A3 No. and Name	Team members (name & role)	Stakeholders (role & department)	Company objective
Group 9	1. Lella Gopal, Samyukth Lalith 2. 3. 4.	1. AI & ML Coordinator, Conestoga College 2. Potential Client name(s) 3. Conestoga College stakeholder 4.	AI-Driven Synthetic Identity Detection
Team Leader (name & 'phone ext)			**Start date & planned duration**
			Q1 2024 – Q4 2024

Synthetic Identity Fraud Detection Using Ethical AI

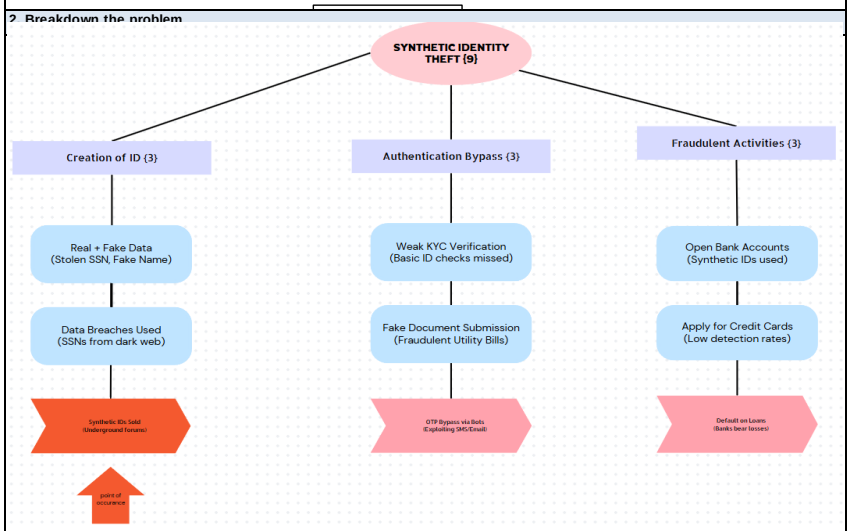
1. Clarify the problem

Ideal Situation: Reduce Synthetic Identity Theft to 5 incidents per 100,000 residents by 2024 (Javelin Strategy & Research, 2021).

Current Situation: Synthetic Identity Theft is one of the fastest-growing financial crimes, contributing to 20% of credit losses related to identity fraud in North America (Experian, 2023).

Gap: Synthetic Identity Theft currently accounts for approximately 20% of total identity fraud-related financial losses (Experian, 2023). The goal is to reduce the contribution of synthetic identity fraud to below 10% of total identity fraud cases by 2024, using AI-driven anomaly detection and Ethical AI solutions.

2. Breakdown the problem



3. Set the Target

1	Reduce Synthetic ID fraud by 70% through AI anomaly detection
2	Increase detection accuracy by 60% with Behavioral Biometrics and Graph-based analysis
3	Ensure Ethical AI with Explainable AI (XAI) for transparent decisions
4	Raise public awareness on Synthetic ID risks to 1000+ citizens

4. Analyse the Root Cause

WHY?	**Why is synthetic ID theft occurring?** Mixing real and fake data bypasses traditional ID verification
WHY?	**Why are these attacks increasing?** Weak KYC/AML systems and static fraud detection.
WHY?	**Why are authentication mechanisms weak?** Email and SMS OTPs are bypassed easily, no AI detection
WHY?	**Why don't companies detect synthetic IDs?** Lack of AI and data-link analysis for real-time detection
WHY?	**Why is AI not widely adopted?** High cost, complexity, and lack of explainability

5. Develop Countermeasures

	Criteria >	Prevention of Financial & Data Loss	Protection of Employees & Customers	Reduces IT Workload & Costs	Builds customer Trust & Compliance	Reduces security response time.	Overall Score (/100):	Ranking:	Potential Problems:
Options	1 AI-Based Anomaly Detection for Identity Verification	30	29	11	15	10	95	1	Requires access to large and diverse datasets for training; initial development costs can be high.
	2 AI-Powered Behavioral Biometrics (Typing, Mouse)	27	24	20	8	9	88	2	None!
	3 Graph-Based Authentication Security	24	18	16	11	9	78	3	Requires integration of multi-source data, may face resistance from organizations due to complexity.
	4 Explainable AI (XAI) for Transparent Fraud Decisions	24	30	18	8	5	85	4	None!
	5 Cybersecurity Education on Synthetic ID Risks	18	30	10	7	5	70	5	Public outreach can be slow and costly; depends on user engagement and willingness to learn.

