Communication Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 277–285
DOI: 10.15439/2022F35
ISSN 2300-5963 ACSIS, Vol. 32

Elaboration of Financial Fraud Ontology

1stAdamu Hussaini

Université de Reims Champagne Ardenne, France.
(CReSTIC EA 3804, 51097 Reims, France.)
adamu.hussaini@etudiant.univ-reims.fr

2ndZahia Guessoum

Université de Reims Champagne Ardenne, France
(CReSTIC EA 3804, 51097 Reims, France.)
zahia.guessoum@univ-reims.fr
0000-0003-2303-7263

3rd Eunika Mercier-Laurent

Université de Reims Champagne Ardenne, France
(CReSTIC EA 3804, 51097 Reims, France.)
eunika.mercier-laurent@univ-reims.fr
0000-0003-2303-7263

***Abstract*—Financial Frauds have dynamically changed, the fraudsters are becoming more sophisticated. There has been an estimated global loss of 5.127 trillion dollars each year due to various forms of financial frauds. Industries like banking, insurance, e-commerce and telecommunication are the main targets for financial frauds. Several techniques have been proposed and applied to understand and detect financial frauds. In this paper we propose an ontology to describe financial frauds and related knowledge. The aim of this ontology is to provide a semantic framework for the detection of financial frauds. Theoretical ontology has been elaborated exploring various sources of information.**

After describing the research objectives, related works and research methodology, this paper presents details of theoretical ontology. It is followed by its validation using real datasets. Discussion of the obtained results gives some perspectives for the future work.

***Index Terms*—Fraud, Detection, Ontology, Concepts, Class, Entities, Validation, Dataset.**

I. Introduction

FINANCIAL frauds exist in different aspects of our life. Detection and prevention of those frauds represent an important question relevant to many academic disciplines. Fraud is an economically significant issue, that leads to depressions, bankruptcies and unfortunately suicide.

Ontology is used to study concepts, it has been applied to solve various problems in several domains such as medicine, computing, and economics. The objective of this research work is to understand the existing frauds and to propose a system able to detect them in order to prevent damages. This challenge is ambitious because the addressed problem is complex and dynamic - the fraudsters learn from failures and invent new ways of fraud. We can find in the literature various approaches to face this threat.

The paper is organized as follows. Section one is the introduction of financial frauds and the use of ontologies to detect financial frauds. In Section 2, we discuss the research challenge. Research methodology is described in Section 3, and related works in Section 4. While elaboration of the ontology, which includes theoretical development, validation, and discussion, is presented in Section 5. Finally, conclusion and some future works are proposed in Section 6.

II. Research Challenge

Financial institutions are facing several challenges in preventing and detection of frauds. Many firms proposed to use artificial intelligence and other advanced detection approaches.

In this paper the main challenges are to develop and validate a theoretical ontology for financial frauds that will understand and identify patterns for fraud detection and prevention. While undergoing the research, we must identify and provide a comprehensive solution to the problem of domain modeling, knowledge extraction, as well as creation of concepts.

III. Methodology

The research methodology focuses on heavyweight research methods that provide comprehensive detail of the research work process. Several development stages proposed in this research work are heavily influenced by one another. These stages suggest a routinely certain activities, which include:

- The understanding of the various types of frauds and related fraudsters' methods: The research work relies on publicly available information, consulting experts with the aim of understanding different types of financial frauds, while at the same time learning and deducing various fraud patterns used by the fraudsters. This methodology facilitates the organization of the frauds into different classes.
- Hypothesis, the research pointed out a possible expectation or prediction expected from the ontology, which include understanding of financial frauds' behaviors. The hypothesis of the research noted a possible change of behavior from the fraudsters. Hence, modification and change in the ontology will be needed from time to time.
- Study the state of the art: The section studies the domain of financial frauds, as well as other related entities of financial frauds that have a direct impact with the aims and objectives of the research. Understanding the existing

works makes it possible to achieve the goal of this research.

- Search for existing ontologies for possible reusing: Existing ontologies help to reposition the research as well as getting the right document we need for the creation of concepts, entities and other semantic content of our ontology. Some elements of the knowledge extracted are to be used as concepts that can be mapped to an ontology, some of the extracted knowledge are more complicated to be transformed into a concept.
- Adaptation and elaboration of theoretical ontology considering the types of frauds: This stage provides us with a theoretical implementation of the concepts and sub concepts of the ontology, the section iterates the overview of the ontology. Adaptation of this approach is the key point to the successful development of the ontology.
- Validation using datasets: The validation is proposed to help users to develop a confidence in the semantic content of the ontology. We use the dataset-based validation process due to the available resources and quality of the validation technique. We use various kinds of datasets for different types of frauds to validate the content of this ontology.
- Discussion of results and improvement: This stage of the research work will consider the aim and objective of the research work, compared with the validation result. The discussion of the result highlights if the goal of the research has been achieved. This section will give us a prediction of possible future improvement and modification of the ontology.

