Leveraging knowledge sharing for preventing and investigating on-line banking frauds:
On-line Fraud Centre

Salvatore Camillo Zammataro

Technical Report
RHUL–MA–2014–2
01 September 2014

www.ma.rhul.ac.uk/tech
Leveraging knowledge sharing for preventing and investigating on-line banking frauds:
On-line Fraud Centre

MSc Information Security Project Report

Supervisor: Dr John Austen

Submitted as part of the requirements for the award of the MSc in Information Security at Royal Holloway, University of London.
EXCLUSIVE SUMMARY

For social, technological and political reasons electronic transactions are now the most used payment method in Europe. As such, fraudsters have been focusing on On-Line transactions to gain money through Phishing and Crimeware. These kind of frauds generates losses for EU citizen of an estimated value of 250M€/year.

Banks and Law Enforcement Agencies are engaged in the prevention, detection and prosecution of this crime. Some limit of actual legislation (i.e. Data Protection, International Treaties on cyber-crime, Fraud prosecution laws), low speed of communication between Banks and LEAs, and the fraudster’s speed in taking advantage of weaknesses of the system leave the space for improvement.

To improve the countermeasures that Banks and LEAs have deployed, this paper suggests the adoption of an InfoSharing Service between national banking system and Law Enforcement Agencies. This service uses a “hub-and-spoke” framework, where LEA is the hub and banks are the spokes.

Target of the service is to shorten the time needed to communicate from Banking Fraud Managers to LEAs, and to share the relevant information on fraudster accounts in the whole banking industry.

The first target will permit to block in shorter time a larger amount of fraudulent transactions.

The second target will help all the banks to update their Fraud Management Systems’ “watching lists” and “black lists” to better cope with fraudster attempts and block them before the money is transferred.

This paper is focused on the description of the context and of the design of the proposed InfoSharing Service in terms of roles, law constraints, organization and technology.
# Table of Contents

Executive Summary ........................................................................................................... 2

Table of contents .............................................................................................................. 3

List of tables ...................................................................................................................... 4

List of figures .................................................................................................................... 5

1 Preface & Objectives ..................................................................................................... 6

2 Context Analysis ............................................................................................................ 8

   2.1 On-line fraud methodologies ................................................................................. 9

   2.2 The dimension of the problem ............................................................................. 14

   2.3 The reaction: how banks and law enforcement work today? ............................... 17

3 The Online Fraud Centre ............................................................................................... 21

   3.1 Operational Framework Definition ....................................................................... 22

   3.2 Legal Constraints ................................................................................................. 24

   3.3 Definition of the centre operations ...................................................................... 25

   3.4 Platform Description ......................................................................................... 27

4 Possible evolutions ....................................................................................................... 64

5 Conclusions .................................................................................................................. 65

Bibliography ..................................................................................................................... 67

Glossary ............................................................................................................................. 68
LIST OF TABLES

Table 1 - some European countries estimation about on-line banking frauds – M€.......................... 14
Table 2 - On-Line Losses in M€ in UK 2004-2011 (16)................................................................. 15
Table 3 - Polizia Postale - On-Line Fraud data - 2009-2012.......................................................... 16
Table 4 - Provided Functionalities for each user interface............................................................ 23
LIST OF FIGURES

Figure 1 - Proposed Fraud Taxonomy Framework .......................................................... 9
Figure 2 - Area of interest of this project in terms of frauds methodologies ........................................... 10
Figure 3 - Time needed to close a Phishing Website (11) ............................................. 12
Figure 4 - Example of Paypal phishing webpage (Phishtank.com 2012) .................................... 12
Figure 5 - typical cross-functional fraud organization in a bank ..................................................... 17
Figure 6 - LEAs typical Investigating procedure .............................................................................. 19
Figure 7 - External Interface Home Page ....................................................................................... 31
Figure 8 - External Interface - Level 2 ......................................................................................... 31
Figure 9 - External Interface - Level 3 .......................................................................................... 32
Figure 10 - External Caption - Level 4 .......................................................................................... 32
Figure 11 - External Interface utilization schema ............................................................................ 33
Figure 12 - Internal Interface - Level 1 – Homepage ..................................................................... 33
Figure 13 - Internal Interface - Level 2 – Tracker ........................................................................... 34
Figure 14 - Internal Interface - Level 2 – Correlation ...................................................................... 35
Figure 15 - Internal Interface - Level 3 – New Correlation .............................................................. 35
Figure 16 - Internal Interface - Level 3 - Notification .................................................................... 36
Figure 17 - Architecture Overview ............................................................................................... 39
Figure 18 - Single Firewall / Dual Firewall DMZ .......................................................................... 40
Figure 19 - Architecture Hypothesis 1 ............................................................................................ 41
Figure 20 - Architecture Hypothesis 2 ............................................................................................ 41
Figure 21 - Architecture Hypothesis 3 ............................................................................................ 42
Figure 22 - Architecture Hypothesis 4 ............................................................................................ 42
Figure 23 - Architecture Hypothesis 5 ............................................................................................ 43
Figure 24 - Architecture Hypothesis 6 ............................................................................................ 43
Figure 25 - Database Hypotheses .................................................................................................. 46
Figure 26 - ER Model (Part 1) ........................................................................................................... 48
Figure 27 - ER Model (Part 2) ........................................................................................................... 49
Figure 28 - ER Model (Part 3) ........................................................................................................... 49
Figure 29 - ER Model (Part 4) ........................................................................................................... 50
Figure 30 - ER Model (Part 5) .......................................................................................................... 50
Figure 31 - ER Model (part 6) ............................................................................................................ 51
Figure 32 - Database tables (Part 1) ............................................................................................... 52
Figure 33 - Database tables (Part 2) ............................................................................................... 52
Figure 34 - Database tables (Part 3) ............................................................................................... 53
Figure 35 - Database Tables (Part 4) .............................................................................................. 53
Figure 36 - Database tables (Part 5) ............................................................................................... 54
Figure 37 - Database tables (Part 6) ............................................................................................... 55
Figure 38 - Database tables (Part 7) ............................................................................................... 56
Figure 39 - Database UML (Part 1) ............................................................................................... 57
Figure 40 - Database UML (Part 2) ............................................................................................... 57
Figure 41 - Database UML (Part 3) ............................................................................................... 58
Figure 42 - Database UML (Part 4) ............................................................................................... 58
Figure 43 - Database UML (Part 5) ............................................................................................... 59
Figure 44 - Communication schema 1 ............................................................................................ 60
Figure 45 - Communication schema 2 ............................................................................................ 61
Figure 46 - Communication overview (Hypothesis 1) ...................................................................... 62
Figure 47 - Communication overview (Hypothesis 2) ..................................................................... 63
1 Preface & Objectives

The strong will to trace financial transactions followed by the European Community in the last years, mainly in order to minimize bribery and money-laundering, and the steady progression of the digital age, pursuing the ability to interact remotely through a digital seamless network, has been pushed forward and have obtained many successes. This will be further regulated when the EU directive on electronic transactions will be released in the next months (1) and the effects of PSD2 (2), recently adopted, will spread. Capabilities to follow the electronic trace of money transactions is leading to a more accurate investigation of economic crimes, and “electronic transactions” in Europe have reached and surpassed the number of traditional cash/paper transactions, either in number than in value (3).

The ever increasing quantity of money transferred and transferrable by electronic means has had a drawback in terms of the number of frauds leveraging the possibility to mis-behave digital transactions, often in automatic or semi-automatic approach.

The convergence of electronic money and digital age has brought to a new era of frauds, with different methodologies, described in next pages, used to move the money to criminal accounts without the knowledge or will of the owner.

Global, European and single state regulations are trying to cope with this kind of behaviour, in order to give to police forces and to judges the tools needed to block and prosecute such crimes, but many problems arose when dealing with these “smart-frauds”: Criminal people tend to act cross-border, using the gaps between the different legislations and with the protection of companies that get revenue from these illegal activities.

The cyber-crime raise, apart from the state-backed cyber intelligence, is now an industry that has an estimated turnover of some billions Euros, with an estimated direct damage to EU citizens in the order of 0,08% of EU total sales transactions (4 p. 145).

It’s now difficult to follow the electronic traces of these frauds without a correct cooperation between cyber security experts, present mainly in private companies (i.e. banks, telcos, ...) and in law enforcement agencies.

The correct interaction between private and public can foster great results in prosecuting cyber-crime, on one hand leveraging the focused competencies that have grown in specific sectors (i.e. money-laundering professionals in banks, voice and data fraud managers in telcos, ...) with the broad and full of different experiences Law Enforcement Professionals, now more than ever called to combat the day-by-day activity of the digital criminals.

Being able to have a deep competence coupled with broad crime experience, and eventually acting as a node into to the whole EU anti-fraud network, can detect, block and follow a higher percentage of crimes.
This project is born in this context in order to propose the possibility to:

- provide a Point of Contact on the Cyber Frauds for Banks, Financial Institutions and Credit Card issuers
- collect Information on the suspicious activities, the frauds and the attacks detected by Banks, Financial Institutions, Credit Card issuers, other Police Forces;
- support the Investigations and the Response to Frauds (shutdown requests, evidence collection);
- analyse new threat scenarios, frauds and attacks techniques;
- create an Information Exchange system with national and international organizations (other law enforcements, banks, credit card issuers) in order to multiply the efficacy of the services of the Fraud Centre.

The project has the objective to evaluate and propose an approach to fight the e-crime through a strong cooperative network of actors involved in the detection and the management of specific electronic crimes, especially on-line frauds; this approach needs a common standard and a sharing platform to exchange information between financial institutions, law enforcements agencies and financial associations.

Law enforcement agencies could also develop procedures to analyse the information collected by the different partners.

The structure of the report is as follows:

- Chapter 2: Analysis of the Scenario
  In this section will be discussed the AS-IS Scenario based upon data gathered from the different stakeholders and the direct interviews with relevant experiences in the banks and the law enforcement.

- Chapter 3: The proposed approach
  This section will elaborate how a specific “knowledge sharing solution” could solve some of the issues found in the analysis. It will describe some possible implementation model to support the proposed approach in terms of roles, law constraints, organization and technological solutions.

- Chapter 4: Possible evolutions
  During the analysis of the context and the design of the InfoSharing Service, 5 areas of improvement have been identified.
2 Context Analysis

Fraud is an essential part of human nature, being born with the business of trading thousands of years ago. We have information that long time ago biased balances were used in markets, in order to fraud on the weight of food or gold. Later, with the invention of money and banks, fraudsters tried to print false banknotes and coins. When banknotes became bank letters and subsequently cheques, identity theft became the usual attack in order to obtain the value printed on the (original or counterfeit) cheque.

Little has changed from a conceptual point of view about identity theft: the target for a fraudster is to be trusted by the banking/credit system, thus obtaining access to a credit line or to a banking account.

The different techniques used by criminals in order to gain money from banking accounts and credit cards are mostly based upon the illegal access to credit card details (Name/Surname, PAN, SID, Expiration date) or to the access to banking accounts (ability to interact and dispose of transactions). In the pre-digital age this was accomplished with the physical interaction at the desk or with the usage of telephones.

What has really changed in the last 20 years is the capacity to scale and the tools used. In the ‘60s the technical mean in order to get the trust of the banks was to print counterfeit banknotes, cheques or identity cards. Now the ability has become to act as a trusted user on a banking system based upon a computerized system. This change in banking tools, from paper to personal computers, has permitted to scale the same attack from a single, manual episode to multiple, automated attack clones.

This generated the strong need to evaluate, study, invent new ideas to prevent and avoid the losses.

The sections of this chapter are meant to:

- Section 2.1 – Describe On-line fraud methodologies

  Having a clear idea on the different fraud methodologies helps to understand information and time constraints needed to detect, prevent and react. Furthermore this can help to focus on On-Line banking losses, as a subset of the overall on-line frauds, to verify and test the solutions proposed in the next chapters.

