
Governing Your Way to Better Long-Term Returns

Higher-quality alpha from higher-quality data

By Dane Rook and Ashby Monk

1. Welcome

Data governance isn't sexy, but it can lead to better investment decision-making, which in turn can lead to better investment returns ... and better returns are always sexy.¹ This chain of improvements is driven by the essential activity of data governance (DG): **ensuring that the right data — and in particular, the right quality of data — underpins each investment decision.**²

However, few investors are doing DG well ... or at all. This DG deficit is partly due to a lack of clarity regarding best practices for DG in investing. In this piece, we outline some DG best practices that can help investors increase the quality of their data, decisions and long-term returns.

A deep understanding of DG can't be achieved through external observation. To really "get" what's going on with an organization's data governance, one must be immersed in that organization and study its technology, its people and culture, and its processes for generating returns. Each of these resources feeds the success/failure of the organization's approach to DG; therefore, any performant DG strategy must take account of them, individually and collectively. We are uniquely positioned to discuss the need for a holistic understanding of DG best practices with investors as our team, by virtue of its ties with the Stanford Research Initiative on Long-Term Investing, has, over the last two decades, used close-dialogue methods to complete hundreds of research interviews and in-depth case studies with the world's biggest asset owners — possibly more than any other group on Earth. We're therefore confident that the findings here are not only accurate,

¹ A decade ago, data science was called the "sexiest job of the 21st century" (Davenport and Patil 2012). We're not convinced that's true today (or was true even back then). However, the sex appeal of data science is certainly lessened whenever data scientists are compelled to assume responsibilities for data governance (which they often aren't).

² For those readers unfamiliar with data governance, an informal/working definition is: the resources and practices needed for ensuring the proper quality and use of data in investment decision-making. We will be providing a more extensive definition later; what's key to realize here is that data governance is not the same thing as data management.



but differentiated, and hopefully valuable to any investor, regardless of their size, location or ambitions.

2. NTK

Here are the key takeaways from this brief that you need to know (NTK):

- **Decision mapping:** Data governance is about mapping data to decisions and ensuring the quality of both. This makes data governance distinct from data management. Simply managing data isn't the same as governing data for better decision-making, which is a point that's misunderstood by many investors.
- **Tailoring:** Data governance isn't one-size-fits-all. To work well, it needs to align with an investor's broader set of resources and processes, including its portfolio strategy, people and technology — in short, its contextualized "identity" (see Monk and Rook [2023]). Failing to get this alignment right can lead to malfunctioning data governance, which can limit the quality of investment returns.
- **Empowerment:** Well-crafted data governance empowers investors to more easily (and prudently) unlock new, game-changing capabilities, such as the nimbleness to quickly embrace novel asset classes, responsibly adopt transformative AI and conduct advanced knowledge management.

The best practices that will allow investors to capitalize on these insights are described in detail in the Findings section of this paper.

3. Significance

Recently, it's become fashionable to point to data management (DM) as a source of "operational alpha."³ This is warranted because a large fraction of most investment organizations' internal data is un(der)-utilized. Improvements to DM can make more of that data discoverable and usable, while also reducing storage costs, eliminating redundant data subscriptions, etc.).⁴

³ Popularly, operational alpha is defined as portfolio (out)performance that's attributable to improvements in an investor's internal operations (e.g., reducing its data costs or enhancing its tech stack).

⁴ In the present piece, we avoid any deep discussion of the specific architectures and tools that support good DG. We do so for two reasons: first, to concentrate attention on the people and process elements of DG; and second, because the question of architecture as it relates to DG is a complex and extensive one (which we will be covering in a subsequent paper). That said, readers should bear in mind that the quality of DG is tightly coupled with the quality of technology/architecture (both the tools and systems that directly support DG, and the data tools and systems that DG seeks to govern).



