

Big Data & International Trade:

Creating Transparency through Information

P.G. Swartz

peter@panjiva.com

Panjiva, Inc.

www.panjiva.com

K.W. Soura

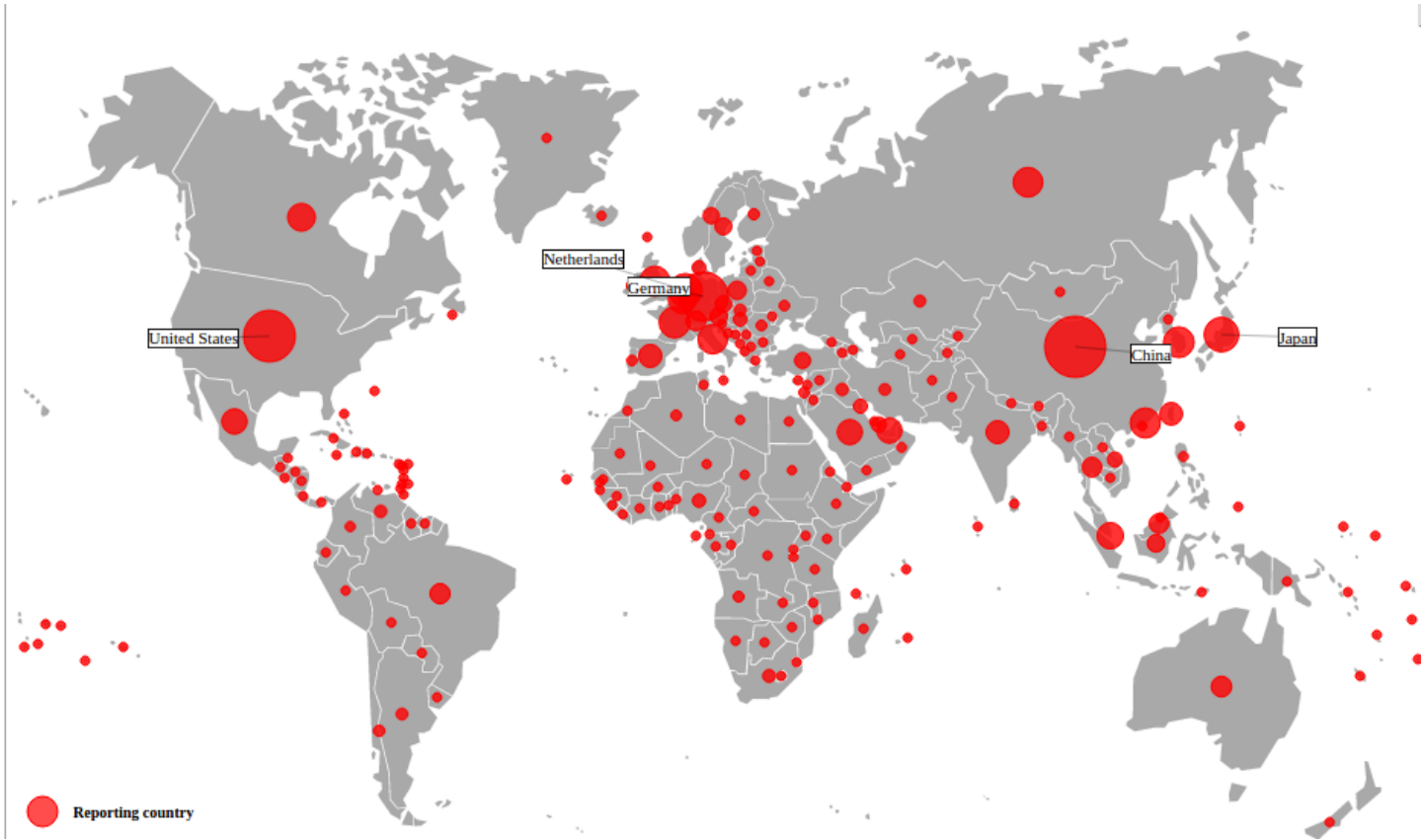
keith@panjiva.com

Overview

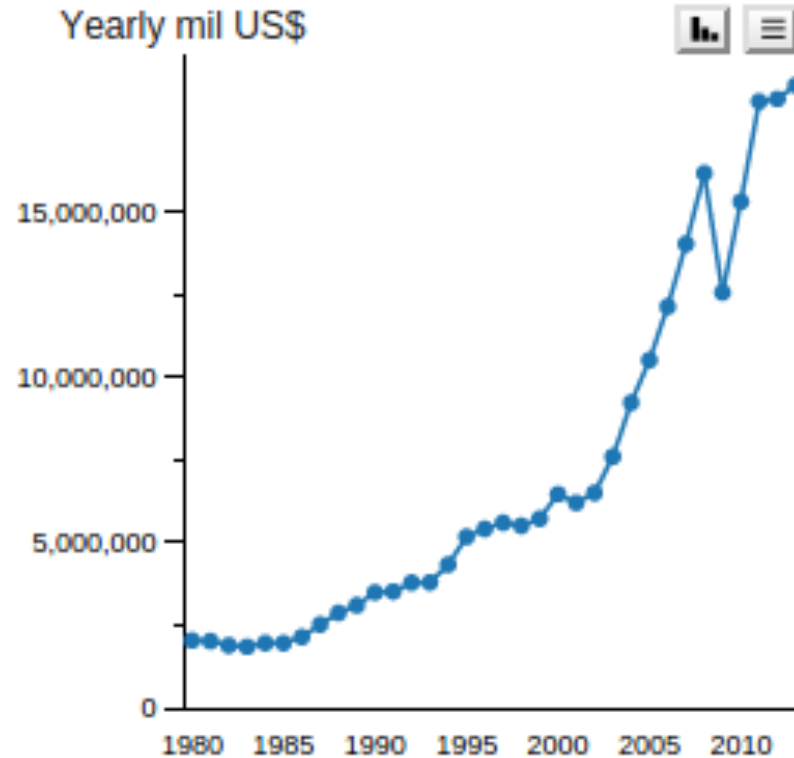
- Importance of global trade
- Promise of trade data
- Technical challenges and solutions
- Business applications
- Our take-aways

The Concept

Global Trade Touches All Countries



Global Trade Is Big And Growing



A Map of this Network is Very Useful

- Manufacturers/Retailers
- Logistics companies
- Governments
- Finance

The Opportunity

Valuable Information, if Extractable

- Hugely valuable information buried in a mountain of paperwork
- Complex system of privacy laws that impede acquisition or dissemination

Panjiva: Mapping the Trade Network

- 8 governments
 - US, China, Mexico, Chile, Colombia, Panama, Peru, Uruguay
- 190 countries
- 450 million shipping records
- 5 million companies

Showing 6 of 35 columns		Group by Shipments		Sort by Date (newest first)	
	Arrival Date	Raw Shipper	Raw Consignee	US Port Of Unlading	Container Info
					Commodity Description
🔍	2014-08-28	SAMSUNG ELECTRONICS CO., LTD 129, SAMSUNG-RO, YEONGTONG-GU SUWON 443-742 KYONGGI-DO SOUTH KOREA	SAMSUNG ELECTRONICS AMERICA, INC. AMERICAN DISTRIBUTION CENTER 4921 DIRECTORS PLACE SUITE 200 SAN DIEGO, CA 92121 TE858-642-5100	Port of Long Beach, Long Beach, California	REFRIGERATOR
