

BARC Spotlight

A Data Marketplace Is What Your Agents Need

Author: Florian Bigelmaier

Publication: June 2026

Provided by



Entropy Data

Trust, Access, and Accountability: How Data Marketplaces Enable AI Agents

Making data trustworthy and discoverable has been a persistent, unsolved challenge in enterprise data analytics over years. According to the BARC study *Data Mesh and Data Fabric – From Theory to Application*¹, 73% of respondents report that relevant data is not easy to find, and 75% confirm that their data and analyses lack reliability and interpretability.

These figures predate the agentic AI wave. As AI agents begin automating critical business processes, the stakes rise sharply: the infrastructure gaps that slow data consumers² down can cause agents to act on incomplete or untrustworthy data. A data marketplace is part of the answer: it makes data discoverable, trustworthy, and accessible in a way that works for both humans and AI agents.

Humans and Agents Require the Same Four Conditions From Enterprise Data

Before we come to the proposed solution, let us dive one level deeper into the diagnosis. BARC’s research consistently points to four foundational data challenges facing organizations today:

- Data consumers cannot find the right data (discoverability)
- They do not understand what data means because it is poorly explained (context and semantics)
- They cannot access data quickly enough due to complex and manual data access workflows (data access governance)
- They cannot assess whether the data they are looking into is trustworthy (quality signals)

We argue that AI agents face the same four barriers, just with less tolerance for failure. Let me give you two examples:

Situation	Human Behavior	Agent Behavior
Data Quality does not fit purpose	Can pause, investigate, or ask a colleague	May not recognize the quality gap at all. Outcomes range from proactively seeking better data to aborting the task, or silently proceeding with unsuitable data and hallucinating results.
Lack of access rights for the intended purpose	May request, wait, ask the owner via a phone call	Either stalls or, if metadata is missing or unclear, draws on a data asset it has physical access to – but does not have the right to use it for the new, deviating purpose.

¹ BARC Survey "Data Mesh and Data Fabric" 2024, sample sizes: For the survey in total n=197, for the cited questions: n=197

² "Data consumers" is simply a term for people who need data to do their job

Unfortunately, these are not edge cases, but rather the predictable consequences of deploying agents on top of data infrastructure that was not designed with agentic AI in mind.

The implication is straightforward: data must be discoverable, contextualized, and accessible in a way that works for AI agents. Then it will work for humans, too, because they are more adaptable.

Arbitrary Data Assets Will Not Do

Addressing these four barriers is not only a matter of the data platform or user interface design. It starts with how data is packaged for consumption. Let us use an analogy: A bulk active pharmaceutical ingredient, such as ASA, is not a product. Aspirin³ is, because it is the dosed, labeled version of it, packaged with contraindications, from a responsible manufacturer, and coming with a usage insert. The analogy is simple: if you want to offer something on a marketplace, make a product out of it that is easy to evaluate, consume, and trust.