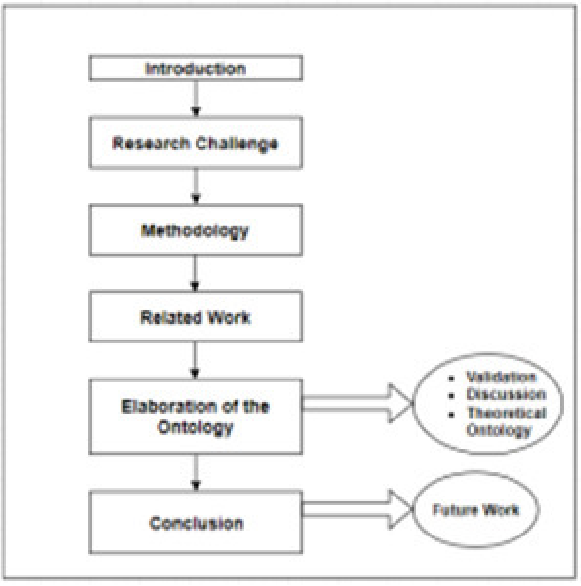


Fig. 1. The structure of the paper.

IV. Related Work

There are a considerable number of ontologies for finance, and fraud detection and prevention that have been proposed and developed.

Alexopoulos et al. [1] this propose provide an ontological basis which serves as a base on which the classification of ontology layers can be created. Generic, domain and case specific layers proposed. These three independent but interconnected layers, each one defining its own set of ontologies. We use the approach proposed in this ontology to implement to design our ontology. For example Generic layer of financial fraud, domain layer of banking fraud and the case specific layer of email fraud.

Tang et al. [2] propose an ontology to detect e-mail frauds. The ontology aims to tackle financial frauds, using formal ontological repositories as well as multilingual terminological resources. This ontology makes it possible to detect a fraud indicator within email content no matter how the fraudster tries to manipulate the content. The knowledge extraction process used to extract fraud indicators within emails is updated regularly. Hence, automatic rules are inserted in the knowledge extraction process. The approach provides a methodology of knowledge extraction, the extracted knowledge is used to built lexicons for semantic annotations. While this is important, the words processing without going into conceptual modelling and systematic linguistic engineering and without going deeper into semantic models of frauds, can be problematic In the cause of ontology validation.

Ontology-Based Framework Applied to Money Laundering [3] By Carnaz et al. This work provides a unified approach to represent and reason with investigation into money laundering. The approach is based on the knowledge of three money laundering steps. Placing, circulation, and integration. The ontology supports knowledge representation related to money laundering composed by several components. However, the paper does not present any framework of using real-world scenarios.The framework composed other components, preprocessing, data handling, knowledge and visualization. It has not been implemented yet. Hence, there is not test result with real-world scenarios.

Ahmed et al. [**?**] developed an ontology for electronic payment fraud prevention. The aim is detecting and preventing suspicious transactions on various electronic payment systems. The approach allows users to share an adaptive preventive rules of fraud on an ontology based on multiple and various payment systems. The paper discusses problems, issues in fraud detection and proposed prevention of financial fraud using ontology technology, among the concepts represented are legislation, legal rules from multiple countries, support reasoning about compliance with these legal rules and it represents interpretation of legislation and legal rules useful to the user.

In their article Kingston et al.m [**?**] propose a legal financial fraud ontology-based approach. The first phase of the proposed approaches aims building financial fraud concepts and the second phase aims building an ontology for laws against frauds. The ontology models the relationships where a two-way exchange is promised but only one of the two sub-transactions takes place while the second is to identify the relations between fraud laws, facts and inferences about facts present, facts presumed or facts missing.

[?] introduced an ontology for detecting suspicious fraud activities by monitoring individual transactions. The authors combined expert system and ontology with three components: Ontology construction, ontology reasoning as well as query on inferred ontology. The presented approach uses anti-money laundering guidelines provided by the State Bank of Pakistan. The transactions are decomposed into groups of transactions, graphical representation of while few sample transactions updated after the application of the following steps. Rules were built on top of the OWL ontology to deduce new knowledge (suspicious transactions) from the existing knowledge. The approach used protégé built in reasoner to deduce new knowledge about the nature of a financial transaction. One of the important tasks missing in this approach is the system capability to suggest transactions that need to be further analysed. So that its strength to identify suspicious transactions can be validated.

The related works provide several unique ways of treating financial fraud problem with ontology. e.g., deducing knowledge from the user behavior for analysing words in email and basic structure of ontology. Some of the methods used in these approaches have been reused in our ontology. We reused the ontological basis approach proposed by alexopoulos et al. [1] to design the structure of our ontology. The domain of money laundering in the second layer of our ontology was reused from the framework of detecting money laundering [3] by Carnaz et al. this ontology give us an important background knowledge of money laundering life cycle. The class of email fraud in the third layer of our ontology, was reused from the approach proposed to detect e-mail frauds by Tang et al. [2]. While these ontologies have exhibited great performance of knowledge discovery in the domain of fraud detection.They are mostly single layer ontologies with focus on domain or case-specific frauds. The aim of our work is to converge these knowledge and approaches, to provide an ontology that covers the whole area of financial frauds, treating the problem of financial fraud in a wider and deep approach. The ontology will have multiple layers from generic to domain and case-specific. It will serve as a knowledge repository for the various kinds of financial frauds,

V. A NEW FINANCIAL FRAUD ONTOLOGY

This section discusses the elaboration and overview of our proposed ontology. In this section we elaborated the knowledge in-closed in the multiple layers of the ontology, such as banking frauds, insurance frauds, identity theft and so on. The collective concepts, properties and relations matched together to give a comprehensive theoretical ontology. The properties of the concepts are keys to understanding the fraud patterns exhibited in a particular fraud domain.