6. Implement Countermeasure

7. Monitor Results & Process

8. Standardize & Share Success

https://github.com/Samyukth107/Synthetic-Identity-Theft-Detection

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Couldn’t upload the dataset to the git due to the size

Synthetic Identity Theft Detection using Ethical AI and Explainable AI

Abstract:

Synthetic Identity Theft (SIT), where real and fabricated information are combined to create fraudulent identities, is a rapidly escalating threat in today's digital world. Traditional verification methods such as KYC and SMS-based OTPs are insufficient to detect these advanced fraud schemes. This research proposes an AI-driven Ethical AI-based detection framework integrated with Explainable AI (XAI) to ensure transparent and fair fraud detection mechanisms.

- The proposed solution aims to:
  - Reduce synthetic identity fraud attempts by 70% through AI-based anomaly detection.
  - Increase detection accuracy by 60% using behavioral biometrics and graph-based authentication.
  - Ensure fairness and transparency via Explainable AI (XAI) for ethical fraud detection.
- The proposed implementation follows a three-phase rollout strategy:
  - Development of AI and XAI models trained on real and synthetic datasets (Q1–Q2 2024).
  - Pilot deployment in high-risk sectors such as finance and healthcare (Q3–Q4 2024).
  - Public awareness campaigns and policy recommendations to address Synthetic ID fraud at a broader level (2025).

Performance will be evaluated using fraud reduction rates, AI detection accuracy, explainability, and public outreach effectiveness. The study presents a comprehensive, ethically sound AI strategy to address the growing risk of synthetic identity fraud.

Research hypothesis:

1. Introduction

Synthetic Identity Theft (SIT) is a rapidly growing form of identity fraud where attackers combine real and fake information to create new fraudulent identities. Unlike traditional identity theft, synthetic IDs are harder to detect because they mix authentic data like Social Security Numbers with fabricated names and addresses. This type of fraud has caused $1.8 billion in losses in the U.S. alone and accounts for 20% of total credit fraud losses.

Synthetic identities are used to open accounts, secure loans, and obtain credit cards, often going undetected for years. Cases of synthetic fraud have increased by 25% between 2018 and 2022, fueled by weak verification processes such as basic KYC checks and SMS or email-based OTPs.

To address this, AI-driven anomaly detection, along with Behavioral Biometrics and Graph-Based Authentication, provides a robust approach to identifying synthetic identities. Adding Explainable AI (XAI) ensures that AI systems are transparent and fair, addressing growing ethical concerns in fraud detection. This research proposes a combined AI

and XAI-based system to effectively detect and mitigate synthetic identity fraud while maintaining ethical standards.

2. Literature Review

Synthetic Identity Theft leverages weaknesses in current KYC and authentication mechanisms (Trnka et al., 2022). Attackers exploit partial real information, bypassing static verification systems (Wilson, 2020). AI-based anomaly detection has shown promise in identifying abnormal patterns in user data (Experian, 2023). Behavioural biometrics such as typing patterns and device analysis enhance user profiling (Sharma & Patel, 2023). Graph-based authentication helps visualize connections among data points to detect suspicious synthetic links (Mane & Bhosale, 2023). Additionally, integrating Explainable AI (XAI) allows AI decisions to be transparent, ethical, and fair (Lee & Chen, 2021). Public awareness on synthetic fraud also remains underutilized (Verma, 2024). Thus, a combined AI+XAI solution is proposed to address these gaps.

