MICRO DATA SCIENCE

DANMARKS NATIONALBANK

26 JANUARY 2021 — NO. 1

Data-driven effort strengthens the fight against money laundering

- Nationalbanken, in collaboration with other institutions, has assessed the potential of intensifying the usage of granular transaction data to strengthen the fight against economic crime.
- An innovative risk-based approach based on multiple data sources not only improves the identification of suspicious activities, but also identifies them faster and more efficiently.
- The realisation of the full potential of a risk-based approach requires combining data across banks and public institutions. This combination would require a revised legal framework.

Economic crime and money laundering have serious economic and social consequences, threaten the integrity of the legal system and erode trust in the financial sector.

To counteract economic crime, the Danish financial sector currently employs more than 4,300 workers for compliance and anti-money laundering (AML) purposes.[1] Their efforts result in more than 50,000 annual reports of suspected money-laundering practices to the authorities, a number that has been steadily increasing over the years. While significant resources are being invested in the fight against money laundering, the system has weaknesses that are not solvable in the current setup.

Danmarks Nationalbank, in cooperation with the Danish Business Authority, the State Prosecutor for Serious Economic and International Crime (SØIK), a major Danish bank and other government agencies, has assessed the potential of applying a data-driven approach to countering economic crime, specifically money laundering and VAT fraud, on the basis of transaction-level data.[2]

1 Finance Denmark, Anti-money laundering and counter-terrorist financing in the Danish financial sector – report by Finance Denmark's Anti-Money Laundering Task Force, November 2019 (*link*).

2 The combination of information from the participating bank and public institutions has been performed through a dispensation by the Ministry of Industry, Business and Financial Affairs. See Box 2 at page 7 for further details.

MICRO DATA SCIENCE

Data in new ways

Data volumes have grown exponentially. By 2025, an estimated 450 exabytes of data will be created each day.

This is equivalent to hundreds of millions of personal computers being filled with data on a daily basis. The vast volumes of data are highly diverse, but new and sophisticated methods enable analysis of this data in new and more efficient ways.

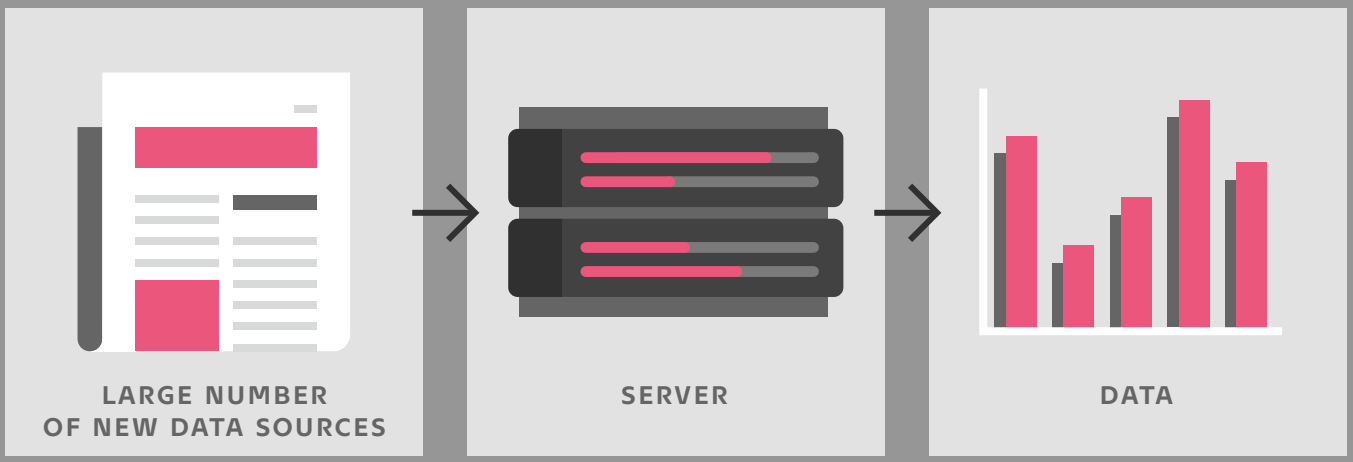
New data types and new data collection methods may be used in various contexts in Danmarks Nationalbank's ongoing work.

