quantexa

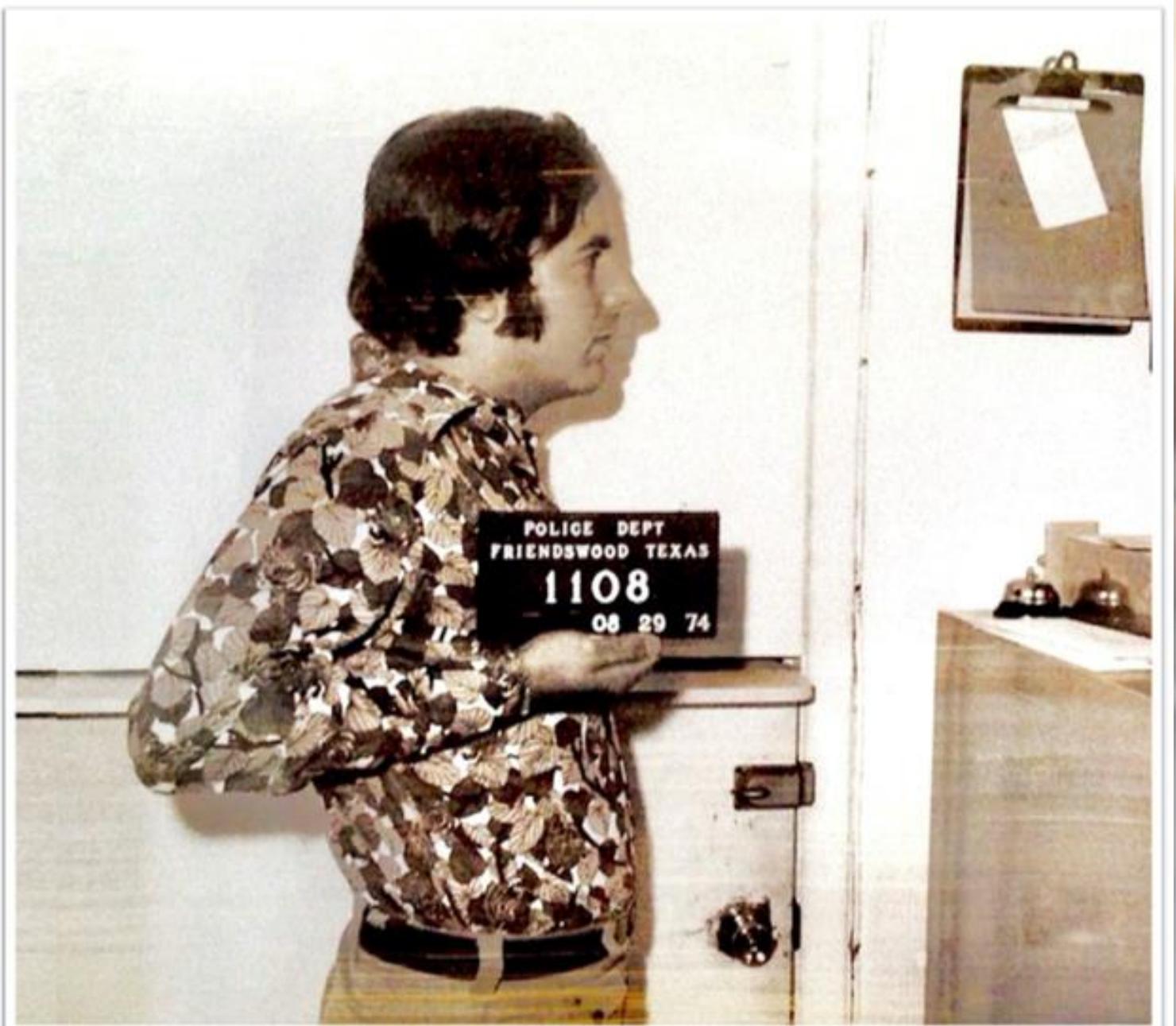# Catch me if you can!

Connecting the Dots by Leveraging generative A.I.

Daan Voets

Global Lead Product Gov. Solutions

Hallucinations

"generated content that appears factual but is ungrounded"

A US judge has fined two lawyers and a law firm $5,000 (£3,935) after fake citations generated by ChatGPT were submitted in a court filing.

Schwartz had admitted that ChatGPT, a chatbot that churns out plausible text responses to human prompts, invented six cases he referred to in a legal brief in a case against the Colombian airline Avianca.

Hallucinations

"generated content that appears factual but is ungrounded"

PA summarise this article https://www.nytimes.com/2023/03/11/technology/chatgpt-prompts-to-avoid-content-filters.html

The article discusses how ChatGPT, an AI language model, is being used to circumvent content filters that are designed to detect and block certain types of content. By generating prompts that are less likely to trigger these filters, ChatGPT can be used to create content that might otherwise be censored. While this technology could be useful for those seeking to express themselves freely, it also raises concerns about the potential misuse of AI-generated content, as well as the difficulty of regulating such content in the future. The article concludes by noting that there is still much to learn about the impact of AI language models like ChatGPT on the way we communicate online.

محمد

Muhammad

300

Muhammad
Muhammed
Muhamad
Mohammad
...
Md.
Mohd.
Muhd.
Mhd.
...

CHINESE 穆罕默德

JAPANESE ムハンマド

KOREAN 무하마드

Cyrillic Мухаммад

quantexa

A Global Leader in Decision Intelligence Solutions

OUR MISSION

To empower every organization with the right data to make the right decisions.

