
Improving Data Governance – and the Quality of Your Returns

Authors: Dane Rook and Ashby Monk

1. The Bedrock of Intelligent Investing

Data governance (DG) is about making sure that the right data (and the right *quality* of data) underpin investment decisions.¹ It is the bedrock of intelligent investing, and the world's top investors excel at it² – for good reason: the quality of an investor's DG significantly dictates the quality of investment returns.³ This is one of our two key findings from a long-term study of institutional investors' approaches to DG (see Rook and Monk [2024]). The other is that most investors do DG poorly—which is surprising, given that **any investor, regardless of size and resources, can do DG well as long as they know where to focus their efforts.**⁴ To that end, we provide summarized guidance in the next section. First, we discuss the reasons why poor DG is prevalent across the investor community.

The first reason is that the performant DG is not one-size-fits-all. To work well, it must align with an investment organization's overall *context* (i.e., its resources [such as talent, budget, culture, technology, infrastructure], strategy, goals and constraints; see Monk and Rook [2023]). Many investors do not take care to foster this alignment between DG and context, either because they believe it is too difficult to achieve (which is untrue) or because they do not fully grasp their own

¹ Good DG entails three essentials: 1) clarity on the provenance of data (that is, its original source, the chain of custody and the transformations performed on it before its final application to some decision, 2) clarity on responsibilities and expectations with respect to data quality and 3) guardrails on the use of data, such that a particular decision is based on data of a pre-established level of quality.

² The fact that leading hedge funds excel at DG, and that this excellence is a key ingredient to their serial high performance, is something of an open secret. Clearly, their governance practices grant them a competitive advantage; thus, most are reluctant to share their specific governance activities in a detailed way. Hence, we instead draw our insights primarily from the governance practices of the world's top institutional investors (that is, public pension funds, sovereign wealth funds and endowments). Doing so is not problematic, however, because the governance practices of those organizations are more likely to reflect best practice ideals for the vast majority of (non-hedge fund) investors.

³ Notably, there is more to returns quality than just the gross magnitude of returns. Other aspects, such as stability, comprehensibility, efficiency and manageability also demonstrably matter (we discuss this more extensively in the parent paper to this piece; see Rook and Monk [2024]).

⁴ It is worth mentioning that DG is practically a zero-sum game for investors: one investor doing DG well does not increase the difficulty of other investors doing likewise.



context in the first place. Another reason for pervasively weak DG is uncertainty about best practices due to a lack of research on DG as it relates to investors. Third, DG is often confused or conflated with data management (DM).

DG and DM are *not* the same thing: DM concerns how data is accessed, handled and stored in the organization, but it does not cover how data is used for decision-making—that is the domain of DG. The misconception that good DM necessarily equates to good DG has caused many investors to under-prioritize and under-resource their DG efforts.⁵ A fourth reason why DG tends to be so poor is that there is no single technology solution that investors can buy to “solve” their DG needs. This is in part because good DG must be individually tailored, and DG needs are ever changing as markets and data sources/systems continually evolve. While this dynamic nature may make DG seem daunting, it should hearten many investors. This means that there are always new opportunities to improve and differentiate oneself on DG.

To reiterate our most important point, **good data governance is not simply something that is nice to have; it is what every investor needs to make genuinely intelligent investment decisions.** Every investor is capable of good DG if they start by focusing on the elements discussed below.

2. Focusing Data Governance: What, Who and How

Fundamentally, good DG is like all good governance; it stems from formalized and transparent internal processes, structures (such as oversight committees) and tools.^{6,7} Further, as with other forms of governance, DG is a scarce resource, and investors can only do so much of it due to limited budgets, time, in-house skill and employee goodwill. Hence, in designing a sound DG program, investors must determine what, who and how to govern data.

What: The first thing an investor must focus on is **deciding what data to govern** versus what data will be allowed to go ungoverned (or minimally governed). This decision is greatly clarified by starting—for any given dataset—with the following questions: For what purposes will this dataset be used (and are those purposes vital enough to justify the cost of governance)? When is it appropriate to start governing it (the answer might not be straightaway)? The success of a good DG program rests on not making it too extensive or onerous. Otherwise, operations will slow

⁵ Good DM is a necessary, but not by itself sufficient, condition for good DG.

⁶ For detailed discussion of best-practice governance (of all forms), see Clark and Urwin [2008].

⁷ Culture can also play a meaningful role in good DG. However, culture can be difficult to control precisely and is characteristically slow to change. Hence, culture is best thought of as an informal (and secondary) lever for governance.



excessively, and employees will become exasperated—and may seek ways to circumvent governance.

