

Red Hat
Summit

Connect

Fraud Detection:

use case innovativi e approcci scalabili
con **Galileo.XAI**

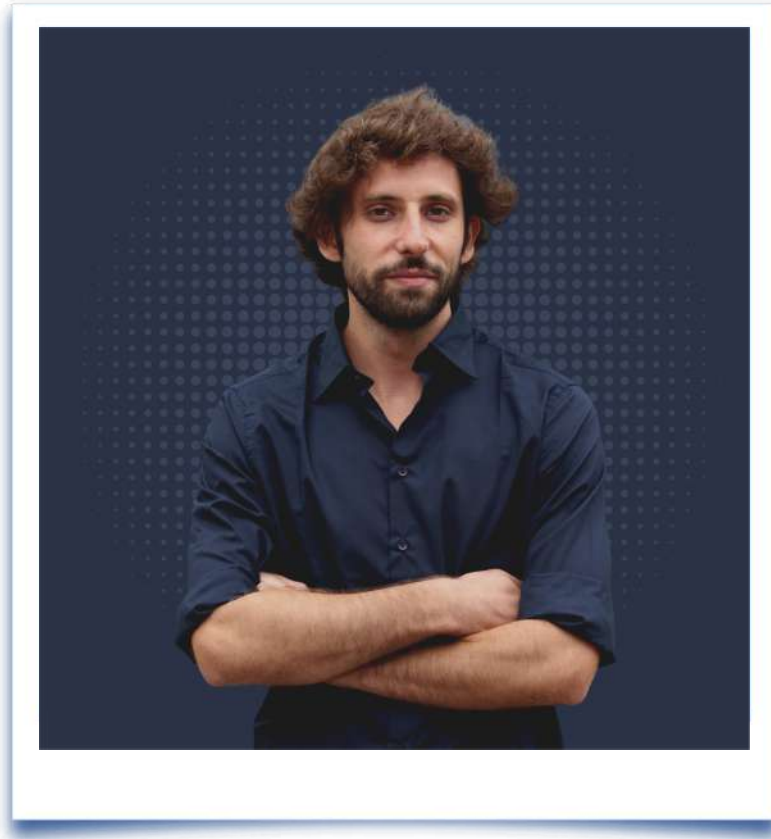
Roberto Sannino, Product Owner
LARUS Business Automation

Milano, 19 novembre 2024



Agenda

- LARUS
- Fraud Detection
- Galileo.XAI
 - Soluzione graph based per l'Explainable AI
 - Casi d'uso di successo per la Fraud Detection
- Q&A