OUR VISION

To bring innovation and confidence in decision making to every industry in the world by driving the switch to contextual data.

Founded in 2016
grown to 750+ innovators, thinkers, solution finders, problem solvers

Offices
in London, Dublin, New York, Boston, Washington DC, Toronto, UAE, Malaga, Amsterdam, Brussels, Melbourne, Sydney & Singapore

Live in 70 countries
with customers processing hundreds of billions of records with 1000's of users

100% ARR growth since close of Series D

Acquired Aylien Q1 Calendar Year 2023

$1.8B Valuation

OUR INVESTORS

GIC
WARBURG PINCUS
dawn.
British Patient Capital
evolution EQUITY PARTNERS
HSBC
BNY MELLON
ABN·AMRO
AlbionVC

OUR PARTNERS

MOODY'S
KPMG
accenture
EY
Google Cloud
Deloitte.

SOME OF OUR CLIENTS

HSBC
Standard Chartered
Allianz
vodafone
Federal Public Service FINANCE
Public Sector Fraud Authority

How Quantexa enables genAI (Knowledge Graph RAG)

Biographies
Criminal Records
Transactions
CDR

Structured

Unstructured

Case Notes
OSINT
Court Proceedings
Social Media

Data Ingestion

No-code data ingest, cleansing, parsing, and enrichment.

Entity Resolution

Create complete 360° views of people, organizations and places.

Knowledge Graph

Use the context rich graph representations to build and transform knowledge graphs.

Contextual AI

Uncover entity, event, and graph level insights for context rich scoring models and explainable AI.

Data Exploration & Visualization

Support, augment, and automate decision-making.

Contextual Grounding (Q Assist)
Combine the power of knowledge graphs and LLMs for trusted Gen AI applications.

Platform capabilities and differentiators

| Scalable | Fast to deploy | Transparent | Highly accurate | Granular security | Open Architecture |

How Quantexa enables genAI (Knowledge Graph RAG)

Biographies

Criminal Records

Transactions

CDR

Structured

Unstructured

Case Notes

OSINT

Court Proceedings

Social Media

Data Ingestion

No-code data ingest, cleansing, parsing, and enrichment.

Entity Resolution

Create complete 360° views of people, organizations and places.

Entity Alias/Matching

Use exact values
- First/Last Name = same
- Passport#= same

Use fuzzy values
- Firstname = alias (john, jon, juan)
- Lastname = Soundex/Metaphone

Arabic / CJK translations
Last 4 digits of SSN
Date-of-birth
- +/- years
- Match MM/DD … DD/MM

Similar Phone/Address
Same Bank Branch Locations
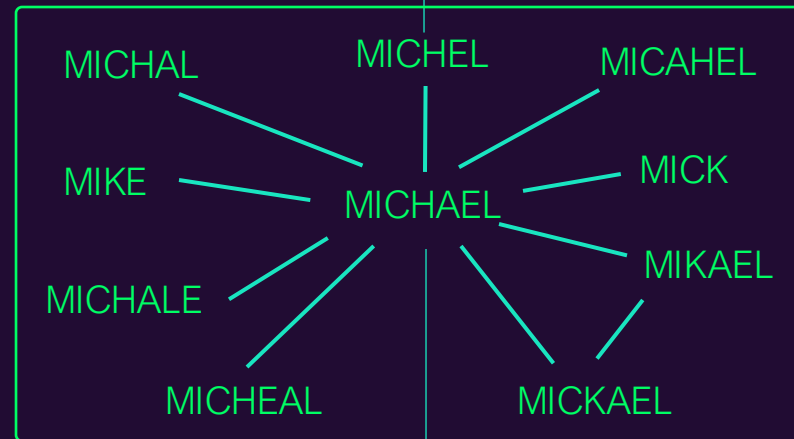## Within 5 miles of Zip Code
Common account numbers
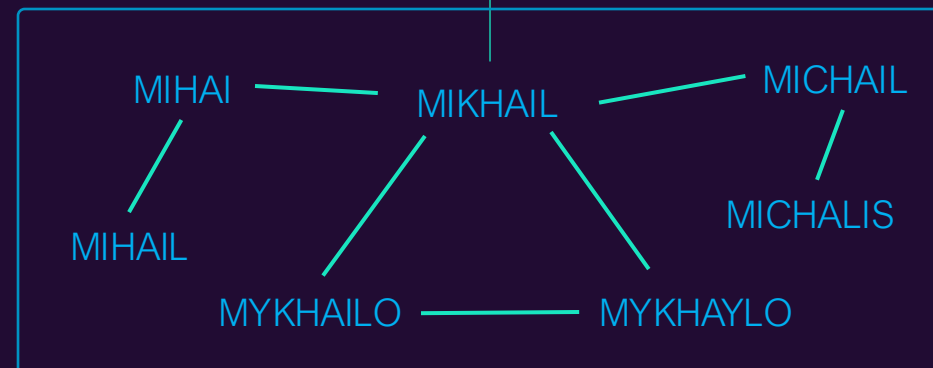## eMail Variations (x.y.z@com)
Q-Python