A. Banking Fraud

In conceptualized context, bank fraud is using deception to steal money or assets from a bank or financial institution including credit unions and other financial institutions that are legally insured. This includes Federal Reserve banks, Deposit Insurance Corporation, Mortgage lending agencies, and other institutions that accept deposits of money or other financial assets. While in legal context, bank fraud is the use of potentially illegal means to obtain money, assets, or other properties owned or held by a financial institution, or to obtain money from depositors by fraudulently posing as a bank. In all instances, bank fraud is a criminal offense.

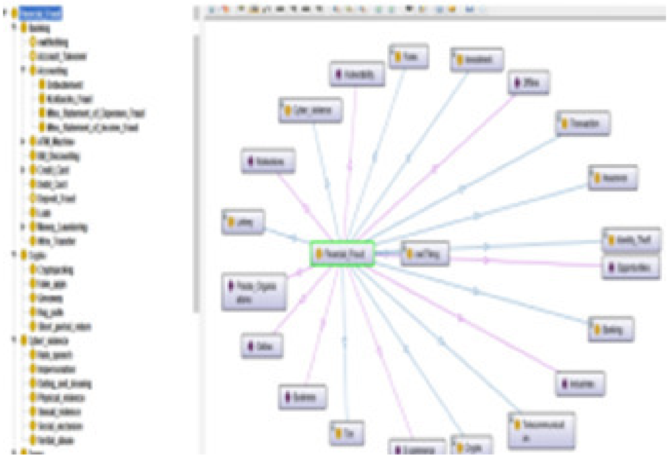


Fig. 2. The ontological structure of financial frauds.

In this research we created the concept of banking fraud with other several sub concepts together with their properties, including accounting fraud, bill discounting fraud, credit card fraud, debit card fraud, ATM machine fraud, money laundering transfer, loan fraud, and wire transfer fraud.

Fig. 3. The banking fraud concept.

1) *Accounting Fraud:* Accounting fraud is described as a deliberate manipulation and alteration of financial records or statements. [4] In this context, bank documents are being

manipulated, the account statement of a client is deviated to represent a false financial position of the client. This could be done to an individual or company account. Accounting fraud aims to convince shareholders and investors by displaying false financial statements of the company. Attempting unlawful tax evasions by misstating assets and liabilities and revenues and expenses.

In this research we classify accounting fraud into several classes, which include embezzlement, kickback, misstating income, expenses and so on. Accounting fraud affects various sections of financial fraud, including tax fraud, money laundering, and so on. Hence, we embedded properties of accounting fraud with relation annotation. The latter describes that an entity has another relation with another entity outside its root concept.

2) ATM Machine Fraud: The automatic teller machine fraud (ATM) is becoming a target for attacks globally because fraudsters and other criminals are trying to find the machine vulnerabilities. Several entities, tools, and dimensions used for ATM frauds. Some of these are, card jamming, shoulder surfing and stolen cards which constitutes 65.2 percent of ATM frauds in countries like Nigeria. [5] The ontology proposed a joint relation between ATM fraud and other forms of financial fraud. Including identity theft, wire transfer fraud, debit card fraud and ATM card fraud. In so many cases ATM fraud occurred due to the machine design structure. For example, the machine user interface serves as the communication medium between the user and the machine. It provides the resources for the user to input his card details, some of the ATM machines as we investigated, provide a window for shoulder surfing, and card jamming, and hidden cameras. We uncovered many reported cases of ATM fraud using some of the above-mentioned cases. Shoulder surfing fraud, refers to observing someone entering the PIN at an ATM machine, it is most effective in crowded places; fraudsters stand too close to the victims, so they can see card details like pin, card number, expiry date, and cvv number. Another case of ATM fraud is using hidden cameras to capture users' pin codes. In this case the fraudsters put a concealed camera to capture the details of the user while using the ATM machine.

3) Credit Card Fraud: Credit cards are one of the most famous targets of frauds but not the only one [6]. Credit card fraud is a healthy and growing means of stealing billions of dollars from credit card companies, merchants, and consumers. Due to the rise and rapid growth of e-commerce, use of credit cards for online purchases have dramatically increased and it caused an explosion in credit card fraud. Credit cards become the most popular mode of payment for both online as well as regular purchase [7].

In this research we investigate card fraud and the patterns used by the criminals to defraud the victims. Credit card fraud provides detailed ontological understanding of credit card fraud. From entities, terminologies, relations, instances as well as data types and classes, including card-not-present fraud, counterfeit, and skimming fraud, lost or stolen card fraud, and card-never arrived-fraud. The above-mentioned cases are the most occurring credit card fraud. For example, card-not-present fraud contributes to overall credit card fraud, accounting for 75 percent. Hence, we provide several properties and relations between the card not present entity with other credit card and non-credit card fraud. The relation we proposed to make it possible to generate reasoning for other credit card fraud detection.

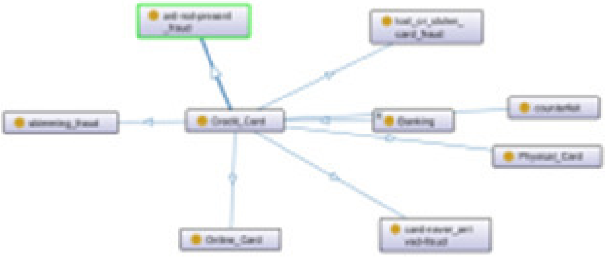


Fig. 4. The credit card fraud class.