Roberto Sannino

Solution Lead @LARUS

Software Architecture
Product Lifecycle Management
Full-Stack Development

roberto.sannino@larus-ba.it



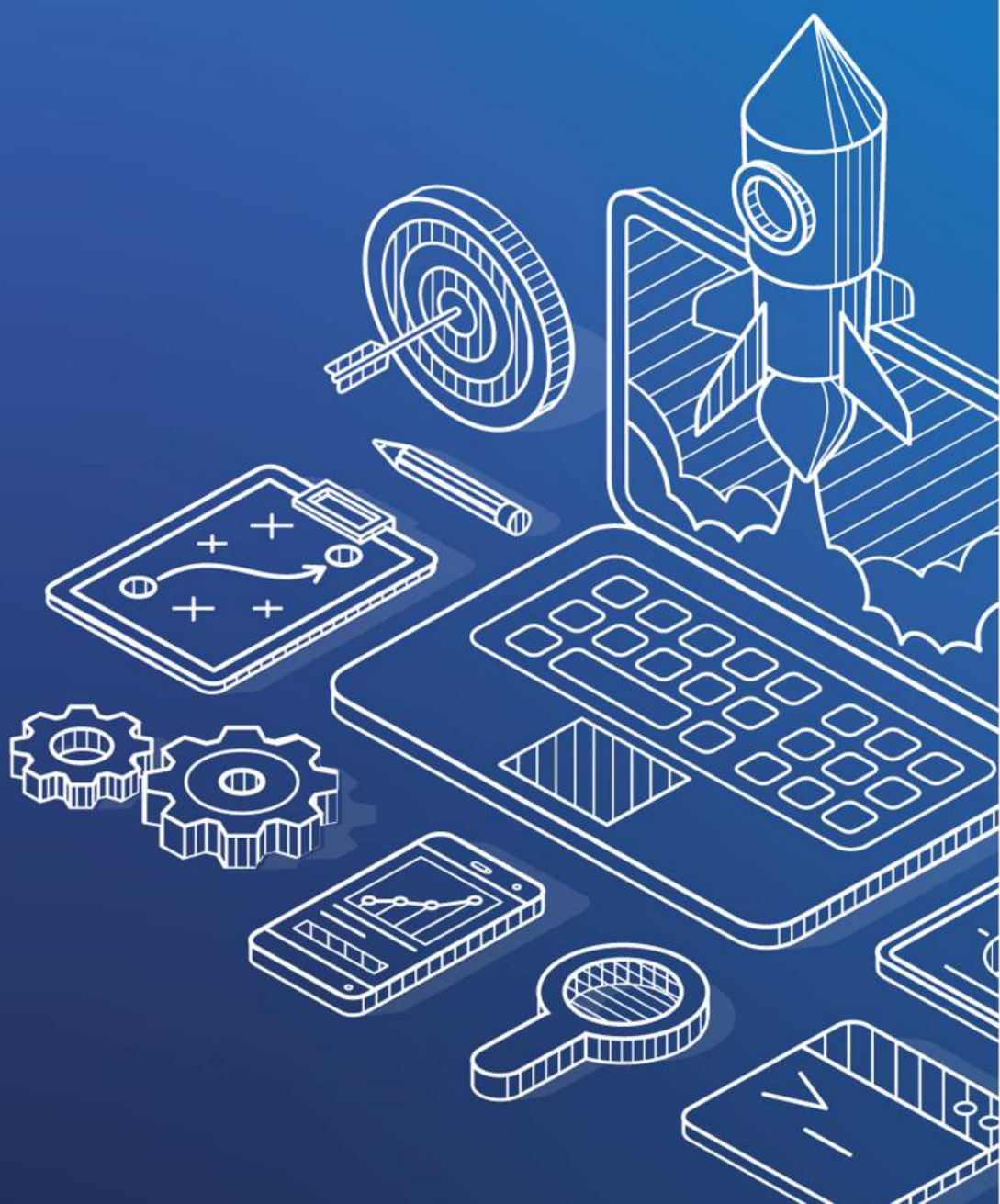
LARUS



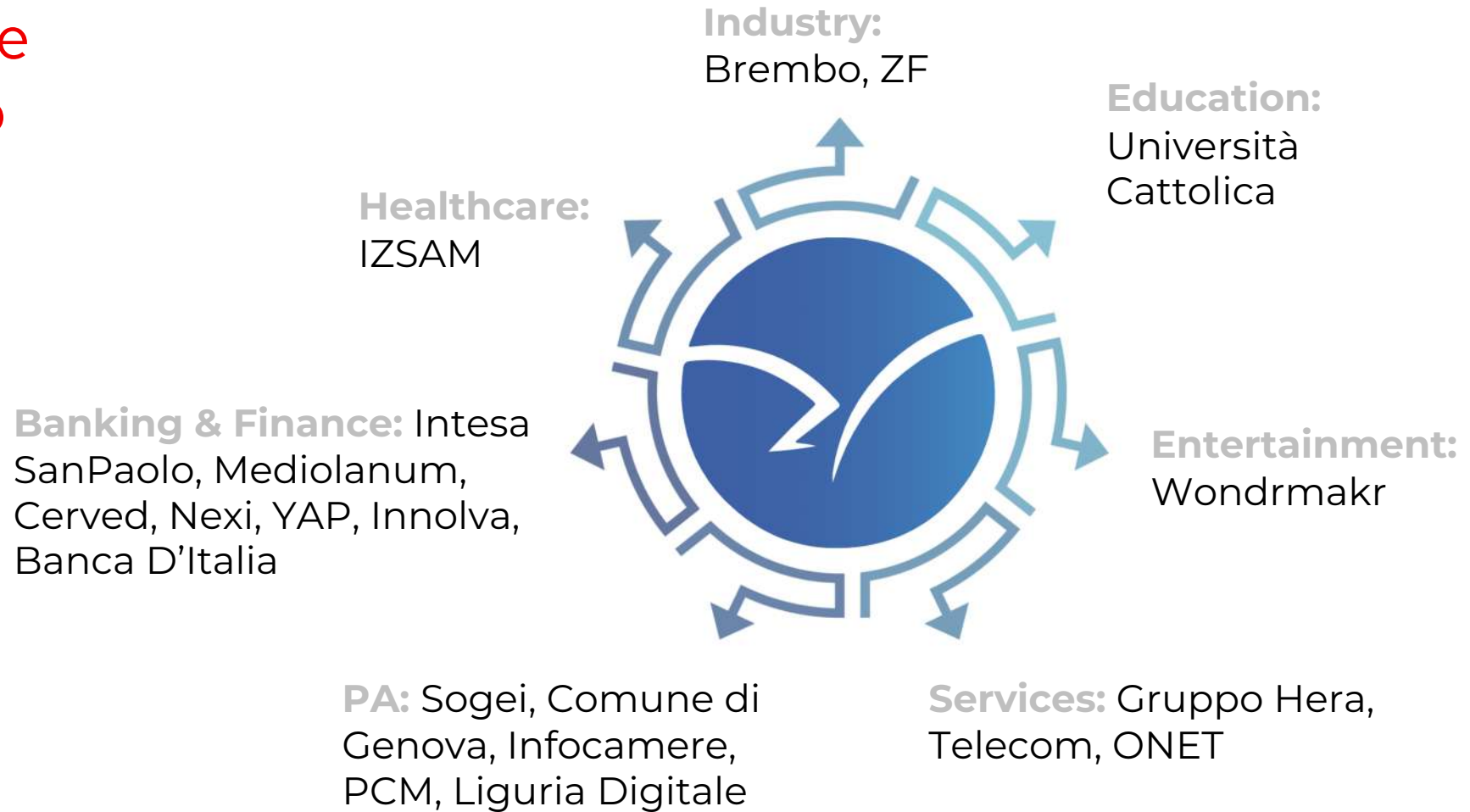
LARUS è una **PMI innovativa**, orientata, attraverso l'approccio **AGILE**, verso lo **sviluppo**, la **consulenza** e la **formazione** su **tecnologie e approcci innovativi**.

Grazie alla consolidata esperienza maturata con database **NoSQL**, nel campo dell'**Explainable AI** e della **Connected Data Science**, aiuta i clienti ad ottimizzare il loro processo decisionale, aumentare i profitti e rimanere competitivi.

Grazie alla divisione R&S e Labs ha intrapreso, inoltre, una **proficua collaborazione con il mondo accademico** nazionale ed internazionale per essere sempre al passo sul piano della ricerca.



I clienti che si affidano a LARUS





Fraud Detection



Ambito

Il contesto della Fraud Detection è incredibilmente complesso, in genere la percentuale di **frodi effettive in un dataset** è inferiore all'**1%**, ma dato l'alto numero di transazioni/incidenti registrati ogni giorno, l'impatto in termini economici può essere devastante se non si dispone di strumenti di contenimento.

Scovare l'1% di frodi in un dataset è come trovare un ago in un pagliaio, con la differenza che l'ago in questione si nasconde attivamente: i frodatori sono intenti a nascondere i loro pattern perchè risultino standard e **le tipologie di frode** messe in atto **evolvono costantemente**, pertanto non è possibile basarsi totalmente su pattern passati e regole di business.

L'attività di Fraud Detection quindi, richiede l'identificazione di pattern sospetti non solo in maniera statica, manuale e a posteriori, ma anche grazie a modelli predittivi, automatizzati e real-time.

