Tracking Proliferation through Trade Data
By Matthew Godsey and Valerie Lincy
January 2017

Introduction
The past decade has seen an expansion in the availability of trade data and a reduction in its cost. As a result, it is becoming a resource for researchers, including those seeking to track proliferation-related transfers. These data, which are derived directly from export declarations made to national customs agencies, are now available through commercial services that aggregate the information, deliver it in a searchable format, and provide enhancements to it. Trade data can take two forms: transactional data, which includes specific information about individual transactions; and statistical data, which is focused on commodity flows. Although commercial services provide trade data from official sources, there is a good deal of variation in the type and quantity of the information provided, as well as in the subscription price.

This paper examines four prominent trade data services as well as organizations that have used trade data to uncover proliferation-related transfers, sanctions violations, and other activities of concern.

Trade Data Providers

Datamyne
Founded in 1992, Datamyne describes itself as the largest searchable import/export database in the world.1 The company offers some form of trade data on over 50 countries and transactional-level data, including party names, for the countries listed below. The data is sourced from government customs agencies and bills of lading.2 Users can monitor shipments, research market trends, and search by buyers, sellers, and service providers. Datamyne receives updates from U.S. Customs daily and updates international trade data within two months.3

Matthew Godsey is a Senior Research Associate at the Wisconsin Project and Editor of the Risk Report Database. Valerie Lincy is Executive Director of the Wisconsin Project.

2 Ibid.
historical data is available back to 2004 for a negotiated price.\(^4\) Datamyne also acquired one of its primary competitors, Zepol, in 2015, allowing it to incorporate Zepol’s archived trade records and business intelligence tools. In December of 2016, Datamyne was acquired by Descartes Systems Group, a global software and logistics company that provides (via MK Data\(^5\)) extensive restricted and denied party screening tools for companies.\(^6\)

**Countries Covered:** Argentina, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, India, Nicaragua, Panama, Paraguay, Peru, Philippines, Russia, United States, and Uruguay.\(^7\)

**IHS/PIERS**

PIERS provides transactional-level U.S. trade data going back to 2003. The database is updated daily with information from U.S. Customs and includes export bills of lading and approximately 45 million company records. PIERS also provides trade data from about 12 other countries. For these countries, information is drawn from import and export bills of lading and includes related company names. The database features searchable data fields that include bill of lading number, vessel name/IMO code/voyage number, carrier line, consignee name and address, shipper name and address, ports traveled through, and weight/description of cargo.\(^8\)

**Countries Covered:** Argentina, Chile, China, Colombia, Costa Rica, Ecuador, India, Mexico, Panama, Paraguay, Peru, Uruguay, and the United States.

IHS also publishes a Global Trade Atlas (GTA) focused on commodity-level trade data to learn about broad trade patterns or specific product flows at the regional, national, or specific port level. This resource includes data fields such as Harmonized Tariff Schedule, trading country, merchandise value, and quantity that can be aggregated and sorted, as well as exported into dynamic files and used to generate data visualizations. It includes official monthly bilateral trade statistics for over 80 countries and annual data for some 200 countries.\(^9\)

IHS markets PIERS and the Global Trade Atlas as complementary products that may be used together.

---


Import Genius
Founded in 2006, Import Genius aggregates approximately 110 million ocean freight records to create its trade database. Data fields provided include bill of lading number, consignee name, shipper name, product/cargo description, cargo weight, port of origin, port of destination, arrival date, vessel IMO number, and voyage number. Import Genius updates its database with information from U.S. Customs within 5-7 days, and users can search records dating back within the past three months for the lowest priced option and back to November 2006 with the most advanced subscription.

Countries Covered: Argentina, Chile, Colombia, Costa Rica, Ecuador, India, Mexico, Panama, Paraguay, Peru, Russia, Ukraine, United States, Uruguay, and Venezuela.

Panjiva
Founded in 2006, Panjiva offers import and export data from the United States and 12 other countries. Data is sourced from government customs agencies. It is compiled from shipment information taken from bills of lading (fields include shipper name and address, consignee name and address, vessel name and transport company name, cargo description and weight, cargo HS code/description, cargo value, port of origin and destination, and port calls), and includes company profiles (contact information, aliases, revenue, employee count, and subsidiaries). Panjiva’s interface allows the user to search by buyer, supplier, product, or shipment and to examine historical trade records. Users can then identify trends, including a company’s or country’s primary export/import, its main trading partners, and the frequency of shipments. In total, Panjiva’s database incorporates approximately 600 million shipping records and approximately nine million company profiles.

