

How predictive models can lose the plot – and how to keep them on track

Vern L. Glaser

Associate Professor, Department of Strategy, Entrepreneurship and Management,
University of Alberta — Alberta School of Business

vglaser@ualberta.ca

Omid Omidvar

Associate Professor, Department of Organisation and Work, University of Warwick —
Warwick Business School

omid.omidvar@wbs.ac.uk

Mehdi Safavi

Senior Lecturer in Strategy and Organisation, Department of Strategy, Cranfield University
— Cranfield School of Management

mehdi.safavi@cranfield.ac.uk

Author Accepted Manuscript for publication in MIT Sloan Management Review (Magazine)

Organizations are increasingly turning to sophisticated data analytics algorithms to support real-time decision-making in dynamic environments. However, these organizational efforts often fail—sometimes with spectacular consequences.

In 2018, real-estate marketplace Zillow launched Zillow Offers, an “instant buyer” arm of the business that leveraged a proprietary algorithm called Zestimate, which calculated the estimated sales prices of real estate. Based on these calculations, Zillow Offers planned to purchase, renovate, and resell properties for a profit.¹ While it had some success for the first few years, the model failed to adjust to the new dynamics of a more volatile market in 2021. Zillow lost an average of \$25,000 on every home they sold in the fourth quarter of 2021—resulting in a write-down of \$881 million.² This is an instance of what we call *algorithmic inertia*: when organizations use algorithmic models to take environmental changes into account, but fail to keep pace with those changes. Here, we explain algorithmic inertia, identify its sources, and suggest practices organizations can implement to overcome it.

A Credit-Rating Catastrophe

To understand the phenomenon of algorithmic inertia, we conducted an in-depth study of an organization that failed to respond to changes in the environment: Moody's, a financial research firm that provides credit ratings for bonds and complex financial instruments such as Residential Mortgage-Backed Securities (RMBSs). These securities aggregated bundles of individual mortgages into distinct tranches with unique characteristics during the period leading up to the Global Financial Crisis of 2008.³

Moody's made a concerted effort to account for environmental changes in their credit ratings by developing a proprietary algorithmic model in 2000 called M3 Prime. The model analyzed data about properties, mortgage holders, and the economy to estimate two parameters central to calculating a credit rating: expected losses for the mortgage pool and the loss coverage protection required for a security to maintain a AAA rating. An analyst would present a recommendation to Moody's credit rating committee, which assigned a publicly posted rating for the security. Moody's monitored these ratings and upgraded or downgraded RMBSs as the environment changed. The M3 model achieved early success, and Moody's expanded its scope of algorithmic analysis by introducing a derivative model, M3 Subprime, in 2006.

Between 2000 and 2008, Moody's provided credit ratings for thousands of RMBSs, but ended up downgrading 83% of AAA-rated mortgage-backed securities valued at billions of dollars by 2008. The U.S. government, along with 18 states and the District of Columbia, held Moody's responsible for the role its inflated ratings of these and other products played in precipitating the financial crisis, and in 2017 the agency agreed to pay \$864 million to settle the allegations.⁴

This is a particularly illustrative example of algorithmic inertia with devastating societal consequences. Moody's provided an excellent context for exploring algorithmic inertia because the organization was explicitly responsible for analyzing environmental changes as part of its core service. Moreover, we were able to access detailed information about its algorithmic model from a report produced by the Financial Crisis Inquiry Commission, including extensive interviews conducted under oath with Moody's executives who were involved in the business at the time.⁵

Our analysis enabled us to identify the most significant contributing factors to algorithmic inertia: buried assumptions, superficial remodeling, simulation of the unknown future, and specialized compartmentalization. (See Figure 1, "4 Sources of Algorithmic Inertia.")

Figure 1: 4 Sources of Algorithmic Inertia

Caption: Organizational and process issues can result in algorithms that don't keep up with environmental changes in the decision domain.

Buried Assumptions	The organization uses legacy data inputs for the algorithmic model despite recognizing significant changes in the environment.
Superficial Remodeling	The organization makes only minor modifications to the algorithmic model in response to substantive changes in the environment.
Simulation of the Unknown Future	The organization overconfidently relies on the algorithmic model to predict the future environment.
Specialized Compartmentalization	Responsibilities for the algorithmic routine are divided between team members in distinct roles based on their specialized expertise.

