Unleashing the Power of Collaborative Intelligence with Generative AI, Zero Trust Encryption, and Knowledge Graph Technology

Author: Jack Singer - Head of Business Development for US Government and Defense @ Data²
Contributors: Brian Wane - CEO @ XQ

data²

Every day, the civilian and military personnel working for the United States Intelligence Community (IC) monitor tens of millions of unique datapoints, collecting valuable information to help solve the nation's most pressing intelligence challenges. In 2019, it was estimated that upwards of 4.3 million people are working in roles that require a government issued security clearance, today that number is even higher. (https://sgp.fas.org/othergov/intel/clear-2019.pdf). This number includes the personnel working for one of the 19 agencies that fall under the oversight of the Director of National Intelligence, in addition to the many government contractors working cleared roles within private industry. The IC is without question the most sophisticated and robust intelligence apparatus the world has ever seen. However, even the mightiest of institutions have their Achilles heels, and the IC is no different.

One of the pressing challenges for the IC as it stands today is the age-old problem of "stovepiping." In the context of intelligence work, stovepiping is a term used to describe the many ways in which a lack of clear communication channels and interagency collaborative tools can obstruct the creation of quality intelligence products for warfighters and policymakers. High-priority integrated intelligence tasks require rapid, collaborative action. Despite the urgency and importance of their mission, when major events require multiple agencies, directorates, and personnel spanning all clearances to collaborate on the same mission, the IC can find itself making the difficult decision between sacrificing the quality of their intelligence products or waiting too long to deliver intelligence products to the stakeholders that need it most.

The next step of the Intelligence Community's evolution as an institution is the elimination of stovepiping by any means necessary.

In order to best accomplish their strategic objectives, intelligence agencies need a tool that can accelerate the intelligence collection and analysis cycle while also maintaining the source data's individual security requirements. Data Squared's reView platform and XQMsg's Zero Trust encryption capabilities combine to provide a comprehensive integrated intelligence suite that puts the unparalleled power of Zero Trust and Graph Generative AI technology directly into the hands of the IC's analytic brass. Data Squared's leadership in Generative AI (GenAI) for traceability and explainability, combined with XQmsg's advanced secure communication capabilities represent a pioneering force capable of transforming how intelligence is shared and analyzed within the IC.

With reView and XQ, different personnel holding different clearances from different agencies can all collaborate seamlessly on the same Intelligence challenges without sacrificing the quality or timeliness of their intelligence products.

For this white paper, we will use a synthetic dataset generously provided by Ed Waltz, a professor of Strategic Intelligence at Patrick Henry College. This synthetic dataset mirrors the characteristics of real data used in a targeting nomination briefing before a real targeting board. During the target nomination process, analysts must develop a list of well profiled targets alongside their personal connections, group associations, location, and pattern of life activity, to deliver to the targeting board. The Joint Special Operations Command defines a target as:

> "an entity (person, place, or thing) considered for possible engagement or action to alter or neutralize the function it performs for the adversary. A target's operational importance is determined by conducting an assessment to determine if engaging the target is consistent with planned operations and will help achieve the commander's objective(s) and the end state."The emphasis of targeting is on identifying resources (targets) the enemy can least afford to lose or that provide him with the greatest advantage (high-value target [HVT]), then further identifying the subset of those targets which must be acquired and engaged to achieve friendly success (high-payoff target [HPT]). Targeting links the desired effects to actions and tasks.

(https://jfsc.ndu.edu/Portals/72/Documents/JC2IOS/Additional_Reading/1F4_ jp3-60.pdf)

In an intelligence activity such as targeting, each agency must have access to the same comprehensive set of data (with respect for clearance considerations and need-to-know bases) to meaningfully contribute towards IC objectives. This article will demonstrate how Data Squared's reView platform and XQmsg's zero trust encryption suite can enable cutting-edge collaboration between multiple intelligence agencies with highly-sensitive military and intelligence-related workloads.

What is reView?

The reView platform is a tool that uses Graph Artificial Intelligence Analysis (GAIA) technology to surface specific insights from a large body of evidence with speed and precision. The reView platform has three distinct subordinate functions that help power the tool itself:

Evidence Loader

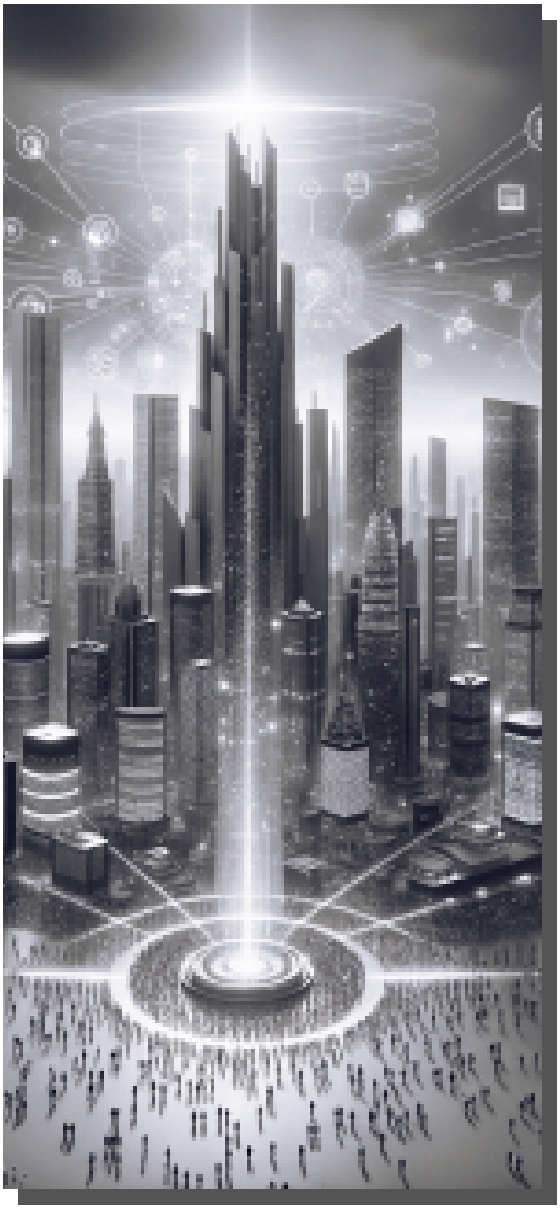
Loads the raw data provided by the user into the reView tool.