For U.S. data, the database is updated daily and company profiles are current within one week. Most international data is reported to Panjiva with about a two month delay, and the database is updated monthly. The availability of historical information varies depending on

---

18 “Data Sources,” brochure provided to Wisconsin Project by Panjiva.
19 Ibid.
the subscription level, with the lowest price option permitting searches within the past year and the premium subscription allowing users to search U.S. data back to 2007.\(^\text{20}\)

**Countries Covered:** Brazil, China, Chile, Colombia, Costa Rica, Ecuador, Mexico, Panama, Paraguay, Peru, United States, Uruguay, and Venezuela.

**Reliability and Pricing**

The trade data providers described above receive their data from authoritative government sources. U.S. import and export data, for example, is sourced from Customs and Border Protection declarations, the Automated Manifest System (AMS), and import-export customs statistics.\(^\text{21}\) Similarly, foreign data is sourced from relevant customs authorities in each country. Data integrity and reliability is judged to be consistent across these providers. The additional value provided by these aggregators is in collecting data from multiple sources, collating it, and presenting it in a readily searchable, highly functional platform, and in allowing the subscriber to conduct searches tailored to a certain area of interest.

The price for trade data varies widely depending on the amount of information provided, the time period covered, the countries included, and the level of freedom the user has in accessing and manipulating the data. For example, at the lower end of the price range, a monthly subscription to a service from Import Genius allows the user to make individual queries of trade data from the last three months of U.S. imports, limited to 10 searches per day, for approximately $100. At the upper end, a subscription service from Datamyne costing more than $600 allows users to make unlimited queries of the last two years of U.S. import and export data and to export several thousand lines of data with each search. Adding additional country data adds to the cost.

Trade data providers may offer additional features, such as company profiles for individual exporters and importers. These profiles are created by collating millions of bills of lading, enabling the user to view the products a company deals in, the export markets it serves, and a summary of its full trading activity. It is also possible to purchase bulk trade data from these providers for integration into another system. The user can then combine this data with other sources, such as corporate registries, satellite and sensing data, vessel movement data, and entries from business-to-business websites, to conduct network or trend analysis. The cost of these bulk data purchases can be significant. For example, a bulk export of six months of U.S.


data can cost $25,000; exports of foreign country data can be even more expensive. Also, the terms and conditions attached to such purchases may limit how widely the data can be shared.

**Other Sources of Trade-related Data**

The four services listed above are among the most prominent trade data providers. Additional providers include firms like Tradessparq, WiserTrade, Infodrive India, and SICEX. They offer more limited information, in terms of countries covered or the quantity of transaction-specific data available.

In addition to fee-based commercial providers that aggregate trade data, it is also possible to access "raw" data directly from certain government websites, often free of charge. In Pakistan, for example, the import general manifests of all vessels arriving in port are available and searchable via a website for Pakistan’s Federal Board of Revenue. These documents contain the names of the consignors and consignees of each shipment, as well as a description of the goods, the vessel name, port of origin, and arrival date. Another site, PakTradeInfo, aggregates data from these documents and provides it in a searchable format. This service can be accessed for a modest subscription price, with the option to pay daily, monthly, or yearly.

In India, public tenders for procurement of items and technology for use in the military and nuclear programs are made available by the Defence Research and Development Organization (DRDO) and the Nuclear Power Corporation of India Limited (NPCIL), among others.

In order to efficiently make use of these datasets for research, or to use them for statistical analysis studies, it is often necessary to use data scrapers like Import.io, Data Miner, or Web Scraper to first extract data.

**The Use of Trade Data in Tracking Proliferation**

A number of arms control organizations and nonproliferation experts use commercially available trade data to enrich their research.

