Money laundering, terrorist financing and how to contrast them: data and text mining in business intelligence solutions

F. R. Federici
Compliance Daily Control, Lugano, Switzerland

Abstract

The application of data and text mining technologies in business intelligence solutions can give a great contribution to fighting money laundering and terrorist financing. These crimes are multiform and so laws are important, but they can't be the only tool to contrast them. Money laundering is a well known offence, but it abandons classic methods in favour of new schemes, thanks to informatics: so called "cyber-laundering". Money laundering has assumed great relevance since September 11th 2001, because it shows many links to terrorist financing. Data and text mining can be useful in understanding links between people and suspicious transactions in order to track real money laundering and terrorist financing activities. There are American and European initiatives in which data/text mining and business intelligence assume a strategic rule. However, these applications have a critical aspect that involves privacy.
Keywords: money laundering, terrorist financing, data and text mining, business intelligence.

1 Introduction

Money laundering offers a lot of competitive advantages and, in particular, criminal organizations use "cleaned" money in order to infiltrate themselves in the vital points of the lawful economy in order to strengthen their base of power and prestige. The recent evolution of IT technologies offers new opportunities to people who want to launder goods: this is cyberlaundering. Money laundering and financing of terrorism prosper also thanks to the complex financial systems and to nations with weak, ineffective and corrupted structures that should fight these two crimes (Schott [1]).

WIT Transactions on Information and Communication Technologies, Vol 38, © 2007 WIT Press
www.witpress.com, ISSN 1743-3517 (on-line)
doi:10.2495/DATA070311

After 11th September, the interest in the financing of terrorism has grown: this crime is often indicated as “reverse money laundering”, because in many cases the process begins from the presence of financial stocks of licit origin in order to hide the destination finalized to commit a criminal or a terrorist action. These goods often pass through the Informal Value Transfer System (IVTS), networks of people based on confidence and personal relationships (Schneider [2]). This way supplies a good capillarity, because money must be available to the operating cells all over the world. However, the International Monetary Fund [3] has estimated that the amount aggregate of laundered money in the world can be comprised between the 2 and 5% of the world-wide Pil, that in monetary values can be translated in approximately 1,5 trillions of $.

2 Is it all ok?

Present anti-money laundering (AML) strategy was conceived at the end of the 80s. The hinge of such repressive system is in the cooperation of financial professionals and banking system, including more recently and extensively, operators like lawyers and accountants, but also operators that are not financial (real estate mediation, commerce of jewels, etc.) which are called to support the activity of monitoring and protection of the legal system from the exploitation by international terrorism. The international AML standard is represented by the Forty Recommendations prepared by the Financial Action Task Force on Money Laundering (FATF) [4]. Since October 2001 there are also Nine Special Recommendations on Terrorist Financing (FATF [5]).

Some factors are important: identification and acquaintance of the customer; recording of the operations above a threshold, both with a single operation and with a sequence of operations in the arc of a determined period; the communication of suspect operations to the national AML agency. There are some weaknesses in this strategy: they can be found, in particular, in the systems of payment via Internet, that make possible to carry out “peer-to-peer” payments with cryptographic applications and different levels of anonymity. Great confidentiality is obtained moreover, historically, through tax and financial havens. These methods don’t make easy controlling activities, especially when there are several passages between more users before the demand for conversion. Surely it’s more difficult to control IVTS, Meldrum and Ciotti Galletti [6]. That’s way it’s impossible to define a "passpartout" money laundering prevention.

3 A new strategy

Money laundering and terrorist financing are phenomena that are not explained only by laws or by technique, but they have a social dimension. A multidimensional approach occurs: it should show the complexity and the systemic nature of the two crimes. Improving the informative databases of several institutions can lead, thanks to business intelligence (BI), to a progressive

convergence and integration of knowledge management systems, in order to investigate not only unusual money transactions or a high recurrence of the transactions, but also to discover unexpected relations that can gush from data. Links between a person, or some people, and the transactions are the crucial fact. Fighting money laundering and the financing of terrorism needs meticulous examination of these links as a fundamental process: it must include networks' analysis, in particular Internet, wire-tapping and travels' analysis, with the fusion of not classified sources and hidden intelligence, like the electronic interceptions. BI is important to give meaningful links in order to trace criminal activities. BI helps the user to take effective and timely decisions, thanks to a comprehensive vision of the activity in analyzing. For example, the fact that a transaction has exceeded a determined threshold can be integrated with other pieces of information that define the context in which the operation takes place.