Buried assumptions — failing to revisit fundamental assumptions undergirding inputs of the algorithmic model in light of changes in the environment — contributes significantly to algorithmic inertia. For example, loan originators were increasingly underwriting mortgages based on lower credit standards and substantially less documentation. So, an original assumption undergirding Moody's M3 model — that the technology mortgage originators were using to streamline the loan application process was also enabling a more accurate assessment of underlying risks — wasn't modified in accordance with the changing lending environment. The managing director of credit policy at Moody's told a federal inquiry that he sat on a high-level structured credit committee which would be expected to deal with issues such as declining mortgage underwriting standards, but the topic was never raised. "We talked about everything but ... the elephant sitting on the table."⁶

Moody's model also assumed that consumers' FICO credit scores were the primary predictive factor in loan defaults. But the quality of this data input significantly diminished over time: As the use of these credit scores became increasingly common, individuals found ways to artificially inflate them. As a result, low- and no-documentation mortgages carried more latent risk that was not being taken into account in Moody's algorithmic model.⁷

Superficial remodeling occurs when organizations make only minor modifications to the algorithmic model in response to substantive changes in the environment. At Moody's, some major changes to the environment included a growing number of loan originators, increasingly low-quality mortgages, and an unprecedented decline in interest rates.

Moody's response to these changes was to seek to capture more business in the rapidly growing market, so it fine-tuned the model to be more efficient, more profitable, cheaper, and more versatile, according to its chief credit officer⁸ — not to be more accurate. When they modified the M3 model to introduce the M3 Subprime model, they extrapolated loss curves for subprime loans based on premium loans, rather than developing fresh loss curves for subprime loans.

Simulation of the unknown future, or relying on an algorithmic model to produce viable scenarios for the future environment, can also leave organizations vulnerable to algorithmic inertia. Moody's constructed a simulation engine featuring 1,250 macroeconomic scenarios that enabled them to estimate possible future losses based on variations in economic markers such as inflation, unemployment, and house prices.

However, the simulation engine was limited by its underlying structure and assumptions, so analysts did not consider that changes were occurring, didn't update scenarios, and failed to accurately represent the changing macroeconomic environment. Based on the belief that detailed performance histories could more precisely reveal causal links between economic stresses and loan behavior, Moody's used estimates based on historical parameters rather than expected pool loss distributions to examine behavior in stress scenarios.⁹

Specialized compartmentalization arises when experts in different domains are involved in an algorithm's design and use, and there is no overarching single ownership or shared understanding of the model. At Moody's, responsibilities for the credit rating routine were

divided between the domain experts (i.e., credit rating committee members) who used the quantitative model and the quantitative analysts who developed it.

Because ownership and use of the model were distributed, and Moody's didn't strictly define how to use it, credit rating committee members created ad hoc rules to adjust the results of the model when outputs didn't conform to what their expert judgment led them to believe it ought to produce. Model outputs weren't considered final; rather the models were seen as tools to be used in conjunction with other approaches, and there was much divergence in how ratings committees made their determinations.

The models were developed and modified by individuals distant from the domain in which they would be applied; disparate groups of domain experts then used the models in inconsistent ways without understanding their underlying logic. The managing director for rating RMBSs described the model as so technically complex that few people understood how it worked.¹⁰ This issue is at the heart of what makes algorithmic inertia hard to tackle: The modeling and algorithms are often so complex that domain experts can hardly grasp the details of their functioning, while data scientists are disconnected from how models are being used in the real world.

Develop Practices That Combat Algorithmic Inertia

We see above how each of the causes of algorithmic inertia played out in Moody's use of an algorithmic model to dynamically incorporate changes in the environment into their credit ratings. Despite recognizing flaws in the model and making active attempts to change it, the organization was unable to effectively adapt to the environment, thereby substantially contributing to the Great Financial Crisis.

In order to prevent similar degradation of critical algorithms' predictive value, we suggest that organizations can implement four practices, described below. (Also see Figure 2, "Keeping Algorithms Relevant.")

Expose Data and Assumptions. Organizations should articulate and document the data used in their algorithmic models, including data sources and fundamental assumptions undergirding their data selection decisions, which can have deleterious effects.¹¹ Models often include operationalizations of many concepts, and it is easy for organizations to lose track of these parameters, which can be buried in layers of software code. Parameters representing the environment need to be documented to ensure they remain visible. Similarly, the fundamental assumptions undergirding the model should be articulated and periodically revisited.

Figure2 : Keeping Algorithms Relevant

Caption: Organizations can increase the likelihood that algorithmic tools will effectively adapt to changes in the environment by following four practices.

Expose Data and Assumptions	Articulate and document the data and the fundamental assumptions in the algorithmic model.
Periodically Redesign Algorithmic Routines	Periodically redesign—and be willing to overhaul—the algorithmic model and how the model fits into broader organizational routines.
Assume the Model Will Break	Actively consider scenarios outside the scope of algorithmic models by challenging predictive assumptions, assuming that the model will be fundamentally flawed, and envisioning qualitative alternative predictions of the future.
Build Bridges Between Data Scientists and Domain Experts	Ensure data scientists and domain experts work closely together to design the algorithmic model and the organizational routine in which it is embedded.