- Section 2.2 – List some numbers on frauds and losses suffered by Banks and their customers

  Understanding (or, better, estimating) the amount of transactions and related money illegally transferred by fraudsters on banking accounts can give some numeric indicator about the dimension of this criminal activity.

- Section 2.3 – Describe actual approach to prevent losses by Banks and LEAs (Law Enforcement Agencies)

  Either banks than LEAs are acting in order to prevent that citizens and customers suffer these losses: this section describes the actual model these two actors adopt, based upon interviews with Fraud Managers and LEAs Officers.
2.1 **On-line Fraud Methodologies**

2.1.1 **Taxonomy**

Criminals use many different techniques to gain access to money. In order to have a clear model to classify and study the different techniques and related countermeasures, a fraud taxonomy should be adopted by the project. Many different approaches have been proposed, such as in (5) but, as well analysed by (6 p. 6-10) and (7 pp. 2-4) a comprehensive taxonomy that could be used for our purpose is still in development. Existent academic taxonomies are inapplicable to LEAs’ and Banks’ needs; Market taxonomies, tailored to specific needs, like (8), are too broad/generic. For these two reasons the project will adopt a simplified model, inspired to the UK SOCA taxonomy of fraud methodologies (as presented in (8) and (9)) adapted to the schema presented below.

![Proposed Fraud Taxonomy Framework](image)

*Figure 1 - Proposed Fraud Taxonomy Framework*

This framework helps to focus on the study area of this project:

- **On-line**
  - Banking

Due to the banking customer segmentation model, Enterprise, GE and NGO are usually referred as “Corporate Customers”, while Individuals are considered “Consumers”.

To further clarify the boundaries of the fraud methodologies considered in this study, next image will be taken as reference.
2.1.1.1 Out-of-scope Fraud methodologies

The project will not consider the following fraud methodologies:

- Advance fee frauds
  o Career opportunity scams
  o Clairvoyant or psychic scams
  o Cheque overpayment fraud
  o Dating or romance scams
  o Fraud recovery fraud
  o Impersonation of officials
  o Inheritance fraud
  o Loan scams
  o Lottery, prize draw and sweepstake scams
  o Racing tipster scams
  o Rental fraud
  o West African letter or 419 fraud
  o Work from home and business opportunity scams
  o Vehicle matching scams
- Business directory fraud
- Bank card and cheque fraud
- Internet auction fraud
- Mass marketing fraud
- Online shopping fraud
- Plastic card fraud
- Rental fraud
- Domain name scams
2.1.1.2 Fraud Methodologies considered in this project
The project will be focuses on these fraud methodologies:

Account takeover
“An account takeover can happen when a fraudster or computer criminal poses as a genuine customer, gains control of an account and then makes unauthorised transactions. Any account could be taken over by fraudsters, including bank, credit card, email and other service providers” (Action Fraud, 2010). The techniques adopted in order to takeover an on-line banking account are based upon:

• Phishing
• Spyware/Crimeware
• Malware scams.

Identity theft frauds
As defined by (10) “Identity theft occurs when someone appropriates another’s personal information without their knowledge to commit theft or fraud. Identity theft is a vehicle for perpetrating other types of fraud schemes.”

Usually the on-line fraud starts with an identity theft, used actively to gain access to a credit line or to a banking account.

The different techniques used by criminals in order to gain money from banking accounts and credit cards are mostly based upon the illegal access to credit card details (Name/Surname, PAN, SID, Expiration date) or to the interactive access to on-line banking accounts (ability to interact remotely and dispose of transactions).

For the subsequent chapters, in order to focus and analyse a clear boundary of fraud methodologies actors, victims and data that has to be exchanged, the project will consider the banking-related frauds:

• Suspect fraudulent transactions done through internet banking (typically originated by phishing or crimeware)
• Suspect IPs detected by banks as potential mules or C&C systems
• Banking clone sites

Money muling
“Money mules are often recruited by fraudsters to receive money into their bank account, then withdraw the money and wire it overseas, minus a commission payment. Even if the money mule is not involved in the fraud to generate the money, they are acting illegally by laundering the funds” (Action Fraud, 2010).
2.1.2 Some example of fraud

2.1.2.1 Fraud by Phishing

Phishing has historically been the first methodology adopted to fraud users of an on-line banking system. Technically a phishing attack is based on a 3 step process:

1. Spread through SPAM mail of phishing email, resembling the emails sent by the financial company. Aim of these SPAM is to deceive the user in clicking on the link in order to access the...
2. Fake website, with the specific purpose to clone the real on-line banking website. This is needed in order to convince the bank customer to insert her/his credentials (user/pwd and/or other authentication mechanism in place) in this fake page. The attacker at this point can...
3. Use captured credentials to launch independently from the real user fraudulent transactions stealing money from the attacked account

With this “simple” approach millions of credentials were stolen in the period 2005-2009. Nowadays phishing websites are usually closed in terms of hours when detected by the bank, the specialized internet security companies or the interested LEAs, as depicted in the next figure by (11)

![gTLDs Average Phishing Uptimes 1H2012](image)

*Figure 3 - Time needed to close a Phishing Website (11)*

This screenshot depicts one example of phishing website (many can be found on-line on phishtank.com):

![Example of PayPal phishing webpage](image)

*Figure 4 - Example of PayPal phishing webpage (Phishtank.com 2012)*
Phishing attempts are now less frequent and effective than 3 years ago, due probably to the fact that the general awareness on this threat has risen in the internet community and to the fact that the monetization of straight phishing attempts is no more working (see section 2.3.1 to understand the reasons behind this).

That is why most attacks now are based upon Crimeware, described in the next section.

2.1.2.2 Fraud by Crimeware
As defined by (12), the Crimeware is “a class of malware designed specifically to automate cybercrime” built to “perpetrate identity theft in order to access a computer user’s online accounts at financial services companies and online retailers for the purpose of taking funds from those accounts or completing unauthorized transactions that enrich the thief controlling the crimeware. Crimeware also often has the intent to export confidential or sensitive information from a network for financial exploitation”.

The aim of crimeware is to steal credentials and/or modify actively the behavior of a user client (PC, Smartphone, ...) in the process of launching a bank transaction, usually with a “man-in-the-browser” approach.

Crimeware have been changing in the last 5 years in order to be effective and not discoverable by standard Antivirus program, as clearly dissected in (13). These toolkits, named Zeus, Citadel or Spyeye, are now so configurable and easy to use that anyone having light basics of informatics and programming can customize a standard version in order to attack a specific bank.

Nowadays, Crimeware is by far the most effective tool used to commit on-line banking fraud (14 pp. 77-78), while what remains of phishing attempts is focused only in the deception of users by email in order to force them to click in some “Trojan Downloader & Dropper” link that will start to install the crimeware.
2.2 THE DIMENSION OF THE PROBLEM

2.2.1 Data gathered from institutions

Focusing on European market in the last 7 years, precise public data has been difficult to find, mainly due to the fact that in many countries (i.e. Italy, Germany, France) the banks tend to restrict access to these data or to hide actual data for marketing reasons. In the non-Anglo-Saxon culture banks tend to avoid any information to the public related to loss or fraud, considering it a driver of loss of trust from the customer base. In some bank the secrecy (or “agreed silence”) on these data is so strong that in the same bank not even the single department relating to fraud activities has the whole picture. The fraud management department usually has the view on technical data (exploits, rules for detection, ...); customer care department has the view on the single fraud, being responsible for the communication to the customer; the fraud related law department deals with the cases that arise to a formal suit; typically only the financial department has a complete number on attempted frauds and real losses (after settlements / agreements).

This is one of the main motivation why even banking associations are not aware of the full dimension of the problem.

Other countries like UK, Belgium, and Nederland prefer to present the overall numbers, thus permitting to have a rather sharp idea on the on-line fraud amounts and trends. In these countries usually banks that behave better than others stress the fact that they combat fraud in line or better than the market, using this information for marketing.

Some number from these countries, compared to an estimation of Italian frauds derived from anonymous interviews to banking fraud managers and related LEAs (see 2.2.2 for specific details), are presented in the next table:

<table>
<thead>
<tr>
<th>Year</th>
<th>Belgium (15)</th>
<th>Nederland (15)</th>
<th>UK (15)</th>
<th>Italy¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>0,5</td>
<td>-</td>
<td>26,6</td>
<td>4,0</td>
</tr>
<tr>
<td>2008</td>
<td>0,2</td>
<td>-</td>
<td>61,7</td>
<td>11,8</td>
</tr>
<tr>
<td>2009</td>
<td>0,1</td>
<td>1,9</td>
<td>70,2</td>
<td>6,5</td>
</tr>
<tr>
<td>2010</td>
<td>0,0</td>
<td>9,8</td>
<td>55,0</td>
<td>3,3</td>
</tr>
<tr>
<td>2011</td>
<td>0,2</td>
<td>35,0</td>
<td>41,6</td>
<td>2,8</td>
</tr>
<tr>
<td>2012</td>
<td>3,0</td>
<td>35,0</td>
<td>47,0</td>
<td>4,1</td>
</tr>
</tbody>
</table>

Table 1 - some European countries estimation about on-line banking frauds – M€

What has been found in Italian studies by BA (14) coherently with what UK-FFA (16), (17) and Febelfin (15) found, puts in evidence the lifecycle of fraud management by banks in last 3 years, demonstrating

¹ Estimation based upon ABI, Polizia di Stato and interviews data
that the high numbers seen in 2006-20010 brought banks to react and adopt countermeasures that permitted the decline seen in 2009-2013; this is clear in the next histogram (16)

The estimation based upon the above data of the total loss suffered by citizens of 28 members of the European Union is between 250 and 280 M€ per year, with a prospective of light decrease in the next years due to the fact that immature or “niche” markets, that will be targeted by organized crime, will undergo to initial losses and then will decrease due to the countermeasures banks will put in place (see in example the trend of Italian and UK losses in Table 1 and Table 2).

A percentage of 1 economic loss every 1M transactions, in absence of configuration errors or specific vulnerabilities in the single bank, could be achievable in the long term for mature markets. This objective can be derived from the actual results obtained by banking fraud departments in mature markets like Italy (14 p. 41) and in recently attacked market like Belgium (15 p. 4).
2.2.2 Drill-down on specific interview with Fraud Managers and Experts
Due to the restrictions provided by banks on the publication of internal frauds data, the interviews have been anonymized. Data referred under the anonymizing agreement have been double checked with Law Enforcement Agencies and Banking Association, and are consistent.

2.2.2.1 Bank “α” – Fraud Manager
This bank has about 300M on-line transactions per year (2012) and suffered of phishing attacks since 2004. Due to very high losses in 2008 (5M€), mainly on consumer customer, this bank invested in a 2-factor authentication architecture and in Fraud Management System. Losses dropped to 2M€ in 2009 and then practically zeroed in subsequent years, with spikes of 100k€ of loss in 2012. The same Fraud Management System is now adopted either for traditional banking, than for TV banking and Mobile banking.

2.2.2.2 Bank “β” – Fraud Manager
This bank has about 200M on-line transactions per year (2012) and started to suffer of attacks in 2009, when they lost 3,5M€ (prevalently on consumer), even if 2-factor authentication was in place. Fraudster focused this bank with a crimeware attack based upon Zeus, circumventing the strong authentication mechanisms. These losses brought the management to build new anti-fraud capabilities, introducing a Fraud Management System specific for the on-line environment. This helped to lower the amount of losses to a value between 200k€ and 500k€ per year.

2.2.2.3 Bank “π” – Fraud Manager
This bank has about 50M on-line transactions per year (2012) and their on-line presence was not strategic until 2010. This is why they were not targeted by fraudster campaign until 2011, when their total loss by crimeware was 1,7M€. They introduced 2-factor authentication in 2010 and are now introducing an On-Line Fraud Management System, shared with the Card Processing System (they had a loss of >15M€ on Credit Card scams in 2012). This bank had more problems with corporate clients than with consumer.

