INTEL:ID
NEXT GENERATION DATA DRIVEN INTELLIGENCE SOFTWARE

Innovations and emerging technologies bring new tools and methods to law enforcement and intelligence agencies for identifying patterns and relationships in data to provide improved insights for criminal intelligence. This digital transformation is the result of demands being made that data be shared among agencies to improve collaboration in fields such as public threats, counterterrorism, child exploitation and organised crime. Increasingly, these demands are extended to data sharing and further collaboration with other agencies such as social services and health.

Entity resolution (the process of matching and linking individuals, locations, assets, etc. across multiple data sources) and the identification of non-obvious relationships are perhaps the most critical functions in law enforcement data management. And yet, many systems in law enforcement do not possess even the most basic tools to support these functions.

Duplicate records can be created by intentional deception on the part of those providing the data or via unintentional data entry errors and omissions. The very real impact of duplicate records in a system is that front-line officers are only looking at a part of the information. One record might reflect a non-violent occurrence while the second record may contain details of a violent crime that could include firearms and a relationship (affiliation) to a gang. Any front-line officer accessing an incomplete picture of the offender will not be sufficiently prepared and is placed at risk. Intelligence-led policing uses data and crime intelligence to support objective, informed decision making that both maximizes police efficiency and officer safety.

Recognizing that record matching is difficult is an understatement. In law enforcement and intelligence one must effectively deal with attempts at data obfuscation, poor and/or incomplete data, and countless variations in the data for the same entity. For example, multicultural names confound the ability of simple deterministic matching logic to effectively match records.

In our
20+ years
of experience, we have seen duplicate records in the range of
5–15%

NAMES CAN HAVE AS MANY AS 6 TO 20 LOCALIZED VARIANTS

Name variants occur for many reasons, including spelling variations (Geoff, Jeff), nicknames (Bill, William), cognates (Peter, Pierre) and culture (Juan – Hispanic, Jwan – Chinese).

Intel:ID allows investigators to manage, search, analyze, and compare names taking into account variations, nicknames, abbreviations, and cultural differences. Intel:ID uses name data sets that leverage culture-specific name data and

linguistic rules that are associated with the name's culture. The need for more sophisticated, flexible searching methods has become more pressing as our population diversity increases, as information is shared across jurisdictions, and as individuals provide incorrect or misleading information. The table below shows the potential name variations for one name.

Abdulmutallab, Umar Farouk

Summary		
Culture	Arabic	More...
Gender	Male	More...
Order	Correct	More...
Countries or Regions	Pakistan Egypt Maldives Niger Burkina Faso Syria Nigeria Indonesia Algeria Saudi Arabia	More...

| ABDULMUTALLAB	UMAR	FAROUK
MOTALEB	OMAR	FAROUK
ABDELMOTALEB	OMER	FAROUQ
MOTELEB	AOMAR	FAROOQ
ABDELMOTELEB	OMR	FAROUG
ABDELMOTELB	UMMR	ELFAROUK
ABDELMOTTALEB	OMAER	FAROQ
MOTTALEB	ALOMAR	ALFAROUK
MOTELB	OMAAR	FARROUK
ABDEL MOTALEB	ABDULLAOMAR	FAROK
ABDULMUTALAB	ABDULLAOMER	FAROOK
ABDULMUTALEB	ABDULLOMAR	FARAUK
ELMOTALEB	ABUOMAR	FAROCK
MOTALB	ALAAOMAR	FAROWK
MOTTELEB	ALLAOMAR	FARUUK
ABDEL MOTELEB	AOMER	ELLFAROUK
ABDEL MOTTALEB	BEN OMAR	FAROKE
ABDELMOTALAB	OMAERE	FAROUCK
ABDELMOTALB	OMARR	FAROWC
ABDELMOTTELEB	OOMAR	FARUOK
ABDELMOUTELEB	UMR	FRAUQ
ABDULMOTALEB	WLOOMAR	FROOCK
ALMOTALEB	UMAR	PAROUK
MOTTELB	UMER	FARQ
MOUTALEB		FARUK
ABDULMUTALLAB		FARUQ

