IIAC ACCVM
Investment Industry Association of Canada Association canadienne du commerce des valeurs mobilières

Through a Glass Darkly

"The Death of 1000 Cuts" redux

The following article is by R. "Tee" Willams, Principal and Owner, Tee Williams Associates. A well-known consultant for various international exchanges, Tee has built industry recognition as an expert on market data operations and strategy, and in creating a capacity planning methodology for the Composite Tape Association for all U.S. listed securities, and the Options Price Reporting Authority for all U.S. options. His firm has specified the methodology used to allocate revenues among exchanges contributing data for consolidation by the U.S. Securities Information Processors (SIPs) under a mandate from the Securities and Exchange Commission (SEC). Tee also teaches on market data, market structure, and automated trading.

(edited – R. Tee Williams, August 30, 2023)

The title is an appropriate literary reference to the current state of the market data industry. More specifically, I used the reference as a sidehead on the future of market data in an article I wrote about twenty years ago titled “The Death of 1000 Cuts” . Looking back at the past is interesting because many predicted things were wrong, but a few were correct. Given the current developments and interest, this is the right time to revisit this topic.

By early 2003, I had become fascinated by how much the market data industry had changed since the late 70s. And so, I began to think about the things that had changed. I concluded that, particularly from a vendor's perspective, it was a much less attractive environment than had existed nearly 25 years before. At that point, many vendors were in serious trouble. Bridge and Telerate, major industry players then, were dead or dying. Other prominent vendors were a goulash of acquired products with little coherent vision.

Twenty years later, in 2023, there is good and bad news. The good news is that many vendors in the Bloomberg mould have expanded niche services into top-tier vendor offerings, not by throwing random stuff into the mix but by expanding on a coherent base. For example, both FactSet and Morningstar have grown from niches to significant vendors. In the 1990s, FactSet was one of their several analysis tools. It is now a primary vendor. These developments are justified and are good for the industry. However, other developments have led to inefficiencies, contributing to some major issues concerning market data related to cost and access.

Since 2003, the mainstream vendors have protected themselves with a strategy that relies on “coercive monopoly” practices. In the earlier paper, I referred to coercive monopoly as the “irritation defense”: vendors attempt to prevent customers from moving to a better or cheaper competitor not by improving products or lowering prices but by making the switch too complicated or time-consuming.

Exchanges in 2003 were protected even though they had stayed the same since inception, beginning with the Amsterdam Effectenburse in 1602. In the 1990s, the regional German stock exchanges created Deutsche Borsa. Deutsche Borsa, from its founding, is a publicly listed company. In 2000, Euronext, The Hong Kong Exchanges, Nasdaq, The New York Stock Exchange, The Tokyo Stock Exchange, and the Toronto Stock Exchange all chose public listing in a five-year period.

These exchanges transformed from “membership companies” to for-profit companies by publicly listing. This change had a profound impact on the cost of market data.

Exchanges, in their purest form, have only four primary revenue sources:

1. Trading fees (under intense competitive pressure);
2. Clearing revenues (often separated into different entities);
3. Listing fees (not available to many exchanges); and
4. Market data.

As membership companies, exchanges had simple goals: to provide a low-cost trading facility, to invest enough in technology to ensure continued operation in the most stressful conditions, and to ensure that quality information was available to members and their customers.

In stark contrast, for-profit exchanges are responsible for the “care and feeding” of investors. Data sales are the only source of exchange revenue that can consistently provide investor-grade income. Increasingly, exchanges have depended on market data as the dominant revenue source.

The insatiable consumer demand for more and faster data allowed exchanges to increase fees rapidly, depend on inconsistent and confusing licensing terms and definitions, and use audits to boost revenues. By the late 2010s, consumer dissatisfaction with high market-data costs in all major trading markets and the absence in Canada, Europe, and the United Kingdom of a consolidated tape resulted in organized demand for reform.

Certain exchanges may have made a strategic misstep, potentially driven by a focus on profit. This brings to mind a well-known saying in the trading markets, 'The Parable of the Pigs': 'Bulls make money, bears make money, and pigs get slaughtered.' This adage is a cautionary reminder of the risks of overreaching in financial pursuits.

The issue has caught the eye of the regulator and, most recently, the Canadian Securities Administrators (CSA). The CSA issued the paper in November of 2022 and sought public input on possible changes in the provision of Real-time Market Data (RTMD). There has been much interest in the topic, garnering twenty comments on the various aspects of the proposal. Given the complexity of the topic and the multitude of issues being covered, this paper will focus on the most important element: cost.