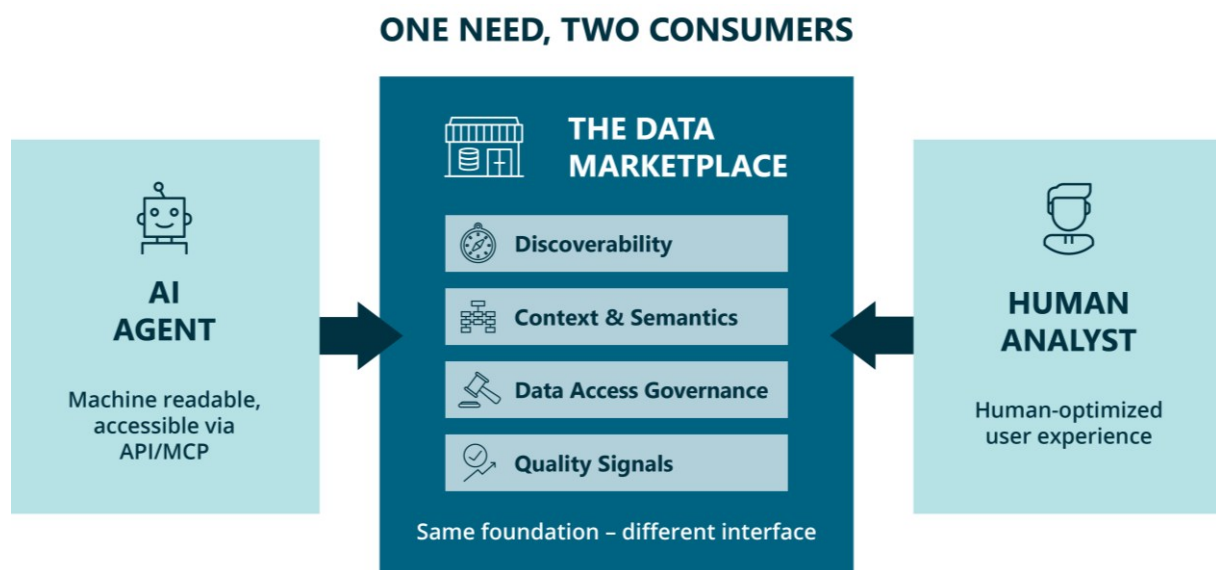


Figure 1: One need, two consumers

Data Product Definition – Deep Dive for the Curious

A data product is a shoppable, reusable, active, and standardized data asset designed to deliver measurable value by applying product thinking principles. It includes one or more artifacts enriched with metadata such as governance policies, data contracts, and optionally a SBOM (Software Bill of Material). Aligned to a specific domain or use case, it ensures accountability, continuous evolution, scalability, and compliance with business and regulatory standards.

– Jacqueline Bloemen, Florian Bigelmaier (both BARC), 2025, slightly adjusting the Bitol definition of data products.

³ Aspirin® is a registered trademark of Bayer AG. This report is independent and not affiliated with or endorsed by Bayer.

Find out more about data products at barc.com/data-products

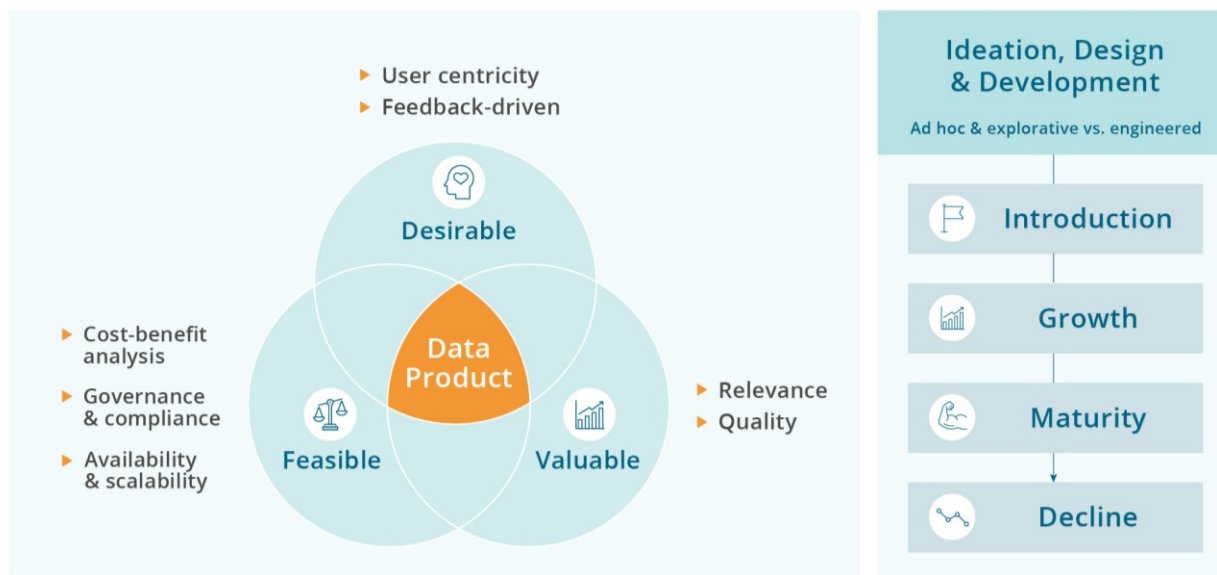


Figure 2: The most powerful aspect behind the data product concept is product thinking: creating a data asset that is truly desirable and valuable but also feasible to build and maintain.