Nevertheless, DM isn't the same thing as data governance (DG). **DM mostly deals with how efficiently data is handled, whereas DG is concerned with how the quality of data affects the quality of investment decisions.** Distinguishing between DM and DG matters quite a bit because the cost of bad decisions can far outstrip the costs of inefficient data handling.⁵ From what we've observed (based on both what investors tell us and the outcomes we have witnessed), many investors don't appreciate the differing roles that DM and DG play, and so their data quality suffers, as does the quality of their investment decisions.⁶

Given the severity of this issue, it's useful to be even more concrete on the distinctions between DM and DG. For the sake of concreteness, DM and DG can be distinguished as follows:

- **Data management** involves the resources and practices related to how data enters, moves through, is stored and is accessed within the organization.
- **Data governance** involves the resources and practices related to the quality and proper *usage* of data in investment decision-making. Essentially, DG aims to ensure that:
 - The quality of data (i.e., its accuracy, completeness, etc.) isn't degraded as it moves through the organization; indeed, good DG strives to enhance data quality.
 - The provenance and chain of custody for a given data item, whether it's a single data point or an entire dataset, is transparent at all times.
 - Responsibilities for maintaining data quality are understood and enforced.
 - Mechanisms exist (and work as intended!) to discourage the use of the wrong types of data as the basis for particular categories of decisions (e.g., data that isn't of sufficient quality can't be used as the basis for trades).⁷

While the foregoing list is by no means exhaustive, the pivotal takeaway is that **DG affects more than just operating alpha: It impacts the quality of overall alpha, since high-quality alpha relies on high-quality decisions.**⁸ However, "quality" is a multidimensional concept, and there's more that

⁵ Bad decisions might occur, for example, when data isn't properly reconciled or becomes inconsistent across the organization, such as when different parts of the organization perform different transformations on the same dataset, which can possibly lead to differing "views" of the same data; these views may conflict with each other if one set of transformations is less valid than the other, and this conflict can lead to harmful dissonance at the portfolio level. Bad decisions also tend to happen when there's a mismatch between the importance of a decision (in terms of the magnitude of its consequences) and the quality of the data that underpins it.

⁶ Failure to properly appreciate the distinctions between DM and DG, in terms of the different functions they serve, can cause an investor to mistakenly think that good DM practices can replace a solid DG strategy, which is a dangerous mistake to make.

⁷ Notably, these target activities for DG only apply to data that an investor has actively decided to govern. But the decision of which data deserves governing is itself part of a DG system! (More on this below.) It's also worth mentioning that there are other facets of DG beyond these example activities; however, these are the ones that leading investors have cited as being most crucial.

⁸ This holds true regardless of whether such decisions involve in-house trades, risk-management choices or decisions about which external managers to hire or fire, etc.



goes into high-quality returns than just their (net) magnitude. For example, returns can be of higher quality when they're more:

- Stable — They're less volatile and/or more predictable.
- Comprehensible — They're generated by processes that are easier to identify and are more explainable.
- Efficient — They have lower cost ratios.
- Manageable — An investor has more control over them.

Different aspects of DG can have distinct impacts on each of these dimensions of returns quality.

Below are some examples (conveyed to us by major institutional investors) of these impacts:

- **Stability:** Better DG generally reduces “noise” trading (i.e., trading on false signals) because it leads to more accurate detection of true signals. Portfolio volatility tends to be lower because a significant amount of volatility comes from the reactionary nature of noise trading. In contrast, proactive trading tends to be more purposeful and planned, with better performance from fewer trades. This tends to create stabler performance because portfolio adjustments and course corrections can be smoother.
- **Comprehensibility:** Data quality and genuine statistical significance tend to be positively correlated: better data generally leads to clearer identification of valid relationships between variables, which translates to a better understanding of what truly drives returns.⁹
- **Efficiency:** In seeking to improve data quality, DG concentrates on the processes by which that quality is controlled. Better quality-control processes generally decrease the relative expensiveness of data in the long run. This is true not only in terms of the outright costs of the data itself, but also through reducing errors in investment decisions, with greater efficiency in returns being the end result.
- **Manageability:** Managing an asset involves more than just making choices about when/how much of it to buy and sell; it also entails controlling the risk-return profile of that asset *in the context of the overall portfolio*. Higher data quality allows an investor to better manage that profile on a relative basis (e.g., higher-quality data can permit better identification of relationships, so an investor can make changes elsewhere in the portfolio that alter the relative risk-return profile of the asset without making any changes to their position in the asset itself, thereby making the asset more “manageable”).¹⁰