Evidence Enricher

Runs a standard set of queries against each datapoint (or node) in the graph, 'enriching' each node with additional context in order to draw connections between nodes.

AI Reasoner

Receives user questions which are passed to the Large Language Model (LLM) of the user's choice. The LLM will answer the user's questions, provide citations to specific nodes on the graph, and then store the answer that was generated back into the graph for further enriching. The reasoner bases its insights only on the data provided by the user, a design choice that is proven to mitigate the risk of hallucinations in AI reasoned answers to user questions.



reView is a tool designed to learn and evolve alongside the user. Because of the modular architecture of reView, the tool itself becomes more powerful as the user continues to engage with its AI Reasoner. Each question asked by a user will grow the AI Reasoner's subject matter expertise in your industry. The more a user interacts with reView, the more tailored and useful the answers reView generates will become.

What is XQ Encryption?

XQ offers a Zero Trust Architecture-based data protection platform, ensuring data is safe wherever it travels. It provides data provenance, residency, compliance, and logging, making it ideal for data lakes, digital twins, IoT, and critical infrastructure. XQ's technology keeps everything fully compliant with even the strictest DoD security requirements.

XQ encryption focuses on zero trust architecture wherein it assumes that every attempt to access data, whether from within or outside the network, should be verified and authenticated. XQ also allows for data encryption and access control. Encryption keys are generated on edge systems, ensuring that data is encrypted at the source. These keys are managed by XQ's key cache, which handles policy- based key management and authorization without ever handling the data itself. This ensures secure data transmission by providing each data transaction with a unique key, rotating it using traditional encryption methods.

reView's integration of XQ's zero-trust technology allows secure file management and user verification processes, bolstering cybersecurity measures within collaborative environments. This setup has been tested and verified, ensuring reliable performance across development and production settings.

XQ Compliance Suite

- GDPR
- CMMC 2.0
- CUI
- HIPAA
- FERPA
- CJIS
- ITAR

In order to demonstrate how an XQ protected reView instance can deliver an intelligence advantage for its users, let's turn our attention to the reView platform itself.

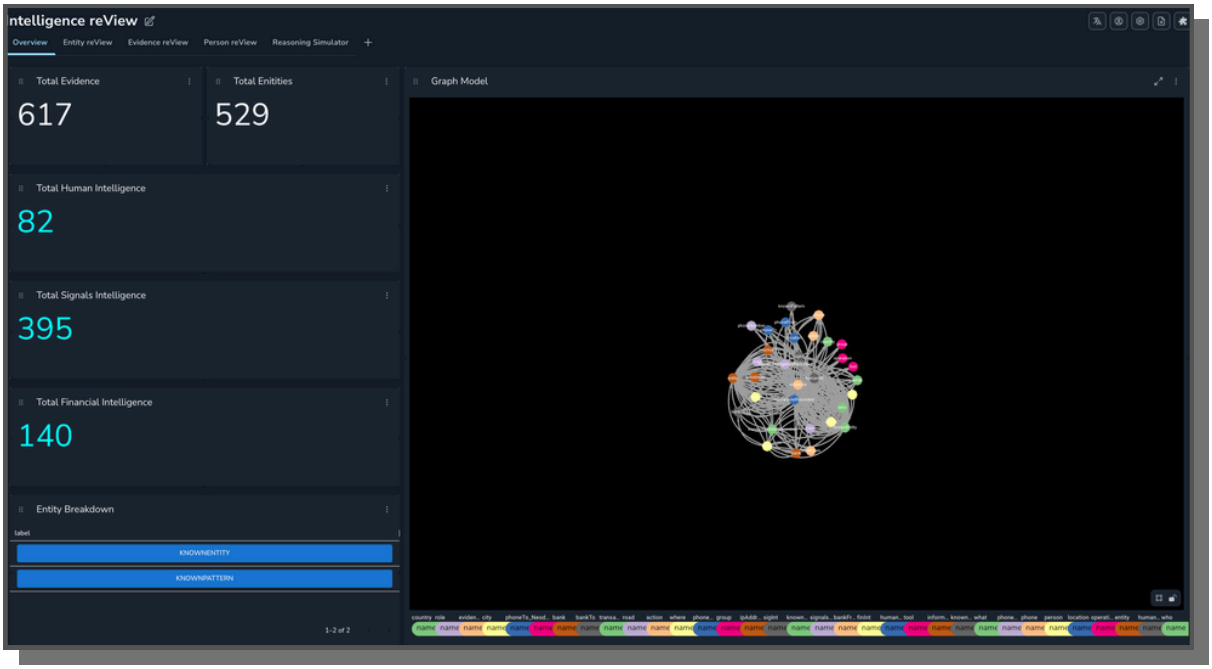
reView Functionality Guide

The reView tool is divided into five tabs that support 5 different functional components.

- Overview
- Entity reView
- Evidence reView
- Person reView
- Reasoning Simulator

