InfoSphere™
software

Trusted Information

IBM Platform for Threat Prevention & Prediction
Identity Insight Software

Tony Curcio
InfoSphere Product Manager

Security Threats Are Intensifying

Today's intensifying challenges mandate a fresh approach to managing threat information

- □ Complexity, frequency, and scope of security threats continue to grow.
- □ Increasing asymmetrical nature of threat.
- □ Threats can come from individuals, foreign and domestic terrorist group, foreign governments, and even agency employees.
- □ Large and ever-increasing volumes of data need to be analyzed.
- □ Multi-cultural nature of threat.
- □ Must adhere to privacy regulations and concerns of the public.

Information Must Become a Strategic Asset

Data Types Across the Information Landscape

Name

- Last name
- First name
- Middle name
- Other name parts
- Generation
- Org name
- Aliases

Identifiers

- Credit card
- Driver’s license
- Bank account
- Tax ID
- Passport
- Loyalty club
- Phone

Attributes

- Date of birth
- Circa date of birth
- Nationality
- Citizenship
- Place of birth
- Height, Weight
- Eye, Hair color

Location

- Address
- City
- State/province
- Postal code
- Country
- Latitude/longitude

Digital

- Email address
- Cookie
- IP address
- Audio
- Video

Events

- Criminal Activity
- Phone Calls
- Money Transfers
- Crossing Border

Affiliations

- Organized Crime
- Gang
- Terrorist
- Sympathetic Organizations

The Identity Problem: Ambiguous, Misrepresented

Variety of challenges in understanding how to establish the true identity of an individual

Data Islands/Silos

- Data is spread across the enterprise
- Multiple data structures force variations

Business Processes

- Variant data collection methods and quality of training as well as individual style and personality

Data Drift & Degradation

- Life Events introduce natural change
- Common entry errors (misspelling, miskeying, abbreviations…)

Cultural Influences

- Geographically and ethnically diverse population
- Specific cultural knowledge s at a minimum

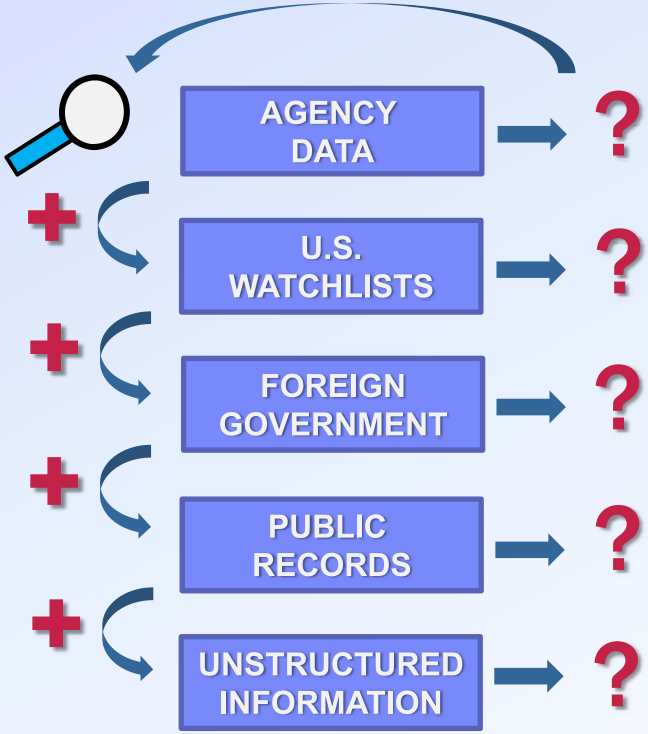
Purposeful Misrepresentation

- External parties probe all channels to defraud
- Trusted internal resources take advantage of their position

Challenges to Investigative Efforts

The "Re-search" Cycle

Name, Passport, Address, Phone, Associates, Vehicle, ...