MIGUEL, MIKEL, MIQUEL

weak name connection

MICHAEL, MIKE, MICKAEL, MIKAEL, MICK, MICHEL, MICHAL, MICHEAL, MICHALE

weak name connection

MICHAIL, MIKHAIL, MIHAI, MYKHAYL, MYKHAYLO, MYKHAILO, MICHALIS

The Challenges

92 Avenue des Champs-Élysées, Paris, France
92 Avenue des Champs-Élysées, 75008 Paris, France
92 Ave des Champs-Élysées, Paris, France
92 Champs-Élysées, Paris, France
92 Avenue des Champs-Élysées, Paris, Île-de-France, France
92 Avenue des Champs-Élysées, 75008 Paris, Île-de-France, France
92 Avenue des Champs-Élysées, 75008 Paris
92 Champs-Élysées, 75008 Paris, France
92 Avenue Champs-Élysées, Paris, France
92 Champs-Élysées, Paris, Île-de-France, France
92 Avenue Champs-Élysées, 75008 Paris, France
92 Avenue Champs-Élysées, Paris, Île-de-France
92 Avenue Champs-Élysées, 75008 Paris
92 Ave Champs-Élysées, Paris, France
92 Ave des Champs-Élysées, Paris, 75008 France

Av. des Champs-Élysées 92, 75008 Paris, France

Address Standardization / Normalization
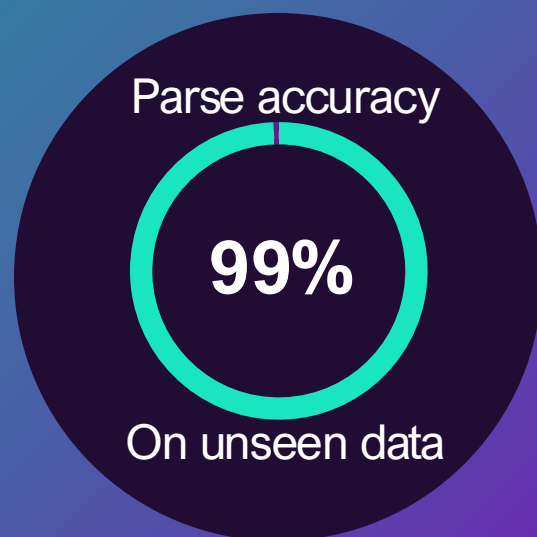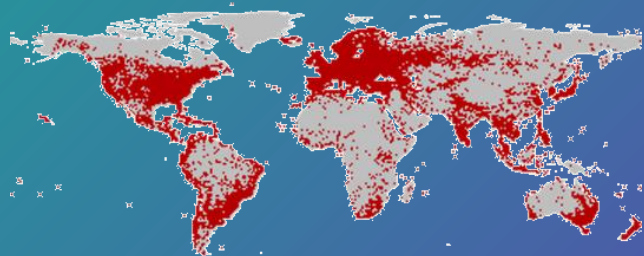
Convert messy real-world addresses:

Input	Output
One-hundred twenty E 96th St	120 east 96th street
C/ Ocho, P.I. 4	calle 8 polígono industrial 4
V XX Settembre, 20	via 20 settembre 20
Quatre vingt douze R. de l'Église	92 rue de l eglise

Normalization also to multiple variants:

92 Ave des Champs-Élysees	92 avenue des champs-elysees 92 avenue des champs elysees 92 avenue des champselysees

Transliterations:

Marktstraße 14	markt strasse 14
ул Каретный Ряд, д 4, строение 7	улица каретныи ряд дом 4 строение 7
ул Каретный Ряд, д 4, строение 7	ulitsa karetnyy ryad dom 4 stroyeniye 7

Parse accuracy

99%

On unseen data

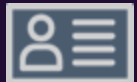# Cleansing / Parsing – Out-of-the-box Functionality

DATA/EXTRACT

Last Name
First Name
Middle Name
Date of Birth

Street
City
State
Zipcode
Country

ID-Type
Number
Description
Country
Issue-State

Number
Type
Ext

Email

TRANSFORMS

Name parser

Name synonyms

Gender guesser

Address parser

Address standardizer

Country Imputation

ID-Parser

SSN-Validator

Phone standardizer

Email parser

ELEMENTS

Soundex
Metaphone
Double Metaphone
MinHash
Synonyms/Alias
Transliteration
Known Variants

House number
County
Suburb
PO-box check
Congressional district
Zip-centroid
Postal Zip->City

Number of characters
Sequence checks
formatting
validation (SSN/ITIN)
Check-digits

Cell/fixed Classification
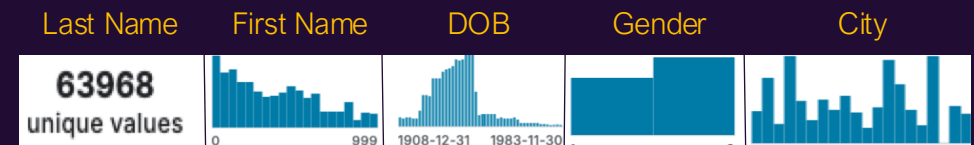Carrier
Country/Region
Validity check Geocoding