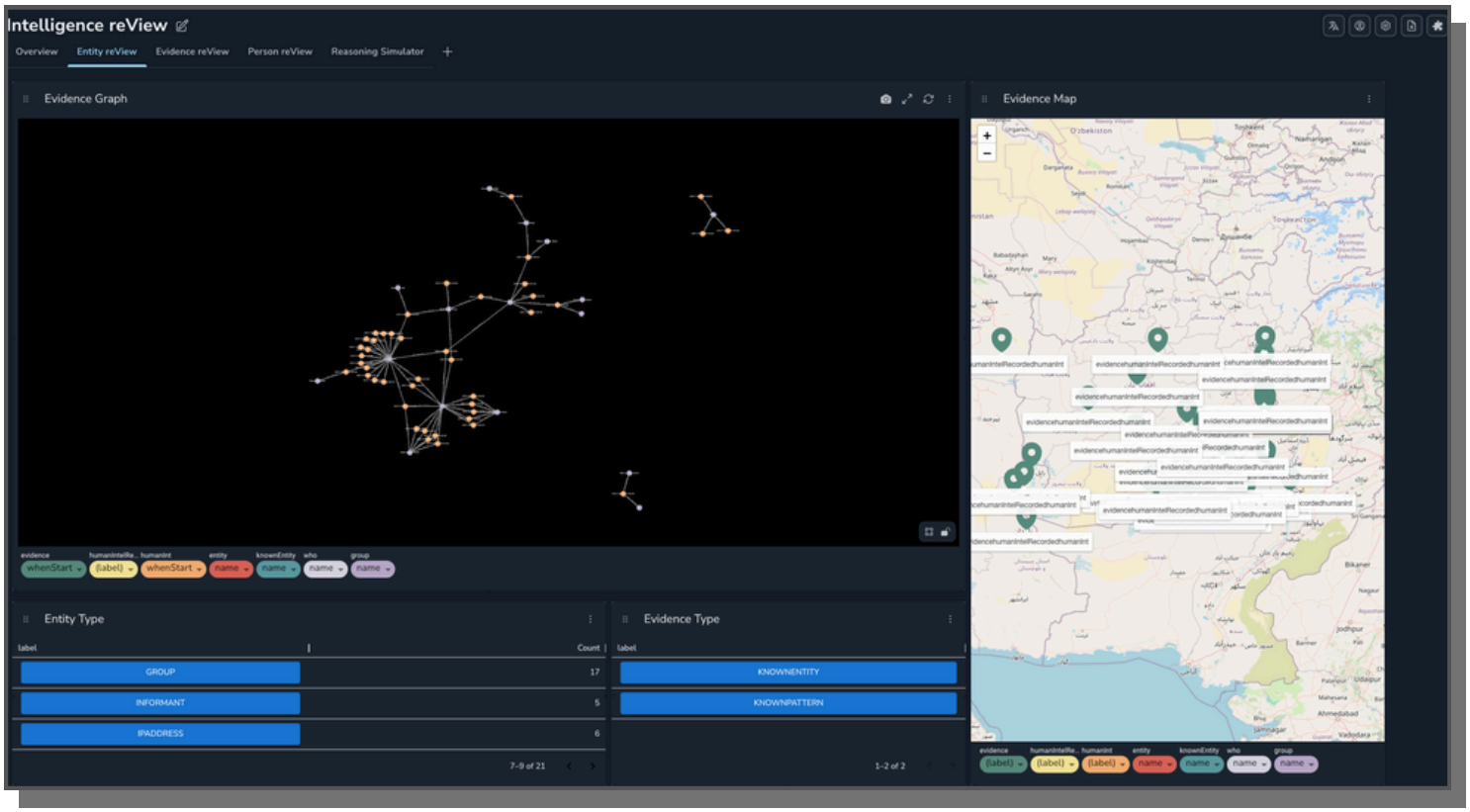
Each of these components drive reView's analytic engine and can be engaged with using embedded Knowledge Graph Data Models (KGDMs) within reView.

Entity reView



Shown above is the Overview tab of the reView tool. The aforementioned synthetic dataset has already been loaded into reView. The contents of the dataset itself are ultimately unimportant and have only been included in order to better illustrate reView's analytic functionality. A summary of the data loaded into reView is displayed on the lefthand side of the screen. On the right is a graph model displaying the entity and evidence types that were pre-generated by the reView Evidence Loader. This dataset's data is limited to three intelligence types (HUMINT, FININT, SIGINT) while the reView platform itself can load any type of structured or unstructured data, including high frequency and telemetry data. The connections between nodes were then generated by the reView Evidence Enricher.

Entity reView



This is the Entity reView tab. Before importing any data into the graph, reView's Evidence Loader divides each datapoint into two categories:

Entities

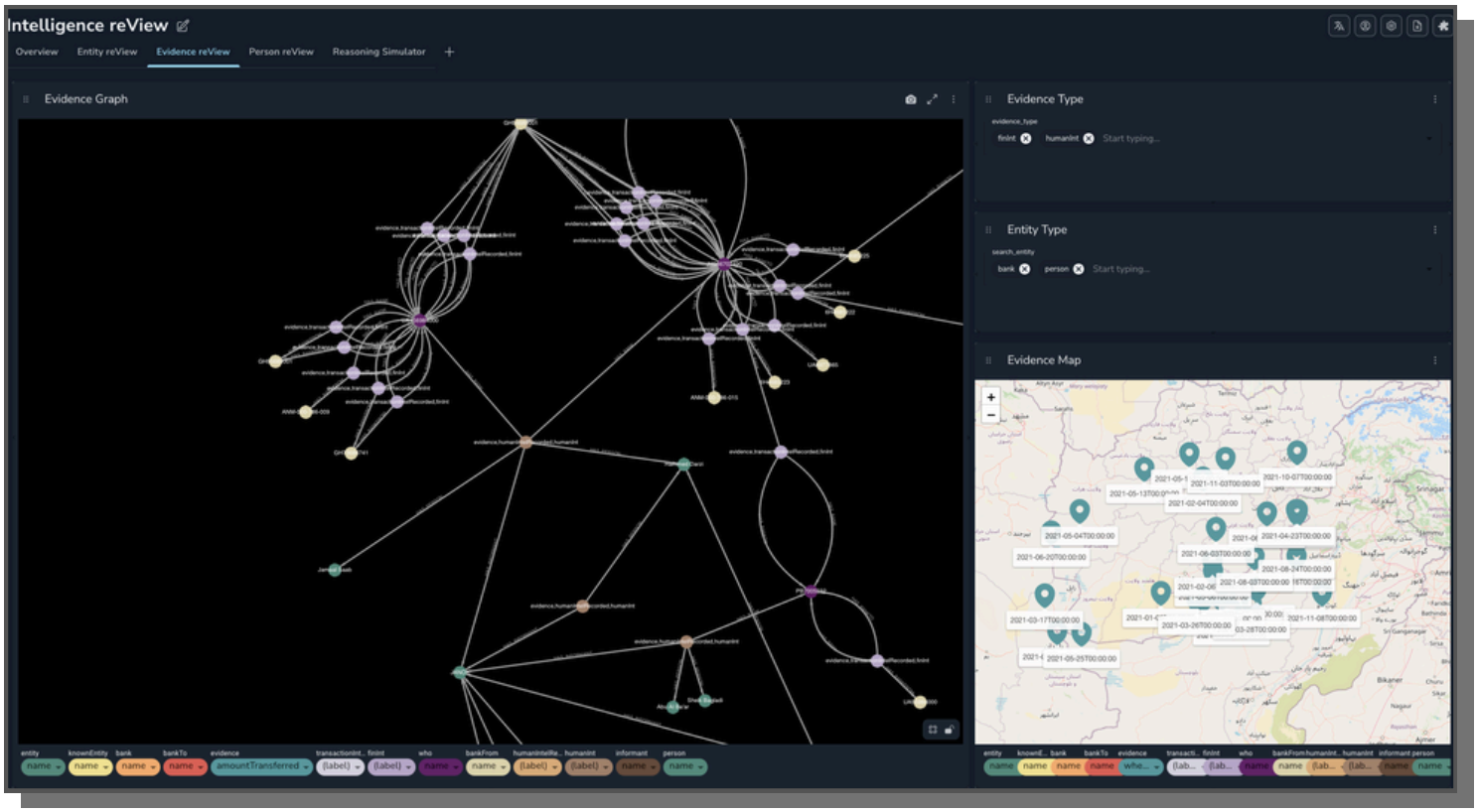
- People
- Vehicles
- Groups
- Countries
- Equipment
- Phones
- Banks
- etc.