DATA + CONTEXT = INSIGHT

An incomplete picture of an individual is only the tip of the iceberg. Social context through both known and unknown relationships allows investigators to expand the depth of inquiry from a simple request on a person to also include those persons, organizations or objects around the subject of inquiry. Using IMT's Intel:ID, along with i2 Analysts Notebook, an investigator can plot two seemingly unrelated

records to determine if a path exists and see the degrees of separation or relative strength of the relationship path. This allows investigators to operationalize the threat assessment on organized crime and identify looser gang affiliations. This is also true of any law enforcement-led investigation. Non-obvious relationships enhance the information required to make informed decisions and promote officer safety.
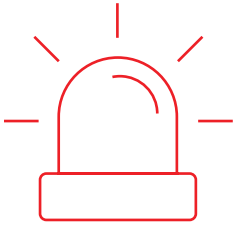
RETROSPECTIVE ANALYSIS IS GIVING WAY TO REAL-TIME CRIME CENTERS AND ALERTS

Retrospective analysis using entity resolution and non-obvious relationships is very important when crime analysts are supporting active investigations or when detecting patterns in crime through accumulated knowledge. Increasingly in law enforcement and intelligence analysis, there is a paradigm shift away from retrospective analysis towards a more centralized technology center with the advent of real-time crime centers. By increasing data from Open-source information (OSINT) and Human-source intelligence (HUMINT), the real-time data supports active alerting which brings intelligence-led policing to the front-line officer. Intel:ID supports mobile applications that include a full dossier based on all the shared data as well as external data and can provide alerts based on pre-set criteria such as gang affiliation or history of violence.
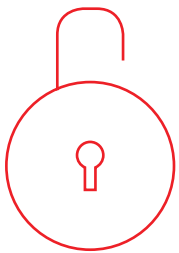
NATURAL LANGUAGE PROCESSING AI UNLOCKS THE INSIGHTS TRAPPED IN UNSTRUCTURED TEXT

A great deal of data in law enforcement is either not codified or is recorded as "unstructured text" (e.g., incident reports, witness statements, etc.). Historically, attempts to establish the complete context around a person of interest has been extremely difficult and time consuming and has often not included the mining of unstructured text. Natural Language Processing (NLP) is an AI approach that provides the capability to identify and extract critical entities from unstructured data (names, locations, etc.) and include these in the entity resolution and non-obvious relationship detection process. Intel:ID provides NLP Artificial Intelligence trained to extract person names in context from existing repositories of unstructured or textual information. This NLP Entity Extraction can train and learn on a specific type of document or data store to find names that are associated with an incident/occurrence report. The names and metadata are loaded through a pipeline that checks for redundancy in the structured content and are loaded as part of the underlying incident/occurrence as an explicit relationship.
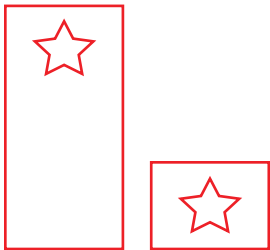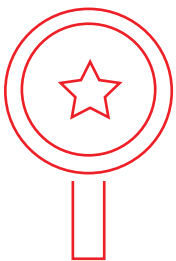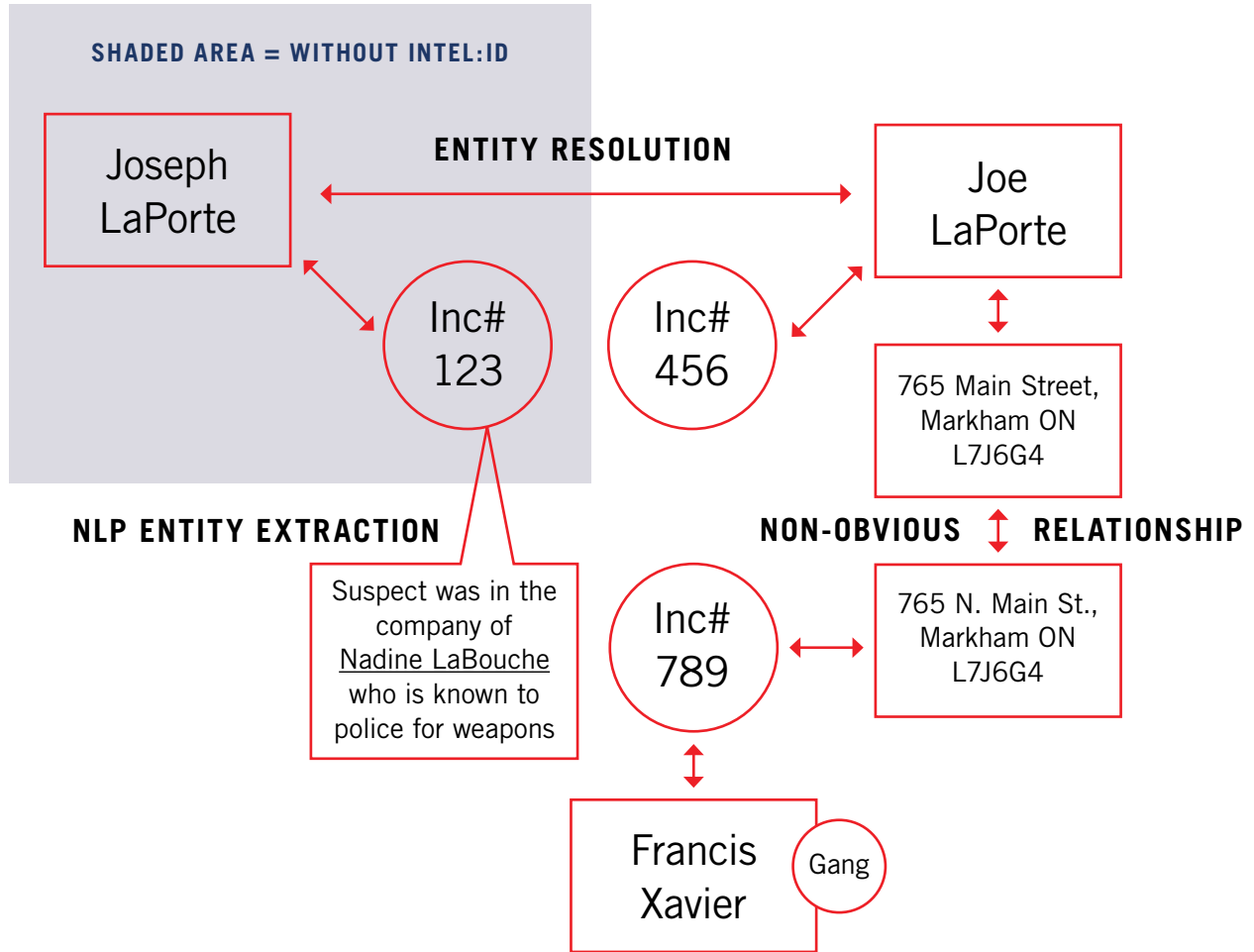
INTELLIGENCE-LED POLICING IS BECOMING ACCESSIBLE FOR AGENCIES OF ALL SIZES