---

The Wisconsin Project on Nuclear Arms Control
The Wisconsin Project makes use of trade data in conducting research for the Risk Report, a database of nearly 5,000 entities of concern for proliferation. In building individual profiles of entities of concern, the Project reviews transaction-level information to identify a company’s suppliers and/or customers, the countries or regions in which it is active, what products or commodities it trades in, and what names and addresses it uses when conducting business. It is sometimes also possible to learn the contact person in a transaction, relevant phone numbers of parties to be notified, and even a detailed description of the goods being shipped, including contract numbers or the name of a project or site for which the cargo is destined. For trade data, the Project relies on Panjiva, the Indian and Pakistani resources described above, and has also used a resource called Trade Navigator.

In the course of conducting this type of research, the Wisconsin Project has uncovered instances of sanctions violations. For example, in 2009 the Wisconsin Project discovered that China Precision Machinery Import-Export Corporation (CPMIEC), which was then subject to a U.S. import ban, continued to make dozens of sales to the United States using a then-unknown subordinate company to avoid detection. Following this work, the new name variation was added to the Treasury Department’s Specially Designated Nationals (SDN) list, blocking further exports by this entity.

In another study, the Project used Pakistani import manifest data to identify shipments to companies known or suspected to be procurement fronts for Pakistan’s nuclear and missile programs.

The Center for Advanced Defense Studies (C4ADS)
C4ADS is a nonprofit organization that specializes in using data driven analytics and new technologies for security policy research. In August 2016, C4ADS collaborated with the Asan Institute for Policy Studies and used tools like Panjiva, Palantir, and the Windward ship-tracking service to produce a report on North Korea’s strategic trade networks called “In China’s Shadow: Exposing North Korean Overseas Networks.”

The authors were able to construct their own database of North Korean entities on Palantir’s Gotham network analysis platform by using ship names identified by both the United Nations

---

and NK News. They then compared these names with records in Equasis, a maritime database that maintains shipping registration records, to map out a network of companies, owners, and shareholders. The end result was a dataset of 147 ships, 167 individuals, and 248 companies, for a total of 562 entities. The authors then used their original dataset to single out a conglomerate, Liaoning Hongxiang Group, run by Ma Xiaohong, a Chinese businesswoman. Using records obtained from Panjiva, the authors discovered that one of Ma Xiaohong’s companies imports 99.9% of its goods from North Korea, and that North Korea also serves as the destination for 78% of the company’s total exports, amounting to over $500 million in trade. Further investigation of Panjiva records revealed that, in addition to the huge sums of money flowing into North Korea, the company has sent at least two substantial shipments of aluminum oxide, a dual-use good that can be used to prevent corrosion in gas centrifuges that are used for uranium enrichment. The study also discovered that several of Ma Xiaohong’s companies were established in cooperation with U.N.-sanctioned North Korean entities.

Following their work, the Treasury Department sanctioned Ma Xiaohong and her affiliates and the Department of Justice indicted Ma and other individuals linked to her companies.

King’s College – Project Alpha
Operating under the Centre for Science and Security Studies at King’s College London, Project Alpha focuses on researching illicit trade in support of the non-proliferation regime.

In a November 2016 study, “Pakistan’s Strategic Nuclear and Missile Industries: A Baseline Study for Nonproliferation Efforts,” Project Alpha used trade data, combined with information

---

29 Ibid.
from Pakistani academic publications, government memoranda, U.S. diplomatic cables, entity watch lists, and business-to-business websites, to analyze Pakistan’s procurement network for dual-use goods. They concluded that Pakistan’s nuclear and missile programs “rely on a network of at least 20 trading companies in mainland China, Hong Kong, Dubai, and Singapore” with China being the major supplier. Project Alpha was able to categorize these companies as either witting or unwitting by examining whether a front company acted as buyer for Pakistan’s strategic industries or if real names of the Pakistani organizations were used in trade records.

In a separate study, Project Alpha used trade data as a tool to analyze an Indian news agency’s allegation that Pakistan had funneled dual-use items from a Chinese company, Beijing Suntech Technology Co Ltd, to North Korea. The study confirmed through shipping records that Beijing Suntech had made multiple shipments of dual-use goods “with potential utility in the nuclear fuel cycle and WMD-related applications” to Galaxy Corporation PVT LTD, a likely front company for the Pakistan Atomic Energy Commission (PAEC). These items included “polypropylene filter cloth, ventilated plastic suits, motors for cooling towers, humidity chambers, seamless pipes and sheets, and mixing equipment.” Project Alpha was unable to confirm, however, if the goods were then transshipped to North Korea.

