



WHITE PAPER

**DELIVERING ON ESG WITH
TECHNOLOGY, THROUGH
TECHNOLOGY**

Author: **Alex Weinrich**

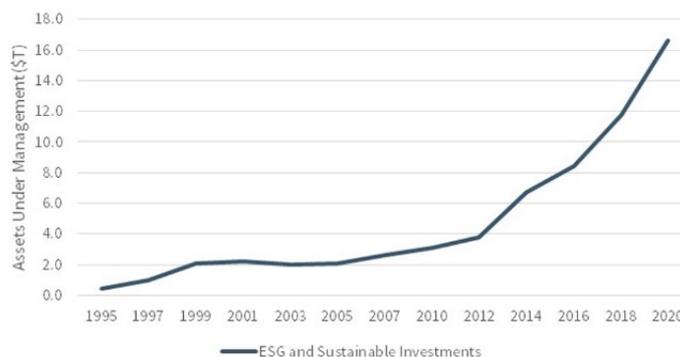
Delivering on ESG with Technology, through Technology

By Alex Weinrich

ESG as a Business Imperative

Environmental, Social and Governance (ESG) has become a dominant issue in the financial world. It has fueled a rapidly growing market sector as illustrated in the graph below, where, because of regulation and social pressure, financial institutions are being held accountable for their ESG actions and principles. As an example, a buy-side firm does not want to be associated with “green washing” and the [reputational damage](#)* it can cause. Similarly, banks need to ensure their issuing and lending practices are in line with their ESG principles. To do so they must process an increasing range of disparate ESG data to confirm they are being regulatory compliant and upholding their stated ESG policy. A 2020 survey by SustainAbility reflects this trend where ESG ratings were cited as the most frequently referenced source of information that institutional investors rely on to gauge ESG performance – ranking at 55% and tying with direct company engagement.

ESG Mutual Funds: Assets Under Management (1995-2020)



Sources: Google Trends, “ESG,” (2004 to July 2022); U.S. SIF Foundation, “Sustainable and Impact Investing—Money Managers,” (2020).

As ESG data increases in volume and becomes more readily available, market participants recognise that their commercial and investing decisions must be similarly linked to it to ensure compliance and probity of operation. Completeness, accuracy and timeliness of ESG data is as important as any other data source. What happens between reporting periods can cause considerable damage to an Institution if they are not diligent about having the most current data and analysis on the measures that are critical to them. That poses a data challenge.

The ESG Data Challenge

ESG data and its integration are still relatively new, with a lack of standardization that poses challenges for institutions seeking to create a single source of truth across departments and regions. Unlike some of their other business processes, there are few digital solutions that can ingest and analyze fragmented data sets to create a reliable “truth” and common language to base investment decisions and risk assessments on. Social media channels and more informed investors are forcing

*login required

financial institutions to stay current with ESG trends and the data to back up their policies and practices. Further complicating the challenge for organisations is that different departments will have differing requirements:

- **Trading** – requires data in real-time to understand mandates, current scores and other ESG metrics when taking on risk or rebalancing portfolios.
- **Quantitative** – It is essential data is complete and accurate as models and algorithms are developed and back tested.
- **Loan departments** – Writing loans and the bank's accountability with their ESG mandates requires data to be current.
- **Senior Management** – current reports are needed to ensure ESG protocols are being followed.

This leads organisations to ask some fundamental question of their systems:

- 1) On the ability to ingest the right data sets from the necessary ESG sources to create reliable policies. channels, streaming and batch - can your systems support multiple, changing data feeds, over multiple protocols?

Capturing and integrating ESG data from its many sources presents significant challenges. They may come in different formats, over different protocols and in streams, files or batches. The need for (and cost of) storing its historical data means it may reside across multiple databases. Some providers use Natural Language Processing to scrape internet and social media channels to form opinions or at least social impressions on a company or sector.

In addition to there being multiple sources of ESG data and scores, a further complication is that they have different methodologies. Within these sources, the range of input parameters can vary significantly. FTSE Russell claims its model uses 300 indicators; Refinitiv uses 630 ESG metrics; S&P Global uses 1,000 underlying data points. As a result, there is a consensus that multiple providers are needed to form aggregate metrics that reconcile the differences. Managing this number of variables in non-standardized data sets can be a monumental task. These challenges combine when attempting to analyze different sources and timeframes concurrently as most modern databases struggle to combine historical, real-time and streaming data together. The right database architecture that can deal with fragmented- data coming from multiple sources in historical and real-time series is imperative

- 2) On the ability to establish a single source of truth for ESG data- can your systems capture, process, validate and store streaming data that can be normalised and made easily available across the enterprise?

With all this complexity, it is little wonder that institutions struggle to have a single source of truth with ESG data. Ideally, data would come in one format, from one source, in an efficient well-defined database structure. It doesn't. That is the struggle data engineers face daily in trying to present relevant stakeholders with access to the ESG data they need when they require it.

A first step may be assessing which data is material, which can entail significant modelling and quant work. Another consideration is how to deal with missing data (which, incidentally, provides another reason for using multiple providers). Some providers may make changes their methodology over time that must be taken into account while still preserving the integrity of previous measures (similarly to managing stock splits from Corporate Actions). Compounding these issues will be how to add weightings to this potential output and being able to support the reasoning behind it. This demands analytics and testing capabilities to model, pre-process and validate data being consumed and published to users.

3) On the ability to create real-time reports based on the data across your organization - can your users easily access accurate, validated data for scheduled and adhoc reporting to regulators, investors and throughout the enterprise?