2.2.2.4 Credit Card Circuit “Ω” - Fraud Manager
This information is not related to the overall objective of this paper, but these data are useful to compare a different payment method to on-line banking. The overall fraud losses by this credit card circuit in Europe in 2012 is 6 basis point (0,06%) of the total number of transactions. Of these, about 30% are on-line, and the losses with these are at 13 basis point. Considering that the medium credit card transaction value is about 1/6 of the medium on-line banking transaction value, we can deduce that total losses in terms of value are comparable.

2.2.2.5 Italian Law Enforcement Agency – Polizia di Stato
Polizia Postale, the department of the Polizia di Stato that handles on-line frauds, during the interview declared these numbers (cumulating on-line banking frauds and credit card frauds):

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2011</th>
<th>2010</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal reports</td>
<td>65000</td>
<td>46000</td>
<td>21000</td>
<td>5000</td>
</tr>
<tr>
<td>Reported loss</td>
<td>17.1M€</td>
<td>15.3M€</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*Table 3 - Polizia Postale - On-Line Fraud data - 2009-2012*
2.3 THE REACTION: HOW BANKS AND LAW ENFORCEMENT WORK TODAY?

2.3.1 Banks: Fraud-related organization and anti-fraud procedures

In the banks interviewed the typical organization that deals with frauds is structured as in the diagram below:

![Diagram of a typical cross-functional fraud organization in a bank]

Each department has one specific responsibility:

- The Fraud Operations office has to define the anti-fraud rules, govern the anti-fraud techniques and spread the approach all over the company. They deal with suspect cases and investigate internally upon the data they receive from IT department. Usually the fraud operations is split in different offices, each focused on a functional type of fraud (On-line frauds, ATM frauds, Identity Fraud, ...); the IT-related group usually is in charge of the communication with IT and with LEAs in the first steps of the discovery of a fraud. This office coordinates the actions for the first 2 weeks of the fraud lifecycle (account blocking, transaction blocking, early return from fraudster’s bank, ...). When money has been moved from the fraudster’s bank, the probability to recover the theft drops and the responsibility passes to legal department.

- The Customer Care office has a fraud competence centre that deals with the communication with the customer, either directly than through the branches employee/directors.

- The Fraud Legal Office, in case that a formal report must be followed by and a magistrate incriminates a fraudster, will take the lead. The timeframe expands to months/years, following the trial timeframe and the slow or inexistent communication with third parties, like LEAs and Banks in “non-mature” countries.

What has been discovered during the interviews is that, at least in Italy, the real loss the bank sustains is less than 10% of the total loss: customers defrauded through phishing or crimeware usually do not have a refund.

This area could be improved by adopting an insurance model at the level of the banking system.
2.3.2 Law Enforcement Agency: Organization and Fraud Investigation Procedures

Law Enforcement Agencies usually split the competencies between the local agents, in charge to receive the reports from citizens/companies and to investigate on the case, and the central Fraud Management administrators, in charge to define the methodologies, spread the investigation framework and maintain the relationship with the formal stakeholders (Banks HQs, banking associations, international bodies, ...).

In order to support different stakeholders involved in the fraud, Law Enforcement Agencies usually have a twofold approach to the management of On-Line Frauds:

- Maintain an official record for every fraud that has been officially reported by the victims;
- Try to centralize and spread all-over the banking network the fraudulent account numbers and credit card numbers (usually not within a formal framework agreement).

As an example, as of today, the Italian Law Enforcement Agency named “Polizia di Stato, dipartimento di Polizia Postale e delle Comunicazioni” (Postal Police), in order to prevent and prosecute on-line banking frauds, follows this specific investigation procedure:

a. Receive of the report from the bank of a suspect transaction, including the destination IBAN, to the LEA’s central service (bigger banks) or to local LEA compartment (Small & medium banks). If signalled centrally, the Law enforcement agency, Central Postal Service, coordinates and delegates the correct compartment to manage the investigation and to act as the local interface for the victim and the associated bank branch. If the fraud is international, it’s managed centrally.

b. Identification by LEA of the data of the destination account, by direct interrogation of the bank (if in Italy) or through Interpol (trans-national); correlation on different DBs in order to understand if the fraudster and the fraud methodology are known.

c. If the reporting has been transformed in a formal “denuncia-querela” the relevant Magistrate can issue an international request (“rogatoria internazionale”).

d. With these data, Law enforcement agency blocks the transaction on the destination account and starts to investigate on the fraud: methodology, correlation to other similar frauds (sender, destination, methodology, transactions value, geography, timing, ...)

e. The result of the transaction block and of the investigation usually is positive, and investigation is considered closed when the Magistrate incriminates the fraudster (“rinvio a giudizio”).

To fulfil these tasks, different Roles and Responsibilities are identified:

a. Victim (Consumer or Corporate)

b. Bank Department in charge to interface with LEA and Victim

c. LEA’s Administrators/Central Department

d. LEA’s Investigators/Compartment

e. Magistrate

f. (in case fraud is international) External stakeholders needed to reach the fraudster banks in case these are located outside the nation (i.e. Interpol, Europol, Internet providers, ...)

18
Law enforcement investigators act usually as a pivot between banks and helps them to recover the money that has left the bank. In order to start the process is needed a formal report by the final customer that will enable the LEA to force the block of the final account and/or the return of the stolen money.

In order to be more effective reporting to the LEA is crucial because the sooner the report (formal or informal) takes place, the sooner the fraudulent transaction can be blocked, especially for inter-banking or international transactions.
2.3.3 The actual interaction between Banks and Law Enforcement Agencies

As described earlier, the banking system usually has a relation to LEAs through an informal channel: on one side there’s the Fraud Management Department; On the other side many LEAs all over Europe are specializing specific departments on On-Line Frauds prosecution, so the natural communication flows to this “On-Line Frauds” Department.

Two structured examples of correct communication are present in UK and in US that for different reasons structured a central database for reporting financial crimes. Only UK is reflecting this data to the financial community, while US is more focused on anti-money-laundering activities.

The communication, out of these 2 specific examples, typically is not strictly formalized and uses media such as telephone calls and emails to accomplish the task of seeking help from LEAs to prosecute crime.

Banks usually do not formalize a report to the Police, in order to avoid bad publicity.

When a formal report from the final victim (customer) is filed (not always in time to recover the amount stolen) then the bank law department takes the lead with the local investigation team, in order to follow-up the theft. At this point the Bank’s Fraud Management team usually is disengaged, leaving the room to customer care team and the law enforcement team.

This relationship has many areas of improvement, either between the single bank and the LEA, than at level of the whole banking system. These improvement could be obtained with an Information Sharing System that will be described in the next chapter.
3 THE ONLINE FRAUD CENTRE

Good results obtained by banks and LEAs in contrasting on-line frauds can be confirmed and maintained with less resources with the adoption of a cross-communication framework in whole banking system. This solution can improve:

- The communication between the single bank to the LEA, in order to design and build a formal reporting mechanism that is legally safe and effective in terms of fraud blocking/amount retrieval.
- The proactive approach of the whole banking system to on-line fraud, through an alerting list of potential fraudster shared under the control of the LEA.
- Shortening the timeframe of formal communication within the banking system (through LEA) so that more fraudulent transaction can be blocked before the commitment of the money transfer.

Target of the service is to shorten the time needed to communicate from Banking Fraud Managers to LEAs, and to share the relevant information on fraudster accounts in the whole banking industry.

The first target will permit to block in shorter time a larger amount of fraudulent transactions.

The second target will help all the banks to update their Fraud Management Systems’ “watching lists” and “black lists” to better cope with fraudster attempts and block them before the money is transferred.

For obtaining these results, the InfoSharing Service, designed to be coherent with existing procedures, has the following requirements:

- A clear organizational framework and related roles/responsibilities
- A clear legal context in which it operates
- Defined procedures to be followed by the service actors
- A technological platform designed to support this service

The goal of this chapter is to study and define a common model for sharing information on cyber fraud between law enforcement and financial institutions.
3.1 Operational Framework Definition

The roles this information sharing service considers are 3: law enforcement agency, banks and banking association. Generic people/citizen could be considered as a fourth potential element that does not participate directly to the network but could be involved as final addresser of some cyber-fraud alerts. As an example a person, target of a fraud, can be warned by LEA after the related malicious activity is signalled by the service. This role will not be described.

- **Law Enforcement Agency (LEA)**, to guarantee the maintenance and the evolution of the communication model, and to prosecute eventual crimes found through the analysis and correlation of the data shared from financial institutions. Inside the LEA two different roles can be identified: **Investigators** and **Central Administrators**.

- **Banking Association (BA)**, as a representative of the banking market, in order to support their associates in representing the category needs to LEA and to use the platform to communicate with the different parties.

- **Financial institutions (Banks or credit card issuers)**, that will share their information with the LEA with the guarantee of accessing recent fraud data without disclosing specific bank attacks (in order to avoid cross-communication between banks of Data Protection Act protected data or commercial relevant data).

The Law Enforcement Agency is the entity that will host the centre, in order to define the common language and maintain the procedures and the platform used to communicate.

Other supporting competencies that can be useful to further specify, build and let the information sharing system and start the operations are:

- **Experts**, chosen between banking sector security officers, fraud officers (in banks and telcos), security professionals with strong background in ICT Security, fraud management and investigation. These will help LEAs and Banks to define
  - the final framework,
  - the data flows and data specification
  - the platform requirements
  - the application and related programming

The dimensioning of the service can span from a group of 15-20 banks in little countries with limited evolution of the financial market, to many hundreds of counterparts, for example in more financially evolved countries (In Italy more than 1000 banks are registered).

Furthermore, the dimensioning could consider the potentiality to interchange information between similar system installed in different countries or in different LEAs boundaries (i.e. Telco Frauds managed by economical police or Banking Frauds managed by criminal police, that in some country are split – see for these chapter 4).

LEA, Banks and BA have a quite different view of the service and, subsequently, on the supporting platform.

**Law Enforcement Agency** will use the platform to collect and retrieve information on criminal activities. It will be capable to insert new data in the platform as the banks but this will be not the primary interest. LEA will perform data retrieval through data correlation. The platform will allow LEA operators to describe customized research on information stored. Moreover, LEA operators can communicate with
banks and BA through the InfoSharing Platform infrastructure sending information about data previously inserted in the platform. In LEA departments there will be different kind of InfoSharing Platform operators depending on task permitted or assigned: these can range from the simple utilization to administration activities. Some functionalities of the platform will be accessible only through specific physical machines and/or through the utilization of one or more authentication factors. LEA will dedicate some specific and physically protected workstations to communicate with the InfoSharing Platform as well will allocate and train operators to use the technology.

**Banks** will use the platform as an information container and they will primarily push data in it. The information unit that banks send to the platform can be related to a specific money transaction episode that its other internal systems have signalled as a fraud attempt or a fraud related activity. Banks could also have the possibility to see or modify information previously stored in the database but this will be done through secure methods and never letting the protected information to be shown outside as plaintext. The platform should be accessed only through specific physical machines and through the utilization of one or more strong authentication factors. The bank should dedicate at least a specific workstation to communicate with the LEA platform as well as allocating and training operators to use the technology.

**The Banking Association** will use the platform both to push statistical data that banks do not have and to retrieve some information. Particularly BA can use the InfoSharing Platform to dispatch information among banks. BA will have the possibility to see statistical data related to banks participating to the association but cannot obtain details on specific transactions or malicious episode related to them. It would be possible to do some kind of soft correlation gaining statistical trends on banks behaviour and criminal activities. As in the other cases some functionalities of the platform will be accessible only through specific physical machines and/or through the utilization of one strong authentication factor. BA should dedicate some specific and physically protected workstations to communicate with the InfoSharing Platform as well as allocating and training one or more operators to use the technology.

