Anti-Money Laundering in the Real Estate Sector

Overview and recommendations for data models relating to money laundering in the real estate sector for British Columbia

BC-Canada Working Group on Real Estate

Work Stream 1: Data Collection and Sharing

December 9, 2020

ACKNOWLEDGEMENTS

This report was compiled under the direction of Haig McCarrell, Data Work Stream co-lead; the principal authors are Jean-Philippe Deschamps-Laporte, Project Lead; Ellen Bekkering, Project Manager; and, principal analysts, Annik Gougeon and Jason Jensen, all from Statistics Canada.

Jonathan Baron, British Columbia Ministry of Finance, Data Work Stream co-lead, and the members of the Data Collection and Sharing Work Stream, listed in Appendix 5, contributed to the report as well.

TABLE OF CONTENTS

GLOSSARY

AML	Anti-Money Laundering
AMLC	Anti-Money Laundering Centre (the Netherlands)
AUSTRAC	Australian Transaction Reports Analysis Centre
BC	British Columbia
BCA	British Columbia Assessment
BC FSA	British Columbia Financial Services Authority
BC LOTR	British Columbia Land Owner Transparency Registry
BC LTSA	British Columbia Land Title and Survey Authority
BC MBTS	British Columbia Mortgage Broker Tracking System
BC TACS	British Columbia Tax Administration Compliance & Services
BDRF	Banking Data Referral Portal (the Netherlands)
BR	Business Register
BSA	Bank Secrecy Act (US)
CBCR	Cross Border Currency Report
CBSA	Canadian Border Services Agency
CCJCSS	Canadian Centre for Justice and Community Safety Statistics
CDR	Casino Disbursement Report
CHSP	Canadian Housing Statistics Program
CRA	Canada Revenue Agency
CSAIR	Condo and Strata Assignment Integrity Register
CSIS	Canadian Security Intelligence Service
CSP	Corporate Service Provider
CSV	Comma-Separated Values
DAML SAR	Defence Against Money Laundering Suspicious Activity Report (UK)
DATF SAR	Defence Against Terrorist Financing Suspicious Activity Report (UK)
DEA	Drug Enforcement Agency (US)
DEFRA	Department for Environment, Food, and Rural Affairs (UK)
DHS	Department of Human Services (Australia)
EDTF	El Dorado Task Force (US)
EFTR	Electronic Fund Transfer Record
FATF	Financial Action Task Force
FEC	Financial Expertise Centre (the Netherlands)
FBI	Federal Bureau of Investigation (US)
FinCEN	Financial Crimes Enforcement Network (US)
FINTRAC	Financial Transactions and Reports Analysis Centre of Canada
FIOD	Fiscal Information and Investigation Service (the Netherlands)

FCI-net	Financial Criminal Investigation network
FIU	Financial Intelligence Unit
FLSC	Federation of Law Societies of Canada
FRFI	Federally Regulated Financial Institutions
GTO	Geographic Targeting Order (US)
HMRC	Her Majesty's Revenue and Customs
ICE	Immigration and Custom Enforcement (US)
ICCS	Integrated Criminal Court Survey
iCOV	Infobox for Criminal and Unaccountable Assets (the Netherlands)
IM	Information Management
iRR	Relations Report (the Netherlands)
IRS-CI	Internal Revenue Service – Criminal Investigation (US)
iRT	Thematic Report (the Netherlands)
iRVI	Capital and Income Report (the Netherlands)
ISED	Innovation, Science and Economic Development Canada
IT	Information Technology
ITN	Individual Tax Number
JMLIT	Joint Money Laundering Intelligence Taskforce (UK)
LCTR	Large Cash Transaction Record
LEA	Law Enforcement Agency
LIEC	National information and Expertise Centre (the Netherlands)
LLC	Limited Liability Company
LOTA	Land Owner Transparency Act (BC)
LOTR	Land Owner Transparency Registry (BC)
LTSA	Land Title and Survey Authority (BC)
MBTS	Mortgage Broker Tracking System (BC)
ML	Money Laundering
MSB	Money Services Business
NCA	National Crime Agency (UK)
NDEC	National Data Exploitation Capability (UK)
NECC	National Economic Crime Centre (UK)
NGO	Non-Governmental Organization
OPBAS	Office for Professional Body Anti Money Laundering Supervision (UK)
OCDETF	Organized Crime Drug Enforcement Task Force (US)
OFC	OCDETF Fusion Center (US)
OSFI	Office of the Superintendent of Financial Institutions
PCMLTFA	Proceeds of Crime (Money Laundering) and Terrorist Financing Act
PEP	Politically Exposed Person

PET	Privacy Enhancing Technology
PPSC	Public Prosecution Service of Canada
PSPC	Public Services and Procurement Canada
RCMP	Royal Canadian Mounted Police
RDBMS	Relational Database Management System
RE	Real Estate
RECBC	Real Estate Council of British Columbia
RIEC	Regional Information and Expertise Centre (the Netherlands)
SAR	Suspicious Activity Report (UK and US)
SCTF	Serious Crimes Task Force (the Netherlands)
SIN	Social Insurance Number
SMR	Suspicious Matter Report (Australia)
SQL	Structured Query Language
STR	Suspicious Transaction Report
T1	Canada Revenue Agency T1 personal tax filings
T2	Canada Revenue Agency T2 corporate tax filings
T3	Canada Revenue Agency T3 trust tax filings
TACS	Tax Administration Compliance & Services (BC)
TPR	Terrorist Property Report
UCR	Uniform Crime Reporting
UK	United Kingdom
UKFIU	United Kingdom Financial Intelligence Unit
US	United States of America
USSS	United States Secret Service
UTR	Unusual Transaction Reports (the Netherlands)
UWO	Unexplained Wealth Order (UK)
VAT	Value-Added Tax
VIR	Voluntary Information Record

1 EXECUTIVE SUMMARY: KEY FINDINGS AND RECOMMENDATIONS

1. INTRODUCTION

This Executive Summary of the Data Collection and Sharing Work Stream of the BC-Canada Working Group on Money Laundering (ML) in Real Estate[1] report focuses on the Work Stream's key findings and recommendations. The outcomes described herein are based on research and collation of money laundering typologies, an assessment of data holdings and data sharing practices of work stream members, a review of data practices of foreign anti-money laundering (AML) regimes, data management best practices, a literature review and consultations with subject-matter experts in Canada.

The Work Stream, co-lead by Statistics Canada and the British Columbia Ministry of Finance, considered the data needs of government authorities and data models for research and analytical functions, based on the feasibility of producing a data framework to facilitate information sharing among relevant government bodies. Although the focus of the Work Stream was on ML in British Columbia residential real estate, the findings and recommendations can support tax enforcement, which could benefit from the inter-agency collaboration and data sharing discussed.

AML data frameworks can be applied for both strategic and tactical purposes. Applying the framework for policy purposes could produce general estimates of money laundering in real estate at the aggregate level, useful for relevant governmental entities and policymakers. An application for strategic purposes could focus on more narrow metrics or trends that inform emerging patterns of illicit activity. An application for tactical purposes would focus on enforcement, analysing information with the intent of identifying and apprehending suspected money launderers. These complementary applications could benefit from the establishment of a data model facilitating the identification of potential money laundering and greater data access and sharing among AML regime partners.

The application of an enhanced AML data framework for real estate would rely on data sharing among organizations in the AML regime and other authorities. Three different data model options are suggested for the ongoing provision of data, with varying impacts to privacy and legislative constraints. The report presents opportunities and challenges in applying the models without being prescriptive of a preferred option.

Legal aspects such as changes to agency mandates, in depth privacy and Charter implications, or resource requirements associated with data and data framework implementation were not explored. Furthermore, the AML data framework is a conceptual model and will require empirical testing to determine its applicability to AML operations.

[1] The Work Stream members include the Bank of Canada, the British Columbia Ministry of Finance, the British Columbia Financial Services Authority (BCFSA), the Real Estate Council of British Columbia (RECBC), the British Columbia Securities Commission (BCSC), the Canada Revenue Agency (CRA), the Canada Mortgage and Housing Corporation (CMHC), Finance Canada, the Financial Transactions and Reports Analysis Centre of Canada (FINTRAC) the Royal Canadian Mounted Police (RCMP) and Statistics Canada.

1.1 KEY FINDINGS FOR AN ANTI-MONEY LAUNDERING FRAMEWORK IN THE REAL ESTATE SECTOR

1. **Laundering money through the Canadian real estate market uses a diverse array of methods.**
   A systematic review of relevant literature reveals 8 schemes and 22 sub-schemes for laundering money in real estate. These schemes represent different ways of obscuring the origins of the illegally acquired funds through real estate transactions and the placement of these funds in real estate investments

2. **Court records do not reflect the full extent of ML efforts in Canada.**
   In examining court outcomes of money laundering offences from 2009 to 2016, it was found that many police-reported incidents involving money laundering did not result in a money laundering charge. Furthermore, the majority of money laundering charges that made it to court were stayed, withdrawn, dismissed or discharged. However, guilty verdicts were often reached on other charges in these cases.

   Similarly, the extent of reported money laundering incidents is frequently obscured in public reporting as more serious offences occurring in the same incident become the principal statistical record.

   The low counts of money laundering charges proceeding in Canada's criminal courts may be related to the complexity of prosecuting money laundering offences. The *Criminal Code* has recently been amended to facilitate the prosecution of money laundering[2]. A future assessment of the impact of this change would reveal whether the difficulty in prosecuting a money laundering offence had an impact on the low number of charges associated with police-reported incidents.

   FATF Mutual Evaluation Reports are evolving to include more measure of outputs of AML efforts such seizures, charges and convictions. Although court proceedings are objective indicators, they do not reflect the full extent of AML efforts in Canada, and international comparisons should take into consideration the different operating contexts of national AML regimes.

3. **Anti-money laundering efforts would be more effective by enhanced partner collaboration and data sharing.**
   Recent reports[3] have drawn attention to the ways in which money is laundered in the BC real estate market, and have estimated the extent of money laundering in real estate by using data held outside the AML regime partners[4]. These analyses have highlighted the need for more data sharing

[2] *Criminal Code,* 1985, as amended 2019-06-21
[3] German 2019; Maloney et al. 2019
[4] Finance Canada, Department of Justice Canada, Global Affairs Canada, Public Safety Canada, Financial Transactions and Reports Analysis Centre of Canada (FINTRAC), Office of the Superintendent of Financial Institutions Canada (OSFI), Innovation, Science and Economic Development Canada (ISED), Office of the Privacy Commissioner of Canada, Royal Canadian Mounted Police (RCMP), Public Prosecution Service of Canada (PPSC), Canada Revenue Agency (CRA), Canada Border Services Agency (CBSA), Canadian Security Intelligence Service (CSIS), Public Services and Procurement Canada (PSPC), as well as other partners including provincial, territorial, and municipal Law Enforcement Agencies (LEAs), financial sector regulators, and self-regulatory organizations.

to facilitate a broader approach and to better understand the extent of money laundering in real estate.

3.1. Canadian organizations engaged in AML initiatives use real estate data for specific investigations and case-based approaches, and have participated in partnerships to facilitate a broader approach.

Targeted projects have brought together public and private-sector parties to overcome problems of differential access to data for AML purposes. For example, Project Athena brought together public and private sector institutions to exchange knowledge and information on casino-related money laundering. Recently, the scope of the cooperation has expanded to include real estate and other high-value items. Implementing an AML data framework supported by money laundering real estate typologies could supplement ongoing public-private partnership efforts.

3.2. Effective AML initiatives could benefit from the participation of other relevant organizations, particularly those in the real estate sector.

Regulatory and administrative institutions in the BC real estate market collect and acquire data on properties, real estate transactions, real estate financing, and the actors involved, which could be used for the detection of ML in real estate. Some of these organizations are not required to undertake or aid in AML efforts, so information is largely not kept, maintained, or shared for this purpose. Canadian law enforcement agencies may access real estate data on a case-by-case basis, but mainly encounter money laundering in real estate as a by-product of their investigations or prosecutions.

4. An AML data framework for real estate can contribute to identifying money laundering in real estate.

An approach to systematically identify instances of money laundering in real estate through data can generate insights to drive policy and enforcement efforts. Money laundered through the acquisition or use of real estate often employs methods inconsistent with standard real estate market practices, and as a result, can be identified using data about real estate owners, buyers, and their transactions.

The AML framework presented systematically defines indicators of money laundering by enumerating money laundering schemes and the ways the schemes are visible within data. The availability of data required to construct each indicator was also assessed, and this information is summarized at the sub-scheme level in Table 6-1 (p. 87).

4.1. The comprehensive data framework presented requires accessing 160 data points within a data model for ML in real estate.

The 8 schemes and 22 sub-schemes for money laundering in the real estate sector break down into 160 individual data points required for their detection. Based on the analysis of public data holdings, some sub-schemes would not be detectable as they rely on data which does not exist, or would be difficult to detect.

4.2. An AML data framework could provide information on the extent of money laundering in real estate, and inform policies and programs.

Contingent on data of sufficient quality and comprehensive data sharing, a data model for AML in real estate could be used to perform consistent and systematic red flag analyses of indicators. These analyses could:

- Provide indications of the extent and incidence of money laundering in real estate, including the approximate number and value of properties that may be affected.
- Provide benchmarking indicators to assess the impact of new AML policies and efforts.
- Inform analytical and enforcement efforts of geographic hot spots or prevalent schemes.
- Advise law enforcement agencies and policy makers on emerging money laundering schemes.

4.3. The efficacy of a data framework for AML in real estate is contingent on extensive and high quality data coverage.

Identifying money laundering using data indicators would require comprehensive coverage of real estate transactions and ownership arrangements within a given geographic area over time. The process relies on detecting methods and practices atypical to standard real estate practices. However, irregular or non-standard behaviours are not always a result of money laundering attempts. Prudent money launderers will use schemes that more closely mimic regular market behaviour, thereby increasing the difficulty of detection. Data of higher quality will be more able to detect smaller anomalies or peculiarities in real estate transactions.

5 The AML data framework for real estate described would be internationally innovative.

The framework presented is novel in terms of the scope of data linked and analyzed, in its focus on real estate, and in its potential to detect real estate purchased using criminal proceeds going back in time. This contrasts with existing approaches in Canada and abroad which generally detect illicit funds as they are being laundered.

5.1 Models for data approaches and extensive public-private partnerships to detect money laundering can be found in other countries.

The AML regimes examined – namely the UK, US, Netherlands, and Australia – allow access to information and data on suspicious transactions to a wider range of public institutions and analysts than is permitted in Canada. In some cases, they allow this data to be merged with their own datasets. Constitutional and privacy protections differ among countries, which must be considered in assessing the possible adoption of foreign AML regime data approaches in the Canadian context.

Most of the countries studied have developed well-functioning public-private partnerships in which data on cases are exchanged and typologies are developed. Public-private partnerships in Canada, such as Project Athena, have been successfully implemented as responses to specific threats, such as money laundering in casinos. Greater data availability and enhanced information sharing could support deeper public-private partnerships.

5.2 Data-driven approaches are an emerging strategy in AML efforts.

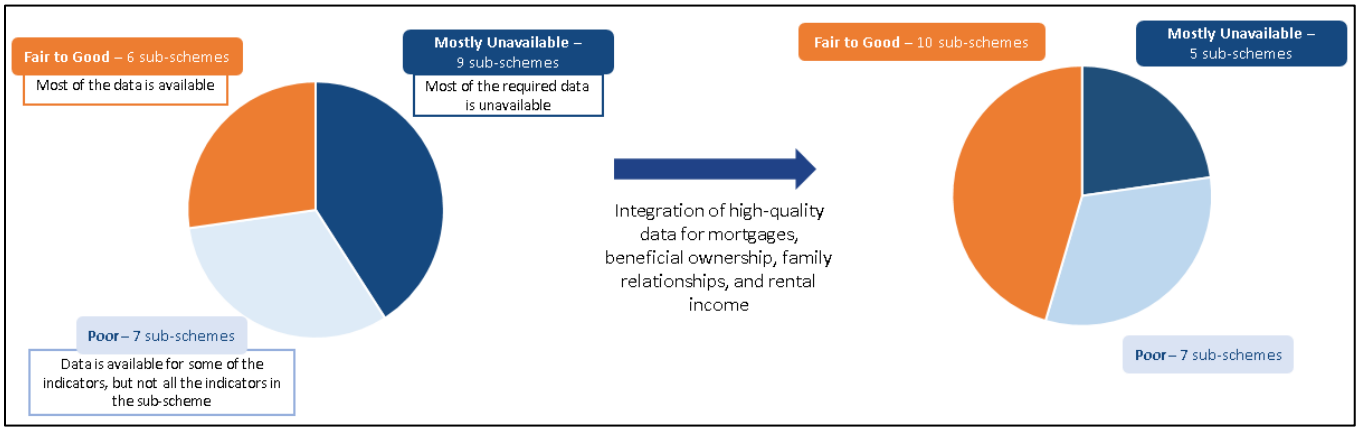
There is a broad acknowledgement of the threat posed by ML and the need for data to detect it. This has spurred initiatives aimed at expanding capacities for large-scale data analysis at AML

institutions abroad, as well as the formation of partnerships for approaches employing multiple datasets. The Dutch system, iCOV, exchanges administrative data relating to financial crime among public institutions with customized reports to partner agencies.