- ❑ Did I ask the right question?
- ❑ How should I cast a broader net?
- ❑ Knowing this, now what question should I ask?

Reactive
Analyst must first know what they need to look

Labor Intensive
Each alert requires significant investment

Subjective
Analysts may recognize "leads" differently

Temporal
Discovered information may not be shared

Senate Committee on Homeland Security and Governmental Affairs

"This was not a failure to collect information, and … it was not a failure to share it. All the dots were on the same table but our government was unable to connect them – to separate this information out of the enormous mass of information the government collects and shares -- so that this terrorist could be stopped before he acted."

"I think we also need automated mechanisms to connect disparate data points 24/7, 365 days a year, and flag potential threats for analysts to examine."

Senator Joe Lieberman
March 10,2010

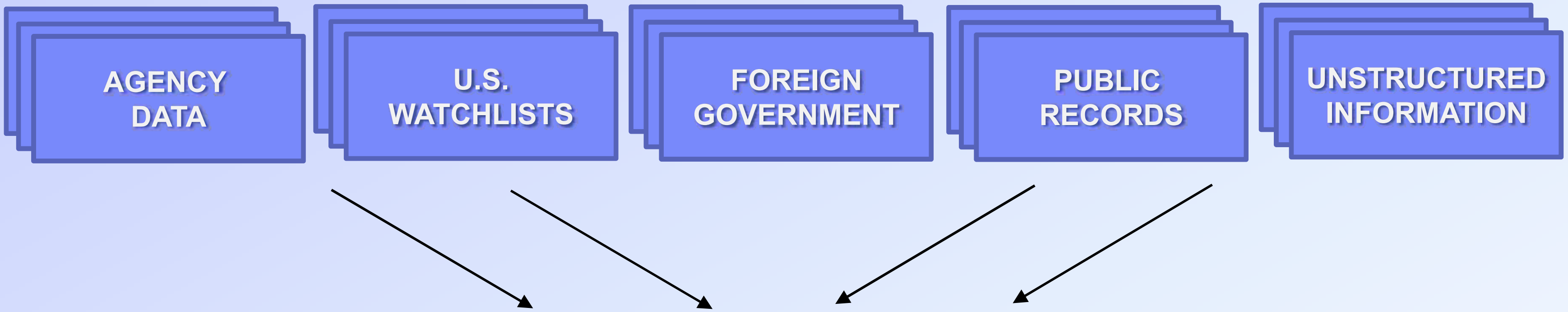
Connect related info from massive data volumes

Analyze data to find and respond to potential threats

Automated discovery that is real-time, all the time

Connect Related Information From Massive Data Volumes

Establishing Situation Context



What do I already know?

- ❑ Is this a person already in the information universe?
- ❑ Is he related to other people?
- ❑ In what location Is this activity happening?
- ❑ Is there anything unusual about when and how the data is arriving?

How do I find out?

- ❑ Examine all available pieces of incoming data to consider what to look for.
- ❑ Apply advanced algorithms to determine where this new observation connects within the information universe
- ❑ Retain enhanced information state so it is available to strengthen the next observation.

Entity Resolution establishes unique Identities and Relationships

Multi-silo Integration
Identity information is resolved across multiple data sources to create a single profile/dossier.

Sequence Neutrality
New data observations automatically redress former data that was originally not a threat.

Multi-cultural Name Recognition
Automatic application of linguistic rules that finds matches with regard to specific cultural context via patented IBM technology.

Advanced Transitive Matching
Discovers where individuals mix their identity packets by analyzing the combination of attributes from across source records..

Full Attribution
A full history of each individual is accumulated retaining all attributes so that the solution has the ability to find weak signals in the data.

Attribute & Call-out Extensibility
Adaptive data and integration models to provide for solution expansion without re-engineering.