Validator
Username
Domain
Sub domain

COMPOUNDS

➢ First Name + Last Name + House Number + Street + Zip Code
➢ First Name + Last Name + Phone
➢ Fuzzy First Name + Fuzzy Last Name + Phone
➢ Date of Birth + House Number + Street + Zip Code
➢ Date of Birth + Company Name
➢ Date of Birth + Phone
➢ National Identifier
➢ etc...
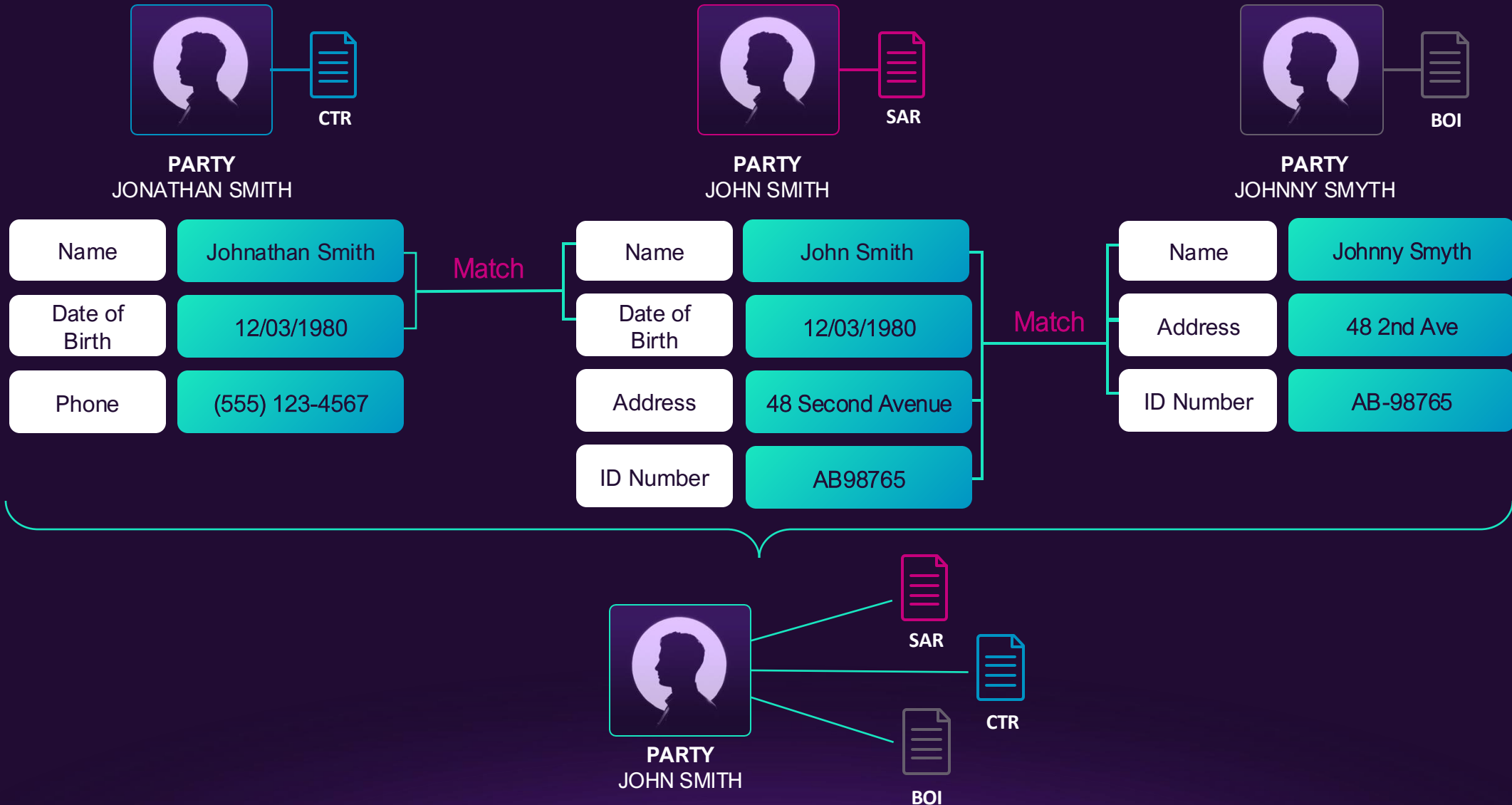
Over +400 Compounds

How unique is the value?

Last Name	First Name	DOB	Gender	City
63968 unique values	0 ... 999	1908-12-31 ... 1983-11-30	1 ... 2	

X =?= Y

Entity Resolution – are they the same?

PARTY
JONATHAN SMITH

Name	Johnathan Smith
Date of Birth	12/03/1980
Phone	(555) 123-4567

Match

PARTY
JOHN SMITH

Name	John Smith
Date of Birth	12/03/1980
Address	48 Second Avenue
ID Number	AB98765