⁹ Some elements of DG focus on augmenting data by marrying a datasets with other data that provides additional ‘context’, and thus enables more situated understanding. Likewise, DG that seeks to improve the granularity of datasets can lead to discovery of more micro-level relationships, which leads to refined understanding of market/returns phenomena.

¹⁰ Another example of benefits of DG to manage-ability can come via having better visibility into the impact of decisions made by external management teams (whether those teams are executives at companies in the portfolio, or general partners at externally managed funds). This visibility can enhance communication with those teams, and such communications are another way in which an asset can be ‘managed’.



Some of the impacts of data quality on returns quality might not appear immediately, but instead will manifest over longer horizons.¹¹ Therefore, **choosing to improve data governance is mostly about opting for higher long-term performance.** The ways in which DG increases long-term performance are generally different from the ways in which DM affects long-term performance. Based on what we've learned from top-performing investors, it's almost always best for organizations to cultivate distinct DG and DM strategies so that they function better together. Thus, the central message of this piece is that **every investment organization needs a well-articulated and well-tailored DG strategy.**

Admittedly, the boundaries between DG and DM aren't universally crisp, and there's often some overlap between DG and DM, in part because any performant approach to DM must take account of the organization's approach to DG, and vice versa. Another challenge is that many investors lack a DG strategy that fits their specific needs and aims or, worse, lack any DG strategy whatsoever, which leads to another key point: Data governance isn't one-size-fits-all. To work well (as we'll discuss below), **DG must align with the technology, culture, organizational structure, leadership and other resources an investor uses when it makes investment decisions.** Any misalignment with these resources generally reduces the quality of the investor's data, and hence the quality of its returns.

Why are ill-fitting DG strategies so common? Our belief is that it's due to the excessive focus that the investment industry — or, perhaps more accurately, the technology companies and consultants who support the industry — has placed on DM relative to DG. This imbalanced focus has largely come down to incentives. Historically, selling DM solutions to investors has been an easier, juicier business than solving their DG problems, in part because DG best practices for investors have remained insufficiently understood.¹² This paper aims to illuminate some of these best practices.

¹¹ It's worth remarking that the true costliness of poor DG is generally only fully realized under abnormal conditions - for example, a financial crisis, or other situations in which an organization's demands on its data (in terms of scope, quality, etc.) change suddenly (e.g., during the COVID-19 crisis, when many investors had to ask themselves questions they didn't routinely ask - and turned to their data to do so). Otherwise, poor DG is often more of a quiet drag on the organization - one that limits upside performance, increases costs, and poses a day-to-day trial of people's patience.

¹² Recently, various technology companies (both established and startup) have begun marketing singular DG "solutions". These tools may meaningfully improve an investment organization's DG capabilities, but almost never will any one of them *solve* DG, in the sense of letting technology substitute for well-crafted policies and processes. As we've been at pains to mention earlier, performant DG draws across resource boundaries, and relies on technology, but also people, culture, and operating procedures.



4. Context

Here, we cover a few concepts that will help readers more deeply understand our findings and subsequent discussion.

First is the *primacy of data* in investment decisions. It's not uncommon for many experienced asset managers, especially traders, to make decisions based on their "intuitive" understanding of market behavior, and it can be difficult, if not impossible, for them to completely articulate their mental models. This can make it very challenging to audit, and thereby improve, their decisions. This difficulty can be partly resolved by compelling managers, whether internal or external, to point to specific data points that drive them to make particular decisions. This tactic is analogous to a best practice in health care: In evidence-based medicine, doctors and nurses are compelled to cite specific studies or particular past cases that support the treatments they choose for a given patient.