4) Money Laundering: Money laundering is described as a technique for hiding proceeds of crime including transporting cash out of the country, purchasing businesses through which funds can be channeled, buying easily transportable valuables, transfer pricing, and using underground banks. [8] Money laundering is an organized crime with sophisticated multinational financial operations that transform proceeds of drug trafficking and other crimes into clean money. [9] This ontology represents the components used in money laundering such as people, organization, portfolios, messages, communication medium, and documents. We categorized people as sub concepts. These are individuals who participate in the money laundering cycle. It consists of several entities. Portfolio: The portfolio represents financial assets and products, it has many subclasses such as investment, transport, and other businesses. This class is associated with "organization", people, companies, and banks". For instance, people own accounts linked to a bank that will be used to purchase or invest in a particular business. Organization: This class describes organizations that are engaged in criminal acts of money laundering, mostly they are well-structured organizations, with many entities such as banking, securities trading, and so on. Messages: Represents all messages exchanged in the domain between people and organization. This is considered as the breaking ground or first stage of money laundering. All the activities in the domain are performed via messages, such as bank messages, trade messages. This class is associated with entities, people, organization, portfolio, and communication medium. Communication medium: This class of communication medium consists of standard and encrypted communication with subclasses such as anonymous proxy server, electronic payment server, and mail server. This entity is in relations with class of message, people, and organization. Document: The class contains all documents that can be provided by the person for identification purposes. It has many subclasses such as national card ID and passport. This entity is only associated with the entity people.

B. Identity Theft

Identity theft is considered as a mind game that favors hackers for the purpose of tricking users. There are various classes of identity theft, including email fraud. Most fraud detection researchers focused on non-semantics features of identity theft. In this type of fraud, the fraudsters try to steal none identity convincing the victims they are communicating with the intended person. In this area we classified several classes of identity theft from email to phishing and other scams.

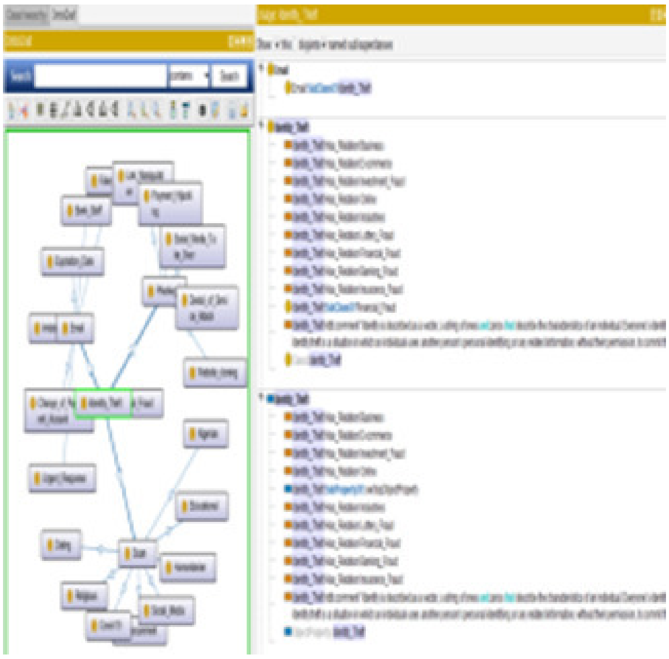


Fig. 5. The concept of identity theft.

1) Email Fraud: The use of email as a means of official and non-official communication is increasing worldwide. The number of emails sent and received globally has increased each year since 2017. While roughly 306.4 billion emails were estimated to have been sent and received each day in 2020, this figure is expected to increase to over 376.4 billion daily mails by 2025. [2] Claim that email fraud is more structured and more successful that in each 10 fraudulent emails sent, seven are delivered in the inbox, and estimated three are responded. More precisely, the incremental trust and usage of the service provide an opportunity to fraudsters to exploit the vulnerabilities. These vulnerabilities can be from both ends: sender, receiver as well as the email service providers. In our implementation we classified email fraud into various subclasses.

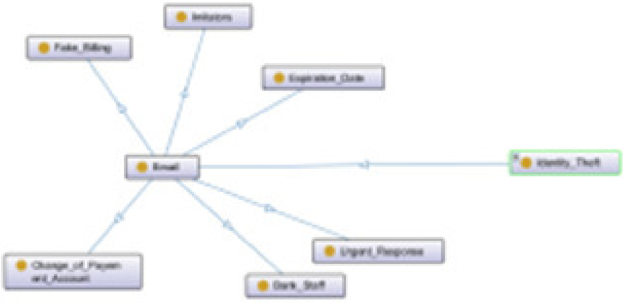


Fig. 6. The sub concepts of email fraud.

2) Phishing fraud: Refers to the spam that are sent by fraudsters in order to deceive their victims and obtain their personal information. Fraudsters can impersonate a service provider or institute that victims collaborate with. The fraudsters can make use of convincing techniques to obtain the victim's personal information including credit card details. The spam may also include links to fraudulent websites which again can deceive victims into revealing their personal information. In this class of ontology, we have several entities.