6 A data framework for AML in real estate would require the creation or acquisition of additional key data sources.

Several data gaps in current data holdings would need to be filled before a framework for AML in real estate would be able to produce indicators for most of the money laundering schemes and sub-schemes identified. Six of 22 sub-schemes could be detectable without filling these gaps, and an additional four sub-schemes could be detectable once these data gaps are addressed.

Figure 1: Improvement in detectable sub-schemes if key data limitations were addressed



6.1 Beneficial ownership information is a key data gap.

The obfuscation of ownership is a facet of the most commonly used schemes for laundering money in real estate. Two of 22 sub-schemes (using a shell company/trust and lending money to yourself) have indicators relying on accurate data of beneficial ownership of legal persons.

In BC, the beneficial ownership information respecting corporate interest holders of BC companies is required to be maintained by the company itself. Retrieval of this information is tedious for enforcement authorities and the information itself may be falsified, rendering it of little use for detection.

The province of British Columbia is addressing this data gap. The Land Owner Transparency Registry (LOTR) will house beneficial ownership information on all indirect interests in land registered with the BC LTSA. Additional beneficial ownership information on companies which do not own real estate property would need to be more readily available to better detect the sub-scheme of lending money to oneself.

A beneficial ownership register of corporations could provide key data points for the AML data framework in real estate. Experience with the UK's beneficial ownership register indicates that continuous verification of the information in the register is required to improve and maintain accuracy. In the UK, the public nature of the register has enabled the public to verify the

accuracy of the information. However, a public registry could also introduce risks of identity theft, fraud, and harassment.

6.2 Mortgage information is a key data gap.

Eight sub-schemes involve the utilization of one or more mortgages. Mortgage data are held mainly by financial institutions and consumer credit agencies, and by some public institutions.

The Land Title and Survey Authority of British Columbia collects mortgage information for properties purchased in British Columbia as a part of their work administering the provincial land title system. However, the data is not currently employed for the systematic detection of money laundering and would be largely unsuitable within the AML data framework given that mortgage data is generally entered only at the time of a land transfer and is not otherwise updated. Furthermore, data holdings may not be precise on key mortgage characteristics and is frequently archived in analog formats.

6.3 Relationship information is a key data gap.

Three sub-schemes rely on the use of family members, close friends, business associates and real estate professionals involved in a real estate transaction, which could constitute networks of interest for understanding ML. The use of family members as nominee owners appears to be a common sub-scheme, making the ability to identify such relationships useful for ML detection**.** Additionally, in police-reported incidents of money laundering where an accused was identified, nearly half had been involved in recent other criminal incidents and more than two-thirds of incidents involved multiple accused.

If real estate professionals involved in a transaction were identified by a unique identifier associated with the property being transacted, networks could be better understood, helping to reveal clusters of ML activity.

Commercially available databases on Politically Exposed Persons (PEPs) and other relevant databases could be leveraged to build familial and criminal network information.

6.4 Wealth information is a key data gap.

Nominee real estate owners can be identified by determining that the individual cannot reasonably afford the property based on their present and historical wealth holdings. An accurate appraisal requires knowledge of the nominee's level of income and wealth prior to the acquisition of the property.

Data on individual wealth are not available from any known centralized data source. Data from many financial institutions, credit rating agencies, and public and private pension funds could be combined to produce wealth estimates for all homeowners or could be a required filing for purchases that are self-financed or not involving registered lenders.

6.5 Rental revenue is a key data gap.

Property owners who rent their properties are required to file a statement of real estate income when completing their income taxes (income tax Schedule T776). Individuals could pad their rental income with the proceeds of crime, effectively leveraging their properties to launder money.

Data to detect this technique is currently scarcely available in British Columbia. Landlords are required to declare rental income on their tax returns, but are only obligated to list the addresses of their rental properties and the gross rents received that year. Additionally, renters are excluded from the requirement to list their rental expenses, preventing any authentication of the values entered by the landlord.

7 Three different data models have been identified to facilitate greater access and sharing of data in the implementation of an AML data framework

7.1 A "Distributed" model is currently the modus in Canada's AML regime.

Data is held by specific organizations and shared on a limited basis, typically with law enforcement agencies, when there are reasonable grounds to suspect that financial transactions or attempted financial transactions are related to the commission of a money laundering offence. This satisfies privacy considerations, but does not enable existing data to be combined and leveraged effectively for AML purposes. An enhancement of sharing under this framework would increase data utilization, but would duplicate work throughout the participating organizations.

7.2 A "Centralized" model would designate one institution to be responsible for holding and managing all data within the framework.

This would enable the systematic analysis of consolidated data for the detection of money laundering. Operational efficiency is a key advantage of this model. A centralized model would help to ensure that the processing, cleaning, quality, and documentation of the data is consistent throughout the various sources, resulting in reliable and accurate outcomes. However, this model comes with potentially greater privacy implications and may require extensive legislative changes.

7.3 A "Hybrid" model would seek to optimize and enhance data holdings within organizations where data can be shared, coupled with an organization or unit to coordinate or lead data efforts for strategic and tactical use.

This model combines advantages of a centralized model while addressing some implications surrounding privacy and legislative changes. It would leverage current public data holdings by centralizing subsets at designated custodial institutions, which would serve as key sources of data on a given subject – such as real estate or taxes. The model would coordinate access and analysis of data by AML regime partners.

Some data holders have no AML mandate, and face operational constraints related to the organization of their data holdings that limit database search and research functions. If expanded functionality is to be realized, additional resources may be required. An AML mandate, as a complement to a non-AML regime institution's existing mandate, could enable the institution to expand its current data collection to close critical AML data framework gaps. Data custodians would be responsible for subsequent data cleaning and processing, as well as maintaining the quality of the data, and for organizing data to enable search and access functions.

In addition to creating key AML data custodians, this model should designate a coordinator to lead data linkage and analytical efforts related to ML in real estate, which could be fulfilled by an existing regime partner, a special unit tasked with this responsibility, or a consortium of AML regime partners. Adequate data base functionality within all data custodians would enable the designated coordinator to facilitate and/or undertake strategic and tactical research in cooperation with other AML regime partners. Authorized enforcement and research institutions could have access to relevant data through the coordinating unit or directly from the data custodian – as deemed appropriate.

Key to the implementation of each of these models is the development, availability and sharing of data to close identified data gaps in the AML data framework presented.

8 Applying the AML data framework for real estate requires sounds data management from contributing data holders.

A key challenge for the use of data from multiple sources is the assurance of data quality and standardization across data providers. Sound data management practices can help improve the overall accuracy of the identification of money laundering in real estate and can contribute to the avoidance of flawed inferences. As such, the implementation and maintenance of adequate data management practices among public data holders ensures that collected data maintains its value both within and outside of the institution holding the data. This includes adequate data structuring, data quality, data retention and information management practices.

9 Privacy and other legal protections need to be analyzed further when considering ways to enhance data-sharing on AML framework for real estate.

While the potential of an AML data framework is demonstrated at the conceptual level, a thorough review of the practical implications in the context of existing legislation is necessary. As such a model requires the collection, integration, storage, and analysis of personal data stewarded by multiple data holders, an equally extensive assessment of the legal compatibility and privacy impacts of such an approach is warranted.

Whereas data in the public domain greatly enhances the shareability of data, it can raise privacy concerns. As noted previously in Finding 6.1, the publicly accessible UK BOR enables citizen involvement to verify data, aiding AML efforts, however it can have unintended consequences. These consequences, including the risk of commercial data exploitation, could be mitigated by restricting data to authorized users and uses. It is not intended that data maintained for AML purposes be available to financial service providers to market their products or to solicit business.

9.1 Current legislation limits data sharing.

The Privacy Act, governing federal institutions, broadly prohibits the sharing of personal information except for some purposes such as law enforcement, research and other purposes consistent with the statute that authorized the data collection. Additional acts govern the use of data holdings, such as the PCMLTFA for FINTRAC and the Statistics Act for Statistics Canada. The Freedom of Information and Protection of Privacy Act (FOIPPA), governing provincial institutions in BC, mostly mirrors the Privacy Act.

Many statutes contain their own information disclosure provisions, which may allow for broader information sharing between and among provincial and federal institutions.

9.2 The use of personal data is regulated, and considerations of a data framework will need to elucidate and assess privacy impacts.

A thorough assessment will be needed to consider the potential benefits to combatting money laundering by implementing the three data models presented, against the privacy impacts of an AML data framework.

9.3 Separate assessments are required for the generation of strategic and tactical insights.

The use of personal data for the purposes of research or law enforcement is addressed independently in prevailing legislation. The legal requirements and privacy impact of any implementation of an AML data framework in real estate vary depending on its intended use. The generation of strategic insights – aggregated results to inform policymakers, is subject to legislative restrictions governing the supporting organization. The generation of tactical insights, findings on individuals and organizations which inform law enforcement actions, will be regulated by legislation applied to law enforcement agencies. Both have extensive privacy implications for Canadians. Separate assessments of the legal restrictions and privacy impacts are required for each use case.

10 Data framework approaches may be relevant beyond the real estate sector.

While the findings of the Work Stream members has focused on money laundering in real estate, financial crimes in Canada span a multitude of other avenues such as tax evasion and money laundering through luxury items or false invoicing. The approaches discussed may be relevant to data and information analysis and sharing with respect to other sectors.

1.2 RECOMMENDATIONS FOR ANTI-MONEY LAUNDERING IN THE REAL ESTATE SECTOR

More and better data can help raise flags on the illicit use of real estate to launder money.

Specifically, the Data Work Stream recommends that:

1) **A hybrid data model consisting of custodial and co-ordinating functions could support high-quality data on BC real estate for either strategic and/or tactical purposes and merits further consideration.**
   a) A custodial function is mostly applicable to non-regime partners but regime organizations may perform this role as well. Formal recognition of custodial roles could be considered to enable and/or foster these functions within designated organizations, which could include:
      i) expanded collection or receipt of AML related data
      ii) verification, processing, storage and organization of relevant data
      iii) sharing data or otherwise facilitating its access, including creating and maintaining data base search functions
      iv) ensuring data is protected and used only for authorized uses. Commercially exploitable data relevant to the AML data framework, such as mortgage terms, should only be accessible to authorized data users.
   b) A co-ordinating function within a hybrid data model could be performed by an existing AML regime partner, a newly constituted unit or organization, or a consortium of organizations. This function could include:
      i) determining options for sourcing, linking and sharing AML framework data among regime partners;
      ii) the maintenance and development of the AML data framework
      iii) AML related data analysis and research including research to determine the most effective indicators for detecting money laundering;
      iv) setting data standards to facilitate AML data sharing and system inter-operability for authorized data users
      v) reducing and eliminating duplication of activities among hybrid model partners
      vi) the development and production of metrics such as suggested in Recommendation 7.

2) **Organizations holding AML data could further consider how their data holdings can be expanded or improved within the purview of their current operations with a view of better filling the AML data framework data gaps.**

   Additional data points that would enhance the application of the AML data framework could include:

   a) the names of all significant actors[5] in real estate transactions. This could be facilitated if a unique identifier, such as a property ID, was required for all actors in a real estate transaction;
   b) specific and accurate terms of financial contracts to acquire real estate, including lender(s), principle amounts, interest rates, amortization periods;
   c) updated financial records when real estate is refinanced or the contract terms are otherwise altered;

[5] Including lawyers, realtors, real estate firms, brokers, financial institutions, lenders, and guarantors.

d) the retention of land title applications that are rejected or returned to the applicant due to possible irregularities.

Data that is publicly accessible will facilitate (a data custodian's) data sharing and data organizations that hold publicly accessible data will be better placed to assume AML data custodial functions.

3) **BC registries aimed at improving corporate transparency consider:**
   a) including all Canadian (extra-provincial and federally incorporated) and foreign corporations registered in BC;
   b) adopting practices to regularly update and verify submissions including allowing public access to enable citizen participation to verify submissions;
   c) that registry information, at a minimum, be made available to law enforcement, tax and other competent authorities.

4) **Key data gaps on professional and personal relationships should be filled through further work to establish the feasibility of linking proxy ownership and criminal networks or politically exposed persons.** This could include:
   a) building and maintaining network information;
   b) leveraging information on co-accused in registered criminal cases and other public records to help detect cases of potential ML;
   c) leveraging information on PEPs and familial relations in public and commercial databases to help detect cases of potential ML.

5) **AML regime organizations should review their data stewardship and management practices.** This includes maintaining data that is validated, and kept in machine-readable and searchable formats, along with up-to-date data documentation, adequate data retention and protection strategies.

6) **Comprehensive and objective indicators relevant to Canada be created for strategic uses to measure the effectiveness of AML programs, processes, strategies and initiatives related to real estate.** These could include recording and tracking the following type of information:
   a) resource inputs such as salaries, full-time equivalent employees, program budgets;
   b) outputs such as charges, convictions, number and value of seizures, number of reports generated or information exchanges (including by source), etc.;
   c) outcomes such as estimations of price impact of illicit flows in real estate markets, consumer welfare loss and ML activity trends in real estate;
   d) comparing efforts and outcomes of AML efforts related to ML across provinces and internationally, taking into account the Canadian operating environment (e.g., administrative regulatory context, the Charter of Rights and Freedoms, federal system, etc.).

7) **Further work** could be conducted to/by**:**
   a) evaluate the potential to fill data gaps using data holdings not studied in this report, including holdings by the private sector;

b) investigate the potential to fill the rental data gap pertaining to rents received and paid;
c) investigate the application of new privacy enhancing technologies and artificial intelligence applications to identify and share relevant data and information with responsible authorities;
d) investigate ways to collect beneficial ownership information concerning the private financing of real estate to address schemes involving loans among non-arms-length parties;
e) Finance Canada to further assess the feasibility of collecting information on the purpose of international financial transfers;
f) develop models to measure ML in Canada and the flows of illicit funds to and from Canada, possibly engaging Statistics Canada to enhance its measure of the underground economy and undertake a feasibility study to build an economy-wide framework to measure ML in aggregate. Consideration be given to reactivate its financial performance ratio program by industry and geography to help identify outliers consistent with ML activity;
g) the RCMP and/or police departments to consider estimating the value of proceeds of crime for all incidents as an input into measures of ML in Canada.

2 OVERVIEW OF MONEY LAUNDERING ACTIVITIES IN CANADA

The following chapter provides an overview of official police and court statistics regarding offences of money laundering (ML). The data used in this chapter is not restricted solely to money laundering in real estate, and instead focuses on all types of ML offences, as the source data does not include variables to identify specific types of ML.

Money laundering has been acknowledged as a widespread problem, both in Canada and internationally. Also known as laundering the proceeds of crime, money laundering is a chargeable offence under sections 462.31 and 462.33 of the *Criminal Code of Canada*. Money laundering occurs when an individual or group uses, transfers, sends, delivers, transports, transmits, alters, disposes of, or otherwise deals with any property or proceeds of any property that was obtained as a result of criminal activity. This is done with the intent to conceal or convert illegal assets into legitimate funds.

Reports have given insight to the extent of ML occurring in British Columbia (BC) and Canada by estimating the overall amount of ML at the national and provincial level (Maloney et al., 2019) or by estimating the occurrence of various real estate-related ML schemes in BC (German, 2019). This chapter extends this understanding by analyzing justice statistics on money laundering.

There are challenges in estimating the true nature and extent of money laundering in Canada, given the complexity of these crimes. Proceeds of crime in Canada are mainly generated from human- and, illicit drug trafficking, and various types of fraud including payment card fraud, mass marketing fraud and mortgage fraud (FATF, 2016a). Further, money laundering in Canada has also been found to be associated with organized criminal organizations (FATF, 2016a).

This chapter provides an overview of the nature and extent of money laundering visible to police and courts, including insight into the trend of police-reported crimes and the resulting court outcomes. Due to the complexity and hidden nature of ML, the number of ML incidents in official crime and court statistics is not representative of its true pervasiveness. Regardless, these indicators are still useful to monitor the impact of changes in policy or operations geared towards greater detection and prosecution of money laundering in Canada.

The chapter is divided into two main sections. Section 2.2 examines the characteristics of ML incidents[6] and the characteristics of those accused of ML, between 2009 and 2018. This is done with data from the Uniform Crime Reporting (UCR) survey conducted by the Centre for Justice and Community Safety Statistics at Statistics Canada (CCJCSS), which collects incident based data from all police services across the country. Although many incidents of money laundering go undetected by authorities, police-reported data provide important metrics and information on the nature of those incidents that do come to the attention of police in Canada.

Section 2.3 examines the court outcomes of ML incidents reported by police, by joining the UCR data with data from the Integrated Criminal Court Survey (ICCS), which provides information on completed cases in Canada's criminal courts. For more details on the data and methods, see Box 1-1.

[6] The definition of a money laundering incident for the purposes of this chapter is any police-reported incident that had a money laundering violation whether the most serious violation or not.

2.1 KEY POINTS

- The number of ML incidents reported by police in Canada slowly declined over a 10-year period, reaching an all-time low in 2018.
- Police were able to identify an accused in 6 in 10 money laundering incidents, but this proportion varied on whether or not the ML offence was committed in combination with other offences.
- Where an accused was identified, nearly half had been involved in other criminal incidents in the preceding two years.
- In examining the court outcomes of ML offences, it was found that many police-reported incidents involving money laundering did not result in an ML charge in Canada's criminal courts. Furthermore, the ML charges that did make it to court were stayed, withdrawn, dismissed or discharged in just over two-thirds of cases.