Such data-focused practices support the *insight pyramid* (see Monk and Rook [2020]), whereby data underpins information, which underpins knowledge, which ultimately underpins intelligence (i.e., the differentiated understanding that is the foundation of investment outperformance). By deconstructing investment hypotheses and choices into the specific data elements that support them, investors can more objectively rationalize and validate their thinking. However, this entire process fails if the data quality is poor: the pyramid crumbles. **Solid data governance is, therefore, the bedrock of intelligence in investing.**

Another vital concept is the *quality of governance* itself. Basically, any form of (good) governance is a system of formal *processes* (e.g., due diligence protocols), *structures* (e.g., investment committees) and *tools* (e.g., investment memos; see McEvilley et al. [2023]). To work properly, these components of governance must be compatible not only with each other, but also with the organization overall (e.g., its culture, its employee headcount and skill sets, the complexity of its portfolio and the commitment and suitability of its leadership). Clark and Urwin (2008) correctly observed that governance is a scarce resource in investment organizations since the amount of governance that an organization can "do" and still fulfill its goals is limited. In aggregate, this all means that investors must be judicious when designing their governance systems and must strive



to ensure that such systems fit with their unique identity (see Monk and Rook [2023]). This need for tailoring is especially true for data governance (in ways elaborated below).¹³

A related concept is the *scope of governance*, which refers to the things an organization aims to govern in the first place. There's some tricky nuance here. For example, to be effective at DG, an investment organization must govern more than just its data. It must also govern things that modify and make use of its data, such as financial models and data-processing tools, whether these are merely spreadsheets or sophisticated machine-learning pipelines. Many investors (wrongly) don't treat such things within the scope of their DG strategies, which eliminates a crucial link between data quality and decision quality (since, nowadays, multiple tools and models usually sit between data and decisions).

Another resource that many investors often wrongly exclude from DG is their assumptions (e.g., on interest rates, inflation, growth rates for specific industries/asset classes, default probabilities or any other meaningful, forward-looking expectation that drives performance and risk predictions). Put bluntly, **few investors effectively/efficiently govern their assumptions; they lack robust processes for formulating them, refreshing them and harmonizing them across their organization.** While it's true that not every tool, model and assumption needs to be formally governed, any performant DG system should be *explicit* about its own scope to avoid things that ought to be governed going ungoverned, and vice versa (i.e., not wasting resources on governing things that don't deserve to be governed). Above all, top-notch data governance (and, indeed, all good governance) is about striking a balance and being clear about how that balance is struck, which is an exercise that must be continual (not "set-and-forget"), as we explain below.

5. Approach

The insights from this paper (and, chiefly, the best practices identified in the next section) draw on intense observations of several dozen institutional investors (e.g., public pension funds, sovereign wealth funds and endowments) since 2016. These observations have taken the form of research interviews, in-depth case studies and attendance at closed-door gatherings/meetings (to which we were given privileged access). The institutional investors we've studied are mostly large (most have between \$USD 10 billion and \$USD 1 trillion in assets under management), and they're based

¹³ The consequences of doing too little governance (e.g., by giving it inadequate structure or resources or by implementing overly lax policies) are various and vary in severity. They include 1) a lack of confidence in data across the organization (thus seeding tendencies toward low-confidence decisions or overreliance on "gut feeling" decision-making); 2) painful and costly reconciliation processes; 3) mistakes in decision-making (which sometimes remain undetected in the absence of proper DG); and 4) hard limits on how effective data management can be.



across a wide range of geographies (principally, North America, Europe, Asia, the Middle East and Oceania).