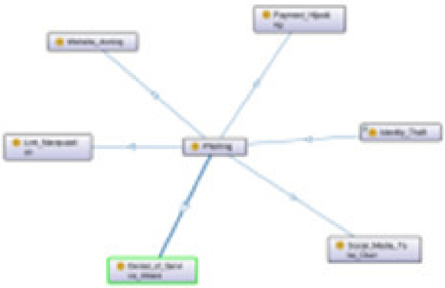
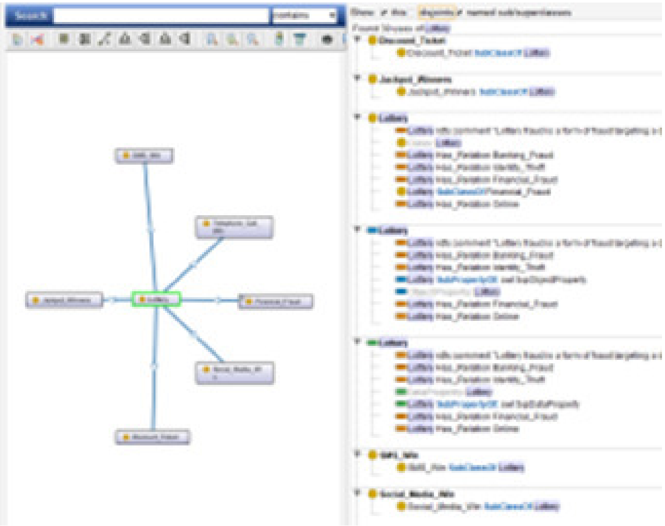


Fig. 7. The sub concepts of phishing fraud.

C. Lottery Fraud

The current trend of lottery fraud cannot be ignored, especially amongst elderly people. In this research, there are several behavior classifications of lottery fraud victims. On the other hand, the ontology provides a comprehensive semantic property to the kind of lotteries where fraud is more likely. For example, social media lottery fraud is more concurrent in developing countries than developed countries. This could be due to the fact that in developed countries people tend to have more available information in these kinds of fraud. While the fraudsters target elderly people, there is a concern the new trend of lottery fraud is shifting toward younger and desperate individuals.



Found Usages of Lottery

Discount_Ticket
Discount_Ticket SubClassOf Lottery

Jackpot_Winners
Jackpot_Winners SubClassOf Lottery

Lottery
Lottery rdfs:comment "Lottery fraud is a form of fraud targeting [illegible]
Class: Lottery
Lottery Has_Relation Banking_Fraud
Lottery Has_Relation Identity_Theft
Lottery Has_Relation Financial_Fraud
Lottery SubClassOf Financial_Fraud
Lottery Has_Relation Online

Lottery
Lottery rdfs:comment "Lottery fraud is a form of fraud targeting [illegible]
Lottery Has_Relation Banking_Fraud
Lottery Has_Relation Identity_Theft
Lottery SubPropertyOf owl:topObjectProperty
ObjectProperty: Lottery
Lottery Has_Relation Financial_Fraud
Lottery Has_Relation Online

Lottery
Lottery rdfs:comment "Lottery fraud is a form of fraud targeting [illegible]
Lottery Has_Relation Banking_Fraud
Lottery Has_Relation Identity_Theft
DataProperty: Lottery
Lottery SubPropertyOf owl:topDataProperty
Lottery Has_Relation Financial_Fraud
Lottery Has_Relation Online

Gift_Win
Gift_Win SubClassOf Lottery

Social_Media_Win
Social_Media_Win SubClassOf Lottery

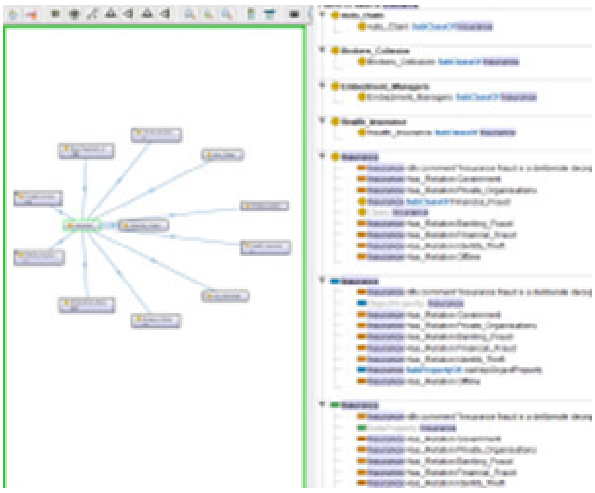
Fig. 8. The concept lottery fraud.

D. Insurance Fraud

Insurance fraud in its many forms is seen as a deliberate deception perpetrated against or by an insurance company or agent for the purpose of financial gain. Fraud may be committed at different points in the transaction by applicants, policyholders, third-party claimants, or professionals who provide services to claimants. We classified insurance fraud into internal and external, the sub concepts of this fraud are classified into health, vehicle, life insurance. We also presented common patterns of insurance frauds including padding, inflating claims, misrepresenting facts, and fake license agents.

E. Investment Fraud

The comprehensive classification of investment fraud is provided in this ontology. There are various existing investment frauds, some are old enough while some are new due to technological advancement. In investment fraud, existing investors are paid with funds collected from new investors. For example, Ponzi schemes whereby organizers often promise investors a huge return on investment with little or no risk. We proposed several sub concepts of investment fraud, some are targeting big investment, while others target small and medium investors. These are Ponzi scheme, pyramid scheme, pump and dump, offshore, gold as well as boiler room and advance fee.



