



**NICE**  
ACTIMIZE

# Trading Communication Surveillance in Different Languages

How to monitor trading calls in multiple  
languages



30 November 2017

# New Stricter Rules

The screenshot shows the NFA website's 'Rule Submissions to the CFTC' page. The page header includes the NFA logo and navigation links for 'About', 'News & Notices', 'Members', 'Registration', 'Investors', and 'Careers'. The main content area lists several rule submissions, with the following details for the highlighted entry:

- Date:** November 27, 2017
- Title:** [Proposed Amendments to NFA's Interpretive Notice: Compliance Rule 2-9: Enhanced Supervisory Requirements](#)

Other visible entries include:

- November 17, 2017: [Increase to NFA's Assessment Fee—NFA Bylaw 1301 Regarding Schedule of Dues and Assessments](#)
- August 30, 2017: [Proposed Technical Amendments to NFA Bylaw 301, Certain Registration Rules and Interpretive Notice Compliance Rule 2-9: Special Supervisory Requirements for Members Registered as Broker-Dealers under Section 15\(b\)\(11\) of the Securities Exchange Act of 1934 \(Effective September 15, 2017\)](#)
- May 25, 2017: [Proposed Interpretive Notice to NFA Compliance Rule 2-49: Swap Valuation Dispute Filing Requirements \(Effective January 2, 2018\)](#)

The screenshot shows an 'INTERPRETIVE NOTICE' from the NFA. The notice is titled 'II. OBLIGATIONS OF MEMBERS SUBJECT TO THE ENHANCED SUPERVISORY REQUIREMENTS' and includes the following section:

**A. Recording of all conversations and maintaining electronic written communications with existing and potential customers**

Those Member firms meeting the criteria requiring them to adopt the enhanced supervisory requirements will be required to make complete audio recordings of all telephone conversations that occur between their APs and both existing and potential customers, including existing and potential retail forex customers of Members subject to NFA Compliance Rule 2-36. Additionally, those Members will be required to maintain a record of all electronic written communications that occur between their APs and customers or potential customers. Electronic written communications include, but are not limited to, email, text messages, instant messages conducted via any web-based messaging system (including instant messages sent via a social media application), and any other communication that occurs in a chat room or on any social media platform. The Board believes that recording these conversations and requiring Members to maintain records of electronic written communications provides these Members with the best opportunity to monitor closely the activities of their APs and also provides these Members with complete and immediate feedback on each AP's method of soliciting customers. Members subject to the enhanced supervisory requirements that are required to record their conversations must retain such audio recordings and records of electronic written communications for a period of five years from the date each recording is created or written electronic communication occurs and the recordings and/or records of electronic written communications shall be readily accessible during the first two years of the five-year period. In retaining the recorded conversations or records of electronic written communications, Member firms must catalog the recordings and electronic written communications by AP and date. Additionally, any Member firm meeting the criteria must require all its APs to maintain a daily log for sales solicitations which reflects at a minimum the identity of each customer or prospective customer the AP spoke with or transmitted electronic written communications to on each day and the method of communication. A Member firm must be able to promptly produce, upon request from NFA or the CFTC, all conversations or records of electronic written communications relating to a specific AP, and only that AP, for a given date. Members that are required to record or maintain records of electronic communications under this Interpretive Notice are further required to promptly provide NFA or the CFTC with appropriate resources for listening to their recordings or viewing the records of electronic communications upon request.

Source: <https://www.nfa.futures.org/news/newsRuleSubList.asp>

REGULATION

# Looking to confuse compliance? Hold a moment while I change languages

Technology lags the rulebook as City firms struggle to monitor any telephone call not conducted in English



Pick up the phone to any City trader, and compliance will be listening. Speak anything other than English, and they may struggle to understand.

Under pressure from regulators to improve staff behaviour, financial institutions in the UK have been ramping up efforts over the last 18 months to better monitor telephone conversations.

City firms have turned to new technology to help them analyse staff phone calls, emails, instant messages and texts against trading data to alert them to patterns of behaviour which may set alarm bells ringing. But the systems currently available are no panacea - much of the tracking software for phone calls can currently understand only English.

Monitoring voice communication is central to the requirements of Europe's new market abuse regulation, which came into effect in early 2016, and to revisions of the European trading rulebook, which are due to come into effect in the UK and Europe in 2018.

Under the second **Markets in Financial Instruments Directive**, in order to detect manipulative trading, firms must systematically record and monitor traders' phone calls, emails, instant messages, texts, a wide range of executed transactions and every single quote produced by the bank or trading firm.