Moody's used a dataset on premium mortgages to train their model that was used to rate RMBSs composed of subprime mortgages. Initially, this might have been a reasonable choice due to the availability of data. But when a model's initial dataset isn't refreshed, algorithmic inertia can result. As the Moody's case suggests, data is never completely accurate, objective, and flawless. Therefore, making the sources of data and assumptions about those sources transparent to algorithm users, and continually reflecting on the appropriateness of that data, are critical practices for organizations seeking to avoid algorithmic inertia.

Organizations must keep data sources clearly organized and evaluate them periodically. Different data sources have different qualities and characteristics. Ensuring that these sources are distinguished before they are fed into algorithmic models, processed, and constantly compared against each other enables data scientists to identify and eliminate algorithmic inertia sooner rather than later.

The assumptions underpinning the use of an algorithm should also be documented and articulated. Any attempt to model the environment involves quantification — transforming aspects of reality into numerical data. Such quantification inevitably involves making assumptions about how the environment works. However, while quantification is necessary for algorithmic models to work, details about how it is done can get lost in the complex process of designing and using algorithmic models. Therefore, maintaining a living record of such assumptions may prevent the emergence of algorithmic inertia.

Periodically Redesign Algorithmic Routines. Organizations should regularly redesign—and be willing to overhaul—the algorithmic model and how it fits into broader organizational routines. The initial design of an algorithmic model can take a lot of work, and it is natural for an organization to want to reap the benefits of that work. However, in a dynamic and quickly

changing environment, it's important to be willing to not just make incremental changes to a model, but to fundamentally overhaul it if necessary.

Of course, organizations face a trade-off when it comes to overhauling an algorithmic routine: it can be very expensive to completely re-architect an algorithmic model. However, the consequences of failing to do so can be disastrous.

For example, when Moody's had to rate an increasing number of subprime-dominated RMBSs, the organization chose to incrementally modify the M3 model. However, it may have been more effective to specify the distinctions between the prime and the subprime markets and do a deeper overhaul of the original model. In addition to re-thinking the algorithmic model itself, an organization can consider how it is deployed in practice: hypothetically, Moody's could have applied the M3 prime model differently to different types of RMBSs—perhaps simply requiring more human intervention for tranches composed of lower-quality loans.

Redesigning and overhauling an algorithmic model is contingent upon understanding what organizational processes interrelate with the model and analyzing the implications that changes in the environment have for it. If it becomes clear that either the model or the processes that it relies on or feeds into have been rendered obsolete or ineffective, an overhaul should be seriously considered.

Assume the Model Will Break. It can be dangerous for an organization to think of potential future scenarios only through the prism of what algorithmic models predict: all assumptions embedded in a model limit the potential futures that can be considered. To address algorithmic inertia associated with the simulation of an unknown future, it is important to assume that the model will break. Consider scenarios beyond the scope of the algorithmic model; doing so requires challenging predictive assumptions and operating from the presumption that the model is fundamentally flawed.

An active practice of considering scenarios that are *outside* the model can help motivate and inspire the prior two practices—exposing data and assumptions, and periodically redesigning algorithmic routines—by forcing team members to actively consider the limitations of algorithmic models.

One particularly useful approach might be to make use of qualitative predictions of the future, instead of quantitative predictions that rely on available data from the past. These forms of scenario planning offer opportunities to consider radically different visions of what the future may hold. This might also entail developing hybrid algorithms that do not precisely rely on past data to predict scenarios but also embed in them qualitative measures and expert rules introduced by domain experts.

Build Bridges Between Data Scientists and Domain Experts. Organizations must create processes where data scientists and domain experts work closely together to design their algorithmic routines.¹² Practically speaking, data scientists and AI specialists approach

problems very differently than domain experts. Domain experts focus on organizational routines and idiosyncratic situations, while data scientists focus on developing generalizable constructs based on mathematical principles. To overcome algorithmic inertia, data scientists and domain experts must work closely together to understand how characteristics of organizational routines and idiosyncratic situations map to the mathematical parameters used in an algorithmic model.

When the worlds of data scientists and domain experts are completely separate, there is also the danger that data scientists and domain experts shift responsibility by superficially trusting each other's work. Such assumptions can actually prevent crucial dialogue between the two worlds. For instance, Moody's ended up subverting the results of their M3 credit rating model because the credit rating analysts didn't attempt to understand why the model might be generating results that didn't fit with their intuitions. Building bridges between data scientists and domain experts enables domain experts to obtain an intuitive grasp of how the algorithmic model works. Such common ground could enable them to create and use models that better adapt to changes in the environment.