Alcune Tipologie di Frode

Ambito Assicurativo

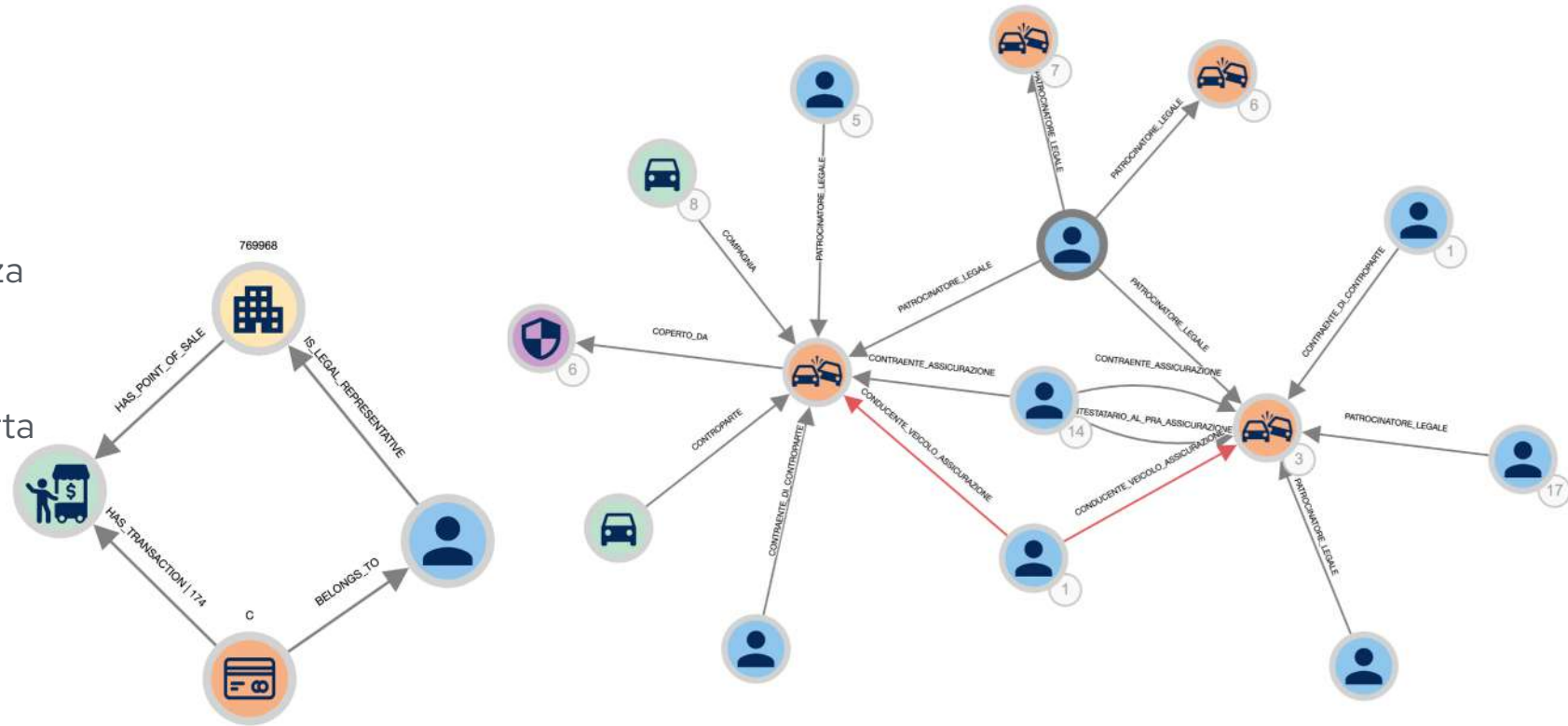
- Sinistri falsi o danni gonfiati
- Fraud Rings
- Modifica retroattiva della polizza

Ambito Bancario

- Furti di identità/credenziali/carta
- Transazioni sotto soglia
- Caroselli

Punti in comune:

- Documentazione falsa
- Informazioni fittizie e/o rubate
- Pattern di azione che eludono regole di business
- Schemi sociali e di interazione complessi ed in continua evoluzione



Sfide

Il **danno da frodi**, per le sole aziende assicurative americane nel 2023, si aggirava intorno ai **308 miliardi di dollari**. Le aziende assicurative, di digital payments, le istituzioni bancarie come anche i siti di e-commerce, devono e dovranno affrontare sfide sempre più difficili.

Da una parte, l'alta competitività del mercato, spinge le aziende ad erogare servizi sempre più rapidi ed accessibili, per garantire la posizione di mercato. Di contro, sempre il rapido accesso a servizi di pagamento e reclami, aggiunge **livelli di criticità aggiuntivi** in ogni step dell'erogazione che devono essere affrontati **senza impattare sulla soddisfazione dei clienti**.

Un'ulteriore sfida è quella di contenere i costi dell'unità anti-frode massimizzando i risultati ottenuti.

Due macro-sfide fondamentali:

- **Alta velocità** del processo di verifica senza compromessi sul rischio.
- **Efficienza** dei processi anti-frode - costi contenuti, analisi automatiche e basso numero di falsi positivi.



Riflessioni

L'ambito della Fraud Detection è un **ambito fortemente connesso**, reti di utenti interconnesse creano complesse reti di informazioni, scambi, ruoli e transazioni. Sistemi di **Network Analysis** per il rilevamento di Fraud Rings risultano **essenziali**.

Le aziende del settore operano su larga scala, con un numero di utenti elevato ed un **alto numero di operazioni al secondo**. Le tecnologie convenzionali, come i database relazionali, fanno fatica a gestire questi numeri e a proporre una modellazione/rappresentazione del fenomeno efficiente.

L'alta velocità con cui i pattern fraudolenti mutano, rende necessario l'uso di **sistemi predittivi che tengano in considerazione le relazioni tra i dati**, non sono i dati di una transazione ad essere fraudolenti o meno, è il contesto in cui è inserita ad essere fondamentale.

La tecnologia evolve rapidamente (pensiamo alla GenAI), le nuove tecnologie non aiutano solo le aziende a migliorare i propri processi, anche i frodatori possono sfruttarle. E' fondamentale che le **soluzioni** siano **facilmente integrabili ed estendibili**.

La Nostra Soluzione - Approccio a Grafo

Per rendere performanti le operazioni di Network Analysis in un contesto estremamente connesso e sfruttare le informazioni di relazione tra le entità, i **Graph Database** sono lo strumento più adatto.

Memorizzando fisicamente le relazioni, assicurano **query di pattern matching in real-time** e prestazioni superiori per gli algoritmi di Data Science, utili per Scoring e Community Detection.

Inoltre questo contesto arricchito dalle relazioni e dai risultati degli algoritmi, permette di sfruttare le **GNN** per definire modelli predittivi che si adattano meglio alla costante evoluzione dei pattern fraudolenti.

Il lavoro degli operatori antifrode fa ancora troppo spesso leva su metodologie ed interfacce inadeguate, l'**approccio a Grafo** risulta vincente anche in termini di **visualizzazione** ed analisi, rendendo **facile ed intuitiva** la comprensione dei fenomeni.

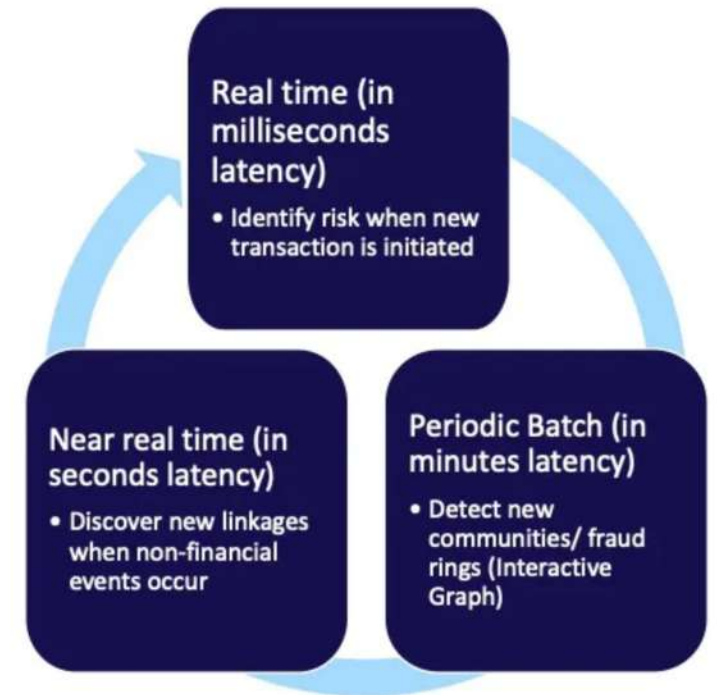
Cosa fanno i grandi - Caso PayPal

PayPal ha creato una **piattaforma basata su Database a Grafo** che ad oggi gestisce circa **9 PetaBytes di dati**, che vengono acceduti in near-real e **real-time**.