Many financial institutions still struggle with legacy systems and the data silos they spawn. The data may be sitting in various databases across different departments, some may be in datastores, some in files, some in internal systems while others may be streaming in via Kafka or similar channels. Access may vary too. Some may be accessible via SQL, some via Python, others via BI tools. These variations pose a significant challenge for users creating reports across the full range of data sets.

Integrating the different data sources, formats and programming languages in order to create reports on, say, the “ESG-ness” of a portfolio or rating an ESG bond based on an institution’s rank poses significant problems for data engineers and solution architects. Beyond reporting, analysis can prove especially difficult, with joins across streaming and historical dates being particularly challenging for determining correlations and trend analysis. The problem can be exacerbated as new sources are added and sometimes requiring expensive and risky rewrite of existing functionality.

The ESG Processing Challenge

There are a number of areas in which firms can act more consciously in their daily operations. They need to be aware if what they are building is carbon efficient, if the technology stacks demand excessive electricity usage, if devices being disposed of properly and not prematurely? These are examples of evaluating Green Computing.

Carbon intensity changes depending on your location and how much renewable energy is creating the electricity it takes to power your task. Are there processes you can run when carbon intensity is at its lowest? One solution is to shift workloads to times and locations where there is more supply of renewable energy. This is called “demand shifting”. An alternative approach is “demand shaping” where, rather than moving location, the consumption of energy is aligned to when it is in most supply. So, an additional question organisations must ask is if they are being “ESG conscious” about its data management and processing workflows? Are they being performed efficiently and with minimum energy footprint?

The Part Technology Plays

KX has long been at the forefront of financial innovation, particularly in trading, risk management and surveillance providing technology for real-time analysis of any data, at speed or at rest, providing firms with actionable insights. KX Insights is its integrated data management and streaming analytics platform that provides that ability. Advanced analytics, with seamless Python integration and SQL querying of data, allows users to detect anomalies and derive insights – all in real-time.

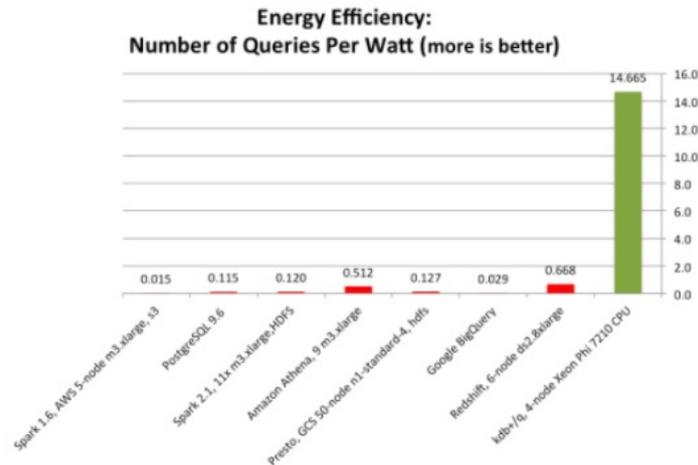


But beyond data management alone KX is committed to practicing greater sustainable software engineering and lowering energy usage. Its kdb+ database platform has been designed from the start to be extremely efficient by bringing compute to the data (rather than move data around), lowering the ‘energy’ needed in storage and processing. Kdb+ sets records for speed at performing complex analytics due to its vector-based algorithms, which have been optimized to make best use of the hardware they run on. As a result, kdb+ makes fewer demands on the hardware and requires less electricity. This lowers its total cost of ownership, while earning high marks for green computing.

So, Watts Different about KX?

Choosing efficient software, especially for the most demanding computing problems, is another effective, yet often overlooked way to manage energy consumption. But how do you compare the energy efficiency of different database platforms? To help answer that we developed a simple method, based on recent benchmarks run on our software by tech blogger Mark Litwintschik who published an article on where he tested kdb+ on a set of standard queries: [1.1 Billion Taxi Rides on kdb+/q & 4 Xeon Phi CPUs](#). Mark’s article focused on the time taken to compute the results of several queries, where each query had to scan each and every one of the taxi ride records. We also looked the power consumption it involved.

Not only did KX come out with the fastest query times by far of any Intel platform, when we then looked at the underlying energy required to calculate these results, we discovered we were doing this whilst at the same time we were the most energy efficient solution by far. We completed these queries using an average of 14.67 queries per watt of energy, with each query operating against the full set of taxi ride records. The closest competitor to this came in with a result of 0.67 queries per watt of energy. In other words we used a fraction (35%) of the power they needed to perform a query AND at the same time completed it in a fraction (12%) of the time!



These results are based on averaging all four query times in the tests that Mark Litwintchik ran on the taxi ride dataset from the New York City Taxi and Limousine Commission. Energy consumption for this exercise was calculated using maximum power for each of the server platforms, which included CPU, GPU, memory and all ancillary equipment. So these are true apples-to-apples comparisons, using real-world numbers.

ESG, A Corporate Duty

Companies need to start treating these workflows as if they were running a household. “Turn lights and other machinery off when not in use!” “Run the washing machine, when electricity prices are at their lowest or ideally when it is being generated by renewable energy!”. Public statements like the following from Blackrock reflect the importance of committed ESG principles

Blackrock quotes “Our investment conviction is that climate risk is investment risk, and that integrating climate and sustainability considerations into investment processes can help investors build more resilient portfolios and achieve better long-term, risk-adjusted returns.”