2.2 POLICE-REPORTED MONEY LAUNDERING INCIDENTS IN CANADA, 2009 TO 2018

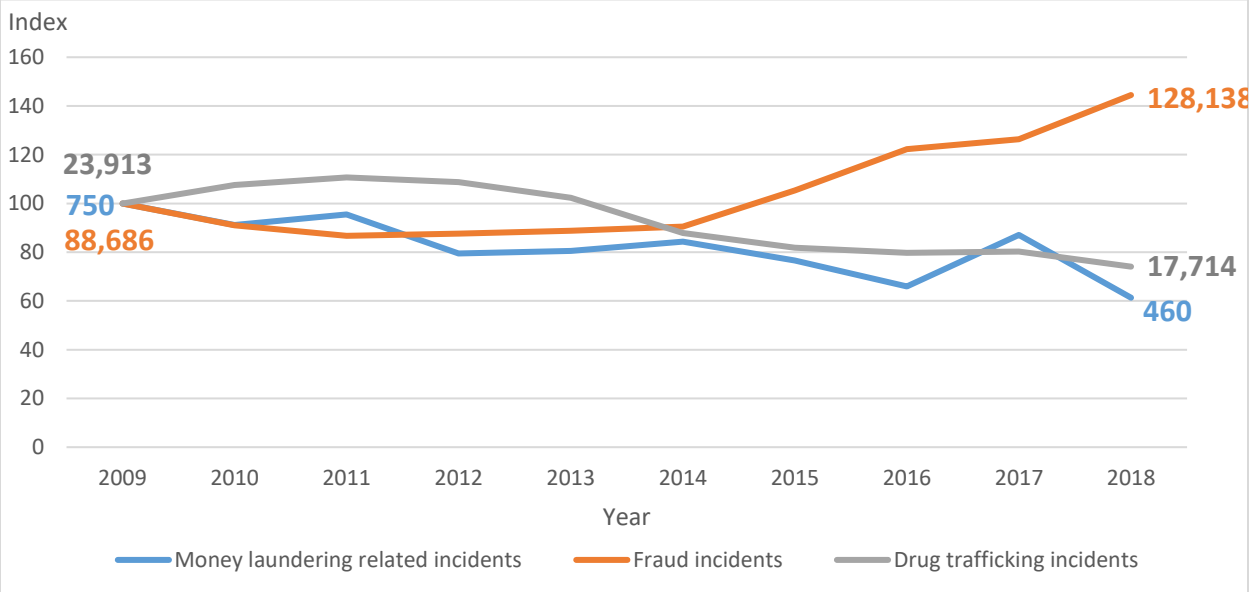
2.2.1 Overview of police-reported incidents of money laundering

Between 2009 and 2018, police services in Canada reported a total of 6,165 money laundering incidents[7]. Overall, the number of ML incidents reported by police decreased over the 10-year period from a high of 750 incidents in 2009 to a low of 460 incidents in 2018, with fluctuations year over year (Figure 2-1).

Comparisons of trends for police-reported incidents of fraud and drug trafficking, crimes generally associated with money laundering over the same 10-year period, showed the number of fraud incidents steadily increasing from 2014 to 2018, while reported incident of drug trafficking declined from 2012 to 2015, and remaining somewhat stable thereafter (Figure 2-1). The decrease in the number of police-reported incidents is at least partially coincident with the reorganization and reallocation of financial crime resources within the RCMP.

[7] Any police-reported incident that involved a money laundering charge covered by sections 462.31 and 462.33 of the Criminal Code. This can include an incident where money laundering was the most serious charge or was not the most serious charge.

Figure 2-1: Indexed fraud, drug trafficking and money laundering incidents, 2009 to 2018, Canada



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey.
Note: Incidents have been indexed to the number of incidents in 2009.

BOX 2.1: DATA SOURCES AND METHODOLOGY

Uniform Crime Reporting (UCR) Survey Trend file, 2009-2018

The Uniform Crime Reporting (UCR) survey collects essential incident-based data on the nature and extent of crime in Canada. Incidents reported to the UCR survey consist of a set of connected events which usually constitute an occurrence report and may contain one or more violations of the law. These may be grouped into the same unique incident if and only if they are committed by the same person or groups of persons and if they occur at the same time and place.

Counts from the UCR trend file presented in this chapter are based upon all incidents with at least one ML offence in the incident. Coverage of the UCR trend file between 2009 and 2018 is estimated at 99% of the population of Canada and includes only those police services who have consistently responded in order to allow for comparisons over time.

An incident can contain one to four violation codes. The most serious violation as determined by the maximum penalty, is referred as the "most serious violation" (MSV) within an incident and when reported, the MSV is always the first violation code listed on record. The MSV is used in the annual publication of Crime Statistics in Canada.

For the purpose of the chapter, 'money laundering incidents' included police-reported incidents that had ML charges as any violation in the incident, not just the MSV.

Investigators in the field do not always have the time to score their investigations appropriately in the record management system which feeds into the UCR, because they primarily work on a different records system which better suits their purposes. Adding the data to the UCR-linked system is an extra step which may sometimes be neglected or done less judiciously.

Integrated Criminal Court Survey (ICCS), 2009/2010-2016/2017

The Integrated Criminal Court Survey (ICCS) collects statistical information on adult and youth court cases involving *Criminal Code* and other federal statute offences in Canadian courts, and their characteristics.

The ICCS collects appearance records and constructs two main units of count including: completed charges and completed cases (the primary unit of analysis). A completed charge refers to a formal accusation against an accused person that has been processed by the courts and received a final decision. A completed case is one or more charges against an accused person, which were processed by the courts at the same time and received a final decision.

In cases with more than one charge it is necessary to determine the charge that will represent the case for tabulation purposes. A completed case that has more than one charge is represented by the charge with the "most serious offence" (MSO) and used to describe the characteristics of the case. The most serious offence is determined using rules involving court decisions, offence seriousness and offence sentences.

Linkage of UCR and ICCS records, 2009-2016

The UCR and the ICCS contain personal identifier information such as sex, Soundex (phonetic algorithm that translates name into an alphanumeric code), and date of birth which allows for

records to be linked from each source. The linking of records from both sources will provide the court outcomes of ML incidents.

All accused individuals that had charges laid in relation to an ML incident were included in the linkage, with the expectation that the laying of charges would have resulted in a court case. A hierarchical deterministic linkage was used to link accused between both data sources using multiple rules involving personal identifiers (sex, Soundex, date of birth), offence dates and geographic information. The linkage rate was 65% in which a total of 3,241 out of 4,978 accused individuals linked to 2,962 completed court cases.

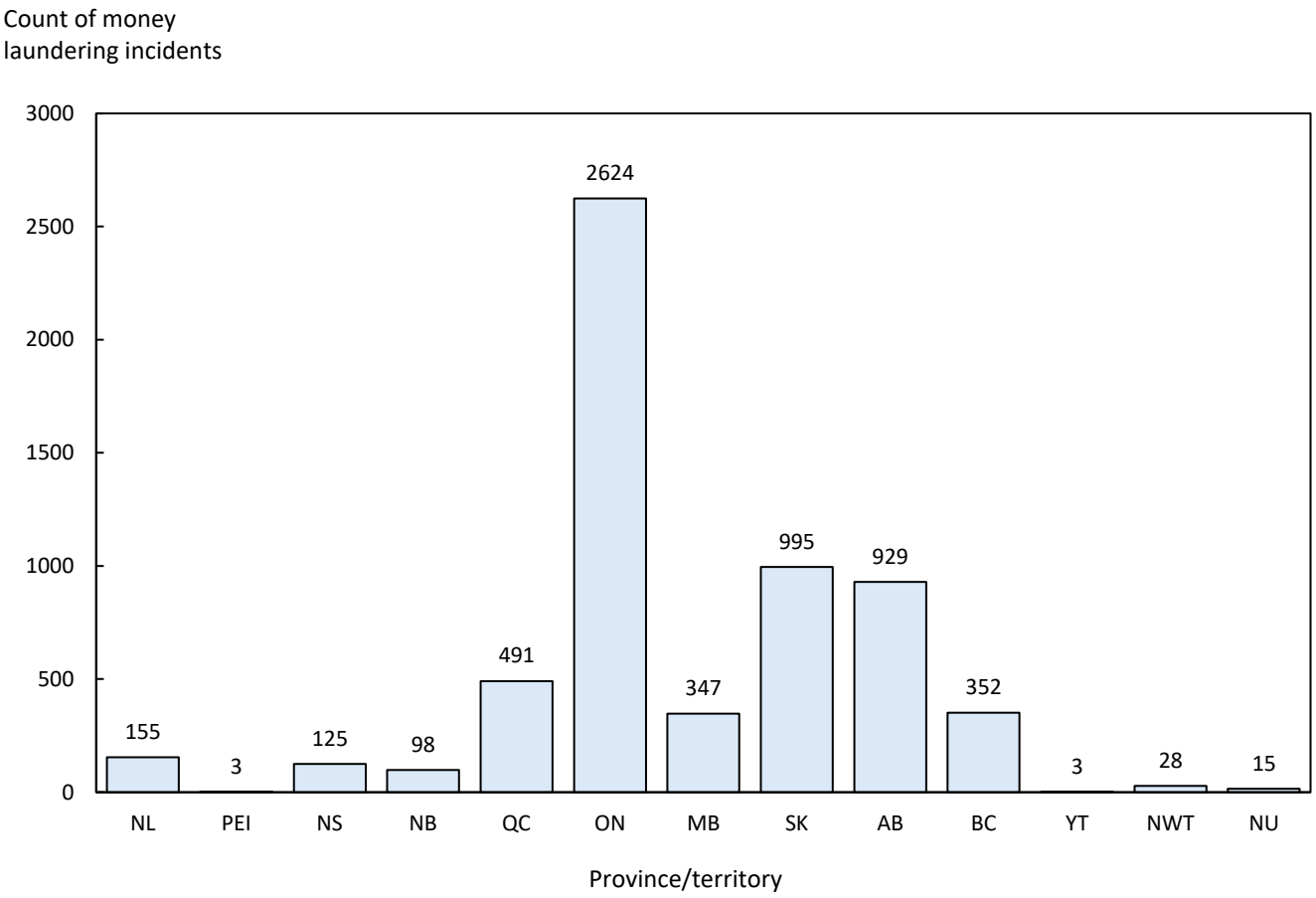
2.2.2 Provincial rates of money laundering incidents

Almost half (43%) of all money laundering incidents reported by police between 2009 and 2018 occurred in Ontario, which is correlated with its population size relative to other jurisdictions (**Error! Not a valid bookmark self-reference.**). After Ontario, Saskatchewan and Alberta accounted for 16% and 15% of ML incidents, respectively. British Columbia accounted for only 6% of the ML incidents reported to police between 2009 and 2018, roughly half of its relative population weight.

However, when taking population size into account, Saskatchewan reported the highest rate of ML incidents, with a rate of 9.24 incidents per 100,000 population (Figure 2-3).

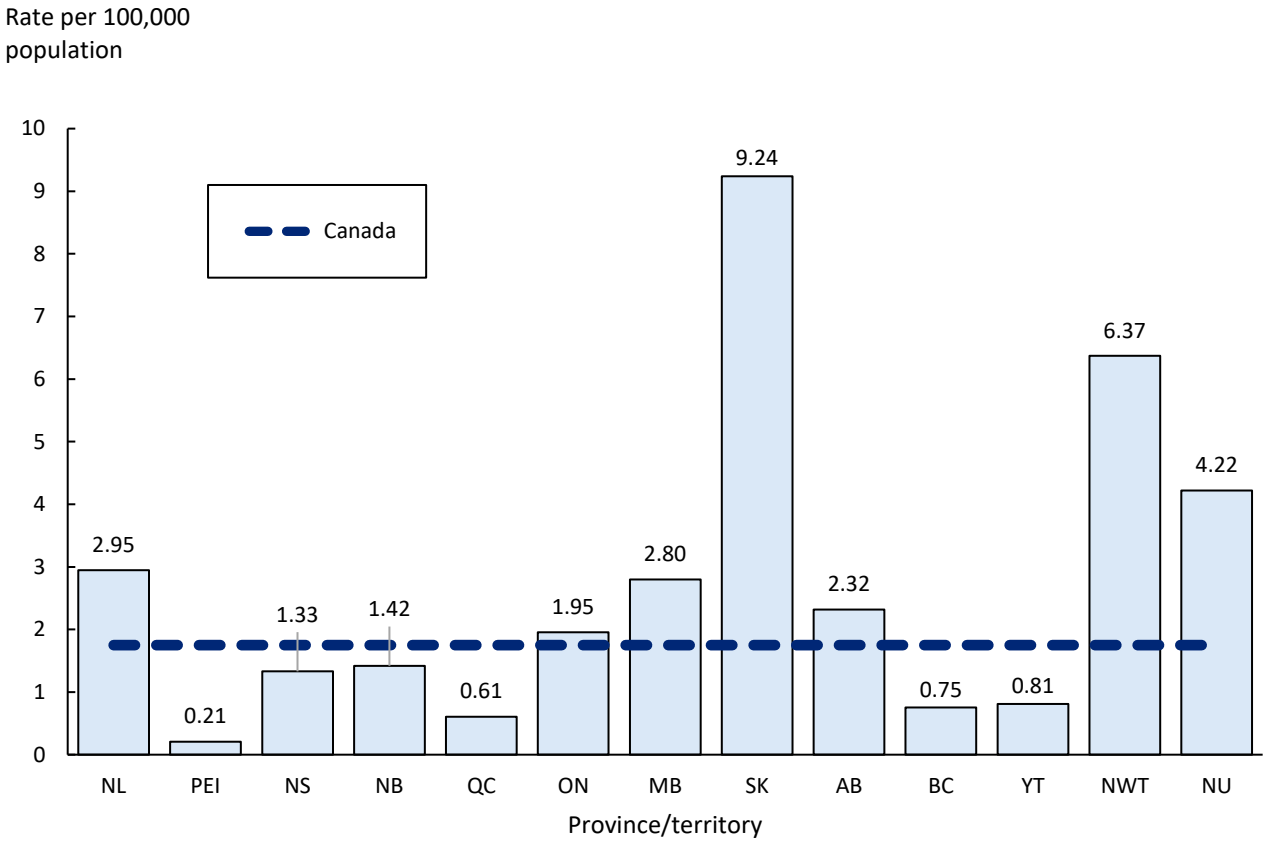
As these are police-reported data, counts may vary depending on operational differences such as resources to proceed with ML investigations. In addition, in British Columbia police do not lay charges without Crown approval, which may impact the number of police-reported incidents, especially relative to other provinces where Crown approval is not required to lay a charge.

Figure 2-2: Number of police-reported money laundering incidents, by province/territory, 2009 to 2018



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey.

Figure 2-3: Rate of money laundering incidents per 100,000 population, by province/territory, 2009 to 2018



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey.

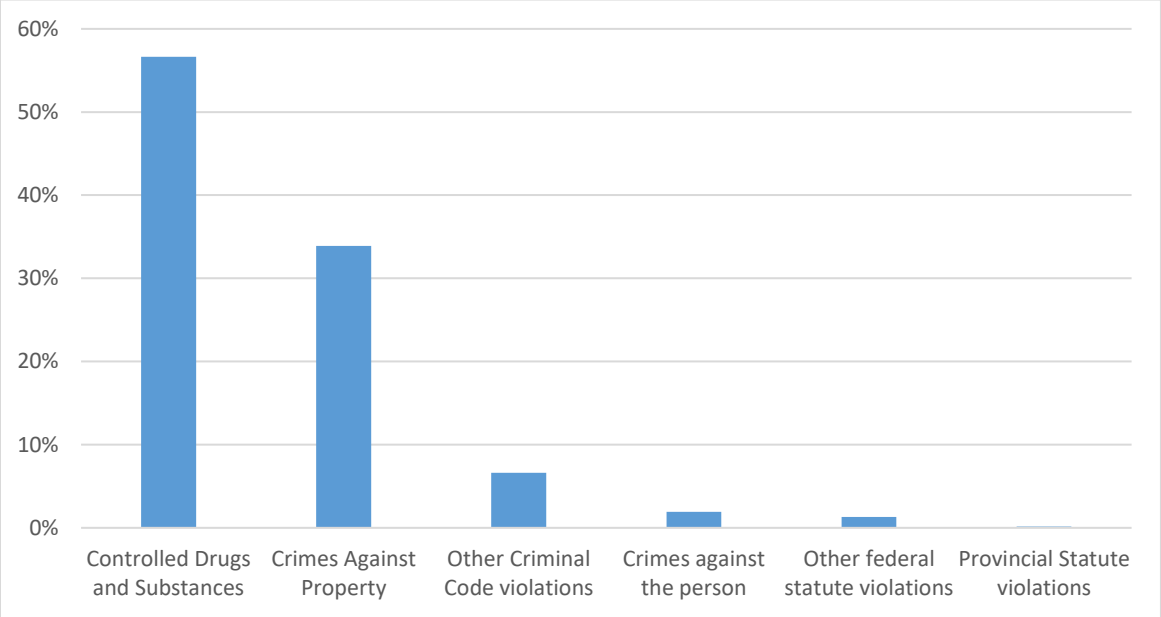
2.2.3 Money laundering offences in conjunction with other *Criminal Code* offences

Money laundering describes a variety of activities that seek to integrate funds from criminal activities into the legitimate economy by obscuring its origins. Data included in police records may not always reflect other associated crimes and this must be taken into account when considering a data framework to inform the nature and extent of money laundering. In the UCR survey, police can report up to four offences per incident. This reflects the reality that criminal offences are rarely carried out in isolation, and that multiple offences can be committed by the same person in the same time period (for more information, see Box 1-1: Data Sources and Methodology). Money laundering incidents with a single violation (i.e., where ML was the only violation) accounted for 43% of all ML incidents reported by police between 2009 and 2018, while the remaining 57% of ML incidents included additional offences, such as drug offences.