According to Graham Ure, a partner in the forensic technology team at professional services consultancy PwC, the compliance software that helps firms monitor the channels

Monitoring  
telephone  
conversations not  
conducted in  
English is becoming  
a problem for  
trading compliance

# 6909 World Languages

Rank	Language	First Language	2 <sup>nd</sup> Language	Total
1	<b>Mandarin Chinese (incl. Standard Chinese)</b>	897 million	193 million	<b>1.09 billion</b>
2	<b>English</b>	371 million	611 million	<b>983 million</b>
3	<b>Hindustani (Hindi/Urdu)</b>	329 million	215 million	<b>544 million</b>
4	<b>Spanish</b>	436 million	91 million	<b>527 million</b>
5	<b>Arabic</b>	290 million	132 million	<b>422 million</b>
6	<b>Malay (incl. Indonesian and Malaysian)</b>	77 million	204 million	<b>281 million</b>
7	<b>Russian</b>	153 million	113 million	<b>267 million</b>
8	<b>Bengali</b>	242 million	19 million in Bangladesh	<b>261 million</b>
9	<b>Portuguese</b>	218 million	11 million	<b>229 million</b>
10	<b>French</b>	76 million	153 million	<b>229 million</b>
11	<b>Hausa</b>	85 million	65 million	<b>150 million</b>
12	<b>Punjabi</b>	148 million	-	<b>148 million</b>

Rank	Language	First Language	2 <sup>nd</sup> Language	Total
13	<b>Japanese</b>	128 million	1 million	<b>129 million</b>
14	<b>German</b>	76 million	52 million	<b>129 million</b>
15	<b>Persian</b>	60 million	61 million	<b>121 million</b>
16	<b>Swahili</b>	16 million	91 million	<b>107 million</b>
17	<b>Telugu</b>	80 million	12 million in India	<b>92 million</b>
18	<b>Javanese</b>	84 million	-	<b>84 million</b>
19	<b>Wu Chinese (incl. Shanghainese)</b>	80 million	-	<b>80 million</b>
20	<b>Korean</b>	77 million	-	<b>77 million</b>
21	<b>Tamil</b>	67 million	8 million in India	<b>75 million</b>
22	<b>Marathi</b>	71 million	3 million in India	<b>74 million</b>
23	<b>Yue Chinese (incl. Cantonese)</b>	72 million	-	<b>72 million</b>
24	<b>Turkish</b>	71 million	<1 million	<b>71 million</b>

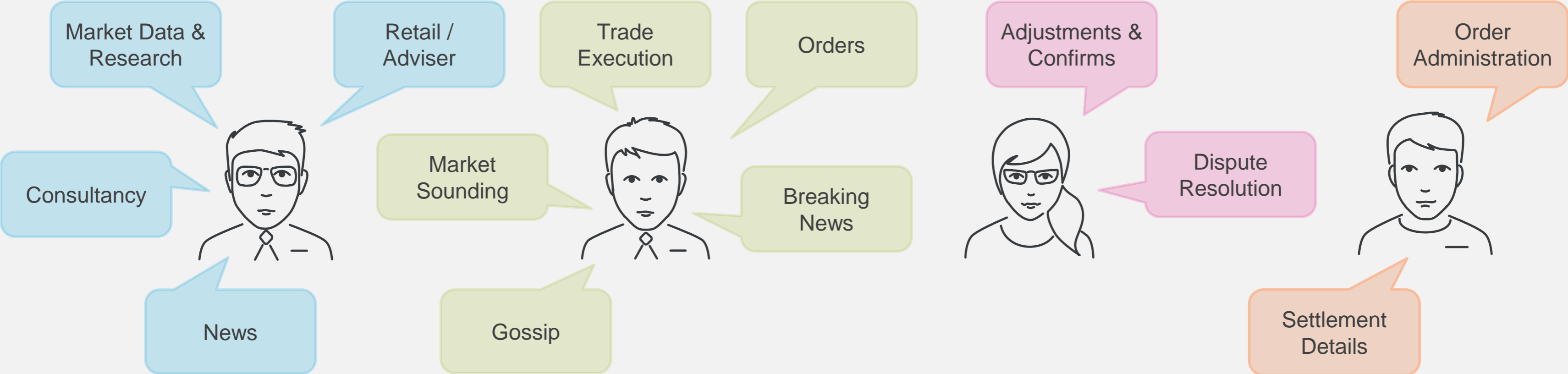
Source: Wikipedia - List of languages by total number of speakers

# Top 10 Trading Languages

Rank	Language	First Language	2 <sup>nd</sup> Language	Total
2.	Chinese(Mandarin)	897 million	193 million	1.09 billion
1.	English	371 million	611 million	983 million
3	Hindustani (Hindi/Urdu)	329 million	215 million	544 million
8.	Spanish	436 million	91 million	527 million
5	Arabic	290 million	132 million	422 million
6	Malay (incl. Indonesian and Malaysian)	77 million	204 million	281 million
7.	Russian	153 million	113 million	267 million
8	Bengali	242 million	19 million in Bangladesh	261 million
10.	Portuguese	218 million	11 million	229 million
4.	French	76 million	153 million	229 million
11	Hausa	85 million	65 million	150 million
12	Punjabi	148 million	-	148 million