🔍	2014-08-28	BEAUTIFUL WINDOW ENTERPRISE. CO., L NO.168,HU-FU-HUANG RD,LIUHE TAICANG JIANGSU CHINA	WESTERN REFLECTIONS,LLC 1385 CHEERS ST.SUITE BROWNSVILLE, TX 78521 BROWNSVILLE TX USA	Port of Long Beach, Long Beach, California	BLINDS

The Technical Challenge

Raw Data is Raw

- A range of data partners
 - Wide variety of data formats
 - Wide variety of regulations and rules
- Messy data:
 - Few unique identifiers
 - A pathological aversion to spell check
 - Records in multiple languages

Cleaning Raw Data

- Receive data
 - CSV, PDF, arcane formats
- Upload to raw data to database
 - Store efficiently
- Structure raw text
 - Geographic locations of parties
 - Ports
 - Units

Clustering Raw Data into Profiles

Notify Parties	Raw Shipper	Raw Consignee
ATLANTIC TRADING AND MARKETING INC 5847 SAN FELIPE,STE.2100 HOUSTON, TX 77057 US TOTAL PETROCHEMICALS & REFINING USA INC 1201 LOUISIANA STREET SUITE 1800 HOUSTON TX 77002 US	PMI COMERCIO INTERNATIONAL SA DE CV AV MARINA NACIONAL NO 329 TORRE EJECUTIVA COL.HUASTEC 11311 MEXICO CIUDAD, MEXICO CIUDAD 11311 MX	ATLANTIC TRADING AND MARKETING INC 5847 SAN FELIPE,STE.2100 HOUSTON, TX 77057 US

- Machine-learning based comparison of shipment records
- Intelligent feature extraction
- Matching to other data sets

The Technical Take Away

Generalize from Domain Expertise

- Data Science is great
- Combined with domain expertise: even better
- Research general ways of understanding the domain
- Example: address similarity
 - Semantic understanding
 - Intelligent robust features

The Platform

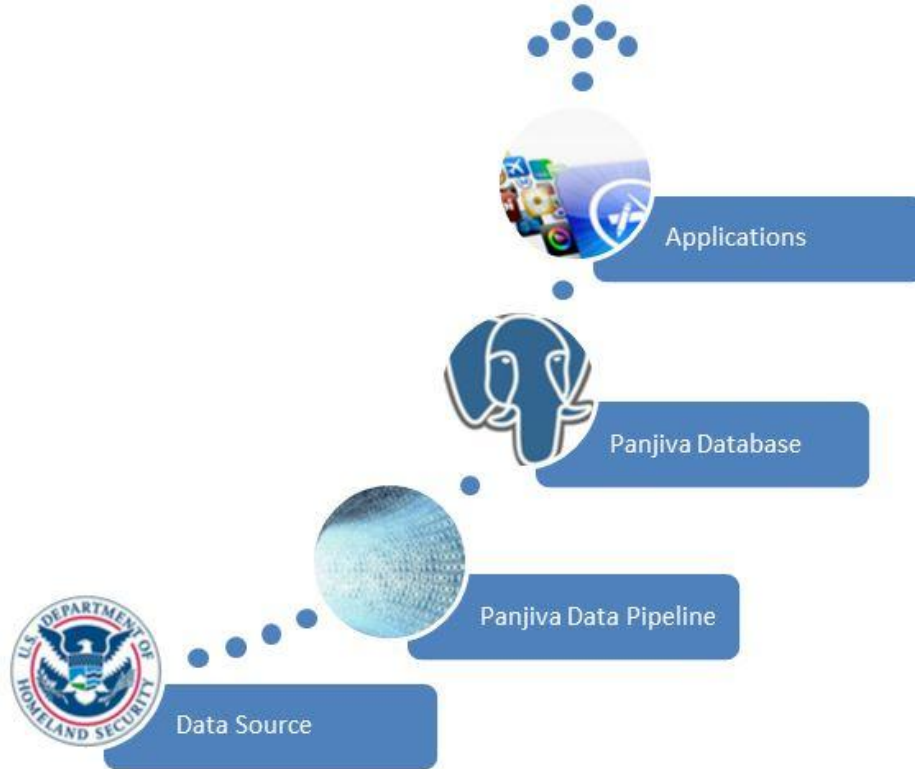
Actionable Data

- Actionable data = data must fit client use profiles:
 - contact information
 - corporate data
 - financial reports
 - product information

Delivery to Clients

- Subscription website
- Customer Relationship Management tool
- Application Programming Interface (API)
- Custom Solutions

Made-to-Measure Tools



The Business Cases

Simple Case: Supply Chain

- **Problem:** I need to find / vet a new trading partner.

Zhongde Equipment Co., Ltd.Shandong

[★ FAVORITE](#)[GROUP](#)[SHARE](#)[ALL LOCATIONS - CHINA ▼](#)[SUPPLIER ✓](#)[Summary](#)[Products](#)[Shipments](#)[Chinese Exports](#)[Customers](#)[Corporate](#)[Contact](#)

Products

[See more product information](#)

Shipments

126

Shipments sent

TOP PRODUCT TERMS

beer

brewing

beer brewing

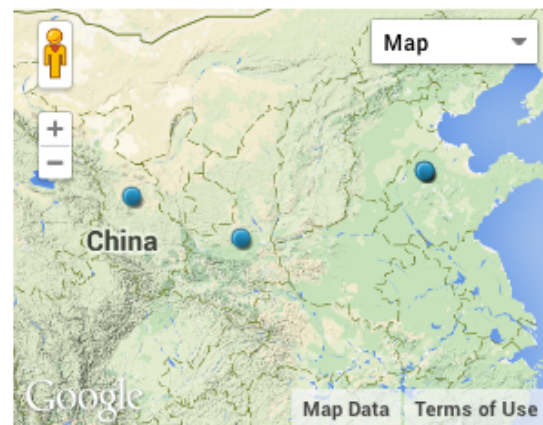
brewing

equipment

beer brewing

equipment

COMPANY LOCATIONS



DATA SOURCES






Intermediate Case: Logistics / Shipping

- **Problem:** I need to optimize shipment flows between ports / trade lanes.