High level functionalities available for banks’, Banking Association’s and Law enforcement agency’s interfaces, depending on the set of the accessible InfoSharing Platform services, is presented in the following table:

*Table 4 - Provided Functionalities for each user interface*
3.2 **Legal Constraints**

Two main legal constraints can be considered in the design of the InfoSharing Service:

- The legal basis on which a LEAs must prosecute fraud
- The Data Protection law that must be considered for the users of the banking system

LEAs of EU member states must prosecute frauds as defined in (18) Treaty, specifically in Art.85, Art 310.6, and Art.325. Member states then have to implement such indications in their respective local rules. As of today, all member states formally comply.

EU legislation for protecting personal data is based upon Directive 95/46/EC (Data Protection Directive) and, even if with slight differences, every member state has adopted the DPD concept. These differences will be superseded in 2014 with the adoption of 2012 proposal of the General Data Protection Regulation (GDPR) that will be applied to all member states for the end of 2016. This Regulation will consider different aspects that today EU and local regulation do not cover like: globalization; technological developments (social networks and cloud computing); EU enlargement; new EU cyber-threats for citizen and society.

In the meanwhile, great attention must be adopted in the treatment of personal data of customers of the bank, especially in the pre-fraud phases. When a potential fraud has been committed, the Fraud Office of the Bank has all the legal rights to manage customer data, as stated in the treatment instructions published by the Bank.

The legal problem that has to be managed in the proposed InfoSharing Service is twofold:

- The Privacy Directive adopted in any bank states that the communication to LEA and subsequent treatment of personal data in case of suspected fraud must be declared in the Persona Data Treatment declaration of the Bank;
- This data cannot be communicated from the LEA to other banks either for lawful motivation (DPD) than for commercial reasons (banks do not want to share with competitors data about their fraud)

What happens next depends on the fact that the suspected fraud has been confirmed or not.

If the suspected fraud has been confirmed, data communicated can be treated as criminal data, so (apart from the rules applicable to criminal data, used to combat and prosecute crime) there are no concern about privacy.

If the suspected fraud is not confirmed, data owned by the suspect victim and the suspect fraudster should be deleted or treated specifically.

The suggestion, in order to avoid any legal risk, is to transmit to the LEA the data of the suspect transaction, so they can verify fraud attempts, covering this treatment in the Personal Data Treatment Declaration of the Bank.

LEA then can share only the beneficiary data with (potential fraudster) with all the banks in order to confirm or not its involvement in one or more fraudulent transaction. Potential victim data must not be shared in any case.
3.3 DEFINITION OF THE CENTRE OPERATIONS

3.3.1 Bank
As introduced in the previous paragraph each bank that accesses the InfoSharing Platform has to allocate at least a terminal for that specific purpose. This workstation has to be isolated from other bank’s network infrastructures and it should be physically protected. Only predetermined operators will have access to the terminal and, consequently, to services provide by the InfoSharing Platform. At least 3 people with the same tasks and authentication levels should be identified in each bank. Accessing those services will always require the operator to strong-authenticate himself/herself to the platform. Security tokens, smart cards or software certificates can be used for this purpose.

The access to the InfoSharing Platform should be a consequence of the fact that the internal fraud detection service of the bank has signalled some episodes of potential phishing/crimeware fraud. An internal report will be given to the bank operator and he/she will insert that information on the InfoSharing Platform. Once authenticated a web application will provide a simple interface for data insertion. This interface will be composed by a list of forms labelled with information on what kind of data has to be inserted. The compilation of some part of the data sharing will be mandatory since the presence of unusable data in the database must be avoided. Once the data input is completed, the operator will send the information to the InfoSharing Platform. In the meanwhile he/she could download the updates of the list of fraudulent transaction accounts (such as IBAN, Credit Card numbers, on-line betting account numbers, …) shared by LEA.

A unique identifier linked to an episode previously stored in the platform will be sent to the operator after the insertion. This can be used to retrieve data from the InfoSharing Platform later.

From time to time, banks could receive some alerts from the platform. These alerts could be notifications on specific cyber-fraud, on elements (IBAN, host, email, etc.) involved on malicious activity related to that or other banks. Moreover the platform will provide some statistic of that bank’s behaviour.

3.3.2 Banking Association
The Banking Association (BA) behaves like any other bank in the active role of pushing data in the platform but it has many less available services for downloading or investigating data present on the platform.

BA owns specific data (different from the other actors) that wants to share with the platform and, for this reason, it will be able to access the platform in the same way banks’ operators do. However, BA and Bank interfaces have two important differences: one is related to communication processes and the other linked to visible statistics and data analysis. Communication services are a fundamental part of the InfoSharing Platform since all the information related to correlation and suspected IP, URL, IBAN, etc. have to be shared among the actors. Banks cannot be able to communicate directly with each other. This is done to avoid false information spreads among the actors without LEA can intervene. However the BA could be capable, in some specific cases, to make a role of intermediary similarly to LEA. For this reason, BA interface will be customized, allowing a simple and direct way to send messages and data to Banks. For what concerns statistics and data analysis it can be useful for the BA to have a view a little more detailed of what the service has discovered on fraud episodes. As it will be explained in the Interfaces section, banks will see platform statistics without any further detail while LEA will always be capable to analyse correlations in deep and to know which phishing episodes or elements are involved. BA will see
just few details of each LEA correlation avoiding details that can be useful only for investigative purposes. BA will not be able to define or launch correlations. LEA operator will be the only actor able to create correlations.

3.3.3 Law enforcement agency – Investigation Departments/Compartments
As already introduced, LEA use case has different operational aspects. The main work LEA Operators will accomplish is to prosecute crime reported by banks through investigation. LEA detectives will be allocated and trained to interact with the platform in response to inputs by the banks, in order to query the database and select interesting information. This operation could be performed by generic LEA workstations, provided the detective has authenticated to LEA infrastructures. The system will provide a specific interface to make data analysis, obtain reports and data interpretations. Once authenticated, a simple interface of querying will allow LEA Operators to get data from the platform and/or make correlations among them. No other services are needed in this case.

3.3.4 Law enforcement agency – Central Fraud Service & Administration
The maintenance of InfoSharing Platform application, hardware and software is demanded to specific LEA Administrator. These users will have physical access to the platform’s machines and could have authentication credential for the software. Times and modes of action should be established depending on the infrastructure’s complexity and on the services availability needs.

These administrators will create, modify and disable users of the InfoSharing Platform and will provide the technical knowledge and operation for maintaining the platform updated (patches, virus updates, ...), to define the correlation rules templates, and to compose overall KPIs on the platform (i.e. total number of frauds, total value of frauds, ...). For these special users a specific interface for administering the whole InfoSharing Platform will be provided.

The InfoSharing Platform LEA Administrator uses a dedicated machine called InfoSharing Platform Control Terminal. This workstation should be isolated from the rest of the LEA internal network infrastructure and must be physically protected. Only a predetermined LEA’s administrator will have access to the terminal and, consequently, manage services provided by the InfoSharing Platform. The InfoSharing Platform LEA Administrator will always be asked to authenticate himself/herself to the platform in order to access the management panel. This authentication must be stronger than those related to bank accesses. Once authenticated an application or a web application will provide a simple interface for controlling InfoSharing Platform’s functioning. Administrator will see general information on platform’s usage and detailed information on data stored in the database. For every new data unit sent by banks the terminal will display an alert describing data inserted in the platform. An important task concerns communication. The dedicated terminal could be used also to query the platform for retrieve any kind of information.
3.4 Platform Description

The main achievement of the proposed project is the realization of an ICT infrastructure supporting information sharing among banks and Law enforcement agencies. The development of the InfoSharing Platform includes hardware and software solutions for communication, access and storage of confidential information related to malicious money transactions.

The proposed platform is an advanced system for data storage and analysis and should be a useful instrument in supporting cyber-fraud investigations. The target is to provide a secure communication channel encouraging banks to give as many information as possible to law enforcement departments in order to realize data mining with the purpose of providing new investigative insights.

There are two important guidelines that this InfoSharing Service is following:

- **Full Security**: intended as providing full protection on information and identities of involved actors. Banks as well as Law enforcement agency departments will access platform services in a secure way and will use them without any risks on databases’ protection.
- **Full Usability**: intended as providing easy-to-use software and full availability of services for both banks’ and Law enforcement agency’s operators.

This chapter contains the technical analysis needed to support the development of the InfoSharing Platform.

**NOTE**: variable names and screenshots contained in chapter 3.4 are derived from the EU project OF2CEN, and have been maintained in Italian. Credits to Marco Caselli (GCSEC) for the graphic schemas of the proposed platform, based upon the definition of the InfoSharing Service.

3.4.1 Information and data required

Law enforcement agencies, banking association and banks must agree on a specific formalization of data sent to the platform.

Information are managed through four different signalling categories that are:

- **Mirror site signalling**: this identifies a signalling of a phishing web site.
- **Suspect IP signalling**: this identifies a signalling of a suspect IP used to make a malicious money transaction or to host a phishing web site.
- **Fraudulent transaction signalling**: this identifies a signalling of phishing episode intercepted by a bank.
- **Generic signalling**: this identifies a wide set of signalling concerning suspected elements like a specific media (IBAN, Credit Card, Mobile number account, ...).

Banks and Banking association (as well as LEA) can use one of these four data-forms to insert new information on the database. It is appropriate to store the raw information in a separate section of the database (the so-called raw tables). This can be useful to debug or in situation in which it is needed to change the database’s structure.

3.4.1.1 Mirror Site Signalling

The form **mirror site signalling** describes a phishing web site.

The following attributes should be compiled:
• **Signalling owner ID**: identifier of who make the signalling.
• **Date of detection**: date in which the mirror site is found.
• **URL**: the URL of the mirror site.
• **IP address**: the IP address of the machine that hosts the mirror web site.
• **IP nationality**: the location of the machine that hosts the mirror web site.
• **Provider IP**: Provider for the machine that hosts the mirror web site.
• **Provider email for abuse signalling**: email owned by the Provider where to forward information and requests.
• **Registrar email for abuse signalling**: email owned by the registrar where to forward information and requests.
• **Active URL**: indicate if the URL of the mirror site is active or not.
• **Stolen credentials forwarding mode**: attack’s typologies for circumvent the victim.
• **Receiving email of stolen credentials**: email address where the attacker receives the stolen credentials.

Optional fields that could be compiled are:

• **Original URL**: the URL of the real web site copied by the attacker.
• **Original IP address**: the IP address of the machine that hosts the real web site copied by the attacker.

### 3.4.1.2 Suspect IP signalling

The form for **suspect IP signalling** describes a machine that has performed malicious activities (e.g. hosting a phishing web site, performing malicious money transaction).

The following attributes should be compiled:

• **Signalling owner ID**: identifier of who make the signalling.
• **IP address**: IP address of the suspected machine.
• **Date of use**: date in which the machine was used.
• **Time of use**: time in which the machine was used.
• **IP nationality**: the location of the machine.
• **IP Provider**: Provider for the machine.
• **Activity typology**: description of the malicious activity.
3.4.1.3 Fraudulent transaction signalling

The form for fraudulent transaction signalling describes a malicious or suspected malicious money transaction.

The following attributes should be compiled:

- **Signalling owner ID**: identifier of who make the signalling.
- **Date of the operation**: date in which the money transaction was performed.
- **Scammed customer’s name**: name of the victim.
- **Relation source**: source where the money was maintained.
- **Money Amount**: amount of money transferred.
- **Transfer mode**: type of money transfer.
- **Beneficiary name**: name of the money recipient.
- **Relation destination**: destination where the money was sent.
- **Destination institute**: institute (e.g. bank) responsible for the Relation Destination.
- **Formally reported**: it indicates if the transaction has been denounced or not.
- **Reporting Compartment**: Law enforcement agency office responsible for the denunciation.
- **Recovered money**: amount of money recovered by the Law enforcement agency action.
- **Reversed money**: amount of money relocated by the bank to the account of the defrauded.
- **Blocked import**: amount of money in the transaction blocked by a magistrate
- **IP related to the fraudulent transaction**: IP of the machine from which the transaction was performed.
- **Criminal procedure number**: id, assigned by LEA, related to the criminal procedure.
- **Attorney**: office delegated to follow the case.
- **Notes**: any further information related to the fraudulent transaction.

