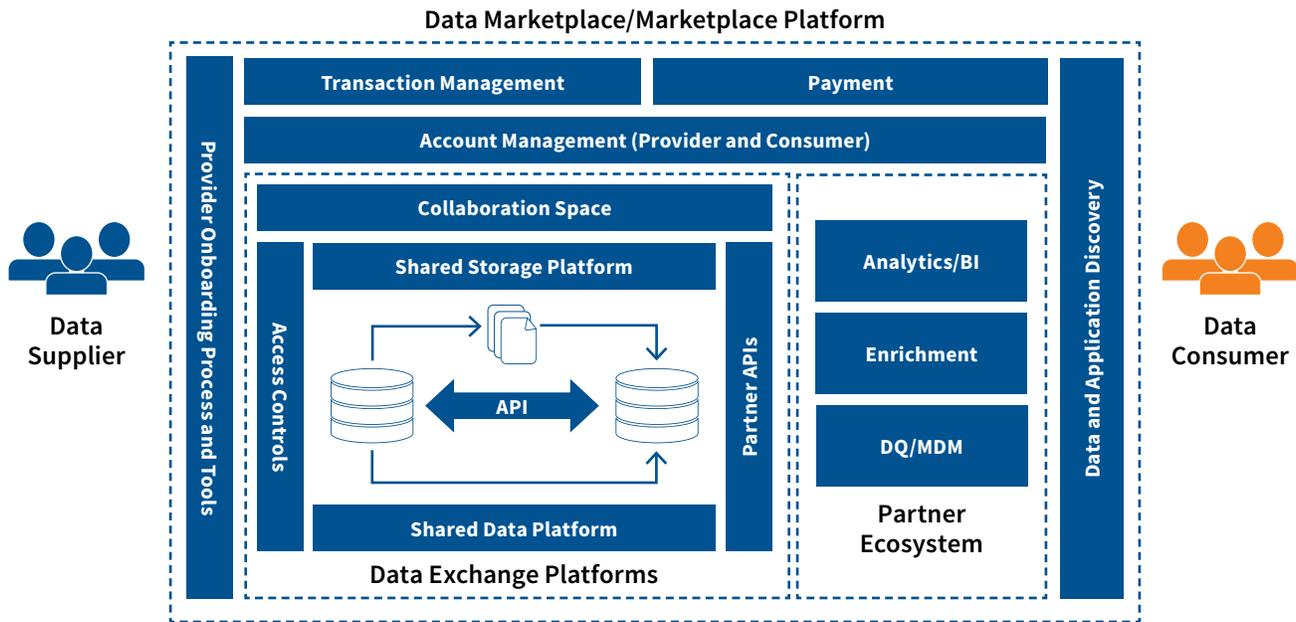
Baseline Functionality

A full-featured data exchange enables data suppliers and consumers to share data in a seamless, secure manner. To do this, it must offer baseline functionality that makes it easy for suppliers to publish data assets and for data consumers to shop for them. This baseline functionality is depicted in figure 1 and described in this section.

> **Searchable Catalog.** Data consumers can search a catalog of data assets. Each entry in the catalog contains a description of the data set along with appropriate details, such as number of records, file type (i.e., relational, JSON, CSV, image), profile statistics, pricing (if applicable) and ratings. Some catalogs let consumers peruse sample data.

> **Asset Management.** A data exchange makes it easy for suppliers to upload, describe, manage, publish, and update data assets. It also enable suppliers to define licensing, access rights, and other terms and manage their inventory. In a private data exchange, this foster distributed data ownership and closes the gap between data domain experts and business stakeholders.

> **Access Control.** An exchange gives data suppliers the ability to control who can access or purchase a data asset, in part or whole, and restrict access until an agreement is in place or payment is complete. Mechanisms include simple file permissions or the exchange of encryption keys.

> **Data Transfer.** A data exchange supports one or more ways for data suppliers to transfer data to consumers, including file transfer, APIs, and multi-tenant data sharing. Cloud-based file transfer uses object storage to hold files, simplifying access. With multi-tenant data sharing, suppliers and consumers run on the same data management system so they can share data transparently via permissions. Some vendors, such as Harbr, create data pipelines that automatically update data at the user's preferred endpoint.

> **Subscription Management.** Data suppliers can offer subscriptions to their data assets and set the price and duration, if desired. For example, on the AWS Data Exchange, suppliers offer subscriptions ranging from 1 to 36 months, and charges are billed on the consumer's AWS account. Many exchanges also let consumers bring their own subscriptions (BYOS) to the exchange via tokens or other mechanisms. Some data marketplaces that offer advanced services, such as Crux Informatics and Narrative, require consumers and suppliers to purchase a subscription to access the platform. Dawex and Harbr allow their operators to offer a host of additional subscription options including free, tiered, or transactional.