The additional crimes included with money laundering offences were largely trafficking and drug possession, followed by crime against property, mainly consisting of fraud (Figure 2-4). Other offences related to *Criminal code* violations were also noted, namely involving administration of justice offences such as failing to comply with bail conditions and breach of probation conditions.

In contrast, when looking only at the 352 UCR incidents reported by British Columbia law enforcement, 72% included only an ML violation with no other associated violation. For the remaining 98 incidents, the most common associated violation was a drug offence (45%), followed by a property crime (17%).

Figure 2-4: Associated offences of money laundering incidents, by offence categories, Canada, 2009 to 2018



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey.

2.2.4 Identification and clearing of money laundering incidents by police

Of the 6,165 money laundering incidents reported by police between 2009 and 2018, 60% of incidents resulted in law enforcement identifying an accused. This was higher than all police-reported incidents, in which police were able to identify an accused in 44% of incidents.

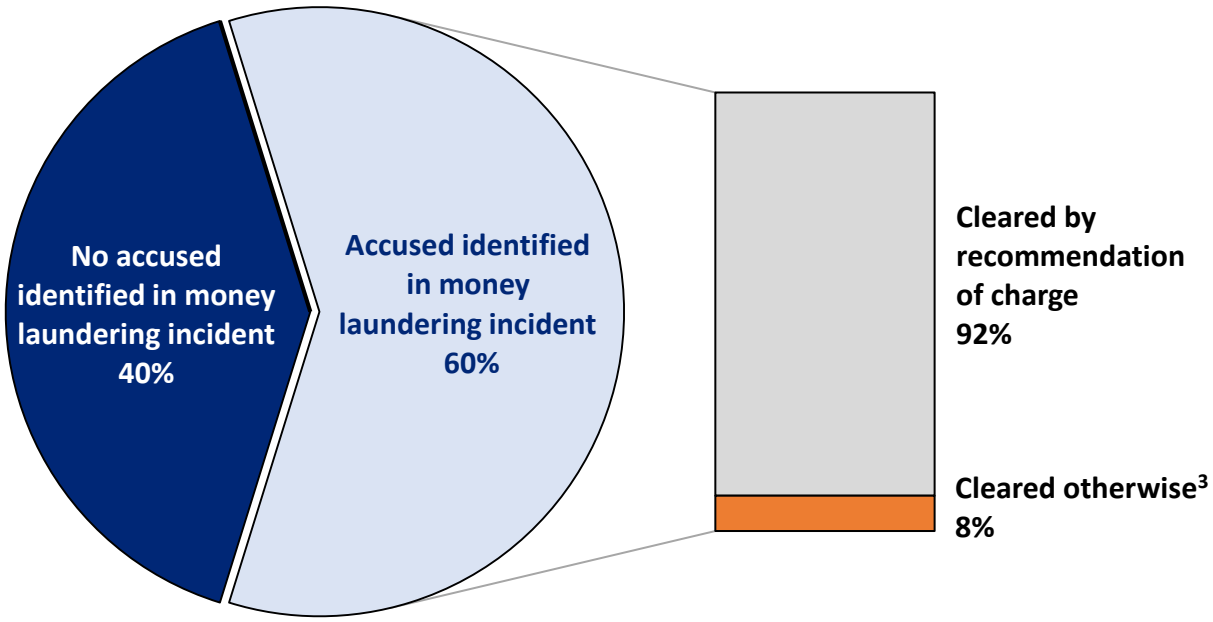
Money laundering incidents generally had an accused identified when the incident involved other violations. When an ML incident was devoid of any other offence, an accused was identified in only 15% of incidents. In contrast, ML incidents containing other offences identified an accused in almost all instances (94%), which highlights the often hidden, even faceless nature of ML.

Police services in Canada have discretion on how to process an incident when an accused is identified in relation to a criminal offence. They can formally charge or recommend charges against the accused. Alternatively, they may clear the incident by other means.[8] The majority (92%) of ML incidents between 2009 and 2018 with an identified accused resulted in charges being laid or recommended to be laid

[8] The following are reasons for clearing an incident: death of the accused, suicide of the accused, death of the complainant/witness, reason beyond the control of the department, diplomatic immunity, an accused under the age of 12, accused committed to a mental health facility, the accused is outside Canada, victim/complainant request that no further action is taken, the accused was involved in other incidents, the accused was already serving a sentence in custody, departmental discretion, diversionary programs, cleared by a lesser statute, cleared by other municipal/provincial/federal agency.

against the accused[9] (Figure 2-5). In contrast, among all incidents that were cleared by police, only 63% were cleared by the recommendation of a charge.

Figure 2-5: Money laundering incidents by identification of accused and clearance status, 2009 to 2018



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey.

2.2.5 Characteristics of the accused in money laundering incidents

A total of 6,355 accused[10] were identified from 3,666 ML incidents[11]. Similar to police-reported crime overall, the majority of individuals accused of money laundering were men (77%), while women comprised 23% of the accused.

The majority of those accused of money laundering (97%) were adults aged 18 years and older, while youth aged 12 to 17 years accounted for the balance of the accused. The median age of those accused of ML incidents was 29, the same median age of accused involved in all UCR incidents.

Police-reported data indicated that individuals accused of money laundering do not work in isolation, and often work in collaboration with other individuals in the commission of these offences. Of the 6,355 accused individuals identified by police, more than two-thirds (67%) worked in collaboration with at least one other person. Women were more likely than men to be involved with a co-accused, as a higher

[9] Charges laid or recommended can be for any violation on a money laundering incident.
[10] Accused involved in multiple incidents, are counted multiple times as one accused for each separate incident. In 40% of incidents, no accused was identified.
[11] Money laundering incidents linked to accused information that had unknown sex information or identified as a company. These accused records were excluded from counts.

proportion of women (78%) were linked to an incident involving multiple accused, compared to that of men (63%). Almost all women were paired with men as only 3% of pair/group incidents involved a woman paired with another woman.

The accused's criminal histories and previous contacts with police were also examined. In total, just under half of the accused[12] (48%) had at least one prior contact with police in relation to a criminal offence in the two year period preceding their ML incident. The majority of the previous offences[13] were typically an ML predicate crime: 32% were drug offences, 32% were property offences and 11% were other criminal code offences. Previous offences unrelated to money laundering accounted for almost one quarter of offences and were related to crimes against the person (17%) and traffic violations (5%).

2.3 COURT OUTCOMES OF MONEY LAUNDERING INCIDENTS

The process of money laundering moving from a police charge to a court case is complicated and impacted by elements which are particular to each case. Investigations into ML offences are often at the discretion of police officers and can be influenced by a number of factors such as the priorities, resources, and enforcement practices of a police service, and by court requests for financial records (FATF, 2016a). The following section provides information on the court outcome (decisions and sentencing) of accused persons in police-reported incidents who were successfully linked to court cases.

Between 2009 and 2016, police reported a total of 4,978 accused who were involved in an ML incident that resulted in the laying or recommendation of charges.[14] Of the 4,978 persons that were involved in an ML incident, 3,241 were linked to a completed court record (a linkage rate of 65%), leading to 2,962 linked completed court cases[15]. Completed court cases are shown on the left hand side of Figure 2-6.

Figure 2-6 shows the outcomes of these 2,962 cases. It divides them into three groups:

1. cases where no ML charges were laid (the blue path);
2. cases where ML charges were laid as well as more serious charges (the orange path);
3. cases where ML charges were laid as the most serious charge (the grey path).

For each of these groups, the chart shows the number of cases which resulted in a finding of guilt, and the number in which a custodial sentence was issued.

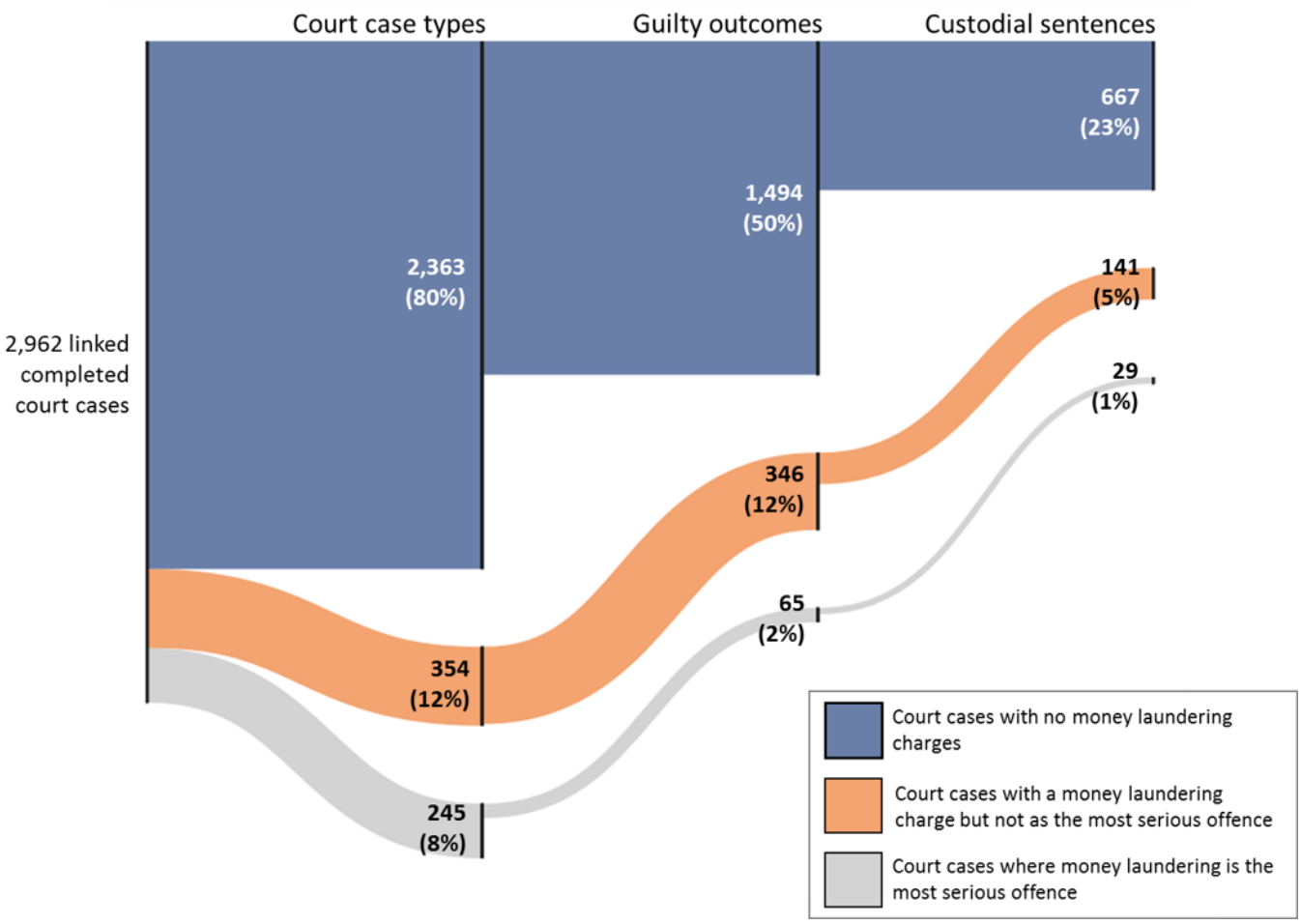
Figure 2-7: Court outcomes of accused involved in money laundering incidents, 2009 to 2016

[12] Excludes 6 accused who had a unknown date of birth

[13] Previous offences include the most serious violation from the incident most recent to the money laundering incident. Offences related to the administration of justice have been excluded. Links to previous offences for accused in 2009 and 2010 were less likely as full incident information for previous the previous 2 years was not available.

[14] Incidents that were cleared by laying charges involving multiple accused may not have charges laid against all accused based on the discretion of the police.

[15] Persons may be counted multiple times in the incidence counts when involved in multiple incidents, but they resulting charges may have been combined into a single court case. This is why the number of persons linked to court cases (3,241) is greater than the number of court cases (2,962).



Source: Statistics Canada, Canadian Centre for Justice and Community Safety Statistics, Uniform Crime Reporting Survey, Integrated Criminal Courts Survey

2.3.1 Accused involved in police-reported money laundering incidents unlikely to see money laundering charges against them in court

The integrated data reveal that, while many accused were involved in money laundering, 80% of resulting cases faced no ML charges in their court case, although these cases included other charges (the blue path in **Error! Reference source not found.**). As noted in section 1.2.3, 57% of police-reported incidents of money laundering involved other offences, with drug-related offenses being the most common.

Of the 80% of cases without an ML charge, 1,494 (or almost two thirds) had a guilty outcome and 667 (or almost one third) were sentenced to custody. The low counts of ML charges proceeding in Canada's criminal courts may be related to the complexity of prosecuting ML offences. In fact, the *Criminal Code* has recently been amended to facilitate the prosecution of money laundering (*Criminal Code,* 1985 as amended 2019-06-21). The future assessment of the impact of this change may help to reveal whether

the difficulty in prosecuting an ML offence contributed to the low number of ML charges associated with police-reported incidents.

Just over 1 in 10 (12%) of cases linked to police-reported ML incidents included an ML charge in court, along with other more serious offences, such as fraud or drug trafficking (the orange path in **Error! Reference source not found.**). Of these cases, almost all (346) had a guilty verdict and 40% were sentenced to custody. It is notable, however, that of the 354 court cases that contained an ML charge and other more serious charges, the ML charge was upheld in only 0.3% of cases. In nearly all the cases (99.7%), the ML charges were acquitted, stayed, withdrawn, dismissed or discharged. This implies that, while the charges related to money laundering were dropped, or the accused was found not guilty, almost all cases (98% or 346 cases) did result in a finding of guilt for charges deemed more serious, such as fraud or drug trafficking. ML charges are often dropped as part of a plea bargain in part because both parties are incentivized to not pursue this angle; prosecutors face a more difficult burden of proof on money laundering charges and money launderers are in a better position not to forfeit their assets when ML charges are dropped.

Finally, 8% of the linked cases had a money laundering charge as the most serious offence in the court case.[16] Among these 245 cases, 27% or 65 cases resulted in a finding of guilt for the ML charge (representing 2% of all completed court cases linked to ML, the grey path in **Error! Reference source not found.**). This is lower than the average for all court cases, where 63% resulted in a finding of guilt. However, in these cases, the accused was more likely to be given a custodial sentence, than the average for accused persons in a criminal court case. Of the 65 ML cases that resulted in a finding of guilt, 45% of accused were sentenced to custody. In comparison, among court cases overall, 34% were sentenced to custody.

Overall, of all the cases that were completed in court, 64% resulted in guilty outcomes and 29% resulted in a sentence to custody which is comparable to the rates for all court cases (where 63% result in a finding of guilt and 34% result in custody).

2.4 PROPERTY SEIZURES

A supplemental analysis of property seizure data provided by Public Services and Procurement Canada shows a similar use of PCMLTFA provisions in enforcement. Fewer than ten residential properties have been seized between 2009 and 2019 in cases related to the PCMLTFA. More than half of these properties were the result of a single enforcement action.

As with the court cases, it is possible that more properties are seized in cases related to other sections of the criminal code.

[16] The most serious offence is determined using rules involving court decisions, offence seriousness and offence sentences.

3 A CONCEPTUAL MODEL FOR MONEY LAUNDERING IN THE REAL ESTATE SECTOR

Money laundering describes a variety of activities that integrates funds from criminal activities into the legitimate economy by obscuring its origins. The methods used to launder funds are complex and ever evolving, often incorporating technological advances to muddle the trail that connects the proceeds back to criminal activities. In the residential real estate sector, money laundering operations are varied and range from simple schemes, such as the quick purchase and sale of a property, to more intricate schemes involving the collusion of real estate professionals such as lawyers and real estate agents. This chapter presents a conceptual model for ML in the real estate sector, detailing the stages separating predicate crimes and laundered funds.

The real estate sector is an ideal medium for money laundering as transactions are generally of large value, a property is an investment that typically retains or gains value over time, and a property can be used as a physical base for further criminal activities.

Recent studies, including Maloney et al. (2019) and German (2019), concluded that more needs to be done in terms of red flag analysis to distinguish between legitimate real estate activity and money laundering real estate activities. In that context, this chapter employs the theoretical understanding of the stages of money laundering to outline a data framework which would enable such a red flag analysis.