Data products bundle everything a consumer needs: purpose-oriented data, metadata, SLAs, clear ownership, sustained maintenance, and terms of use. Arbitrary datasets without ownership and semantic documentation do not generate reliable demand.

However, this is not something you can build overnight. BARC’s research shows that 49% of organizations identify building a data product organization and mindset as an active challenge⁴, because it requires changing from a service-oriented model to a product-oriented model. A marketplace can be a catalyst in this challenge, but the cultural transformation this requires is not to be underestimated.

Ownership of data products is the strongest organizational tool we have to increase trust in data. Or put the other way round: Data products will only succeed if business-side owners accept accountability for the data their processes produce. Data engineers and application developers can build infrastructure for digitizing and automating those processes; they can be responsible for the technical pipelines that bring the data from A to B and refine them. But they cannot be held accountable for the process that creates the data.

One simple example: A data engineer rarely has the power to tell a sales agent how to enter data in the CRM system.

AI raises the stakes of data quality. Business managers now have unusual organizational leverage: the data quality issues that teams long tolerated as "imprecise dashboards" will, in an agentic world, translate directly into broken automation and unreliable agents. That consequence is hard to ignore.

⁴ BARC Survey "Data Mesh and Data Fabric" 2024, sample sizes: For the survey in total n=197, for the cited questions: n=121

When Data Products Don't Speak the Same Language

There is a further dimension that arbitrary datasets fail on: interoperability. A single well-packaged data product is useful but the combination of two or more is where insight compounds. But combining products from different domains requires shared semantics. Does "customer" in the sales domain mean the same as "customer" in the logistics domain? Does "revenue" follow the same recognition rules in finance as in commercial reporting? Packaging data as a product can, paradoxically, reinforce this fragmentation: Domain ownership gives teams less incentive, not more, to align their definitions with the rest of the organization.

One approach to solve this is to work with more explicitly documented definitions. Data products need to explain not just what they contain, but how their key terms relate to, extend, or *contradict* equivalent terms in other products. A data marketplace that surfaces these relationships, explicitly flagging where definitions align and where they diverge, turns passive discovery into informed combination. Consumers and agents alike can then make explicit choices about which definition to use and why, rather than unknowingly combining incompatible representations of the same real-world concept.

Data Contracts: The Missing Item to Automating Data Access

While establishing true ownership is the most common organizational method to increase trust in data, there is a way to further strengthen the relationship between data producer and consumer: data contracts.

Data contracts spell out what is delivered and under what conditions. This serves several purposes: First, a contract formalizes the relationship between producer and consumer. In practice, this means a consumer knows exactly what to expect from a data product, who is accountable for it, and whom to contact if something breaks. Instead of an informal dependency on whoever built the pipeline, the relationship becomes explicit: a named owner with defined obligations, and a consumer with clearly stated usage rights, bound to a specified purpose.

Second, it is a structured source of information about a data product, comparable to the package insert that comes with medication, to stay with the analogy above. Having quality signals embedded in the contract lets consumers, and agents especially, assess fitness for purpose before they consume.

Third, contracts make data quality checkable. A contract defines what a data product is supposed to deliver. They may include format, completeness, and freshness expectations. So, actual data can be validated against those specifications automatically. For agents especially, this matters: rather than propagating quality issues downstream, an agent can verify fitness for purpose before processing.

Fourth, contracts make access management automatable. If the rules about who may use a data product and for what purpose are already written down, the system can approve access requests automatically without human involvement. For a human, this means faster access to the data they need. For agents, this matters even more, because they will eventually request access to data far more often than humans ever did.

Data contracts, however, are not sufficient on their own. Much like in the real world, where contracts govern individual relationships but operate within a broader legal framework, data contracts need organization-wide policies to give them context. Policies define what is universally permitted or prohibited and which standards to follow; contracts specify additional rules within those boundaries for one specific data product.

The Data Marketplace: Where It All Comes Together

Every mature data platform has at least a technical metadata repository, just as every library has an index. Such tools help with discoverability. What they cannot do is create trust. A metadata index tells you that a dataset exists but it does not tell you whether it is reliable, what it means in your business context, or whether it fits your use case. Trust, built through context, semantics, and quality signals (as explained above), is what turns passive discovery into actual demand and what makes an internal data marketplace so valuable.