Rank	Language	First Language	2 <sup>nd</sup> Language	Total
6.	Japanese	128 million	1 million	129 million
3.	German	76 million	52 million	129 million
15	Persian	60 million	61 million	121 million
16	Swahili	16 million	91 million	107 million
17	Telugu	80 million	12 million in India	92 million
18	Javanese	84 million	-	84 million
19	Wu Chinese (incl. Shanghainese)	80 million	-	80 million
9.	Korean	77 million	-	77 million
21	Tamil	67 million	8 million in India	75 million
22	Marathi	71 million	3 million in India	74 million
23	Yue Chinese (incl. Cantonese)	72 million	-	72 million
5.	Italian	61 Million	10 Million	75 Million

# Trade Conversations Take Place Everywhere



Complicated products such as Swaps, Fixed Income and Cross Asset are traded **over the phone**.

# Considerations for Surveillance of Trade Conversations

- Manual vs. Automated
- Text vs. Voice
- Trading has its own unique set of 'slang' that is not typical in normal conversations
- Traditional speech engines do not speak 'trading' language
- Compliance analysts may struggle with reviewing text for languages they do not speak



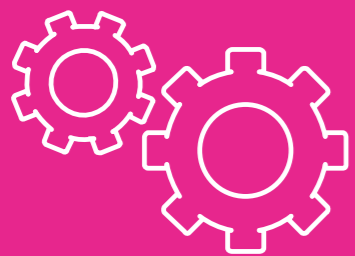
# The Challenges with Multiple Language Communications



Speech Recognition – How to ‘understand’ voice calls



Language Identification – How to ‘detect’ which language is spoken

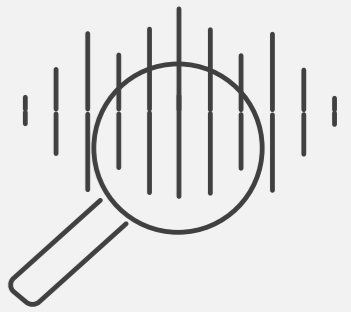


Analytics – How to ‘interpret’ the conversations

# Speech Recognition



# Automated Speech Recognition (ASR)



## What is it?

- Automated Speech Recognition (ASR) consists of linguistic models that enable converting spoken language into text

## Why use ASR?

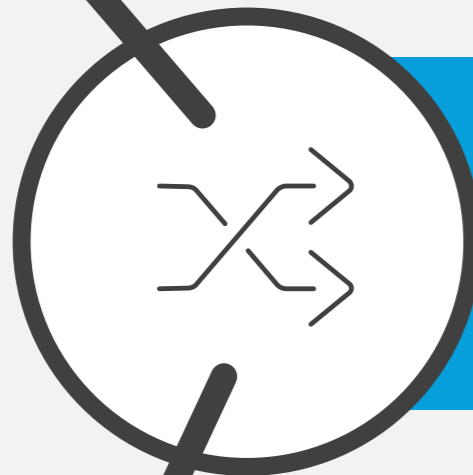
- Provides a single query interface across all communication channels
- Allows for easily using synonym dictionaries to increase recall of relevant recordings
- In combination with contextual queries improves result accuracy
- [Accuracy of ASR Models has significantly improved over earlier models](#)

# Effective Voice Analytics Process



## Basic Voice Transcription

- Transcription may miss specific industry or product set verbiage.



## Phrase Packs

- Go beyond language transcription
- Incorporate industry terminology and product information into transcription