Match

PARTY
JOHNNY SMYTH

Name	Johnny Smyth
Address	48 2nd Ave
ID Number	AB-98765

PARTY
JOHN SMITH

SAR
CTR
BOI

CTR   SAR   BOI
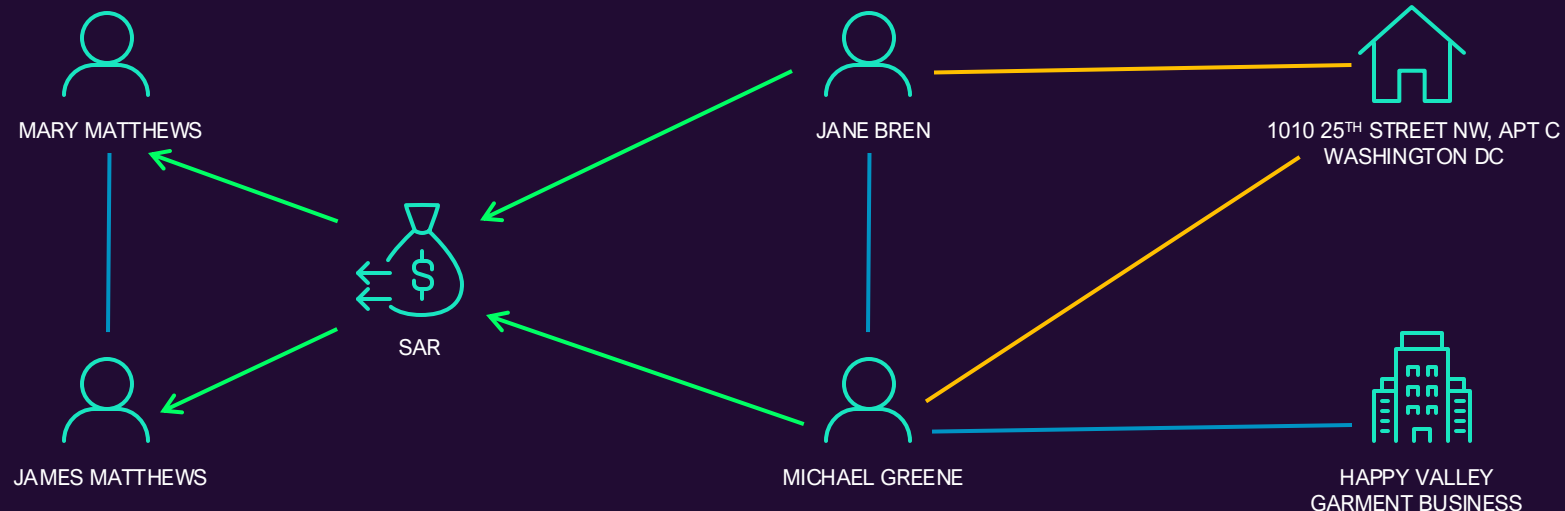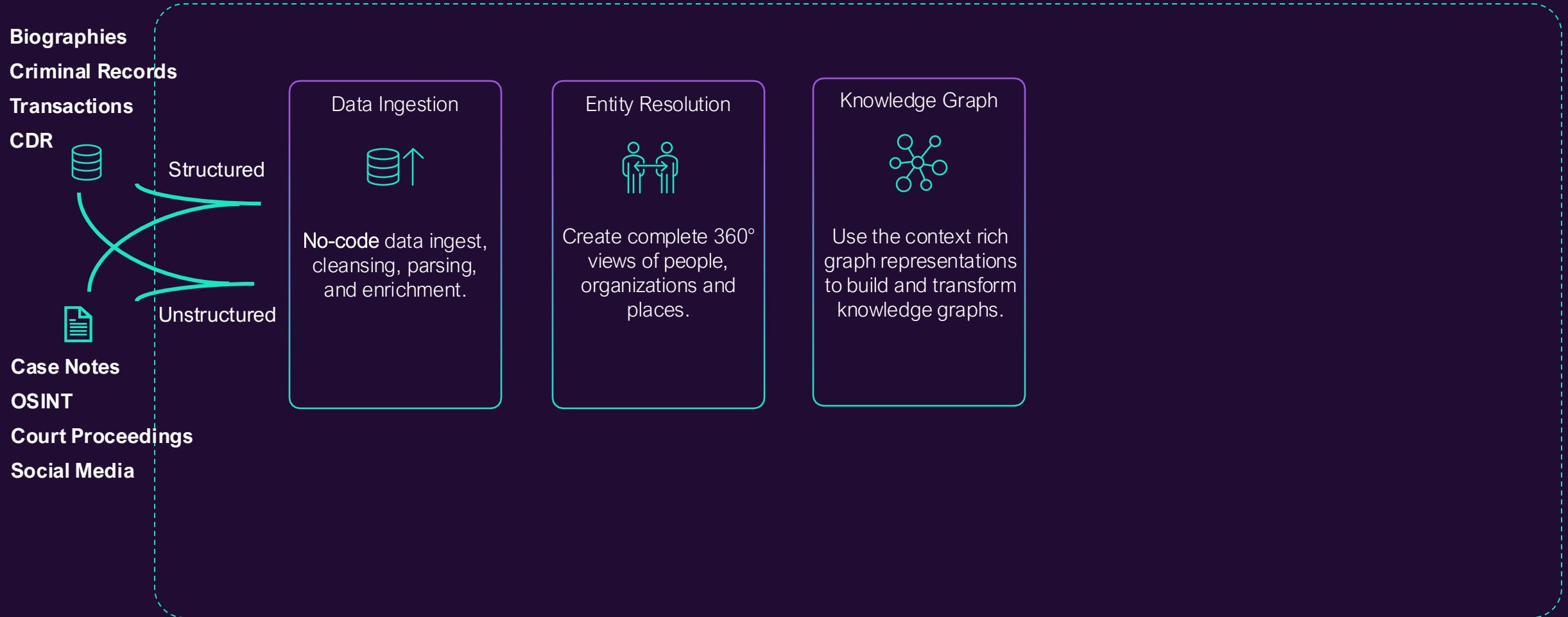
Entity Extraction from SAR Narratives