Given the large volume and diversity of this research pool, our findings are focused on synthesis (i.e., drawing on commonalities rather than looking at individual cases or circumstances of single funds). This focus on synthesis also follows the research practices advised by Clark (1998) and Clark and Urwin (2008), under which confidentiality is preserved by not disclosing the names or other identifying information of research subjects (due to the fact that much of our research involved accessing privileged information, and a usual condition for such access was anonymity and non-disclosure).¹⁴

6. Findings

Here, we concentrate on existing and emerging best practices in DG, rather than describing what's commonplace across the investment industry, because 1) most investors are doing DG poorly or not at all, and 2) those who are (even somewhat) successful at doing DG agree that building a solid DG strategy is well worth it, which is an opinion that's backed by their returns performance.

Best-practice DG possesses three hallmarks: It's ever-evolving, it's people-centric and it's properly resourced. We treat each of these successively. To start, **good DG is never outright "solved", in the sense that it can reach a point where its policies, resources, etc. no longer need to be updated to remain effective.** Markets, data, technology and investment organizations themselves are always evolving; to fulfill its purpose, an investor's approach to DG must do likewise.¹⁵ This fact may dishearten some, but there's a flipside; **ever-changing requirements for DG pose a continuing opportunity for investors to gain (or extend) an advantage over their competition.** This possibility should be particularly attractive to smaller investment organizations (in terms of headcount) that are able to more readily adapt their governance systems (or may have more degrees of freedom in doing so) relative to their larger counterparts.

Moreover, there's also the prospect of *adaptive* DG, whereby the system itself doesn't require retooling in response to changes (e.g., moving into new asset classes or onboarding significant

¹⁴ Some readers may wonder why our research pool excludes hedge funds, a group that is commonly seen as leaders in the financial services industry with respect to their approaches to data quality and decision-making. The reason for this exclusion is comparability: most investors are not hedge funds, and the differences between their structures, resources and objectives differ enough from those of hedge funds to merit focusing on a more representative group of investors: asset owner institutions.

¹⁵ More explicitly, most investment organizations are continually adding new elements (even if slowly and incrementally) to their technological architectures, portfolios and ambitions. Therefore, a once-and-for-all, set-and-forget DG system is doomed to failure for all except the most static and unambitious of investors (who can hardly be considered 'successful').



new datasets/tools); but instead, DG policies are set such that they dictate how the system will be reorganized to accommodate these changes — one can think of this as being something like “meta governance.”¹⁶ The key idea is that **ideal DG should be forward-looking and should reflect not how an organization functions now, but how it hopes to operate in the near-to-mid-term future.**

Perhaps more than anything else, top-flight DG is geared around an organization’s people. This idea has many facets, but a major one is cultural; **best-practice DG should aim to influence an organization’s culture, creating a mindset that “all data is an organizational asset” and, like other assets and resources in the organization, it shouldn’t be hoarded, abused or wasted.** Like other assets and resources, using governed data (and data tools, models and assumptions) comes with a set of rights and responsibilities, which we’ll discuss below. First, it’s imperative to note that there’s a tradeoff between effectiveness and onerousness regarding data responsibilities.¹⁷ Pointedly, people will resist and shirk governance if it hinders their ability to perform other tasks, namely, those considered core to their roles.^{18,19}

The designation of who is responsible for ensuring the quality of which datasets and what quality standards are applicable is a matter of *data stewardship*, which is the practice of assigning **specific duties** for the upkeep of **specific datasets** to **specific people** in an organization. As is likely obvious, the key here is *specificity*: Who is responsible, and what those responsibilities are, must be made explicit to be effective. Most successful DG systems clearly define the following:

- **A process for assigning stewardship responsibilities.** In most cases, every unit of governed data is assigned to a particular individual or team. Ideally, this process should account for the fact that fulfilling stewardship responsibilities will consume some amount of time and should, therefore, incentivize the steward or adjust their non-DG responsibilities accordingly.²⁰

¹⁶ We don’t have sufficient space here for a detailed treatment of adaptive DG; that’s a subject for a future ARB or other piece of research. For the present, interested readers might consult Clark and Urwin (2008) for a discussion of adaptive governance in general.