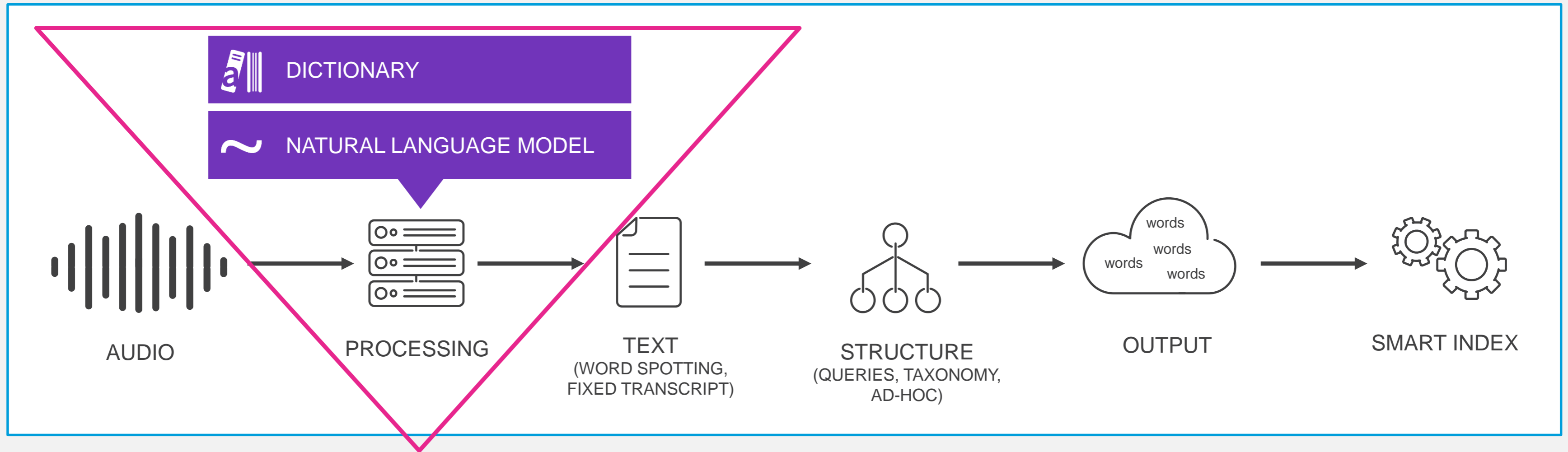
## Ongoing Management

- Consistent updates necessary to ensure translation models continue to perform at optimum efficiency.

# Language Identification

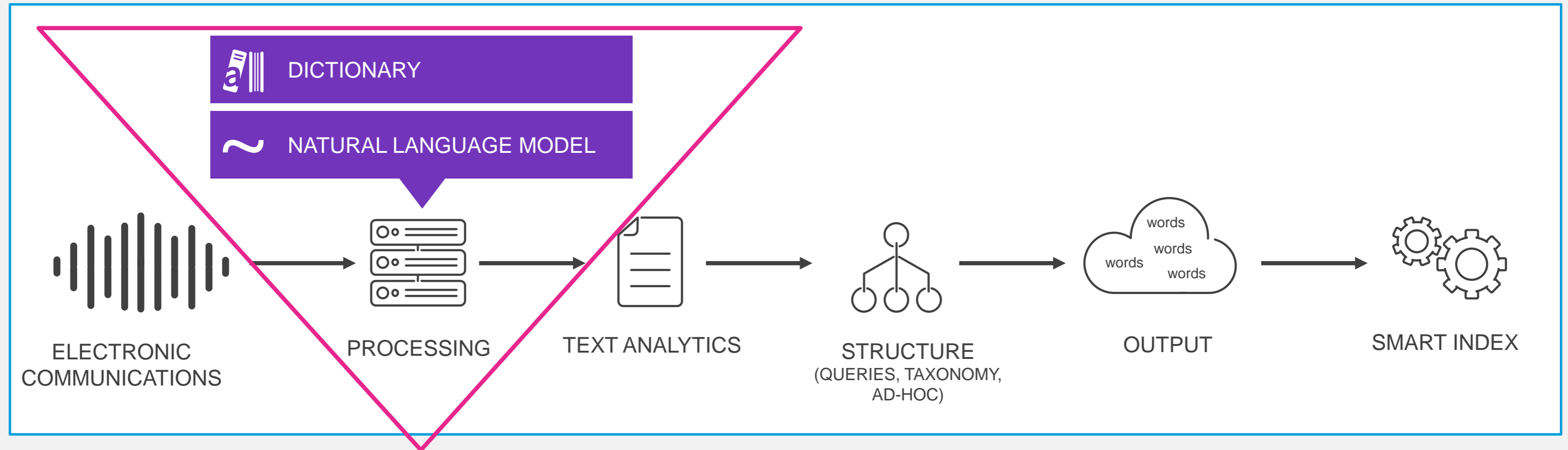


# Language Identification – Voice



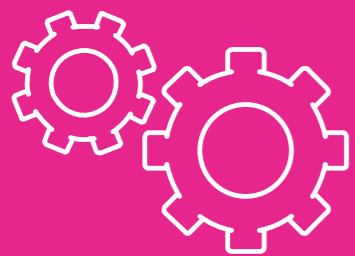
- Reduction false positives by applying analytics to relevant languages
- Language ID support for multiple languages simultaneously simplifying the configuration for Speech Conversion

# Language Identification – Electronic Communications

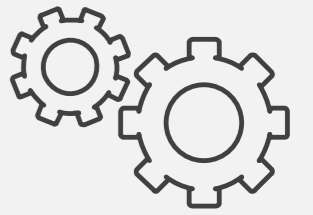


- Character set identification
- Language Composition Heuristics
- Tokenize communication content