Specialized Analytics within the Name Domain

Powered by Global Name Recognition technology

Leading worldwide multi-cultural name analytic technology

- ❑ 1B Name Facts in the Repository
- ❑ 200 Cultures Linguistically Studied
- ❑ Detailed Engine Parameter tuning
- ❑ Culturally Aware rules engine with a statistical name knowledge base
- ❑ Only name analytic product on DHS SafetyAct list

Common Data Issues

Typos:	Smith / Msith, Gomez / Bomez
Noise:	Smith/S^9mith
Truncation:	Vasconcellos / Vasconc
Concatenation:	Smith / SmithBeth
Variant spellings:	Mohammed / Imhemmed
Variant Syntax:	Maria Jose / Jose Maria
Titles and Affixes:	Hajj / Abdul
Equivalent Forms:	Abd El Salaam / Abdussalaam
Partial Names:	Maria Garcia Sanchez / Ma Garcia

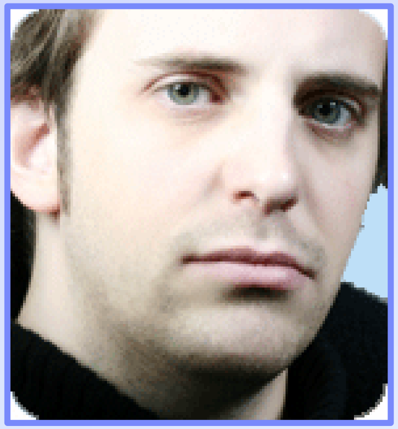
Multicultural Name Expertise

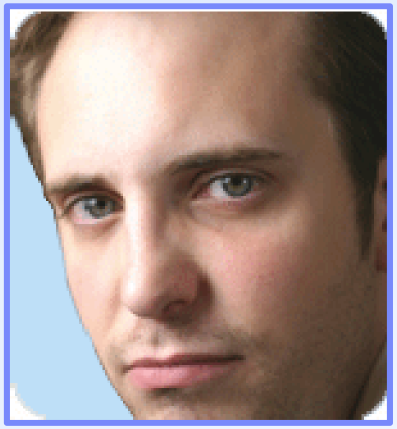
Cultural Classification	Identifies cultures of a name
Parsing Sequence	Culturally sensitive person name parser
Name Variants	Name variants based on culturally sensitive rules
Name Gender	Identifies the gender of a name
Name Comparison	Culturally sensitive name comparison

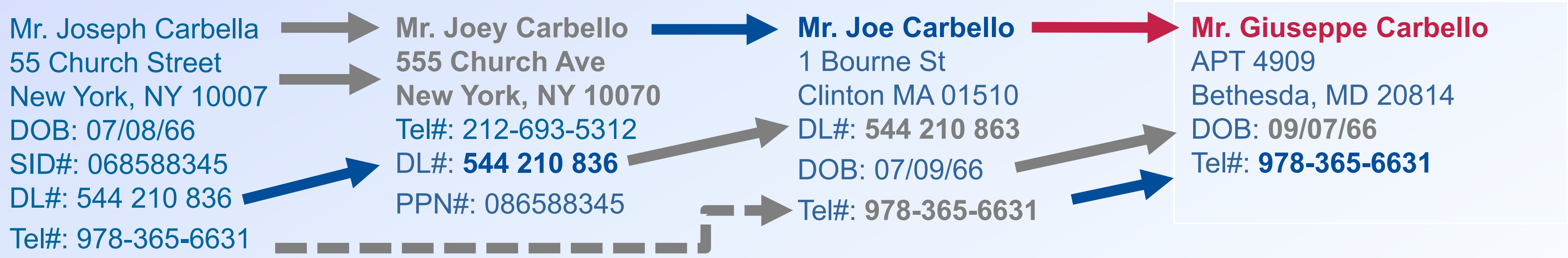
Business Name Expertise

Variable spacing in acronyms	ADT Incorporated / A D T Inc
Abbreviations and variant forms	Manufacturing / Mfg
Fuzzy matching	Coralee Inc. / Cotalee Inc.
Mis-ordered tokens	Candles by Amy / Amy's Candles
Cross-category matching	David Smith Enterprises Inc. / David E. Smith
Digits and special chars	1st Union Bank / First Union Bank
Business legal terms	Inc. / Incorporated, S.A. / Sociedad Anonim
Stop words and particles	the, of, de, la
Professional qualifiers	LCSW, Attorrney at Law

Examine identity data to understand who is who







Mr. Joseph Carbella
55 Church Street
New York, NY 10007
DOB: 07/08/66
SID#: 068588345
DL#: 544 210 836
Tel#: 978-365-6631

Mr. Joey Carbello
555 Church Ave
New York, NY 10070
Tel#: 212-693-5312
DL#: **544 210 836**
PPN#: 086588345

Mr. Joe Carbello
1 Bourne St
Clinton MA 01510
DL#: **544 210 863**
DOB: 07/09/66
Tel#: **978-365-6631**

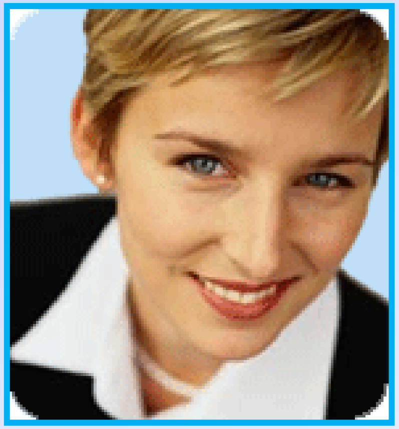
Mr. Giuseppe Carbello
APT 4909
Bethesda, MD 20814
DOB: **09/07/66**
Tel#: **978-365-6631**

→ Close match
→ Exact match
⇢ Transitive Match
→ Multi-cultural match

Analyze and uncover disclosed, fuzzy and hidden relationships

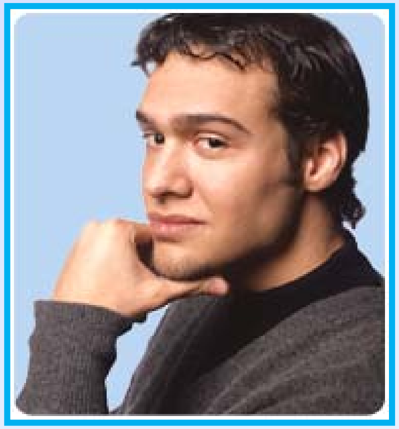
Joey Carbello

Has a car registered to the same address as . . .



Andréa Duval

Whose emergency contact is …



David Travers

Who is a suspect in the armed robbery of a state armory and shares a phone number with…