¹⁷ From here on, for the sake of brevity, interpret “data” to mean not only data but also tools, models and assumptions that are governed under a DG system.

¹⁸ A pearl of advice that one sovereign wealth fund stressed to us is the need to move at a measured pace when implementing new DG measures. To paraphrase the anecdote they conveyed to us: showing folks the value in caring for a specific dataset in a new way takes time; if you ask too much too quickly, people will just find workarounds to evade governance and the slow-downs they believe it brings - even when the new measures speed up the organization overall.

¹⁹ One way to mitigate this resistance is to add participation in DG as an activity on which employees’ performance is assessed, with appropriate incentives attached. Technology can also lessen the burden placed on people. Tools that (semi-)automatically improve data quality are multiplying and improving, and we expect these to play an upsized role in best-practice DG in the future. However, for now, robust policies (rather than automation) remain the centerpiece of good DG.

²⁰ Many DG systems adopt a “primary user” ethic on stewardship, whereby the team or individual that is the most frequent (or first) user of a dataset becomes the default steward. Norms are different, however, for “centralized” DG systems, which we discuss later in this paper.



- **A process for prescribing the responsibilities a steward has for a particular dataset.** In most cases, this will entail stipulating a standard for data quality that the steward has a duty to maintain (e.g., keeping the dataset refreshed and free of certain types of errors, reconcilable with some reference dataset). However, in some instances, other responsibilities may apply, for example, tracking who in the organization is using the dataset and policing its proper use. A good DG system will make clear for any user (and for any given dataset) what responsibilities the steward has and what “buyer beware” caveats may exist (e.g., where the applicable standards might be lax, and the onus for quality control is more on the user than the steward). Moreover, there are times when processes that stipulate how standards (and therefore the steward’s responsibilities) need to change, for example, when users and use cases for the governed dataset change. (This applies most often when the dataset shifts from being used solely by one team to being consumed by multiple teams.)
- **A process for monitoring and enforcing the quality standards prescribed for a given (governed) dataset.** Sometimes, key datasets are formally quality-checked (either by people or technology) at regular intervals.²¹ Other times, no scheduled checks are implemented, and the responsibility for secondary quality checks (i.e., other than those by the steward) falls on data users themselves. Most good DG systems have some mechanism to force compliance with standards whenever a dataset deviates from its designated quality standard.

Data stewardship is a lynchpin of performant DG systems, and improving stewardship policies can be a big “unlock” not only for data quality, but also for the quality of data-driven decision-making. However, several necessary (although not by themselves sufficient) conditions must be in place to allow any stewardship to be effective and efficient. These include:

- **The ability of data users to identify stewards and standards for particular datasets.** For smaller organizations with limited data needs, this might be facilitated by way of a shared spreadsheet. Larger, more data-sophisticated investors will almost always need a more advanced tool for this purpose (e.g., some institutional investors have platforms that serve as “data directories”).
- **A common language for speaking about data quality, responsibilities and other elements of DG.** It’s not enough for data specialists to be versed in DG terminology; anyone who uses

²¹ The rigorousness and frequency of such checks should be dictated jointly by the effort involved and the importance of the data.



governed data should be fluent. This can sometimes be achieved in part by eliminating unneeded jargon in DG documentation, tools, etc., in favor of plain-English descriptions.²²

- **A need for buy-in from senior leadership.** It's common for stewardship to trickle down the organizational structure, so that stewardship activities are mostly performed by junior employees. These activities will usually flounder (and data quality along with them) whenever senior people don't value — and insist upon — compliance. Senior leadership should be obligated to “sow demand” for DG by continually interrogating and championing the quality of all data that enters their decisions.