> **Transaction Management.** Data exchanges facilitate monetary transactions and offer a host of payment terms, including bank transfers, credit cards, and account billing. Data consumers can track their history of data purchases and subscriptions, see renewal terms, and cancel subscriptions. Data suppliers see a history of consumer transactions and activity by data set.

> **Account Management.** A data exchange lets data consumers and suppliers create and manage accounts that contain information about users, administrators, authorized buyers and sellers, billing information, account activity, and payment mechanisms. Large organizations with multiple parties interested in acquiring external data can easily manage permissions and expenditures.

> **Administer.** A data exchange provides a means for data exchange operators to manage the platform, monitor activity, and troubleshoot issues. For example, a company wanting to set up a data exchange for its partners can quickly implement and support the exchange without much more expertise than they would need to support normal data operations.

> **Collaboration.** Finally, a data exchange enables suppliers and consumers or multiple suppliers to collaborate together to create new or targeted data sets. For example, Harbr establishes a secure and trusted area for parties to come together to evaluate, compare, combine, and processes multiple data sets into new data sets. For the consumer, they can evaluate supplier data sets ahead of commiting to a subscription and only select the relavent data for their needs. Suppliers can work together to create a new data set more valuable than the sum of its parts.

Value-Added Features

The capabilities listed here are standard features of a modern data exchange platform. But competition is heating up and platform providers are racing to develop advanced features. The goal is to make it as easy as possible for data suppliers to package, publish, sell, secure, and distribute data assets and for data consumers to browse, purchase, load, and query those data sets. A frictionless data exchange platform is easier to sell and use, making success more likely for everyone involved: data operators, suppliers, and consumers.

The following emerging features and functions will soon be required to do business for most data exchange platform providers.

> **Data Enrichment Services.** A data exchange can offer a variety of data quality services, including deduplication, address correction, standardization, data cleansing, file merging, and validation.

> **Selective Sharing.** Some data exchanges let data suppliers selectively share data, either by filtering data for select consumers or by selecting consumers with whom to share data. Selective sharing works well on a multi-tenant cloud platform, such as Snowflake, where data suppliers configure permissions for select consumers to

access data subsets. Dawex and Harbr enable selective sharing in the data discovery process to limit visibility of available data sets to only targeted customer types.

> **Data Mapping.** Data exchanges, such as those from Narrative and Ringlead, analyze consumer data, if allowed, and recommend supplier data within the exchange that best supplements or enriches that data. Dawex, in order to meet the highest degree of data security compliance, uses metadata mapping to accomplish the same goal.

> **Multi-Tenant Data Sharing.** Some cloud data platforms, such as Snowflake, replace FTP and API-based sharing mechanisms with cloud-based, multi-tenant solutions that eliminate the need to copy or move data, erasing the infrastructure headaches that accompany traditional forms of data sharing.

> **Connector SDK.** Data exchanges offer a software development kit that makes it easy for data exchange operators or suppliers to create a custom connector to less common data platforms.

> **Derived Aggregate Data.** Rather than give consumers access to raw data that may contain sensitive data, data suppliers can instead offer user-defined functions (UDFs), such as market basket analysis. Consumers run the UDFs against a described data set and receive the aggregated output.

> **Enhanced Onboarding.** Many data exchanges now offer workflows and wizards to simplify the onboarding process for data suppliers. The workflows evaluate supplier data, assess compliance, and manage data ingestion, among other things.

> **Alerts.** Some data exchanges alert data consumers when a newly published data set matches their interests. Some also alert data suppliers when potential customers join the exchange.

> **Pipeline Management.** Some exchanges, such as Crux Informatics, build and maintain data pipelines that deliver third-party data to predefined endpoints, sometimes combining and integrating data along the way to the consumer.

> **Enhanced Reporting.** Data suppliers want a clear picture of their sales performance in the exchange. Data exchanges, such as Narrative, provide detailed reporting on supplier performance, including elevating top "buyers" to allow the suppliers to focus their efforts on the right consumers.

> **Custom Data Products.** Some data exchanges enable providers to create a data product by blending, segmenting, or engineering data to suit use cases and then publish that product to the exchange.

> **Prohibit Change of Custody.** Some exchanges, such as Harbr, enable administrators or product owners to restrict data access to the exchange platform, enabling users to preview and trial third-party data without violating license terms or collaborate on sensitive data without change of custody.