The Total Cost to Use Market Data

At the beginning of the project, it was observed that the cost of market data to the consumer is more than the fees exchanges charge. In financial accounting, the concept termed “the total cost of ownership” is not just the purchase price of an asset but all the attendant costs involved in owning the item. In the case of market data, data is not “owned”; it is “licensed for use.” This led to us agreeing on the term “Market Data Total Cost of Usage” (MDTCU).

There are three primary components when we dissect MDTCU:

1. Market data fees

While not the only cost, the fees exchanges and other data providers charge are significant and should be reduced. The IIAC proposed that Canadian market data fees be an explicit function of the cost to produce the data and the delivery cost. In each of the four markets we monitor — Canada, Europe, the UK, and the US — there is a current effort to tie fees to the cost of producing the data. The eligible costs and the definition of a reasonable markup will cause fierce debate. In addition to cost-based fees, the IIAC and most other respondents to the CSA paper requested that fees be transparent. However, we argue that posting fees on an exchange website is not sufficiently transparent. We believe that the prices reflected in invoices must be consistent as well.

Data licenses contain terms that establish how fees apply and often limit data usage. Current licensing procedures result in double or multiple charges for a single user to see the same data on different vendor

services. The term for this practice is: “Multiple Instances, Single User” or MISU. Exchanges charge multiple times for the same data.

Case 1: if a user has two separate display fees for the same data provided by different vendors, each vendor entitles data usage. Each vendor records the usage and charges for each “use.” Two fees for the same data.

Case 2: a trader is charged once for a visual display and again when the same data is used to create a dynamic blotter, which displays the value of trading positions updated with real-time prices.

The exchange asserts position values can be “reverse-engineered” to recreate the original instrument prices. We propose that a user — individual or application — is only charged once to access data.

A final way to reduce market data fees is to redefine 'delayed data.’ Delayed data is the term applied to data when the creator ceases to be able to assert an intellectual property interest in the data. A US court case established that an exchange has an intellectual property interest in the data produced from trading; however, the interest extends only for a fixed period. After that time elapses, the information enters the public domain. In 1905, the period was set to 15 minutes. Given today's market conditions, a 15-minute period is no longer relevant. (Any delay has no legal basis outside the US but is assessed by licenses.) Therefore, after a period where information is no longer “actionable” — a period likely measured in milliseconds — its only cost basis should be the delivery cost. Most retail investors and much of the institutional buy-side would be happy with data with reasonable delays. One measure to start negotiations could be the amount of the “speed bumps” used to protect “pegged” buy-side orders — a 350-microsecond delay.

2. Direct data costs

Second, there are direct data costs. Direct costs are ordinary business expenses, but streamlining data distribution and license administration will reduce these costs dramatically. Direct costs include charges by market data vendors for delivering the data, formatting it, and administrative fees to satisfy the demands of exchanges. Data consumers incur additional direct costs for acquiring data, frequently referred to as “onboarding,” storing the data, and maintaining it by continuously applying capital changes and updates referred to as corporate actions.

Standards for definitions, units of count, and the measures used in licenses, such as usage groups of individuals and applications, are the primary methods for determining fees. Beyond managing the data, licenses demand continuous reporting to vendors and content owners, responses to administrative messages, and periodic audits, ostensibly to ensure compliance with license terms.

3. Friction

Finally, the market data business is rife with what an economist calls “friction” — costs that produce no legitimate benefits to anyone. Market data distribution and licensing evolved from antiquated technology and business policies formulated in a much less global or integrated market environment. If you track market data through each step in the distribution process, you find that each message handoff usually

entails a change in message format and a new instrument identifier. Reformatting the same data item is friction.

How do we "fix" market data?

Paul Romer, the Stanford economist, once said: "A crisis is a terrible thing to waste." Likewise, focusing only on the mechanics of creating a consolidated tape or purely on the level of market data fees risks ignoring structural costs that we can reduce as we address the tape and fees. The industry needs to seize the opportunity to improve — to re-engineer — the market data process.

To address MDTCU, we chose to address each component cost explicitly. The issues outlined above that have caused a global crisis in market data lead some proposed solutions.