3.1 KEY POINTS

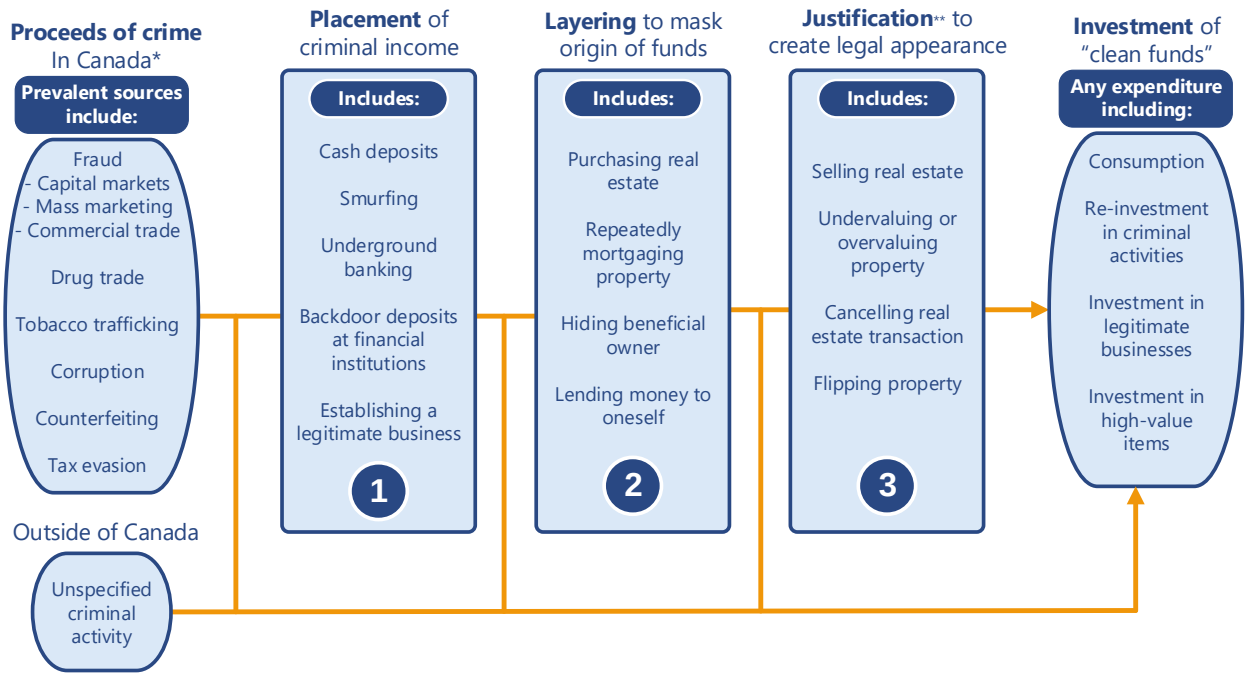
- Money laundering proceeds in four stages from the generation of the criminal proceeds: placement, layering, justification, and investment/reintegration.
- Two stages of ML are in scope for this report: layering and justification, as these stages involve real estate transactions and are potentially identifiable via strategic data linkage and analysis as money laundering through real estate involves unusual owner, property, and mortgage characteristics.
  An extensive database with complex indicators can be developed by considering money laundering in real estate as a multi-dimensional issue spanning multiple data sources.

3.2 THE STAGES OF MONEY LAUNDERING

This chapter presents the four-stage version of the well-known stages of money laundering model. A three-stage model is generally employed in Canada (Levi, 2013), but the addition of the “Justification” aids in explaining some of the more overt maneuvers that money launderers undertake (van Koningsveld, 2013).

Figure 3-1 presents a visual overview of the stages of ML and provides a representation of the activities at each stage. Stages 2 and 3 (layering and justification) are the stages where housing transactions principally occur and can be indicated with data, as indicated in Appendix 1: Data points and potential sources.

Figure 3-1: Conceptual framework of money laundering in real estate



***(FATF, 2016a). ** (van Koningsveld, 2013)**

3.2.1 Stage 0: Sources of proceeds of crime

A money laundering scheme is preceded by an accumulation of funds through criminal or illegal endeavours referred to as predicate offences. The most common predicate offences associated with ML in Canada are fraud, corruption and bribery, counterfeiting and piracy, illicit drug trade, and tax evasion (FATF, 2016a). Fraud in particular encompasses many different types, including capital markets fraud, mass marketing fraud, commercial trade fraud and mortgage fraud. In addition to these domestic sources, significant levels of foreign proceeds are also laundered in Canada, in part through the Canadian real estate market (Maloney et al., 2019).

The identification of data sources for this stage is not in the scope for this report. These predicate offences are, by nature of being illegal activities, covert and difficult to measure. Furthermore, the methodology of laundering is not tied to the type of predicate offence, and identification via data at this stage does not provide insights into laundering activities in the real estate sector.

3.2.2 Stage 1: Placement of criminal income

The placement stage initiates the money laundering process. This can be accomplished by placing the illegal money into the legitimate economy, from which it can be transferred elsewhere and used in many types of transactions.

Financial institutions are commonly at the forefront of the placement stage. The seemingly innocuous task of placing money into legitimate financial systems can be challenging for money launderers due to

the large amounts of cash that can be generated through their illegal activities, especially as legislation and policies are designed to prevent and expose the entry of the proceeds of crime into the financial system.

Lawyers can be similarly involved in this stage due to their ability to accept large amounts of cash in exchange for their services (German, 2019, p. 132), and can provide trust accounts where money launderers can park their cash.[17]

Similar to Stage 0, the identification of data sources for this stage of the ML process is not in scope for this report as this stage does not yet entail a real estate transaction. Money laundering efforts relative to this stage typically involve external players, such underground or unregulated banking networks, whose involvement is difficult to identify and who may not be willing to share data pertinent to money laundering.

3.2.3 Stage 2: Layering to mask the origins of funds

The second stage of the ML flow aims to conceal the origin of the funds by adding layers of financial maneuvers. This is one of the two stages in scope for this report as layering can be identified via red flag analysis and can be identified with existing data.

Layering involves a multitude of transactions through various channels such as electronic funds transfers, withdrawals and deposits into other accounts, trading in financial markets and purchasing high value items like precious metals, luxury cars and real estate. The transfers often move the funds between multiple jurisdictions in order to frustrate any investigation. This is particularly true when lawyers facilitate these maneuvers as it is more difficult to obtain information on transactions moved through a lawyer's trust account.

For the real estate sector, the purchase of a property to conceal criminal income belongs to the layering stage since it involves a substantial financial transaction with the ability to conceal the origin of the funds. Purchases can be enabled or supplemented with a mortgage, which can add layers, complexity, lowers the chances of detection and provides another vehicle to absorb the proceeds of crime. One strategy involves repeatedly mortgaging the same property to continue laundering activities after the initial loan is paid off. Additional layers are often added by using a formal or informal nominee owner, or by lending money to oneself through a web of financial transactions, and then paying these back at a regular rate with additional proceeds of crime.

3.2.4 Stage 3: Justification to create a legal appearance

The justification stage of the ML process was proposed by van Koningsveld (2013) in order to introduce the key component of creating the semblance of legally obtained proceeds. While not every ML scheme will include this stage, it helps avoid detection with a veneer of legitimacy. As with the layering stage, this stage involves various real estate transactions and the strategies of justification can also be made visible within the data. As such, it is in scope for this report.

In the real estate sector, the sale of a property provides a legitimate explanation for recently acquired funds, so a property can be purchased one day as part of the placement and layering stages, and sold

[17] Law societies, such as the Law Society of BC have rules curtailing lawyer's exploitation of these loopholes, but these do not stop lawyers who willfully aid their clients in breaking the law, or in circumventing the reporting requirements imposed on all other financial intermediaries.

the next day to create a justification for their funds. On paper, the funds received by the seller appear legitimate. While this is a rudimentary method, quick resales can be detected through data. More creative schemes may also be used and identified, such as flipping a property repeatedly or cancelling a real estate transaction in order to receive a deposit or cheque refund from the lawyer or broker involved.

3.3 DATA FLOW OF MONEY LAUNDERING IN THE REAL ESTATE SECTOR

Money laundering transactions made through the layering and justification stages will often have distinctive characteristics that can be flagged through the use of various data sources. A real estate from stages 2 and 3 leads to data that can be collected and analyzed on both legitimate transactions, and those involving ML schemes. Five different dimensions of data are required in order to identify the suspicious combinations of characteristics produced by ML schemes:

1. Buyer characteristics;
2. Seller characteristics;
3. Financing characteristics;
4. Property level data; and
5. Auxiliary information.

In combination, these can comprise a comprehensive collection of data capable of identifying and estimating the prevalence of money laundering in the Canadian real estate sector through analyses of the congruence of characteristics. Such a collection of data could be used for more accurate red flag analysis, as well as in more sophisticated analyses of the prevalence and character of money laundering real estate markets.

The following sections briefly illustrate the pertinent data points of the first four categories. The fifth category, auxiliary information, is comprised of reference materials which brings context to the characteristics in the other categories.

For a more detailed list of the data points included in the data framework, see Appendix 1.

3.3.1 Buyer characteristics

Buyers are either natural persons or legal persons[18] which purchase real estate, and both have data points of interest that can be used to identify characteristics of a person or business involved in real estate transactions related to money laundering.

Pertinent characteristics for individuals include income and citizenship, particularly if the country of origin is listed as a high-risk jurisdiction by the Financial Action Task Force (FATF). The characteristics of spouses and close relatives are also of interest, as many launderers aim to hide the beneficial ownership through trusted individuals, such as their partner or child.

For legal persons, pertinent information relates to the business structure and activities, such as the beneficial owner(s), country of registration, date of incorporation, revenues, and number of employees.

[18] Legal arrangements, such as trusts, are included under this term.

Low levels of activity and multiple levels of ownership can, for example, indicate that a given legal person is a shell corporation intended to create layers between the beneficial owner and the property.

A full mapping of owners to properties is also essential, as an owner might possess multiple properties, and a property can have multiple owners. It is also important to understand the total real estate holdings of each owner, and the linkages between co-owners.

3.3.2 Seller characteristics

Sellers may be easily overlooked when considering money laundering in the real estate sector but some schemes rely on the seller being a co-conspirator. Seller characteristics of interest are similar to the buyer's characteristics listed above, namely income and citizenship for natural persons, and ownership structure and activity for legal persons. For legal persons selling property, information on the complete ownership structure is necessary to, for example, identify when the buying and selling entities are controlled by the same legal person.

3.3.3 Financing characteristics

The arrangements for financing a real estate purchase can form a central part of the justification step of the ML process. As such, mortgage data are a key component of a money laundering data framework.

Mortgages allow for the purchase of a large property by paying back loaned funds over an extended period with dirty funds. The smaller amounts and regularity of the payments make them more difficult to detect as part of an ML scheme. Cautious launderers may so closely mimic regular buyers that careful data analysis would be required to discern them. However, several schemes rely on uncommon mortgage arrangements including unusually high interest rates, accelerated repayments, multiple or repeated mortgages, or loans provided by unregulated or previously unknown lenders. Pertinent mortgage characteristics thus include the interest rate, loan-to-value ratio, term length, scheduled and actual repayments, and the identity and characteristics of the lender.

The Financial Transactions and Reports Analysis Centre of Canada (FINTRAC) collects a wealth of information on the financial transactions of individuals and businesses in Canada. For the purposes of money laundering, it is of interest to discern buyers who have a history of suspicious transactions, namely by being the subject of a Suspicious Transaction Report (STR). Additionally, information available through Large Cash Transaction Reports (LCTR) and Electronic Funds Transfer Reports (EFTR) could also be pertinent, including the sender, recipient, country of origin, and destination country.

3.3.4 Property characteristics

Pertinent property characteristics include the type, location, history of sales dates and prices, and current use. A typical home buyer may consider real estate purchases as a mid- to long-term investment, having more potential for capital gains as the years go by. This contrasts with those purchasing as a part of an ML scheme, who are often motivated to clean dirty funds as efficiently as possible, leaving various clues such as repeated sales, and purchases at artificially inflated or deflated sales prices. Moreover, properties purchased as a part of a money laundering scheme may be vacant for a prolonged amount of time.

3.4 FINAL RESULT: INVESTMENT OF CLEAN FUNDS

The goal of any money laundering strategy is the use of seemingly clean funds at liberty in the economy. Once the money appears legitimate, it can be spent like any legally acquired income on everyday consumption, or on more extravagant purchases like jewelry, vehicles and boats, as well as investments in legitimate businesses, the stock and bond market or the real estate industry. The money can also fund further criminal activities, especially if those require purchases in the legal economy, such as the purchase of a property intended as a base for an unlicensed cannabis growing operation or for lawyer and bail fees (German, 2019). This stage is also called the integration stage.

Most of the investment of seemingly clean funds is out of the scope of the report as the money can be spent as legitimate income in any country and might not include a real estate transaction.

4 TYPOLOGIES FOR MONEY LAUNDERING ACTIVITIES IN THE REAL ESTATE SECTOR

This chapter presents the typologies for money laundering activities in the real estate sector. In the AML context, the term "typologies" refers to the techniques used to launder money (IMF, 2011). Money laundering typologies can be studied to derive a comprehensive list of indicators of ML activities. Indicators, in turn, inform the collection of data necessary for constructing a system that can be used to detect money laundering in real estate.

4.1 KEY POINTS

- Money laundering typologies are the approaches and tactics employed to launder money.
- AML efforts are clarified by clearly understanding ML schemes and sub-schemes, and subsequently deriving how these can be detected via indicators and required data.
- A list of schemes, sub-schemes and indicators of money laundering in the real estate sector is enumerated and will inform the choice of data points required in the data framework.

4.2 THE STRUCTURE OF THE DATA FRAMEWORK

There are many listings of ML typologies in real estate available, although these can vary widely in their level of specificity (Irwin et al., 2011). The framework in this report incorporates and expands upon existing typologies and indicators from the literature[19], and organizes them into a multi-level framework.

Each typology is organized into a hierarchical structure with four levels: schemes, sub-schemes, indicators, and data points:

1. A **scheme** encompasses a general goal or key ML mechanism.
2. A scheme is organized into one or more **sub-schemes**. A sub-scheme refers to the particular tactic employed towards the given goal, or scheme.
3. A sub-scheme is associated to one or more **indicators** which point **to** the use of a sub-scheme, allowing for its detection.
4. Each indicator is populated by **data points**, allowing for their construction and which comprise the fourth level of a typology.

For example:

1. The goal of the **scheme** in the typology to "unlawfully obscure beneficial ownership" is to conceal the true owner of a property.
2. The three **sub-schemes** for obscuring beneficial ownership are: "by using a shell company or trust", "by using a formal or informal nominee owner", and "by having an unclear address". These sub-schemes consist of different approaches, but all have the common goal of concealing the beneficial or true owner of a property. These tactics will yield distinct indicators for their detection.

[19] Chief amongst these: FINTRAC 2016, Unger et al. 2011, Schneider 2004, German 2019, and FATF 2007.

3. The sub-scheme “by using a shell company or trust” has sixteen distinct **indicators** including “owner is a company or other legal person” and “owner company was just established”. These indicators are markers pointing to the use of this sub-scheme.
4. Finally, these indicators lead to a subset of **data points** that are used in the construction of the indicator. The indicator “owner company was just established” requires data points such as the date of incorporation for the company and the date of sale of the property.

The typologies identified for money laundering in real estate are grouped into eight schemes and 22 sub-schemes. An additional 11 indicators were placed in a “miscellaneous” category. These may indicate money laundering activities, or additional grounds for suspicion, and work well in combination with the other indicators but are not derived from a particular typology. One example of such an indicator is whether the owner or buyer of a property has a criminal record.

As money laundering is continuously evolving, the data framework can be expanded to allow for new techniques and organizing of new observations into this model. Thus, the data framework can be kept evergreen as a tool for AML uses.

4.3 AML DATA FRAMEWORK SCHEMES AND SUB-SCHEMES

The top three levels of the data framework structure are presented in this section. A full list of indicators by scheme and sub-scheme can be found in Appendix 2. A full list of data points can be found in Appendix 1.

Table 4-1 lays out the full list of typologies broken down by scheme and sub-scheme. It is worth noting that these typologies are often not mutually exclusive. A money laundering tactic may, for example, involve both buying a house in the name of a nominee and at the same time attempting to use a large amount of cash in the transaction.

Table 4-1: Typologies for laundering money in the real estate market

Scheme	Sub-schemes
To unlawfully obscure beneficial ownership	…by using a shell company or trust …by using a formal or informal nominee owner …by having an unclear address
To funnel cash/money through mortgages	…by acquiring many mortgages …by repeatedly mortgaging a property …by acquiring an outsized mortgage …by lending the money to oneself …by acquiring very valuable property with a large mortgage …by taking over existing mortgages …by making mortgage payments in cash
To quickly convert dirty funds into a real estate asset	…by undervaluing and paying the difference under the table …by buying the house outright with dirty or mixed funds …by defaulting on a loan to one-self or an associate
To flip houses to turn dirty funds into clean funds	…by buying a property and selling it as soon as possible …by conspiring with the owner and selling it back to them …by renovating a property with cash/dirty funds

	…by purchasing a pre-construction condo and reselling it before the completion date …by selling an existing property on assignment
To hide capital in other jurisdictions	…by purchasing property in a country with strong property rights
To purchase property for criminal use	…by purchasing property in an area of interest
To launder money through cancelled real estate transactions	…by paying a deposit, reneging, and receiving a clean refund
To acquire an income source and mix in dirty money	…by purchasing rental property and padding rental income
Miscellaneous	Miscellaneous

The following sections explain the functioning of each scheme and sub-scheme and lists the indicators which can be used for flagging them. For the purpose of completeness all indicators are included, even in cases where currently available data does not allow for their detection.