For at least a year DATE, beginning on May 20, 2002 DATE, two CARDINAL customers. Michael Greene PERSON (DOB:04/10/1992 DATE). Michael Greene PERSON is currently employed at Happy Valley Garment Business ORG, and his wife. Jane Bren PERSON (DOB:05/04/1990) DATE, with fixed address: 1010 25th St NW. Apt C LOC/STREET Washington LOC/CITY DC LOC/STATE 20037 CARDINAL have been using our money transmitter service to send large amounts of money (in cash) to receivers named: James Matthews PERSON and Mary Matthews PERSON located in Antigua LOC/COUNTRY. Funds are sent to the Bank of Hope ORG in St. Johns LOC/CITY Antigua LOC/COUNTRY. The amount of money presented each time by Michael Greene PERSON and/or Jane Bren PERSON is usually $ 5000 MONEY and the transmittals are sent bi- weekly DATE. During one CARDINAL particular incident earlier this month DATE. Greene PERSON attempted the following transaction: - June 12, 2003 DATE. Michael Greene PERSON attempted to send $ 10,000 MONEY without proper identification. We refused to send the funds and Mr. Greene PERSON left the premises. He returned later in the afternoon TIME with identification but only sent $ 5000 MONEY. All incidents / transactions have occurred at our store in Greater Bank of New York ORG. The office has documentation of the driver's licenses of both customers. Suspicion lies in the Greenes PERSON ' occupation and the amount of money leading me to suspect possible tax evasion. Records related to the money transmittals are maintained in our store in Yonkers LOC/CITY NY LOC/STATE.

Identify entities and connections in unstructured data

- People
- Org
- Address/Loc
- ID/SSN
- Phone
- Account
- IP-Address/URL
- Email

MARY MATTHEWS

JANE BREN

1010 25TH STREET NW, APT C WASHINGTON DC

SAR

JAMES MATTHEWS

MICHAEL GREENE

HAPPY VALLEY GARMENT BUSINESS

How does Quantexa prevent hallucinations?

Biographies

Criminal Records

Transactions

CDR

Structured

Unstructured

Case Notes

OSINT

Court Proceedings

Social Media

Data Ingestion

No-code data ingest, cleansing, parsing, and enrichment.

Entity Resolution

Create complete 360° views of people, organizations and places.

Knowledge Graph

Use the context rich graph representations to build and transform knowledge graphs.

Knowledge Graphs

Organize data from multiple sources, capture information about entities of interest in a given domain or task (like people, places or events), and forge connections between them
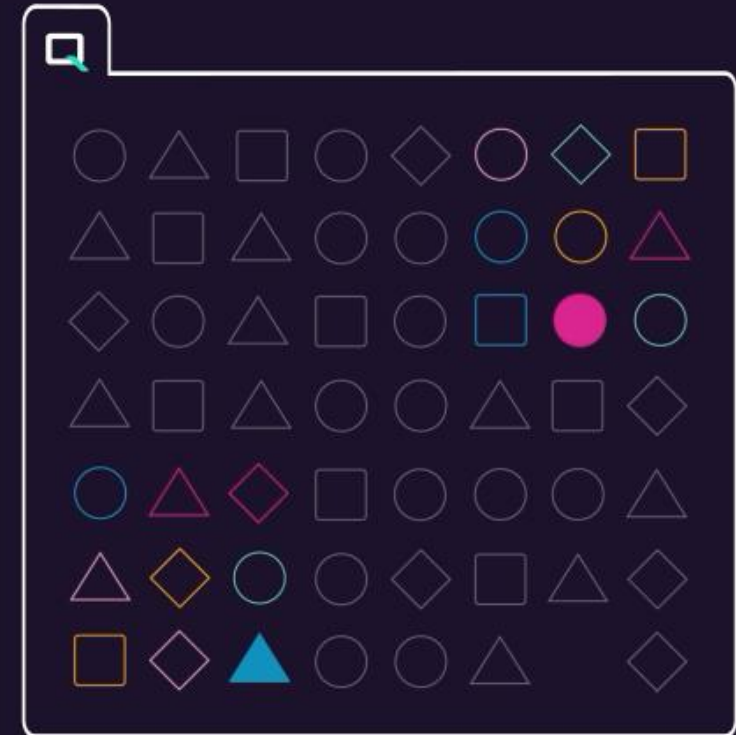
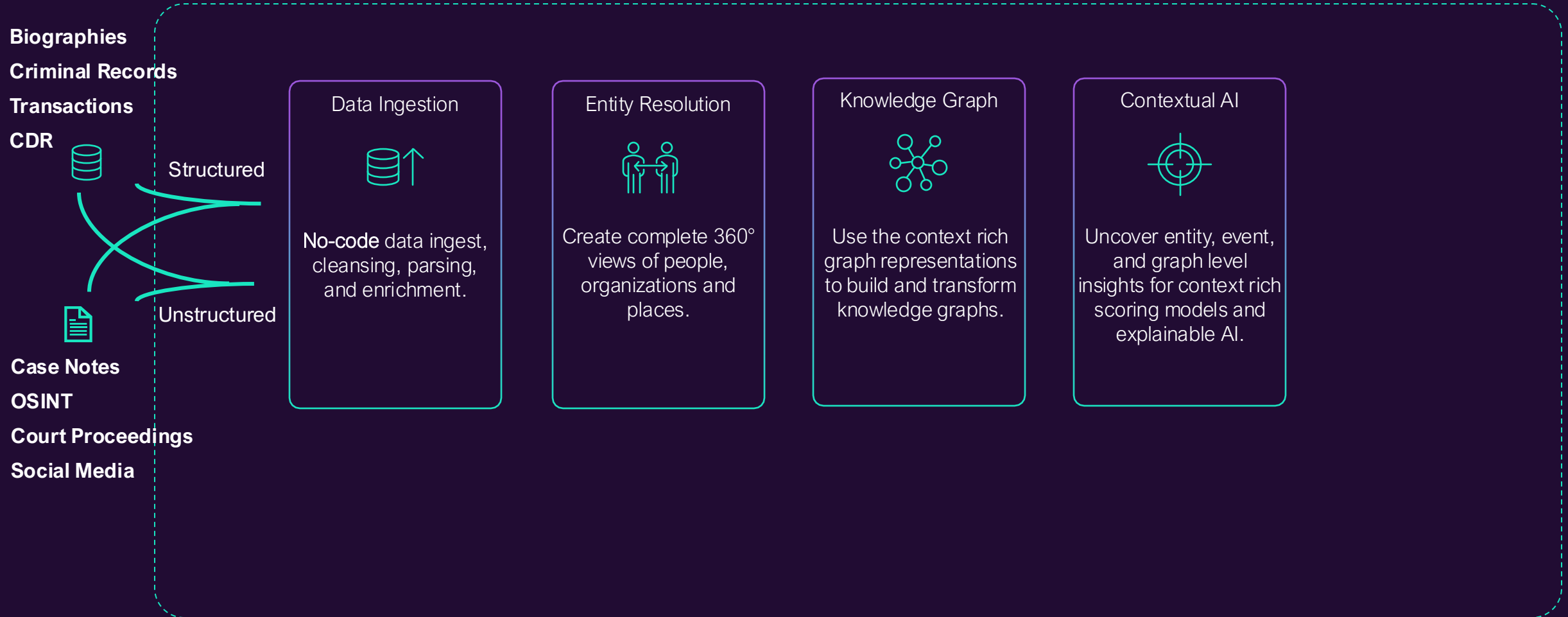# Graphs surface the relevant data for your unique problem