A reliable step toward right-sizing the scope of governance is to start backward and identify the most critical types of decisions the organization makes and trace those back to the datasets on which those decisions are based (or should be). Determining the necessary quality for those critical decisions can help in establishing the necessary DG (i.e., the data quality that must be delivered). For instance, different levels and types of data quality are required for reporting vs. informal/quick ideation vs. trading.⁸

Who: The second area of focus lies in **determining who is responsible** for governing the data selected for governing. At a general level, the organization should actively decide whether it aims to have *centralized* or *federated* governance. Under centralized setups, a single person or unit (often an IT or specialist data team) is solely responsible for designing DG policies (standards, roles, etc.) and enforcing them. Under a federated system, responsibility for policymaking and enforcement is distributed throughout the organization. For example, individual teams may be allowed to set their own standards for the datasets they create or manage.

In reality, almost every DG system falls somewhere between full centralization and complete federation. This makes it necessary to have clear rules for designating stewardship, such as assigning specific tasks to ensure data quality and usage to specific individuals. Such rules should articulate:

- The level of discretion a steward has in determining the quality standards for datasets under their care
- The obligations the steward has in upholding those standards (which may include policing how the dataset is used by others)

The key to successful stewardship is *specificity*. Stewardship of specific datasets must be assigned to specific people, who are given specific obligations regarding those datasets.

How: The third area to focus on is **how data is governed**. This set of decisions entails the following:

- What standards should be applied (and how should these be tested)?

⁸ See Monk et al. [2019] for more detail on the various dimensions of data quality.



- What enforcement mechanisms should be in place to ensure compliance with standards and make certain that the correct data is used for a given investment decision?
- What assistive structures (e.g., DG committees) and tools (e.g., automation of quality checks) are helpful or essential?

If the “what” and the “who” of governance (for a particular dataset) are well specified, making these “how” decisions often becomes straightforward. Yet, all three sets of decisions (what, who and how) must take the organization’s context into account if the resulting DG system is to work well. That said, investors should treat their context as static and immutable; various resources can be altered to better accommodate a new or improved DG system. These include deeper buy-in from senior leadership, creation and promotion of a common language for talking about DG within the organization and fostering a data culture—particularly one in which all datasets are treated as an organizational resource and not the “exclusive territory” of specific individuals or teams.

By tackling the three focus areas above, **any** investor will be well on their way to having performant DG capabilities.

We regularly publish reports on investment trends we see across high-net-worth investors. Please reach out to us at research@addepar.com if you’re interested in discussing these topics.



References

Clark, G., and R. Urwin [2008] "Best-Practice Pension Fund Governance," *Journal of Asset Management*, 9(1).

Monk, A., M. Prins, and D. Rook [2019] "Rethinking Alternative Data in Institutional Investment," *Journal of Financial Data Science*, 1(1).

Monk, A., and D. Rook [2023] "The 'Investor Identity': The Ultimate Driver of Returns," *Social Science Research Network*, available at:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4324537.

Rook, D., and A. Monk [2024] "Governing Your Way to Better Long-Term Returns," *Addepar Research Brief*.



Notice and Disclaimers

All information provided by Addepar, Inc. or its subsidiaries (collectively, “Addepar”), including without limitation, all text, data, graphs and charts (collectively, the “Information”) is the property of Addepar and is provided for informational purposes only. The Information may not be modified, reverse-engineered, reproduced or re-disseminated in whole or in part without prior written permission from Addepar. All rights in the Information are reserved by Addepar.

The Information may not be used to create derivative works or to verify or correct other data or information without prior written permission from Addepar. For example (but without limitation), the Information may not be used to create indexes, databases, risk models, analytics, software or in connection with the issuing, offering, sponsoring, managing or marketing of any securities, portfolios, financial products or other investment vehicles utilizing or based on, linked to, tracking or otherwise derived from the Information or any other Addepar data, information, products or services.

The user of the Information assumes the entire risk of any use it may make or permit to be made of the Information.

Addepar makes no express or implied warranties or representations with respect to the information (or the results to be obtained, but rather the use thereof), and to the maximum extent permitted by applicable law, Addepar expressly disclaims all implied warranties (including, without limitation, any implied warranties of originality, accuracy, timeliness, non-infringement, completeness, merchantability and fitness for a particular purpose) with respect to any of the information.

Information containing any historical information, data or analysis should not be taken as an indication or guarantee of any future performance, analysis, forecast or prediction. Past performance does not guarantee future results.

The Information should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. All Information is impersonal and not tailored to the needs of any person, entity or group of persons.

None of the Information constitutes an offer to sell (or a solicitation of an offer to buy), any security, financial product or other investment vehicle or any trading strategy.

Addepar does not recommend, endorse, approve or otherwise express any opinion regarding any issuer, securities, financial products or instruments or trading strategies and Addepar’s research products or services are not intended to constitute investment advice or a recommendation to make (or refrain from making) any kind of investment decision and may not be relied on as such.

Addepar, Investment Sentiment Index and other Addepar brands and product names are the trademarks, service marks or registered marks of Addepar or its subsidiaries in the United States and other jurisdictions.

© 2024 Addepar. All rights reserved.