However, the impact of senior leadership on the success of DG transcends stewardship and calls for more than just leadership's approval of DG. There's also a need for senior leaders to invest their time, which is typically the scarcest of their individual resources, on particular DG workflows. Some of these are infrequent. For example, one DG best practice is the design of a *data hierarchy*, which identifies successive levels of data quality (e.g., “gold,” “silver,” “bronze,” and “raw”) and stipulates what purposes each level can be used for (e.g., trading, long-term planning, risk management and rapid ideation).²³ Any specific dataset can then be mapped to a particular quality level, which helps clarify what types of decisions it can inform.

Designing such a hierarchy is usually not something that must be done repeatedly and so might involve a concentrated, but not ongoing, time commitment from senior leaders. Other DG workflows, however, can pose regular claims on senior leaders' time, such as serving on DG committees or working groups. Many institutional investors who excel at DG have such committees and groups that meet regularly to oversee and assess DG initiatives (e.g., they may make determinations on stewardship and standards for new datasets, resolve internal disputes related to data quality, be involved in diligence exercises on new data technologies or participate in a myriad of other DG-related activities). **For larger organizations, having one or more committees and working groups dedicated to DG should be considered a best practice.**²⁴

Committees, working groups and stewards are all important ingredients in any successful DG system, but for these ingredients to work properly (both separately and collectively), concerted decisions are needed regarding what type of DG “model” best fits the organization. There are two

²² Some organizations that we've studied go a step farther, with (what some call) *data glossaries*, which are essentially formal descriptions of what various datasets 'mean' for the organization, in terms of how they should be interpreted at the organizational level. Such interpretations are often stipulated within the context of certain models, strategies or frameworks for risk management.

²³ Many investors, especially institutional investors, have increased needs for data “experimentation”; that is, they need the ability to quickly test and validate new ideas with data, possibly using new datasets. How to fit such experimentation into DG setups is an open question for many investors, because simply opting to not govern experimentation seems a poor choice.

²⁴ Some institutional investors also find it beneficial to have DG expertise/officials on their investment and risk committees.



main types of model — centralized and federated — which are distinguished by where the main *authority* for DG policy-setting and enforcement resides. In centralized models, the authority is concentrated in a single team (usually a data/IT team), while in federated models, the authority is distributed across the business so that different teams have some autonomy in setting and enforcing DG policies (e.g., they have some leeway in deciding what data they'll govern, as well as how it should be governed).

In practice, most organizations use a hybrid model that blends facets of centralized and federated approaches. For example, a core DG team might issue high-level guidance, and other teams are free to make their own decisions within the boundaries of that guidance, subject to the approval of the DG team or a relevant DG committee. Under such hybrid systems, it's important to achieve the appropriate balance between coordinated control, which comes more readily from centralization, and localized efficiency, which comes more easily through federation. A vital point to appreciate here is that authority for policy-setting and authority for enforcement aren't identical; a DG system can be centralized in its policy-setting but federated in terms of policy enforcement (i.e., responsibility for ensuring that policies are followed is distributed across the organization).

Yet, no matter whether a given system leans more toward centralization or federation, it's vital to realize that doing DG well hinges on suitable awareness of (and properly limiting) the degrees of freedom that exist in how data is accessed, maintained and used across the organization. Part of that awareness (and the ability to act on it) will inevitably be driven by what DM strategy and tech stack an investor has. We've already mentioned that, for DG to work well, it must fit with an investor's DM approach and technology architecture, but this street is bidirectional; sometimes, an organization's tech and DM pose unacceptable limitations on its DG ambitions, because even the best-designed DG policies and incentives can fail if they sit atop poor technology or DM policies. Hence, being better at DG may mean allocating more resources to improve technology and DM in the organization.

The foregoing is by no means an exhaustive specification of best practices in DG; instead, it covers what we understand to be the fundamental best practices that an investor needs to enable other best practices (which we'll be covering in upcoming work). However, even if an investor pursues only these fundamental practices, it's likely that they'll reap serious rewards in terms of the long-term quality of their returns.



7. The ARB-itrage

Data governance is likely to become increasingly vital in the future, given the extreme speed with which (generative) artificial intelligence (AI) is progressing. Many investment organizations are looking to empower their teams with AI tools, but there's enormous risk in doing so without proper data governance in place.

One of the chief value propositions for investors in using AI is to make better use of their internal data (e.g., extracting less-obvious insights from old pitch decks and investment memos or finding nuanced relationships that hide in disparate databases). However, AI tools are not flawless, and investors will — for the foreseeable future — need to audit AI outputs in terms of being able to trace those outputs back to the underlying data that informed them. This traceability will be helped by sound DG. DG can also help in improving the quality of AI outputs by driving higher-quality data inputs.²⁵

AI proficiency isn't the only novel activity that can be aided by better DG. A solid DG strategy can also be a huge boost to portfolio nimbleness, specifically the capacity to move quickly into emerging asset classes. Whenever an investor chooses to invest in a new asset class (and do so prudently), it must also acquire and manage new datasets and develop processes for converting these data into investment decisions. The basic components of a good DG system can be speedily applied to new assets and datasets (via "cut-and-paste"), which means that decision quality for new asset classes is generally easier to ensure if an investor already has a well-functioning approach to DG. This ability to become quick and confident early movers can be a source of massive outperformance for investors.

Additionally (but not finally, as there are many other capabilities that DG can enable; we simply lack space to cover them all here!), DG is also a strong enabler of knowledge management. Advanced knowledge management has been recognized as a reliable path toward long-term outperformance and one that all investors should consider (see Rook and Monk [2018] and van Gelderen and Monk [2015; 2019]).

²⁵ Even if an investor doesn't plan to use AI internally, it will likely need to make better use of its data (and have higher quality data) to effectively compete with those who do.



8. Coda

Good data governance (DG) is key to improving the quality of an investor’s data, decisions and long-term returns. Most investors, however, lack a cogent DG strategy and fail to appreciate the crucial differences between DG and data management. Therefore, savvy investors who are able to implement DG best practices can build serious competitive advantages, as long as they ensure that their chosen approach to DG aligns with their resources and objectives. Getting that alignment right, however, is rarely a simple matter. The best practices identified in this ARB are a prudent starting point, but there is far more that investors can do in optimizing DG. We’ll be detailing more of these optimization steps in coming publications.

9. Compass questions

- Which dimensions of returns quality does your organization prioritize/value the most (e.g., stability, comprehensibility, efficiency or manageability)?
- What fraction of your organization’s data is “governed” — in the sense of each dataset having identifiable quality standards, specific users tasked with ensuring that quality and some articulation of the types of decisions the data can be used for?
- How visible are data-quality responsibilities in your organization? That is, is it easy to tell who “owns” a given dataset?
- On the spectrum of federated to centralized, where would your organization’s ideal system for data governance sit?

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. [1998] "Stylized Facts and Close Dialogue: Methodology in Economic Geography," *Annals of the Association of American Geographers*, 88(1).

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

Davenport, T., and D. Patil [2012] "Data Scientist: Sexiest Job of the 21st Century," *Harvard Business Review*, 90(10).

McEvelley, B., A. Monk, and D. Rook [2023] "Investment Memos and Decision-Making," *Addepar Research Brief*.

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 [2020] *The Technologized Investor: Innovation through Reorientation*. Stanford University Press.

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 [2018] "Managing Knowledge Management: Towards an Operating System for Institutional Investment," *Social Science Research Network*, available at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3277989.

van Gelderen, E., and A. Monk [2015] "Knowledge Management in Asset Management," *Social Science Research Network*, available at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2642467.

van Gelderen, E., and A. Monk [2019] "Knowledge Management for Investors: Filling Knowledge Gaps via Networking and External Partnerships," *Social Science Research Network*, available at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3331643.



Authors

Dane Rook, Head of Research at Addepar (dane.rook@addepar.com)

Ashby Monk, Strategic Advisor at Addepar (ashby.monk@addepar.com)



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.