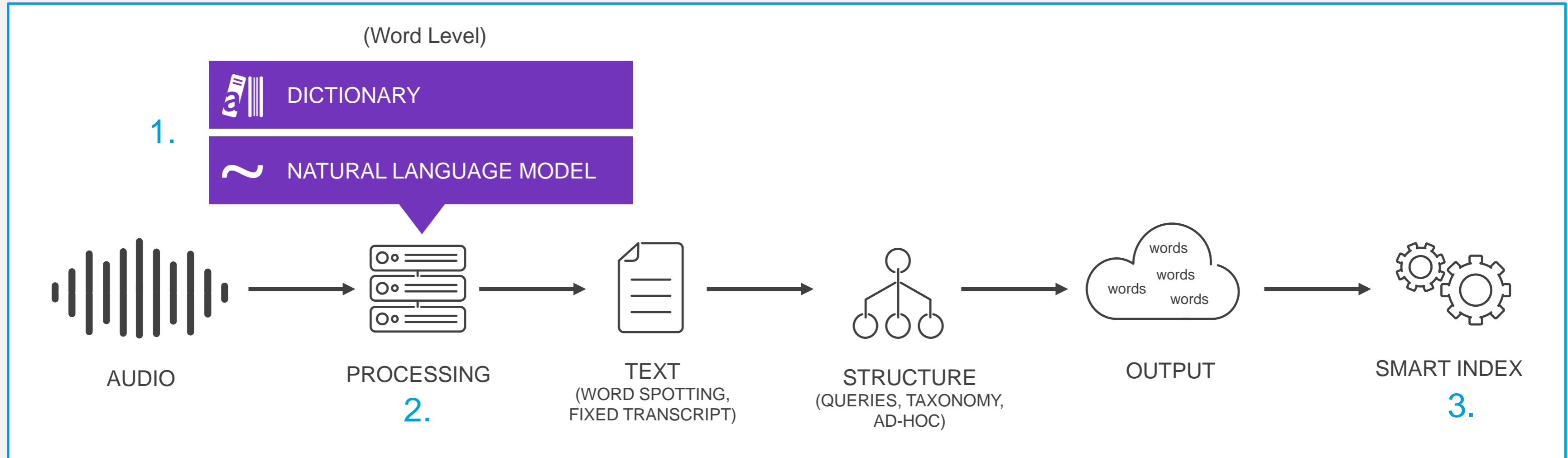
# Analytics



# How Analyzing Speech Works

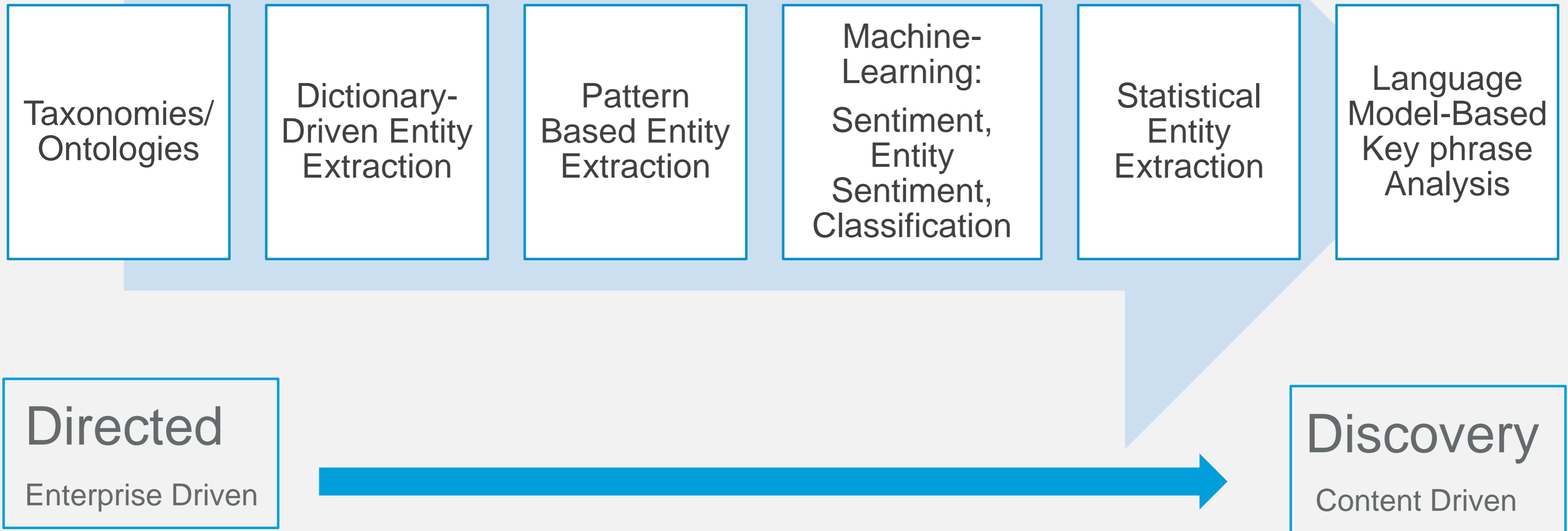
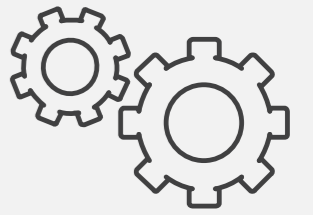