THE DATA MARKETPLACE AS FACILITATOR

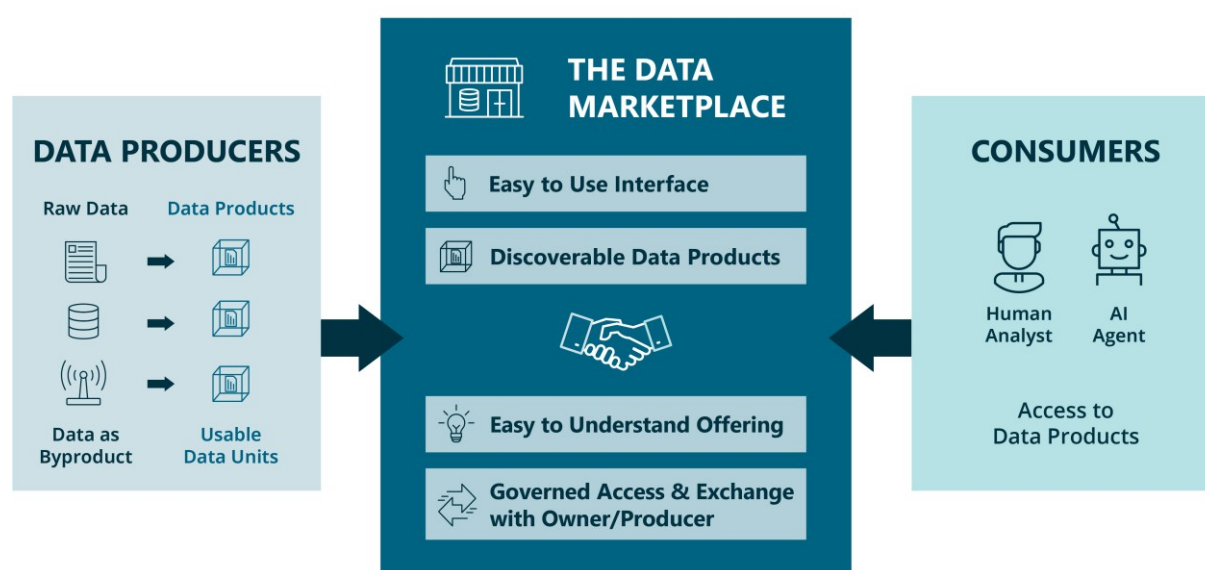


Figure 3: How a data marketplace facilitates data exchange and consumption

Many data marketplaces feel like an Amazon shopping experience – but they begin to break as soon as you want to actually get that product. Once a consumer has found a trustworthy data product, the next expectation is immediate access. This is where many organizations still fall short, with manual approval workflows that take days or weeks. A data marketplace addresses the discovery part of this loop. At its core, it is a curated metadata search engine that surfaces data products with the context needed to assess their fit and trustworthiness. Most marketplaces today stop there. The more ambitious vision, and the direction the most mature implementations are moving toward, is to extend this into access provisioning as well. Rather than simply pointing a consumer toward the right data product and then leaving them to open a ticket, the marketplace would also govern and automate the access request itself.

Trust is the market mechanism that is missing in many companies with failed marketplace implementations. **Successful implementations have one thing in common:** consumers find an item and can quickly assess to what degree a data product is suitable for their use case and rely on this judgment.

The more complex an organization becomes, the more valuable a data marketplace is. More business domains, more data platforms, more teams working with data: all of this increases the need to know what data exists, where to find it, and how to access it. It needs one reliable entry point that provides sufficient context to select and consume the right data product.

⚠ A word of caution: This only holds if the marketplace is kept current. As the number of registered products grows, so does the maintenance overhead. A poorly maintained marketplace quickly becomes noise rather than signal, and consumers stop trusting it. Every marketplace project has a point where complexity starts working against discoverability. This brings us back to what was said earlier: ownership for data products is a powerful concept. It does not only render data products trustworthy; it is also the force that keeps the marketplace up-to-date.