Evidence

Reports that mention entities.

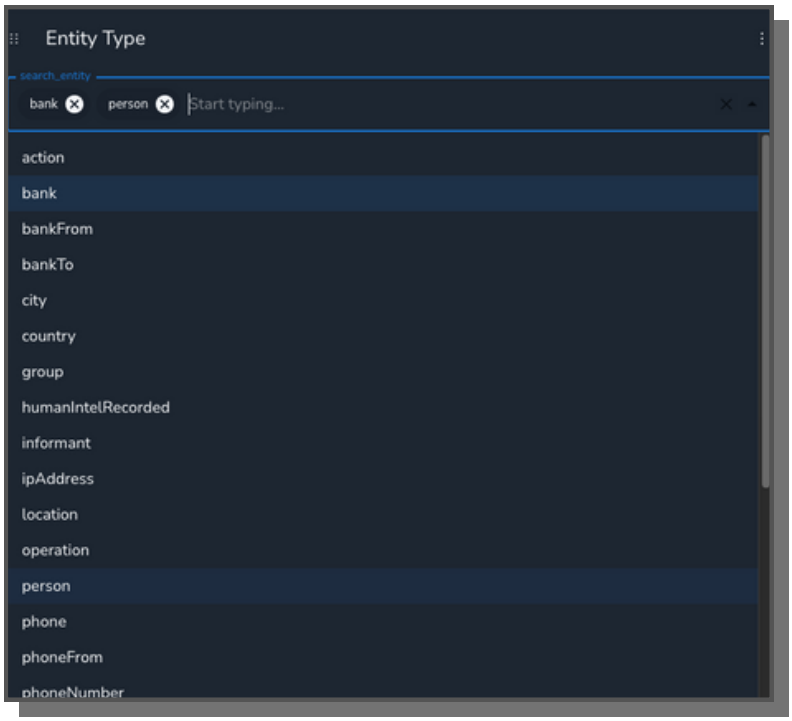
Upon separating the entities from the evidence, a user can choose from a list of entity types in the bottom left of the screen. For example, if a user were to click the **'Group'** button, it would inject all entity nodes labeled **'Group'** into the above knowledge graph. In addition to displaying the corresponding entities in the graph, the graph will also load each evidence report mentioning at least one of the groups into the graph. Some reports mention multiple groups, the interface illustrates this through the lines that connect each evidence node to the entity node of any groups mentioned. There is also a map included on the righthand side of the screen that displays evidence nodes on a map based on the Latitude and Longitude information contained within the report.

Evidence reView



This is the Evidence reView tab. The top right of the screen includes two filter selection fields. For the 'Evidence Type' field, a user can select one (or all three) of the intelligence types in this dataset. In the 'Entity Type' field, a user can choose from a variety of different entity options.

By experimenting with the parameters in the Evidence and Entity fields, a user can craft an entirely unique perspective with which to explore the data they have imported into reView. If a user is curious to investigate what connections (if any) exist between the phone call logs and financial transaction records, they can load the 'phone', 'bank', and 'person' entities into the graph. This set of entities would allow a user to visualize the overlap between SIGINT and FININT nodes within the graph in real time