1. Dictionary and natural language models are created and tuned for each deployment
2. Audio converted into text at a rate of 5-15x real time per core



Technology well-suited for detecting patterns, aggregating data, and uncovers topics you might not know to look for

# Using Advanced Text Analytics and NLP



The background features abstract geometric shapes. On the left, a large blue triangle points downwards. On the right, a blue shape curves upwards from the top. A light grey shape is positioned at the bottom right, partially overlapping the white background. The central area is white.

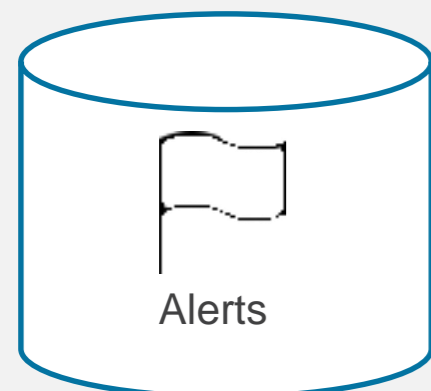
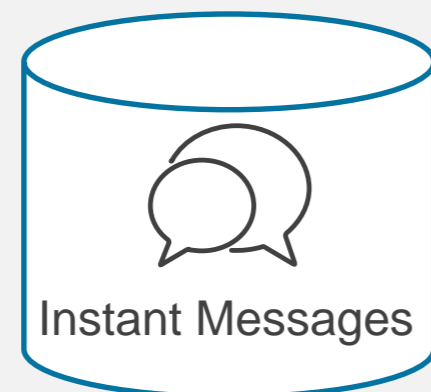
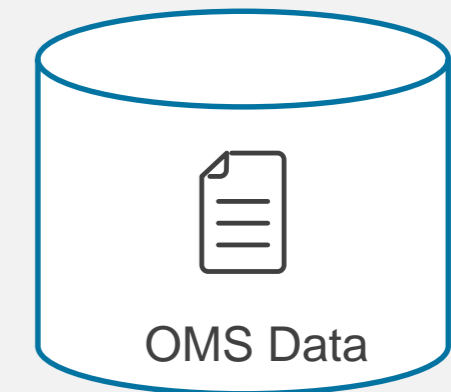
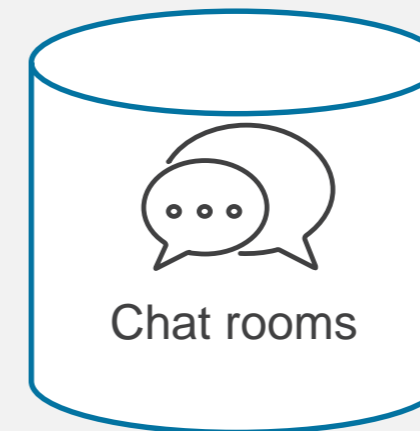
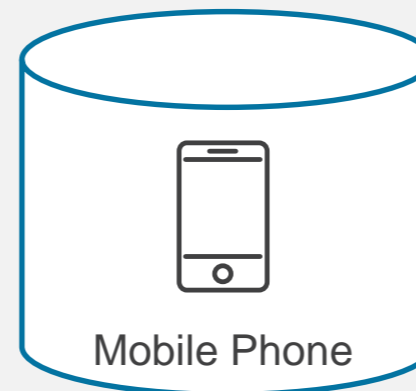
# BEST PRACTICES

# Trade Reconstruction – its complex

Disparate systems and multiple data types complicate the identification of relevant information