In order to acquire more knowledge and a better basis for assessing the Danish economy, Danmarks Nationalbank focuses on new data types and methods in a series of publications.

ABOUT THIS ANALYSIS

Danmarks Nationalbank has collaborated with other organisations to investigate how to strengthen the effort against economic crime. The suggested approach builds upon advanced analytical tools and multiple data sources, among which granular transaction data. This advanced risk-based approach detects more suspicious transactions than today, and more effectively.

New data creates new knowledge



This analysis presents Danmarks Nationalbank's conclusions on the project.

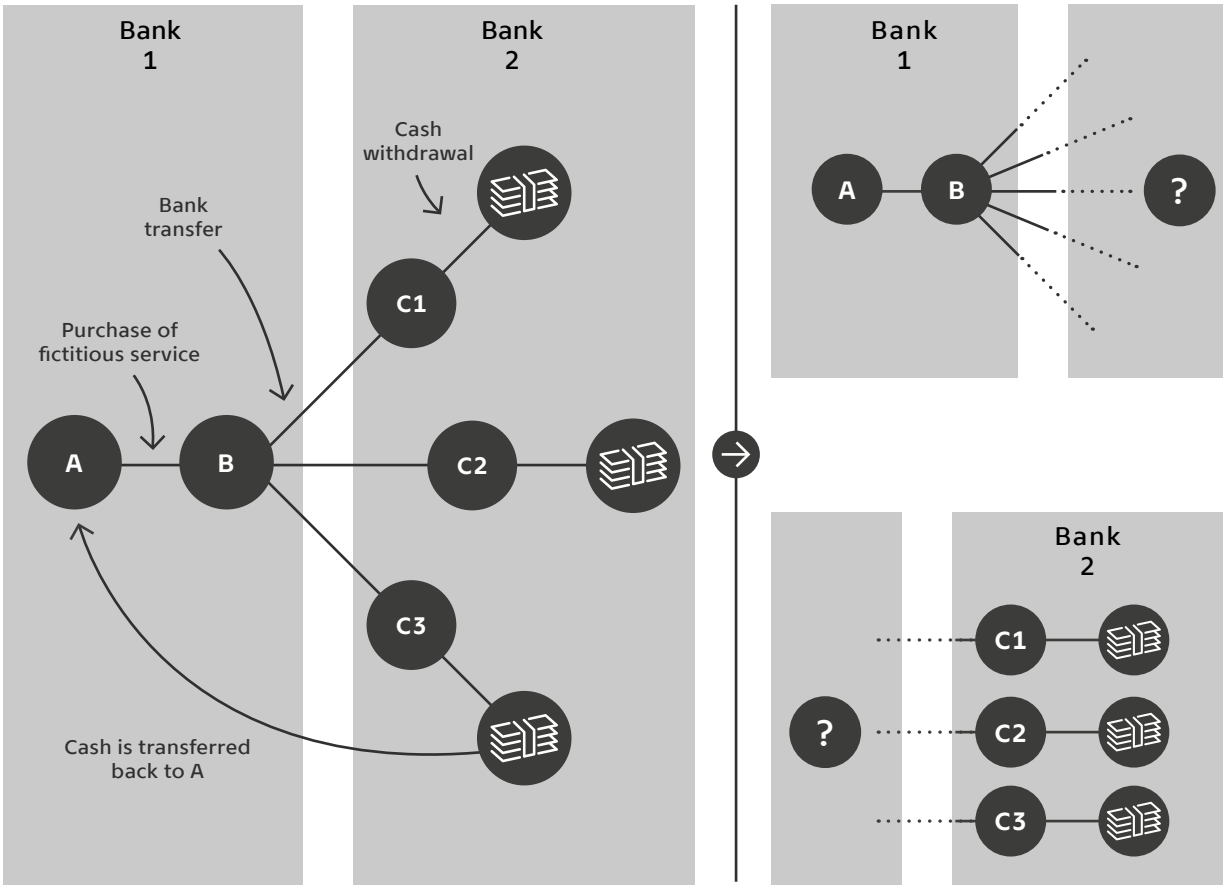
Weaknesses of current anti-money laundering system

Currently, most banks employ automated trigger systems that flag suspicious transactions based on predefined scenarios. For example, such a system might flag the withdrawal of large amounts of cash within a short period of time. While these scenarios might take into consideration basic client characteristics, e.g. whether a client is a natural person or a legal entity, they are not flexible enough to precisely detect high-risk client behaviour.

This system has two main weaknesses. First, it is very labour-intensive. Second, it relies exclusively on internal data sources within each bank.

The inspection is extremely labour-intensive both for the banks and for the public institutions. The predefined scenarios generate a large number of false alarms, which AML investigators in the individual banks must manually discard. Alarms that, after manual investigation, are deemed suspicious are sent to the Financial Intelligence Unit (FIU) at the State Prosecutor for Serious Economic and International Crime, which performs a further layer of

Cash factories are hard to detect using data from a single bank Chart 1



Note: A needs cash. To obtain it, A purchases a fictitious service from player B and transfers a corresponding amount. B then uses a series of accomplices (C1, C2 and C3) to withdraw small fractions of the transferred amount in cash. After keeping a cut, the cash is collected and sent back to A. Bank 1 cannot see that the transferred amounts are withdrawn in cash. Bank 2 cannot see the connection between the many withdrawals and the single payment source. A similar system has been applied by the money-laundering scheme behind "Operation Greed".

controls and eventually forwards them to the competent authorities (e.g. the police) for further investigation. Only about 5 per cent of all alarms raised by the automated system of a large Danish bank in 2019 were ultimately included in a report sent for further investigation in the same year.[3]

The large number of false alarms implies that the current system is not geared to handling the increasingly large volumes of transaction data that need to be analysed, nor the resulting higher number of suspicious transactions that need to be monitored and possibly investigated.

The second weakness is that the existing automated detection systems rely exclusively on internal bank data. Banks have only detailed data on their own clients and are therefore unable to assess the risk associated with the counterpart in a specific transaction, unless such counterpart happens to be a client of the same bank. Because of this limitation, complex money-laundering chains involving networks of multiple banks and accounts are unlikely to be flagged by the automated monitoring systems (see chart 1). Similarly, banks do not have access to detailed information from the authorities. Therefore, the current system is particularly vulnerable when criminals use multiple banks to blur their tracks. As a consequence, many undetected money-laundering transactions currently fly under the radar.

An innovative data-driven approach can improve detection, efficiency and speed of economic crime detection

Nationalbanken, in cooperation with other institutions, has investigated and assessed whether the weaknesses of the current system can be mitigated through an innovative data-driven and risk-based approach that exploits enriched granular transaction data.[4] Box 1 provides additional details on the data and approaches used in the analysis.

Based on multiple independent approaches, the analysis estimates that twice as many cases that would have eventually been sent for further investigation by the FIU could have been detected.

The project also shows that a data-driven approach can optimise the current manual investigation pro-cess. Currently, the automated system triggers a large number of false alarms, which need to be manually discarded by AML investigators. An innovative risk-based and data-driven approach allows a risk score to be estimated for each generated alarm. Equipped with these estimates, a risk-based system could automatically discard low-risk cases that are routinely not included in reports sent to the FIU, and as such considered false alarms (see chart 2).

When sorting alarms triggered by the current system according to their estimated risk score, only 1 per cent of all cases being sent for further investigation originated from the bottom 27 per cent of alarms with the lowest imputed risk score. In other words, one could discard 27 per cent of all alarms that are currently manually investigated, and still retain 99 per cent of cases that are sent for further investigation, as shown in chart 2.

This result shows that both banks and authorities are using a lot of resources to examine low-risk alarms which ultimately prove to have no relation to or little risk of economic crime. These resources could instead be redirected to the investigation of transactions with a high estimated risk score that do not trigger an alarm in the current system, improving the efficiency of manual investigations.

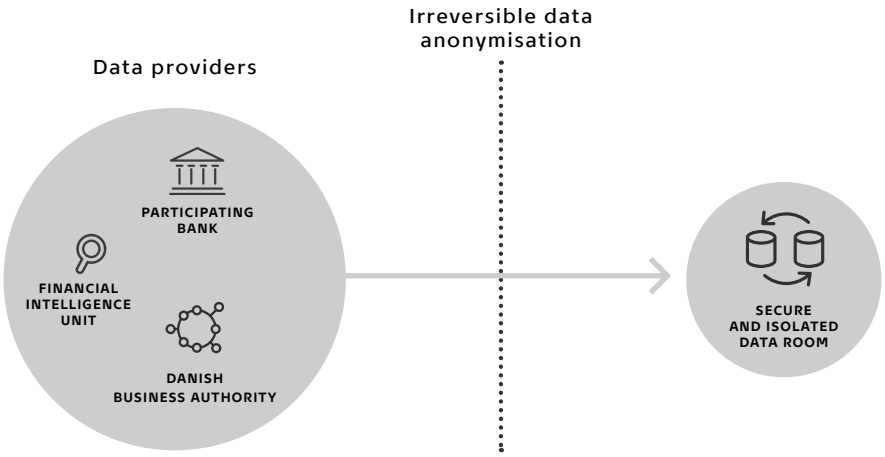
The innovative data-driven approach can detect 84 per cent of the alarms that were reported to the FIU earlier than the current system. More than half of these clients would be detected more than ten weeks before they trigger an alarm in the current system.

3 The participating bank is representative of the sector both in terms of the number of cases sent to the FIU relative to the number of clients, and in terms of the proportion of cases that end up being sent for further investigation.

4 This type of data contains detailed information on every single electronic payment taking place in Denmark, e.g. cash operations at ATMs, MobilePay transactions, domestic and international credit transfers and card payments. At the national level, the data would cover about 3.6 billion transactions in 2019.

Data and approaches used in the analysis: The data from the participating bank were enriched with information from multiple authorities, anonymised and analysed in a secure data room.

Box 1



The analysis was based on detailed transaction data from a major bank, which contained detailed information on more than 1 billion transactions over a period of approximately four years. The transaction data were combined with information from the following institutions.

- The participating bank: Information on all alarms triggered by the current system.
- State Prosecutor for Serious Economic and International Crime: Information on the reports sent to the FIU by the participating bank, including whether these were sent to further investigation.
- Danish Business Authority: Information on all firms established in Denmark organised in a graph database, which includes both information on the connection across firms and major stakeholders, and information on VAT fraud investigations by the Danish Tax Agency.

The data were anonymised and analysed in a secure data room by analysts with security clearance. Data and hard drives were destroyed after completion of the analytical work.

The analysis employed a combination of two approaches:

- A model-based approach, which consists in applying advanced analytical tools and machine learning to predict the likelihood of a bank client to be eventually reported for further investigation by the FIU. This estimation is based on recent and past transaction history coupled with background information on the bank's clients, mapped against what has earlier been considered suspicious by the authorities.[1]
- A scenario-based approach consists in combining transaction data from a single bank with other data sources to construct scenarios, designed in collaboration with AML experts, flagging suspicious behaviour that ought to be investigated.

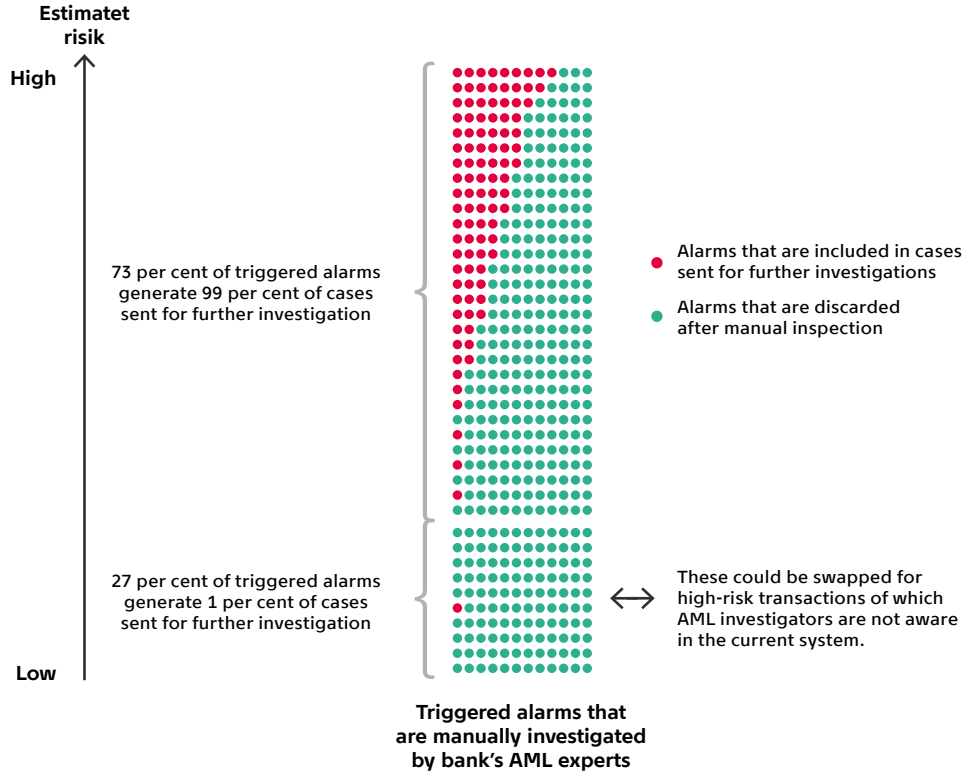
The two approaches complement each other. A model-based approach employs data to recognise transaction patterns and detects high-risk cases, triggering an alarm that AML investigators should then investigate manually. However, a model is dependent on past history and can only learn patterns of behaviour that have been detected and processed in the past. Therefore, a model will be unable to discover new patterns of money laundering by itself.

A scenario-based approach allows new scenarios to be developed to detect money-laundering patterns in collaboration with AML experts. While it can be refined by patterns detected through a model-based approach, it allows AML experts to actively contribute to the investigation. The scenario-based approach is very similar to the current automated flagging system: The additional value added is provided by the innovative combination of multiple data sources (banks and public institutions) and, in particular, by developing scenarios that utilise data sources across organisations, e.g. when investigating chains of money-laundering networks as described in chart 1.

1. Machine learning models are ideal for recognising suspicious behaviour. These models do not need to be overly complex and uninterpretable. While model performance increased with model complexity, at least initially, simple regression scoring models also worked satisfactorily. A combined approach would be to let the complex model inform an analysis of which variables and patterns the model is using to formulate predictions. The analyst can then incorporate these insights into a simple and more interpretable model..

Capacity can be freed up by discarding low-risk cases in order to focus on the most risky ones

Chart 2



Note.: A model-based approach allows the ranking of flagged cases according to the estimated risk given the outcome of past reports. When resources are scarce, automatically discarding very low-risk cases that are extremely likely to constitute false alarms allows the examination of other cases with a higher estimated risk, or the testing of novel scenarios designed by AML experts.

Realising full potential requires combining transaction and client data from multiple sources

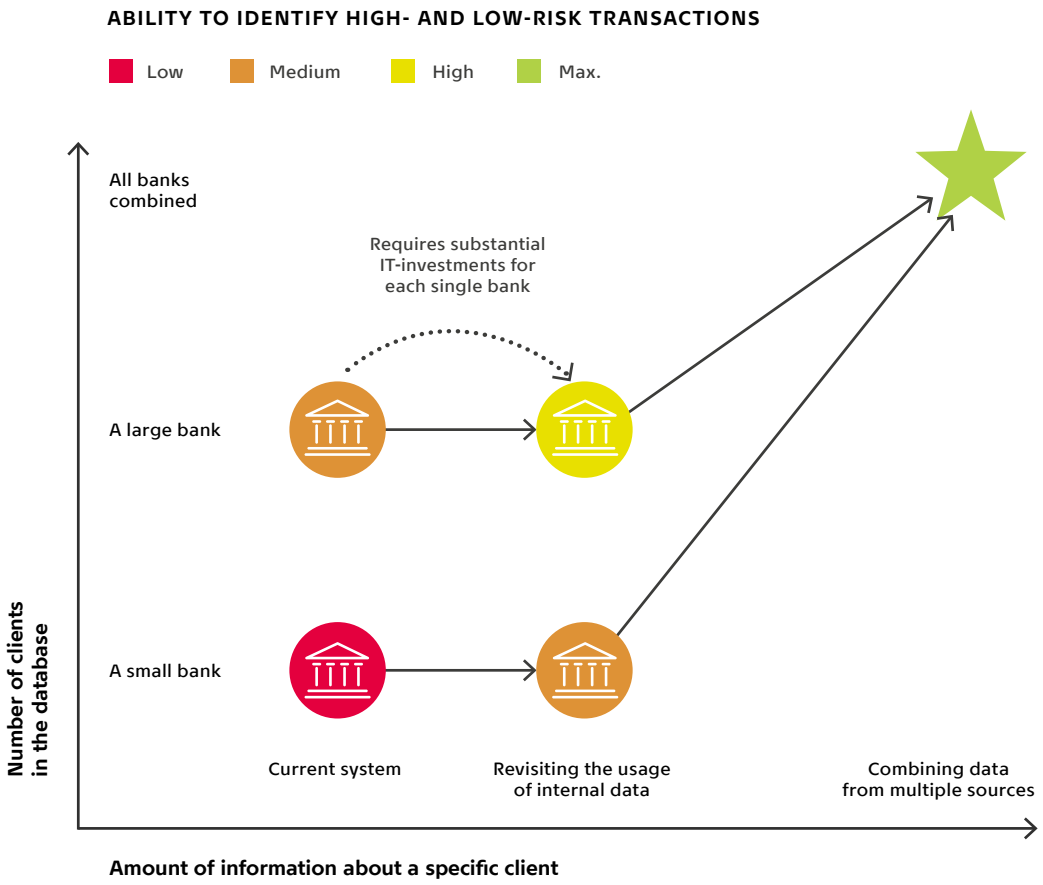
The analysis has shown that considerable resources are spent on false alarms, that many suspicious transactions fly under the radar, and that those detected are often detected with a delay. A common system can improve the joint efforts against economic crime, but it requires combining data across multiple banks and public institutions for three reasons.

First, in most cases banks do not have detailed information about the counterpart of a transaction and are therefore unable to properly evaluate its riskiness. This limitation constitutes a large obstacle in the detection of complex networks spanning multiple banks (see chart 1). Moreover, information is not shared across banks, and criminals can therefore erase their history of suspicious behaviour simply by changing bank. More than half of the reports submitted by the participating bank involve one or more players that are known to the authorities from other reporting sources.[5] In contrast, a single platform for collecting

5 This result refers to a supplementary analysis conducted in collaboration with the FIU at SØIK. The analysis also shows that 13 per cent of the players directly mentioned in the reports by the participating bank were also directly reported by other reporting institutions in the same period.

The more data are available, the better economic crime can be identified

Chart 3



Note: The data volume can be defined both in terms of the amount of information available about a single client, and the number of clients in the database. The project shows that the ability of an anti-money laundering system to detect suspicious transactions increases as the available data volume grows.

and processing data from multiple banks would allow the financial sector to share information in order to more effectively counteract criminal behaviour.

Second, joining data across banks creates synergies that are impossible to achieve for a single bank. The project shows that the potential gains from a data-driven approach increase with the volumes of data available. Specifically, the gains improve both in terms of client numbers in the database and in terms of the amount of information available about a single client (see chart 3). Combining data across banks will increase both the number of clients in the database and the amount of information available about each client, and their relations with other clients across banks. This type of information is particularly useful for identifying criminal networks.

Third, small banks have a considerable disadvantage since they do not have enough internal bank data to learn to recognise suspicious patterns in their transactions. The discrepancy between small and large banks risks amplifying structural weaknesses in the fight against economic crime, as a disparity of resources risks creating safer havens for criminals where the probability of detection is smaller.

Transaction data can also support the fight against VAT fraud

The International Monetary Fund estimates that VAT fraud results in a revenue loss of over DKK 10 billion

for the Danish state budget every year.[6] The project shows that granular transaction data can also improve the detection of VAT fraud by imputing the amount of VAT that firms should pay given their transaction history. By comparing firms' reported VAT with the VAT imputed on the basis of their transaction history, tax authorities would be able to better target their controls at firms with large discrepancies, or at firms that are not registered for VAT, but ought to be given their income flow.

Public institutions and financial sector to evaluate the possibility of joint solution for countering money laundering and economic crime

The project has shown that an innovative data-driven approach based on granular transaction data can greatly improve the fight against economic crime. Nonetheless, realising the full potential of such an approach requires collaboration between public institutions and the financial sector to collect and process data from multiple institutions in a single platform.

The project has consciously chosen not to focus on how a common system would work in practice, the IT infrastructure required, and the need for a revised legal framework that would make such a solution feasible and effective. These crucial steps and what they would entail in terms of resources and legal framework will be addressed by a working group led by the Ministry of Industry, Business and Financial Affairs composed of representatives from both public institutions and the financial sector.

Background Box 2

The project originates partly from Nationalbanken's ongoing work with the collection and production of payment statistics, and partly from an increased focus by public institutions on anti-money laundering and VAT fraud. Furthermore, Finance Denmark's Anti-Money Laundering Task Force presented its 25 recommendations in December 2018, which included the financial sector's proposals for enhanced collaboration in the field of IT and data handling to combat economic crime. Several high-profile economic crime cases led to a joint vision by politician and bankers that Denmark should be at the forefront in the fight against economic crime.

The aim of the project was to identify whether a public-private partnership could improve the detection and prevention of money laundering and VAT fraud by combining transaction data and client background information such as civil registration numbers and business registration numbers, as well as information on business purpose and scope, with data from the relevant public institutions. The data comprise client background information and transaction data from a major Danish bank, information on the outcome of the FIU's assessment of money-laundering reports, and VAT and business information not available to the public, see box 1.

6 IMF, Country report no. 16/59 – *Technical assistance report – Revenue Administration Gap Analysis Program – The Value Added Tax Gap*, February 2016.

PUBLICATIONS

NEWS

News offers a quick and accessible insight into an Analysis, an Economic Memo, a Working Paper or a Report from Danmarks Nationalbank. News is published continuously.

ANALYSIS

Analysis from Danmarks Nationalbank focuses on economic and financial matter. Some of the analyses are published with a regular frequency e.g. *Outlook for the Danish economy and Financial stability*. Other analyses are published continuously.

REPORT

Report comprises recurring reports and reviews of the functioning of Danmarks Nationalbank. For instance Report includes the *Annual report* and the annual publication *Danish government borrowing and debt*.

ECONOMIC MEMO

Economic Memo is a cross between Analysis and Working Paper and it often shows the ongoing study of the authors. The publication series is primarily targeted at professionals. Economic Memo is published continuously.

WORKING PAPER

Working Paper presents research projects by economists in Danmarks Nationalbank and their associates. The series is primarily targeted at professionals and people with an interest for academia. Working Paper is published continuously.

The analysis consists of a Danish, Greenlandic and an English version. In case of doubt regarding the correctness of the translation the Danish version is considered to be binding.

DANMARKS NATIONALBANK
LANGELINIE ALLÉ 47
DK-2100 COPENHAGEN Ø
WWW.NATIONALBANKEN.DK

Thais Lærkholm Jensen
Head of Data Analytics and Science
tlj@nationalbanken.dk

Bjarke Mørch Mønsted
Data Scientist
bmm@nationalbanken.dk

Alessandro Tang-Andersen Martinello
Senior Data Scientist
alem@nationalbanken.dk
STATISTICS

DANMARKS NATIONALBANK

CONTACT

Teis Hald Jensen
Communications and Press Officer
tehj@nationalbanken.dk
+45 3363 6066
SECRETARIAT AND COMMUNICATIONS