People who work using these technologies have developed tools that facilitate the collection and the analysis of great amounts of structured and not structured data: many solutions of BI, in fact, include tools of data and text mining. Data mining (DM) is the process that, through mathematical techniques and statistics, extracts valid and exploitable information, previously unknown, from great amounts of structured data, that are already included in databases, organized and ready for being managed easily from IT applications (Zanasi et al. [7]). The DM, in fact, is distinguished from the traditional approaches of data analysis, because the obtained information is completely new. It must be also valid in order to avoid wrong or not useful data. DM can be used for many purposes: market management, risk management and fraud management. The principal operations are predictive modelling, database segmentation, link analysis and deviation detection, together with different techniques, Figure 1.

Text mining (TM) is the process that extracts information from not structured data, as they can be texts, in analogical and digital format (Zanasi [8]), Figure 2.

	Market Management	Risk Management	Fraud Management
Applications	✓ Target marketing ✓ Customer relationship management ✓ Market basket analysis ✓ Cross selling ✓ Market segmentation	✓ Forecasting ✓ Customer retention ✓ Improved underwriting ✓ Quality control ✓ Competitive analysis	✓ Fraud detection

	Predictive Modeling	Database Segmentation	Link Analysis	Deviation Detection
Operations				
Techniques	✓ Classification ✓ Value prediction	✓ Demographic clustering ✓ Neural clustering	✓ Associations discovery ✓ Sequential pattern discovery ✓ Similar time sequence discovery	✓ Visualization ✓ Statistics

Figure 1: DM applications, Zanasi et al. [7].

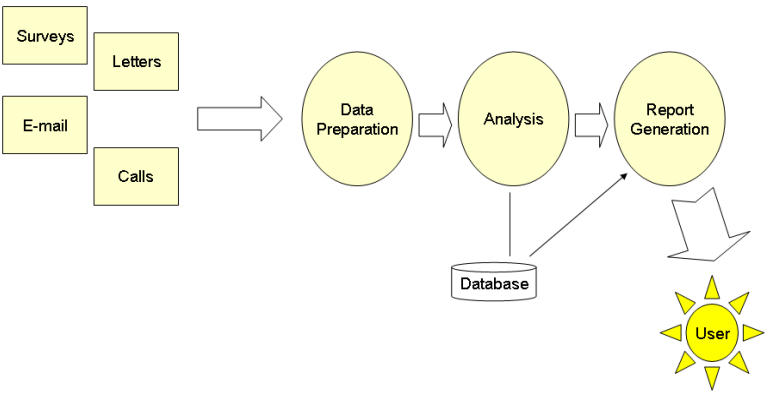

Figure 2: TM process, Zanasi [8].

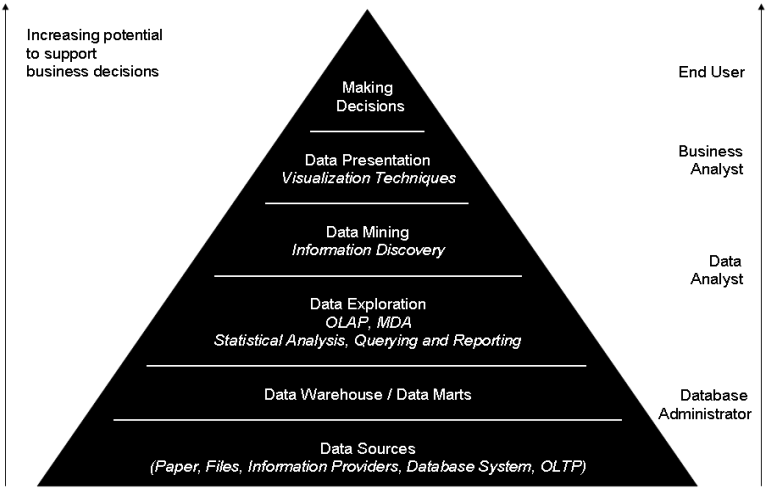

Figure 3: Mining processes and BI, Zanasi et al. [7].

When such information is available, it can be analysed and compared with the rest of the informative patrimony of BI. There are different technologies, classified in accordance with their potential value to support business decisions. Lots of people give their contribution to exploit the power of BI systems: database administrators, data analysts, business analysts. They work together to provide the end user the capabilities to make well pondered decisions trough user-friendly interfaces, Figure 3.

4 The potential of text mining

TM is an interdisciplinary field that it is based on the DM, statistical-mathematical techniques and linguistics. The first TM applications appeared in the second half of 80s, but developed in the 90s. So this technology is recent, but has however great potentialities and many uses: in fact, it can extract information from great amounts of textual data, classify it in thematic categories, but also show unexpected correlations of high informative value. Typical organizations that can need TM are credit institutions, assurances, intermediation societies, financial institutions, research centres and generally all those agencies that have to manage great volumes of documents. Many experts emphasize human participation in the development of the process: users must be aware of some limits of this technology and, therefore, they must integrate them with human discernment, in order to avoid wrong conclusions on the base of the presented data. It isn't useful working with this software "as it is".

4.1 Factiva

To fight money laundering and terrorist financing, Factiva TM [9] offers an interesting function with the program "Public Figures & Associates" (PFA): it supplies detailed profiles - political, economic, social condition and other biographical details - of 500.000 politically sensitive people all over the world and of those people that are linked to them, like relatives and friends, with the recording of orthographical variations and alias. PFA includes also an internet crawler. The list of individuals implemented in PFA presents the names of known terrorists and criminals of different species, but also potential criminals.

4.2 RDC

Something similar to the PFA is the service that is in the RDC suite for AML [10], GRID, the Global Regulatory Information Database, a collection of data that contains approximately 1,5 million people: everyday there are updates and the new entries are approximately 700-800 everyday. The sources of data are governmental or Interpol lists but also newspapers, reviews, transcriptions of TV and radio, Internet, etc.

4.3 TEMIS

"Text Mining 360° Skill Cartridge" is the suite of TEMIS [11], an advanced tool for text analysis that concurs to the automatic content exploration of all types of textual data. XeLDA, the multilanguage engine, standardizes unstructured documents in order to explore their content. Two key factors of TEMIS suite are the dictionaries and the rules of entity extraction. The Insight Discoverer Extractor is a server for the information extraction in textual documents that includes automatic identification of the language, morpho-syntactic analysis of words, and knowledge extraction, that is the identification of relations between

entities: this demonstrates great usefulness in finding connections between people and financial transactions. A clear example of these networks of connections is shown in Figure 4.

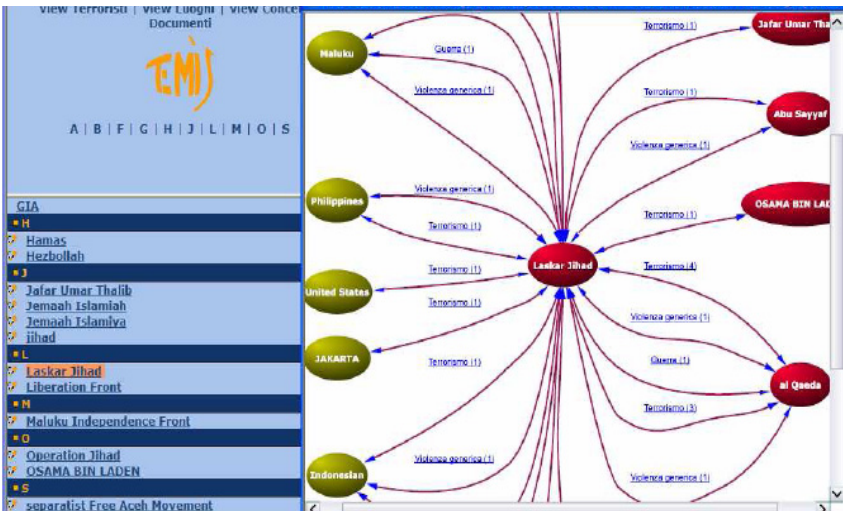

Figure 4: Connections between people and transactions, TEMIS [11].

5 American and European initiatives

There are, in particular, two wider projects that include such technologies: the first, ADVISE (Analysis, Dissemination, Visualization, Insight and Semantic Enhancement), is American, and the second is the European Security Research Programme (ESRP), to whose implementation contributes ESRAB, the European Security Research Advisory Board.

5.1 USA: ADVISE

The Homeland Security Department [12] often manages an enormous amount of data, whose value cannot be explored without technologies for the knowledge management, like BI. There are a lot of objectives: for example, the identification of terrorist threats and vulnerabilities and the discovery of terrorist financing activities. The user-friendly interfaces, linked to BI, are important for the entire architecture of ADVISE: they, together with fast simulations, are fundamental in order to discover the complex relations between people and transactions and so they can obtain useful information to fight terrorism.

Analyzing text data, in particular, is difficult and highly labour intensive: for this reason ADVISE cannot be based exclusively on a manual extraction, but it is equipped with automatic tools that can arrange typical DM and TM techniques. To find information from databases and to make link analysis on the structured

information, data must be converted and formatted in order to be stored in datawarehouse, Coffman et al. [13]. The first step of the text analysis process is the transformation of not structured documents into structured ones. ADVISE implements therefore sophisticated technologies, Figure 5. In particular TM shows great potential in the selective study of e-mail traffic in order to foil terrorist plans.

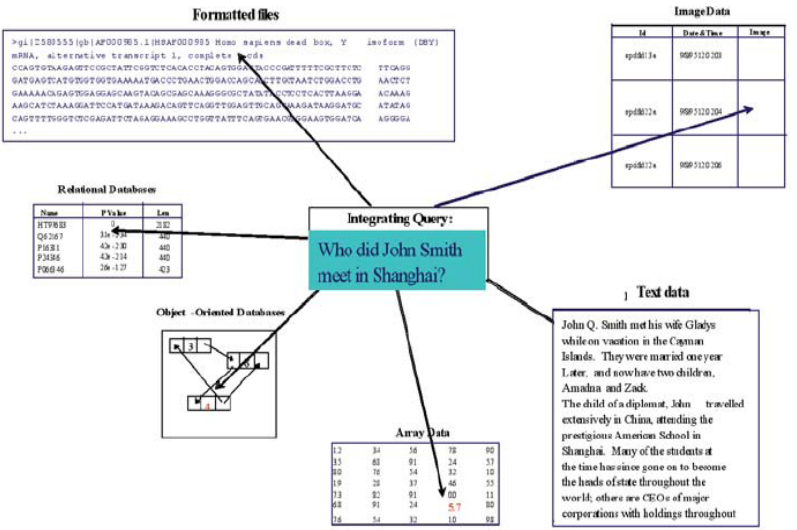

Figure 5: Different typologies of data for a query, DHS.

5.2 UE: ESRP

In the 2003 the European Commission called a Group of Personalities in the field of security research. The primary aim was to propose principles and priorities of ESRP. From the report "The Research for a Secure Europe", introduced from the Group in 2004, it's clear that the Communication of Commission (COM 2004 590 final, entitled "Security Research: the Next Steps") [14] proposed the creation of ESRAB. ESRAB should contribute to the content and the implementation of ESRP, already financed with 3 billions of Euro, inside of 7^FP-Framework Program [15]. ESRAB is composed by fifty experts coming from several fields: 25 from Institutions and 25 from private companies, industry and research organizations. The Commission can consult ESRAB on every issue relative to security. ESRAB presents recommendations in many areas, in particular: strategic missions, warm areas and priorities established for ESRP on the base of the report "Research for a Secure Europe" and considering the creation of the European Defense Agency, together with national activities; the technological abilities are strategic in order to improve the technological base of the European industry in the field of security research and competitiveness.

Inside ESRP there's a financing to the development of AML technologies and solutions. The mission areas in which ESRP acts are: protection against terrorism and organized crime, borders security, protection of critical infrastructures and security restoration in crisis situation [16,17]. In all these fields information management systems are fundamental in order to supply the greatest knowledge and ability to manage crisis to decision maker, Table 1.

Table 1: First two priority technology areas of ESRP.

Technology Domain	**Priority Technology Areas**
Signal and information technologies	• data fusion techniques • data collection/classification • image/pattern processing • information fusion technology • data & information management technology (DB, etc.)
Artificial Intelligence and decision support	• data & text mining • IKBS/AI/expert techniques • knowledge management • modelling and simulation • optimisation & decision support technology

The result is a system that analyzes information deriving from financial, demographic data and from investigative agencies to draw better conclusions thanks to: discovery of common behavioural characteristics in terrorism financing and money laundering; data extraction and discovery of pattern and hidden correlations that indicate a specific threat and the presentation of the data in order to help the decisional process; models of forecast and correlation in order to generate warning on eventual future threats; automatic tools to see criminal behaviours.

6 Privacy

The applications I have described have a critical aspect, linked to the privacy. In fact, secrecy is a remarkable barrier for the relative intelligence activities. The relationship between the collaboration of the institutions and AML activity is undoubtedly based on a game of delicate equilibrium: defence of the society from the criminal phenomena, protection of business and respect of privacy. It will be difficult, therefore, to conciliate the development of the intelligence operations with the respect of the honest citizens' privacy. Zanasi [18] sharply says that unfortunately there's little privacy in digital society and protective measures have appeared not effective, because teenager *crackers* have penetrated into "protected" systems all over the world.

Mark Rotenberg says that DM and TM are useful to fight money laundering and terrorist financing, but there's the idea that they can be used for other purposes (industrial espionage, customer profiling, political sentiment analysis, etc.), Kelley [19]. Public authorities' interference into citizens' private life shouldn't be set by generic discretion, as it happens in many situations Mulinari [20], but it must be practised only with in front of defence necessity of our society under control measures and with guaranties to avoid misuses: a "checks and balances" system must be adopted (Palmieri [21]).

Winston Churcill said: "We live in democracy when who knocks on the door at 7 in the morning can only be the milkman" (Palmieri [22]).

7 Conclusions

DM and TM, implemented in BI solutions, can give a remarkable contribution for an effective strategy in the fight against money laundering and terrorist financing. These tools have some limits. They are not completely self-sufficient and immediately usable "as they are". It's necessary that expert analysts may know how to structure, analyze and interpret the output that the software has created. Although these instruments can help to find patterns and connections between many entities, they often don't present their value and meaning in an immediate way: the user must conduct the analysis. Terrorists usually use IVTS and financing sources that are difficult to discover. Greater success would be probably obtained using DM and TM techniques, like link analysis, with data fusion, in order to create specific behaviour models thanks to a massive use of sensitive data, but in this case privacy is over. The effectiveness of these technologies has been demonstrated, but the eternal dilemma is still difficult to resolve: security or freedom?

References

[1] Schott P.A., "Reference guide to anti-money laundering e combating the financing of terrorism", chapter 1, 2006.
[2] Schneider F., "The hidden financial flows of Islamic terrorist organisations: some preliminary results from an economic perspective", Johannes Kepler University of Linz, 2003.
[3] International Monetary Fund http://www.imf.org/external/data.htm.
[4] FATF Financial Action Task Force on Money Laudering – Groupe d'Action Financière International (GAFI) - The Forty Recommendations, http://www.fatf-gafi.org/document/28/0,2340,en_32250379_32236930_33658140_1_1_1_1,00.html#40recs.
[5] FATF Financial Action Task Force on Money Laudering – Groupe d'Action Financière International (GAFI) - Nine Special Recommendations on Terrorist Financing (October 2001 versions, updated 22 October 2004) http://www.fatf-gafi.org/document/9/0,2340,en_32250379_32236920_34032073_1_1_1_1,00.html.

[6] Meldrum C., Ciotti Galletti S., "The logistics of Terrorism and Organised Crime: Networks, Convergences and Implications", European Society of Criminology 6th Annual Conference, University of Tübingen, August 26-29 2006.
[7] Zanasi A., Hadjinian P., Stadler R., Verhees J., Cabena P., "Discovering data mining – From concept to implementation", Prentice Hall Ptr, Upper Saddle River, NJ, 1998.
[8] Zanasi A., "Text mining and its Applications to Intelligence, CRM and KM", Wit Press, 2005.
[9] Factiva Dow Jones & Reuters Company, http://www.factiva.com.
[10] RDC, Regulatory Data Corp, http://www.regulatorydatacorp.com.
[11] TEMIS Text Intelligence, http://www.temis-group.com.
[12] DHS Workshop on data sciences, "Data sciences Technology for Homeland Security Information Management and Knowledge Discovery, p.21, 22-23 September 2005.
[13] Coffman T., Greenblatt S., Sherry M., "Graph-based Technologies for Intelligence Analysis", Communication of the Acm, vol.47, n.3, March 2004.
[14] COM 2004 590 final, "Security Research: the Next Steps", Brussels 7/9/2004, http://europa.eu.int/eur-lex/en/com/cnc/2004/com2004_0590en01.pdf.
[15] 7° FP – Framework Program, http://cordis.europa.eu/security/home.html.
[16] Commission of the European Communities, "Green paper on detection technologies in the work of law enforcement, customs and other security authorities", p.5, Brussels 1/9/2006, COM(2006) 474 final
[17] "Meeting the challenge: the European Security Research Agenda. A report from the European Security Research Advisory Board", pp.29-33, September 2006.
[18] Zanasi A., "Nuove forme di guerra, nuove forme di intelligence: il text mining", Modernizzazione e Sviluppo (Quaderni del Centro Gino Germani), n.3, "Intelligence nel XXI secolo, p.5, 2001.
[19] Kelley M., "Feds sharpen secret tools for data mining", USA TODAY, 20/07/2006.
[20] Mulinari S., "Cyberlaundering. Riciclaggio di capitali, finanziamento del terrorismo e crimine organizzato nell'era digitale", PricewaterhouseCoopers, Pearson Prentice Hall, pp.292-293, 2003.
[21] Palmieri N.W., "Diritti fondamentali a rischio", Pitagora, p.15, 2003.
[22] Palmieri N.W., "Sicurezza o libertà? Introduzione al diritto di internet", Pitagora, p.9, 2004.