## What do you have?

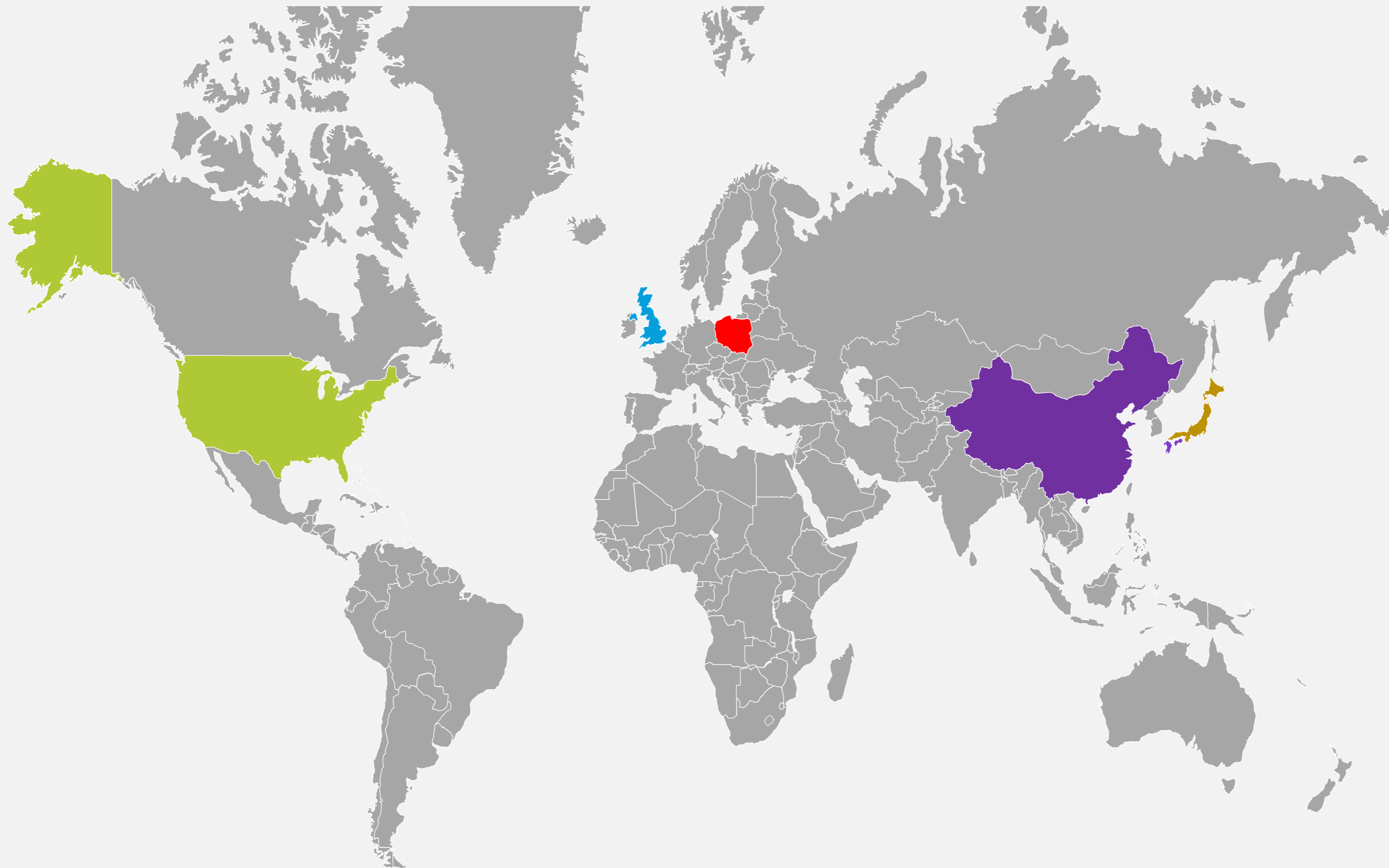
- Siloed data
- Excessive data
- Different meta data
- Different tools to access the data



## Where do you start?

- How do you identify the first place to look?
- How do you find all the relevant data?
- How do I find and connect to all relevant data together?
- How do I understand a valid output