One structural bridge-building practice that organizations can use to facilitate communications between data scientists and domain experts is establishing a position such as a product manager. This should be held by one individual with both domain experience and data science experience who has direct responsibility for overseeing algorithmic routines. For example, some have called for a new organizational structure by adding a position such as an "innovation marshal" respected by both data scientists and field experts¹³. These people, with the knowledge and expertise in both areas, can gain the respect of the organization by developing and maintaining "high-bandwidth, bidirectional communication channels" that help ensure that algorithmic routines are able to adapt to environmental changes.

Another bridge-building practice is called *model explainability*: describing the algorithmic models in a practical and comprehensible manner. For data scientists, such explainability can facilitate access to the expert knowledge needed to counteract the sources of algorithmic inertia; for domain experts, such explainability can help them develop a deep and intuitive understanding of how the model takes environmental changes into account. As such, model explainability establishes common ground between two groups of professionals who have different types of expertise. Such practices enable organizations to build bridges instead of just talking about them.

Organizations seeking to reap the benefits of powerful predictive analytics are increasingly confronted with the problem of algorithmic inertia. Despite leveraging dynamic algorithms to adapt to changes in the environment, organizations may find that results are not keeping pace with new developments. By exposing data and assumptions, periodically redesigning algorithmic routines, and assuming their models will break, and building bridges, organizations can increase the likelihood that their substantial investments in algorithmic solutions will pay off with better decision-making.

References

- ¹ Stokel-Walker, C. 2021. Why Zillow Couldn't Make Algorithmic House Pricing Work. *Wired*. <https://www.wired.com/story/zillow-ibuyer-real-estate/>.
- ² Parker, W. 2022, February 11. Zillow's Shuttered Home-Flipping Business Lost \$881 Million in 2021. *Wall Street Journal*. <https://www.wsj.com/articles/zillows-shuttered-home-flipping-business-lost-881-million-in-2021-11644529656>.
- ³ Omidvar, O., Safavi, M., & Glaser, V. L. forthcoming. Algorithmic Routines and Dynamic Inertia: How Organizations Avoid Adapting to Changes in the Environment. *Journal of Management Studies*. <https://doi.org/10.1111/joms.12819>.
- ⁴ <https://www.justice.gov/opa/pr/justice-department-and-state-partners-secure-nearly-864-million-settlement-moody-s-arising>
- ⁵ United States Financial Crisis Inquiry Commission. (2011) *The financial crisis inquiry report: final report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. [Washington, DC: Financial Crisis Inquiry Commission: For sale by the Supt. of Docs., U.S. G.P.O.] [Web.] Retrieved from the Library of Congress, <https://lccn.loc.gov/2011381760>.
- ⁶ The Financial Crisis Inquiry Report
- ⁷ Rona-Tas, A., & Hiss, S. 2010. The Role of Ratings in the Subprime Mortgage Crisis: The Art of Corporate and the Science of Consumer Rating. In M. Lounsbury & P. M. Hirsch (Eds.), *Markets on Trial: The Economic Sociology of the U.S. Financial Crisis: Part A*, vol. 30A: 115–155. Emerald Group Publishing Limited.
- ⁸ The Financial Crisis Inquiry Report
- ⁹ Moody's Investor Services, Moody's Mortgage Metrics: A Model Analysis of Residential Mortgage Pools, 2003. Retrieved from [https://fcic-static.law.stanford.edu/cdn_media/fcic-docs/2003-04-01%20Moody%27s%20Mortgage%20Metrics%20-%20A%20Model%20Analysis%20of%20Residential%20Mortgage%20Pools%20\(Moody%27s%20Special%20Report\).pdf](https://fcic-static.law.stanford.edu/cdn_media/fcic-docs/2003-04-01%20Moody%27s%20Mortgage%20Metrics%20-%20A%20Model%20Analysis%20of%20Residential%20Mortgage%20Pools%20(Moody%27s%20Special%20Report).pdf)
- ¹⁰ The Financial Crisis Inquiry Report
- ¹¹ Lindebaum, D., Glaser, V. L., Moser, C., & Ashraf, M. 2022. When Algorithms Rule, Values Can Wither. *MIT Sloan Management Review*, 64(2): 66–69.
- ¹² Glaser, V. L. 2017. Design Performances: How Organizations Inscribe Artifacts to Change Routines. *Academy of Management Journal*, 60(6): 2126–2154.
- ¹³ Hoerl, R. W., Kuonen, D., & Redman, T. C. 2020. To Succeed With Data Science, First Build the 'Bridge.' *MIT Sloan Management Review*, October 22, 2020. <https://sloanreview.mit.edu/article/to-succeed-with-data-science-first-build-the-bridge/>.