4.3.1 Scheme: to unlawfully obscure beneficial ownership

The goal of this scheme is to place distance between the money launderer and the property, which serves two main purposes. First, it reduces the appearance of wealth and the need to explain the accumulation of wealth. Second, it reduces the risk that the asset will be lost in the event that a money launderer is caught in the course of other criminal activities. In the absence of evidence linking a crime to an owner of a house, courts are unlikely to seize the assets.

The obfuscation of beneficial ownership can also serve to circumvent restrictions or tax disincentives for purchasing multiple properties or properties outside ones country of residence.

The province of BC has implemented significant measures to collect information on the beneficial owners of real estate, namely through the Property Transfer Tax, the Speculation and Vacancy Tax, and with the upcoming Land Owner Transparency Register. These measures decrease the opacity of real estate ownership in the province by introducing a legal requirement to declare beneficial ownership. While these initiatives aid in bolstering Canada's AML regime, money launderers may continue to seek obscurity when purchasing properties and can attempt to lie about the true owners of a property. Furthermore, money launderers in other provinces face fewer or none of the transparency requirements present in BC. As such, the following sub-schemes and indicators focus on the unlawful concealment of ownership in real estate, and how it may be detected.

4.3.1.1 Sub-scheme: using a shell company or trusts

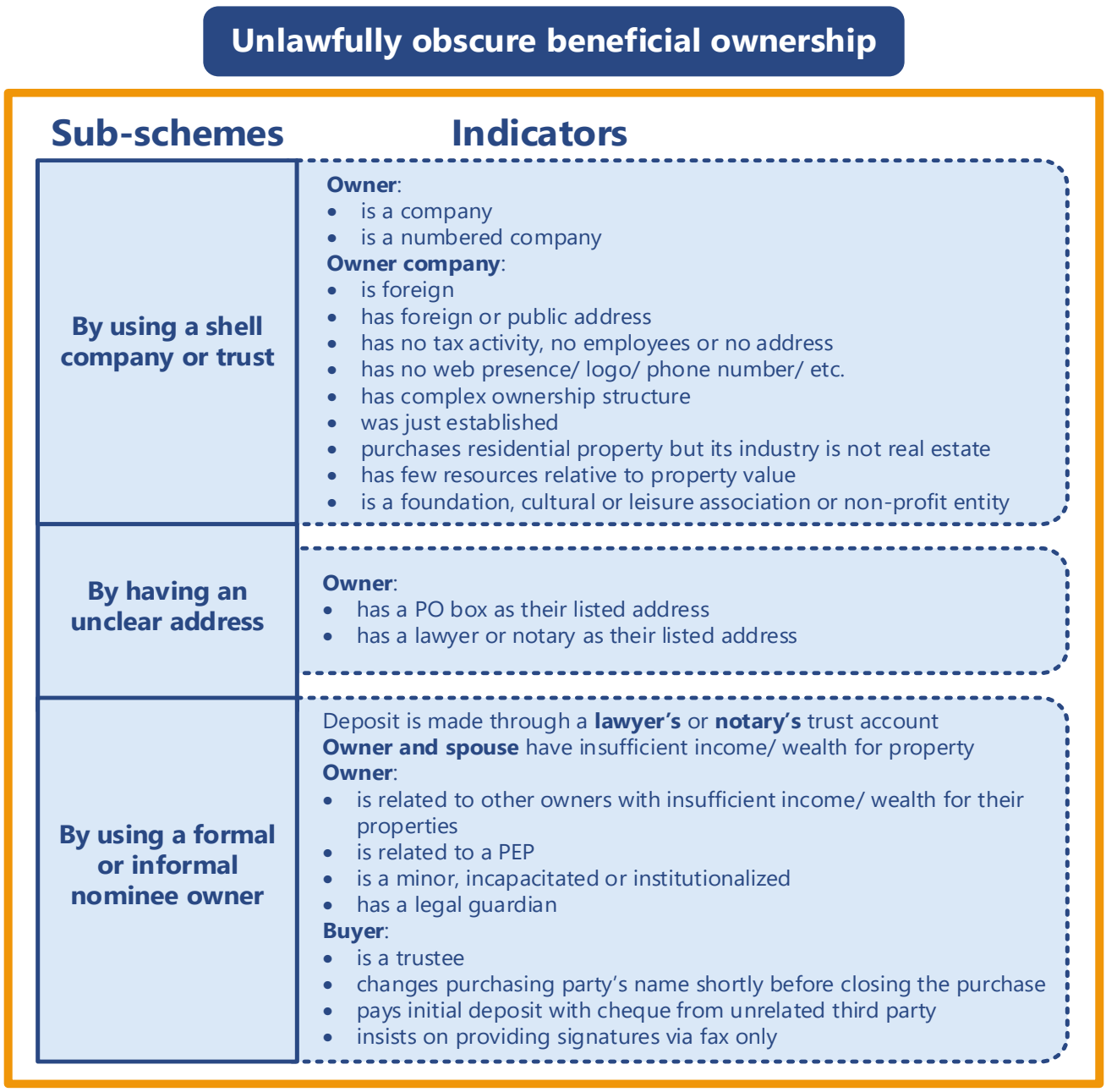
An elementary way to create distance between a person and the ownership of a property is to establish a corporation whose sole or main purpose is to hold assets acquired with criminal funds. Beneficial ownership of the corporation can be obscured through the use of nominee owners and directors or by the layering of the corporation inside a structure of other corporations. Additionally, placing the corporation in a jurisdiction which is less likely to divulge beneficial ownership information will obscure its origin, as will having the corporation issue bearer shares for ownership. These approaches can be readily combined in order to further obfuscate ownership.

In general, law enforcement agencies (LEAs) in Canada have insufficient capacity to trace ownership of legal entities. As such, cases involving such complex structures are often not pursued and assets are left unseized in other related cases (FATF, 2016a)**.**

Statements about the use of Canadian corporations can be found in the so-called "Panama Papers" leak of documents from the law firm Mossack Fonseca, which specialized in tax avoidance. In an internal memo, the managing director in Panama allegedly told colleagues that Canadian companies "are managed in a way that the administrators simply declare annually NO-ACTIVITY. In other words, they cheat a bit", and that "It is impossible for the Canada revenue governmental system to look into such information for every single company formed in Canada … This is risky, but we will try to provide the service." (Cribb & Oved, 2017). They later started offering to set up Canadian corporations for $2000 (Cribb & Oved, 2017).

Trusts can function in much the same way in that they can be used to hold assets and for moving funds.

Figure 4-1 - Summary of scheme: unlawfully obscure beneficial ownership



4.3.1.2 Sub-scheme: using a formal or informal nominee owner

Similar to shell companies, this sub-scheme adds distance between the money launderer and the property acquired. As a result, the accumulated wealth is less apparent and the properties are less prone to seizure in the event a money launderer is arrested and prosecuted. In many cases, the nature of the nominee ownership will be informal and enforced through informal means, making it difficult to link the property to the money launderer.

Typical nominee owners include a money launderer's spouse or close relatives, but associates or unsuspecting third parties with limited autonomy such as children, elderly parents, or incapacitated relatives, can also be used (Schneider, 2004; van Duyne, 2013; FINTRAC, 2016). In one example from the Netherlands, a money launderer purchased a property in the name of his unaware elderly mother, who lost her social allowance as a result of her newfound wealth (van Duyne, 2013).

Schneider (2004) analyzed RCMP cases where real estate had been seized as part of the proceedings and found that property had been registered in a shell corporation or a nominee's name in 51 out of 83 cases. In one of the cases, 23 properties were owned and used for growing marijuana and were registered with nominee owners.

Politically exposed persons (PEPs) – individuals who are entrusted with a prominent public function – often face greater scrutiny when engaging in international financial transactions due to the increased risk that wealth accrued by such individuals might be public assets acquired by virtue of their status or public office. PEPs have a strong incentive to use nominees, so as to hide the evidence of their wealth. It is worth noting that PEPs have not necessarily acquired their wealth through illegitimate means, so ownership of properties by PEPs or their relatives is not necessarily an attempt at money laundering.

The use of nominee owners can also be used to facilitate tax evasion or to circumvent foreign ownership restrictions.

4.3.1.3 Sub-scheme: having an unclear address

A money launderer may decide to have all their correspondence about a property sent to a post office (PO) box, lawyer's office, or corporate service provider (CSP). This creates some distance between the property and their activities and can be used by overseas owners who claim to be residents in Canada. This sub-scheme might also be used in conjunction with a nominee owner to ensure that the de-facto owner receives relevant legal correspondences related to the property. Their lawyer's office or CSP will know the true contact details of the beneficial owner, and so are likely to be complicit in or indifferent to the deception.

4.3.2 Scheme: funnel money through mortgages

Mortgages can serve as a means of money laundering by turning money into clean ownership of an asset. Depending on the size of the mortgage, the individual payments may be lower than the threshold required to trigger the financial institution's reporting requirements, and may be paid with anonymous instruments such as cash, bank drafts, or through a lawyer's trust account.

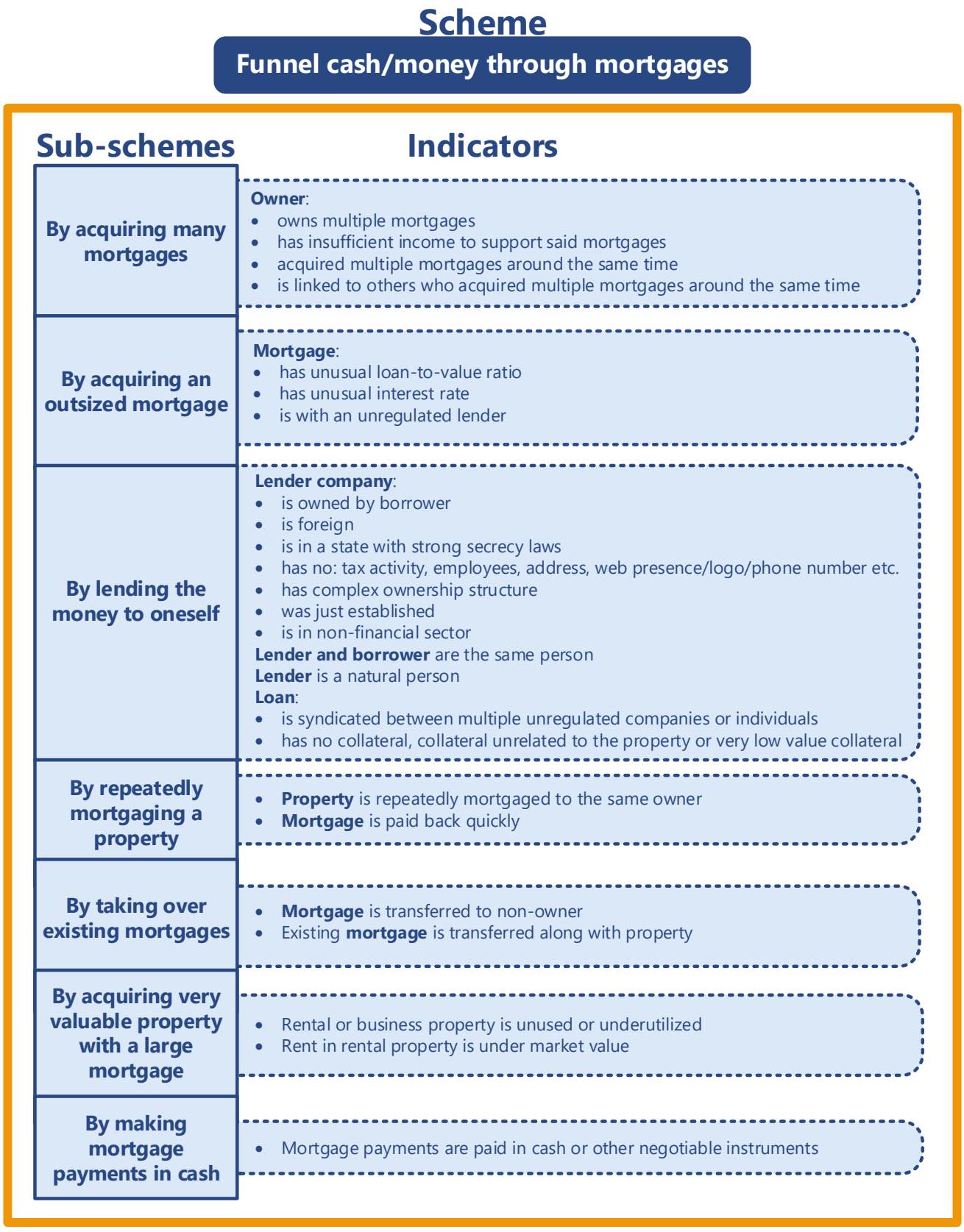
One of the key functions of this scheme, beyond the conversion of dirty funds, is that it creates a paper trail for ownership thereby creating a justification for the payments and assets acquired.

4.3.2.1 Sub-scheme: making mortgage payments in cash

One sub-scheme involves using funds to be laundered in whole or in part to pay a mortgage issued from a bank. These funds can be in the form of cash, bank drafts, or electronic funds acquired through fraud.

The regular interval and moderate size of these payments can make it very difficult to spot as they may fall below the PCMLTFA's reporting requirements, and they are virtually indistinguishable from regular mortgages paid with legitimate income.

Figure 4-2 - Summary of scheme: funneling cash or dirty funds through mortgages



4.3.2.2 Sub-schemes: acquiring many mortgages, an outsized mortgage, or high-value property with a large mortgage

There is an incentive in money laundering schemes to make mortgage payments as large and as frequent as possible without triggering suspicion, as it ensures that a larger amount of money is being laundered on a regular basis.

Acquiring several mortgages from different lenders, even if these are on the same property, can allow the perpetrator to evade the monitoring systems of each individual institution while still laundering a significant amount of money each month. For example, if the payments on the launderer's mortgages totaled $10,000 a month, the lending institution may be suspicious if the launderer attempted to pay this with cash, or if regular international transfers were received to pay the mortgage. These suspicions would be lowered if the payment were a few thousand dollars per month. By spreading out the loans, the risk of each individual institution filing an STR is reduced.

Other schemes involve acquiring mortgages which are larger than the value of the property or purchasing very expensive properties, such as luxurious houses and condominiums, or commercial real estate. This increases the mortgage size and the amount of cash that can be laundered.

The primary goal of this sub-scheme is to funnel as much money through mortgages as possible. A money launderer who purchases a property for this purpose may not bother maximizing the revenue-generating potential of the asset. In instances where the acquired properties are rental units, this underutilization may be visible through unusually large vacancy rates.[20]

4.3.2.3 Sub-scheme: by repeatedly mortgaging a property

Mortgages are eventually paid off, which is a limitation for continued money laundering through a property. This is particularly true if the mortgage is taken out with a short term length or if accelerated mortgage repayments are used. One way to continue laundering money through a property is to take out subsequent mortgages on the property once each has been paid off.

The repeated use of the same property reduces the overall exposure of the individual and the need to use nominee owners and other accomplices in the laundering process.

4.3.2.4 Sub-scheme: lending the money to oneself

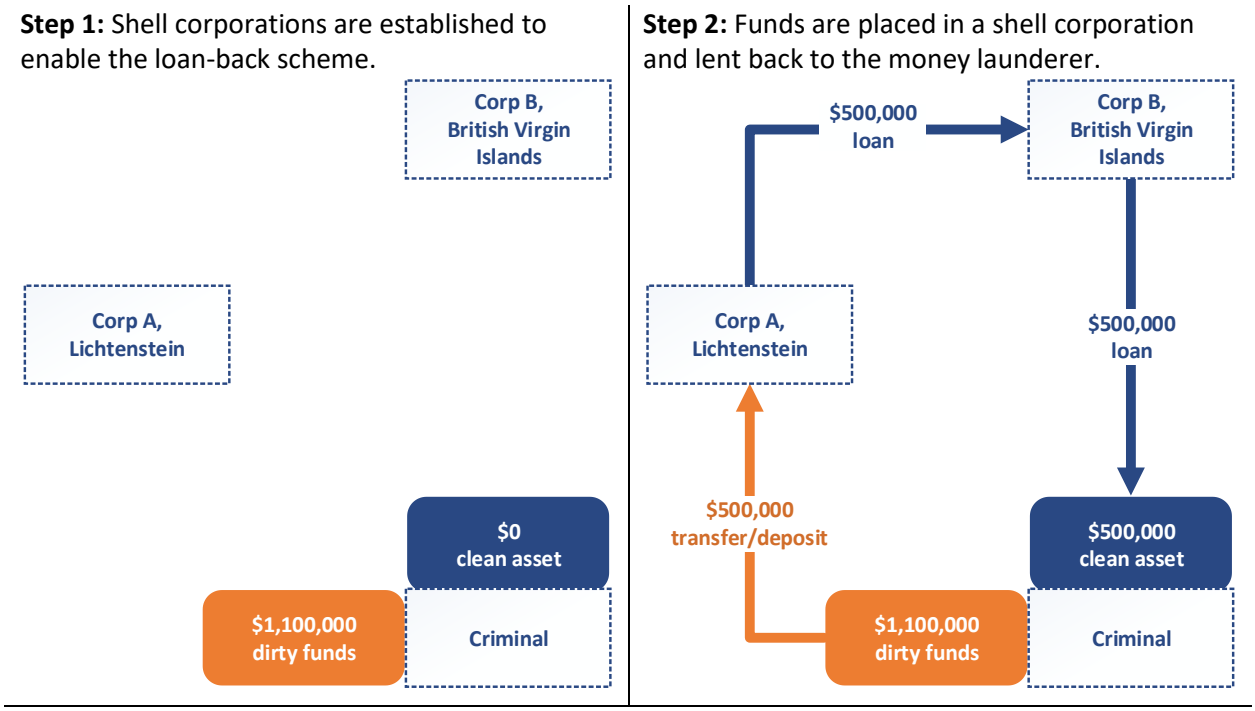
A loan-back scheme is one of the better-known money laundering techniques which involves an individual lending criminal proceeds to themselves to acquire a property, creating the appearance that the funds come from a legitimate source. The advantage of this scheme is that the money is cleaned both ways, through the lending and the repayment of the loan, creating both a cleanly-acquired asset and seemingly legitimate liquid funds.