Optional fields that could be adopted are:

- **Blocked**: it indicates if the transaction has been blocked, if the fact is uncertain, or if it passed.

3.4.1.4 Generic Signalling

The form generic signalling describes information that does not concern any of the previous cases. It can be used to signalling a wide set of information since it is possible to describe it through its general attributes.

The following attributes should be compiled:

- **Signalling owner ID**: identifier of who make the signalling.
- **Signalling type**: it describes what kind of information the signalling is carrying.
- **Signalling content**: information content of the signalling.

3.4.2 Information sharing Interfaces

As said in the introduction interfaces have to be simple and easy-to-use.
Each actor using the InfoSharing Platform will have a different interface customized for specific services and activities. Interfaces can be described in two classes with several sub-types:

- **External Interface**
  It represents the set of web pages used by InfoSharing Platform Bank Operators to inserts data in the database and to access other services. This interface could be further divided in:

  - **Bank Interface**: interface customized for banks
  - **BA Interface**: interface customized for the Banking Association.

- **Internal Interface**
  It represents the set of web pages and/or applications used by LEA to inserts data in the database and to access or manage other services. It could be further categorized in:

  - **LEA Administrator Interface**: this is the application used by the Super User Operator to access directly the InfoSharing Platform database. This access allows him/her to perform any operations. Deletion is allowed in the main DB, not in the RAW DB.
  - **LEA Auditor Interface**: this is the application used by the Audit Operator to access InfoSharing Platform’s log files. This access allows him/her to control communication flows and the platform’s functioning. No deletion is allowed.
  - **LEA Operator Interface**: this is the interface used by the LEA Operator (Detective/Agent). It allows the operator to perform queries on the database or insert new information. No deletion is allowed.
3.4.2.1 **External Interfaces**

The InfoSharing Platform will provide a secure login page where operators will be asked to enter their credentials. Once logged in a homepage the interface will provide a quick overview of all the services provided by the platform.

![Figure 7 - External Interface Home Page](image)

The page is divided in two main parts. The first on the left is an always shown site-map used to facilitate the navigation on the web application. The central section consists of a series of windows that show news, signalling, notifications and statistics. The first window is highlighted LEA-to-Bank communications are the most significant. The two following elements collect LEA news directed to all Banks and news sent by the Banks themselves. The last box links to graphs and maps showing statistical information. This information do not give any specific description of data involved.

The red buttons on the left leads the operator to the signalling section.

![Figure 8 - External Interface - Level 2](image)
When a bank has to insert new data in the database the web application will present to the operator a set of possible signalling forms. Each of them is related to a different kind of information that it is possible to add and consequently leads to a specific input page.

**OF2CEN Signaling**

![Figure 9 - External Interface - Level 3](image)

As described before four possible type of signalling are allowed:

- **Mirror site signalling**
- **Suspect IP signalling**
- **Fraudulent transaction signalling**
- **Generic signalling**

The specific signalling page shows a list of previously inserted elements at the centre.

The four buttons at the bottom allow the user to add a new element, modify an old one, search information or visualize statistics.

**OF2CEN Signaling**

![Figure 10 - External Caption - Level 4](image)
Addition and modification are implemented with similar web pages. Data could be bleared depending on the importance of the related information.

Finally, this is a hypothesis on the utilization flow of the external interface. A possible sequence of operation in the platform is schematized as follows:

![Diagram](image)

*Figure 11 - External Interface utilization schema*

### 3.4.2.2 Internal Interface

In this section is outlined how the internal interface should be presented to LEA investigators and administrators.

The interface presents all the functionalities described in the previous paragraph, plus specific functionalities needed by LEA users.

The InfoSharing Platform will provide a secure login page where users and operators will be asked to enter their credentials. Once logged in the homepage the interface will present a quick overview of all the services provided by the platform.

![Diagram](image)

*Figure 12 - Internal Interface - Level 1 – Homepage*
The page is divided in two main parts. The first on the left is an always shown site-map used to facilitate the navigation of the web application. The central section consists of a series of windows:

- **Tracker Database**: that leads to a page where the information on database utilization is stored.
- **Automatic Correlation**: that leads to a page that presents correlations’ results. The page allows also managing present correlations and adding new ones.
- **Statistical Trends**: that leads to information related to statistics performed on stored data. This is quite similar to the bank’s operators one but it provides more information and details of the elements that gave rise to those statistics.
- **News**: that allows Law enforcement agency users and operators to take a look or create news and reports for banks and for internal use.

The Tracker Database web page has a central part collecting the last elements inserted in the platform. Each element can be selected to visualize its details. The TakeInCharge button allows LEA Operators to signal the start of the Investigation phase. Each LEA operator is free to modify a record and make some researches or statistics. It is worth noticing that, despite banks, LEA can always see all the fields for each record presented.
The layout of Correlation page is quite similar to the one proposed in the previous paragraph. In the central part the page visualizes all the correlations previously inserted in the platform. On the right, there is a coloured flag indicating importance of possible results. It will be possible for a user or an operator to see the detail of each correlation and the elements involved.

LEA users and operators can insert new correlations in the platform thanks to the page “New Correlation rule”. Depending on the complexity of such correlation it will be possible to express it through a pre-established form or directly with an SQL query.
The platform presents the page Correlation Detail both when users click on a specific correlation and press the “Notify” button. This page shows information related to the correlation and details about elements involved.
3.4.3 Analysis and correlations to be performed (examples)

One of the main tasks of the InfoSharing Platform is to provide to LEA officers investigative insights. The InfoSharing Platform core will have to “reason” on data stored in the database and propose hypotheses on grouping together simple phishing episodes as well as criminal activities.

Therefore, the goal of a correlation process is to bring together episodes and/or activities that share something in common. Correlate other elements can be valuable, but episodes and activities are the most important and meaningful objects of the database to be correlated. From a high level point of view correlation processes schematization is as follows:

- **Common Attribute**: this kind of correlation focuses on episodes and activities that share a specific attribute. This attribute belongs to the episodes and activities we are correlating. As example, all the episodes that have the same Owner or Medium as well as search for common Transfer Mode or Host can be linked. This way of correlation can be useful also for other database entities like IBAN, Site, Email, geo-localization, etc.

- **Related Common Attribute**: this correlation concentrates again on common attributes but allows to bring together elements that do not own them. The idea is that some episodes cannot have any attribute in common but different attributes linked to them can have a connection in another part of the database. As an example can be considered two phishing episodes made from different hosts. If these hosts have the same ISP they can be correlated as “near”. This is different from a Common Attribute Correlation as looking only to table Episode cannot be found anything to link the two phishing occurrences.

- **Common Dynamic**: this way of correlation puts aside specific attributes and focuses on similar trends. Differently from previous cases, episodes and activities are correlated even if they do not have any common element within the database. As example ten phishing episodes recorded at intervals of five seconds from each other can be considered. Finding similar sets among data at disposal can lead to a correlation of these elements as “two phishing activities that probably share the same way of action and possibly the same people or criminal organizations”.

- **Related Common Dynamic**: as in the previous case this correlation concentrates on common trends. However those trends are found in different parts of the database. As example phishing episodes related to different hosts that use similar URLs (like xxx.fakesite.com, yyy.fakesite.com, zzz.fakesite.com and xxx.badplace.org, yyy.badplace.org and zzz.badplace.org) will show similar way of action (URLs have been created in the same way).

It is important to remember that not all the automatic correlations represent really connected episodes and activities but it’s important to propose LEA Operators some insights. The InfoSharing Platform will propose Operators to take a look on specific episodes and activities together and confirm or not what the correlation has evidenced.
3.4.4 Technical features of the Centre

3.4.4.1 Infrastructure

The InfoSharing Platform infrastructure has to be compliant with the following characteristics or basic properties:

- **Security** – intended as protection and privacy of information.
- **Integrity** – intended as assurance on the integrity of the information.
- **Coherency** – intended as certainty of maintaining data stored in the database coherent with real data and with each other.
- **Resiliency** – intended as ability to recover from a problem to a safe state.
- **Availability** – intended as continuity of services.
- **Performance/Service Level Agreement (SLA)** – intended as maintenance of an adequate level of services in response to the needs.
- **Scalability** – intended as capable to manage, in time, an increase in traffic.
- **Usability** – intended as simple interfaces for an easy to use.

Basic Elements

Before describing in deep some infrastructure’s hypotheses it is important to identify some basic elements that could be needed or simply useful to meet the requirements listed in the previous paragraph above.

Among them there are:

- **Demilitarized Zone (DMZ)** – that, in this case, is a physical sub-network containing and exposing some web application services to Internet. This can be used to avoid banks accessing directly the InfoSharing Platform core platform.
- **Authentication** – that is the process needed to access some services with the aim of confirming a digital identity or some software/hardware credentials.
- **Access Control** – that is the process of monitoring accesses to the platform in order to be always aware of who is using the services.
- **Traffic Control** – similar to the previous one, it is focalized on controlling traffic and communication flows in order to be always aware of how the platform is used.
- **Logging** – this is the process of recording all the information related to InfoSharing Platform activities. For example those concerning accesses and traffic control.
- **Virtualization** – that represents the ability to replace physical computer to virtual machines in order to cut costs but also to make infrastructure’s management easier.
- **Redundancy** – that represents the duplication of machines and/or information to ensure services and data security.
- **Caching** – that is the possibility to store some information in faster memory to speed up the process of access to them.

Not all the previous are mandatory. For example, InfoSharing Platform infrastructure must provide authentication and data redundancy. It should have a DMZ and implement controlling processes. Virtualization and Caching can be avoided but they could be useful for speed up maintenance and functioning of the whole system.
Architecture

Overview
The whole InfoSharing Platform infrastructure could be logically divided in four main parts:

- **External Components** – that includes banks (and therefore their connection terminals) and the interconnections to the access point of the platform.
- **DMZ** – that includes access firewalls and authentication services needed to use the InfoSharing Platform.
- **Internal network** – that is the Law enforcement agency network infrastructure and includes all Law enforcement agency workstations.
- **InfoSharing Core Platform** – that includes the database and machines that directly access to it.

![Architecture Overview](image)

The shield in the previous figure indicates that LEA will physically protect those specific infrastructures avoiding any un-accredited access.

**External components**
Banks has to organize a small hardware infrastructure to access the InfoSharing Platform. Basically a simple pc can be sufficient to manage InfoSharing Platform services. The workstation has to access the Internet with a static IP address. This could be useful for the authentication phase.

The operator’s machine could be protected behind a firewall. Beyond the obvious security provided by a firewall this infrastructure could be useful to create a Firewall-to-Firewall VPN instead of a Host-to-Firewall one.

**DMZ**
The demilitarized zone is the access point to the InfoSharing Platform infrastructure. It consists of several firewalls and one or more servers. Each connection toward the platform will pass through server in the
DMZ. Banks accessing the platform will use the DMZ servers to authenticate and to establish a secure point-to-point connection.

The schema used to implement a DMZ could be Single Firewall or Dual Firewall. In the first case it uses a firewall with three network interfaces that has to deal with all the traffic. In the second one the first firewall ("front-end" firewall) must be configured to allow traffic destined to the DMZ only while the second firewall ("back-end" firewall) allows only traffic from the DMZ to the internal network.