Auto_Claim
Auto_Claim SubClassOf Insurance

Brokers_Collusion
Brokers_Collusion SubClassOf Insurance

Embezzlement_Managers
Embezzlement_Managers SubClassOf Insurance

Health_Insurance
Health_Insurance SubClassOf Insurance

Insurance
Insurance rdfs:comment "Insurance fraud is a deliberate decep[illegible]
Insurance Has_Relation [illegible]
Insurance Has_Relation Private_Organisations
Insurance SubClassOf Financial_Fraud
Class: Insurance
Insurance Has_Relation Banking_Fraud
Insurance Has_Relation Financial_Fraud
Insurance Has_Relation Identity_Theft
Insurance Has_Relation Offline

Insurance
Insurance rdfs:comment "Insurance fraud is a deliberate decep[illegible]
ObjectProperty: Insurance
Insurance Has_Relation [illegible]
Insurance Has_Relation Private_Organisations
Insurance Has_Relation Banking_Fraud
Insurance Has_Relation Financial_Fraud
Insurance Has_Relation Identity_Theft
Insurance SubPropertyOf owl:topObjectProperty
Insurance Has_Relation Offline

Insurance
Insurance rdfs:comment "Insurance fraud is a deliberate decep[illegible]
DataProperty: Insurance
Insurance Has_Relation [illegible]
Insurance Has_Relation Private_Organisations
Insurance Has_Relation Banking_Fraud
Insurance Has_Relation Financial_Fraud
Insurance Has_Relation Identity_Theft

Fig. 9. The concept of insurance fraud.

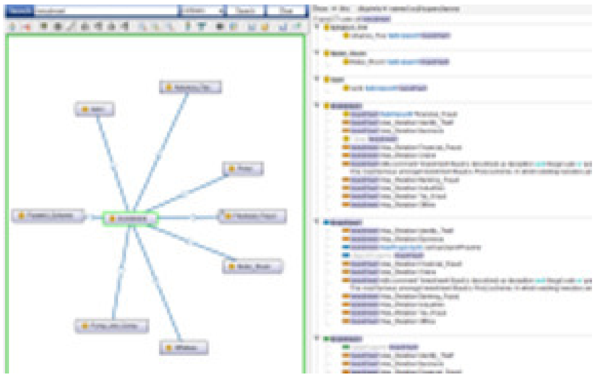


Fig. 10. The concept of investment fraud.

F. Tax Fraud

This section includes the process of classifying, constructing the various ways of tax evasion and as a means of tax fraud. The aim is to understand the tax collection structure and the main challenges they faced. Hence, we classified tax fraud into several sub concepts, from tax evasion, fake income, report, corporate tax fraud. We conclude that while avoiding tax burden is legal, there are several ways that individual, and organizations avoid tax illegally. These provide us with other entities, properties, and individual tax noncompliance activities.

G. Transaction Fraud

The transaction fraud was extracted from one of the datasets we obtained for the purpose of validating the ontology. Initially this concept of fraud was not in the ontology. We proposed

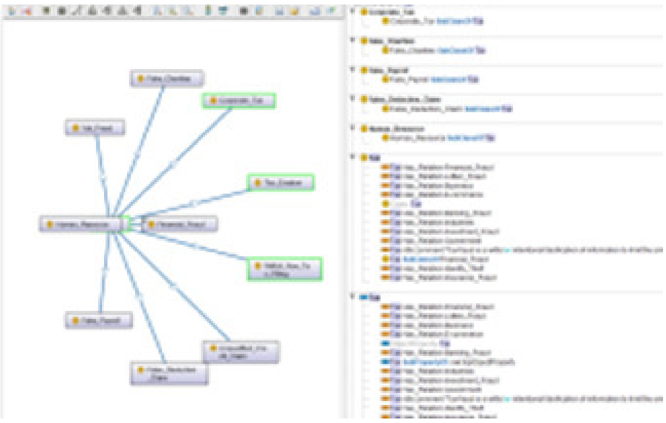


Fig. 11. The concept of tax fraud.

transaction fraud to separate it from the payment fraud, which already exists in the ontology. The concept of transaction includes various sub concepts. From fake alerts, to fake items, items not delivered, service not rendered, etc. The properties of this concept are inline and related with various concepts and sub concepts of frauds in the ontology. Hence, it is one of the most relatively related concepts in this research.

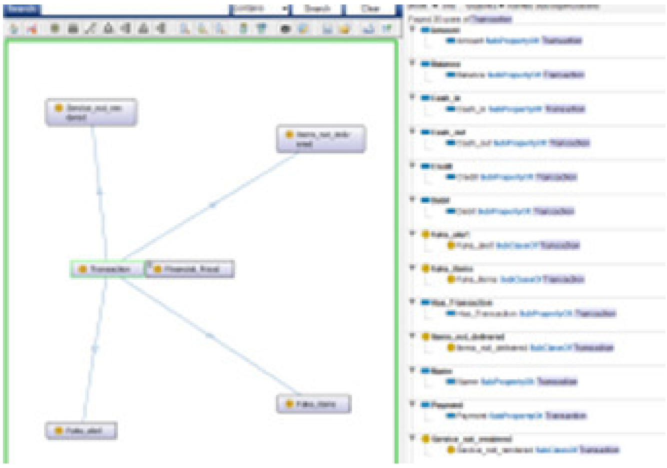


Fig. 12. The concept of transaction fraud.