Il loro flusso si basa su 3 servizi:

- Un servizio batch che utilizza algoritmi di Network Science per scoring, community detection, link prediction e Fraud Rings Detection - gli operatori anti-frode inoltre taggano i ring sospetti (Human in the Middle).
- Un servizio near-real-time su dati temporali che immagazzina gli eventi di login, registrazione, transazioni, controversie, ...
- Il servizio di query real-time che interagisce con il modello di AI istruito col primo servizio e con i dati temporali per bloccare le transazioni sospette.

Grazie a queste tecnologie e metodologie, PayPal ha **ridotto di 30 volte i Falsi Positivi** e del 98% l'esposizione alle frodi, arrivando ad un Fraud Rate del 0.17% contro l'1.86% di media delle aziende del settore.





Galileo.XAI

Insight graph data platform

Galileo.XAI

PER utenti business ed esperti di dominio

CHE NECESSITANO di ottimizzare i processi aziendali e di derivare nuova conoscenza e predizioni da i dati

Galileo.XAI è una Graph Data Platform velocemente integrabile

CHE rende automatiche, facili ed intuitive ricerche, analisi ed investigazioni complesse

AL CONTRARIO della gran parte delle Data Platform

IL PRODOTTO utilizza avanzati algoritmi di network science, XAI ed una versatile visualizzazione a grafo, il tutto presentato in modalità low-code/no-code facilmente fruibile e configurabile anche da figure non tecniche.



Product CORE

Galileo.XAI permette la comprensione di dati complessi, l'identificazione di relazioni nascoste e la presa di decisioni informate basate su previsioni trasparenti.



Graph Approach

Galileo.XAI si basa su una struttura dati a grafo

Business Friendly

Galileo.XAI offre una interfaccia user-friendly orientata ai profili business

Explainable AI

Galileo.XAI è progettato per fornire spiegazioni sulle previsioni dell'IA

Vertical Solution

Galileo.XAI dispone di funzionalità specifiche per ogni caso d'uso

Dal dato alla conoscenza



IMPORTA

Definisci come modellare il tuo dominio. Importa facilmente i dati da una qualsiasi fonte. Rileva le modifiche e gli aggiornamenti esterni e importali in modo automatico

ANALIZZA

Analizza i tuoi dati, trova nuovi insights e modelli significativi e lascia che sia Galileo.XAI a notificarti quando saranno nuovamente rilevati

SCOPRI

Combina diversi algoritmi a grafo per creare il tuo mix perfetto. Identifica le strutture del grafo e le dinamiche della rete per individuare nodi o cluster

PREVEDI

Utilizza algoritmi di Graph-AI per apprendere direttamente dalle strutture a grafo come scoprire nuovi insights e risolvere i tuoi problemi di business più urgenti

COMPRENDI

Dai un contesto ai tuoi risultati ed evita l'effetto black box dell'AI. Incrementa la fiducia verso gli algoritmi grazie alla spiegabilità del processo decisionale

Importa

The screenshot displays a database import wizard interface. At the top, a progress bar shows four steps: 1. Connessione JDBC (checked), 2. Visualizza lo schema (active), 3. Conferma, and 4. Stato avanzamento import. The main area features a schema diagram with nodes for Rooms, Sections, Schedules, Teachers, Subjects, Courses, and Grades. Relationships are shown with arrows and labels like 'SECTIONS', 'INCLUDES', 'TEACHES', 'SUBJECT MANAGER', 'GRACE_DISTRIBUZIONE', and 'GRADUATION'. A sidebar on the left shows a list of properties for 'Sections' and 'uuid', including their types (VARCHAR, STRING) and options to import or set as a key. A 'CONFERMA' button is located at the bottom right of the main area.

Analizza

No-Code

Nodo	Proprietà	Filtro		
Actor	born	Minore di	02/10/1968	+
Seleziona il Nodo	Seleziona la proprietà	Seleziona il filtro	Inserisci una data nel formato gg/mm/aaaa	
Relazione				
ACTED_IN->	Proprietà	Filtro	+	
Seleziona la Relazione	Seleziona la proprietà	Seleziona il filtro	Valore	
Nodo	Proprietà	Filtro		
Movie	title	Uguale a	Matrix, The	+
Seleziona il Nodo	Seleziona la proprietà	Seleziona il filtro	Seleziona il valore	
Relazione	Proprietà	Filtro	+	
Seleziona la Relazione	Seleziona la proprietà	Seleziona il filtro		

Ricerca valore
mel gibson

Proprietà trovate:
 name - Mel Gibson
 bio - An actor, film director, producer and screenwriter. Born in Peekskill, New York, Gibson moved with his parents to Sydney, Australia when he was 12 years old and later studied acting at the Australian National Institute of Dramatic Art. After appearing in the Mad Max and Lethal Weapon series, Gibson went on to direct and star in the Academy Award-winning Braveheart...

Visualizza tutte le proprietà

Actor:Director:Person

Full-Text

Scopri

Aggiungi

Nome Proiezione *
Nuova Proiezione

Sovrascrivi proiezione con lo stesso nome

Nodi
Actor X Movie X

Relazioni
ACTED_IN X

Proietta Relazioni come non orientate

Configurazione

Proprietà Nodi
imdbRating X

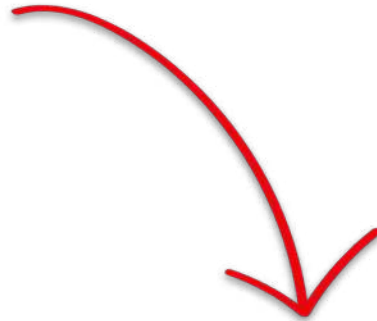
Si possono selezionare solo proprietà numeriche