![Figure 18 - Single Firewall / Dual Firewall DMZ](image)

Inside the DMZ there could be one or more servers. The server has to provide a web application front-end interface to banks to allow authentication and it communicates with the InfoSharing Platform core in order to verify credentials. When an access to the platform is granted this server can continue to operate as proxy server to avoid direct connection between banks’ PCs and the InfoSharing Platform core. This will take place within a strong authenticated VPN.

**InfoSharing Platform Core**

The InfoSharing Platform Core is the main part of the InfoSharing Platform infrastructure. It consists of several machines, one or more databases, a terminal and any hardware needed to connect them. Servers of the InfoSharing Platform Core provide a back-end interface to the database for machines in the DMZ, for LEA workstations and for the core terminal. InfoSharing Platform Core databases will store data sent by banks, BA and LEA departments as well as log information. Finally, the terminal will be used by the LEA Administrator for any management operation that requires direct access to back-end servers and databases.

The core infrastructure could be implemented in several different ways depending on levels of availability and recovery we want to provide. It is worth noticing that machines could be either physical or virtual. In the next section some possible hypotheses are described.
In this schema all the previous actors access services through a single machine. This is connected to the database through a dedicated line and information between the two servers flow in both the directions. The database is directly connected to a backup database that is inaccessible from any other machine. This will intervene actively only in case of disaster recovery.

Figure 19 - Architecture Hypothesis 1

In this schema all the previous actors can access services through two different machines with the same characteristics. Both the two servers are connected to the database through a dedicated line and information between the two couple of machine flow in both the directions. The database is directly connected to a backup database that is inaccessible from any other machine. This will intervene actively only in case of disaster recovery.

Figure 20 - Architecture Hypothesis 2
Different Interfaces / Simple Backup

This schema is similar to the previous one except for an important characteristic. The two servers that implement the interface to the database are used differently to insert or retrieve data in/from the database. Banks will be connected only to the push server while both Law enforcement agency departments and the core terminal could access both of them. Both the two servers are connected to the database through a dedicated line and information between the two couple of machine flow in both the directions. The database is directly connected to a backup database that is inaccessible from any other machine. This will intervenes actively only in case of disaster recovery.

Different Interfaces / Double Database

In this schema there are always two different interfaces to the database but since one of them is used just to retrieve information it could be connected to the backup server. With this schema the system can offer better performance. Therefore the two access servers are connected to different databases. Between the Push machine and the master database information flow only downwards while between the Pop machine and the slave database they flow only upwards. The master database is directly connected to the slave one. This last will always intervenes in case of disaster recovery.
Different Interfaces / Double Database / Simple Backup

This schema is similar to the previous one. The two interfaces to the database are always connected to different database. The first database stores only data sent by banks while the second stores both that data and other information (for example logs and correlation schemas). The Push interface will manage the two connections depending on who is using the platform (Banks or Law enforcement agency departments respectively). Links between the Push interface and the databases are possible only for incoming while the connection between the Pop interface and the database is the reverse. Since it is possible to intervene in both the databases a new backup database is needed (in the middle) that will manage the synchronization among the previous two and will intervene also in case of disaster recovery.

More hypotheses are possible and this is just a view about design problems and solutions. LEA Administrator should choose the Architecture compliant to their Internal Policy.

Different Double Interfaces / Double Database / Simple Backup

This schema is logically equal to the previous one. Accessing interfaces are doubled to take into account the need of full availability.
High Availability
For reasons related to Disaster Recovery procedures, in case the need of recovery of the system in case of unavailability is less than 4hrs, two locations should be used, where back-end servers and databases should operate.

Hardening procedures and operation
The InfoSharing Platform infrastructure will comply with hardening instruction defined by the LEA. The system will be operated by LEA administrators, chosen by the level of trust and competence LEA defines.

3.4.4.2 InfoSharing Platform Software
The InfoSharing platform will be based on a custom InfoSharing software, specifically developed on open-source or commercial framework.

The starting point of this discussion is related to the choice between open source and proprietary software. What follows is a description of some areas where it is important to define accurately software characteristics.

Programming Languages
Programming language should be chosen based upon:

- Developers skills and capabilities
- Identification of operations and processes of the InfoSharing Platform that can discriminate between a language and another
- Infrastructure technology chosen

Any of the following development framework can be suitable for the InfoSharing Platform infrastructure.

**PHP**
PHP is a general-purpose server-side scripting language. It is widely used for developing Web platforms and it is integrated into the LAMP (software bundle for Linux, Apache HTTP Server, MySQL and PHP/Pearl/Python). Basically, any PHP code in a requested file is executed by the PHP runtime, usually to create dynamic web page content or dynamic images used on Web sites or elsewhere. It is fully supported by the community and there are several good frameworks for web platform developments such as CodeIgniter, Zend, Yii, etc.

**Java**
Java is a very famous and widely used programming language developed in 1995 that reaches now its seventh standard edition. It is a general-purpose, concurrent, class-based, object-oriented language. Java applications are typically compiled to byte code that can run on any Java Virtual Machine (JVM) regardless of computer architecture. It can be used on the Java Platform Enterprise Edition (J2EE) that provides an API and runtime environment for developing and running enterprise software, including network and web services. It is not specifically suited for a web application but for larger and more
complicated infrastructures. However, a solution that uses both PHP and Java could be considered for the InfoSharing Platform.

ASP
ASP is a Microsoft's server-side script engine for dynamically generated web pages. More interesting to notice is the possibility of using its successor ASP.NET that is a Web application framework developed and marketed by Microsoft to allow programmers to build dynamic Web sites, Web applications and Web services. ASP.NET is built on the Common Language Runtime (CLR), allowing programmers to write ASP.NET code using any supported .NET language.

C#
C# is a multi-paradigm programming language encompassing a lot of programming disciplines such as object-oriented and component-oriented ones. It was developed by Microsoft within its .NET framework. As in the case of Java this is more than needed for InfoSharing Platform development. C# can be used to integrate specific functionalities in and ASP.NET web application.

Development and testing
Development environment and testing environment must be dedicated and physically separated from production systems. Development data and testing data cannot be copied from production, and must be anonymous. The installation in production and the passage in operation will follow LEA’s rules on IT change management and IT operation management.
3.4.4.3 **Data Base Management System**

The InfoSharing Platform database is probably the most important element of the whole platform. It is used to store data sent by banks, BA and LEA departments as well as correlations coming from LEA Investigators. It is also a critical part of the infrastructure since it has to be secure, always consistent and available (must guarantee Confidentiality, Integrity, and Availability).

Database technology to use is connected to data stored inside. Some elements to be considered are: the number of relations in the database, the complexity of the objects to represent and the queries that will be developed. Many well-defined DBMS are available and each of them has specific characteristics. Possible choices considered in this paper are three:

- **Relational database**: RDBMS are the most common choice for data storage. They are based on a relational model where a relation is defined as a set of tuples that have the same attributes. Since they are widely used, the related standards are very stable and many tools are available. The language used to manage data inside an RDBMS is the SQL.

- **Object-Oriented database**: OODBMS are database management systems in which information is represented in the form of objects as used in object-oriented programming. Respect RDBMS this kind of paradigm provides a more dynamic access to data through the use of pointers. The language most commonly used to query OODBMS is the OQL.

- **Object-Relational database**: this is a database management system similar to a relational database, but with an object-oriented database model. In this kind of paradigm objects, classes and inheritance are directly supported in database schemas and in the query language.

- **Deductive database**: this is a database system that can make deductions and conclude additional facts based on rules and facts stored in the database itself. This kind of paradigm takes advantage of the logic programming integrating it with the concept of relational database. Deductive database often used Datalog to define and describe facts, rules and queries.

The InfoSharing Platform database does not contain complicated entities that justify the use of OODBMSs. Moreover dynamic access to database’s data is not strictly necessary. It might be interesting to take advantage of logic languages expressiveness, and therefore use deductive databases to make correlations on information stored. However, most of the queries that will be executed on the database
is quite simple and it can be probably expressed easily with SQL. For this reason the choice goes in the direction of Relational databases.

**DBMS products**
The landscape of the database management systems is wide and varied. Among open source solutions, the two most famous and used DBMS are MySQL and PostgreSQL. Both have advantages and disadvantages.

**MySQL**
MySQL is probably the world's most used relational database management system (it is used by Google, Wikipedia, Facebook, Twitter, etc.). Its celebrity is mainly due to its integration in the software bundle LAMP open source web application software stack. Its main feature is to use multiple storage engines (MyISAM, InnoDB, etc.) allowing to choose the one that is most effective for each table in the application. However, it is not compliant with SQL standards and it is worth noticing that it implements a semi-synchronous replication.

**PostgreSQL**
PostgreSQL (or simply Postgres) is an object-relational database management system widely used. Respect MySQL it provides more features and best performance on complex queries and objects. It is deployed with a single storage engine with good characteristics as a fully ACID compliance and the ability to manage huge amounts of data (in the order of petabytes). It is compliant with the last SQL standard (2008) and it provides an asynchronous replication.

Both MySQL and PostgreSQL are excellent products. The utilization of MySQL, due to its integration into the LAMP, has a definite advantage in the development of software for the platform. On the other side PostgreSQL is a more mature application with better performance on average. Replication mechanisms is a pretty important aspect in the InfoSharing Platform database. The speed in the synchronization between the main database and the backup one is certainly a factor to take into account when choosing one of the two DBMS.

Among proprietary DBMSs, Microsoft SQL Server can be a good solution.

**MS SQL**
MS SQL Server is a relational database server developed by Microsoft. It is distributed in different editions aimed at different audiences and for different workloads. Query languages supported are T-SQL and ANSI SQL and database operations are ACID-compliant. Even if this is proprietary software both commercial and freeware editions are available.
ER Model

InfoSharing Platform DB has to contain information on malicious money transactions focalizing especially on phishing activities. Beginning by this information may be useful for designing a coherent and targeted entity-relation model. Entity **Episode** (Episodio) describing a specific money transaction can be defined. This transaction must take place between two actors and two monetary media storage. The previous two will be two entities as well. Since the actor could be a physical person or a company a general entity **Owner** (Titolare Rapporto) of the relationship will be defined. This has to be linked to more specific elements in the database. Each actor could have one or more contacts like telephone numbers, addresses, emails, etc. This will be represented by the entity **Contact** (Contatto). For what concerns the monetary media storage the **Relationship medium** (Mezzo Di Rapporto) is defined. This entity will be better specified with the aid of other tables. Relationship medium could be linked to a specific bank or institute and therefore to an entity **Institute** (Istituto). Every institute has a security operator represented by the entity **Security operator** (Operatore Sicurezza).

Another important section of the model is related to the ICT elements involved in the money transaction and, in a more general way, to the malicious episode causes and parts. Three different ICT elements interconnected with each other are identified. These are the entities **Host** (host), **Web Site** (Sito) and **Email** (Email). The first represents a pc or any other device that has been used in an electronic money transaction or host for a malicious web site. Email is used to describe phishing episodes linked not to the web pages but to email spam. Email and Web Site are specialization of a more general entity called **Phishing Element** (ElementoDiPhishing). Both hosts and phishing elements are discovered in a specific moment (this could be also the moment in which the information is inserted in the database). This data is recorded in the table **Takeover** (Rilevamento). Finally, it is possible to record information on where the hosts are through table **Geo-location** (geo-localizzazione).
At this point most important elements connected to the entity Episode are defined. Since one of the goals of the InfoSharing Platform is making correlations among those episodes it could be useful to define an entity intended as a group of episodes sharing a particular characteristic. Grouping some episodes is a result given by both correlations made by LEA operators during an investigation. A set of linked episodes is represented in the entity Activity (Attività). Both episodes and activities are discovered in a specific moment (this could be also the moment in which the information is inserted in the database). Also in this case data is recorded in a table Takeover (Rilevamento).
Another part of the model includes elements related to reports that flow from the platform to Law enforcement agency, institutions or people. Since notifications to the Law enforcement agency are very different from the others signalling from internals and from externals will be split. Therefore we have **Internal Report** (Segnalazione Interna) and **External Report** (Segnalazione Esterna) that specify the entity **Report** (Segnalazione).