In the early 1990s, a particular individual working at Boston Consulting Group, a consulting company, authored a book that introduced a concept he labelled "Business Process Re-engineering" (BPR). The idea proved to be highly profitable in the consulting industry. While I made a superficial attempt to trace the origins of BPR through a search engine, I was hindered by numerous advertisements from firms offering BPR as a service and asserting that the concept originated with them. Putting the hype aside, resolving market data issues necessitates re-engineering the market data distribution process and the administrative structure.

To successfully re-engineer market data, we must focus on three major domains:

1. The selection of data to be included
2. The method of physical distribution
3. The underlying commercial process

Defining Market Data

"Market data" for our proposal is limited. I teach a course that identifies 18 categories of market data. The exact number of categories does not matter, but there are too many types of information loosely referred to as market data to address them all at once. Therefore, we should focus on data about transactions collected from the markets. Trading transaction data is the original information offered through vendors in the 1960s before most other types evolved. If we fix the problems for traded-instrument prices, consumer pressure will force other data types to adopt our solutions.

Specifically, our focus is quotes and their size, transaction prices, and the transaction's size. We include all information produced by officially recognized trading sources. We do not include dealer quotes or derived prices.

The Physical Distribution Process

After the market crash in 1987, the Group of Thirty, a consortium of "economic and financial leaders from the public and private sectors and academia," analyzed "Clearance and Settlement Systems in the World's Securities Markets" and recommended reducing the settlement cycle from the then-existing five business days to three by 1995 and then to one day ("next day settlement' or "T+1") in 2000. The three-day

transition in 1995 occurred without problems. The markets accelerated the steps in the five-day cycle, and regulators eliminated unnecessary requirements, such as sending notifications by mail. However, shortening the process to one day required re-engineering the entire settlement process. Each process step necessitated computer-to-computer message transfers between successive applications operating in different entities to complete each step in the settlement. Computer-to-computer message transfer has come to be known as “straight-through processing” (STP). In the US, the SEC recently mandated T+1 settlement for equities next year — 24 years after the proposed date.

To address the problems that make market data expensive, we require a similar re-engineering of the market-data distribution process. (Labeling this “straight-through data” is tempting, but the acronym is unfortunate.) Unlike STP, streamlining the market data process is more straightforward. Most elements needed to streamline the market data process already exist. We must adopt them.

The Commercial Process

Licensing is a part of the process that works together with other parts. In the world of market data, licenses can be confusing. They have a lot of words and rules that don't always make sense together. Some exchanges and providers take advantage of this confusion to make more money.

This confusion is more than just a problem for those buying the licenses. It also makes the companies selling them scared to try new things. They worry that any change might hurt their money-making. If we simplify this whole thing, it will cost less, make audits easier, and people won't be as scared of surprises.

The confusion also leads to problems with checking the rules (audits). Someone looking at these rules for over 10 years once said he had never seen an audit without problems. Some firms might intentionally break the rules, but that's not usual. Some companies that check the rules might look for problems to make money. I even know about a case where an exchange made up a problem to get money they didn't deserve. But mostly, the rules are so hard that even people trying to do the right thing mess up. If we make things simpler in how we share data and licensing rules, we might not need to check the rules as much, or maybe not at all.

Our proposal for market data in Canada

In summary, our proposal to the CSA for Canada is to reduce total market data costs, not just fees. A fundamental feature of our approach is that all market data prices be transparent and justified based on cost.

We next proposed an “Infrastructure” that would make creating a consolidated tape straightforward. The Infrastructure should have three parts describing the functions performed, not necessarily distinct entities.

1. Technical Information Processor: This term in Canada is similar to what's called a “Securities Information Processor” (SIP) in the U.S. and “Market Data Union” in Europe. The U.S. SIP model has worked well since the early 1980s. The only concern for Canada is that exchanges

own these processors in the U.S., keeping costs high. (The U.S. has two SIPs, run by Nasdaq and NYSE.)

2. Administrative Agent: This second part handles data from exchanges and other trading entities and shares this combined data with vendors and users. It will also control access to data (i.e., entitle the data) for both people and computer programs. This new idea solves the problem of a person being charged more than once for the same data by different sellers. The Administrative Agent will also tell vendors who should have access to any services that need this combined data.
3. Financial Agent: The third part of the Infrastructure takes care of billing everyone who gets data, collecting payments, and dividing the money among the places where trades happen.