Hajj Mohamed Uthman Abd Al Ragib

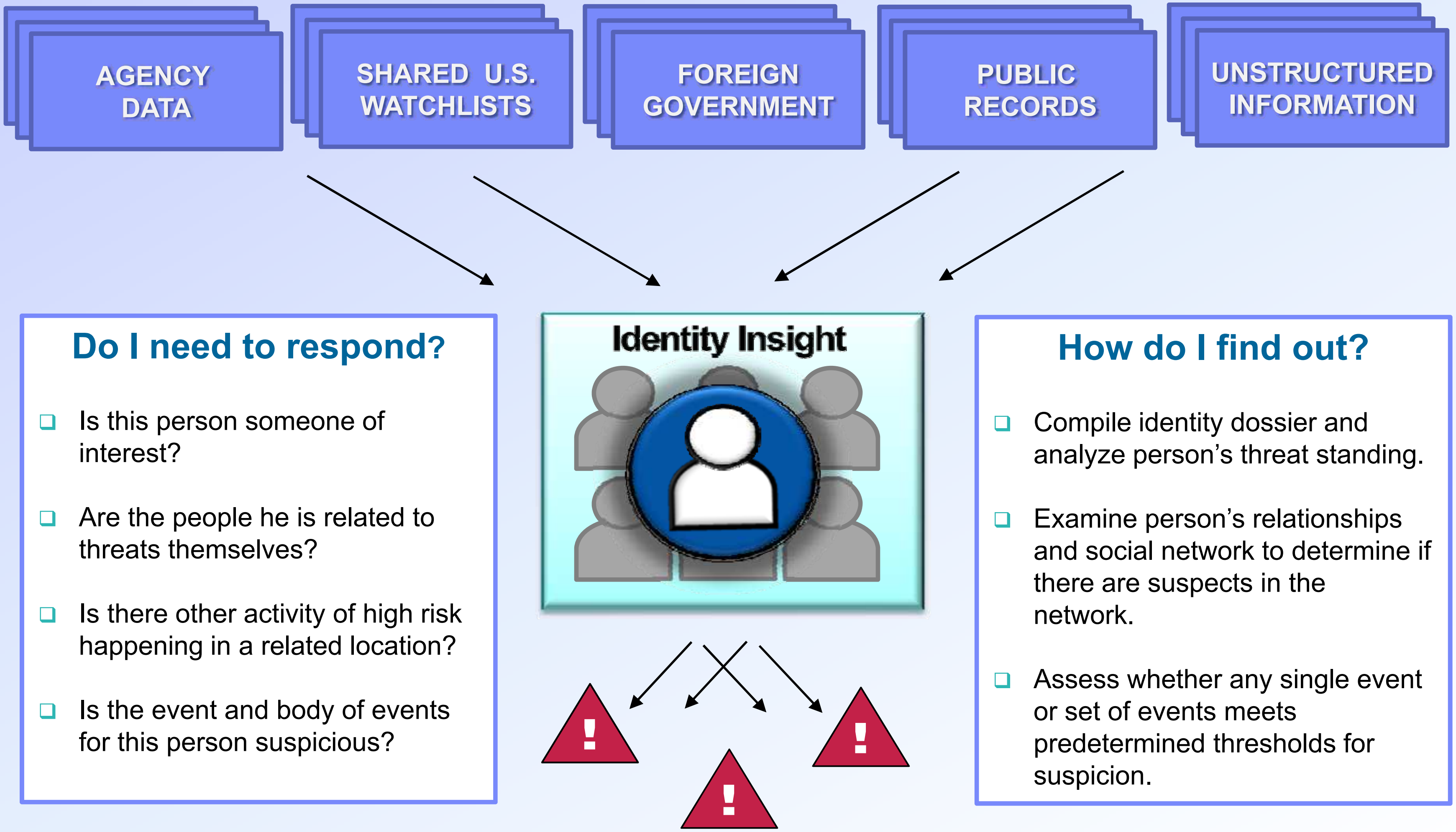
a.k.a Muhamad Usman Abdel Raqeeb

***. . .** On a government watchlist.*

Relationship Resolution creates the social network which can then be used for real-time threat assessment and investigatory link analysis.

Information Analysis to Discover Potential Threats

Automated Real-time Risk Assessment



Do I need to respond?

- ❑ Is this person someone of interest?
- ❑ Are the people he is related to threats themselves?
- ❑ Is there other activity of high risk happening in a related location?
- ❑ Is the event and body of events for this person suspicious?

How do I find out?

- ❑ Compile identity dossier and analyze person’s threat standing.
- ❑ Examine person’s relationships and social network to determine if there are suspects in the network.
- ❑ Assess whether any single event or set of events meets predetermined thresholds for suspicion.

Unique Analytical Abilities to Discover Threat and Fraud

Role Conflicts
User configured rules to provide alert notification whenever data connects conflicting roles.

Multi-Level Event Monitoring
Monitor for suspicious events as they relate to individual sources, resolved entities and relationships.

NORA™
Non-Obvious Relationship Awareness allows for weak signal conflict detection at 2 to N degrees of relationship separation.