![ER Model (Part 4)](image)

The last part of the schema describes correlations’ storing and relations. The platform is capable to link phishing episodes together inside an activity. This process of grouping comes from specific correlation defined by the LEA. **Table Correlation** (Correlazione) describes this information.

![ER Model (Part 5)](image)
Finally, the database will store in a specific table the entire content of messages sent by banks (or by any other different actor of the system). This is quite important since in table Messages (Messaggi) are stored information about digital signatures or other security parameter for authentication purpose. In this way the database has a precise reference for any requests dealing with non-repudiation or similar functions (i.e. keys needed for web authentication or password hashes).

The entity-relation model presented above is designed to be scalable. The generalizations on Owner, Relationship Medium, Institute and Report give the possibility to add new actors, money storage mediums, etc. simply specializing the related entity and leaving the rest of the database unchanged.
Database tables

With the aid of an Entity-Relation model InfoSharing Platform database can be specified. With this schema it will be possible to add new attributes in the future without any particular changes in the database.

*Owner* (TitolareRapporto) has just a distinctive **ID** attribute. Every *Owner* could have an assigned contact. *Contact* (contatto) is defined through attributes *Tipo* (type of contact, i.e. Telephone, Email, Address) and *Value* (valore, i.e. 02123123, contact@totoz.com or Via dei Verri, 12 – Venice – Italy). Two examples of specifications of *Owner* are *Person* (Persona) and *Company* (Società). *Person* has attributes like tax code, name, Surname; *Company* has the name as specifier. The previous two have however an attribute **ID** that links them to their generalization *Owner* (Titolare Rapporto).

Institute provides monetary or other services. It has a unique **ID** and could have some specification such as *Bank* (banca) identified by an **ID** and a name. Every institute has a *SecurityOperator* (Operatore Sicurezza) assigned, identified by an **ID**. This person will be the InfoSharing Platform bank operator.

---

**Figure 32 - Database tables (Part 1)**

<table>
<thead>
<tr>
<th>Contatto</th>
<th>Tipo</th>
<th>Valore</th>
<th>ID (TitolareRapporto)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 33 - Database tables (Part 2)**
Relationship Medium (MezzoDiRapporto), as Owner (TitolareRapporto) and Institute (Istituto), has an ID as unique identifier. Two specifications of Relationship Medium are IBAN (representing International Bank Account Number) and Telephone card (CartaTelefonica). The first has attributes like code, opening date and closing date, while the second has, at least, an attribute number.

Account ownership (ProprietàConto) is the table that connects Relationship Medium to Owner and so the money storage medium to its owners.

![的关系][1]

Figure 34 - Database tables (Part 3)

As described above ICT elements of interest are hosts, web sites and email. The last two are specializations of a generic entity called Phishing element (Elemento di Phishing) that has an ID. Web Site and Email add to that ID URL, URLAttivo, EmailAbuseRegistrar and email address (indirizzo) respectively. Table Host is described with attributes IP Address (Indirizzo IP), Provider, Provider Country (NazionalitàProvider) and EmailAbuseProvider. There is a further table indicating where web sites are stored. Table Hosting links host IP address with site URL. Both Phishing element and Host have attribute information stored in two tables RilevamentoElPhi and RilevamentoHost. Finally, table Geolocalizzazione (geo-location) has attributes: IPAddress (Indirizzo IP), Latitude (Latitudine) and Longitude (Longitudine).

![的关系][2]

Figure 35 - Database Tables (Part 4)

Episode (Episodio) is the main element of the database. It has several attributes assigned:

- **Transfer mode** (ModalitàTrasferimento): It describes the way money has been stolen.

---

[1]: #
[2]: #
- **Amount of transfer** (ImportoBonifico): It describes how much money has been moved.
- **Date of transfer** (DataBonifico): It specifies transfer’s date.
- **Recovered amount** (ImportoRecuperato): If present, it represents the amount of money recovered by the Law enforcement agency.
- **Diverted amount** (ImportoStornato): If present, it represents a situation in which banks involved have intercepted the attempted fraud and have recovered a certain amount of money.
- **Blocked amount** (ImportoBloccato): If present, it represents a situation in which the phishing attempt was blocked by the bank account from which money has been taken.
- **Reporting office** (UfficioDenuncia): It specifies where the complaint was made.
- **Procedure’s number** (NumeroProcedura): It is an identifier of the complaint procedure.
- **Prosecutor** (Procura): It specifies the public prosecutor.
- **Notes** (Note): It allows a more comprehensive description of the episode.

*active holder* (TitolareAttivo), *passive holder* (TitolarePassivo), *active medium* (MezzoAttivo), *passive medium* (MezzoPassivo) and *Host* describe the relations mentioned in the previous paragraphs. Instead, table *Activity* (Attività) has just an ID. The relation between *Episode* and *Activity* is described in table *episode activity* (EpisodioAttività) with an attribute *Sorting* (Ordinamento) used to chronologically order common episodes.

![Database tables (Part 5)](image)

*Figure 36 - Database tables (Part 5)*
The following tables are related to reports management. A generic alert is defined by a unique ID and four other attributes: Date (Data), Type (Tipo), Content (Contenuto) and Inbound/Outbound describing if LEA, institutes or people are recipients or senders. An episode as well as an activity could be linked to an alert message through tables activity reporting (SegnalazioneAttività) and episode reporting (SegnalazioneEpisodio). An alert could be forwarded to people or institutes only if that alert is an external one. The other will reach only the Law enforcement agency. Tables describing alert recipients/senders are institute intervened (IstitutoIntervenuto) e people intervened (TitolareIntervenuto).

![Database tables (Part 6)](Figure 37 - Database tables (Part 6))
Latest tables are related to correlations and messages. Table *Correlation* (Correlazione) is defined with fields: *ID*, *Date* (Data), *Description* (Descrizione) and *SQL*. This last is used to store SQL queries used to correlate events and malicious activities. A correlation that obtains some concrete results generates a new activity. This link is stored in table *Activity Generation* (GenerazioneAttività) in which a correlation id and an activity id are associated in a record. Table *Messages* (Messaggi) will store instead information on date and time the message is sent and both *Content* (contenuto) and *Digital Signature* (FirmaDigitale) of such communication.

![Table: Correlation](image)

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
<th>Descrizione</th>
<th>SQL</th>
</tr>
</thead>
</table>

![Table: Messages](image)

<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
<th>Descrizione</th>
<th>SQL</th>
</tr>
</thead>
</table>

*Figure 38 - Database tables (Part 7)*

Table *Messages* will be managed through four different “raw” tables. The raw tables reflect the four types of messages that banks, BA but also LEA can insert to the platform.
UML
What follows is the UML class diagram for the database.

For now, we don’t use the UML to schematize the application and therefore to translate that schema into programming code but just for data modeling. We omit the description since it is quite similar to the one done before for the entity-relation model.

As in ER model paragraph, the first part of the UML is related to the main actors of the InfoSharing Platform infrastructure.

UML Part 2 concentrates on ICT elements.
UML part 3 describes malicious episodes and activities.

Figure 41 - Database UML (Part 3)

UML Part 4 defines signals and alerts of the platform.

Figure 42 - Database UML (Part 4)
The last part of the UML concentrates on correlations and messages sent by Banks, BA and Law enforcement agency.

![Diagram showing correlations and messages](image)

### 3.4.4.4 Communication

Communication protocols are the backbone of the whole infrastructure.

The InfoSharing Platform is designed to allow a secure sharing of information among banks and LEA and, for this reason, security requirements related to the data exchange of data must be well designed. A schema for managing such interchange and a model using two UML sequence diagrams are proposed.

Most important service of the InfoSharing Platform is to allow banks and BA to store information about phishing episodes and permit LEA to correlate data. During these activities the actors have to exchange message with each other (confirmation, validation, alerting, etc.).

The actors involved in the communication processes are:

- **Bank or BA** – Intended as the **terminal** used by the bank operator to send data to the platform.
- **Platform** – Intended as the InfoSharing Platform core.
- **LEA** (Law enforcement agency) – Intended as the **terminal** used by LEA operators to manage InfoSharing Platform core.

Communications among the previous actors can be split into two different groups:

- **Reporting Communications**: these are messages used to notify an actor for some information (e.g. insertion of a new data in the platform, important information to be notified to banks or BA by the LEA, statistical analysis on phishing trends, etc.).
- **Operative Communications**: these are messages used to ask an actor to do some work (confirmation on some data previously inserted, request for new data, request for new kind of correlation, etc.).

Since each actor can exchange messages with each other there are eight possible ways of communication among the previous.
- Bank/BA->$Piattaforma$ will be the **external signalling**; this describes the communication in which the InfoSharing Platform bank operator inserts new data in the InfoSharing Platform. The external signalling is mainly a Reporting Communication.

- **LEA -> Platform** will be the **internal signalling**; this describes the communication in which Law enforcement agency inserts new data in the InfoSharing Platform or asks to correlate some record stored in the database. Differently from the previous this can be both a Reporting and Operating Communication.

- **Platform -> Bank/BA** will be the **external notification**; this describes information the platform automatically communicates to the InfoSharing Platform bank operator. This is mainly related to some correlation results. These messages should be Reporting Communications.

- **Platform -> LEA** will be the **internal notification**; this describes information the platform communicates to the InfoSharing Platform Law enforcement agency operator mainly related to correlation results. Law enforcement agency User can receive this communication but when this is a consequence to a previous specific internal signalling sent to the platform (e.g. a Law enforcement agency agent asking to be informed on specific information or correlation results). These messages are Reporting Communications.

- **Bank/BA->$LEA$** will be the **semi-external news**; this describes information banks or BA want InfoSharing Platform Law enforcement agency operators to know. This is a direct communication between Bank/BA and LEA and the platform will convey information without changing database data (except for logging purpose). These messages can be both Reporting and Operative Communications.

- **LEA -> Bank/BA** will be the **semi-internal news**; this describes information Law enforcement agency wants banks to know. As in the previous case this is a direct communication between LEA and Bank. In this case, the platform is just a vehicle of information. These messages will be both Reporting and Operative Communications.

- **LEA -> LEA** will be the **internal news**; this describes internal information that flows among InfoSharing Platform Law enforcement agency users. In this case the news can be related to data interpretation. Also in this situation the platform is a vehicle of information. An example of this
kind of situation can be a LEA Detective asking an InfoSharing Platform LEA Administrator information on data correlation or platform’s activities. Also in this case, messages will be both Reporting and Operative Communications.

- **Bank/BA -> Bank/BA** will be the **external news**: this describes information that flows among banks but specifically among InfoSharing Platform bank operators. In this case, three sub-processes can be described:
  
  o **BA -> Bank**: in which BA informs banks about statistical analysis.
  o **Bank -> BA**: in which banks ask BA for statistical data or any other information.
  o **Bank -> Bank**: in which a bank asks another bank for general information. This is a type of communication at risk. It should not be implemented.

This kind of messages will be Reporting Communications. Also in this situation the platform is just a vehicle of information.

In a more general way it could be useful to identify another more simple interaction schema just between banks/BA and Law enforcement agency. This situation presents four different types of communications.

- **Bank/BA -> Law enforcement agency** will be the **Outside Information Signalling**: this describes information and new data that the InfoSharing Platform bank operator inserts and/or notifies to the platform and the Law enforcement agency.
- **LEA -> Bank/BA** will be the **Inside Information Signalling**: this describes notifications that LEA wants the bank to know.
- **Bank/BA -> Bank/BA** will be the **External Communication**; this describes the possibility for banks and BA to share information with each other. As before, this is a type of communication at risk and should not be implemented (at least for what concerns bank-to-bank communication).