Today, data exchanges charge fees for the features and functions described here, or roll them into a platform subscription. In the future, these features will become a requirement for doing business.

Rise of the Data Marketplace

In the past five years, dozens of data companies have launched data marketplaces serving every industry, functional area, and role with a diversity of data sets in every conceivable format. Like homesteaders in the American west during the 19th century, many data marketplace pioneers are startups staking out virgin territory.

Data Exchange Specialists

Crux Informatics, for example, created a data exchage in 2017 geared to financial firms. It currently has 100 data suppliers delivering almost a half billion data files covering commodities, derivatives, equities, fixed income, foreign exchange, indices, and real estate. Other data marketplaces specialize in a specific type of data. **Terbine**, for example, offers machine-generated data collected from government agencies and universities from around the world. The marketplace currently offers 30,000 data sets generated by sensors in mobile phones, automobiles, farm equipment, and various monitoring devices.

Some startups offer not only a data marketplace, but also the underlying platform required to create and run one. **Dawex**, for instance, runs the **Dawex Global Marketplace**, which currently has 12,000 registered companies across all industries and regions of the world. But its core business is selling the **Dawex Data Exchange Platform**. One customer is Api-Agro, an agricultural consortium in Europe, which uses the platform to run a commercial data marketplace for farm sector companies that want to buy and sell agriculture-related data.

Similarly, Narrative offers the **Narrative Data Streaming Platform** as well as a global marketplace with 40 data suppliers. Narrative offers a simple software-as-a-service interface for both data suppliers and consumers, who must purchase a subscription to access the platform. They can access a range of value-added services depending on their subscription plan, such as automatic deduplication of all purchased data, automatic mapping of customer data to relevant supplier data on the platform, and the ability to create custom terms and agreements.

Lastly, **Harbr** offers their data exchange software-as-a-service to data providers launching their own branded data exchange or enterprises looking to deploy an internal exchange or private data exchange with partners to optimize third-party data access. Unique to the Harbr platform is the ability of suppliers and consumers to collaborate securely and combine data assets to create new data products.

Cloud Platform Vendors

The biggest potential data marketplace players are cloud platform vendors. They've already captured the eyeballs that data suppliers covet. They also have mature data platforms and account management systems to facilitate data sharing and transactions between suppliers and consumers.

Amazon Web Services. AWS made a huge splash in November 2019, when it launched **AWS Data Exchange**, an offshoot of its successful AWS Marketplace, which enables software vendors and systems integrators to sell software and services online. The AWS Data Exchange currently lists more than 2,500 data products from more than 160 qualified data suppliers, including Reuters, FourSquare, and Dun & Bradstreet. Unlike most other marketplaces, AWS Data Exchange provides an e-commerce function to conduct financial transactions rather than force buyers and sellers to handle payments offline.

Snowflake. A popular cloud data platform that runs on all three public clouds (AWS, Microsoft Azure, and Google Cloud Platform), Snowflake launched the **Snowflake Marketplace** in June 2019, five months before AWS. The

Snowflake Marketplace has 156 data sets from nearly 100 data suppliers, but Snowflake's real advantage lies in the novel way it supports data transfers. Because Snowflake is a cloud-native, multi-tenant data environment, every customer can seamlessly share data with any other Snowflake customer without moving it, no matter what cloud platform or data processing region they run in. There are no file transfers, no data copies, no latency, just a frictionless data sharing experience.

Microsoft. Not to be outdone, Microsoft in July 2019 launched **Azure Data Share**, which makes it easier for Azure database customers to share data directly with each other. Currently in early preview, the service lets customers of its data lake, data warehousing, and data exploration databases send snapshots of data (i.e., files) to each other. More advanced forms of sharing are planned. Interestingly, Microsoft launched a data marketplace several years ago but abandoned it due to lack of interest.

Google. Last, Google has a very modest data exchange that hosts several public data sets and a number of commercial data sets. In total, the Google exchange has a little more than 100 data sets available on Google Cloud Storage, which is tightly coupled to Google Cloud's data platforms, such as Google Big Query.

The Future. It's too early to tell which data marketplace will become as dominant as Amazon.com in the consumer world. More likely, we'll see specialized data marketplaces serving each industry and functional area (marketing, sales, finance) along with one or two global marketplaces serving up any and all data. Cloud platform vendors, such as AWS, Snowflake, and Microsoft have a built-in advantage here, but numerous startups are working hard to carve out niches in various verticals or as data exchange platform providers.