Proprietà Relazioni
impressiveness X

Si possono selezionare solo proprietà numeriche

readConcurrency 4 Valida Relazioni

ANNULLA SALVA



ESECUZIONE ALGORITMI

Proiezione Grafo
Nuova Proiezione

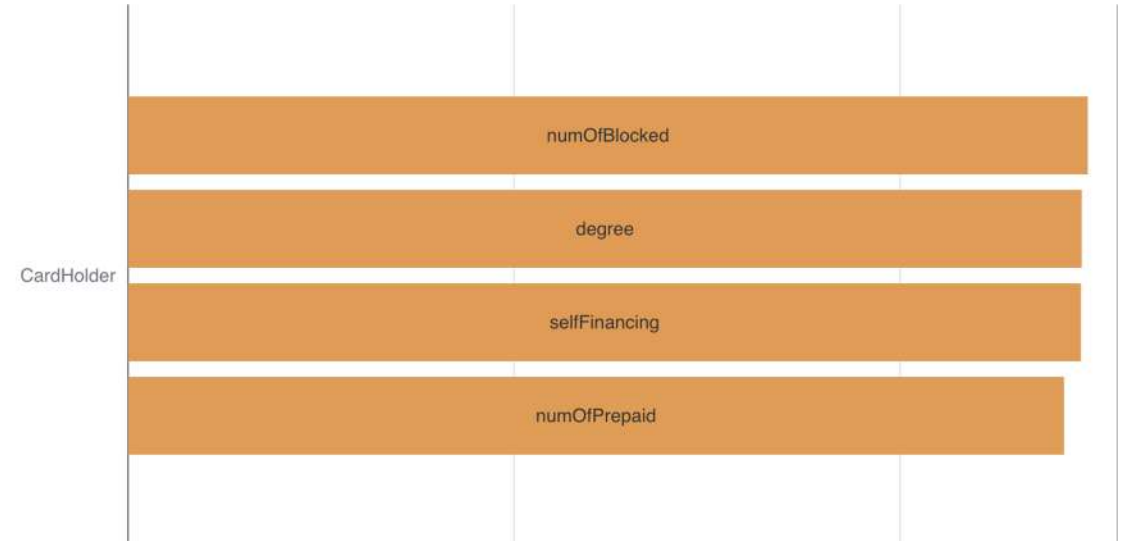
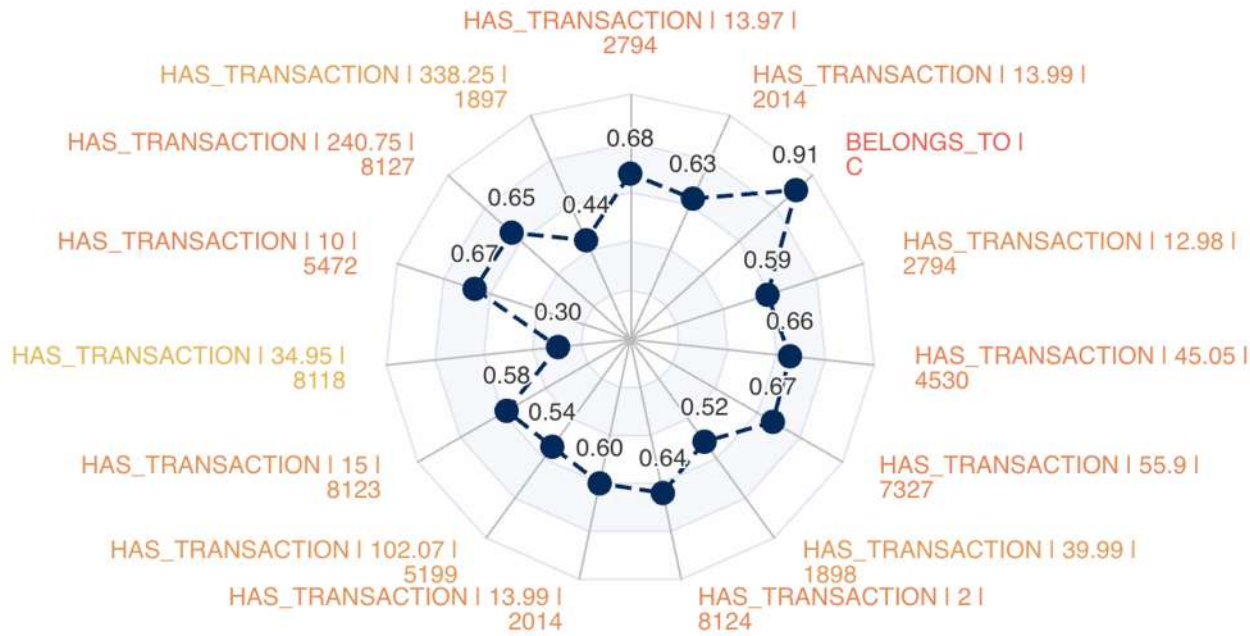
Selezionare la proiezione che si intende usare

Categoria Algoritmo Selezione Algoritmo Tipologia Esecuzione

CENTRALITY articleRank STREAM

Descrizione: Article Rank is a variant of the Page Rank algorithm, which measures the transitive influence or connectivity of nodes.

Prevedi e Comprendi





Galileo.XAI

Caso d'uso di successo per la Fraud Detection

Contesto e Soluzione

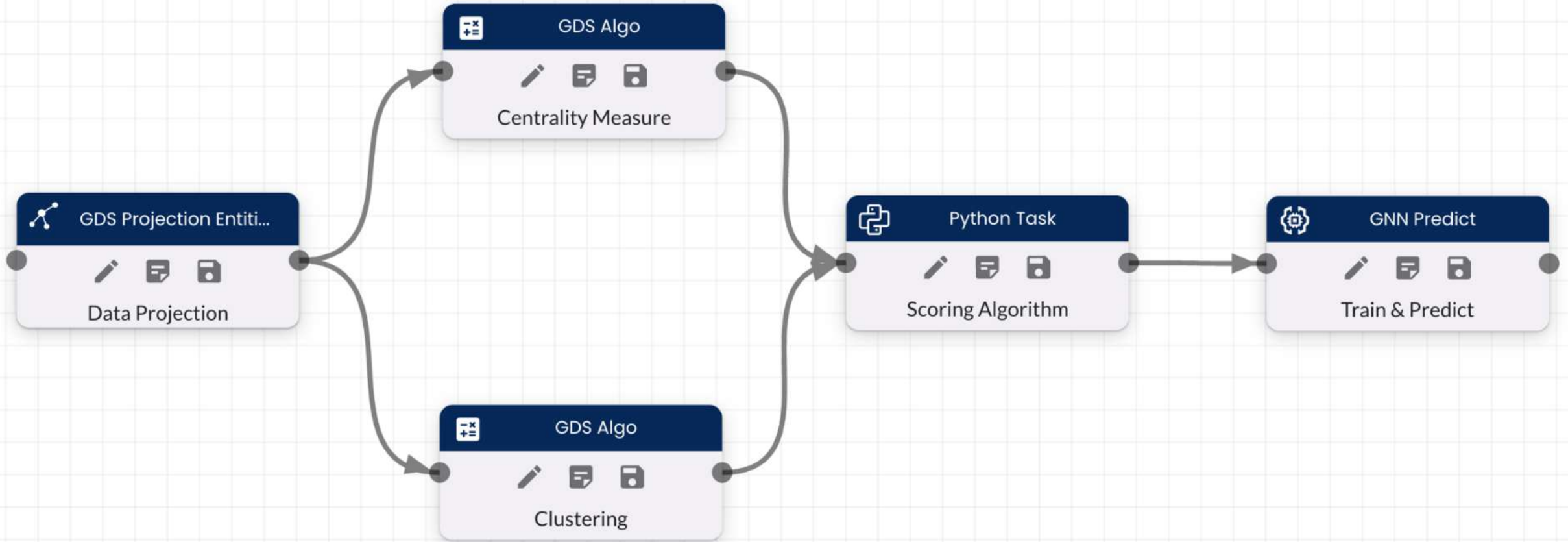
Contesto Cliente

- Importante azienda Assicurativa Italiana
- Processo di identificazione dei casi sospetti **manuale**
 - Un'unità apposita individua i casi sospetti e li assegna agli ispettorati anti-frode di competenza.
- Strumenti di analisi dei casi sospetti
 - Gli operatori tramite strumenti standard, come **Excel e Gestionali** controllano i casi uno ad uno
 - Investigazioni scambiate tramite fogli Excel e **fotocopie** del gestionale

... con Galileo.XAI

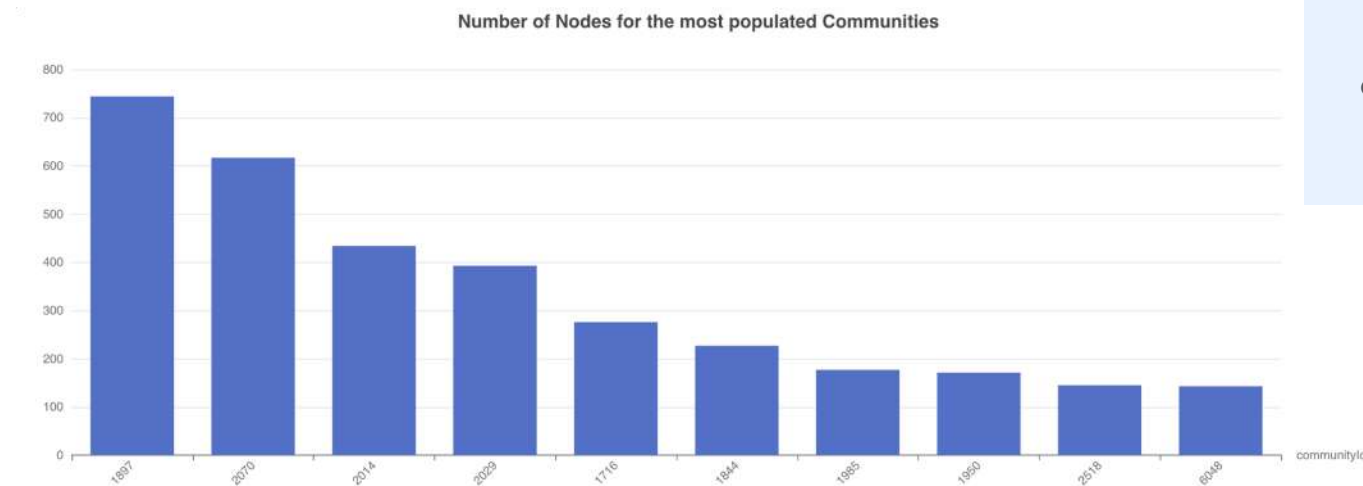
- Processo di identificazione dei casi sospetti **automatizzato**
 - Esiti che rispecchiano le regole di business
 - Esiti derivanti da Network Analysis per scovare i **Fraud Rings**
 - Esiti di **ML classification** ordinati per confidence
- Strumenti di analisi dei casi sospetti
 - Visualizzazione a Grafo **interattiva**
 - Ricerche complesse con un click
 - Ricerca **collegamenti** Soggetti con un click
 - Area dedicata per **condivisione** investigazione e sospetti, con report

Data Pipelines



Data Science

- Interfaccia **no-code** per creare Proiezioni sui dati e lanciare Algoritmi sia in lettura che in scrittura
- Vari algoritmi per Centralità, Community Detection, Similarity, Embeddings e Path Finding
- Risultati tabellari e **Grafici** di distribuzione



- **Calcolo di score** derivanti da Network Analysis essenziali per il calcolo di Score combinati e come **feature** per il processo di **ML**.
- Individuazione Semplice di **Fraud Rings** tramite Community Detection, come ad esempio SCC o WCC
- Primo step per l'Individuazione di nuovi pattern in modalità unsupervised

Graph Neural Network - Classification

- Interfaccia **no-code** per creare modelli sulla base del dataset e delle relazioni, scegliendo le feature più adatte
- Integrato con sistema di **MLOps** per riproducibilità e storia di modelli ed esperimenti
- Capacità di **Explainable AI**, per comprendere il ragionamento del modello e scoprire quali feature e relazioni condizionano maggiormente il risultato.
- **Prediction Confidence**, per prioritizzare le investigazioni

- Individuazione automatica di **nuovi pattern** fraudolenti sulla base della conoscenza pregressa.
- Esiti in near-real-time ordinati per confidence
- **Migliora nel tempo** grazie al normale lavoro di Investigazione degli operatori (Human in the middle)



Individuazione Casi Sospetti - Esiti


- Traduzione di una **regola di business** in una ricerca schedulata di pattern nei dati, esprimibile in modalità **no-code**
- Esiti raggruppati per regola e giorno, **filtrabili** sulla base di proprietà nei dati, **assegnabili** e **risolvibili**
- Esplorazione degli Esiti in cui compare un entità
- Esplorazione delle Investigazioni collegate ad un Esito

Automazione dell'individuazione di casi sospetti:

- **Risparmio del 90%** sui costi del personale
- **Individuazione near-real-time** dei nuovi casi sospetti
- Censimento delle regole di business, **democratizzazione delle procedure**
- Ottimizzazione del processo di Indagine

Investigazione - Ricerche rapide

13667569



Quick Action [X]

Seleziona Quick Action

Co-Participants in Accidents [🔍]