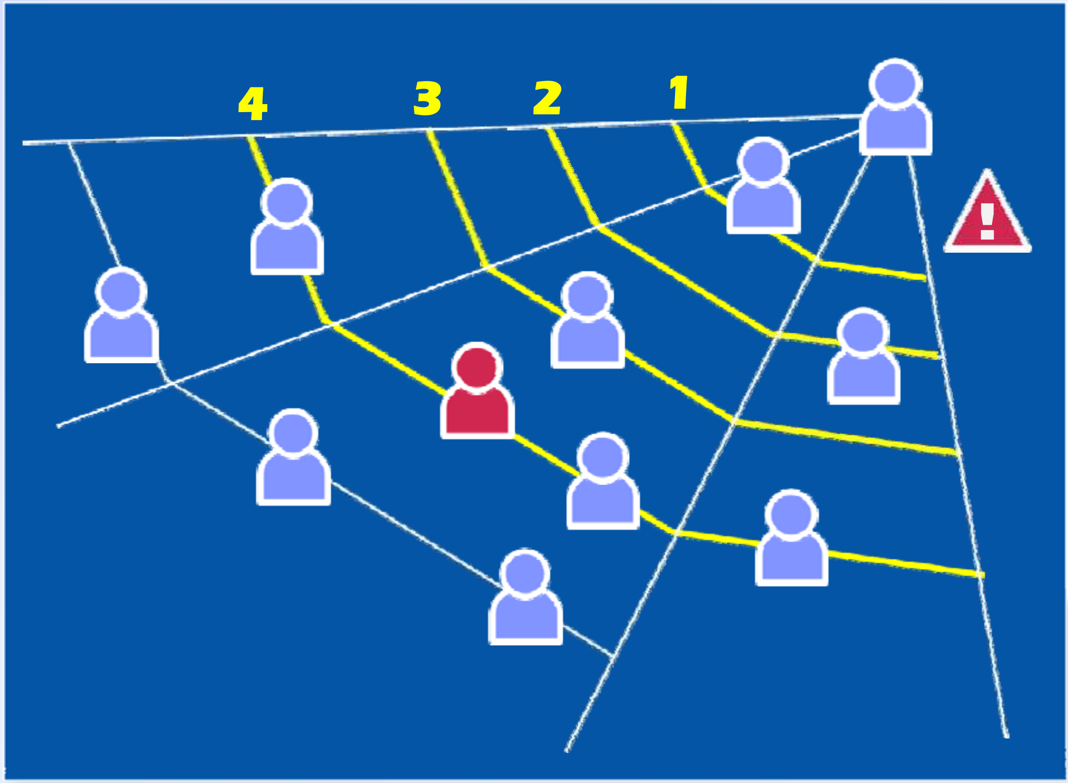
Configurable Event Types
Build and assess multiple event types across a variety of time, geography and attribute dimensions

Real-time Persistent Search
Alerts sent the moment that data an investigator has registered about a person or set of attributes is introduced to the system.

Anonymous Resolution
Allows cross agency collaboration on joint missions through fully privatized data sharing.