Challenges and Issues

Data exchange platforms and marketplaces are a relatively new addition to the data services universe. Many key issues therefore need to be addressed before data exchanges become a fixture of modern data architectures and a driver of the new data economy. The top issues follow:

Trust

Data consumers can cite a litany of bad experiences with third-party data. Modern data exchanges need to overcome this unfortunate legacy by providing data integrity, quality, and consistency. Unless data consumers trust the data housed on a data exchange, they won't use it. There are numerous things a data exchange operator can do to ensure consumers trust the data:

> **Curation.** While challenging to scale, data exchange operators can certify and curate supplier data sets to ensure quality and consistency. For example, Ringlead evaluates all the data suppliers on its exchange, collects metadata for its exchange catalog, and then maps the data into a canonical form to display to data consumers.

> **Profiling.** A data exchange should contain rich metadata about each data asset, including number of rows, names and types of fields, cardinality, and other statistics as well as the lineage of the data set, including source systems, collection methods, and age.

> **Sampling.** Ideally, data exchanges allow consumers to query or sample data sets so they can validate the quality of data.

> **Ratings.** Much in the way e-commerce platforms provide ratings of products, data exchanges can allow consumers to give feedback on the quality and utility of data sets.

Critical Mass

Data exchanges need to reach a critical mass of both data suppliers and consumers to gain traction and become viable, long-term operations. But this is a chicken-and-egg situation: Data consumers won't use a site that lacks sufficient breadth and depth of data assets, while data suppliers won't populate a site that lacks potential buyers.

Unless a data exchange is run by a major cloud platform, it needs to target a specific market with a built-in and demonstrated need for third-party data. Government agencies, industry consortia, and membership organizations have a distinct advantage because they have a built-in audience and already serve as liaisons and intermediaries.

For example, the Republic of Korea is building 10 exchanges across several verticals to support its I-Korea 4.0 strategy to digitize and connect companies so they can fully participate and compete in the fourth industrial revolution. Through subsidies, the government aims to have 30,000 small and midsize organizations trade data on these exchanges by 2023. A market-mover like this is often needed to help kickstart a data exchange.

Security and Compliance

Robust security is table stakes to any platform or application dealing with data. Data exchanges need to secure data transfers and transactions, guard the entry and exit points of data, and apply appropriate access controls so only contracted data consumers can access a supplier's data. This also contributes to building trust in marketplace data.

Compliance is a significant hurdle. Data exchange operators need to ensure that data offered on the exchange complies with all country- and industry-specific regulations. For example, the General Data Protection Regulation requires that a data exchange ensures that data generated inside a country stays in that country and also prohibits or carefully monitors the exchange of personally-identifiable information, among other things.

Data exchanges also must pay attention to corporate end user license agreements. Companies eager to monetize their data may apply terms and conditions that limit the data's availability to certain consumers or restrict the frequency of use or other conditions. There is precedence for companies to claim rights to insights generated from their data. It is advisable to seek legal counsel, especially when exchange data is destined for critical business processes.

Data Integration

Although the primary function of a data exchange is to facilitate the transfer of data between suppliers and consumers, that doesn't add business value for data consumers. They still need to integrate third-party data with their existing data management systems so business users can access and query the data along with all other internal data.

To facilitate integration, data exchanges need to validate supplier data. They need to ensure the data is clean, consistent, and delivered "as advertised." That means the schema doesn't alter mid-stream, or a column field doesn't change its name or type, and the number of records remains consistent. They also need to ensure that the data supplier has documented all the requisite metadata so data consumers can make informed purchasing decisions.

Interestingly, some data exchanges are taking integration challenges to their logical conclusion. For example, Crux Informatics builds data pipelines that distribute data to various endpoints specified by a data consumer on their exchange. Crux will clean, combine, map, and transform the data so customers simply need to validate the output and use it. For companies that struggle to integrate third-party data and broaden access to it, this additional service may be well worth the cost.

E-commerce

Some data exchanges offer a fully integrated e-commerce platform, providing all aspects of account and transaction management with a seamless payment system. Others rely on transactions and payments that happen outside the platform but are verified within it.

Soon, data exchanges will need to support a full range of e-commerce functionality and payment options. This includes subscriptions managed by the data exchange platform and subscriptions executed outside the platform (i.e., bring your own subscription.) It also includes transactions for an individual data file as well as variations on that file, such as some or all of its history or subsets of the file. Ultimately, all data exchanges will need to align with trends in cloud pricing and support "pay-by-query" where users are only charged for the data they actually use.

Summary. The data exchange market is rapidly evolving. New data exchanges already address many of the challenges described here, and we expect others to follow suit soon.

What's Next?

Data exchanges are the harbinger of a global data economy where organizations of all sizes and shapes freely exchange information to improve performance and address problems. Any paradigm shift in the way goods are packaged, sold, or consumed creates market disruption. New winners and losers will emerge in the global data economy. Companies need to understand the new dynamics and prepare accordingly.