Foreign Ports

Showing top 25 foreign ports with shipments to Los Angeles, CA

Trendline	Shipments	Port	Country
	3,316,390	Shanghai	China
	2,724,605	Yantian	China
	1,984,695	Kaohsiung	Taiwan
	1,544,130	Hong Kong	Hong Kong
	1,427,792	Singapore	Singapore
	1,192,303	Ningpo	China
	1,189,623	Busan	South Korea
	754,459	Tsingtao	China
	388,212	Keelung	Taiwan
	360,967	Laem Chabang	Thailand
	339,182	Vung Tau	Vietnam
	308,192	Xiamen	China
	249,702	Tokyo	Japan
	237,383	Chiwan	China
	221,676	Kobe	Japan
	217,875	Dagou/Tanggu	China
	212,822	Yokohama	Japan
	171,629	Nagoya	Japan
	144,759	Manzanillo	Panama
	115,710	Bremerhaven	Germany
	96,529	Taipei	Taiwan
	92,429	Tanjung Pelepas	Malaysia
	72,144	Rotterdam	Netherlands
	71,829	Lazaro Cardenas	Mexico
	68,267	All Other Peoples Republic Of China Ports	China

[See all 613 ports](#)

Suppliers

Buyers

Products

Shipments

Shipment Search Pro

US Imports ▼

1,481 results totaling 14,614,077 shipments, 548,094,640,363 KGs and 32,353,793 TEUs

★ Save this search

⚠ Enable alerts

📄 Export to Excel

2 filters applied.

Clear all filters

Filter results...



Advanced fields...

Date Range

2007-07-01

2014-09-03

Bill Type

clear

☒ House (1,218)☐ Master (50)☒ Simple (298)

Country of Origin

☐ China (951)☐ Hong Kong (856)☐ Taiwan (830)

Showing 7 of 7 columns

Group by Carrier SCAC ▼

Sort by number of matching shipments ▼

Carrier SCAC	Address	Phone	Matching Shipments 14,614,077 shipments	Matching Weight 548,094,640,363 KG (548,094,640 MT)	Matching TEU 32,353,793 TEUs	Matching LCL 3,187,608 shipments
APLU - Apl Co Pte Ltd Nol Group	456 ALEXANDRIA RD NOL BLDG #06-00 Singapore		1,621,690 shipments 11.10%	23,743,530,118 KG (23743530.1 MT) 4.33%	3,951,438.61 TEUs 12.21%	116,981 shipments 3.67%
MAEU - Maersk Line	9300 ARROWPOINT BLVD CHARLOTTE North Carolina United States 28273	7045712606	1,236,886 shipments 8.46%	18,659,363,760 KG (18659363.8 MT) 3.40%	3,583,707.41 TEUs 11.08%	34,706 shipments 1.09%
EGLV - Evergreen Line	EVERGREEN SHIPPING AGENCY (AMERICA) CORP. JERSEY CITY New Jersey United States 07302	2017613361	1,039,291 shipments 7.11%	18,940,003,737 KG (18940003.7 MT) 3.46%	3,340,118.20 TEUs 10.32%	17,637 shipments 0.55%

Advanced Case: Governments

- **Problem:** I need to find anomalies in shipment data at the local, state, regional, port, or national level.

The Business Take Away

Generalize from Domain Expertise

- Understand the exact nature of needs
- Understand these needs as examples of a class
- Construct front-end and back-end infrastructure to solve as many needs as possible

The Future

What's Next?

- More governments
- Higher quality data
- More tools for the wide range of clients
- Custom(s) services



Want Additional Information?

- API access
- Custom services
- Other questions
- Contact us!
 - Peter G. Swartz, Data Scientist
 - peter@panjiva.com
 - Keith W. Soura, Technical Account Manager
 - keith@panjiva.com

Big Data & International Trade:

Creating Transparency through Information

For API Access, send an e-mail to:
keith@panjiva.com

P.G. Swartz

peter@panjiva.com

Panjiva, Inc.

www.panjiva.com

K.W. Soura

keith@panjiva.com