Build views of entities and their connections in your data to uncover relationships and insights that are relevant to your business problem.

Graph Generation finds business domain relevant sets of entities, documents and links within across the whole entity population.
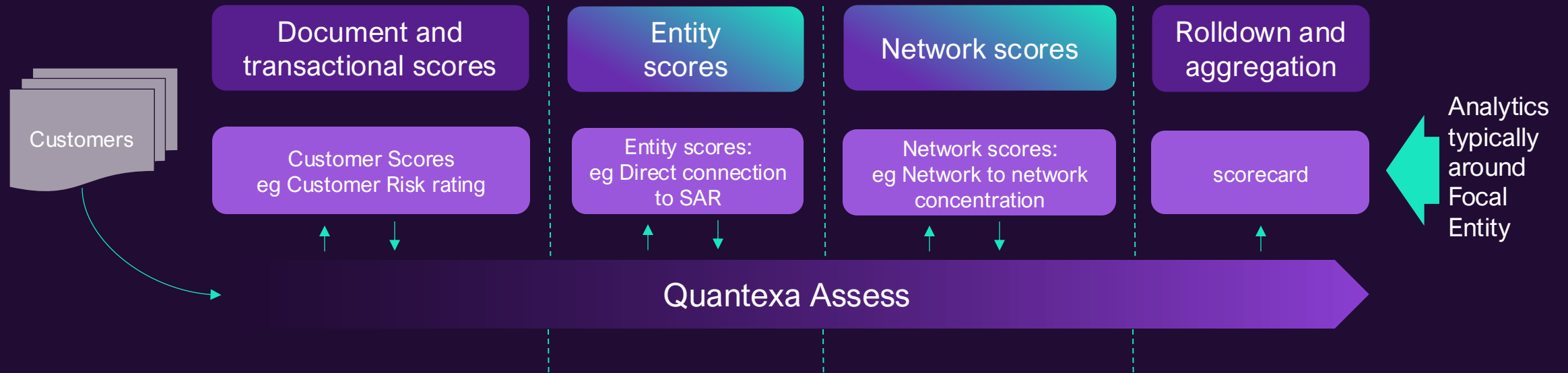
Graph Analytics then applies scores, graph analytics, graph machine learning for finding insights and RAG for contextual grounding of Gen AI.
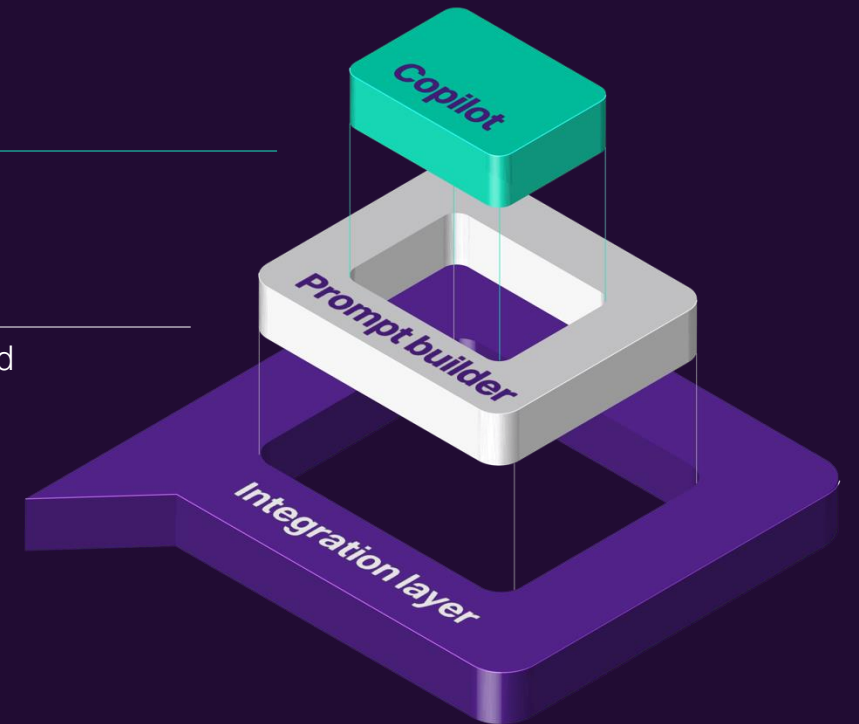
How does Quantexa prevents hallucinations