The marketplace also resolves one of the central tensions of data products as a part of efforts to give the business domains more power: Federated organizations and architectures are not fully decentralized. They should have a well-defined spine. Many organizations rightly push the ownership and creation of data products to business domains, but two things must remain central: a single access and discovery point (because a marketplace follows scaling laws), and a certain set of organization-wide governance policies that apply to all published products. The marketplace can be the natural home for both. Such marketplaces can validate contracts against organization-wide policies at publication and evaluate access requests against those same policies at consumption.

Access Granted Is Not Access Governed

Marketplaces that extend beyond discovery into access provisioning gain a capability classical systems lack: they can record not only who accessed what, but why. Tracking this last dimension (the *purpose* of data consumption) is typically absent from classical access control, yet it is the prerequisite for automating data access at scale.

In most organizations, data access works like an RFID badge: privileges accumulate and rarely expire or get revoked. Three years after switching from the manufacturing team to the engineering office, you may still be able to open the shopfloor door at 3 am on a Saturday.

For example, imagine a large corporation that has data about customer lifetime value. A data scientist gets access to build a sales prioritization model. It works really well. Then a marketing colleague asks to use the same CLV data for campaign segmentation. The data scientist shares it informally without a new access request. The data is now used for a purpose it was never cleared for.

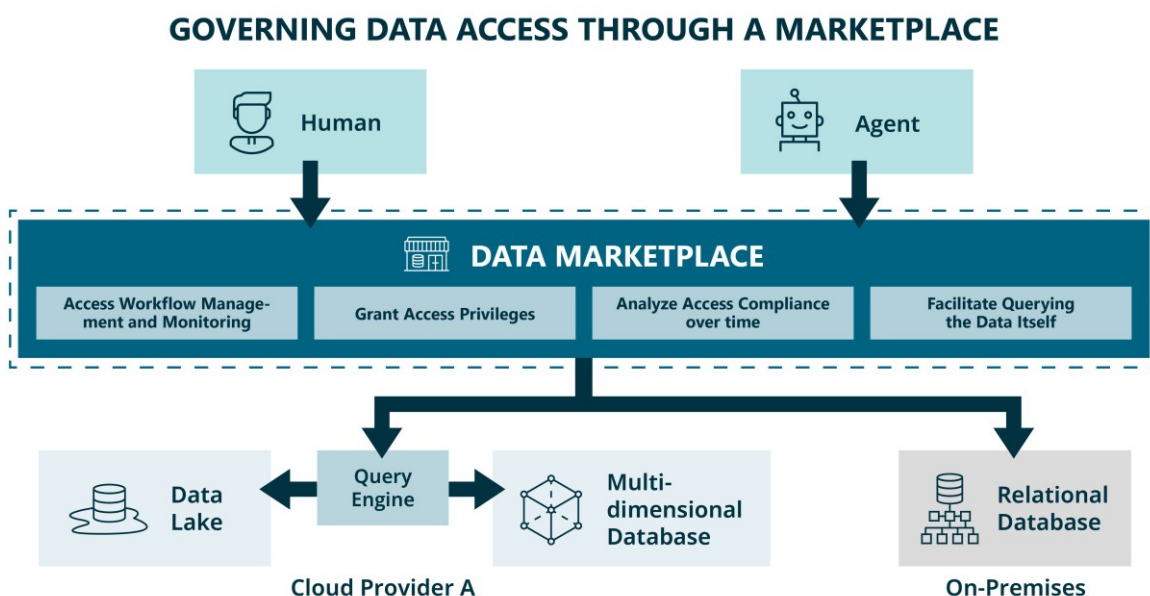


Figure 4: Governing data access through a marketplace

For human analysts, informal signals, review processes, and professional judgment create some natural friction against use without permission and blatant misuse. AI agents operate without moral judgment and without that friction. Trying to build such scruples into them is never 100% successful.

The consequence: Organizations need to protect the data itself. They can no longer treat the granting of access as the end of governance. Classical access control is a point-in-time decision: a request arrives, the requester states a purpose, and the system either grants or denies access. Once granted, the assumption is that usage stays within the declared scope. At best, the access will be revisited once every few years.

Back to the example, but now we swap the data scientist for an AI agent. The same CLV data product and the same declared purpose: sales prioritization. Three months later, the agent starts using the data to feed a risk model that drives the question of whether a customer can purchase via invoice or they should pay in advance. While it might make business sense, this is clearly a different purpose that has neither been declared nor approved. In a compliance audit, no one can explain why the agent processed customer data outside its original scope.