The following are a few of the major implications of this new data-sharing technology:

> **The New Data Warehouse.** Some organizations may use a data exchange to facilitate selective data sharing among business units rather than build an enterprise data warehouse or merge existing data warehouses and data marts. This will likely happen in large universities, healthcare systems, and government agencies with very distinct business units or operating groups.

> **Data Marketplace Explosion and Collapse.** Expect an explosion of data marketplaces in the next few years followed by a collapse. Although data exchange platforms make it easy to start a data marketplace, many will fail to gain critical mass. The ones that succeed will either have huge scale (see next) or target a specialized niche facilitated by a well-connected intermediary.

> **Cloud Platforms Prevail.** The providers to beat in the data exchange market are large cloud platform players who will soon amass a majority of potential data consumers. In some cases, they also have a head-start on e-commerce features (i.e., AWS) and multi-tenant data sharing (i.e., Snowflake).

> **Monetary Transactions Grow.** Today, monetary data transactions largely occur outside data exchanges, as subscriptions signed on the supplier's site. In the near future, data exchanges will bring those subscriptions in house and begin offering more selective payment options, including pay-per-query.

> **Private Data Exchanges Replace Existing B2B Methods.** While public data marketplaces may shrink, private ones will show steady growth. Large companies will replace existing B2B exchange mechanisms, including FTP sites, BI extranets, and even EDI networks, with off-the-shelf data exchange platforms.

> **Data Syndicators Adapt.** Large data syndicators will keep their current exchange mechanisms in the short run while they test the viability of public data marketplaces as new channels for their data products.

> **New Data Suppliers Struggle.** Monetizing data is not easy, even with a modern data exchange to grease the wheels. Companies that want to monetize their data exhaust on a public exchange need to understand the market dynamics before wading in. Just providing raw data may not be worth the effort. They need to think hard about their target customers and the types of value-added data and services they want, to make a monetization strategy worth the effort.

> **Data Marketplaces Interoperate.** Customers will want to align with a primary data marketplace and a handful of private data exchanges. They will naturally force data exchanges to interoperate with each other so users can shop across exchanges while maintaining their accounts and transactions.

> **Private-Label Marketplaces.** Software vendors will piggy back on global marketplaces, providing tailored, private-label access, giving their customer one-click access to external data from within their applications.

> **Regulation.** Governments will move to regulate data exchanges to protect consumer data from proliferation and misuse. Data exchanges will be required to implement strict access controls and police the use and distribution of PII.

About Eckerson Group

Wayne Eckerson, a globally-known author, speaker, and advisor, formed **Eckerson Group** to help organizations get more value from data and analytics. His goal is to provide organizations with expert guidance during every step of their data journey.

Today, Eckerson Group helps organizations in three ways:

> **Our thought leaders** publish practical, compelling content that keeps you abreast of the latest trends, techniques, and tools in the data analytics field. We share best practices that align your team around industry frameworks.

> **Our consultants** listen carefully to craft tailored solutions that translate your business requirements into compelling strategies and solutions.

> **Our educators** share best practices in consulting workshops or external conferences on 30+ topics.

Our experts each have more than 25+ years of experience in the field. They specialize in data analytics—from data architecture and data governance to business intelligence and artificial intelligence. Their primary mission is to help you get more value from data and analytics by using their extensive experience.

Our clients say we are hard-working, insightful, and humble. It all stems from our love of data and our desire to help you get more value from analytics—we see ourselves as a family of continuous learners, interpreting the world of data and analytics for our clients and partners.

Get more value from your data. Put an expert on your side. **Learn what Eckerson Group can do for you!**

Eckerson Group

RESEARCH   CONSULTING   ADVISORY

G E T · M O R E · V A L U E · F R O M · Y O U R · D A T A

About Harbr

Organizations understand the potential value of their data, but face significant barriers to realizing it. Harbr removes those barriers by delivering enterprise and commercial data exchanges that help our customers generate new insights, improve decision making and monetize their data.

The Harbr data exchange platform empowers you to unlock the full value of data by bringing people, data and tools together—securely expanding access, while maintaining control. By converting data into data products, Harbr makes it easy to find, use and manage through a one-stop digital storefront. It empowers domain experts to collaborate on data products with internal and external business stakeholders. Those collaborators can share the output of their work as new custom data products on the data exchange, advancing a cycle of value creation. As a secure SaaS platform, we take care of the technology, so you can take care of accelerating your data-driven business.

Learn more at **www.harbrdata.com**.