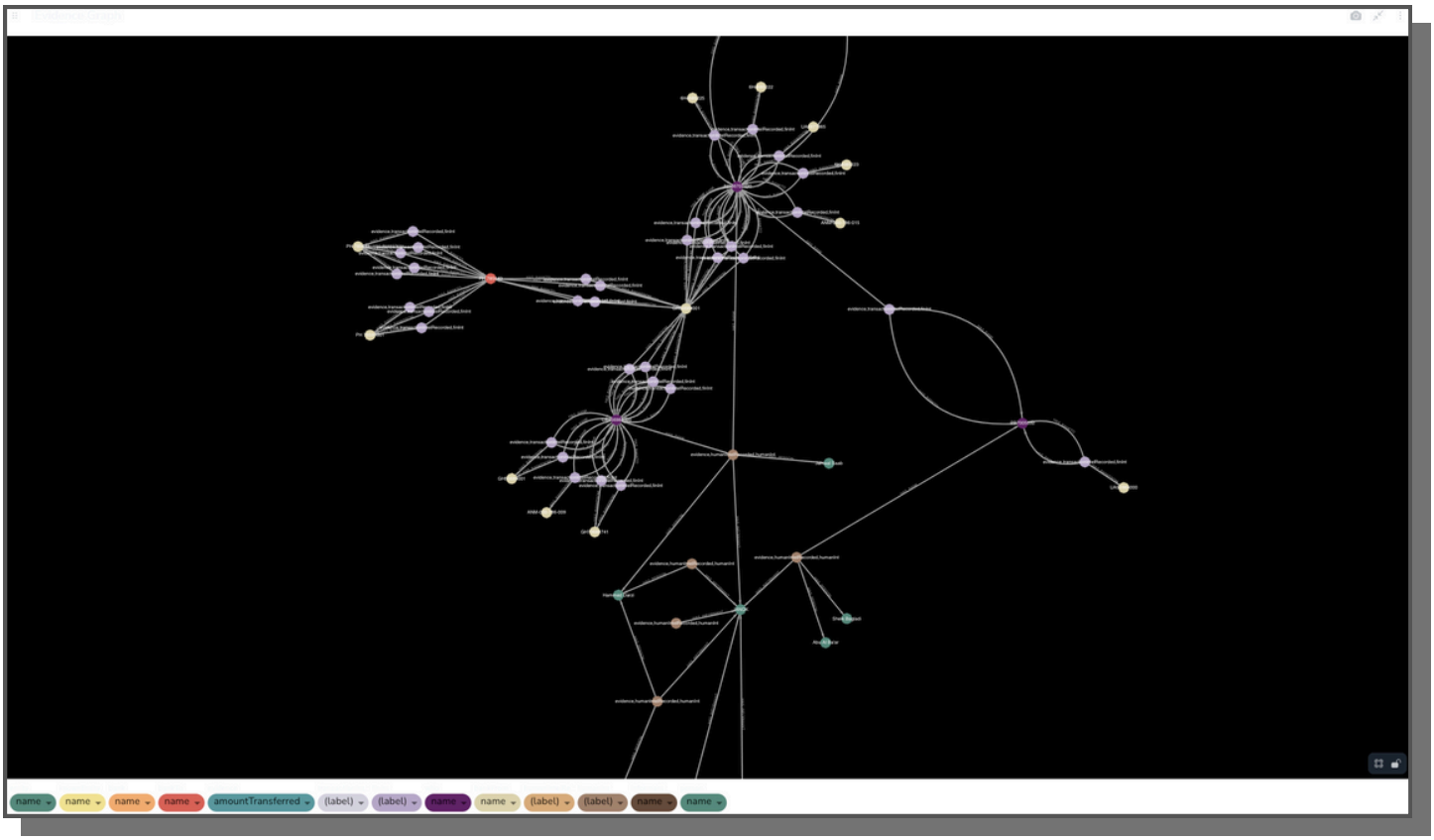


EXAMPLE USE CASE: Identifying Money Laundering Activity Using Graph Visualization

Within this synthetic dataset is a storyline designed by Professor Ed Waltz that is meant to be discovered by students taking his 'Advanced Analytic Targeting Methods' class using legacy network modeling tools such as IBM's Analyst NoteBook 2. The story itself centers around a fictional confidential HUMINT source, codenamed JANOK, who was killed in the Afghanistan/Pakistan region in 2021. Students enrolled in Professor Waltz's class were required to analyze all 1146 datapoints to answer the question "Who killed JANOK?" and recommend targets for additional collection, interrogation, and in some cases, elimination.

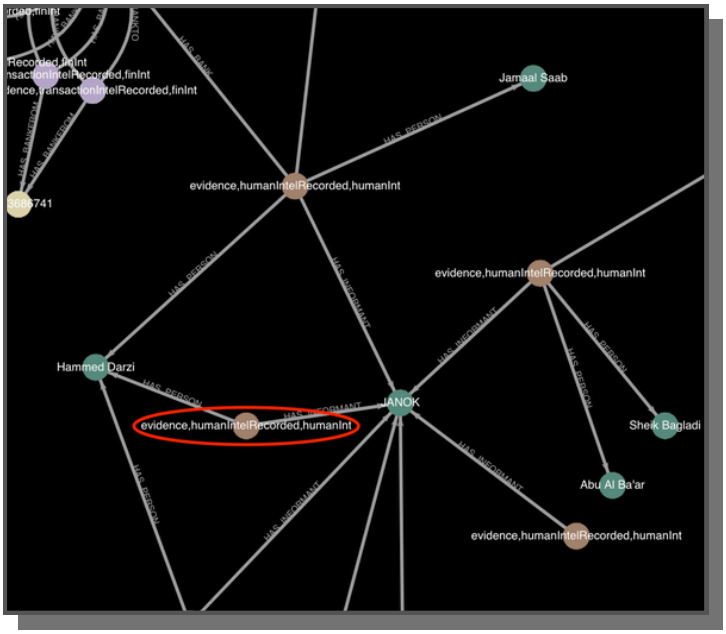
One of the analytic approaches recommended in the investigation of JANOK's death is an in-depth analysis of financial transaction logs retrieved from banks in the Afghanistan/Pakistan border region managing accounts owned by persons of interest to the United States' investigation.

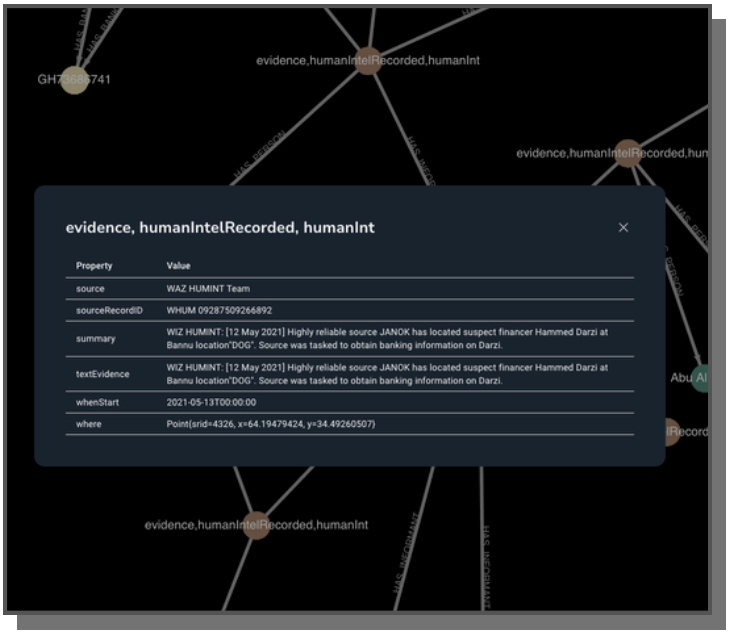
If an analyst wanted to pursue this analytic approach, they can start by loading HUMINT and FININT in the Evidence Type field. Upon loading **'bank'** and **'person'** into the Evidence reView graph, the analyst will be presented with a graph that looks something like this:



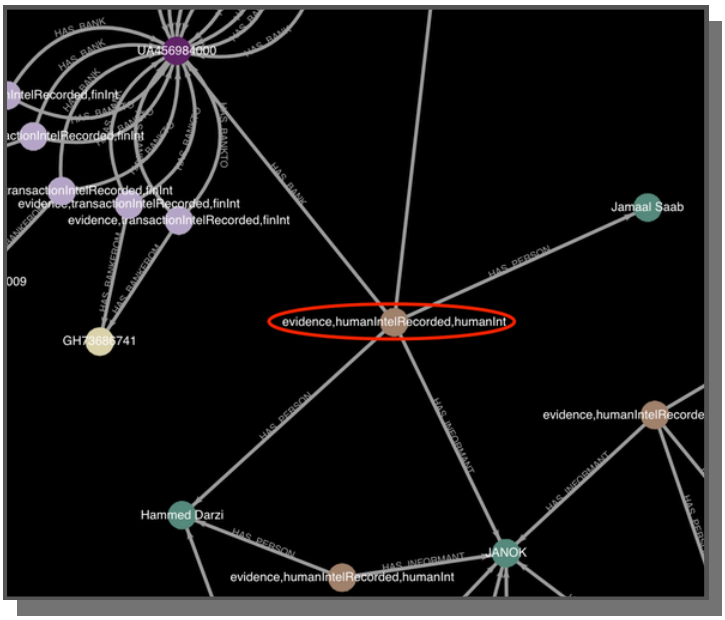
This graph, containing every person and every bank account in the dataset, gives an analyst a perfect starting point for investigating who might have financed the murder of JANOK, in addition to the other terrorist attacks reported in the area.

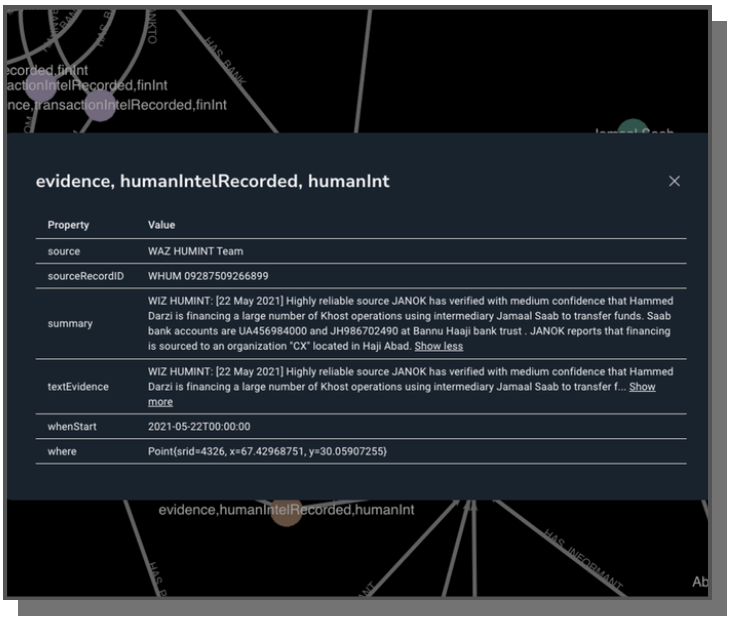
We can start by looking at the report shared by JANOK and a person named Hammed Darzi:





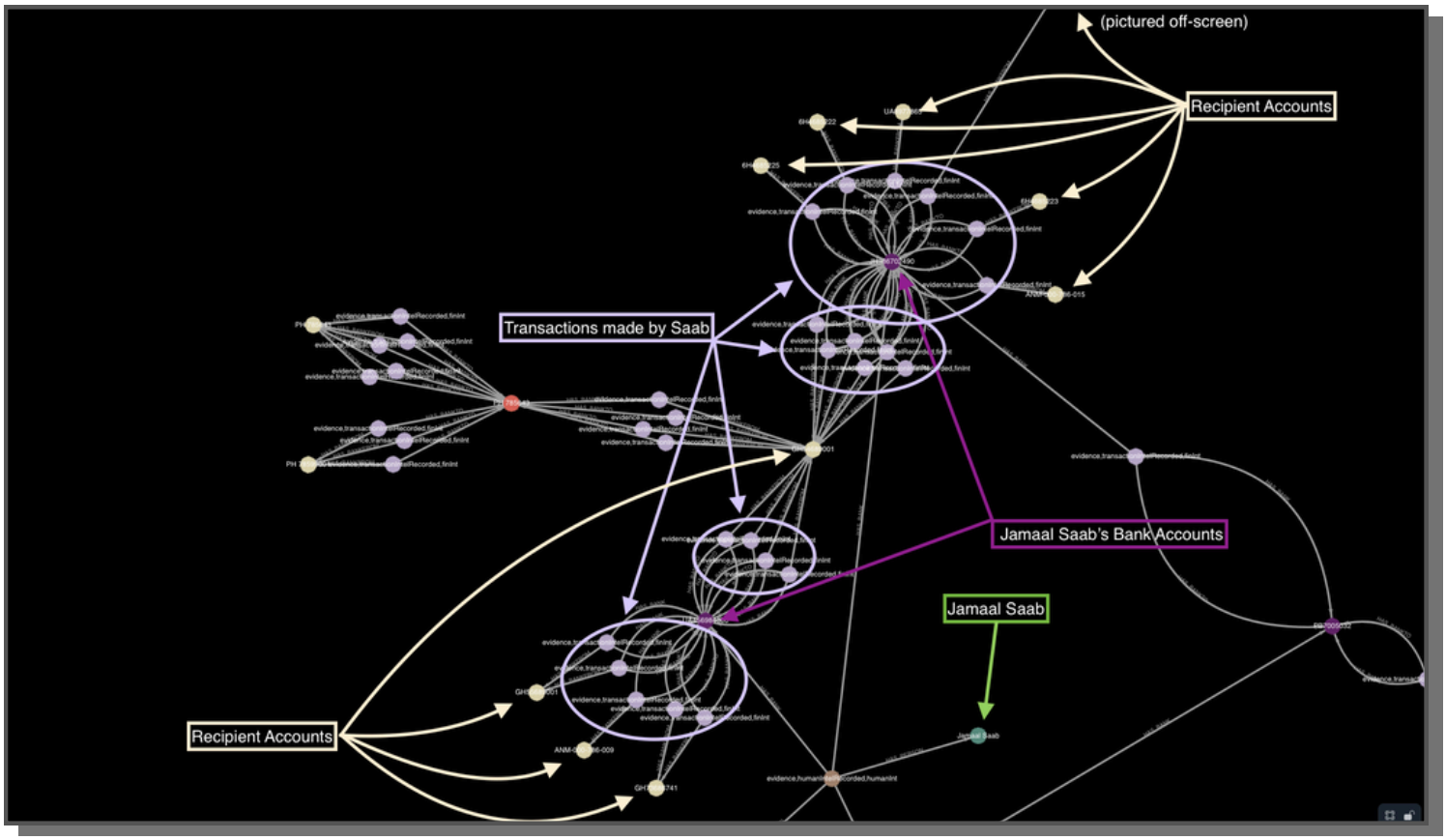
This HUMINT report seems to suggest that JANOK was tasked with gathering Darzi's banking information in the days leading up to his death. We can then turn our attention to the other report node shared by JANOK and Darzi:





The second report tells us that JANOK found information tying a "Jamaal Saab" to Darzi's terror financing operation. The report goes on to say that Darzi uses Saab as an intermediary to transfer funds to terror cells in the region with allegiance to a group known as "CX" located in the city of Haji Abad. The report also includes bank account selectors known to be associated with Saab. As an analyst, this would direct our attention to the two bank accounts owned and operated by Saab on behalf of Darzi. Those accounts, in turn, are sending money to other accounts creating a complex and sophisticated network of transactions tracing back to one person, Jamaal Saab.

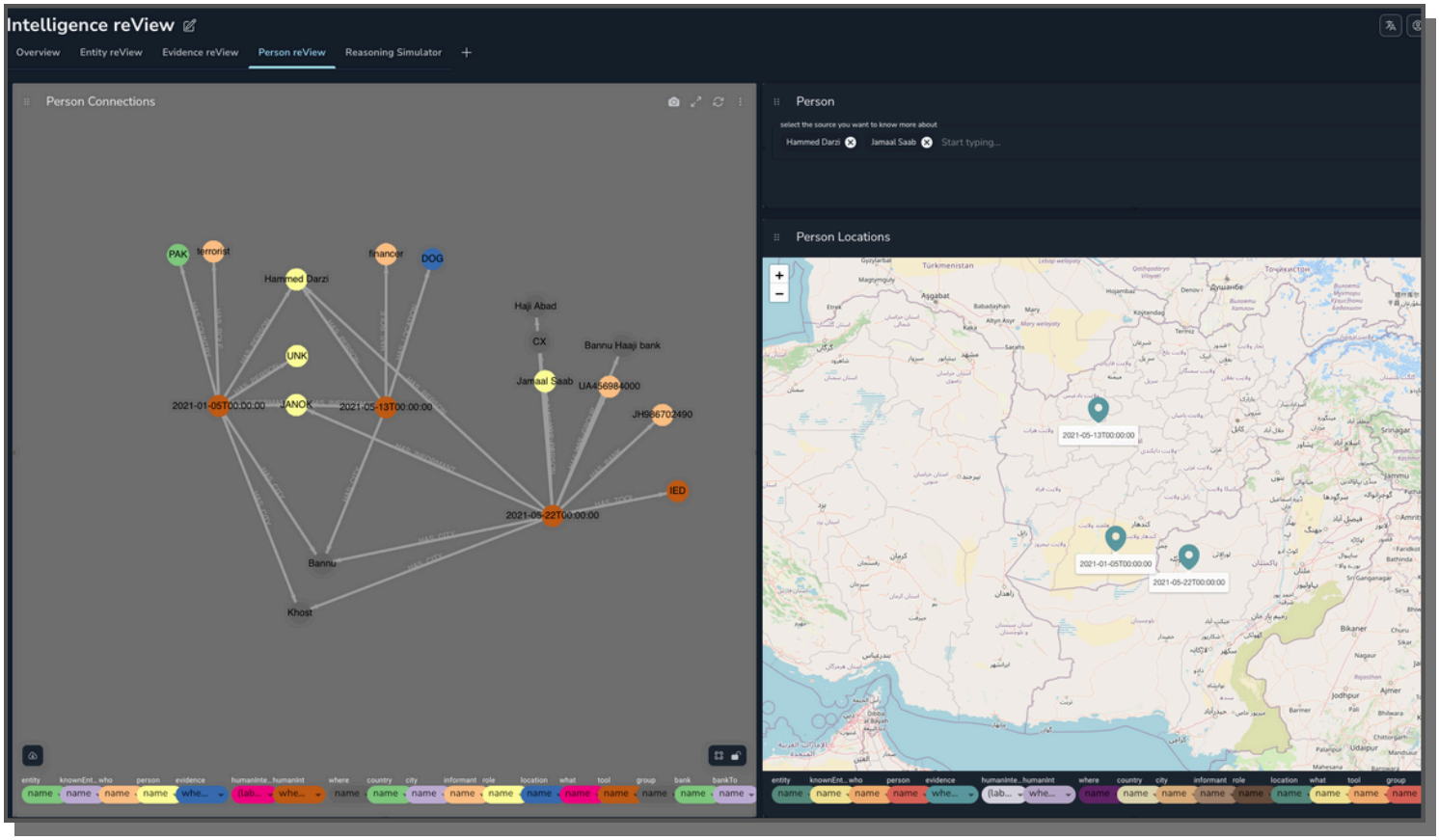
Below is an annotated version of the graph calling out vital information pertaining to Jamaal Saab's money laundering activity:



In the top right of the above image, there is an annotation that reads "(pictured off-screen)." If we take our current view of the graph and zoom it out, we'll get a sense of the scale of this money laundering operation:

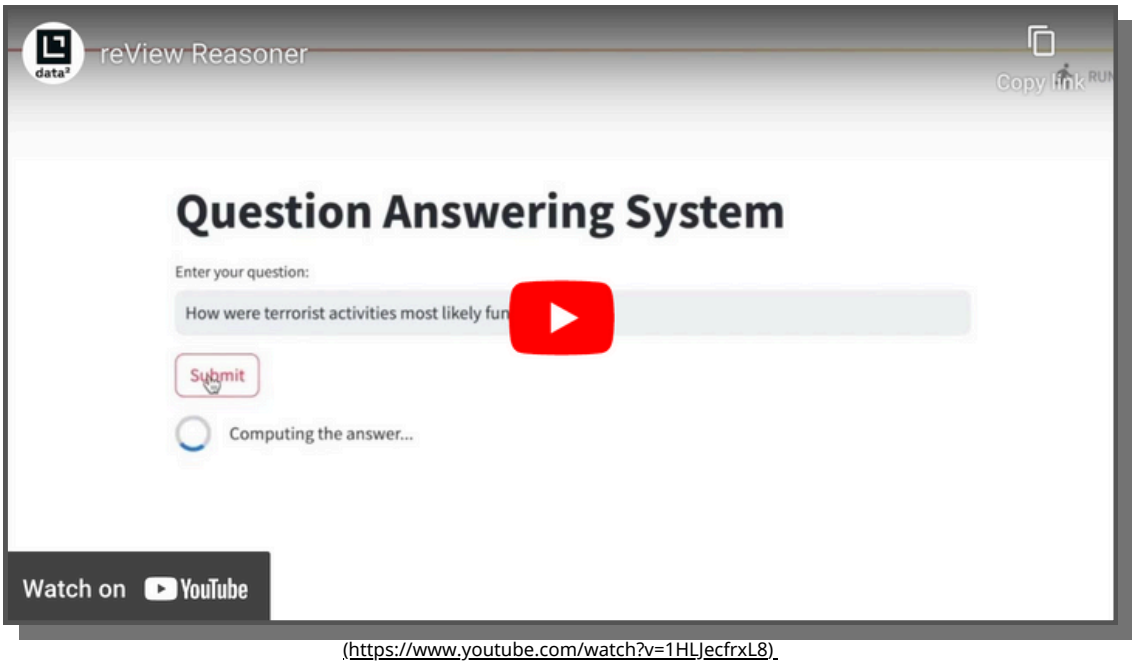
The red box is meant to illustrate the perspective of the above image in light of the larger knowledge graph. The money moved by Jamaal Saab changes hands several times before reaching its target bank account. Without the ability to load and graphically represent AI enriched data in real time within reView, the information included in the images of the reView tool would have to be viewed as static cells in a .csv file or manually loaded into a legacy network modeling tool. With the power of AI enriched graph visualizations, our analyst was able to easily determine that Hammed Darzi is the person responsible for terror financing in the region, and thus is prime suspect in the investigation of the death of JANOK.

Person reView



This is the Person reView tab. On the top right of the screen is a single field where you can enter the names of multiple people and see what connections they share. Person reView serves as a sort of 'magnifying glass' when compared to the Evidence reView tab. Users can interact with the Evidence reView graph, find new insights or connections, and then load any relevant persons from Evidence reView in the Person reView tab for closer inspection. To build upon the prior money laundering example, we can load Hammed Darzi and Jamaal Saab into person reView, and generate a graph of all of the nodes that they share. Person reView gives us yet another way of visualizing the intricate ways in which specific datapoints interact within a dataset.

Evidence reView



(https://www.youtube.com/watch?v=1HLJecfrxL8)

The final function tab in Intel reView is reserved for the AI Reasoner and Question Answering System. The reasoner (shown in link above) functions using the LLM to reason over the knowledge graph containing all the data in a dataset. The LLM will then retrieve information it reasoned over to answer a user's question in natural language. reView's GenAI Graph framework enables its AI Reasoner agent to return complex and factual insights based ONLY on the information provided, it will not infer or make assumptions about information that is not accessible within the dataset it was provided. This approach to reasoning ensures hallucinations are mitigated and reView's AI insights can be trusted. As a user feeds the AI information, reView will become a "subject matter expert" on the intricacies of the user's problem set and be able to deliver in-depth analysis of the data at its disposal in an easy to navigate format.

Data Squared's hallucination resistant graph algorithms and XQ's data provenance and insight capabilities enable users to see data origin points when they generate AI insights. This approach to AI transparency prevents data poisoning that can derail reView's AI Reasoner insights by generating auditable reports detailing the chain of custody for each individual file entering the knowledge graph.

Concluding Thoughts

In a world where the minutes between crisis and action can be the difference between life and death, intelligence professionals require tools that enable seamless collaboration across multiple agencies and stakeholders. When the United States is challenged by its enemies, whether foreign or domestic, the Intelligence Community needs its best minds analyzing the best information at a moment's notice, reView and XQ makes that possible.

Data Squared and XQ's strategic partnership holds groundbreaking transformative potential for the defense and national security spaces. Together, Data Squared and XQ are helping envision a future where intelligence agencies can swiftly and securely access, share, and analyze intelligence information, empowering them to make informed decisions with unprecedented speed, certainty, and control. XQ's industry leading security and compliance features, paired with reView's advanced intelligence analysis capabilities, paves the way for unparalleled decision superiority for the United States Intelligence Community.

About the Authors

Jack Singer is Head of Business Development for US Government & Defense for data² and is based out of the D.C. Metro Area

Brian Wane is Chief Executive Officer at XQ and is based in the San Francisco Bay Area.