For over 30 years, exchanges and vendors have discussed creating an independent group to handle these administrative and financial tasks. One suggestion from the FISD Best Practices committee is to create a "central utility" that can deal with all the paperwork that every exchange and vendor needs to handle. Right now, these duplicated efforts cost extra without providing any benefits. Some exchanges use vendors for these tasks, but the vendors add extra charges to the invoices. Our idea only covers combined market data but could also be used for other types of data. It is a real step towards creating that independent group.

A final note: the methods in this section depend on a series of standards and messages we refer to as “Tools.” I provide a list of Tools that will likely grow in the Appendix at the end of this document.

Ownership and competition

As our proposal evolves into discussion and implementation, Canadian market participants have significant questions to answer. Notably:

- How should ownership of the Infrastructure be structured?
- Should there be a single infrastructure or competing entities?
- Is the Infrastructure a for-profit entity or a regulated monopoly?
- What is the mechanism needed to drive the implementation?

Moving Forward

Canada must depend on what we describe as an “enlightened regulator.” If you imagine a regulator as an entity that sets standards such as electricity voltages, weights and measures, and the allocation of radio frequencies, we need that type of regulation. The goal is not to punish information providers for being profit-oriented. Instead, we are regulating the method and terms of distribution so that markets are fairer and more efficient. IIAC has a role to play where the industry and regulators can discuss proposed solutions and move forward. In this regard, the IIAC will host a series of webinars and a conference on this subject.

Desired Outcome

I do not expect to write the next installment of this two-paper series in 2043, but I hope someone will. I admit my bias, but it is useful to periodically pause and ask: "What are we doing, and does it make sense?" The consumer revolt against market data costs creates a perfect "moment in time" for reflection on the market data industry. My reflection suggests that what Canada is proposing to do is a rational approach. If adopted, we may see:

- The price of market data will be a simple function of the cost to produce it.
- The licenses that assert intellectual property rights in the data and control usage and distribution will be simplified using standard terminology. License administration will be simple and unambiguous.
- An exchange's intellectual property rights in market data will be tied to the length of time that the data remains actionable — certainly much less than a minute.
- Data distribution will use common message formats and standard instrument identifiers to flow from the point of execution to where it is used without repetitive message manipulation.
- Simplified license terminology and standardized terms mean that license administration will be easy to understand and apply.
- Audits may not be necessary, but to the extent they are, they should be straightforward.
- Finally, the cumulative effect of the abovementioned changes will reduce or remove most causes for friction in the market data process.

Our proposal in Canada is for a really limited set of what is broadly called market data. However, if what we propose for this limited data set works as we believe it will, our approach will quickly be extended to other types of data as well.

There will likely be casualties as we reform market data. A vendor's data collection, aggregation, and normalization roles will no longer be needed. That does not mean that vendors go away. It simply means vendors only provide high-value services, particularly developing new content.

As for exchanges, it is hard to argue that there are less differentiated services than matching buy and sell orders. Exchanges will remain critical, but it is hard to see how they can generate returns that satisfy investors. Exchanges will probably revert to utilities and need to operate as such.

Appendix – Necessary tools

The Canadian Infrastructure requires supporting features that we refer to as Tools to operate as we describe it. In this section, we will list the tools, describing them only briefly. We are developing a series of videos showing the mechanics of the tools, the consolidated tape, and the Infrastructure.

Unique identifiers

User ID — To use market data, every individual user and every application must have a unique identifier. While application identifiers should be universal, individual identifiers should be unique by country and independent of any linkage to personal data.

Instrument ID — we need to adopt, from available choices, a standard instrument identifier.

Standards

Usage groups — The industry can define a standard set of usage categories that will satisfy individuals and applications, reflecting both the intensity of demand and the ability to pay.

Units-of-count — A number of units of count will be required no matter how the market data administrative process is structured. However, reducing the set to distinct methods of delivery or unique applications will produce a limited number. The final set should be standardized and well-understood.

Central entitlement

By having the administrative agent of the infrastructure handle entitlements for users of data, we can correct the problems created by multiple entitlements for the same user and data source.

A reasonable limit to intellectual property rights

The concept of delayed data is simply a convention. The convention of 15 minutes should be changed. 15-minute-old market data has no purpose.

A tool to track data

There is a need for a tool that can track a data message as it moves along its distribution path. Such a tool would simplify data redistribution and reduce the problem of data theft.

A standard for administrative messages

We propose creating a standardized set of administrative messages to alert users to service changes. Users could program to handle the messages and automate the changes.