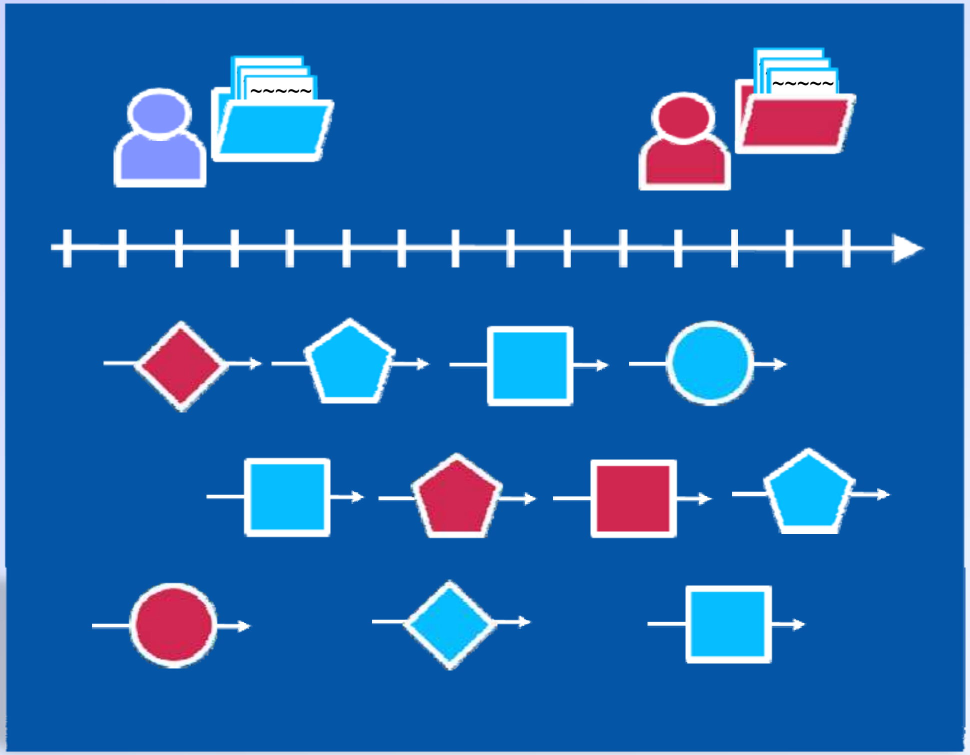
Weak Signal Detection – NORA™

- Casts a web around the chain of people related in the social network to sense activity within that group.
- Agencies can look beyond data directly attributable to an individual's obvious associations to understand their "network value" or "network threat".
- Provides real-time threat assessment at up to 30° of relationship separation.



No other commercially available technology enables streaming link discovery and "n" degree of separation link analysis on the scale of thousands of records-per-second at the moment of data ingestion.

Complex Event Processing on Resolved Entities



- Extract meaningful and actionable information by observing a diverse composition of events, happening at different times, and under different conditions.
- Interrogate traffic to find events of importance that correlate to predefined patterns.
- Assess events in relationship to the full set of identities related to any person through fully integrated Entity Resolution.

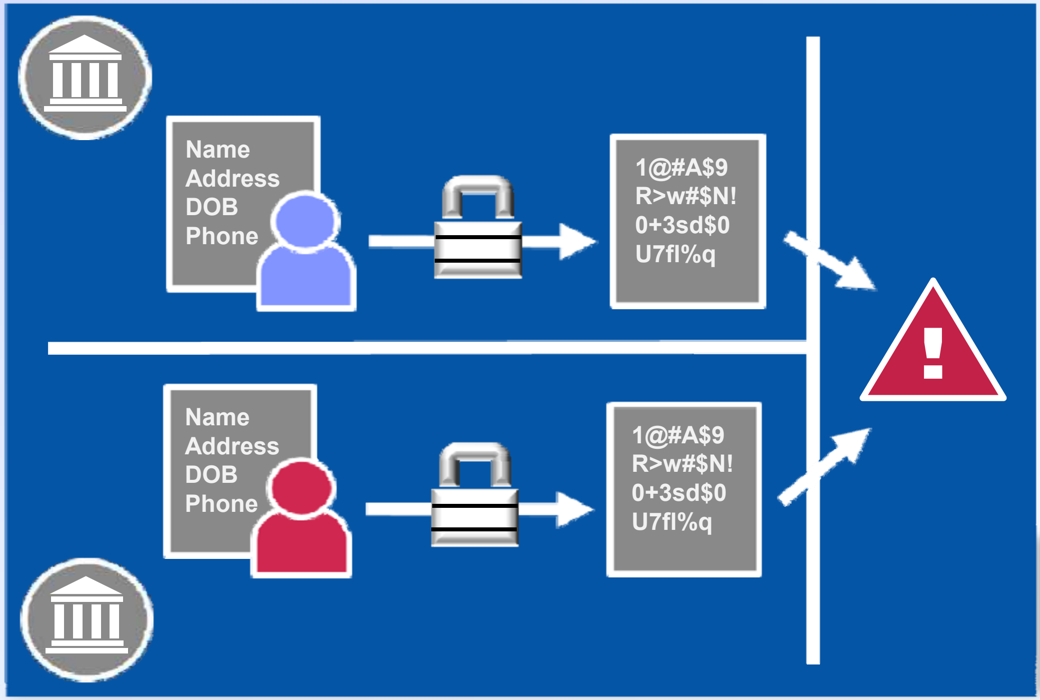
Only commercially available technology that integrates entity resolution and complex event processing into a single solution for threat and fraud detection.