H. Forex Fraud

Foreign exchange is a decentralized market for the trading of currencies, unfortunately the market has been penetrated by fraudsters, in which various frauds are proposed to the customer. In this research we focused on this fraud and declassified them into a semantic context. This concept has very few relations with other concepts in the ontology, all of its sub concepts have no relations with any concept or sub concept in the ontology. Nevertheless, the object properties have some common relation with investment fraud.

I. Cyber Violence Fraud

The act of Bullying was specifically known for teenage ages especially in schools. Cyber violence is often just other sides of the same coin. The act aims at causing emotional and physical violence for the purpose of monetary exploitation from the victims. While conducting the research, we contacted some experts in cyber violence and bullying for understanding the various forms of cyberbullying. This act of fraud is on the rise especially in developed countries. We classified this concept into various sub concepts which clearly diversified the concept into a significant act of violence and fraud. While investigating this concept we understood how gender plays an important role in this fraud, for example a sexism sub concept is mainly targeted at females. While social exclusion is targeted at the male. There are not many relations between this concept and other concepts in the ontology, but identity theft and dating scam have a direct relation with this concept in the ontology.

Fig. 13. The concept of forex fraud.

Fig. 14. The concept of cyber violence fraud.

VI. Validation of Ontology

This ontology was designed and built based on multiple layers. Hence, the validation should be based on layer by layer or component to components. The aim of this section is to determine a concise and acceptance of the ontology to the experts and users. We identified a set of ontology quality criteria and ontology aspects that must be achieved, including accuracy, adaptability, clarity, completeness, computational efficiency, and consistency. There are several methods of ontology validation using tools and other techniques. While tools are good with ontology consistency validation, we consider several techniques that are more consistent with knowledge and content validation. The dataset validation was performed in two ways.

A. Dataset Based Validation

This type of validation is done by using fraud-related datasets to validate the ontology, in this section we elaborate some of the dataset. The challenge with this type of validation is finding the correct data that is in line with the concepts of the ontology. Hence, we explore several options and various data sources. Financial fraud data: The dataset obtained from Kaggle, has various features simulated for mobile money transactions. The technique of concept extraction from the dataset was done by means of distributing every feature into a particular class as ontology. From the features we extracted sub concepts, entities, properties, as well as individuals.

- Matching the semantic content of the ontology concept with the dataset features. This process explored the similarity function to find the matching words that appear in both files. While using a dictionary function to identify words with several meanings to avoid missing matching. This process of validation is considered the most acceptable type of ontology content validation.
- Concept extraction validation. In this process, we studied the features of the dataset in order to extract a possible ontology concept from the dataset, then we checked the ontology if the concept is existing. Else it will be added into the ontology. Using this type of validation, we were able to add several concepts into the ontology, which include transaction fraud, forex fraud, and crypto currency fraud concepts.
- Insurance fraud data: The dataset contains insurance claim data for fraud detection, in the ontology we built several concepts including insurance fraud. Some examples of features in this dataset are employee data which include the data of the employee or agent working in the insurance company. This data is used to validate the claim of internal insurance fraud in our ontology. While the feature of vendor data is for the external agent who assists the insurance company in investigating the claims. As suggested in our ontology vendors are sometimes involved in external insurance fraud. This feature was used to validate the broker's collision class in the ontology.
- Bank Fraud Data: The banks are often exposed to fraud transactions in various forms, we use the dataset to validate the concept of banking fraud in the ontology. From the various features in the dataset, we construct the concept of banking fraud and it exactly tallies with the concept of and entities of banking fraud that exist in the ontology.
- Fraudulent Email Data: This dataset contains criminally deceptive information, usually with the intent of convincing the recipient to give the sender a large amount of money. This dataset is a huge collection of fraudulent letters sent to various recipients. The features include sender, receiver, message id, date, subject, etc. We explore this dataset and extract the concept of email fraud, in our ontology email fraud is a sub concept of identity theft. We also used this dataset to validate Nigerian scam, which is a class of fraud under sub concept termed scam.
- Cryptocurrency Fraud Data: This dataset contains rows of known fraud and valid transactions made over ethereum, a type of cryptocurrency. The various features in the dataset makes it possible to have enough entities to create the cryptocurrency concept in our ontology. Hence, we proposed a class of cryptocurrency fraud concept in the ontology. The properties of individuals as well as entities of this class were not in the ontology before we started the validation process.
- Financial fraud data: The dataset obtained from Kaggle, has various features simulated for mobile money transactions based on a sample of real. We explore these features. Concept extraction from the dataset was done by means of distributing every feature into a particular class as ontology. From the features we have those in sub concepts, entities, properties, as well as individuals. From the features we extracted, the transaction fraud concept is extracted from the dataset.
- Insurance fraud data: The dataset contains insurance claim data for fraud detection. In our ontology we built several concepts including insurance fraud. From the features, employee data include the data of the employee or agent working in the insurance company. This data is used to validate the claim of internal insurance fraud in our ontology. The feature of vendor data is for the external agent who assists the insurance company in investigating the claims. As suggested in our ontology vendors are sometimes involved in external insurance fraud. This feature was used to validate the broker's collision class in the ontology. The claims data in the dataset is for the claim-level transaction details submitted by the customer to the insurance company for reimbursement. The feature was used to validate the auto claim fraud class in the ontology.

