



WHITE PAPER: WHY BUSINESSES SHOULD CARE ABOUT DATA QUALITY

The impact, assessment, and mitigation via manual interventions

Introduction

Coming off the struggles of 2020, 2021 is almost certainly going to see significant economic activity in the second half of the year (<https://www.barrons.com/articles/why-2021-may-at-last-be-the-year-of-the-great-rotation-51606870661>). This means that there will be significant opportunities in many markets as we see new investment (e.g., relief funding from the government) or shifting market dynamics (e.g., possibly permanent changes in the hospitality and tourism industries). As your company monitors markets that you sell into, it's likely that your company will need to be agile to change with the times and take advantage of new opportunities. One key way that successful companies have done this in recent years is through strategic use of their data (<https://hbr.org/2020/01/when-data-creates-competitive-advantage>).

This white paper explores this approach and what is needed at a fundamental level to make it happen. Metamor Systems has used data and technology to help businesses and have many "battle scars" to help you avoid common pitfalls.

Benefits

While many executives may know the benefits of data-driven strategies, it's helpful to set context and review some key benefits to this approach. Regardless of the business you're in, data-driven decision making will invariably allow your organization to: identify new products or markets, find patterns that are not obvious, and lower costs (just to name a few: <https://online.hbs.edu/blog/post/data-driven-decision-making>). Going further, focusing on data strategies enables new technologies (e.g., machine learning and AI) as well as implementing data science in your organization (<https://towardsdatascience.com/recipe-for-a-data-science-project-69ad5c5ecf29>). As the last article outlines, executives can't just simply tell their managers to "make it so" without building the fundamentals. Fortunately, the foundation of all these start with one thing: focusing on good data.

Therefore, the balance of this paper will focus on that aspect - data quality. Understandably, there's more to it, such as hiring the right data engineers. However, forming the right team won't help if the data they work with is of poor quality. Finally, using poor quality data in data-driven decisions will identify the wrong conclusions; and, ultimately, the wrong decisions will be made. So let's dive into ensuring that the building blocks - the data - is ready to be used appropriately.

Assessing Data Quality

Unfortunately in this fast-paced world, knowing the details of an organization's data is often left to lower-level employees who often don't have a strong voice to communicate issues up through the management ranks. In fact, Harvard has identified that most organizations have

challenges with their data (<https://hbr.org/2017/09/only-3-of-companies-data-meets-basic-quality-standards>). There are definitely some quick and easy ways to spot check for data issues, such as this exercise from Harvard Business Review: <https://hbr.org/2016/07/assess-whether-you-have-a-data-quality-problem>. However, to implement machine learning, AI, or any level of data science, a more comprehensive approach is needed. This is called data profiling and it is an area where Metamor Systems can help (<https://dzone.com/articles/what-is-data-profiling>). Our analysts can quickly review huge amounts of data to identify issues with both standard (e.g., zip codes of customers) and business-rule driven data (e.g., product 'x' can only be paired with product 'y'). While too numerous to mention in this paper, we'll go through referential integrity checks and also review key data sets that are used throughout your organization. Through this process, Metamor Systems can provide a summary of issues identified along with a roadmap of ways to resolve them.

Manual Mitigations

Very simply put, resolving data quality problems HAS to happen at the source (<https://hbr.org/2020/02/to-improve-data-quality-start-at-the-source>). If a "fix" is put in place only for reporting (or a data warehouse or other analytics), then issues will continually be a game of "whack a mole" and will never go away. It is understandable that there may be some technical or cultural barriers to resolving data issues at the source. However, it is the only way to ensure that all down-stream systems will be fixed. That is, most source data will be used for multiple reasons. Some of them are: Operations reporting, financial and budget analysis, sales and strategy decisions, new product/service decisions. Within all of these, there may be a myriad of technologies that ingest the data. For example, financial and budgeting might be done on a spreadsheet while operational reporting may be displayed on a dashboard via digital signage. No matter what the technology, the data is the same and coming from the same source (even if it is transformed in different ways).

Given that data quality issues are often deeply embedded, often cross teams (e.g., the application development team is often very different from the analytics team), it is most helpful to bring in a "neutral third party" to work through solutions to the problem. Additionally, everyone still has their "regular job" to do and often won't find time to work on data improvement initiatives. Metamor Systems "speak the language" of application and data people to bridge the gap. We are uniquely qualified to trackdown and resolve data quality problems at the source. As mentioned above, we can identify a roadmap of improvements so that the existing teams can still get their normal work completed in addition to gaining the benefits of improved data quality.

For more information and a free consultation regarding how Metamor Systems can help improve your data quality - leading to improved competitiveness and new technologies - please contact us at: info@metamorsys.com.



12821 East New Market Street, STE #310

Carmel, IN 46032

Phone: (317) 514-6172

Fax: (317) 534-3176

www.metamorsys.com