If an agent can access data it was not supposed to use for a given purpose, it eventually will. Murphy had no idea he was writing data governance policy!

The solution is not to revisit access definitions more often. It is to shift governance logic from access as a one-time approval to access as a continuously enforced state. A data marketplace holds both the contract and the context to evaluate whether a given query is in bounds. It can also surface whether the guarantees stated in a data contract are actually being met. This makes it the natural place to capture the "why" behind every access grant and enforce compliance when queries deviate from the declared purpose.

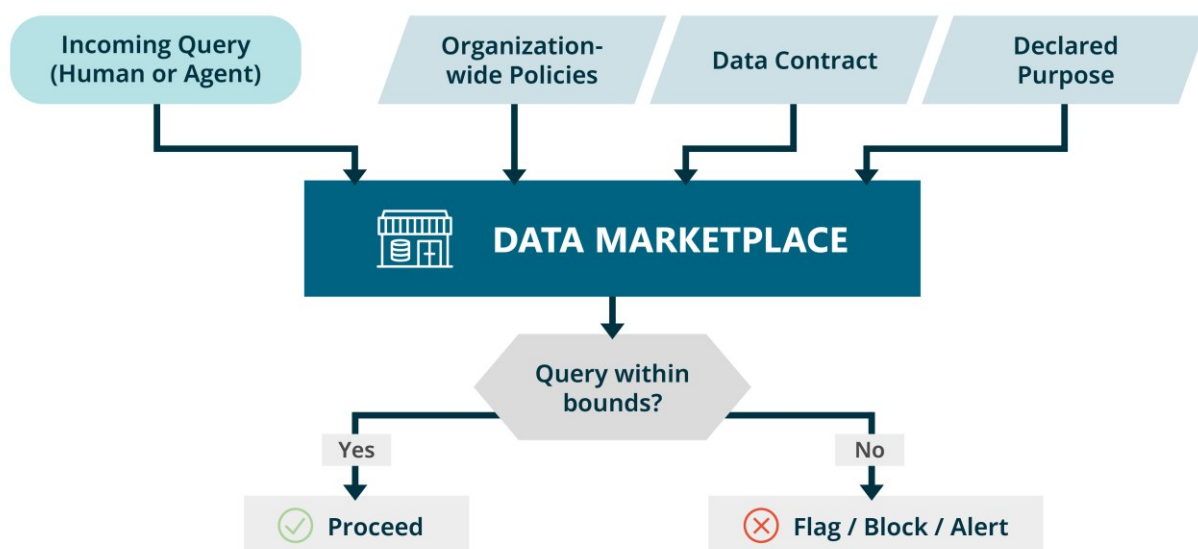


Figure 5: How the Data Marketplace evaluates whether a query is within bounds

Governing data access is one piece of the much larger puzzle called AI Governance. This discipline also covers model behavior, output quality, explainability, observability, human oversight, ethical use, risk management, and regulatory compliance, to name a few. And even within the data dimension, access control is not the whole story: just think of bias, consent, retention policies etc. This article focuses on governing one very important aspect: the data access of agents.

Four Actions to Start

The path forward is not a big-bang marketplace rollout, it is much more an organizational transformation supported by getting one data product on the marketplace at a time, and convincing one team at a time that it will be worth the effort.

1. **Invest in culture:** Data product ownership belongs with the business. Build literacy and accountability in parallel with infrastructure.
2. **Run a narrow pilot:** Many marketplace and catalog initiatives fail because their scope is too broad in the beginning. We have observed that for most initiatives it is most promising to start with one domain that is willing to collaborate and create a joint success story (e.g., by showing that agentic AI has more success when business domains go the extra mile and register their data products on the marketplace).
3. **Design for agents:** Optimize metadata so that agents can find the most suitable data offering quickly. This could mean following industry standards like ODCS for data contracts, but also adding AI instructions per data product, synonyms, and links to deeper semantics.
4. **Put your data under contract:** Start with data contracts on your most critical data domains. Even one contract per domain changes the governance conversation and creates the foundation for "automated policy enforcement": The possibility to automate governance because deterministic algorithms and/or LLMs can decide whether an action by a human or an agent is in compliance or in opposition to the contract and policies that apply.