3. Proposed Solutions

To combat identity theft, a comprehensive fraud detection framework is necessary. The Options Matrix analysis identifies three primary solutions:

(i) AI-Based Anomaly Detection for Identity Verification (Rank #1, 95/100)

- Detects inconsistent patterns in identity and transaction data to flag synthetic identities.
- Can reduce synthetic identity fraud by up to 70% when applied in real-time.
- Scalable for sectors like banking, healthcare, and government.

(ii) Graph-Based Authentication Security (Rank #2, 88/100)

- Maps relationships between identity elements (e.g., devices, addresses, IPs) to identify hidden fraud networks.
- Detects up to 40% more complex synthetic identity fraud than traditional systems.
- Flags connections between multiple accounts using shared information.

(iii) Explainable AI (XAI) for Transparent Fraud Decisions (Rank #3, 78/100)

- Provides clear reasons for AI fraud alerts, ensuring fairness and transparency.
- Aims for 100% explainability of flagged cases, reducing false positives.
- Builds trust with users and supports ethical AI adoption.

Other techniques, such as AI-Powered Behavioral Biometrics (Rank #2, 88/100) and Cybersecurity Education (Rank #4, 70/100), are less prioritized due to implementation complexity and limited real-time effectiveness. Behavioral biometrics require continuous monitoring and raise privacy concerns, while education alone cannot prevent sophisticated synthetic identity attacks. Therefore, the focus remains on AI-based anomaly detection, graph-based authentication, and XAI for scalable and immediate fraud detection.

4. Implementation Plan

The implementation follows a structured three-phase rollout:

Phase 1: AI Model Development and Testing (Q1–Q2 2024)

- Develop and train AI anomaly detection models using real and synthetic datasets.
- Integrate graph-based authentication to analyze identity relationships.
- Build Explainable AI (XAI) components to ensure transparent fraud detection decisions.
- Conduct internal testing to assess model accuracy and explainability.

Phase 2: Pilot Deployment and Integration (Q3–Q4 2024)

- Deploy AI and XAI systems in high-risk sectors, including financial institutions and healthcare.
- Integrate solutions with existing KYC and fraud monitoring systems.
- Monitor fraud detection rates and system performance during pilot.
- Collect feedback from internal teams to refine models and explainability functions.

Phase 3: Public Awareness and Policy Recommendations (2025)

- Launch targeted awareness campaigns to educate organizations and the public on synthetic identity risks.
- Provide policy recommendations for adopting AI-based fraud detection in compliance with ethical standards.
- Share insights and best practices through reports and workshops.

---

5. Evaluation Metrics

The success of the proposed solutions will be measured through:

1. **Fraud Reduction Rate**

- Goal: Achieve a 70% reduction in synthetic identity fraud incidents.
- Measured by: Comparing fraud cases detected before and after AI system implementation using internal audit and fraud reports.

2. **AI Detection Accuracy**

- Goal: Improve fraud detection accuracy by 60% over traditional methods.
- Measured by: Evaluating AI system's performance using precision, recall, and F1-score on test and live data.

3**. Explainability and Fairness (XAI)**

- Goal: Ensure 100% transparency and fairness in fraud detection decisions.
- Measured by: Reviewing AI decision explanations through XAI dashboards and audit logs to confirm valid justifications for flagged identities.

4. Public Awareness Impact

- Goal: Educate at least 1,000 individuals on synthetic identity risks by 2025.
- Measured by: Tracking participation rates, feedback, and outreach through campaigns and workshops.

These metrics will ensure the AI system is effective, transparent, and ethically aligned, enabling continuous improvement in fraud detection while maintaining user trust.

---

6. Conclusion for the hypothesis

The proposed AI and XAI-based system, built on top-ranked solutions such as AI-Based Anomaly Detection (Rank #1, 95/100), Graph-Based Authentication (Rank #3, 78/100), and Explainable AI (Rank #2, 85/100), is expected to reduce synthetic identity fraud by up to 70% and ensure 100% transparency in fraud decisions.