Operationalizing an intelligence-led policing initiative is not always simple. Budgets are stretched to a breaking point and many conflicting priorities exist. Objectives are not always clearly stated and consensus building within the various departments and command staff remains difficult. To further exacerbate this, priorities are not always shared and competing interests can derail any ability to move forward. And yet, it is widely agreed that having incomplete information about a person of interest reduces operational efficiency and places officers at risk.

At IMT we realized that to address these very real challenges and make intelligence-led policing software accessible for medium and small agencies required rethinking intelligence software from the ground up. Using our 25+ years of expertise we have leveraged innovative open-source and AI technologies to create Intel:ID as a fully managed solution that offers quick time to value, reduced costs, and improved operational efficiency while getting ALL the data to front line officers to increase officer safety. This integrated approach enhances intelligence sharing, coordination and strategic deployment.

The shaded area in figure 1 reflects what an officer or analyst can see without the Intel:ID integrated approach for law enforcement and intelligence. In crime and intelligence analysis, the 'analysis' process is 80:20 – 80% preparation and 20% analysis. With Intel:ID, data consolidation is immediate so analysts can spend 100% of their time analyzing the data.

FIGURE 1 – IMPACT OF HAVING AN INCOMPLETE VIEW

SHADED AREA = WITHOUT INTEL:ID

Joseph LaPorte

ENTITY RESOLUTION

Joe LaPorte

Inc# 123

Inc# 456

765 Main Street, Markham ON L7J6G4

NLP ENTITY EXTRACTION

NON-OBVIOUS ⇕ RELATIONSHIP

Suspect was in the company of Nadine LaBouche who is known to police for weapons

Inc# 789

765 N. Main St., Markham ON L7J6G4

Francis Xavier

Gang

"KNOW YOUR PARTNER"

At IMT, we bring to you over 25 years of experience, lessons-learned, and best practices in the Entity Resolution and Relationship management space. Just because someone says that their solution can bring all of the disparate data together into one place does not necessarily mean that the data is being accurately matched and linked. The fact is, many data aggregation solutions simply map data from multiple systems into one consolidated repository…much like a data warehouse. Governance and data stewardship are the most often overlooked design factors with these types of solutions.

Intel:ID as a data-driven criminal intelligence solution automates linking, matching and the creation of non-obvious relationships across ALL your sources of data in your initiative while providing the means to enforce governance and mitigate risks of unintended disclosures.

To learn more about Intel:ID contact Mark Holmes at Mark@imt.ca