The Royal United Services Institute (RUSI)
Andrea Berger, former Deputy Director of Proliferation and Nuclear Policy at RUSI, relied on trade databases for a May 2016 article published in *Arms Control Today* titled, “From Paper to Practice: The Significance of New UN Sanctions on North Korea.” She used Panjiva to discover that China’s Liaoning Province hosted nearly 1,000 different companies that had conducted

---


official trade with North Korea. The fact was used to illustrate the extent of corporate structures set up by the North Korean government used to circumvent international sanctions.

Ms. Berger again cited Chinese trade data in a December 2016 article written for the website 38 North. In “Recycling the Playbook: UNSCR 2321 and Its Coal Caps,” she used records obtained from Panjiva to show that hundreds of companies in China’s Liaoning and Jilin Provinces rely on North Korean coal imports as a key part of their businesses.

**European Commission’s Joint Research Centre (JRC)**

The JRC is tasked with supporting European Commission policies through independent research and analysis, and has been investigating the uses for open source information for nonproliferation efforts since 2007. The JRC has noted that open source trade data could be used to create better export control policy and as an aid to enforcement in detecting cases of trade diversion.

Some of its work has focused on the challenges of using HS codes in proliferation-related research. HS codes were designed by the World Customs Organization as a way to categorize traded commodities, but the categories are often broad, and it can be difficult to connect a specific export with a specific HS code. As a result, the same HS code can sometimes denote either an export-controlled item or a non-controlled item. In the absence of a more detailed cargo description in the export declaration, there would be no way to tell for certain whether the item was of proliferation concern or not. This limits the ability to use trade data in analyzing commodity flows to identify potentially illicit transactions or instances of sanctions violation.

The JRC created The Big Table (TBT), a software tool that assists in trade analysis and correlating HS codes with potentially WMD-related goods. However, the tool is not publicly...
available: it was developed specifically for the IAEA’s Trade and Technology Analysis Unit, and is available to the licensing and customs authorities of EU Member States.  

**Research Group for Biological Arms Control**

The Research Group for Biological Arms Control is an organization operating under the University of Munich with a mission of conducting research and outreach to prevent the development and use of biological weapons. In a report recommending the amendment of the harmonized tariff system to enable improved monitoring of biotechnology trade, the Research Group noted that when certain biological items are connected to a specific HS code it is possible for analysts to use open source trade data to clearly map trade in those items and observe trends, without relying on licenses or other official sources of information. The study suggested linking more dual-use biological items to unique HS codes to improve this capacity to independently monitor biotechnology trade through commercially available trade data.

**Limitations of Trade Data in Tracking Proliferation**

These examples demonstrate the various roles that trade data can play in research on proliferation. However, there are some limitations to using trade data that bear noting. First, trade data is only available from certain countries, and even among these countries there are discrepancies in the data elements available for a given transaction. For example, for a shipment from China to Iran, the name and address of the company in China sending the goods would be provided. On the receiving end, no company name would be provided for the end user in Iran; only the “country of destination” would be supplied. By contrast, the record for a shipment from Peru to Ecuador would contain the names and identifier information for both the consignor and consignee, as well as a more detailed description of the cargo involved. This difference in data availability could make it difficult to use trade data to, for instance, make a balanced comparison of the procurement patterns of two countries involved in proliferation, since data may be more abundant for one than the other.

---


A second challenge is the use of only HS codes for many of the countries for which data is available. As the JRC points out in its study, HS codes do not overlap neatly with codes used for export controlled items, such as Export Control Classification Numbers (ECCN). Therefore, in many cases, knowing an item’s HS code might not be enough to determine whether it is controlled or not, since many HS codes denote a range of goods that might contain both dual-use and non-controlled items. For countries where a more detailed cargo description is available it may be possible to identify dual-use goods with more clarity; however, these descriptions are provided by the exporters themselves, so any company seeking to obfuscate the true nature of its products with a more vague description can easily do so.

Finally, the cost of trade data services can be considerable, especially for non-profit organizations or independent researchers. While a range of pricing options is available for most services, the least expensive subscription plans are usually limited to a subset of data, in many cases only from a single country or for a brief time period. Such plans also may carry restrictions on the number of queries that can be conducted each month or the number of data downloads available.