Sample Complex Event Processing Rules

	Law Enforcement *Parole Violation*	**Financial Services** *Anti-Money Laundering*	**Military Intelligence** *Shipment Analysis*
Composition	*Any touchpoint with the department (ticket, accident, etc…)*	*Two or More payments*	*Ship passes through quadrant*
Correlation	*Person is currently on parole*	*Received by same person (could be various identities)*	*Any ship embarking from foreign port*
Chronology	*Anytime during duration of parole period.*	*24 hours or less*	*Time window of ½ hour*
Conditions	*Incident occurs outside his geographic limit (for instance outside of a 100 mile radius)*	*Total amounts between $9,000 and $10,000 Monies from another institution*	*Other ship in same quadrant*
Strategies	*Notify responder to apprehend*	*Raise alert for investigation by compliance team*	*Raise level of investigation of ship including crew and passengers .*

Anonymous Resolution

- Accelerate knowledge discovery process through inter-agency data sharing.
- Determines both identities and relationships anonymously
- Maintain privacy and security controls by anonymizing personally identifiable information.
- Reduce the risks of unintended disclosure and data being repurposed for use in non-permissible missions.

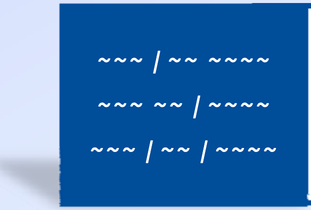


No other commercially available technology allows data used in joint missions to be protected at the originating department before it is co-mingled.

Global Name Recognition

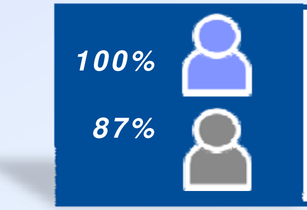
Analytics for when name matters most

Name Classification
Determines most likely cultural context as a baseline to maximize name parsing and scoring results.



Cultural Aware Name Parsing
Parses names into surname and given name components based on cultural rules.

Transliteration
Applies transliteration rules to produce an romanized equivalent forms for cross script matching.



Name Scoring
Returns a ranked, ordered list of results based on combinations of scoring for full name and name parts.

Full List Search
Compares every name in a watchlist to virtually eliminate false negative and false positive results.

Automated discovery that is real-time, all the time

Scalable
Highly optimized and tunable environment through billions of records with sub-second response.

Integration
Extensive set of Web Services to enable simple integration into the organization's business systems.

High Availability
Supports high availability & disaster recovery topologies for mission success despite hardware and other system failures.

Proven
Customers successfully deployed with hundred of millions and billions of records.

Identity Insight Software Portfolio

New Technical Whitepaper!!!

Identity Insight

Threat and fraud detection that establishes unique identity, discovers obvious & non-obvious relationships and monitors events in light of that situational awareness.

Global Name Recognition

Delivers culturally-relevant name analysis for individuals and businesses with patented search and scoring capabilities.

Anonymous Resolution

Inter-agency information sharing that discovers common identities and relationships while maintaining privacy and security by anonymizing PII.

THANK YOU