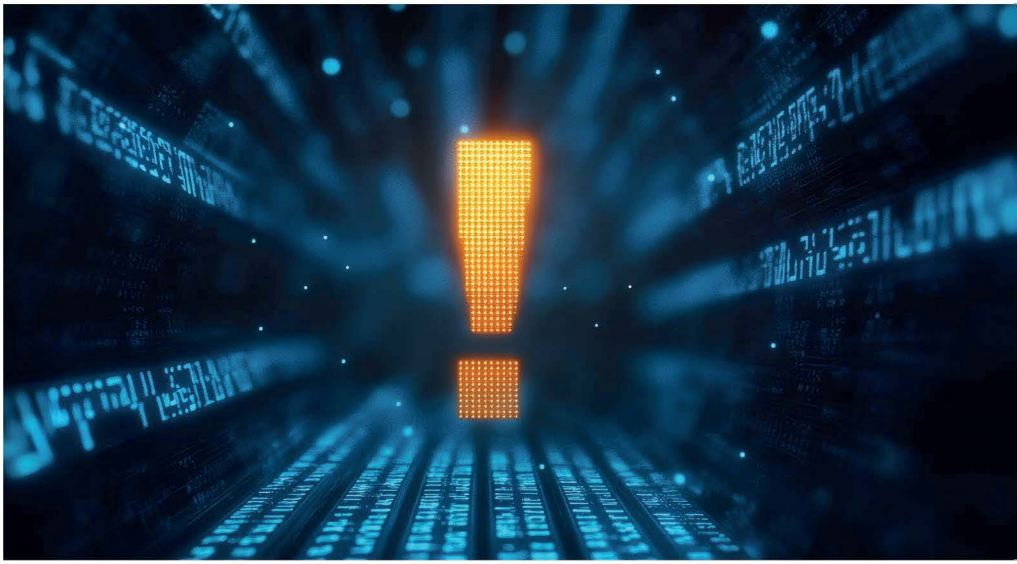
Fraud Repression through EDucation

Edited by

Mario Calabrese and Maria Felice Arezzo



Giappichelli

Fraud Repression through EDucation


IUSTITIAM COLIMUS

Fraud Repression through EDucation

Edited by

Mario Calabrese and Maria Felice Arezzo

Giappichelli

VIA PO, 21 - TEL. 011-81.53.111

http://www.giappichelli.it

ISBN/EAN 979-12-211-1474-4
ISBN/EAN 979-12-211-6372-8 (ebook)

"Funded by the European Union's EUAF programme under Grant Agreement No 101101784 2022-IT-FRED2. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them"

DOI 10.82018/9791221163728

G. Giappichelli Editore

This book was printed on certified
paper, 100% recyclable

Table of Contents

page

Introduction

The urgent imperative to safeguard public resources and maintain the integrity of financial systems has never been more pressing. In an era defined by rapid digitalization, cross-border flows of capital, and increasingly sophisticated schemes of misappropriation, the dual challenge of preventing fraud and fostering innovation demands new paradigms of collaboration, research and practice. It is against this backdrop that the FRED2 ("Fraud Repression through EDucation²") project was conceived and implemented under European Union grant agreement No. 101101784 (2022–IT-FRED2), and it is within the intellectual and organizational ecosystem of the ImpreSapiens Research Centre at Sapienza University of Rome that its outcomes have been generated, curated and now presented in this volume.

Since its foundation in 2009, the ImpreSapiens Centre has pursued a mission that transcends traditional disciplinary boundaries. Promoted jointly by the Faculties of Economics and Political Science and supported by the Departments of Communication and Social Research (CORIS), Business Law and Economics (DEI), Management, and Methods & Models for Economics, Territory and Finance

DOI: 10.82018/9791221163728.00

(MEMOTEF), ImpreSapiens operates as an inter-faculty hub for business innovation, organizational re-engineering, continuing education, workplace safety and the study of entrepreneurial and labour market dynamics. By integrating legal, economic, managerial, social-scientific and quantitative expertise, the Centre cultivates an ecosystem in which public and private organizations can embark upon digitally-enabled transformation with methodological rigor, strategic clarity and sustainable impact.

A flagship expression of this ecosystem is the Business Innovation Hub (BIH), inaugurated on 22 May 2023. The BIH—both physical and virtual—serves as an incubator and accelerator for start-ups and nascent ventures, offering mentoring, training programs, design-thinking workshops and networking opportunities with investors, corporate partners and institutional stakeholders. Initiatives such as the BIH Challenge and the Innovation Award foster a culture of continuous entrepreneurial experimentation, catalysing high-potential projects in technology, social impact and creative industries. In doing so, the BIH demonstrates ImpreSapiens's conviction that innovation flourishes where academic knowledge, professional practice and institutional support converge.

It is in this fertile environment that FRED2 took shape as a transnational, multidisciplinary endeavour to strengthen Europe's capacity to prevent, detect and deter fraud in the management of EU funds. Coordinated by Sapienza University of Rome in partnership with anti-fraud units (AFCOS) and academic and professional actors in Italy, Greece, Finland and Belgium, FRED2 pursued four interlocking objectives: (1) to design and deliver an experiential "co-lab-learning" itinerary—conferences, webinars,

workshops and study visits—uniting academics, practitioners and students in joint problem-solving; (2) to constitute a mixed academic-practitioner task force charged with mapping the conceptual terrain for a pilot study in fraud detection; (3) to establish an Anti-Fraud Observatory with a durable transnational network for information-sharing and best practices; and (4) to produce dissemination materials that amplify the project's insights across Europe. Over two years, more than one thousand participants engaged in eleven events, both online and in presence in Italian and abroad locations, forging relationships of trust, exchanging tacit knowledge and pooling diverse perspectives on the vulnerabilities and controls associated with subsidized finance.

This volume—"FRED2 Final Publication"—collects the four core scientific contributions that emerged from that collaborative journey. Together, they trace a coherent intellectual trajectory from theoretical foundations to empirical evidence and prototype tool development:

1. **Co-Production as a Key Driver in Capacity-Building Processes** (Eleonora Cova) examines the psychosocial mechanisms by which co-production—joint creation of value by providers and users—generates trust, reflexivity and collective learning. Drawing on service-management theory, social-exchange and trust models, this chapter demonstrates how immersive co-lab activities foster legitimacy and new mental models among stakeholders.
2. **Digital Transformation and AI in Fraud Detection: Challenges and Opportunities in Subsidized Finance** (Francesco Bellini) surveys the state-of-the-art in Big Data analytics, machine-learning algorithms, neural networks and blockchain applications for fraud detec-

tion. It balances the promise of real-time anomaly identification against ethical, privacy and operational constraints, and proposes a multi-disciplinary architecture for future decision-support systems.

3. **EU Fraud Risk Profile Analysis: Results of the FRED2 Survey on Italian Registered Chartered Accountants** (Tommaso Di Nardo & Antonia Coppola) presents the findings of a large-scale questionnaire ($N \approx 3\,000$) exploring accountants' direct and indirect experiences of suspected fraud, regional variations in risk perception, and the diagnostic utility of financial ratios (e.g. operating subsidies to production value; capitalized costs to fixed assets).
4. **Pilot Research Study** (Maria Felice Arezzo, Francesca Iandolo, Roy Cerqueti, Domenico Vitale and Giuseppina Guagnano) describes the development and validation of a prototype decision-support system. Employing unsupervised anomaly-detection algorithms, Benford's Law tests and a taxonomy of thirty-seven Key Fraud Indicators, the pilot translates conceptual mappings into an analytic workflow that flags high-risk units for focused investigation.

Each chapter combines rigorous literature review, methodological transparency, empirical illustration and practical guidance. Collectively, they furnish policy-makers, auditors, compliance officers, academic researchers and technology developers with a comprehensive toolkit: theoretical rationales for co-production, architectural principles for AI-enhanced detection, survey-derived risk profiles, and an extensible prototype for anomaly scoring. In this way, the volume exemplifies ImpreSapiens's integrative ethos—melding legal, managerial and technical lenses to address a societal challenge of acute public interest.

We extend our gratitude to the European Union's EU Anti-Fraud Programme for funding FRED2, to the twelve partner entities whose collaboration made the project possible, to the AFCOS authorities for their institutional support, and to the many speakers, participants and peer reviewers whose insights enriched every phase. Above all, we acknowledge the dedication of the ImpreSapiens research center team, whose interdisciplinary commitment has yielded both scholarly advancement and concrete instruments for protecting Europe's financial interests.

It is our hope that this volume will not simply document the achievements of FRED2, but will serve as a springboard for further research, cross-sector partnerships and policy innovation. The complexity of fraud in the digital age demands sustained cooperation, continual refinement of methods and an unwavering commitment to transparency. By sharing these findings, we aim to strengthen the collective capacity of European institutions, professional communities and scholars to anticipate emerging threats, to design resilient control systems, and to uphold the integrity of funds dedicated to the public good.

Mario Calabrese
Director of the ImpreSapiens Research Centre
Sapienza University of Rome,
Co-Editor, FRED2 Final Publication

Maria Felice Arezzo
Coordinator of the FRED2 project
Professor of Statistics, Sapienza University of Rome
Co-Editor, FRED2 Final Publication

1.
Co-Production as a Key Driver in Capacity-Building Processes

by *Eleonora Cova*

1.1. Introduction

This contribution offers a snapshot of the knowledge-building process within the FRED2 project, providing a psychosocial perspective on co-production as a key driver. This chapter aims to explore how co-production has shaped knowledge creation through the active engagement of diverse stakeholders.

The FRED2 project was structured around an "experiential co-lab-learning itinerary", engaging university professors, researchers, experts in EU project fraud prevention, institutional bodies such as AFCOS, and students

DOI: 10.82018/9791221163728.01

from both Italian and European universities. Their interaction proved essential in achieving the project's learning and capacity-building objectives.

Co-production—understood as a collaborative process in which clients and providers jointly generate value—emerges as a crucial mechanism for knowledge creation. Rooted in service management theory (Normann, 1984), the concept has evolved into a multidisciplinary approach, particularly effective in fostering learning and capacity building at both the meso (individual) and macro (organizational) levels.

Three core elements underpin co-production: interaction, trust, and reflexivity. Interaction is the foundation of any collaborative endeavor. Trust, gradually built through information sharing and reciprocal engagement, enables deeper knowledge exchange and the emergence of new power dynamics. Reflexivity allows continuous evaluation of roles and processes, fostering learning and improvement.

The FRED2 project embodied these principles through a range of co-productive initiatives—including webinars, study visits, and collaborative workshops—with the intention of establishing an anti-fraud observatory. This initiative strengthened transnational and multidisciplinary cooperation, enhancing knowledge-sharing and collective capacity building.

In conclusion, FRED2 stands as a prototypical example of how co-production processes enable collaborative learning, foster trust and transparency, and contribute to building legitimacy within complex, multi-stakeholder environments.

1.2. Co-Production

Co-production is a multi-faced concept (Brudney & England, 1983). In fact, it can indicate a large set of instruments and ways of working that improves the quality and efficiency of public and organizational services (Ewert & Evers, 2014) through the contribution of service users to the provision of services (Realpe & Wallace, 2010). It is defined as the voluntary or involuntary involvement of stakeholders in the provision of any of the design, management, delivery and/or evaluation of public services (Ostrom, 1999; Osborne et al., 2016).

Early work on public services and co-production can be traced to the "new public governance" (NPG) scenario that emphasizes partnership and collaboration instead of the "new public management" (NPM) paradigm. This early concept of public services, which approached citizens as "(potentially) active co-producers of the services they receive" (Fledderus et al., 2014), accounted, at least partly, for a definition of co-production and focused on the relationship between users and producers (Bovaird, 2007), the effects caused, the value generated (Prahalad & Ramaswamy, 2004; Vargo & Lusch, 2008; Payne et al., 2008; Svensson & Gronroos, 2008; Spohrer & Maglio, 2008; Gronroos, 2011; Edvardsson et al., 2011) and the motives of co-production (Verschuere, 2012). However, in the past, co-production based on NPM paradigm treated citizens as a passive recipient of public value. Currently, this conceptualization is outdated because the co-production assumes that users have a core role in co-planning, co-designing, and co-delivering services (Bovaird, 2007; Parks et al., 1981; Palumbo et al., 2018). The NPG view

opens up the co-production to look into the processes inside. By focusing on the extended concept of public services, the co-production is the "public sector and citizens making better use of each other's assets, resources and contributions to achieve better outcomes or improved efficiency" (Bovaird & Loeffler, 2012, p. 1120). Building on this initial definition, the concept of co-production has progressively extended to other fields and contexts, including service management and learning processes. It was central to one of the classic texts in service management (Normann, 1984), which highlighted that a key characteristic of services is the dual role of the client: both as a consumer and as part of the service delivery system. specifically, to describe the success of IKEA. The company integrated consumers into its organizational process by involving them in the final construction of its furniture, as assembly is the customer's responsibility.

Over time, co-production has been independently developed across multiple disciplines and applied in various policy and practice fields. This is because it offers several advantages, particularly in knowledge creation and capacity building, making it a key factor in the learning process.

Indeed, co-production is also a process activated in the creation of new shared knowledge and competencies (Verschuere et al., 2012). It can be understood as a practice that enables the construction of a shared setting, built on trust, where actors are encouraged to reconsider power dynamics and trust relationships, with the ultimate goal of learning new mental models and knowledge.

By reshaping relationships and being sustained by reflexive processes, co-production practices foster innovation through the continuous interplay of action and learn-

ing. The value generated by co-production stems from interaction: the active participation of diverse actors is not merely a procedural choice but a fundamental component of the process itself. Such interaction strengthens legitimacy, fosters identification, and facilitates learning by integrating different perspectives, practices, and expertise (Bandola-Gill et al., 2023).

1.3. The Relationship Between Co-Production and Trust

The power relationship between producers and users shifts in the context of co-production, as it enables the sharing of power initially held by the organization (Redman et al., 2021). As suggested by Social Exchange Theory, balancing the relationship creates the necessary conditions for a trust-building process between the organization and its stakeholders. From this perspective, several studies (Fledderus et al., 2014; Geyskens, 1998; Lusch, 1992; Oliver, 2019) define co-production as a generative process of trust. This is because stakeholders, being actively involved in service delivery or product or knowledge creation, perceive themselves as part of a well-defined relationship recognized by all participants, one grounded in reciprocity and shared values (Verschuere et al., 2012). Trust, in turn, strengthens and sustains high-quality relationships across multiple levels, from individual to organizational (Farnese et al., 2022). Therefore, co-production operates within a gradual, multi-level framework where trust is crucial in driving learning and change. Therefore, research on co-production suggests that when an organization involves

external actors in its processes, it necessitates the establishment of a new power dynamic between the parties, facilitated by trustworthiness (Greenwood & Van Buren III, 2010). Co-production facilitates the creation of a shared environment, where the participants can rethink power dynamics (Redman et al., 2021) and develop mutual trust to find new knowledge. This necessitates an initial stance of reflexivity, where participants critically examine themselves, their interactions with others, and the entire process to co-produce meaning, learn, and build a foundation of trust (Möllering, 2006).

The act of sharing information during interactions serves as a fundamental precursor to trust, as it aligns the knowledge bases of both the trustor and trustee and emphasizes the reciprocal nature of their relationship (Farnese et al., 2022). This dynamic not only enables second-order learning (Koole, 2020) but also supports the gradual co-construction of trust through continuous exchanges. As outlined in Shapiro's (1992) tripartite model, this process ultimately leads both parties to rely on the knowledge co-created through shared information as a basis for their actions and sense of self. In doing so, it fosters a strong perception of similarity, which not only helps to dismantle unequal power dynamics but also lays the groundwork for building a shared sense of identity (Kasten, 2018). Sharing information to build new knowledge and skills fosters a trust-based relationship grounded in what has been co-created. This process leads the working group to the final stage of trust, as defined by Shapiro: identification. Farnese and colleagues (2022), reviewing Schapiro's work, distinguish two key processes at play: perceived similarity and reciprocity. The first, perceived similarity, is rooted in both

cognitive and emotional foundations. Trust is strengthened when teammates recognize shared goals, aligned objectives, and common purposes. This alignment facilitates the integration of diverse roles and individual contributions, promotes shared responsibility and, crucially, fosters a shared framework for interpreting events and building new capacities

At the same time, similarity is reinforced by a sense of belonging to the group, manifested through shared behaviors (such as adopting specific methodologies or a distinctive 'style') and implicit norms (for instance, trusting colleagues' judgments or perceiving a decision as inherently fair and right). Participants also described this sense of sameness using embodied metaphors.

1.4. FRED2: A Project where Co-Production was both the Goal and the Means for Building New Capacities and Knowledge

Who?

One of the key objectives guiding the entire FRED2 project was the active involvement of stakeholders from diverse backgrounds, academics, and practitioners with varying levels of seniority and different backgrounds. This also included the participation of students from different European countries, all sharing an interest in the fight against fraud within the EU framework. This diverse participation led to the creation of a task force, a mixed European academic-practitioner group, aimed at developing a conceptual map for a future pilot study designed to detect and predict behaviors and risk profiles associated with the

misuse of European funds, both in qualitative and financial terms.

Moreover, the interaction and engagement of these different actors enabled the first steps for the establishment of an Anti-Fraud Observatory with a European transnational perspective. This initiative has strengthened the sustainability of the newly formed network by fostering relationships that enhance awareness of fraud and other illegal activities while promoting transnational and multidisciplinary cooperation.

The interaction within the group was crucial in achieving the goal.

When and where?

The spaces and moments for co-creating new meaning were conferences, webinars, and study visits. These events played a crucial role in fostering collaboration, sharing insights, and strengthening the overall co-production process. Notably, each event was hosted at the premises of different partner organizations, reinforcing the very essence of co-production.

Welcoming others into one's own institution carries powerful symbolic value: it signifies more than just the exchange of knowledge and experiences aimed at developing new skills. It also reflects the establishment of a deeper connection, marked by a shared sense of belonging and mutual trust. This level of identification becomes even more evident during conferences and study visits, where the act of physically hosting participants further embodied the trust-based relationship built among the actors involved. Thus, in this context, the study visits to some of the Universities involved, were essential because the best and worst practices related to the aspects of the

learning path have been examined to identify common and different elements based on the specific characteristics of the territory visited and realizing a collaborative study using a European and transnational approach. These events have been considered a tool for gathering information and building new knowledge.

The temporal sequence in which the various actors agreed to host the project's events offers meaningful insight. After the Italian AFCOS officially hosted the project's kick-off meeting in an institutional setting, the first event took place at Sapienza University of Rome. This choice reflected the project proponent's intention, Sapienza, to establish itself as a trustworthy actor in the eyes of the stakeholders and project partners. In this initial phase, trustworthiness rested primarily on one of its three foundational pillars: competence. The successful outcome of the event and the competencies activated to build such an ambitious collaborative project legitimized Sapienza's role in continuing to organize the scheduled activities. This also signaled that the other actors recognized Sapienza's research group as a reliable partner, worthy of engaging in a collaborative relationship not only within the university's walls but also, symbolically, within their own institutions.

1.5. Conclusions

The FRED2 project is a prototypical case that illustrates how co-production can serve as a process of knowledge creation, capacity building, and trust development. By engaging diverse actors, academics, practitioners, and students from various European contexts, FRED2 fostered a

collaborative context where expertise, perspectives, and responsibilities were shared. This active participation not only enhanced the learning process but also contributed to reshaping power dynamics, facilitating the development of mutual trust and a shared sense of identity. In a virtuous cycle, trust further fuelled the creation of new knowledge through a continuous learning process.

Through continuous interaction, joint reflexivity, and the symbolic act of hosting one another, the project demonstrated that co-production is more than a methodological choice; it is a relational process that generates value by reinforcing legitimacy, reciprocity, and identification among participants. According to Shapiro's theory, the group called a task force, reached the final stage of trust building: identification, sharing goals, and knowledge. This relationship of trust, in turn, strengthened the entire process of co-creating new knowledge and capabilities.

To summarize, FRED2 stands as a prototypical example of how co-production, when intentionally designed and implemented, fosters innovation, supports transnational cooperation, and strengthens the collective capacity to tackle complex challenges, such as preventing and detecting fraud within the European Union.

References

Bandola-Gill, J., Arthur, M., & Leng, R.I. (2023). What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives. *Evidence & Policy*, 19(2), 275-298.

Bovaird T. (2007). Beyond engagement and participation: user

and community co-production of public services. *Public Administration Review*, 67, 846-860.

Bovaird, T., & Loeffler, E. (2012). From Engagement to Co-production: The Contribution of Users and Communities to Outcomes and Public Value. *VOLUNTAS: International Journal of Voluntary and Noprofit Organizations*, 23(4), 1119-1138.

Brudney, J.L., & England, R.E. (1983). Toward a definition of the coproduction concept. *Public administration review*, 59-65.

Edvardsson, B., Tronvoll, B., & Gruber, T. (2011). Expanding understanding of service exchange and value co-creation: a social construction approach. *Journal of the academy of marketing science*, 39, 327-339.

Ewert, B., & Evers, A (2014). An ambiguous concept: on the meanings of co-production for health care users and users organizations? *VOLUNTAS: International Journal of Voluntary and Noprofit Organizations*, 25(2), 425-442.

Farnese, M.L., Benevene, P., & Barbieri, B. (2022). Learning to trust in social enterprises: The contribution of organisational culture to trust dynamics. *Journal of Trust Research*, 1-26.

Fledderus, J., Brandsen, T., & Honingh, M. (2014). Restoring trust through the co-production of public services: A theoretical elaboration. *Public Management Review*, 16(3), 424-443.

Geyskens, I., Steenkamp, J., & Kumar, N. (1998). Generalizations about trust in marketing channel relationships using meta-analysis. *International Journal of Research in Marketing*, 15, 222-248.

Greenwood, M., & Van Buren III, H.J. (2010). Trust and stakeholder theory: Trustworthiness in the organisation-stakeholder relationship. *Journal of business ethics*, 95, 425-438.

Lusch, R.F., & Vargo, S. (2006b). *The service-dominant logic of marketing: Dialog, debate, and directions*. N.Y: M.S. Sharpe.

Möllering, G. (2006). *Trust: Reason, routine, reflexivity*. Emerald Group Publishing.

Normann, R. (1984). *Service management, strategy and leadership in service businesses*. Chicester: Wiley.

Oliver, K., Kothari, A., Mays, N. (2019). The dark side of co-production: do the costs outweigh the benefits for health research? *Health Research Policy and Systems*.

Osborne, S.P., Radnor, Z., & Strokosch, K. (2016). Co-production and the co-creation of value in public services: a suitable case for treatment? *Public management review*, 18(5), 639-653.

Ostrom, E. (1999). Coping with tragedies of the commons. *Annual review of political science*, 2(1), 493-535.

Palumbo, R., Vezzosi, S., Picciolli, P., Landini, A., Annarumma, C., & Manna, R. (2018). Fostering organizational change through co-production. Insights from an Italian experience. *International Review on Public and Nonprofit Marketing*, 15(3), 371-391.

Payne, A., Storbacka, K., & Frow, P. (2008) Managing the co-creation of value. *Journal of the Academy of Marketing* Science, 36(1), 83-96.

Prahalad, C., & Ramaswamy, V. (2004). Co-creating unique value with customers. *Strategy & Leadership*, 32(3), 4-9.

Realpe, A., & Wallace, L.M. (2010). What is co-production. *London: The Health Foundation*.

Redman, S., Greenhalgh, T., Adedokun, L., Staniszewska, S., Denegri, S., Co-production of Knowledge Collection Steering Committee (2021). *Co-production of knowledge: the future. BMJ*, 372.

Shapiro, D.L., Sheppard, B. H., & Cheraskin, L. (1992). Business on a handshake. *Negotiation Journal*, 8, 365-377.

Spohrer, J., & Maglio, P. (2008). The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-Creation of Value. *Production and Operations Management*, 17(3) 238-246.

Svensson, G., & Gronroos, C. (2008). Service logic revisited: who creates value co-creates*? European Business Review*, 20(4), 298-314.

Vargo, S., & Lusch, S. (2008). Service-dominant logic: continuing the evolution. *Journal of the Academy of Marketing Science*, 36(1), 1-10

2. Digital Transformation and AI in Fraud Detection: Challenges and Opportunities in Subsidized Finance

by *Francesco Bellini*

2.1. Introduction

The digital era has revolutionized various aspects of governance and finance, including fraud detection sector. With an increase in financial transactions and the complexity of economic activities also frauds have evolved, necessitating advanced technological countermeasures (Baumgärtler, Eudelle, & Gallud Cano, 2024). Artificial Intelligence (AI) and Big Data analytics have emerged as critical tools in combating fraudulent activities. However, while technology presents

DOI: 10.82018/9791221163728.02

opportunities for enhanced detection and prevention, it also introduces new challenges, including ethical considerations, data privacy issues, and operational complexities. This discussion explores the scope of fraud and the misuse of European Union direct and indirect funds, the role of digital transformation in mitigating risks, and the potential roadblocks in leveraging digital technologies for fraud detection.

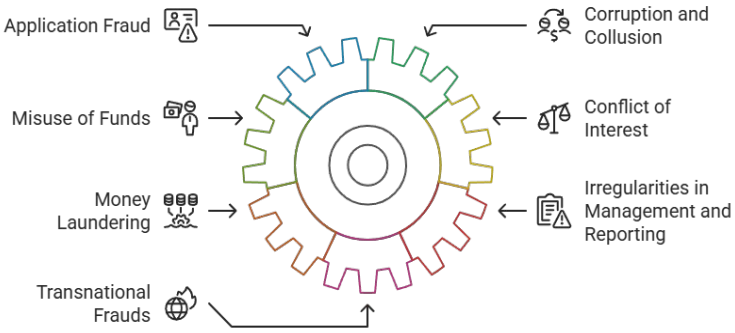
Figure 2.1. – EU Funds Frauds in the News

2.2. The Taxonomy of Fraud in Subsidized Finance

Fraud presents itself in various forms, each distinct yet interconnected, making its detection and prevention particularly challenging. Understanding the different types of fraud is essential for creating effective countermeasures. Recent criminal investigations have evidenced complex and technologically advanced fraud schemes.

The following picture summarizes (not exhaustively) the different categories of financial frauds (Presidenza del Consiglio dei Ministri - Dipartimento per gli Affari Europei, 2024):

Figure 2.2. – Taxonomy of Financial Frauds in Subsidized Finance



One of the most prevalent types of fraud is Application Fraud, which involves submitting false or misleading information to obtain financial benefits, such as EU funds. This can include forging documents, inflating costs, or claiming funding for non-existent projects. The ability to manipulate official records poses significant challenges in fraud detection.

Closely related is Corruption and Collusion (Kállay, 2015), where individuals exploit financial systems by engaging in bribery or manipulating contracts to unfairly secure funds. In such cases, officials or decision-makers may abuse their positions to benefit themselves or their associates, making it difficult to detect fraud through traditional oversight mechanisms.

Another widespread issue is Misuse of Funds, wherein

recipients use allocated financial resources for purposes other than those intended. Often, this misappropriation is for personal enrichment or unauthorized activities, which not only distorts financial distribution but also erodes public trust in financial systems.

Conflict of Interest represents a particularly insidious form of fraud, occurring when individuals responsible for fund allocation have a direct or indirect personal stake in the decision-making process. This form of misconduct can significantly influence financial allocations, skewing resources in favor of select individuals or organizations.

Additionally, Money Laundering is a key concern, wherein criminals use financial transactions—including public funds—to legitimize illicit earnings. By integrating dirty money into legitimate financial streams, perpetrators can effectively disguise their unlawful activities, posing a challenge for regulators and enforcement agencies.

Another critical area is Irregularities in Management and Reporting. Whether intentional or due to negligence, failure to comply with financial regulations and reporting requirements can create loopholes for fraudulent activities to flourish. In many cases, fraudulent actors exploit weak oversight mechanisms to manipulate reports, making it harder to trace financial discrepancies.

Lastly, Transnational Frauds highlight the complexities of fraud schemes that extend across multiple jurisdictions. Criminal networks take advantage of differing legal systems and regulatory frameworks across countries to evade detection, making international cooperation essential in combating financial fraud effectively.

2.3. The Challenges and the Future of Fraud Detection

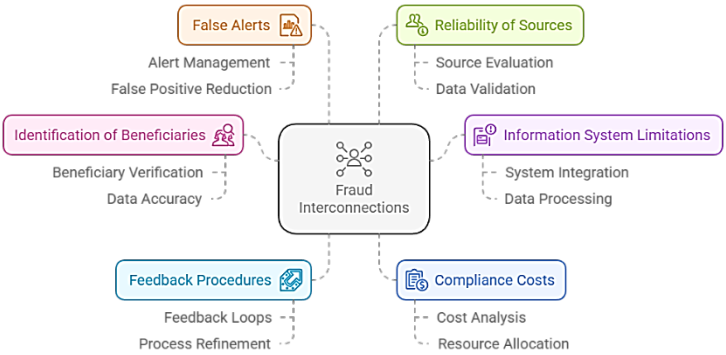
Detecting fraud is an intricate process due to the interrelated nature of fraudulent activities, limitations in existing information systems, and evolving regulatory landscapes. One of the primary challenges lies in identifying the real beneficiaries behind transactions, as fraudsters often employ complex ownership structures to obscure their involvement. Similarly, limitations in information systems hinder comprehensive monitoring, as current frameworks may not be sufficiently equipped to detect sophisticated fraudulent schemes.

Moreover, the effectiveness of feedback-based procedures is often constrained by bureaucratic inefficiencies, while compliance procedures and associated costs can be burdensome, deterring entities from engaging in proactive fraud detection measures (Nato & Bontempi, 2022). Another significant issue is the prevalence of false alerts, where AI-driven fraud detection systems sometimes generate excessive false positives, leading to wasted investigative resources. Lastly, the reliability of data sources remains a major concern, as fraudulent actors continually evolve their tactics to manipulate financial records, complicating efforts to ensure data accuracy.

As fraud techniques grow more sophisticated, the next step in combating financial fraud involves a shift from traditional information systems to decision support systems that can analyse vast streams of financial data in real time. The development of more precise fraud detection mechanisms must focus on identifying criminal organizations and fraudulent schemes before they cause substantial financial damage.

A promising approach is the implementation of multidisciplinary methodologies, integrating artificial intelligence, forensic accounting, and network analysis to examine large datasets and detect fraudulent patterns efficiently. By leveraging Big Data analytics, fraud detection can move beyond simple rule-based alerts to adaptive models capable of recognizing previously unseen fraudulent behaviours.

Figure 2.3. – Challenges in Fraud Detection



Furthermore, fostering greater collaboration among regulatory bodies, financial institutions, and law enforcement agencies will be key to ensuring fraud detection frameworks remain robust and adaptable in an ever-changing digital landscape.

2.4. Digital Technologies in Fraud Detection

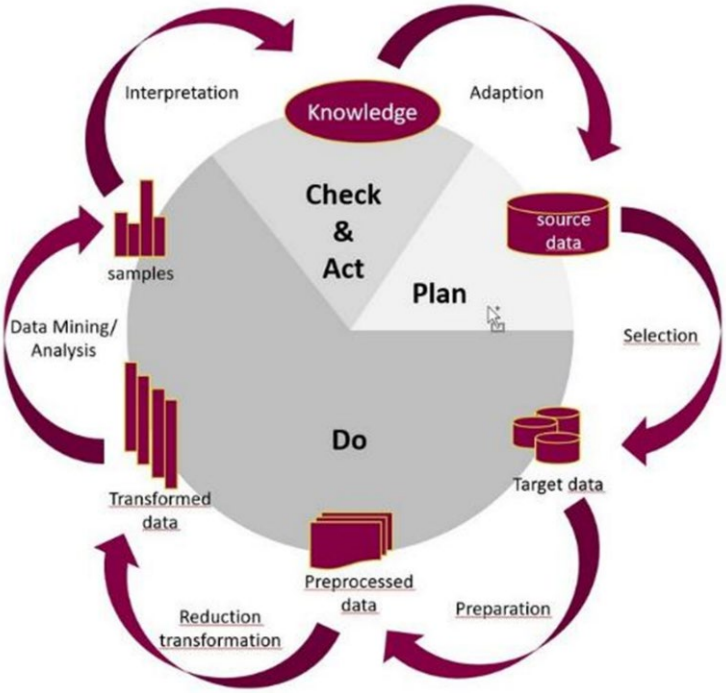
The advancement of digital technologies has significantly transformed fraud detection, equipping institutions

with powerful tools to analyse financial transactions and uncover fraudulent activities more effectively. Several innovative approaches have emerged in the fight against fraud, each contributing to enhanced transparency and security within financial systems.

Figure 2.4. – Digital Technologies for Fraud Detection

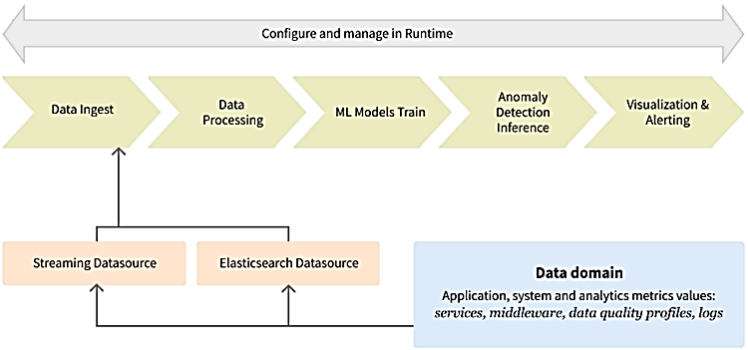
Data Analytics and Big Data (DA & BD) play a crucial role in fraud detection by enabling the analysis of vast volumes of financial transactions (Bellini, 2014). These technologies help uncover suspicious patterns and detect anomalies that would otherwise go unnoticed. By leveraging sophisticated algorithms, institutions can proactively identify fraud risks and take preventive measures before significant financial damage occurs (Ngai, Hu, Yijun, & Xin, 2011).

Figure 2.5. – Demming Cycle for Data Mining in Fraud Detection



Machine Learning and Artificial Intelligence (ML & AI) have revolutionized fraud detection by continuously learning from historical data. Unlike traditional rule-based systems, AI-driven models improve over time, recognizing emerging fraud trends and adapting their detection capabilities. By analysing complex datasets, machine learning algorithms can flag unusual transactions, reducing manual intervention and improving efficiency in identifying fraudulent activities (Ali, et al., 2022).

Figure 2.6. – Machine Learning Training Process



Furthermore, Artificial Neural Networks (ANN) make available different models for detecting anomalies selecting the most appropriate for a given scenario (Edson, Brandão, Torres Fernandes, & Alexandre, 2022).

Blockchain Technology (BC) introduces an added layer of security by enhancing transparency and traceability in financial transactions. Due to its decentralized and immutable nature, blockchain makes it significantly harder for fraudsters to manipulate financial records. Smart contracts, a feature of blockchain, automate compliance processes, ensuring that funds are used for their intended purposes, thereby minimizing the risk of fraudulent claims (Cai & Zhu, 2016).

Network Analysis (NA) is another critical tool in fraud detection, allowing investigators to map relationships between entities involved in financial transactions. By visualizing connections, network analysis can reveal hidden patterns of collusion, uncovering complex fraudulent schemes that operate across multiple organizations or jurisdictions (Pourhabibi, Ong, Kam, & Boo, 2020).

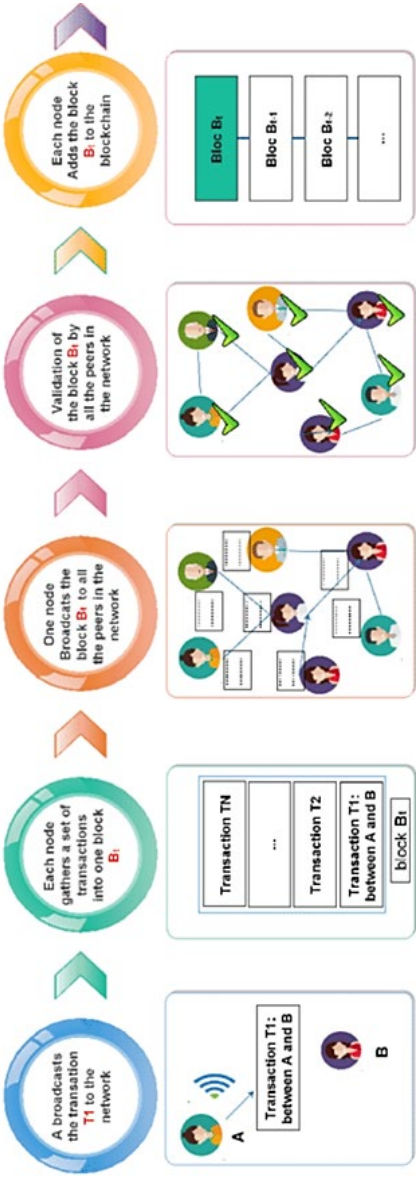
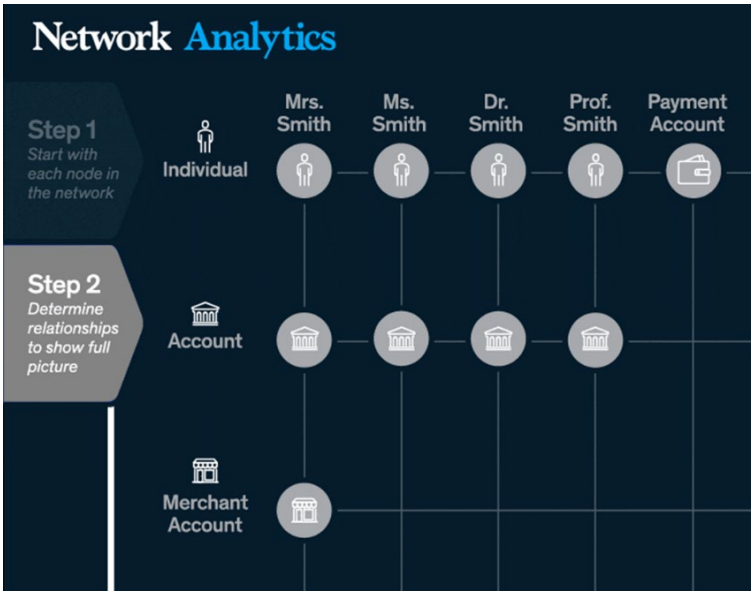
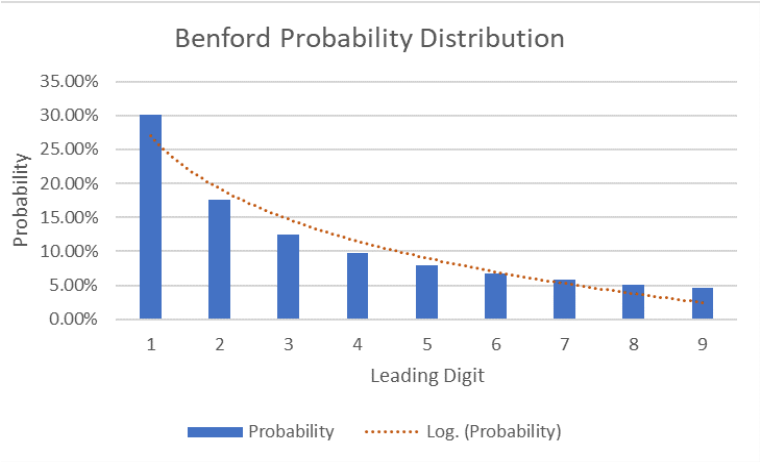

A broadcasts the transaction T1 to the network
A
Transaction T1: between A and B
B
Each node gathers a set of transactions into one block Bt
Transaction TN
...
Transaction T2
Transaction T1: between A and B
block Bt
One node Broadcats the block Bt to all the peers in the network
Validation of the block Bt by all the peers in the network
Each node Adds the block Bt to the blockchain
Bloc Bt
Bloc Bt-1
Bloc Bt-2
...

Figure 2.7. – Network Analytics Process Scheme



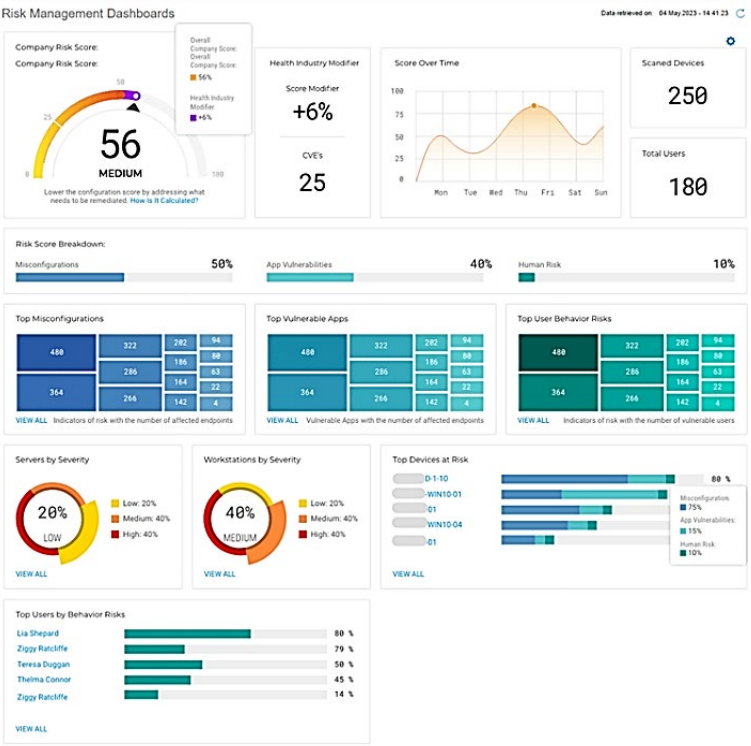
Forensic Accounting Tools (FAT) are instrumental in conducting in-depth financial investigations. These tools help auditors and investigators scrutinize financial records, identify discrepancies, and detect potential cases of embezzlement or misappropriation. Through advanced analytical methods, forensic accounting provides insights that support regulatory compliance and legal proceedings. For example, one the most effective method for detecting anomalies derives from the application of the Newcomb-Benford law. This tells that that in many real-life sets of numerical data, the leading digit is likely to have a value according to the following probability distribution; consequently, exceptions deserve to be further investigated (Barabesi, Cerasa, Cerioli, & Perrotta, 2018).

Figure 2.8. – Probability Distribution of Leading Digits Under Newcomb-Benford's Law



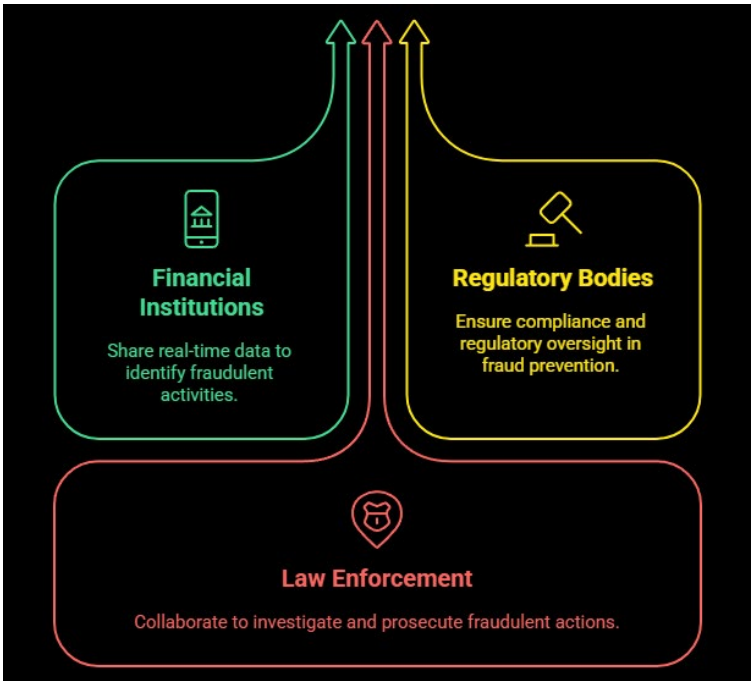
Risk Assessment Software (RAS) enhances fraud detection efforts by evaluating the risk profiles of various financial activities. These systems use predefined criteria and predictive modelling to assess the likelihood of fraud, allowing organizations to prioritize high-risk cases for further scrutiny. By streamlining the investigative process, risk assessment software ensures resources are allocated efficiently (Ilori, Tochi Nwosu, & Nwapali Ndidi Naiho, 2024).

Figure 2.9. – An Example of Risk Assessment Dashboard



IT enforced Collaborative Platforms (CP) facilitate real-time information sharing among financial institutions, regulatory bodies, and law enforcement agencies. These platforms enable cross-border cooperation, ensuring that insights into fraudulent activities are disseminated promptly. Enhanced collaboration strengthens the collective ability to detect and mitigate fraud at a global scale (European Anti-Fraud Office, 2024).

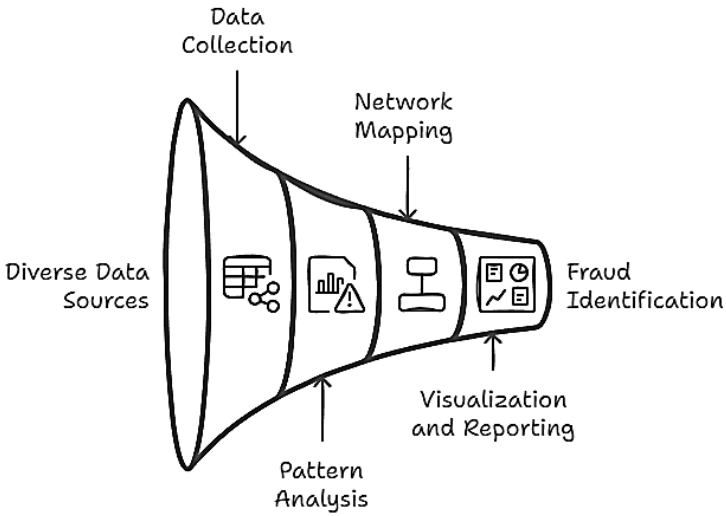
Figure 2.10. – Collaborative fraud prevention



2.5. Process for Fraud Identification

Detecting fraud requires a systematic approach that involves multiple data sources, analytical techniques, and visualization tools. By leveraging various input channels, structured methodologies, and visualization techniques, organizations can identify suspicious activities and intervene promptly (Omair & Alturki, 2020).

Figure 2.11. – Fraud Detection Funnel



2.5.1. Input Channels

Fraud detection begins with gathering data from diverse sources, including unstructured and structured datasets. Social media platforms such as Facebook and X (formerly Twitter) provide real-time insights into fraudulent trends, while business registers and open data repositories serve as essential tools for verifying company ownership and financial disclosures. Additionally, tax and banking system databases help in tracking financial transactions, while news providers and financial intelligence agencies offer critical updates on emerging fraud schemes. Collectively, these input channels provide a broad spectrum of information crucial for identifying irregular activities.

2.5.2. Study of Stylized Facts

Once data is collected, analysts study patterns of fraudulent behaviour across different dimensions, including countries, sectors, individuals, and significant financial events. The process involves defining a time window for analysing historical data and structuring information based on relevance, topic, country, or sector to identify anomalies effectively.

The outcomes of this analytical phase include generating alerts on potential frauds, mapping the ownership structures of beneficial owners, and uncovering complex networks of individuals and companies engaged in fraudulent activities. Moreover, trend analysis is performed to determine whether fraudulent cases are increasing or decreasing over time, allowing for zooming in or out on specific regions and adjusting focus based on different timeframes.

To ensure accuracy in detecting fraudulent activities, various analytical methods and tools are employed, including:

- Cross-lingual data extraction to identify fraud across multiple languages and jurisdictions.
- Multilingual semantic analysis for identifying keywords and patterns indicative of fraud.
- Supervised and knowledge-based machine learning models to improve fraud detection accuracy.
- Network analysis techniques to visualize connections and relationships between entities.
- Advanced visualization tools to simplify complex data for effective interpretation.

2.5.3. Output and Visualization

The final stage of the fraud detection process involves synthesizing findings into comprehensible formats that facilitate decision-making. This includes:

- Comprehensive reports tailored to specific topics, sectors, companies, and individuals.
- Graphical visualizations that represent fraud networks, transaction anomalies, and suspicious patterns according to predefined parameters.
- Links to original sources and supporting background material that provide additional context for investigative teams.

By integrating these structured methodologies with digital technologies, fraud detection can become more efficient and proactive, ultimately strengthening the integrity of financial systems.

2.6. Conclusions

The evolving nature of financial fraud necessitates a paradigm shift in detection and prevention strategies. Traditional oversight mechanisms, while essential, are no longer sufficient to combat increasingly sophisticated fraudulent activities. To effectively mitigate fraud risks, organizations must go beyond formal controls, adopting dynamic and adaptive measures that leverage technological advancements.

One of the most powerful tools in modern fraud detection is Big Data analytics. By aggregating and analysing

vast datasets from various sources, financial institutions and regulatory bodies can detect subtle irregularities that might indicate fraudulent activity. Machine learning models further enhance this capability by learning from historical data to identify emerging fraud patterns in real time. The increased usage of AI-driven Business Intelligence (BI) tools can provide organizations with deeper insights into financial anomalies, allowing for more proactive fraud prevention.

However, fraud detection cannot be solely reliant on technological solutions. A multidisciplinary approach that incorporates expertise from forensic accounting, legal frameworks, data science, and regulatory compliance is essential. By fostering collaboration between different fields, fraud detection strategies can become more comprehensive and effective, addressing financial misconduct from multiple angles.

Furthermore, financial institutions must recognize the importance of discovering the unknown unknowns - fraudulent tactics that have not yet been identified or classified. By continuously refining detection models and staying ahead of fraudsters' evolving techniques, organizations can build more resilient defences against financial crimes.

Ultimately, safeguarding financial systems requires a combination of advanced technology, regulatory oversight, and cross-sector collaboration. By integrating AI, Big Data, and human expertise, financial institutions can create a more secure and transparent financial environment, reducing the prevalence of fraud and reinforcing public trust in financial systems.

References

Ali, A., Abd Razak, S., Othman, S., Eisa, T., Al-Dhaqm, A., Nasser, M., ... Saif, A. (2022). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied sciences*, 12(19). doi:doi.org/10.3390/app12199637.

Barabesi, L., Cerasa, A., Cerioli, A., & Perrotta, D. (2018). Goodness-of-Fit Testing for the Newcomb-Benford Law With Application to the Detection of Customs Fraud. *Journal of Business & Economic Statistics*, 346-358. doi: https://doi.org/10.1080/07350015.2016.1172014.

Baumgärtler, T., Eudelle, P., & Gallud Cano, J. (2024). An international analysis of fraud detection in European structural and investment funds. *European Journal of International Management*, 22(2), 198-229. doi:doi.org/10.2139/ssrn.4561834.

Bellini, F. (2014). Big Data Analytics for Financial Frauds Detection. *Proceedings of the 8th Mediterranean Conference on Information Systems*, 1-10. Verona: MCIS. Retrieved from https://aisel.aisnet.org/mcis2014/21.

Cai, Y., & Zhu, D. (2016). Fraud detections for online businesses: a perspective from blockchain technology. *Financial Innovation*.

Edson, D.A., Brandão, C.L., Torres Fernandes, B., & Alexandre, M.M. (2022). A Review of Neural Networks for Anomaly Detection. *IEEE Access*, 112342-112367. doi:10.1109/ACCESS.2022.3216007.

European Anti-Fraud Office (2024). *OLAF report 2023.* Publications Office of the European Union. Retrieved from https://data.europa.eu/doi/10.2784/56951.

Ilori, O., Tochi Nwosu, N., & Nwapali Ndidi Naiho, H. (2024). Advanced data analytics in internal audits: A conceptual framework for comprehensive risk assessment and fraud detection. *Finance & Accounting Research Journal*, 6(6), 931-952. doi:10.51594/farj.v6i6.1213.

Kállay, L. (2015). *The corruption risks of EU funds in Hungary.* Budapest: Transparency International Hungary Foundation.

Nato, A., & Bontempi, V. (2022). The Protection of the EU's Financial Interests and Pandemic Emergency Tools: An Analysis of the Control Mechanism between the EU and the Member States. *Review of European Administrative Law*, 7-28. doi:doi.org/10.7590/187479822X16669633687975

Ngai, E., Hu, W.Y., Yijun, C., & Xin, S. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559-569. doi:doi.org/10.1016/j.dss.2010.08.006.

Omair, B., & Alturki, A. (2020). A Systematic Literature Review of Fraud Detection Metrics in Business Processes. *IEEE Access*, 26893-26903. doi:10.1109/ACCESS.2020.2971604.

Pourhabibi, T., Ong, K.-L., Kam, B., & Boo, Y. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*. doi:doi.org/10.1016/j.dss.2020.113303.

Presidenza del Consiglio dei Ministri - Dipartimento per gli Affari Europei. (2024). *Relazione Annuale del Comitato per la lotta contro le frodi nei confronti dell'Unione Europea.* Roma: Presidenza del Consiglio dei Ministri.

3.
EU Fraud Risk Profile Analysis: Results of the FRED2 Survey on Italian Registered Chartered Accountants

by *Tommaso Di Nardo & Antonia Coppola*

3.1. Introduction

This paper presents the findings of a survey investigating the behavior of Italian Chartered Accountants regarding fraud at the EU level, specifically in their capacity as advisors to enterprises. The study was conducted by the Fondazione Nazionale Commercialisti – Ricerca (FNC-R) in collaboration with Sapienza University of Rome as part of the FRED2 project.

The Fondazione Nazionale di Ricerca dei Commercialisti (FNC-R) is an instrumental body of the Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Conta-

DOI: 10.82018/9791221163728.03

bili (CNDCEC). The FNC-R's mission is to conduct scientific research aimed at the advancement of the accountancy profession.

In Italy, approximately 120,000 Chartered Accountants are registered. The majority of these professionals operate within organized professional firms, offering a range of services including accounting and taxation assistance, and consultancy.

The FNC-R serves as a technical-scientific partner in the FRED2 project. Within this framework, the FNC-R participated in the design and execution of a survey administered through a questionnaire to Chartered Accountants working in organized professional firms.

The primary objective of the survey was to determine the behavior of Chartered Accountants towards fraud at the EU level in their role as advisors to enterprises.

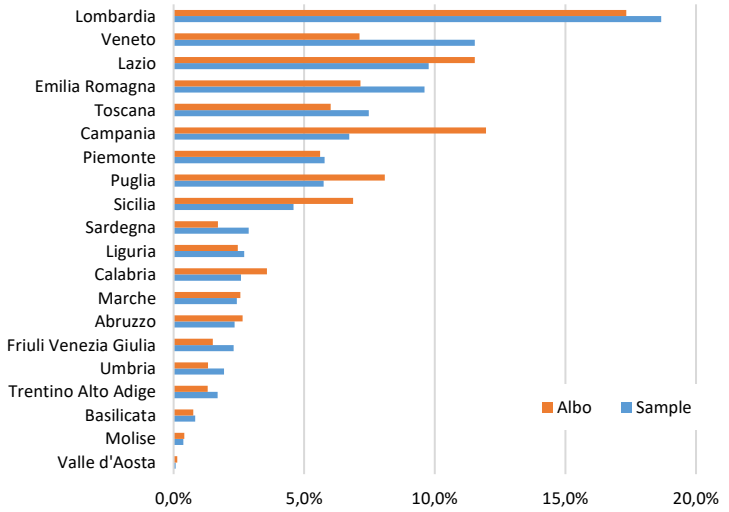
It is important to note that Chartered Accountants provide services to the majority of Italian Micro, Small, and Medium-sized Enterprises (SMEs), while a minority of these businesses are advised by non-registered individuals.

3.2. Survey Methodology and Sample Characteristics

Italian Chartered Accountants were surveyed to gain insights into their experiences and perspectives on fraud, with a specific focus on EU funding. The survey, an online questionnaire, was distributed through the FNC-R mailing list on August 1, 2024, and data collection concluded on September 30, 2024. The study yielded a sample of 3,073 Italian Chartered Accountants. A preliminary evaluation

of the sample structure was conducted to assess potential selection bias inherent in online surveys. Figure 3.1 highlights the variation in regional representation within the sample compared to the distribution of registered members ("Albo"), indicating over-representation in certain regions and under-representation in others. While this discrepancy may introduce some limitations to the generalizability of the results, the survey remains highly valuable due to the particularly sensitive nature of its central inquiries.

Figure 3.1. – Sample Shares in Relation to the Shares of Registered Members per Region



Note: "Albo" is the register of Italian Chartered Accountants.

A significant 67.7% of professional accountancy firms primarily assist micro-enterprises (0-9 employees), followed by those primarily supporting small enterprises (10-49 employees) at 27.6%. Medium-sized enterprises (50-

249 employees) are the primary client size for 4.3% of firms, and large enterprises (250+ employees) for 0.45%.

Analyzing the data by region, as shown in Table 3.1, reveals a clear territorial discrepancy between the North and South. Firms in the North are less likely to primarily assist micro-enterprises compared to firms in the Centre and South. Conversely, firms primarily assisting large, medium-sized, and small enterprises represent a higher percentage in the North.

This distribution reflects the national economic structure, with the North demonstrating a more developed economy than the South.

Table 3.1. – Percentage shares of professional firms by prevailing size (in terms of employees) of client enterprises by territorial macro-area

	NORTH	CENTRE	SOUTH	ITALY
Micro (0-9)	61.0%	75.2%	74.6%	67.6%
Small (10-49)	32.4%	22.5%	22.1%	27.6%
Medium (50-249)	5.9%	2.3%	3.0%	4.4%
Large (250+)	0.7%	0.0%	0.3%	0.5%
Total	100.0%	100.0%	100.0%	100.0%

The survey also collected detailed data pertaining to the specific characteristics of the EU funding received by client enterprises. This included an examination of the amount of funding awarded and the types of funding instruments utilized. In terms of the funding amounts, the data indicates that the majority of contributions fell below €50,000, accounting for 57.5% of cases. A smaller proportion of client enterprises received funding within the range

of €50,000 to €150,000, representing 30.4% of the recipients. Furthermore, only 12.1% of the enterprises received funding exceeding €150,000. These figures provide valuable insights into the scale of funding typically involved and can be useful for risk assessment and audit planning.

Regarding the types of funding, the survey revealed that the most common type of funding instrument was the 'non-repayable contribution', which constituted 47.2% of the funding allocations. This was followed by 'capital grants', representing 33.1%, and 'subsidized loans' accounting for 16% of the funding. The predominance of non-repayable contributions suggests that a significant portion of EU funding is provided without the expectation of repayment, which may have implications for how these funds are utilized and the potential risks of misuse.

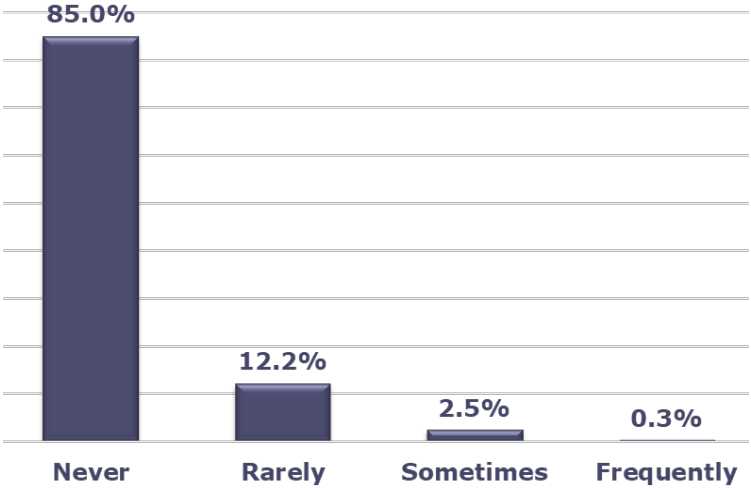
3.3. Key Findings on EU Funding and Fraud Risk

To delve into the investigation of the risk of fraud associated with EU funding mechanisms, the survey included specific questions designed to elicit responses from participants regarding their direct or indirect experiences with potentially fraudulent behavior exhibited by their client enterprises. This line of inquiry is crucial for gaining a comprehensive understanding of the challenges and vulnerabilities inherent in the distribution and utilization of EU funds.

In terms of the overall prevalence of such fraudulent behavior, the survey revealed that 15% of Chartered Accountants reported encountering instances of potentially fraudulent behavior among their clients (Figure 3.2). This

figure, while representing a minority, underscores the existence of fraudulent activities within the context of EU funding, highlighting the need for vigilance and robust control mechanisms.

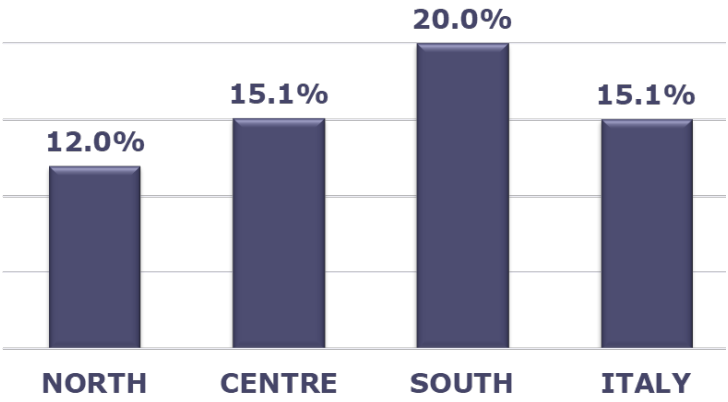
Figure 3.2. – Distribution of Chartered Accountants declaring that They have had an Indirect Knowledge of Potentially Fraudulent Behavior by Their Clients



It is important to note that the prevalence of reported fraudulent behavior was not uniform across all regions; rather, it exhibited significant variation (Figure 3.3). Notably, 20% of Chartered Accountants operating in Southern Italy reported encountering such behavior. This figure exceeds both the national average of 15.1% and the 12% reported by Chartered Accountants in the Northern regions. These regional disparities may be attributable to a complex interplay of factors, including variations in economic conditions, administrative practices, and the effectiveness

of regional oversight. Further research could explore these regional differences in greater depth to identify the underlying causes and inform targeted interventions.

Figure 3.3. – Distribution of Respondents Claiming to have Detected Potentially Fraudulent Behavior by Their Clients by Territorial Macro-Area



Chartered Accountants who reported encountering potentially fraudulent behavior were further asked to describe the actions they undertook in response. The survey revealed that common responses included resignation from the position, with a notable proportion of respondents indicating that their primary action was to resign from their position as advisors to the client enterprise. This measure underscores the severity of the ethical conflict faced by Chartered Accountants when confronted with fraudulent activity and highlights their commitment to professional integrity. In addition, some respondents stated that they reported the identified instances of potentially fraudulent behavior to the appropriate authorities. This course of ac-

tion reflects a willingness to uphold legal and regulatory obligations and to ensure that fraudulent activities are properly investigated and addressed.

The fraudulent behavior most frequently involved over-invoicing, inflated costs, or expenses that were disproportionate to the company's economic structure. In some instances, Chartered Accountants also noted that companies did not meet the eligibility requirements for the funding they received.

3.3.1. Usefulness of Financial Ratios in Detecting Fraud

The survey investigated the potential utility of specific economic and financial indicators derived from company financial statements for the detection and reporting of fraudulent behavior. Participants, comprising Chartered Accountants, were requested to evaluate the efficacy of several financial ratios, including:

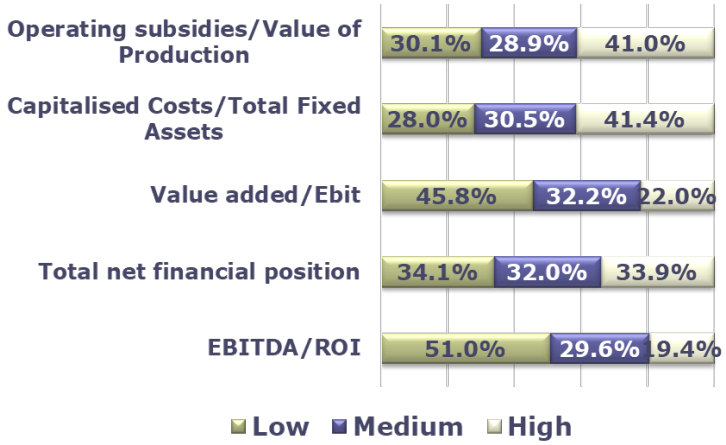
- EBITDA on ROI;
- Total net financial position;
- Value added on EBIT;
- Capitalized costs on total fixed assets;
- Operating subsidies on production value.

Participants assessed the informative potential of each ratio. The evaluation criteria, predicated on the informative potential of each individual indicator in signaling possible fraudulent behavior, comprised the following scale: Not at all, Very little, A little, Quite a lot, Much, Very much, and Completely.

The two ratios deemed most salient for indicating potential fraudulent behavior were 'Operating subsidies on

production value' and 'Capitalized costs on Total fixed assets' (Figure 3.4).

Figure 3.4. – Assessment of the Informative Potential of Selected Financial Ratios in Signaling Potential Fraudulent Behavior: Low = Not at All + Very Little + Little Medium = Quite a Lot and High = Much + Very Much + Completely



3.4. Conclusion

The survey has yielded valuable insights into the perceptions and experiences of Chartered Accountants concerning fraud, with a particular focus on the complexities arising within the context of EU funding. The study's findings illuminate several key areas, including notable regional disparities in the reported prevalence of fraudulent activities, the diverse nature of fraudulent behavior encountered by Chartered Accountants in their professional practice, and the varied actions taken by Chartered Ac-

countants when confronted with such ethical dilemmas. Furthermore, the research provides empirical evidence on the utility of specific financial ratios, derived from company balance sheets, as potential indicators for the detection of fraudulent practices. Chartered Accountants participating in the survey identified 'Operating subsidies on production value' and 'Capitalized costs on Total fixed assets' as particularly informative in this regard.

These findings carry significant implications for various stakeholders. For accounting professionals, the survey underscores the importance of heightened awareness of fraud risks, especially within the context of EU funding schemes, and highlights the need for robust ethical frameworks and decision-making processes when encountering potential misconduct. For businesses, the study emphasizes the necessity of implementing strong internal controls and governance mechanisms to mitigate the risk of fraud and maintain financial integrity. Policymakers and regulatory bodies can leverage these insights to refine existing regulations and develop more effective strategies for fraud prevention and detection, particularly in the allocation and oversight of EU funds.

It is crucial to acknowledge the limitations of the study, primarily those related to potential biases in sample representation, which may affect the generalizability of the findings to the entire population of Chartered Accountants. Therefore, the interpretation of the survey results should be undertaken with due consideration of these limitations.

4.
Pilot Research Study

by *Maria Felice Arezzo, Francesca Iandolo, Roy Cerqueti, Domenico Vitale and Giuseppina Guagnano*

4.1. Introduction

Fraud and corruption in managing financial resources, particularly those allocated by the European Union (EU), pose significant challenges to maintaining transparency and accountability in public and private institutions. The FRED2 project—Fraud Repression through Education2—addresses these pressing issues through a data-driven approach to combating fraud in EU fund management. This initiative combines theoretical frameworks with practical applications, offering valuable insights into fraud prevention and

DOI: 10.82018/9791221163728.04

detection, the importance of identifying fraud indicators, and the development of predictive tools for future applications.

Fraud within the context of EU funds typically involves deliberate deceit aimed at misappropriating resources intended for public benefit. Examples range from falsified financial documentation to manipulating project outputs, each representing a breach of trust and a potential impediment to achieving project goals. The FRED2 project emphasizes the critical need for robust systems to identify and manage fraud risks effectively. These systems are essential not only for preserving the integrity of EU funds but also for ensuring that financial resources reach their intended beneficiaries without compromise.

The added value of projects like FRED2 is to bring together institutions like AFCOS and academia to meet in a common ground where to share their respective expertise and knowledge. The pilot is the result of this challenging process. It is the expression of almost 20 months of meetings (workshops, study visits, informal meetings) where the task force exchanged ideas and experiences. The pilot project aims to go one step further than its initial objective, i.e. to create a map of concepts for detecting and predicting behavioural and risk profiles; in fact, the pilot project was designed to be the first core of a decision support system useful for institutions in the front line of the fight against fraud.

4.2. The Map of Concept

The first step in creating an effective decision support system is to design a map of concepts within which such a system can be built.

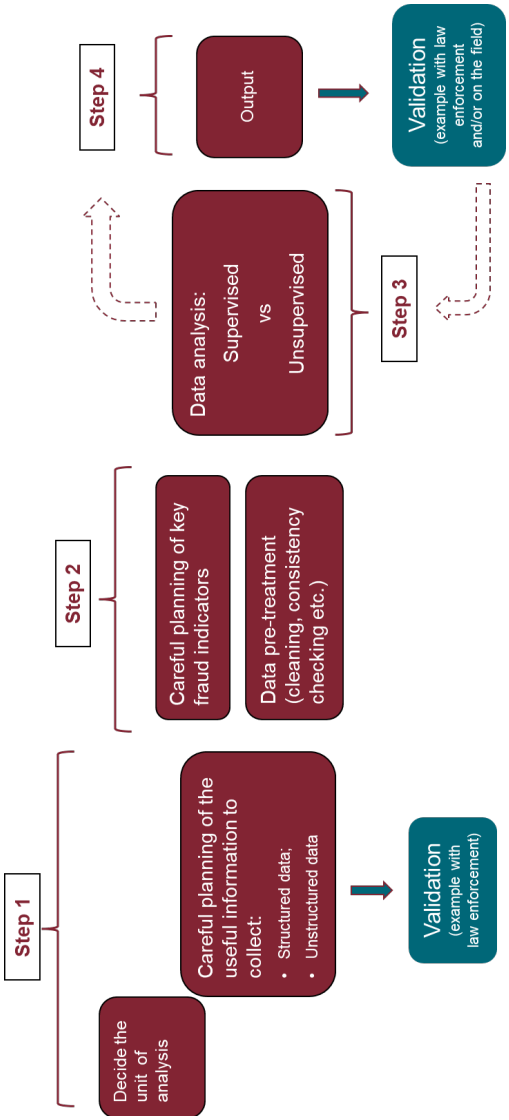
Decision Support System (DSS) is a decision-making tool where data, models, and software are used in combination with individuals to generate efficient solutions. It combines numerous data inputs and offers methodological approaches to evaluation, modelling and display of the information to facilitate decision-making. Their scope is to elaborate a vast amount of data, possibly of heterogeneous nature, to identify few patterns that can be suspicious and that should be validated by human experience, knowledge and competences.

The purpose of a DSS is to collect, analyse and synthesize data to extract relevant information possibly summarized in reports that an organization can use to assist in its decision-making process.

In Figure 4.1. we represent the map of concept useful to detect risky behaviour and propaedeutic for a DSS.

In the first step, it is necessary to decide on the unit of analysis (e.g. the funded project or the submitted project) and the relevant information to be collected. The selection of the information to be collected is arguably the most critical stage, as only relevant information should be selected. In the current era of high-capacity computing, there is a temptation to collect as much information as possible. However, this approach is not only inefficient, but also costly in terms of both financial resources and time, as it necessitates the collection and storage of the information. In addition, the use of models, whether statistical or machine learning, requires additional computing time to process the information, further exacerbating the problem.

Figure 4.1. – The map of concepts to «detect and predict behavior and risk profiles»: toward a decision support system



The second step in the process is to transform the collected information into fraud risk indicators. These indicators are typically created by intelligently combining the collected information. This process is greatly enhanced by the collaboration of data analysis experts and fraud prevention specialists. This is particularly important because within the mathematical framework in which we work, there are several types of fraud, as better explained at the beginning of section 4.3.

The third step involves data analysis and information extraction. The objective is to identify statistical units (those designated in the first step) which may be fraudulent, and which deserve further analysis. This goal is met in step four where the final output is a list of units to watch out for because they could be frauds.

This approach offers a significant advantage by increasing the efficiency of the inspection and control of EU funds. This is because the number of potentially suspect units is considerably smaller than the total number of units, so that the investigative effort would only be focused on a small subset, theoretically made up of units with some form of irregularity, if not outright fraud.

The capability of carry out a smart analysis requires highly skilled professionals. The contribution of FRED2 is, on the one hand, to highlight this need and, on the other hand, to lay the foundations (thanks to the Task Force and the Observatory to be set up) for specific training activities on the tools introduced in the pilot project.

Before going into a more detailed analysis with the pilot, we must clarify that there are three types of algorithms to detect anomalies. They are classified based on the availability of the so-called labels in the dataset. Labels

are values that testifies whether a statistical unit is a fraud. If yes, the label is generally 1, if not it is 0. Furthermore, in the statistical practice of data analysis, the labelled or semi-labelled database is divided into two subsets by a random selection of units. The first, called the training sample, is used to estimate the model and the second, called the test sample, is used to validate the model.

Box 4.1. – Philosophy and Conceptual Framework of a Decision Support System for Fraud Detection

Design and Philosophy

- *A decision support system (DSS) is an information system that supports decision-making activities.*
- *A good DSS comprises* ***Useful information (leg1) + smart analysis (leg2)*** *of the relevant information. If one leg is missing or limping, the decision-making will fail.*
    - *A lot of information does NOT mean useful information.*
    - *A built-in and unchangeable system of data analysis is not a smart way of extracting useful information.*

Conceptual Framework: *The DSS is conceptualized as a two-legged model:*

- ***Leg 1 - Information Acquisition:*** *Focus on gathering useful structured and unstructured data.*
- ***Leg 2 - Smart Analysis:*** *Employ advanced analytical methods to process the data. A lack in either leg impairs the system's effectiveness, emphasizing that mere data abundance doesn't equate to utility, and rigid analysis methods hinder adaptability and insights.*

With this premises we distinguish three types of anomaly detection algorithms:

1. Supervised anomaly detection is characterized by the presence of fully labelled training and test data sets, fol-

lowed by the training of an ordinary classifier that is then applied. Classes tend to be heavily imbalanced, but this configuration closely resembles traditional pattern recognition. Some classification algorithms may not be equally adept at handling this task.

2. Semi-supervised Anomaly Detection utilises training and test datasets, with the training data consisting exclusively of normal data devoid of anomalies. The fundamental premise is that a model of the normal class is first learned, and then anomalies can be detected by deviating from this model.
3. Unsupervised Anomaly Detection eliminates the need for labels, making it the most adaptable configuration. In addition, it does not differentiate between a training and a test dataset. The premise of this approach is that an unsupervised anomaly detection algorithm scores the data using only the "core" properties of the data. The usual method is to use distances or densities to determine what is a normal and what is an outlier, often with the help of distances or densities.

In the case of the pilot, the lack of labels for fraudulent projects meant that unsupervised algorithms had to be chosen. Furthermore, as no information on individual projects was available, information was collected at enterprise level. In other words, the statistical unit chose (step one in Figure 4.1) is the enterprise. Nevertheless, it is imperative to underline the fact that the methodology delineated below remains valid even in the event of a change in the unit of analysis.

4.3. Unsupervised Anomaly Detection Algorithms

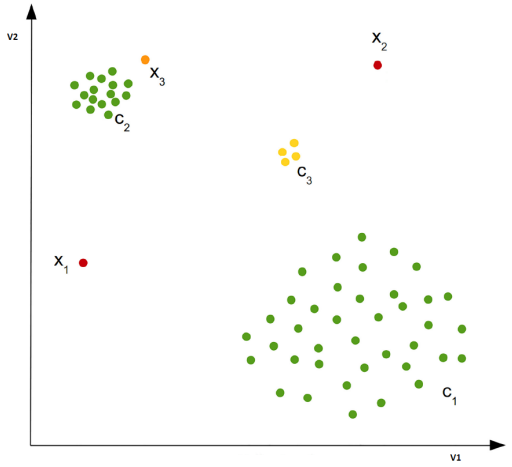
Anomalies have two important characteristics:

1. They differ from the majority of the data with respect to their characteristics;
2. They are rare compared to normal instances.

The primary objective of unsupervised anomaly detection algorithms is to identify items that deviate from the majority. Although this definition might look clear, it needs further clarifications that allow us to introduce the types of anomalies recognized in scientific literature. A graph will simplify the task.

Figure 4.2 plots points in the space identify by two characteristics (variables V1 and V2 represented in the horizontal and vertical axis respectively).

Figure 4.2. – Graphical Representation of Global Anomalies (X1, X2), Local Anomaly (X3) and Micro-Cluster (C3)



Two anomalies, x1 and x2 in Figure 4.2, can be readily identified by visual inspection. These anomalies exhibit significant deviation from the surrounding regions in terms of their attributes and are thus classified as *global* anomalies.

When the dataset is observed on a global scale, x3 can be regarded as a typical record due to its proximity to the cluster c2. However, when the focus is directed exclusively towards the c2 cluster and a comparison is made between it and x3, whilst all other instances are disregarded, an anomaly is revealed. The conclusion is that x3 is a *local* anomaly, given its anomalous nature when evaluated in comparison with its immediate neighborhood. The significance of local anomalies, and whether they are of interest or not, is a matter that is contingent on the specific application in question.

Another intriguing question pertains to the interpretation of the instances within the cluster c3: should they be regarded as three anomalies or as a (small) regular cluster? This phenomenon is referred to as a *micro cluster*, and anomaly detection algorithms are expected to assign scores to its members that are larger than the standard instances, but smaller than the obvious anomalies. This simple example illustrates that anomalies are not always obvious, and that a score is much more useful than a binary label assignment.

The task of detecting single anomalous instances in a larger dataset (as introduced so far) is termed point anomaly detection. Most unsupervised anomaly detection algorithms are of this type.

In the event of an anomalous situation being represented as a set of multiple instances, this is referred to as a collective anomaly. It is important to note that not all instanc-

es of a collective anomaly are necessarily point anomalies; rather, it is the specific combination of these instances that defines the anomaly.

A third category is that of *contextual* anomalies, which describe the effect that a point can be seen as normal, but when a given context is taken into account, the point turns out to be an anomaly.

In any case it remains possible to employ point anomaly detection algorithms to detect contextual and collective anomalies. In order to achieve this, it is possible to incorporate the context itself as a new feature. This is often not an easy task that requires strong cooperation between experts who work in the field and academics.

Collective anomalies are treated by producing a new dataset with a new set of features, by using correlation, aggregation, and grouping of the original data. A suitable basis of knowledge about the dataset is often required to move from a collective anomaly detection task to a point anomaly detection task, resulting in the production of a point anomaly detection dataset characterized by distinctive features and instances that deviate greatly from the raw data.

In the pilot for simplicity, we assumed the scenario of global anomalies.

4.4. Benford Law

In this section we will briefly outline the Benford Law. We have implemented it in the pilot project, so it's useful to give a description and some key facts that led us to consider it as an essential tool for the protection of the EU's financial interests.

In 1881, the astronomer Simon Newcomb noticed that logarithmic tables were more worn on the first pages, and found the same irregularity in every book in the library. He then published an article on the subject in the American Journal of Mathematics. However, this insight went unrecognised for some fifty years until 1938, when the physicist Frank Albert Benford coined what is now known as Benford's Law, also known as the Law of the First Digit.

The Benford Law is mathematical and statistical in nature, enabling the prediction of the frequency with which the first digit from the left is distributed in spontaneously generated number series. As it is an empirical finding, it does not enjoy the definition of a theorem but of a law, and is still being investigated by mathematicians today.

Benford's Law allows for the graphical representation of the distribution curve of the occurrences of the first significant digit in large data sets. The first significant digit of a positive number is the leftmost non-zero digit of its decimal expression. For example, the first significant digit of π is 3, that of 2371.5 is 2 and that of 0.00563 is 5. Benford's Law states that the distribution p(d) of the first significant digit d is given by:

Table 4.1. – Theoretical Distribution of the First Digits of Benford

d	1	2	3	4	5	6	7	8	9
p(d)	0.3010	0.1761	0.1249	0.0969	0.0792	0.0669	0.0580	0.0512	0.0458

The key condition for its application is the random gen-

eration of numbers, with each sample being independent of others and varied enough to span different orders of magnitude (from tens to thousands). It is also important to note that no lower or upper limits must be imposed on the data considered. Finally, the numbers must not be of an identifying nature (such as, for example, telephone numbers or bank details). Consequently, Benford's Law cannot be applied for example to ATM withdrawals because the figure is bound to the denomination of the ATM banknotes, to salaries because the figure is predetermined upstream, or to the stature of a population because it does not span several orders of magnitude. Any other kind of measurement of random origin, regardless of the scale, will respect this statistical principle.

To provide intuition on the reason why Benford's Law occurs, let us make an example. Imagine a nine-storey building with an equal number of tenants per floor and a lift. Each time the lift passes a floor, even without stopping, the LED corresponding to the floor number will light up. How frequently will the LED for each number light up? A logical conclusion would be that the frequency would be equally distributed per floor. However, in truth, number 1 will light up every time a tenant takes the lift, while number 9 will only light up when the lift is called by the ninth-floor tenants. It is evident that the switching frequency of each LED is inversely proportional to the height of the floor.

The first to recognise the potential of Benford's Law in the field of anti-fraud was Mark Nigrini, in his work 'The detection of income evasion through an analysis of digital distributions' in 1992. To illustrate the reasoning behind the intuition, let us consider the hypothetical scenario of establishing a fictitious company (referred to as a 'bad company') with the objective of incorporating the liabilities of other

'more virtuous' companies. At the end of the year, we would be required to present a balance sheet. However, as this is a fictitious company, the invoice amounts would obviously be fabricated. The distribution of the first significant figure will show serious discrepancies with the Benford distribution curve, despite the random selection of invoice amounts. The 'random' processing of numbers by humans will not follow the natural distribution curve of Benford's Law, but it will be biased toward uniformity meaning that the first digits will have roughly all the same frequency of occurrence.

Of course, compliance with the Benford curve does not necessarily imply the genuineness of the data and non-compliance is simply a suspicious indicator that needs more attention.

There are several measures to assess the compliance of a given distribution of first digits with Benford's Law. In the pilot study we used the Minimum Absolute Deviation (MAD) defined as:

$$MAD = \frac{1}{9}\sum_{d=1}^{9} |p(d) - \tilde{p}(d)|$$

where $\tilde{p}(d)$ is the theoretical distribution of Benford reported in Table 4.1. MAD is evaluated as the most reliable test for checking the validity of the Benford Law, with a value below 0.006 as close conformity, between 0.006 and 0.012 as acceptable conformity, marginally acceptable conformity for values between 0.012 and 0.015 and non-conformity otherwise [1].

[1] Nigrini M.J. Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons (2012).

4.5. A Taxonomy of Fraud Indicators

A comprehensive taxonomy of fraud indicators is a pivotal contribution of the FRED2 project, designed to enhance stakeholders' capacity to detect irregularities.

This taxonomy is structured into two main components: descriptive indicators and quantitative indicators, facilitating the identification of fraud in various organizational and financial contexts.

Descriptive indicators focus on qualitative aspects of the beneficiary entities and the nature of the funds, providing context for potential fraud risks. For instance, the typology of funds examines whether the resources are allocated as grants, subsidies, or loans, as each type carries distinct risks and monitoring requirements. Measurement and amount are assessed to identify anomalies relative to project size and industry standards, while temporal extension considers the duration over which funds are utilized, revealing inconsistencies in fund allocation and usage. A distinction between grants and lump sums helps tailor oversight processes to specific funding mechanisms.

The beneficiary company's description also falls under descriptive indicators. Elements such as the company's legal form, shareholder composition, and the country of residence of its legal representative provide insights into its structure and operational legitimacy. The presence of consortia or temporary associations can complicate accountability, necessitating a closer examination of collaborative arrangements. The company's track record, including a history of completed projects, helps assess reliability, while ongoing legal procedures and pre-existing banking relationships indicate its financial and legal stability.

Quantitative indicators, on the other hand, derive from financial statements and provide measurable data to detect potential fraud through anomalies and inconsistencies. Balance sheet-derived indicators, such as capitalization, assess the equity-to-asset ratio, which reflects financial health. Sudden increases in short-term debt or unexplained changes in equity may raise concerns about liquidity and financial manipulation. An analysis of current debt composition, receivables deterioration, and stock rotation indices reveals inefficiencies or signs of mismanagement, while profitability indicators like ROS, ROI, and ROE offer critical insights into operational efficiency and financial stability.

The application of the taxonomy is further illustrated through real-world case studies. One example involves financial statement fraud detection, where an organization receiving EU funds was suspected of inflating its financial performance to meet eligibility criteria. The investigation revealed revenue recognition issues, such as sudden increases in revenue without corresponding operational growth, and expense manipulation that understated costs to inflate profits. Another case study focused on fraud in grant allocations, where a non-profit organization misused EU funds intended for community projects. Audits revealed a lack of documentation, deviations from grant agreements, and discrepancies in reported expenditures, leading to the revocation of the grant and the implementation of corrective measures.

The FRED2 project also emphasizes the proactive management of fraud risks. Predictive modelling based on validated indicators enables the early detection of anomalies, allowing institutions to anticipate and prevent fraud. Continuous monitoring is essential, requiring regular up-

dates to fraud indicators and the integration of advanced technologies for enhanced oversight. Fostering a culture of transparency further deters fraudulent behaviour by encouraging ethical conduct across all levels of engagement.

Fraud impacts extend beyond immediate financial losses; it undermines institutional trust and disrupts operational efficiency. Collaboration among EU institutions, national authorities, auditors, and beneficiaries is vital to addressing these challenges. By sharing best practices and harmonizing regulations, stakeholders can collectively strengthen accountability and promote sustainable fund management practices.

The FRED2 project illuminates the critical role of education, collaboration, and data-driven methodologies in addressing fraud in EU funds. The comprehensive taxonomy of fraud indicators serves as a foundational tool, enabling more effective detection and prevention strategies. By leveraging innovative tools and fostering a culture of integrity, institutions can safeguard financial resources and strengthen the broader ecosystem of trust within the EU's financial framework. This ongoing effort underscores the importance of vigilance and adaptability in an era where fraud schemes continue to evolve in complexity and sophistication.

Another relevant contribution of FRED2 is that part of the indicators used for the pilot were validated by mean of a survey administered by the Fondazione Nazionale Commercialisti, a key partner in the task force and a pivotal body in the future Observatory. The questionnaire was administered during the months of August and September 2024 to all the accountants affiliated to the Fondazione. About 3,300 filled the questionnaire certify-

ing the strong interest of the Italian accountants to the frauds issued. The results were presented during the webinar held in December 2024 and can be listened by accessing the FRED2 website.

4.6. Pilot

The pilot represents the translation in terms of analysis of the learning process developed during the duration of FRED2. It integrates the methodological expertise of Academia with the practical insights and requirements of field-based Professionals. As a pilot project, it is both extendable and open to improvement. However, the pilot project's conceptual validity remains intact.

4.6.1. Step 1

The unit of analysis chosen is the firm (step 1 in the map of concept). As an illustrative example we focused on Italian building and construction companies in the period before COVID. The reason for this is that the pandemic constituted an exogenous market shock, which in turn altered the normal course of business activities. As a result, the analysis would have been strongly affected by it. As part of the same step 1 of the map in Figure 4.1, several quantities were retrieved from firms balance sheet.

4.6.2. Step 2: Key Fraud Indicators

In Step 2, the work group from Sapienza University developed a comprehensive set of 37 Key Indicators of

Fraud (KIF), specifically designed to detect discrepancies in the financial behavior of companies benefitting from public funding. These indicators were derived from intensive workshops held within the FRED2 project, where the taxonomy was presented. The framework reflects an empirical and theoretical understanding of how financial anomalies may relate to fraud risk, especially in the context of EU and national funding programs.

The indicators, validated through a national survey coordinated by the Fondazione Nazionale Commercialisti (see chapter 3 in this book for more details), are grounded in historical patterns of irregular financial conduct, drawing from balance sheet data and broader company profiles. They are grouped into three interrelated domains: equity-based (patrimonial) indicators, financial indicators, and economic performance indicators. This structure reflects the classical approach to financial statement analysis while enhancing its relevance for fraud detection purposes. The Equity-Based Indicators (Patrimonial Analysis) assess the solidity and composition of the company's asset base and capital structure, aiming to detect excessive leverage, undercapitalization, or artificial inflation of equity. The Financial Indicators (Liquidity and Solvency) explore the firm's short- and long-term financial health, with a focus on its ability to cover obligations, manage receivables, and maintain sustainable cash flows. The Economic Indicators (Profitability and Value Creation) assess profitability, value generation, and the sustainability of operational performance over time. This classification, based on the financial statement theory and practice, offers a structured, empirically tested model for identifying early warning

signs of financial misconduct. It draws from standard financial statement logic but adapts it for fraud-oriented risk profiling. When applied across multiple financial years, these indicators provide evaluators, auditors, and funding bodies with a reliable diagnostic grid.

The full list of indicators retrievable from the balance sheet/financial statement is:

1. PN / Imm (Shareholders' Equity to Fixed Assets): This ratio expresses the extent to which Shareholders' Equity (PN) covers Total Fixed Assets (Imm). It serves as an indicator of long-term capital stability, reflecting the company's capacity to support its investments with internal funds. A higher ratio suggests reduced dependence on external debt to finance long-term resources, which may be interpreted as a sign of structural robustness.
2. Imm / A (Fixed Assets to Total Assets): This indicator measures the proportion of Total Fixed Assets (Imm) relative to Total Assets (A), offering insights into the company's strategic orientation toward long-term investment. Elevated values may denote a focus on infrastructure, property, or durable capital, which can affect both financial flexibility and risk exposure.
3. Iimm / Imm (Intangible Fixed Assets to Total Fixed Assets): This ratio evaluates the share of Intangible Fixed Assets (Iimm)—such as patents, trademarks, and licenses—within the broader category of Total Fixed Assets (Imm). High values may indicate investment in innovation and knowledge-based assets, especially in sectors driven by research and development. However, such investments also carry valuation risks due to their non-physical nature.

4. CCap (Capitalised Costs): Capitalised Costs (CCap) are expenditures recognized as assets due to their long-term benefit and revenue-generating potential. While aligned with accrual accounting principles, an excessive or unjustified increase in capitalised costs may suggest an attempt to defer expenses and artificially improve profitability metrics.
5. I.CT (Assets on Behalf of Third Parties): This item reflects the value of fixed assets managed by the company (Immobilizzazioni) but legally owned by external parties. A high presence of such items may suggest operational complexity or reliance on consortia and shared infrastructure, which could introduce additional governance or audit challenges.
6. Debt / Equity (Debt to Equity Ratio): This indicator assesses the extent of financial leverage by comparing total debt to Shareholders' Equity. A high ratio reflects elevated reliance on external borrowing and may raise concerns about solvency and long-term financial risk.
7. D.b. (Short-Term Debt): This item refers to financial obligations that are due within the next 12 months. It provides insights into the company's short-term solvency and liquidity pressure.
8. D.b_DT (Short-Term Debt to Total Debt): This ratio measures the proportion of total debt (DT) comprised by short-term obligations (D.b). Higher values may indicate refinancing risks or insufficient long-term financing planning.
9. D.o (Long-Term Debt): This item includes financial obligations that mature beyond 12 months. It provides an understanding of the company's financial

structure and the long-term burden of its debt commitments.

10. D.o_DT (Long-Term Debt to Total Debt): This ratio reveals the share of long-term liabilities (D.o) within overall indebtedness (DT). A healthy proportion suggests more sustainable financial planning and better liquidity management.
11. AC / PC (Current Assets to Current Liabilities – Current Ratio): This ratio indicates the company's ability to meet short-term obligations using liquid or current resources. Values below 1 suggest liquidity challenges, while excessively high values may indicate underutilization of assets.
12. Giac.scorte (Average Inventory): This indicator shows the average stock value held during the fiscal period, providing insights into inventory management practices and potential liquidity constraints tied to slow turnover.
13. PFNt (Total Net Financial Position): This aggregate measure captures the company's overall net debt position by comparing financial liabilities with liquid assets. It is a proxy for financial independence and resilience.
14. PFNb (Short-Term Net Financial Position): This represents the net balance of short-term financial assets and liabilities, offering a near-term view of financial flexibility.
15. PFNl (Long-Term Net Financial Position): Similar to PFNb, this item measures the net position of long-term financial obligations and investments, shedding light on enduring financial commitments.
16. Sval.Cred (Impairment of Receivables): This indicator

quantifies the value of non-collectible receivables, which may suggest deteriorating client relationships or overstatement of revenues.

17. Deb.fin (Financial Debt): This item includes all interest-bearing liabilities, both current and non-current, and is a direct input in assessing leverage and interest burden.
18. Deb.comm (Trade Debt): This refers to outstanding obligations toward suppliers and trade partners. An increase may signal extended payment practices or liquidity constraints.
19. L.Corr (Current Liquidity Ratio): This liquidity metric evaluates the ability to meet current liabilities using all current assets, including inventory and receivables.
20. L.diff (Deferred Liquidity Ratio): A more conservative liquidity measure, excluding inventories from the asset base, to test how well the firm can meet obligations with liquid assets.
21. L.imm (Immediate Liquidity Ratio): The strictest liquidity test, excluding both inventory and receivables, assessing the firm's capacity to meet obligations using cash or equivalents only.
22. EBITDA / Operating Income: This ratio contrasts EBITDA with operating income (RO), revealing the extent of non-cash or non-operating effects on profitability. High divergence may mask earnings quality issues.
23. ROI (Return on Investment): ROI is calculated as Operating Income over Net Invested Capital, providing a measure of operational efficiency in generating returns on deployed resources.
24. ROS (Return on Sales): This profitability ratio com-

pares Operating Profit to Turnover (revenues), indicating the margin generated from core operations.

25. ROE (Return on Equity): ROE measures the return on shareholders' equity, serving as a barometer of managerial effectiveness in allocating and leveraging owner capital.
26. TCR (Capital Turnover Rate): This indicator evaluates how efficiently the company generates revenue from its invested capital base.
27. VA / RO (Value Added to Operating Income): This ratio compares Value Added (VA)—the firm's contribution beyond industrial cost—to Operating Income (RO), revealing the distribution of economic value and cost structure efficiency.

The key fraud indicators were constructed for all the companies belonging to a random sample of 600 units extracted from the population of building and construction companies operating in Italy in each and every year from 2017 to 2019.

The data source is the AIDA database which provides information on Italian companies and their financial, legal, ownership and management data. The access to AIDA is the result of the acquisition of a number of user licenses by Sapienza.

4.6.3. Step 3: Data Analysis

The database described earlier was analysed using unsupervised anomaly detection algorithms and later the Benford's law.

The study utilized a suite of unsupervised anomaly detection which were:

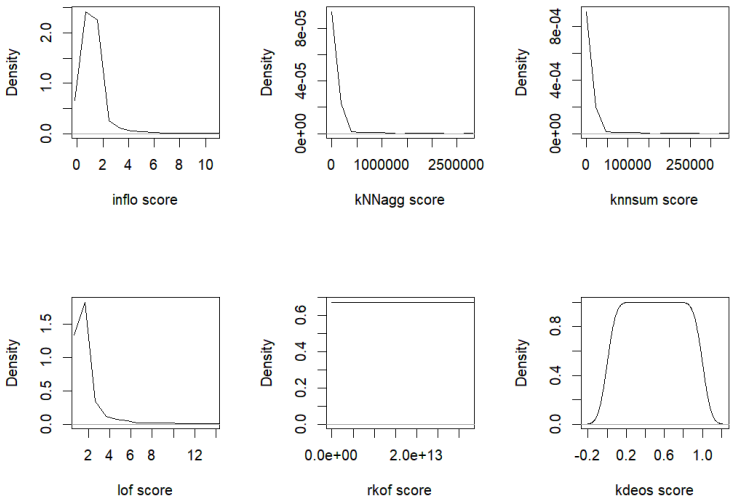
- Inflo
- kNNagg
- knnsum
- lof (Local Outlier Factor)
- rkof
- kdeos [2].

The aim was to score each firm based on the likelihood of anomalous behaviour, where a higher score indicates a higher probability of fraudulent activities. A selection of algorithms was then made based on their observed ability to discriminate abnormal from normal observations. The objective was to identify an algorithm that would assign a high score to a selected number of units. The graph [3] in Figure 4.3 helped in discarding two algorithms (k-deos and rkof).

[2] A description of the algorithms can be found in Durgesh Samariya & Amit Thakkar, 2023. "A Comprehensive Survey of Anomaly Detection Algorithms," Annals of Data Science, Springer, vol. 10(3), pages 829-850, June and references therein.

[3] Each plot is called density plot. On the horizontal axis we find the values of the variable, while on the vertical axis there is the density. The density plot is used to specify the probability of a (random) variable falling within a particular range of values. The higher the density the biggest the probability (for a unit range of values).

Figure 4.3. – Distributions of the Anomaly Scores for Each of the Six Algorithms

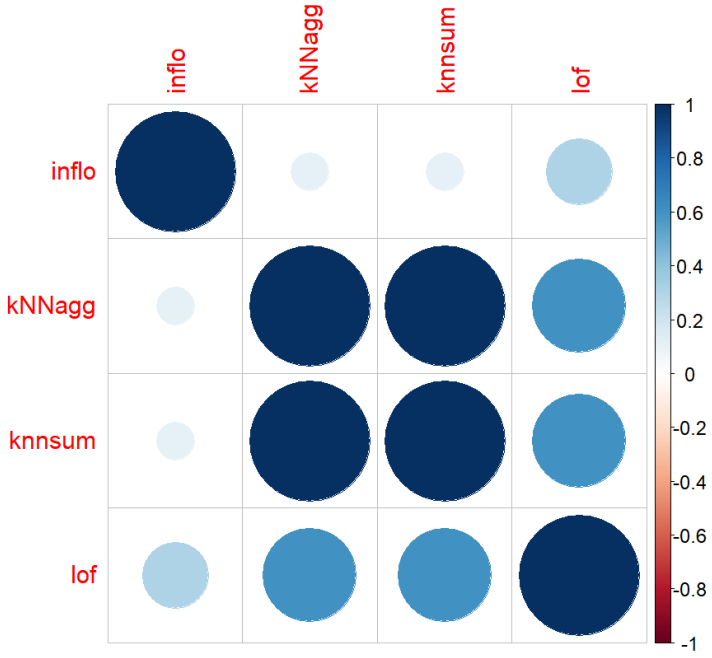


To assess the accordance of the results provided by the algorithms, we have computed the correlation matrix. Strong positive correlations indicate that the algorithms provide consistent responses.

As demonstrated in Figure 4.4, the correlation plot [4] shows that the correlation coefficients tend to be high, which lends further support to the analysis performed.

[4] The correlation plot is a graphical representation where the correlation between each pair of variables is represented through a circle. The bigger the diameter, the highest the correlation. The sign is represented using two different colours.

Figure 4.4. – Correlation plot for the scores of the four anomaly detection algorithms

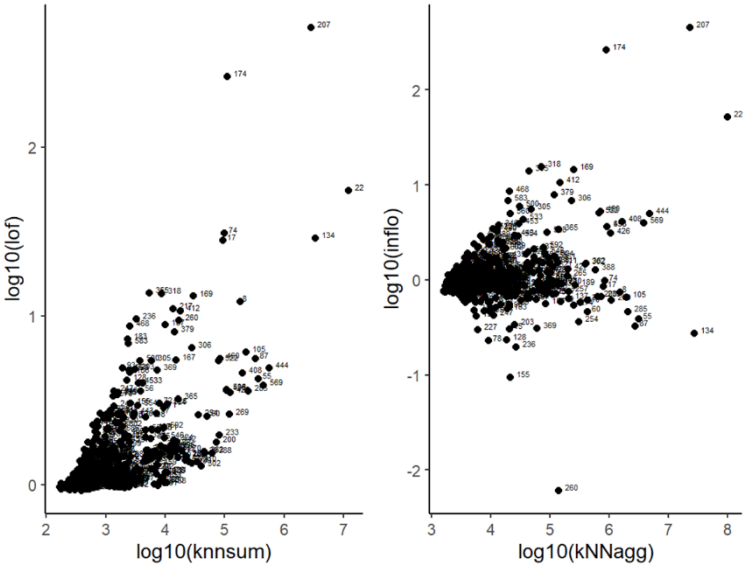


4.6.4. Step 4: Output

At this stage of the analysis, each of the 600 firms has four anomaly scores, one for each algorithm. This enables the identification of firms that deviate from the bulk of the data. As the anomaly score is calculated using key fraud indicators, the firms with anomalous scores may require further investigation.

It is the case for example of units 207, 174, 22 and, in general, of all that points in Figure 4.5 far from most of the observations.

Figure 4.5. – Scatterplot of the Building and Construction Companies According to the Four Anomaly Detection Algorithms



The successive step was to validate the analysis using Benford Law [5]. The refinement required the accomplishment of several steps:

1. We ranked the firms from the most to the less anomalous, based on the average of the four anomaly scores;
2. We extracted the first digit to the Revenues [6] of 2019 to check the conformity to Benford Law of the sample of firms with respect to this variable;

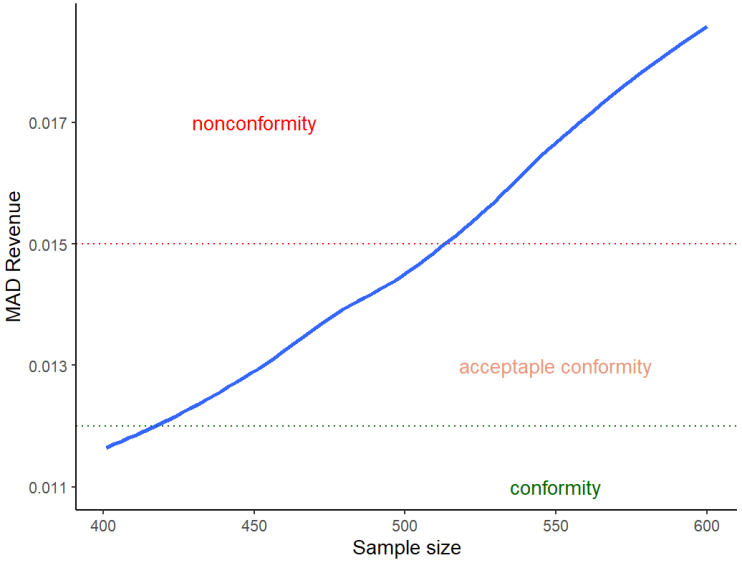
[5] In a real situation the validation process is much more complex because it would require a methodical evaluation of the findings, undertaken in collaboration with the experts working in the field.

[6] We chose revenues because, since it is the most difficult balance sheet items to alter, non-conformity to Benford Law might be due to fraudulent behaviour;

3. We computed the MAD on the overall sample, say MAD600;
4. We excluded from the data the most anomalous firm and computed the MAD for this subset of size n-1=599, say MAD599;
5. We excluded the two most anomalous firms and computed the MAD for this subset of size n-2=598, say MAD598;
6. We repeated the steps 4 and 5 excluding the first 200 most anomalous subsets and obtained a sequence of measures of conformity (MADi, with i=600,599, 598…) for samples that, based on anomaly algorithms are more and more "normal".

The results are reported in Figure 4.6.

Figure 4.6. – Trend in MAD as the Sample Size Changes due to the Exclusion of Outlier Observations



As the most anomalous observations are excluded, the mean absolute deviation (MAD) calculated to assess turnover compliance to Benford's law decreases, which indicates that the companies in the subsample are gradually becoming more compliant with Benford's Law. This interesting result suggests that the joint use of unsupervised anomaly detection algorithms and Benford Law could be a tool that help the Professionals working in the field to contrast frauds. This is because it selects the units to inspect (the most anomalous firms in our example).

References

Benford's Law

Arezzo, M.F., & Cerqueti, R. (2024). A Benford's Law view of inspections' reasonability. *Physica A: Statistical Mechanics and its Applications*, 632, 129294.

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 78, 551.

Cerqueti, R., & Lupi, C. (2021). Some new tests of conformity with Benford's law, *Stats*, 4 (3), 745-761.

Cerqueti, R., & Lupi, C. (2023). Severe testing of Benford's law, *Test*, 32 (2), 677-694.

Mir, T.A., Ausloos, M., & Cerqueti, R. (2014). Benford's law predicted digit distribution of aggregated income taxes: the surprising conformity of Italian cities and regions. *The European Physical Journal B*, 87: 261.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, 4, 39.

Nigrini, M.J. (1996). Taxpayers Compliance Application of

Benford's Law. *Journal of the American Taxation Association*, 18, 72-92.
Nigrini, M.J. (2012). *Benford's law: applications for forensic accounting, auditing and fraud detection*. Wiley Publications, New Jersey.
Nigrini, M.J., & Mittermaier, L.J. (1997). The Use of Benford's Law as an Aid in Analytical Procedures. *Auditing: a journal of practice & theory*, 16, 52.

Anomaly detection

Samariya, D., & Thakkar, A. (2023). A Comprehensive Survey of Anomaly Detection Algorithms. *Annals of Data Science*, 10(3), 829-850.
Chandola V., Banerjee A., & Kumar V. (2009). Anomaly detection: A survey. *ACM Computer Surveys*, 41(3), 1-58.
Agyemang E.F. (2024). Anomaly detection using unsupervised machine learning algorithms: A simulation study, *Scientific African*, 26.
Chalapathy R., & Chawla S. (2019). Deep Learning for Anomaly Detection: A Survey, arXiv:1901.03407.
Goldstein M., & Uchida S. (2016). A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data. *PLoS ONE*, 11(4).