# Case: Global Financial



# Future Developments



- Next level is transcribing all languages within a communication



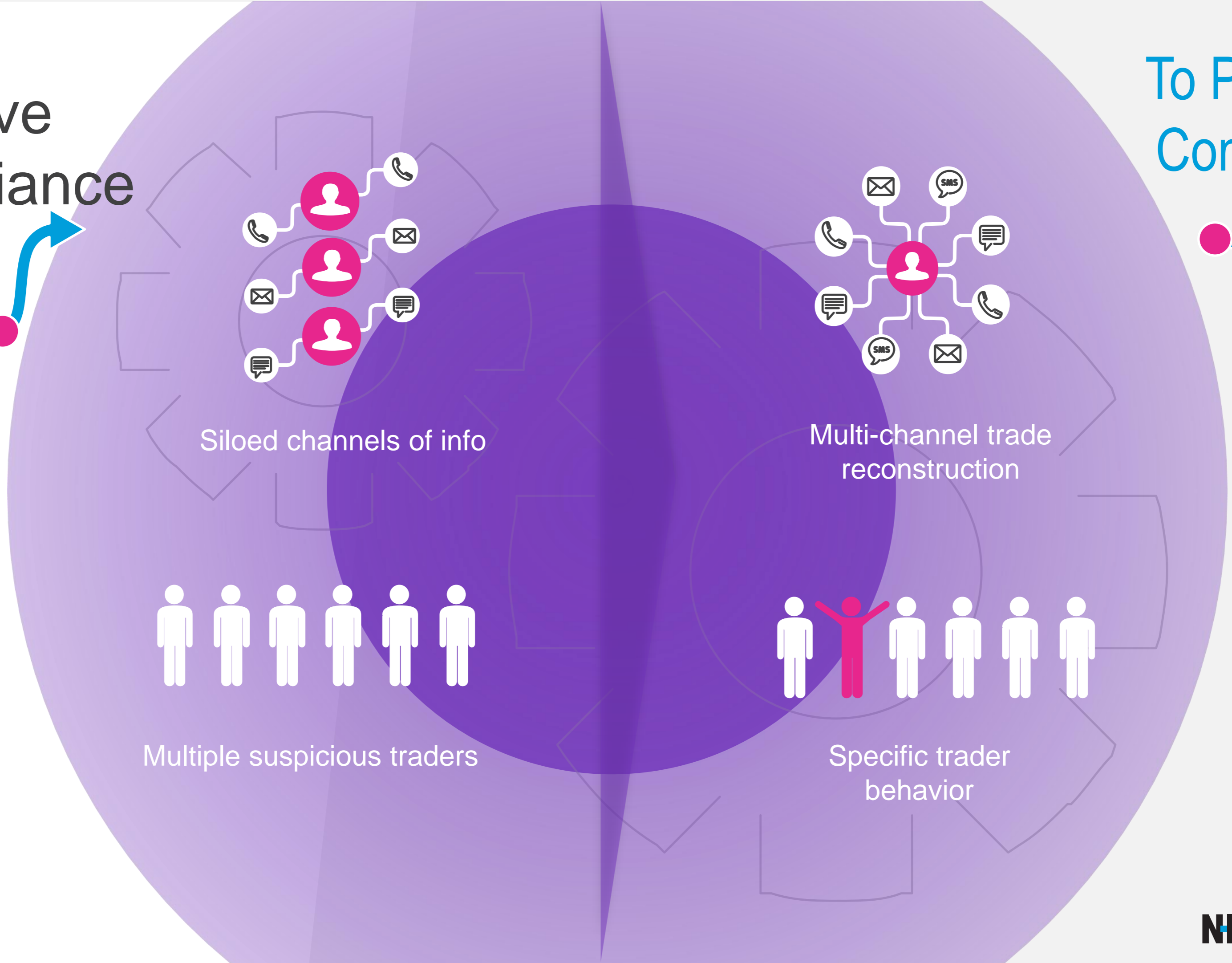
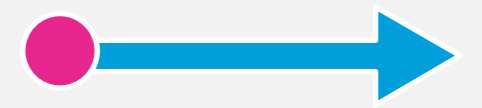
- Speaker identification



- Machine learning will correlate future communications and continue to improve language identification and transcriptions

# From Reactive Compliance

# To Proactive Compliance



# NICE Communication Surveillance

## Supported Languages today

<b>English</b> <ul style="list-style-type: none"><li>• International</li><li>• Australian</li><li>• N.A.</li><li>• UK</li></ul>	<b>Danish</b>	<b>Indonesian</b>	<b>Polish</b>
<b>Spanish</b> <ul style="list-style-type: none"><li>• Castilian</li><li>• Latin American</li></ul>	<b>Dutch</b>	<b>Italian</b>	<b>Russian</b>
<b>French</b> <ul style="list-style-type: none"><li>• Canadian</li><li>• European</li></ul>	<b>German</b>	<b>Japanese</b>	<b>Tagalog</b>
<b>Chinese</b> <ul style="list-style-type: none"><li>• Cantonese</li><li>• Mandarin</li></ul>	<b>Hebrew</b>	<b>Korean</b>	<b>Thai</b>
	<b>Hindi</b>	<b>Brazilian Portuguese</b>	<b>Turkish</b>

Questions?

Thank you, please also join:

**LIVE**WEBCAST

**THE TRADE**  
**NICE** · ACTIMIZE

Webinar – December 7<sup>th</sup>

## MiFID II, MAR and Mastering Trade Reconstruction

Join THE TRADE and NICE Actimize for an informative webinar on how to correlate structured and unstructured data for trade reconstruction

**REGISTER**



[www.HolisticSurveillance.com / webinars](http://www.HolisticSurveillance.com/webinars)

For more information visit:

[www.HolisticSurveillance.com](http://www.HolisticSurveillance.com)