VII. Discussion

The research proposed heavyweight ontology for financial fraud due to the multi-layer structure of the ontology. The elaboration of the ontology was done with a full compliance of the state-of-the-art methodology and advice received from the domain experts in addition to adopting various ontology development articles from several ontology developers' community. The proposed ontology consists of a number of concepts, sub concept, properties, and other semantic annotations.

The proposed domain areas constitute eleven concepts and several sub concepts. In this research work, we are constructively convinced; these areas are of great importance in the financial sector. Hence, understanding the fraud concepts will reduce a significant number of losses in the sector.

A. Semantic contents and related information

- Documentation: It provides a consistent and centrally non-technical description of the semantic content of the concept of the ontology such as concepts, classes, individual resources and properties. The documentation does not consider the proprietary format in which the information is stored in the ontology platform.
- Navigation: Is used to provide links between the information sources established, which can be employed for navigation between resources. Consider for instance when a user tries to navigate from the class of banking fraud object properties to the class of identity theft fraud data properties and vice versa.
- Retrieval: As the annotations are represented in a formal machine-interpretable ontology language, they can be exploited by appropriate browsing and querying tools.
- Definition We added definitions of entities in the individual section of the ontology. The definition will provide users with the most accurate definitions of the entity, including a semantic knowledge of the concept. For example, in Figure 1. We can see the definition of Government as described in the individual section of the ontology.
- Restrictions In OWL there are two kinds of restrictions, existential restrictions, that describe classes of individuals which participate in at least one relationship along a specified property to individuals that are members of a specified class. For example, the concept of banking fraud has a relationship called 'has a part of' with members of the Identity theft fraud concept. While universal restrictions describe classes of individuals that only have relationships along with a property that are members of a specified class.
- Visualization For a user to be able to view the ontology, there is a need for an ontology development, or ontology view platform. We recommend using Protégé ontology platform or Vowl ontology viewer to navigate to windows, then view options, there are several view options from OBO, Owl viz, Onto graph and many more.
- Ontology Instances The instances in ontology are lists of individuals that are asserted to be instances of classes. The classes are selected and assigned to each instance. For example, financial fraud is a concept and 'business' is an instance of that concept. It could be argued that business is a concept representing the different instances of financial fraud and its isotopes, etc. This is a well known and open question in ontology building. However, deciding whether something is a concept of an instance is difficult, and often depends on the application and domain.

VIII. Conclusion and future work

The paper proposed an ontological approach to various concepts of financial fraud, which could serve as first of its kind. The elaboration of the ontology provides a comprehensive and explicit terminology of financial fraud activities, we explored creating an ontological catalogue that defines the various types of fraud committed. This was largely helped by the dataset validation technique adopted. The technique allows the flexibly created fraud ontology to be modified and additional definitions made.

The validation approach used in this research work opens doors to the ontology developer's community to implement such an approach for better accuracy of ontology developed. The availability of datasets makes it possible to achieve this type of validation. But in the case of fraud and other cyber security-related ontologies, there could be a difficulty in obtaining the correct data and information needed, due to the sensitivity of the research.

We found a number of the dataset's attributes in our ontology. Using a dataset classified as transaction fraud, we perform concepts and properties extraction in the dataset, and we find a match with the number of properties in our ontology. But not enough as the unmatched attributes are more than the matched. Hence, we create the concept of transaction fraud in the dataset.

Few improvements could be included in the future work. The paper lacks semantic detection reasoning, this shall be implemented in the future work. The detection reasoning will allow for real time detection of various kinds of financial fraud. Obviously, comparison of this method with other machine learning approaches will as well be implemented. The future work identified in this section (semantic detection reasoning) will make the work an asset to financial institutions globally.

References

[1] Panos Alexopoulos, Kostas Kafentzis, Xanthi Benetou, Tassos Tagaris, and Panos Georgolios. Towards a generic fraud ontology in e-government. In *ICE-B*, pages 269–276, 2007.

[2] Koen Kerremans, Yan Tang, Rita Temmerman, and Gang Zhao. Towards ontology-based e-mail fraud detection. In *2005 portuguese conference on artificial intelligence*, pages 106–111. IEEE, 2005.

[3] Gonçalo Carnaz, Vitor Nogueira, and Mário Antunes. Ontology-based framework applied to money laundering investigations. In *Proceedings of the Seventh Conference on Informatics at the University of Evora*, pages 1–17. University of Evora Evora, 2017.

[4] Steven L Skalak, Thomas W Golden, Mona M Clayton, and Jessica S Pill. *A guide to forensic accounting investigation*. John Wiley & Sons, 2011.

[5] Johnson Olabode Adeoti. Automated teller machine (atm) frauds in nigeria: The way out. *Journal of Social Sciences*, 27(1):53–58, 2011.

[6] Linda Delamaire, Hussein Abdou, and John Pointon. Credit card fraud and detection techniques: a review. *Banks and Bank systems*, 4(2):57–68, 2009.

[7] S Benson Edwin Raj and A Annie Portia. Analysis on credit card fraud detection methods. In *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pages 152–156. IEEE, 2011.

[8] Michael Levi. Money laundering and its regulation. *The Annals of the American Academy of Political and Social Science*, 582(1):181–194, 2002.

[9] Michael Levi and Peter Reuter. Money laundering. *Crime and justice*, 34(1):289–375, 2006.