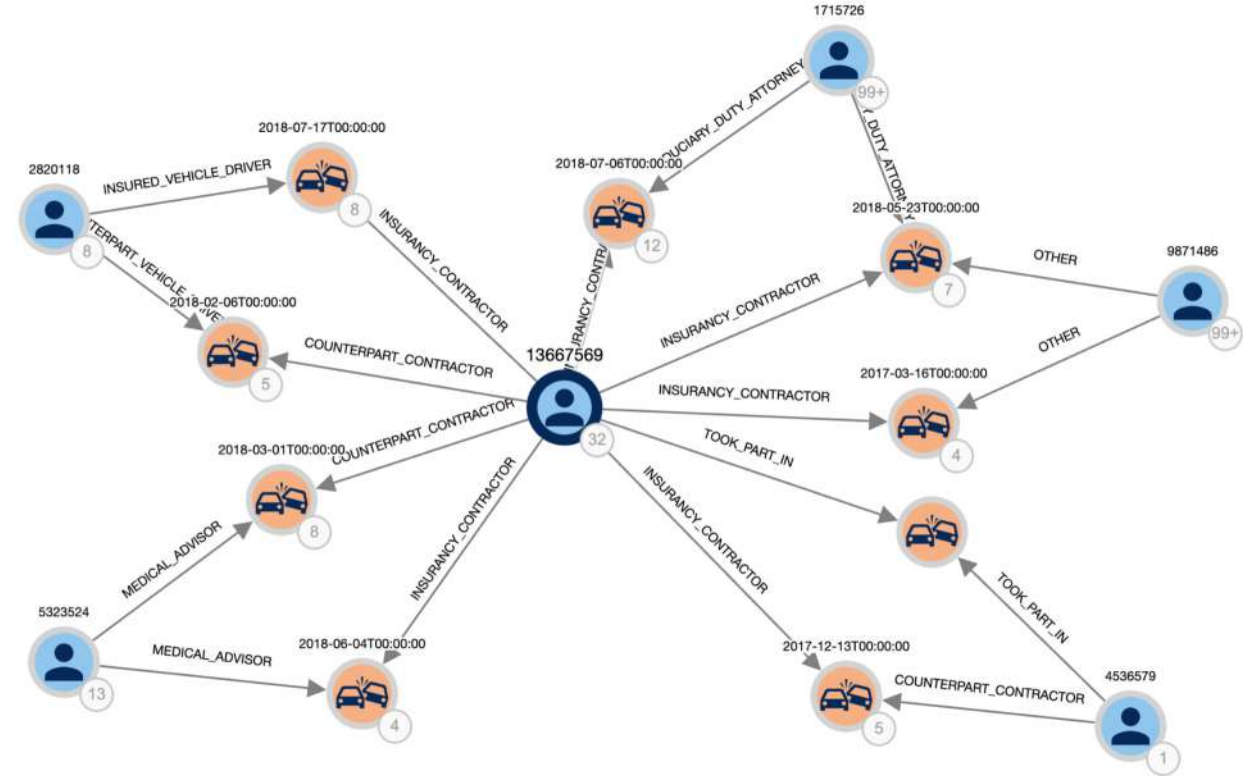
Given a Subject, find in its accidents for any other subject that occurs more than once