Figure 4-3 illustrates an example of such a scheme.

[20] A money launderer may also be padding rental income with additional proceeds of crime, in which case the property would not appear vacant. See section 4.3.8.1. However, employing such a scheme involve additional fraudulent paperwork which could entail an additional risk for the money launderer.

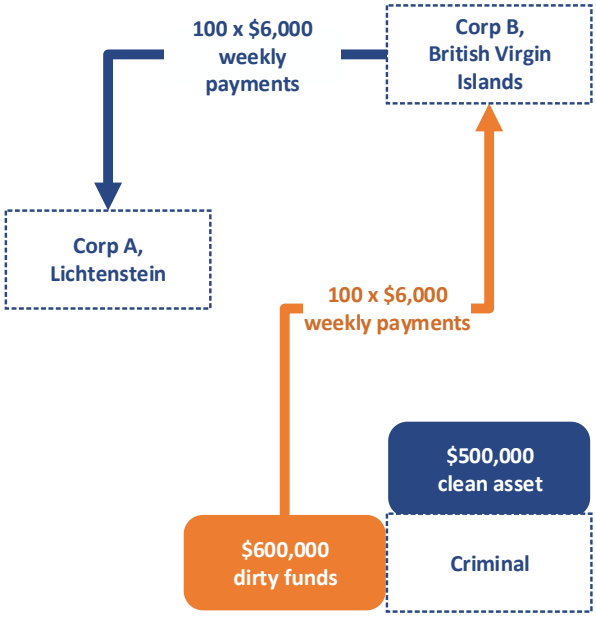
In step one, the individual in possession of a large amount of criminal proceeds sets up two shell corporations, Corp A and Corp B, ideally offshore and difficult-to-probe jurisdictions.[21]

Figure 4-3 - Stages of a loan-back scheme

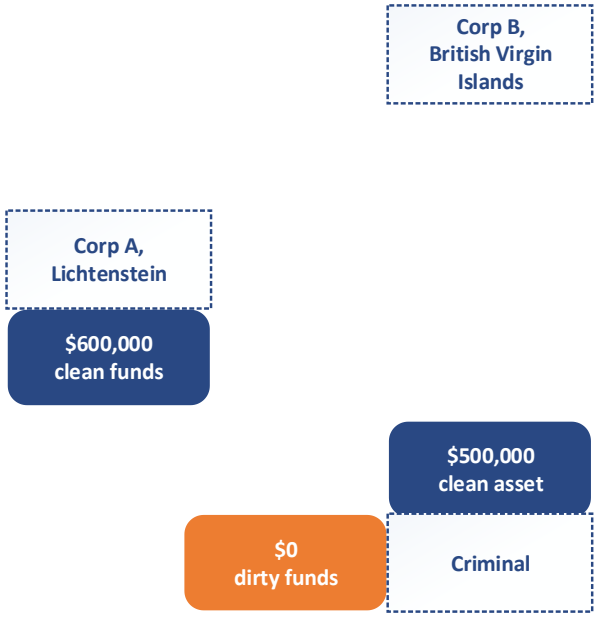


[21] High levels of financial secrecy make it more difficult, time-consuming, or even impossible for law enforcement in Canada to identify the beneficial owners, assets, and bank accounts of a corporation. The Tax Justice Network (2018) measures 20 different indicators of financial secrecy including the manner in which company ownership is recorded, the transparency of limited partnerships, the limits on the use of bearer shares, and the extend of bilateral information-sharing treaties.

Step 3: The loan is repaid with remaining proceeds of crime.



Step 4: Proceeds have been cleaned. A new loan can be used to launder additional proceeds.



In step two, a large amount of proceeds are transferred to Corp A. This step may be the most challenging but can be significantly simplified if the funds are already in electronic form, as is often the case with the proceeds of fraud. Corp A immediately lends the funds to Corp B, which receives them as a "clean" loan. Corp B then lends the clean funds as a clean loan to the individual who uses them to acquire property.

In step three, the launderer funnels the rest of their criminal proceeds through the loan to Corp B, which then uses them to repay the loan from Corp A. Corp A receives the repayments as "clean" funds from Corp B. Through this process, the individual gains ownership over an asset and has documentation of a clean loan used for the purchase. Corp A also gains clean funds as these are funneled first through Corp B.

Steps two and three can be repeated with additional mortgages, but no longer require the initial placement of funds in Corp A. Future laundering can thus be conducted solely through the loans without the need to make large individual transfers.

This sub-scheme may involve a number of shell corporations registered domestically or internationally, leading to more complicated schemes. Criminal proceeds may also be lent directly through a personal loan or by using a nominee.

This type of structure is often necessary for being successful in the justification stage of the laundering process. High quality data on loan arrangements for real estate purchases are therefore essential for the accurate detection or money laundering and for the estimation of its extent.

A family-focused variant of this sub-scheme was illustrated from the case of an apprehended drug trafficker in Toronto:

> "Family members or their companies would be listed as the source of mortgage financing for homes registered in the names of the trafficker's parents or siblings. For example, on August 9 1985 title to one home was transferred to his father for $122,900 and on the same day, a mortgage was registered against the home. The source of the mortgage financing was the mother of the accused. On August 30 1985, title to another property was transferred to the sister of the accused for $176,000. Four mortgages were registered against this property in July 1988, March 1989, May 1990, and September 1990. The last mortgage was registered for $184,000 in the favour of a company whose president was another sister of the trafficker. On September 9, 1985, another home was purchased in the name of the mother for $175,000. On October 22, 1985, a mortgage was registered against this property for $43,000. The reported source of this mortgage was her husband. Police were able to prove that all of the above homes and personal mortgage financing were funded through revenues generated by drug trafficking activities." (Schneider, 2004, p. 111)

The above example demonstrates the use of multiple schemes: the use of nominee owners, repeatedly mortgaging the same property and acquiring several mortgages around the same time.

4.3.2.5 Sub-scheme: taking over existing mortgages

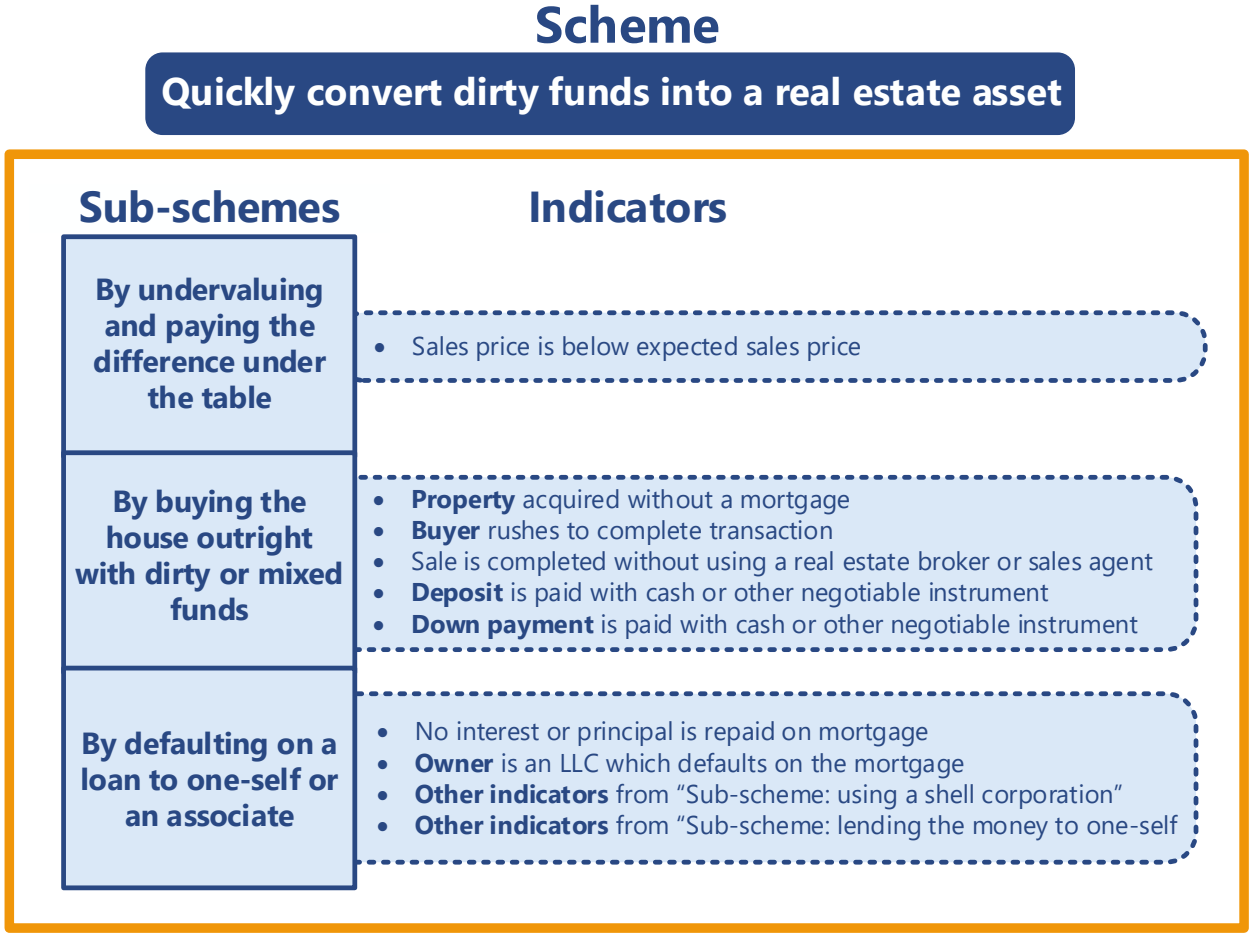
One may purchase a real estate property and take over any existing mortgage liability tied to it in the process (Schneider, 2004). The advantage of this approach is that the perpetrator does not have to justify a full mortgage to a financial institution, but merely the ability to make payments on the remainder of the mortgage. Scrutiny may be even lower if the house has appreciated in value since the signing of the original mortgage, as the bank will be more satisfied with the collateral.

In some cases, the seller will retain the mortgage on the house in their name, thereby becoming a nominee mortgage holder and providing a cover for the launderer.

4.3.3 Scheme: quickly convert dirty funds into a real estate asset

The large transaction sizes typically associated with real estate make it an attractive avenue to launder dirty funds. This scheme and related sub-schemes rely on the ability to use large amounts of criminal proceeds to quickly acquire ownership of an asset. Unlike the mortgage-related schemes above or the house-flipping schemes in the following section, these schemes do not rely on the continued laundering of funds through repeated transactions, but serve instead to give the launderer ownership of the property through a single large transaction.

Figure 4-4 - Summary of scheme: quickly converting dirty funds into a real estate asset



4.3.3.1 Sub-scheme: undervaluing and paying the difference under the table

This sub-scheme involves using a large amount of dirty funds in a real estate transaction by first making the transaction itself using clean funds, but then devising a side-arrangement using dirty funds to pay for value transferred. The nominal sale of the property is transacted below market value and the side-arrangement transfers the difference, thereby guaranteeing that the seller receives the full market price of the property. The seller may also then claim a smaller capital gain on the property, if the official records indicate a lower sales price, resulting in an added benefit for the accomplice.

4.3.3.2 Sub-scheme: buying property outright with dirty or mixed funds

Launderers may seek to simply spend their criminal proceeds directly in a real estate transaction without going through the layering and justification stages. This can be significantly simpler than going through the effort of constructing a loan-back scheme, or making a side-deal with a particular seller, or pursuing other more complicated schemes.

There are few controls on transactions between individuals. Therefore, a money launderer could buy a house from an owner for cash, provided the owner is willing to accept such a payment. There is no obligation on the side of individual sellers to investigate the source of funds, making it possible to complete a transaction with minimal involvement of lenders, lawyers, and notaries. Lawyers or notaries are often necessary for real estate transactions due to the need to keep funds in escrow, but launderers may rely on their own enforcement capacity to ensure that a transaction is honoured.

An individual may also use dirty money in some areas of the real estate deal, such as the initial deposit or the down payment. These sums can be substantial and may be less scrutinized than would a transfer of the entire value of the property.

4.3.3.3 Sub-scheme: defaulting on a loan to one-self or an accomplice

The documented transfer of an asset can also be accomplished by issuing a loan with the property as collateral and then defaulting on the loan. The lending party may then claim the property as their own. This sub-scheme parallels the loan-back scheme, with the main deviation being that the property, rather than the mortgage payments, are transferred to the lender.

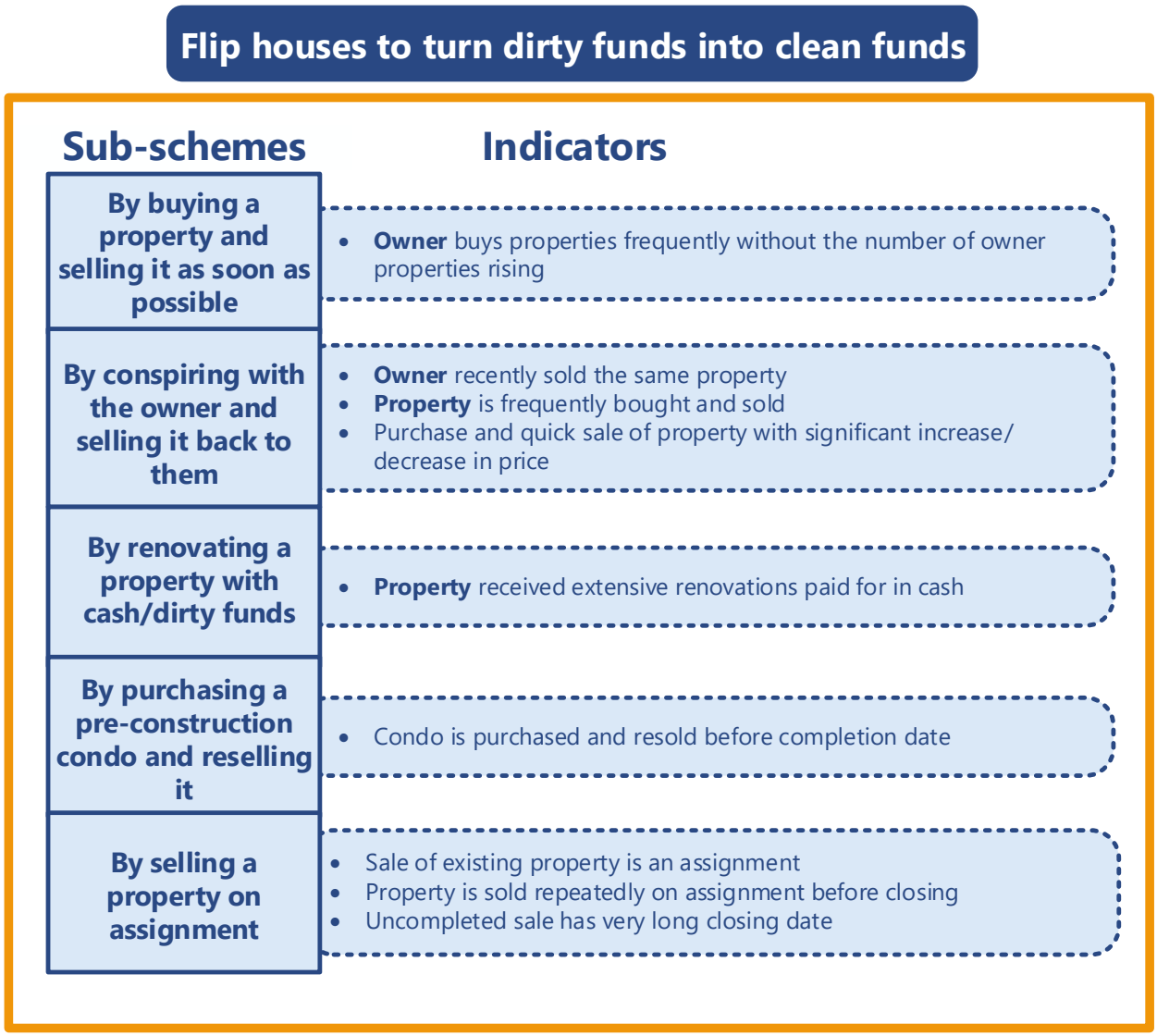
One advantage of this approach is that a default can happen much faster than the typical repayment of a mortgage, making it a relatively fast way to convert dirty funds into equity in a house.[22]

4.3.4 Scheme: flip houses to turn dirty funds into clean funds

House-flipping schemes attempt to engage in as many transactions as possible to increase the total amount of money laundered. However, they differ from mortgage-related schemes by seeking to rapidly finalize the transaction rather than providing ongoing payments, and differ from the single-transaction schemes through their focus on acquiring clean liquid funds, rather than acquiring a real estate asset.