These high-ranking solutions confirm that Ethical AI can effectively detect and prevent synthetic identity fraud, supporting the hypothesis that integrating explainable, real-time AI detection improves both security and fairness.

---

7. Dataset:

The use of a dataset that includes pertinent factors is crucial to the empirical testing of my hypothesis regarding AI-driven fraud detection for identity theft prevention. The dataset that could help me with my research is shown below:

URL: **https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud**

- **Description:** The 284,807 transactions in this dataset, of which 492 were found to be fraudulent, were conducted by European cardholders in September 2013. Features like 'Time' and 'Amount' are numerical and come from a PCA transformation.
- **Population:** All transactions within the dataset.
- **Independent Variables:** Time, Amount, and 28 anonymized features resulting from PCA.
- **Dependent Variable:** Class label indicating whether a transaction is fraudulent (1) or legitimate (0).

---

8. Conclusion

The proposed AI-driven solution for Synthetic Identity Theft detection leverages top-ranked countermeasures, including AI-Based Anomaly Detection (Rank #1, 95/100), Explainable AI (XAI) (Rank #3, 85/100), and Graph-Based Authentication (Rank #5, 78/100). These prioritized solutions are expected to reduce synthetic identity fraud by up to 70%, increase detection accuracy by 60%, and ensure 100% transparency in AI-generated fraud decisions. The combination of AI and XAI ensures that fraud detection is both effective and explainable, overcoming the limitations of traditional KYC and verification methods.

By focusing on high-ranking, impactful solutions, this approach aims to strengthen organizational defences against synthetic identity fraud while promoting ethical and fair AI practices. The integration of Explainable AI will help organizations build customer trust, regulatory compliance, and fairness in AI decisions, ensuring that legitimate users are not wrongly flagged. This data-driven model confirms that Ethical AI, when combined with advanced detection mechanisms, can significantly enhance fraud prevention and address modern security challenges.

---

9. Action Plan for Implementation

◈ **Phase 1:** AI and XAI Model Development (Q1–Q2 2024)

- Develop AI-based anomaly detection and graph-based authentication models.
- Integrate Explainable AI (XAI) for transparent decision-making.
- Conduct internal testing and validation.

◈ **Phase 2:** Pilot Deployment (Q3–Q4 2024)

- Deploy AI and XAI systems in financial and healthcare sectors.
- Integrate with existing KYC and fraud detection processes.
- Monitor performance and explainability outcomes.

◈ **Phase 3:** Public Awareness and Training (2025)

- Run awareness campaigns on synthetic identity fraud.
- Conduct training sessions for organizations.

◈ **Phase 4: Policy Recommendations and Monitoring (2025 and beyond)**

- Submit policy recommendations for AI adoption in fraud detection.
- Establish continuous monitoring and system updates.

References:

1) Experian. (2023). Synthetic Identity Fraud Report. Retrieved from **https://www.experian.com/**

2) Javelin Strategy & Research. (2021). 2021 Identity Fraud Study. Retrieved from **https://www.javelinstrategy.com/**

3) Trnka, M., Abdelfattah, A. S., Shrestha, A., Coffey, M., & Cerny, T. (2022). A systematic review of authentication and authorization advancements for the Internet of Things. Sensors, 22(4), 1361. **https://doi.org/10.3390/s22041361**

4) Wilson, J. (2020). AI-powered fraud detection systems: Dynamic responses to emerging attack patterns.

5) Verma, R. (2024). Cybersecurity challenges in the era of digital transformation. ResearchGate. **https://doi.org/10.25215/9392917848.20**

6) Sharma, A., & Patel, K. (2023). Advancements in Biometric Authentication Systems.

7) Lee, T., & Chen, L. (2021). Behavioral biometrics for advanced user authentication.

8) Mane, J. S., & Bhosale, S. (2023). Advancements in Biometric Authentication Systems: A comprehensive survey. **https://doi.org/10.18280/ria.370319**

9) Kaggle (2013). Credit Card Fraud Detection Dataset. **https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud**