

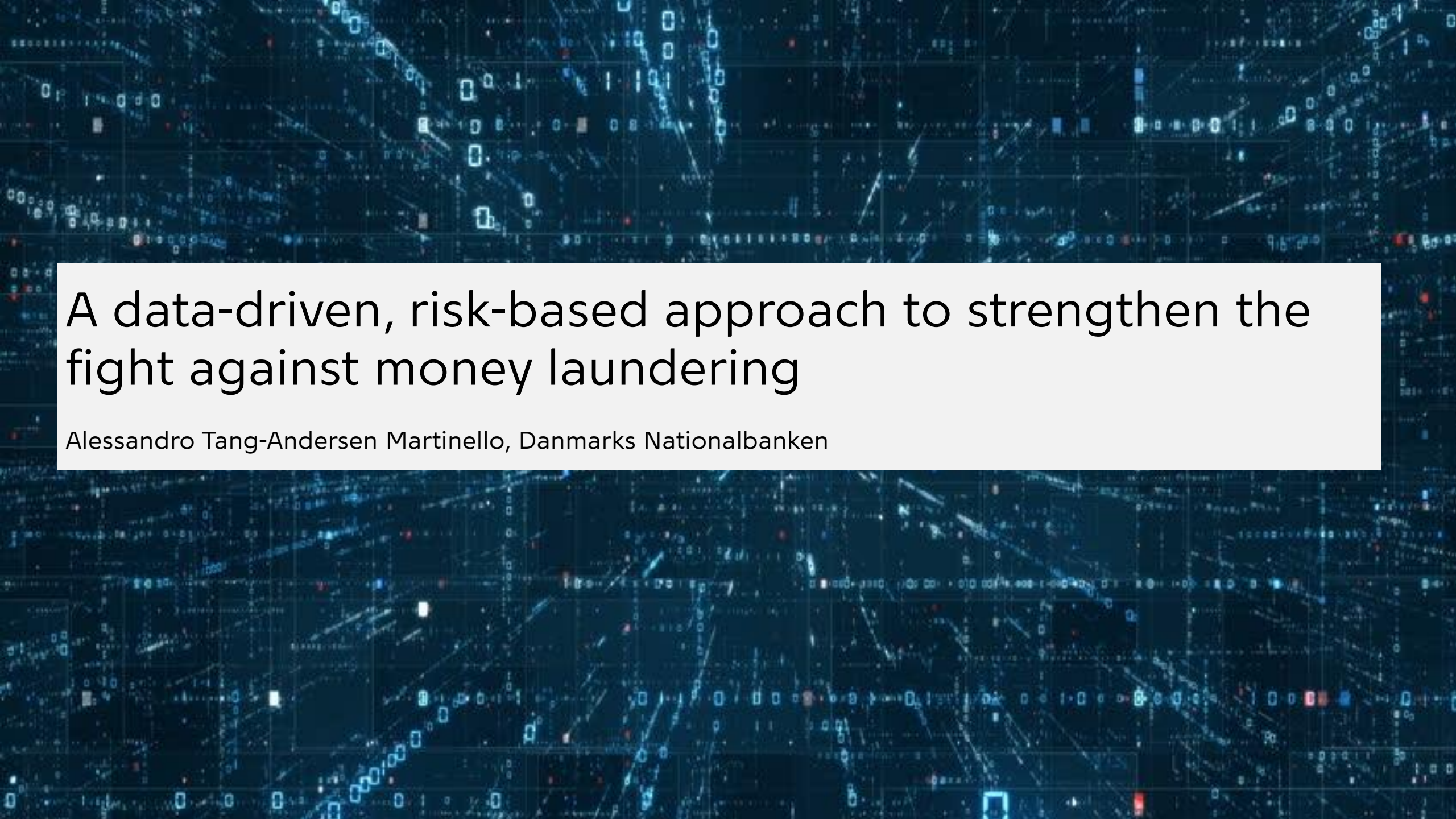
63rd ISI World Statistics Congress (WSC) / IFC-sponsored Invited Paper Sessions (IPS) related to new developments in central bank statistics  
July 2021, The Hague, Netherlands, virtual format

## A data-driven, risk-based approach to strengthen the fight against money laundering<sup>1</sup>

Alessandro Tang-Andersen Martinello, Thais Lærkholm Jensen and Bjarke Mørch Mønsted,  
Danmarks Nationalbank

---

<sup>1</sup> This presentation was prepared for the WSC. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.



# A data-driven, risk-based approach to strengthen the fight against money laundering

Alessandro Tang-Andersen Martinello, Danmarks Nationalbanken

# Granular transaction data can strengthen efforts against money laundering

## Implementation of Proof of Concept (POC) in the first half of 2020

**3**

data sets from SØIK, the participating bank and the Danish Business Authority were combined

**2B**

transactions analyzed as part of the database

**18.000**

lines programming code written during the PoC

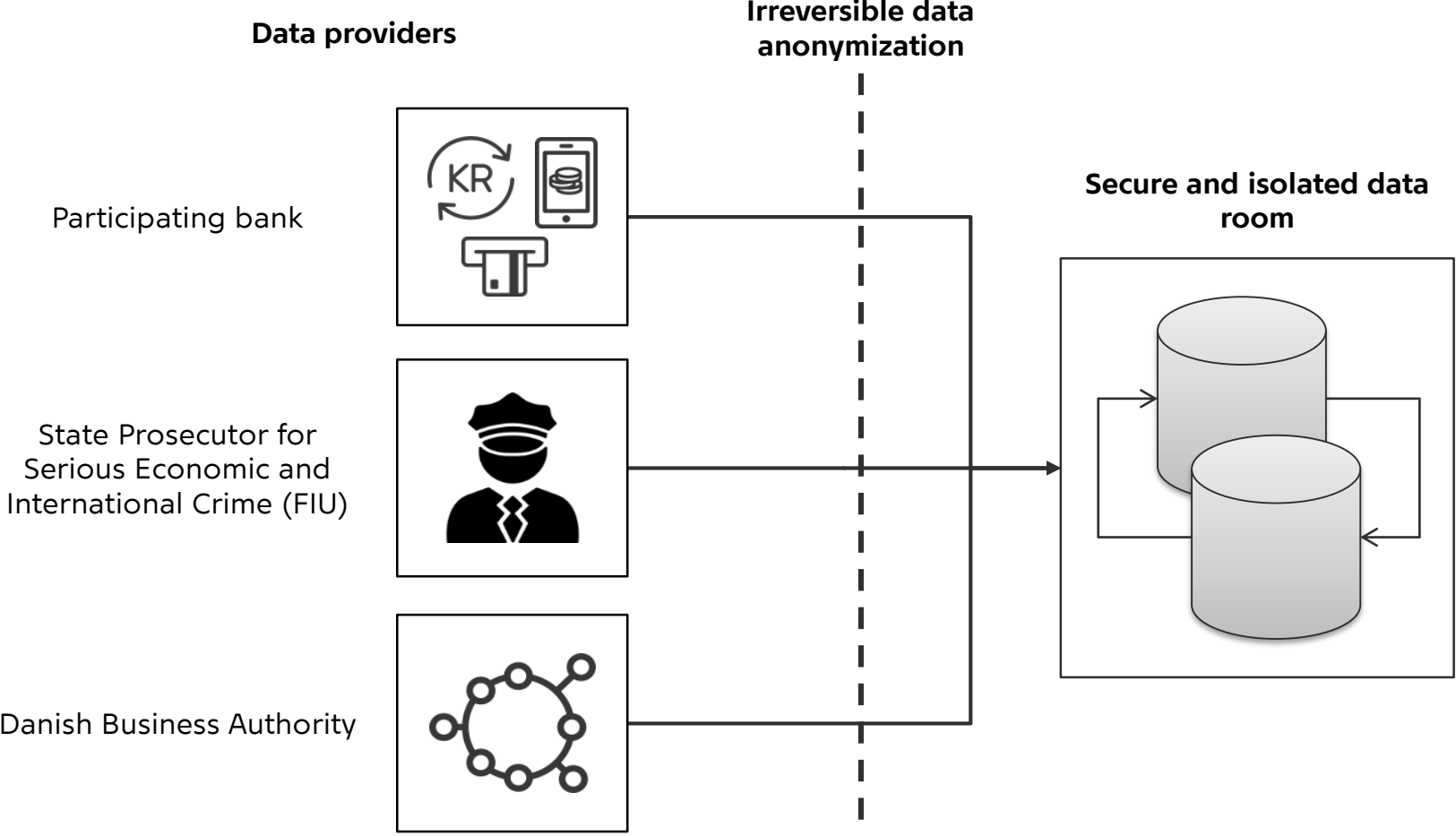
**400.000**

companies and individuals in the data set

**10**

data scientists and experts were involved and worked from the "war room"

# The project required special dispensation from payment law and security clearance of project participants



# Combining data across the financial sector allows better, faster, and more effective countering of money laundering

## Results of the Proof of Concept

### Bottom of the iceberg

Identifying suspicious patterns that are not investigated in the current system



An innovative data-driven approach can identify types of money laundering that are not caught by the current system

### Earlier identification

Flagging suspicious behavior earlier



The same suspicious transactions can be identified earlier

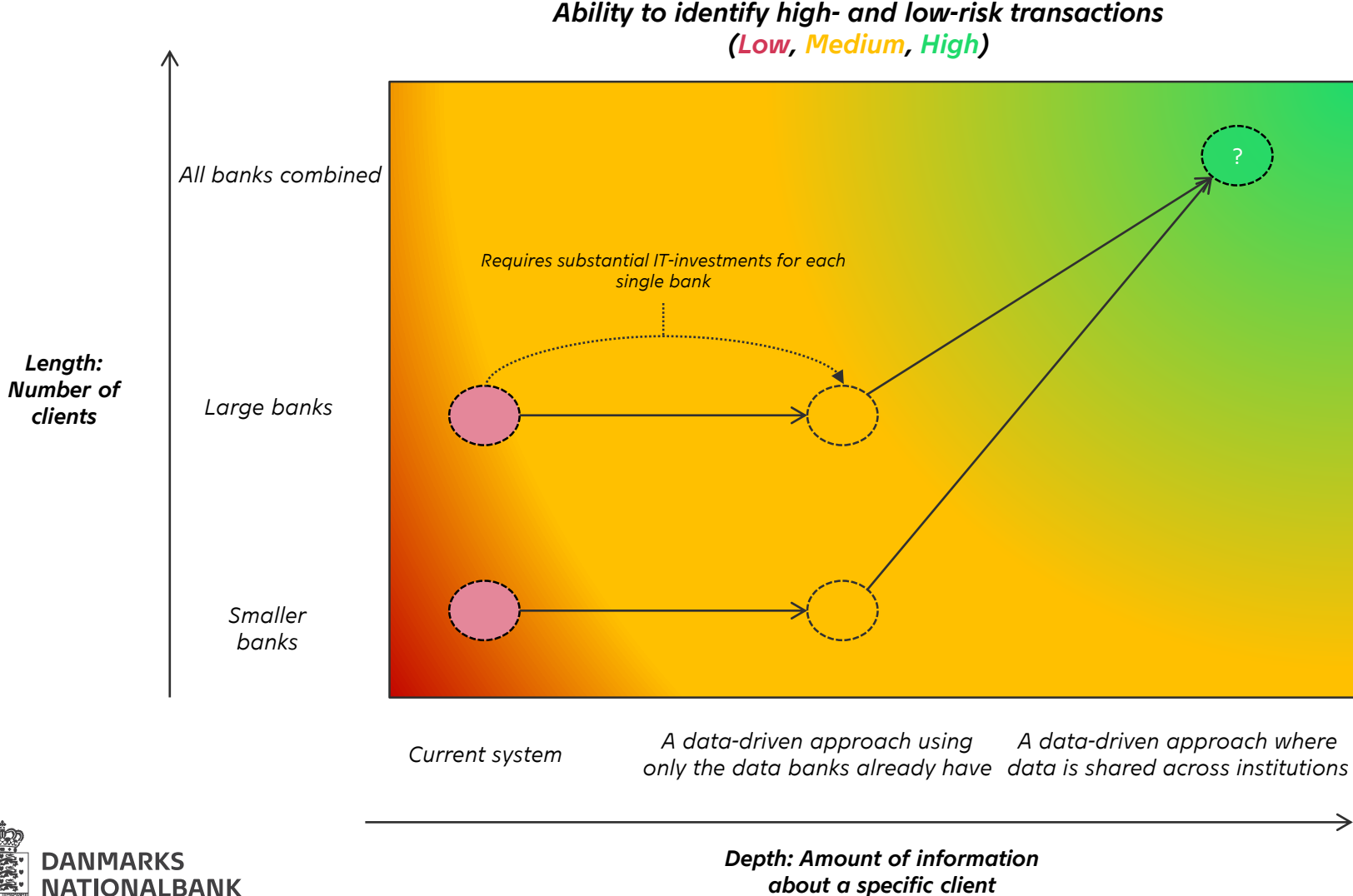
### Better value-chain

Streamlining the processing of suspicious transactions



As low-risk cases can be automatically discarded, resources can be redirected towards high-risk cases

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





## **A data-driven, risk-based approach to strengthen the fight against money laundering**

Thais Lærkholm Jensen<sup>1</sup>; Alessandro Tang-Andersen Martinello<sup>1</sup>; Bjarke Mørch Mønsted<sup>1</sup>

<sup>1</sup> Danmarks Nationalbank, Langelinje Allè 47, Copenhagen

### **Abstract:**

The paper assesses the potential of intensifying the usage of granular transaction data to strengthen the fight against economic crime. We show that 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 however requires combining data across banks and public institutions. This combination would require a revised legal framework.

### **Keywords:**

Money laundering; Transaction data; AML; Machine learning; Network

### **1. Introduction:**

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. 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. Methodology:**

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 from banks included in reports sent for 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.<sup>1</sup>
- 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.

### **3. Result:**

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 be detected.

The project also shows that a data-driven approach can optimise the current manual investigation process. 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 SØIK, and as such considered false alarms.

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.

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.

---

<sup>1</sup> 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.



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.

#### **4. Discussion and Conclusion:**

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. 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. In contrast, a single platform for collecting 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. 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. 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.

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.