Organizations that invest in data infrastructure but neglect data marketplaces, contracts, and product thinking will find that their AI agents inherit the same trust deficit that has plagued human analysts for years. The infrastructure for human-grade analytics and agent-grade automation is, at its foundation, the same. The consequences of getting it wrong are not.

About Entropy Data

Entropy Data offers a data product marketplace, enforced by data contracts, that works for both human analysts and AI agents.

It is built upon open standards, natively supporting the Open Data Contract Standard (ODCS), the Open Data Product Standard (ODPS), and the Open Semantic Interchange (OSI), and we actively contribute to those standards to make them even better. Together, these standards make data products discoverable, contextualized with business semantics, and accessible through automated access workflows, surfacing the quality signals consumers need to trust what they find. The contracts themselves are enforced by our popular open source Data Contract CLI.

Three building blocks power it:

- **Marketplace** handles self-service discovery and access, with Entropy Intelligence answering natural-language questions across data products and their data.
- **Studio** lets teams design data products contract-first (even from a simple Excel template), build them with the Data Product Builder using AI coding agents, and monitor their usage.
- **Governance** enforces organization-wide policies and purpose-based access control, so every grant and query stays within its declared purpose.

Together they give agents a trustworthy entry point via API and the Model Context Protocol (MCP), and give people a marketplace they can rely on.

Entropy Data integrates with all major data platforms and is available as EU-based SaaS, bring your own cloud, or on-prem deployment.

Entropy Data is ISO 27001 certified, and deployed in production by large enterprises across multiple industries such as logistics, pharma, media, and telco, and across multiple geographies, including the USA, Australia, Switzerland, and the EU. This demonstrates that our contract-enforced data product marketplace scales well beyond the pilot stage.

Contact info

Entropy Data GmbH

Krischerstr. 100
40789 Monheim

Germany

www.entropy-data.com
hello@entropy-data.com

Managing Directors

Dr. Simon Harrer
Jochen Christ



About BARC

BARC is the leading analyst firm for data & analytics, AI, corporate performance management (CPM), and ESG with a reputation for unbiased and trusted advice. Our expert analysts deliver a wide range of research, events, and consulting services for the data & analytics community. Our innovative research evaluates software, vendors, and service providers rigorously and highlights market trends, delivering insights that enable our customers to innovate with data, analytics, and AI. BARC's 25 years of experience with data strategy & culture, data architecture, organization, and software selection helps clients transform into truly data-driven organizations.

Research

BARC user surveys, software evaluations, and analyst advisory services along with expert driven content such as research notes, trend analysis, and blogs give organizations the confidence to make the right decisions. Our independent research gets to the heart of market developments, evaluates software, vendors, and service providers thoroughly and gives valuable ideas on how to turn data, analytics, and AI into added value and successfully transform businesses.

Consulting

The BARC consulting practice is entirely focused on translating companies' requirements into future-proof decisions. The holistic advice we provide helps companies successfully implement their data & analytics strategy and culture as well as their architecture and technology. BARC's research and experience-founded expert input sets organizations on the road to the successful use of data & analytics, from strategy to optimized data-driven business processes.

Events

At BARC events, leading minds and industry experts come together to share insights and drive innovation. Our conferences, roundtables, and online webinars attract over 10,000 participants annually, offering a unique blend of information, inspiration, and interactivity. These events provide a platform to exchange ideas with peers, explore emerging trends, and gain expert perspectives on market developments. By engaging with thought leaders and industry practitioners, participants discover actionable strategies to enhance their business and stay ahead in the evolving world of data & analytics.

BARC

Copyright © BARC GmbH 2026. All rights reserved.

BARC

Data Decisions. Built on BARC.

www.barc.com

Germany

BARC GmbH
Berliner Platz 7
D-97080 Würzburg
+49 931 880651-0

Austria

BARC GmbH
Hirschstettner Straße 19 / I / IS314
A-1220 Wien
+43 660 6366870

Switzerland

BARC Switzerland GmbH
Buchhaldenstraße 7
CH-5442 Fislisbach
+41 76 340 35 16

United States

BARC US
13463 Falls Drive
Broomfield, CO 80020
USA