LESSON 1.4: DATA OWNERSHIP VS. DATA USE RIGHTS

LESSON 1.4: SUPPLEMENTAL READING

One of the most common sources of regulatory exposure in data governance arises from a deceptively simple confusion: the belief that having data is the same as having the right to use it. Firms operate under the assumption that if data resides in their systems—whether generated internally, licensed from vendors, or collected from clients—they possess inherent permission to apply it however operational needs or business opportunities suggest.

Regulators don't share that assumption. From a supervisory perspective, custody of data and permission to use data are distinct concepts, governed by different frameworks and carrying different obligations. A firm may lawfully own or store a dataset while still lacking the right to use it for certain purposes. Understanding this distinction isn't a technical matter; it's foundational to defensible governance.

Ownership: Custody, Control, and Stewardship

Data ownership typically refers to control over a dataset's storage, maintenance, and safeguarding. It answers questions about who is responsible for preserving the data, protecting it from unauthorized access, ensuring its accuracy and integrity, and maintaining it over time.

In practical terms, ownership might mean:

- A firm owns internal transaction records generated by its own systems
- A firm licenses market data from a vendor and maintains that data within its infrastructure
- A firm collects and stores client information as part of account opening or servicing processes

Ownership establishes custody and stewardship obligations. It inherits responsibilities to protect the data, retain it according to regulatory requirements, and ensure it remains available for legitimate business and compliance purposes. But ownership alone doesn't answer the question: *What can we do with this data?*

That question requires examining use rights.

Use Rights: Purpose, Context, and Limitation

Use rights define what purposes are permitted, what contexts are allowed, and what limitations apply when data is applied to business activities, decisions, or communications. These rights are shaped by multiple overlapping sources:

Regulatory requirements: Laws and regulations often specify how certain types of data may be used. Consumer financial information, for instance, may be restricted under

privacy regulations or fair lending rules. The fact that a firm owns the data doesn't override regulatory restrictions.

Client agreements and disclosures: When clients provide information, they do so with certain expectations, often formalized through disclosures, terms of service, or privacy notices. These agreements may permit data use for account servicing but restrict use for marketing, credit decisioning, or sharing with affiliates.

Vendor licenses: Third-party data providers typically impose contractual limitations on how their data may be used. A market data license might permit use for internal research but prohibit use in client-facing products or external communications. Once the data is integrated into internal systems, these limitations can become invisible to downstream users.

Internal governance policies: Even when external constraints don't exist, firms may establish internal policies that limit how data is used to manage risk, ensure fairness, or align with ethical standards.

The critical insight is that use rights are **context-dependent**. The same dataset may be permitted for one purpose but restricted from another. For example, transaction history might be appropriate for operational reporting, but not for client segmentation. Behavioral signals might be acceptable for fraud detection but not for determining eligibility for financial products.

When firms fail to articulate these distinctions clearly—when they treat all owned or licensed data as uniformly available for any conceivable use—data gets reused informally, often without deliberate intent to violate constraints. The absence of explicit restrictions is mistaken for permission.

Why Regulators Focus on Use, Not Ownership

Regulators evaluate data governance through the lens of impact. They're concerned with how information influences decisions, communications, and client outcomes. From this perspective, the fact that a firm owns or licenses data is relevant but not determinative. What matters is whether the data was used appropriately.

Examiners ask:

- Was this data used consistently with its permitted purpose?
- Did the use align with client expectations, disclosures, and consents?
- Was the use documented and supervised?
- Who was accountable for validating that this use was appropriate?

Ownership may establish custody and stewardship obligations. Use establishes regulatory responsibility. A firm that owns data but uses it beyond permitted boundaries isn't absolved by the fact of ownership; it's exposed precisely because it failed to recognize the distinction.

This regulatory focus means that firms can't rely on technical access as a substitute for governance. The fact that data is physically present in a system, extractable by authorized personnel, and technically usable does not mean it should be used without review.

Common Points of Confusion: Where Use Rights Break Down

Confusion between ownership and use rights often emerges not in high-risk, deliberate decisions, but in routine workflows where assumptions go unquestioned:

Repurposing without review: Data originally collected for compliance, operations, or regulatory reporting is later reused for analytics, marketing, or client segmentation. Because the data already exists in internal systems, no formal review is conducted to determine whether the new use is permitted. The assumption is that internal data can be freely repurposed for internal purposes. Regulators don't accept that assumption.

Third-party data integration: Licensing agreements for vendor-provided data often impose purpose limitations. These can include restrictions on redistribution, constraints on derivative uses, and prohibitions on combining the data with certain other sources. Once the data is ingested into internal systems and transformed or enriched, these original constraints may no longer be visible. Teams downstream assume that if the data is in the system, it's available for their use case.

Derived data and cascading permissions: As we saw in lesson 1.3, when multiple datasets are combined to create derived outputs—scores, rankings, segments—each underlying dataset may carry different use restrictions. The derived output inherits the most restrictive constraint, but this is often not tracked. For example, a propensity model trained on client transaction data (permitted for internal research) and third-party behavioral signals (restricted from client-facing use) may produce outputs that violate the latter restriction, even though the model itself seems like an internal tool.

Informal data sharing across teams: Business units or functions may share datasets internally under the assumption that since the firm owns the data, internal sharing is unrestricted. But use rights can vary by function. Data permitted for risk management may not be appropriate for sales or marketing. Data collected under one regulatory framework may not be usable under another.

These gaps are rarely malicious. They arise from the absence of clear frameworks that make use rights explicit and enforceable. In the absence of those frameworks, teams rely on informal judgment, and assumptions proliferate.

Frameworks also include requirements for **documentation**, and this documentation isn't static. As data moves through systems, is combined with other sources, or is repurposed for new applications, the use rights narrative must be updated and reassessed. Firms that document ownership but not use rights can demonstrate custody, but they must also demonstrate control.

Why This Matters Before Analytics or AI

Analytics platforms, machine learning models, and **AI-driven systems accelerate data reuse. They make it trivial to apply the same dataset across multiple teams, business lines, and use cases, often with minimal human oversight at each step.** This efficiency is valuable, but only if the underlying governance is sound.

If data use rights are unclear or undocumented before automation is introduced, firms risk scaling unauthorized or poorly supervised use:

- A marketing model trained on transaction data might violate restrictions on using that data for targeting purposes
- An AI-driven recommendation engine might combine datasets in ways that exceed the permissions granted for individual sources
- A risk assessment tool might repurpose client information beyond the scope of original consent or disclosure

By the time an issue is identified—through a regulatory exam, a client complaint, or an internal audit—the data may already have influenced thousands of decisions or communications. The fact that the firm owned the underlying data provides no defense.

Establishing clear distinctions between ownership and permitted use before deploying analytics or AI ensures that automation amplifies appropriate practices rather than inappropriate ones. This is what we mean when we say interpretation moves upstream (to infrastructure and governance) when automation replaces judgment.

Conclusion

Data ownership and data use rights aren't the same. Ownership establishes custody and stewardship obligations. It says who is responsible for protecting and maintaining the data. On the other hand, use rights establish what can be done with it and says what purposes are permitted, what contexts are allowed, and what limitations apply.

Regulators focus on use, not ownership. The fact that a firm possesses data doesn't resolve whether applying it to a particular decision, model, or communication is appropriate. Firms must be able to demonstrate that every use of data aligns with the permissions, consents, and constraints that govern it.

Before introducing analytics, automation, or AI, firms must establish clear frameworks that make use rights explicit, enforceable, and traceable. This isn't a constraint on innovation; it's the governance foundation that allows advanced tools to be deployed safely and responsibly.