Biographies
Criminal Records
Transactions
CDR

Structured

Unstructured

Case Notes
OSINT
Court Proceedings
Social Media

Data Ingestion

No-code data ingest, cleansing, parsing, and enrichment.

Entity Resolution

Create complete 360° views of people, organizations and places.

Knowledge Graph

Use the context rich graph representations to build and transform knowledge graphs.

Contextual AI

Uncover entity, event, and graph level insights for context rich scoring models and explainable AI.

Contextual AI

Customers

Document and transactional scores	Entity scores	Network scores	Rolldown and aggregation
Customer Scores eg Customer Risk rating	Entity scores: eg Direct connection to SAR	Network scores: eg Network to network concentration	scorecard

Analytics typically around Focal Entity
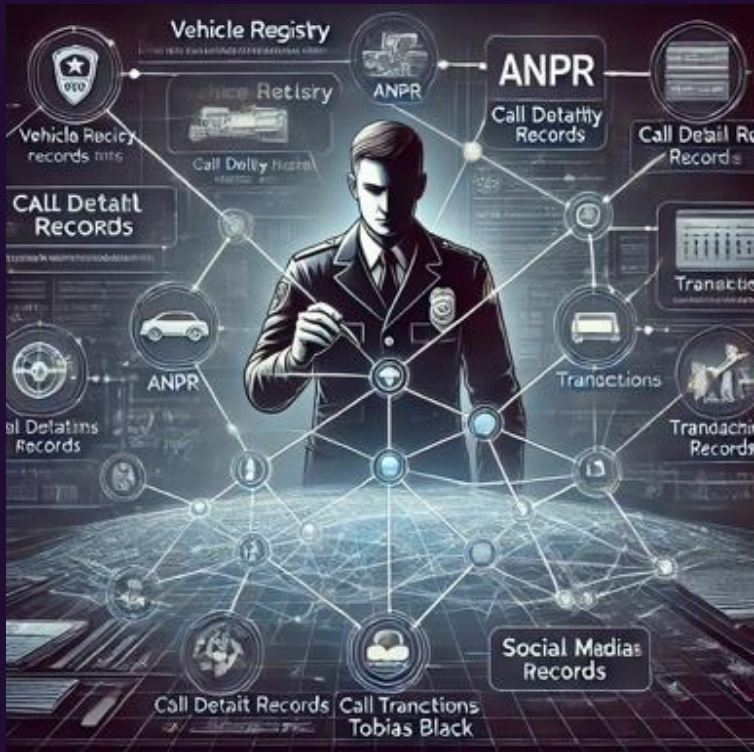
Quantexa Assess

The Context Aware Gen AI Technology Suite

Deploy, control, and operationalize Gen AI that's
grounded in contextual data to augment decisioning

- **Q Assist Copilot:**
  Information retrieval, data analysis, and reporting copilot that streamlines and enhances
  research, Investigation, and decisioning processes

- **Q Assist Prompt Builder:**
  Build, share, and maintain prompts that are grounded in contextual data for more accurate and
  trusted responses

- **Q Assist Integration Layer:**
  Tools, connectors, and APIs designed to securely link Quantexa's Decision Intelligence
  Platform with LLMs and conversational AI systems right out-of-the-box

DEMO

See how Quantexa's AI capabilities help Officer Jerome uncover Tobias Black's criminal network:
- Focus: Exposing a sophisticated drug smuggling operation.
- Data Sources: Integrates information from:
    - Vehicle Registry
    - ANPR (Automatic Number Plate Recognition)
    - CDR (Call Detail Records)
    - Financial transactions
    - Social media activity
    - Travel records
    - Others
- Outcome: Reveals hidden connections and patterns, aiding law enforcement in dismantling the criminal organization.

Prepare Interviews

Based on all relevant information the AI proposes a list of interview questions, a line of inquiry or detects inconsistencies in the different statements.

agentAI

The AI agent acts as a investigators and automatically collects data during the pre investigations based on what it deems a important.

Find Similar Cases

Reduce investigative workloads by using context to generate only high-quality alerts, reduce false positives, and reduce investigation time from weeks to hours.

Write Standardized Reports

Augment and enhance in-house capabilities using an open architecture, which enables rapid scaling to accelerate risk detection across multiple public-sector agencies.

Q&A

Thank you.