Associazione Nodi

Nodo selezionato

subje... node-13667569

ESEGUI

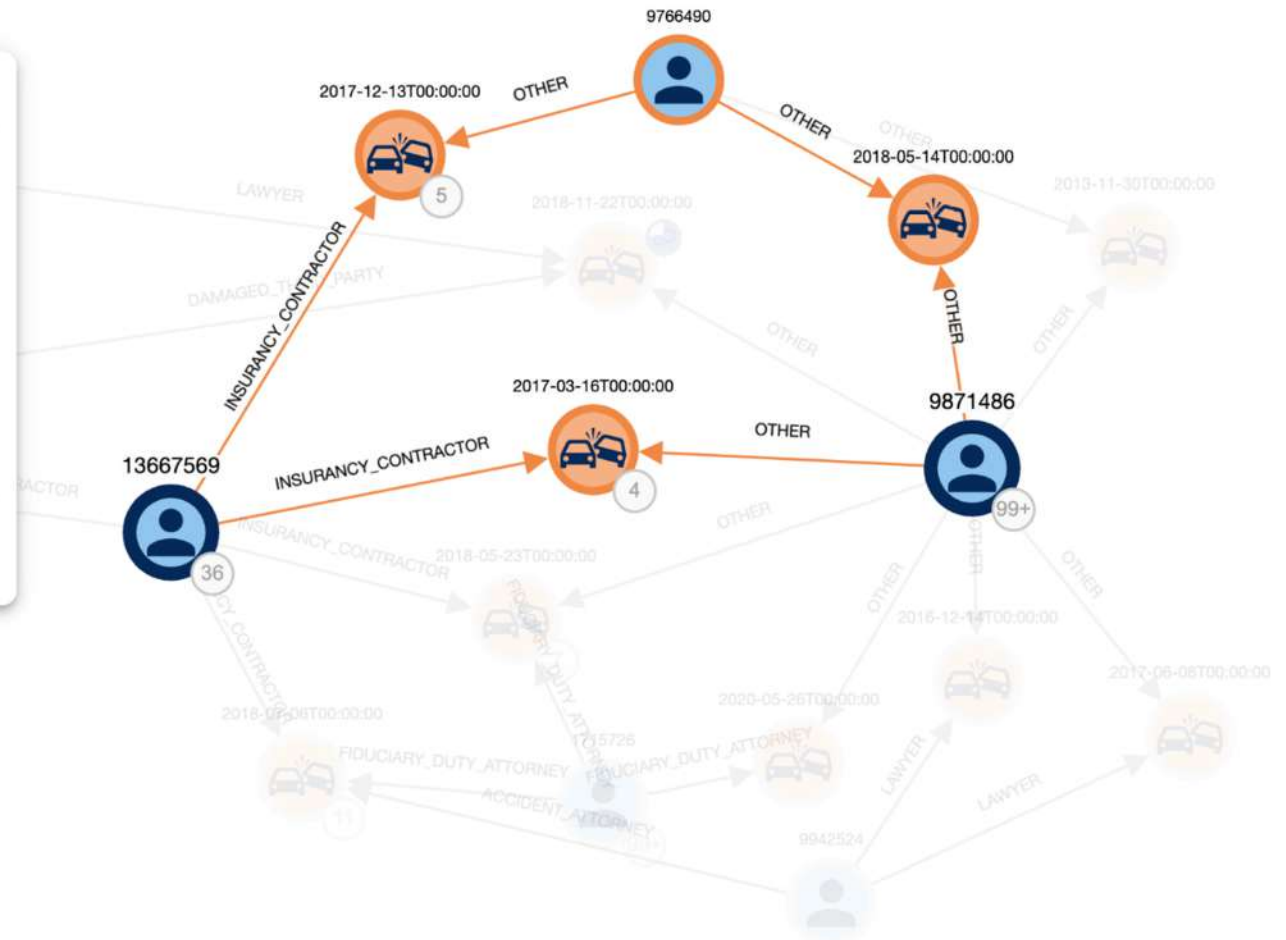


Investigazione - Ricerca Collegamenti

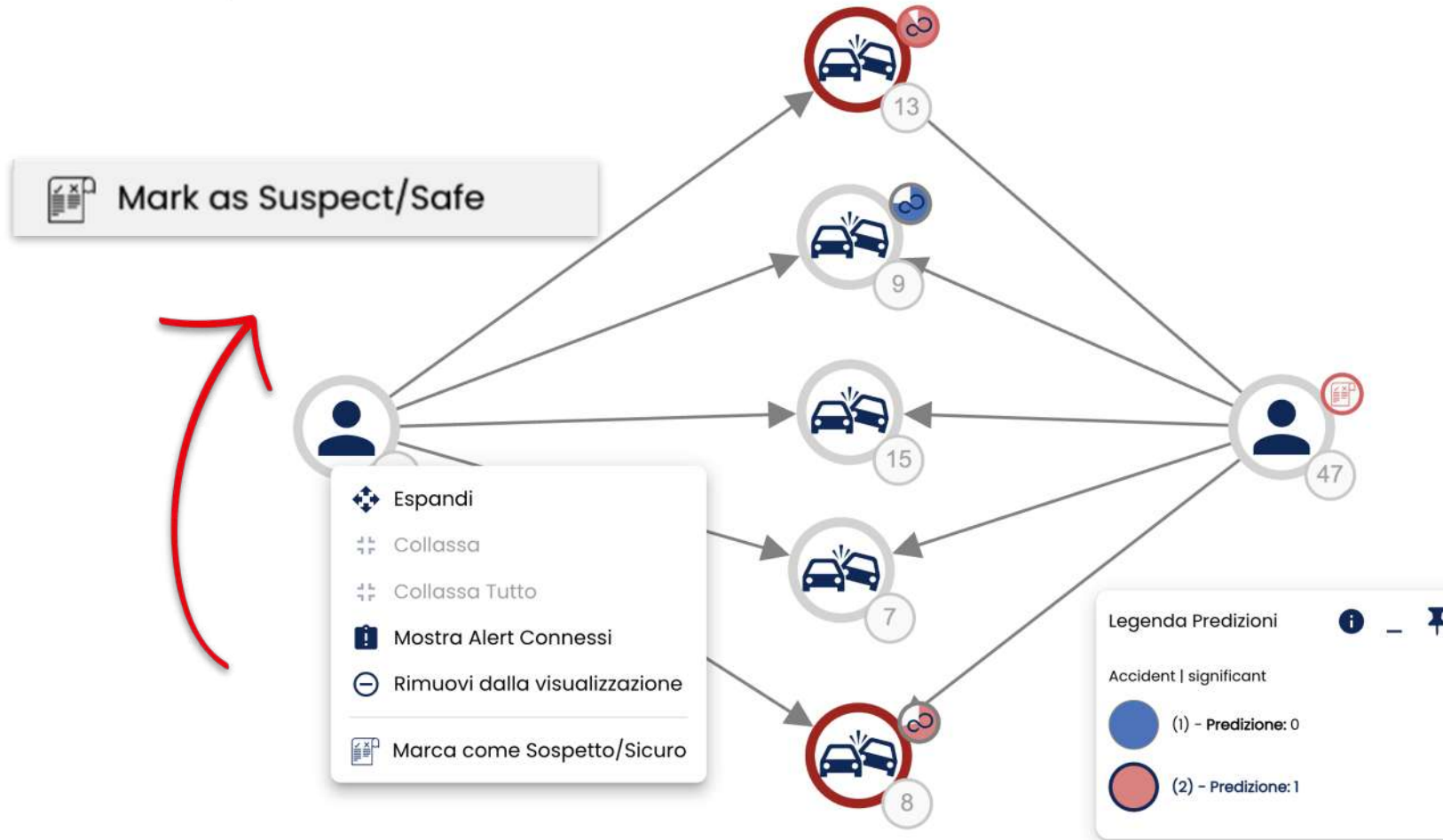
Percorsi trovati:

Selezionati	Collegamenti:
<input checked="" type="checkbox"/>	13667569 → INSURANC... → 2017-03-1... ← OTHER
<input type="checkbox"/>	13667569 → INSURANC... → 2018-05-2... ← OTHER
<input type="checkbox"/>	13667569 → INSURANC... → 2018-02-2... ← DAMAGED...
<input type="checkbox"/>	13667569 → INSURANC... → 2018-02-2... ← LAWYER
<input checked="" type="checkbox"/>	13667569 → INSURANC... → 2017-12-13... ← OTHER

PERSISTI < 1 2 3 4 5 ... 9 >



Investigazione - Sospetti ed AI



Investigazione - Report

The screenshot displays a web application interface for investigation reports. At the top, there are two tabs: "SOSPETTI" (selected) and "SICURI". Below the tabs is a table with columns "Label", "Descrizione", and "Utente". The table contains two rows:

Label	Descrizione	Utente
<input type="checkbox"/> Accident	da verificare	admin
<input type="checkbox"/> Person	sospetto	admin

On the left side, there is a "Ricerca Avanzata" (Advanced Search) panel with sections for "Descrizione" (Description) and "Ricerca Libera" (Free Search). Below the search panels are buttons for "NODI", "RESETTA", and "CERCA".

A modal window is open, showing a conversation between "admin" and "user". The "admin" message, dated 29/09/2023 15:19, says "incidenti in comune con I2784379". The "user" message, dated 29/09/2023 15:20, says "proseguo ulteriori accertamenti". Below the messages is a "Note" field with a text input and a right arrow button.

A red arrow points from a "Revisioni" (Revisions) button to the modal window. Another red arrow points from the "Revisioni" button to the table's action icons for the "Person" row.

Conclusioni e Sviluppi Futuri

Grazie all'integrazione di Galileo.XAI, l'azienda assicurativa in questione:

- Ha **ridotto i costi** relativi alle procedure di Identificazione dei casi sospetti **del 90%**, potendo così destinare più risorse alla fase di Investigazione
- Ha **velocizzato** in modo importante **il processo di Investigazione** grazie alla rapidità ed espressività delle funzionalità messe a disposizione, come la ricerca dei Soggetti correlati e di relazioni tra Soggetti
- Ha **rinforzato** e velocizzato **la collaborazione** tra colleghi grazie allo store condiviso per investigazioni e casi sospetti, tenendo anche traccia del pregresso.
- E' in grado di poter analizzare **nuovi pattern** provenienti dall'**AI** e dagli algoritmi di Network Analysis

La **roadmap** per il futuro immediato, prevederà la messa a disposizione di Pipeline per l'individuazione di certificati, pratiche e testimonianze generate da GenAI.

Non solo Fraud Detection. Tu crei la sfida, Galileo.XAI ti da la soluzione



**Rilevamento
di frodi**



**Governance
dei dati**



**Ottimizzazione
reti di fornitura**



Antiriciclaggio



**Analisi delle
reti sociali**



**Sistemi di
raccomandazione
next-gen**



**Gestione degli
asset aziendali**



**Gestione degli
approvvigiona
menti**



**Gestione dati
di produzione**



**Analisi
Bioinformatica**

Red Hat
Summit

Connect

Q&A



Red Hat
Summit

Connect

Thank you for your curiosity

Roberto Sannino

Product Owner @LARUS

roberto.sannino@larus-ba.it