- **LEA -> LEA** will be the **Internal Process**; this describes all the data that flows inside the LEA’s internal network.

Notifications could be sent in a synchronous or an asynchronous way. In the first case notifications will be the result of a correlation process, generated from the insertion of new information. In the second case the correlation process is continuous and the platform will send notifications any time something will be discovered.

Two different situations are presented as examples:

The first diagram proposes a platform in which correlations are activities generated on the arrival of new information. This does not preclude making correlations in any other moment but concentrates on the synchronous responses of the platform. Once an External or an Internal Signalling notifies a data to the platform, the InfoSharing Platform core starts a storing procedure followed by a correlation process. In case, the new phishing episode has common characteristics with the others previously stored, the platform notifies the fact to the LEA. LEA operator has to take care of the alert and can decide whether or not to report information related to that data to banks and make some other evaluations that can be possibly added to InfoSharing Platform database. In this last case, storing new information in the InfoSharing Platform core leads the platform to another correlation process.

**UML Sequence Diagram**

(hypothesis 1)

![UML Sequence Diagram](image)

*Figure 46 - Communication overview (Hypothesis 1)*
The second diagram proposes to separate, at least from the theoretical point of view, storage and correlation processes. The platform continuously makes correlations on data and there is no temporal relationship related to the External or Internal Signalling. In any moment the InfoSharing Platform core could signal something to the LEA Operator that, instead, acts precisely as described in the previous model. In this situation we have an asynchronous communication system.

**UML Sequence Diagram**

(hypothesis 2)

![Diagram](image)

*Figure 47 - Communication overview (Hypothesis 2)*

This last model is designed to handle very complex correlations. When the database will grow in size, correlation processes will become heavier and slower. This will increase gradually the time gap between the External Signalling and the Internal Notification. During the deploying and the start-up phases the first scheme provides faster responses upon notification of new phishing episodes or elements. Such scheme may be replaced by the second scheme when the platform will become much larger and performances could need an improvement.
4 POSSIBLE EVOLUTIONS

4.1 Fraud Taxonomy
The simplified Fraud Taxonomy proposed in chapter 2.1.1 could be evolved into a more academic and evidence-based model; The ACFE (Association of Certified Fraud Examiner) could be involved in this project.

4.2 Global Data Protection Regulation (GDPR) update
The new EU rules for protecting EU citizen and companies data will probably force an update in functionalities, Identity management and technologies for the InfoSharing platform described in this paper.

4.3 Enlargement of the stakeholders group
The same concepts presented in this paper can be applied to other On-Line Frauds involving Credit Cards, Telcos and On-Line gambling accounts. Basic Data set presented could be enlarged to encompass credit/debit cards data, top-up/roaming mobile numbers/accounts, and on-line gambling/gaming accounts.

This could be achieved involving credit card circuits/service providers, Telcos and On-Line Gambling Companies.

With this evolution, the service could help to identify and prosecute:

- Suspect fraudulent transactions done with credit/debit cards
- Suspect fraudulent top-up or telephone usage
- Suspect fraudulent usage of on-line games account (as used by lotteries, gambling, MMPGs ...)

4.4 International collaboration
The single country/LEAs DB could be connected to others in different countries, in order to create a network of InfoSharing Services. This improvement could be very useful to kill fraud lifecycle at a very early stage. The connection between different InfoSharing platforms will face some legal drill-down, especially for Data Protection in non-EU countries, and some technical engineering.

The latter, in particular, will focus on timing of cross-platform communication, dimensioning of databases and the implementation of the AAA model (Authentication, Authorization, Accounting) of the chain-of-trust in a connected InfoSharing Platform

4.5 Platform KPIs and SLAs
After a limited time of usage of the platform in real environment with real data, some key performance indicator and some hint of service level agreement between LEA and stakeholders could be drafted. Some light modification should be defined in the basic data gathering process to “close the loop” of the Fraud Lifecycle. In fact, with the proposed data model, as of today is not possible to distinguish between fraud that simply has been committed, but no information is present about the recovery of the money at the end the legal process or at the end of the international agreements between banks, and frauds that generated a real loss.
5 Conclusions

Economic transactions are more and more settled through ICT technologies and for political, social and technological reasons this trend will continue.

Fraudsters are adapting to this new model faster than financial institutions and law enforcement agencies, inventing every day new approaches to thief identity and money to banking customer. Focusing on the simplified fraud taxonomy presented in this paper, fraudsters are adopting phishing and crimeware to steal credentials and illegally launch transactions in order to get victims’ money.

These On-Line fraud methodologies produce an estimated loss to EU citizens of at least 250M€ per year, with an actual fraud rate for on-line accounts near to 0,08% of the total transactions.

This number can be lowered to a rate near to 1 loss every 1M transactions, as more mature markets demonstrate. This can be accomplished through usual bank-centric anti-fraud technologies (such as fraud management programme, including 2 factor authentication and fraud detection systems) and prosecuting model by LEAs. A big area of improvement can be achieved through the synergy between financial institutions and law enforcement agencies, with the collaboration of banking associations.

This synergy can be realized through an information sharing programme, focused on spreading in the whole banking system the data related to fraudsters and their accounts.

Due to privacy and commercial reasons banks do not want to share these data with their competitors: that is why the proposal presented in this paper is to let the LEA control the InfoSharing programme. With this solution LEA will obtain information needed to prosecute frauds, while banks will receive the updated lists of fraudulent accounts (IBANs, Credit cards, on-line betting accounts, …) to be inserted in their anti-fraud systems “watching lists” or “black list”

The framework of the “On-Line Fraud Centre”, as the InfoSharing programme has been called, has been analysed from 4 points of view:

- A clear organizational model and related roles/responsibilities
- A clear legal context in which it operates
- Defined procedures to be followed by the service actors
- A technological platform designed to support this service

The organizational framework is based upon the “hub & spoke” model, where LEA act as the hub and banks/BA act as the spokes. LEA will be in charge to coordinate and control the communication, gaining access to the whole data in order to correlate, identify and prosecute fraudulent actions. Banks will be responsible to furnish live fraud data to the LEA and to retrieve fraudulent account list from the service. BA will analyse trends and furnish statistical data.

Attention has to be adopted by banks in the treatment of customer data (especially if they’re victims of fraud) due to DPD (Data Protection Directive), stating clearly in their data treatment declaration that this personal data will be shared with LEAs and with the banking system as a whole in order to protect the customer from fraudulent attempts.

LEA Operators and bank’s operators will follow specific documented procedures, related to the sharing process, in order to guarantee maximum security for customer data and best results for fraud prosecution.
The On-Line Fraud Centre, in order to cope with the quantity of potential data, needs a technological architecture to support effectively the service. The proposed technological solution is a custom web application, developed through standard development tools, based upon a structured Database. The systems supporting the service will be hosted in LEA facilities, in order to protect the physical part of the centre. Even the logical part of the service has to be secure and scalable: every user of the system, either internal from the LEA or external from Banks and BA, will need to strongly authenticate to access their specific interface; The communication will be encrypted through VPN or HTTPS; Access to data will be regulated and audited by LEA Administrators; front end web systems will be protected in a specific DMZ; workstations accessing the service will need to be secured; Application and DB will be documented and designed with security in mind.

The On-Line Fraud Centre can be eventually evolved with a more structured and measured (KPIs) approach, that encompasses new EU rules for Data Protections, new stakeholders (i.e. different LEAs or different economic actors, like credit card service providers, telcos or on-line betting companies), and eventually an international network of similar services.

With such a service in place a big improvement in fighting the fraudsters could be achieved: banks could exit their isolation state and start to share their data in a trusted and safe environment, while LEA could act more on-time and answer to the needs of the banking system and the citizens.

Only with a structured communication between these actors, such as the proposed solution presented in this paper, the ever increasing losses by on-line criminals could be prevented, blocked and prosecuted.
BIBLIOGRAPHY


GLOSSARY

BA Interface: It represents the set of web pages and/or applications that allows a Banking Association operator to access InfoSharing Platform services.

Actor: This term always refers to an entity that accesses the InfoSharing Platform and uses InfoSharing Platform services. InfoSharing Platform actors are banks, banking association and the Law enforcement agency. In some specific cases this set can be enlarged to people.

Alert: This is a way to send messages to actors of the InfoSharing Platform even if they are not logged in the platform itself. It is important to underline that no sensible information are sent within this messages unlike in the Communication process.

Bank Interface: It represents the set of web pages and/or applications that allows a bank operator to access InfoSharing Platform services.

Bank Operator: This represents the bank employee chosen to access and use platform’s services. The term is often used, by extension, as a synonym of the terminal utilized by that employee.

Common Attribute Correlation: This is a correlation on data stored in the InfoSharing Platform database that concentrates on episodes and activities that have the share some equal attributes.

Common Dynamic Correlation: This is a correlation on data stored in the InfoSharing Platform database that concentrates on episodes and activities that show similar behaviours.

Communication: This term refers to all the messages, notifications and signalling sent through the InfoSharing Platform. It is important to differentiate a Communication from an Alert.

Core or InfoSharing Platform Core: It is the most important part of the InfoSharing Platform. It consists of hardware and software related to application backend and databases and also the terminal used to directly access them.

Correlation: This term indicate all the possible operations made on data and information stored. For example, it can be a simple research on the data, an information grouping process on a given common attribute or a data mining procedure.

Database: It indicates the database used to store banks’ and Law enforcement agency’s information.

Data Unit: This expression often refers to the message sent by banks or any other actors to the platform containing new data and information that have to be stored in the InfoSharing Platform database.

DMZ: It indicates the part of the InfoSharing Platform infrastructure where the frontend servers are located.

External Components: This expression indicates those parts of the InfoSharing Platform infrastructures that stay outside the Law enforcement agency DMZ. In other words, the external components represent any machine but also any kind of software that is separated from the InfoSharing Platform Core by the DMZ.

External Interface: This represents the set of web pages Banks and BA use to interact with InfoSharing Platform services.

Full Security: It is one of the two guidelines we want to follow in the development of the InfoSharing Platform. It is related to provide a maximum level of security on information and services.
Full Usability: It is one of the two guidelines we want to follow in the development of the InfoSharing Platform. It is related to provide a friendly and easy-to-use access to information and services.

Internal Interface: This represents the set of web pages LEA uses to interact with InfoSharing Platform services.

News: This indicates all the communications that flow inside the platform that are not notifications neither signalling. In this kind of communication the InfoSharing Platform is neither the sender nor recipient but only a vehicle of information.

Notification: This indicates communications that start from the platform and reach the Law enforcement agency or banks.

Operative Communication: it is a kind of message used to encourage or force an actor to do something. It is mainly related to pushing and retrieving data on the platform.

Law enforcement agency Administrator: This represents the Law enforcement agency employee that has credentials to access directly to the platform. She/he can manage all the services, and has to take charge of the news sent by InfoSharing Platform bank operators.

Law enforcement agency Auditor: This is the Law enforcement agency agent assigned to control InfoSharing Platform’s log files.

Law enforcement agency Interface: It represents the set of web pages and/or applications that allows a Law enforcement agency operator or a Law enforcement agency agent to access InfoSharing Platform services.

Law enforcement agency Operator (or Detective/Investigator): This represents every Law enforcement agency employee that access the InfoSharing Platform to use its services.

Reporting Communication: it is a kind of message used to notify an actor for some information.

Related Common Attribute Correlation: This is a correlation that, as the Common Attribute Correlation, concentrates on episodes and activities that share some equal attributes but it does not imply that those attributes belong to correlated elements.

Related Common Dynamic Correlation: This is a correlation that, as the Common Dynamic Correlation, concentrates on episodes and activities that show similar behaviours but it does not imply that those behaviours are encountered among the correlated elements.

Signalling: This indicates additions of information to the InfoSharing Platform database. The signalling could be external or internal. Four different kind of signalling depending on the given information are identified: mirror site, suspect IP, phishing episode and general.