Figure 4-5: Summary of scheme: flipping houses to clean dirty funds

[22] There is a variation of this scheme which involves lending money to an unwilling accomplice to fund their house purchase or other expense and then registering a fraudulent builder's lien against their property. The lender can file in court to receive the value of the lien when next the property is sold. In essence, this is a loan-sharking scheme which uses a provision of our land title system to secure the loans. While money can be laundered by lending in this manner, the main motive is profit. As such, the use of builder's liens is more of a money-making scheme than a money laundering scheme.



4.3.4.1 Sub-scheme: buying a property and selling it as soon as possible

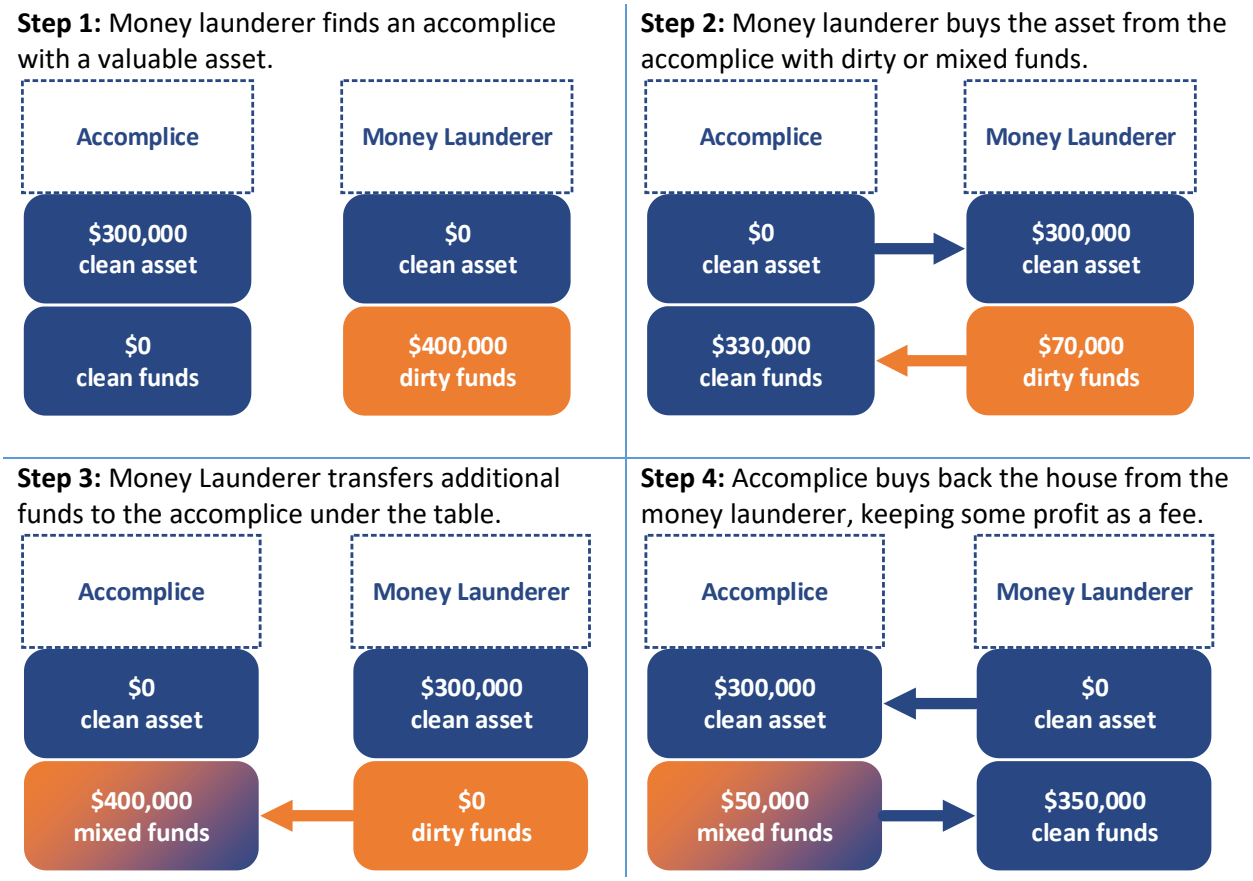
The real estate sector attracts a wide range of investors, including speculators who may engage in buying and selling activity. However, a source of profit is the appreciation of the house despite relatively large transaction expenses, such as a lawyer's fees, assessment costs, land transfer taxes, and others. Some investors and speculators may space out buying and selling activities to ensure that the property will yield adequate returns. Those who are active in money laundering schemes, on the other hand, are prepared to pay a premium to exchange dirty funds into seemingly legitimate income. They may engage in frequent transactions to launder funds through the initial deposit, down payment, or other aspects of the transaction.

It is worth noting here that this scheme may be particularly difficult to detect in overheating real estate markets due to the larger amount of speculation undertaken with legitimate funds in such markets.

4.3.4.2 *Sub-scheme: conspiring with the owner and selling it back to them*

Selling a house can be a time-consuming process that ties up a launderer's funds in the property. Therefore, one sub-scheme consists of pre-arranging the resale of a house before buying it, thus guaranteeing that the launderer's funds remain liquid. The perpetrator may choose to pay a co-conspirator for engaging in the transaction.

Figure 4-6 - Stages of a buy-back scheme



In this scheme, a money launderer first purchases a house from an accomplice and proceeds to transfers additional funds to the accomplice in order for it to be bought back. Then, the money launderer sells the house back to the accomplice, receiving the "appreciated" value of the house in clean funds. Note that the asset in question does not necessarily appreciate, but the selling activity may in fact also increase the market value of the house, providing the accomplice an additional incentive to participate.

A variant of this scheme includes a long string of accomplices or co-conspirators who sell the house from one to the other, repeatedly laundering criminal proceeds. This may also cause the value of the property to appreciate with every transaction. This arrangement is often referred to as carousel fraud, due to its similarity to a well-known Value-Added Tax (VAT) trading scheme with this name (Unger et al., 2011; Keen & Smith, 2006).

4.3.4.3 Sub-scheme: renovating a property with cash/dirty funds

Buying older properties and renovating them can be a profitable activity and includes more real value generation than a simple transaction in the existing housing stock. Contractors or materials suppliers used in the renovation process may be willing to accept cash and turn a blind eye to the source of the funds (German, 2019).

This sub-scheme is difficult to capture through data as cash is often used to pay contractors and suppliers. Supply stores are not required to report or investigate suspicious transactions and may accept unreasonably large amounts of cash.

4.3.4.4 Sub-scheme: purchasing a pre-construction condo and reselling it before the completion date

One sub-scheme focuses mainly on condominium apartments as it entails purchasing one or more units pre-construction, preferably paying the down payment and fees with criminal proceeds. The units can then be resold to legitimate investors before the completion of the building, effectively creating the appearance of legitimately acquired funds. Advantages of this procedure include the relatively low cost and the potential for additional profits if the value of the property increases between transactions.

Because the launderer never took possession of the property in the first place, the sale is technically an assignment; an agreement between the launderer and the new buyer for the latter to take over the obligations and rights of the original agreement between the launderer and the builder. Assignments are often not registered in the same way as are transfers of existing properties, and as such, this approach may involve less scrutiny.

4.3.4.5 Sub-scheme: selling an existing property on assignment

In a process known as shadow-flipping, a buyer may purchase a property and sell it to someone else on assignment before the closing date of the original purchase agreement. This is typically done to turn a quick profit on a house or – from the real estate agent's side – to receive several commissions on a single property.

A money launderer can benefit from shadow-flipping through the decreased transparency entailed. As the original purchaser never takes possession of the property, their name is never registered on the title, but they still received their returned deposit and any proceeds in clean money. The lack of registration of the sale makes these kinds of sales difficult to detect.

By working with accomplices, an assignment sale can be guaranteed. Multiple people can sell the property in succession, thereby creating a variant of the above-mentioned carousel fraud, but with the benefit that the intermediate purchasers would not be registered on title and would not pay property transfer taxes. The assignments may be purchased with dirty money, while clean money is received for the sale of the assignment. Thus, this would allow for greater amounts of laundering than may be possible through the returned deposit.

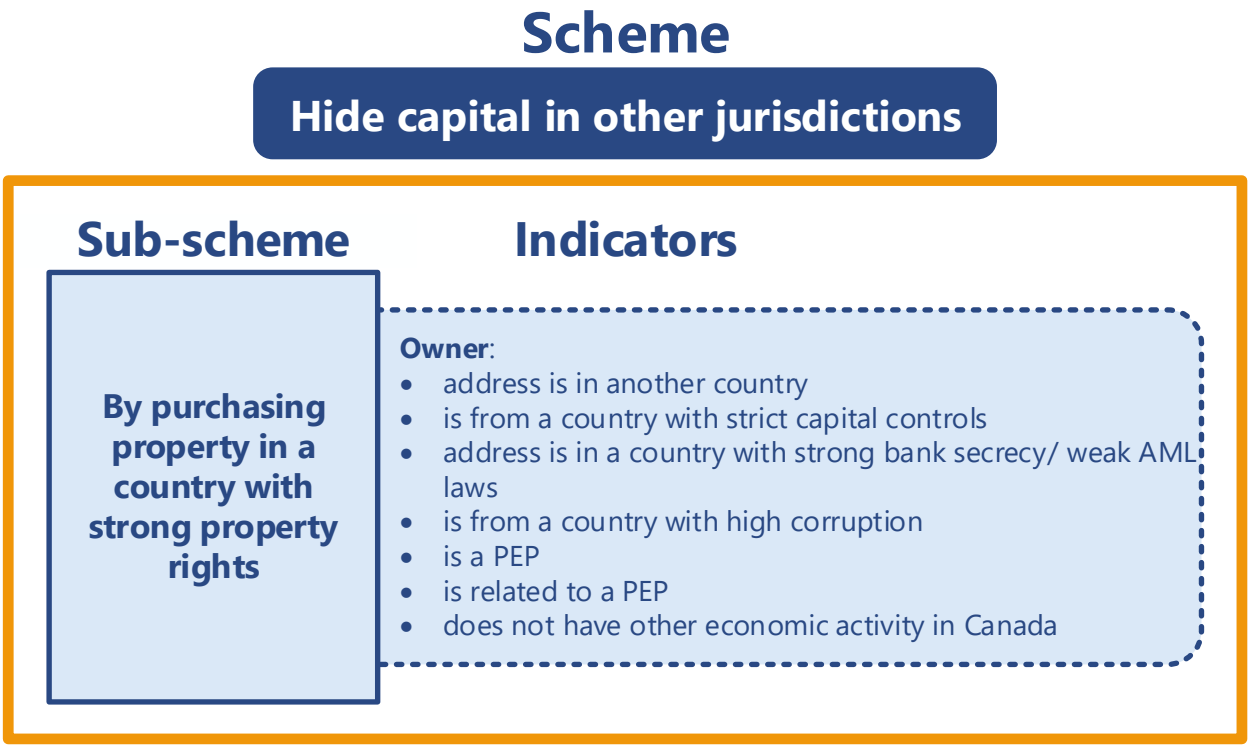
4.3.5 Scheme: hide capital in other jurisdictions

The proceeds from crime that are laundered within the jurisdiction where the predicate crime took place are more vulnerable to surveillance and possible seizure by jurisdictional authorities. Money

launderers and tax evaders therefore may seek to hide their funds in foreign jurisdictions, thereby increasing the difficulty for law enforcement to trace and seize their assets.

Hiding assets abroad may be particularly appealing to individuals who face a lot of scrutiny, such as politically exposed persons. Insofar as their wealth was acquired through exploiting their position, it is more likely to be seized by future governments in their home country, than if it is hidden abroad.

Figure 4-8: Summary of scheme: hiding capital in other jurisdictions



4.3.5.1 Sub-scheme: purchasing property in a country with strong property rights

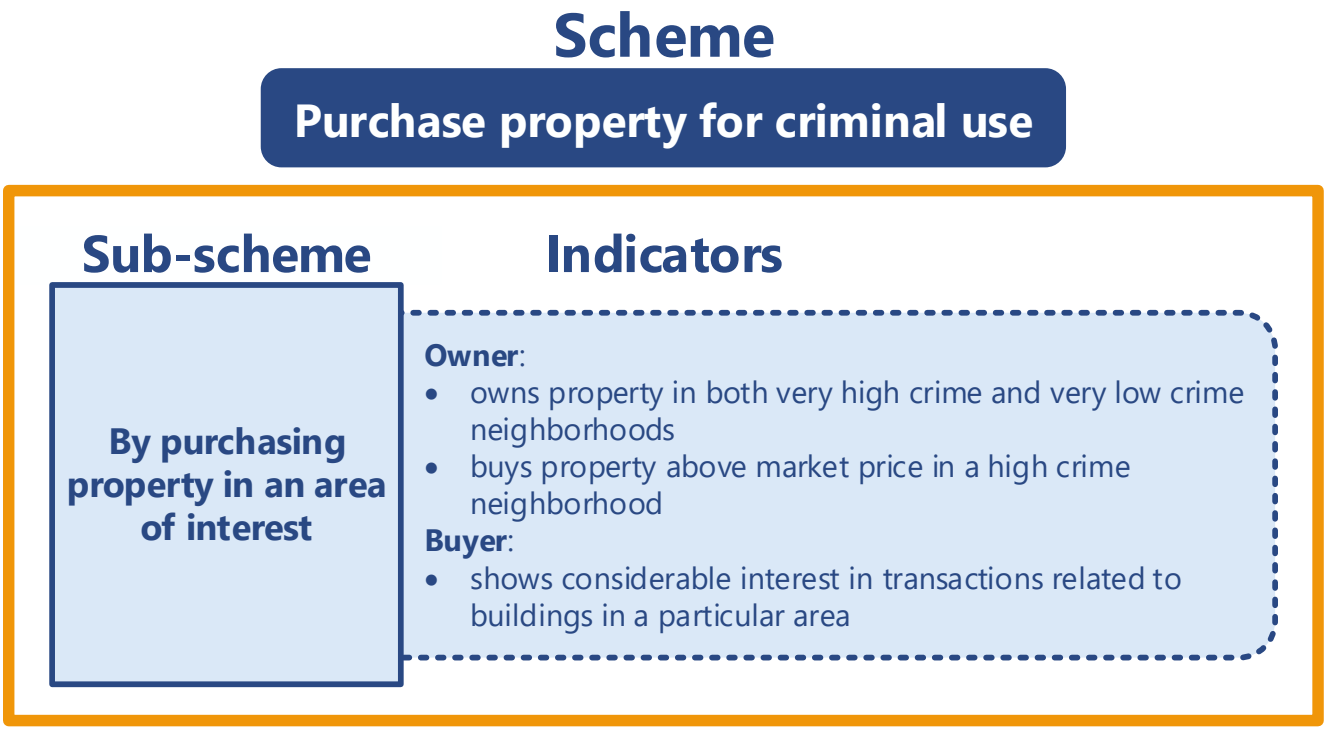
Countries such as Canada, with a long history of political stability and secure financial markets are an ideal place to store wealth. An investment into the Canadian real estate sector is a relatively secure one due to the country's stable economy, especially when compared to jurisdictions with volatile property laws or political instability. An added advantage is that the property can be used for housing or to fraudulently claim residency for tax or immigration purposes.

Legitimately acquired funds may also be hidden overseas if the money is earned in countries with strict capital controls, and economic or political instability. For example, Chinese nationals are forbidden from moving more than US$50,000 out of the country per year, but a former CIBC financial adviser testified to helping a client move $500,000 out of China, by having friends and relatives send 10 transfers to 10 different CIBC account (Tomlinson, 2015).

A more underground placement strategy has been dubbed the "Vancouver Model" by a professor at Macquarie University in Australia (Transparency International Canada, 2019, n. 54; Langdale, 2017). This scheme consists of a network of underground bankers located in Canada and abroad, particularly in Mainland China. This scheme is particular in that no cash moves between countries; instead, the bankers reconcile their accounts by exchanging debts (German, 2019). For example, a Chinese citizen might wish

to receive Canadian money upon arrival to Canada, out of sight of the Chinese currency controls, by paying a Chinese underground banker in Chinese Yuan. The banker then contacts a Canadian associate to ensure that Canadian cash is available upon their arrival. To settle the account between both underground bankers, drugs might be smuggled into Canada or the reverse transaction might be carried out for a Canadian citizen. These underground banking networks are believed to involve multiple bankers in several countries and work for various organized criminal groups.

Figure 4-9: Summary of scheme: purchasing property for criminal use



4.3.6 Scheme: purchasing property for criminal use

Criminal enterprises require space for production, storage, transaction, and administration of their enterprise, similarly to other businesses. Property ownership can allow individuals to launder criminal proceeds while providing a physical location to conduct their business.

4.3.6.1 Sub-scheme: purchasing property in an area of interest

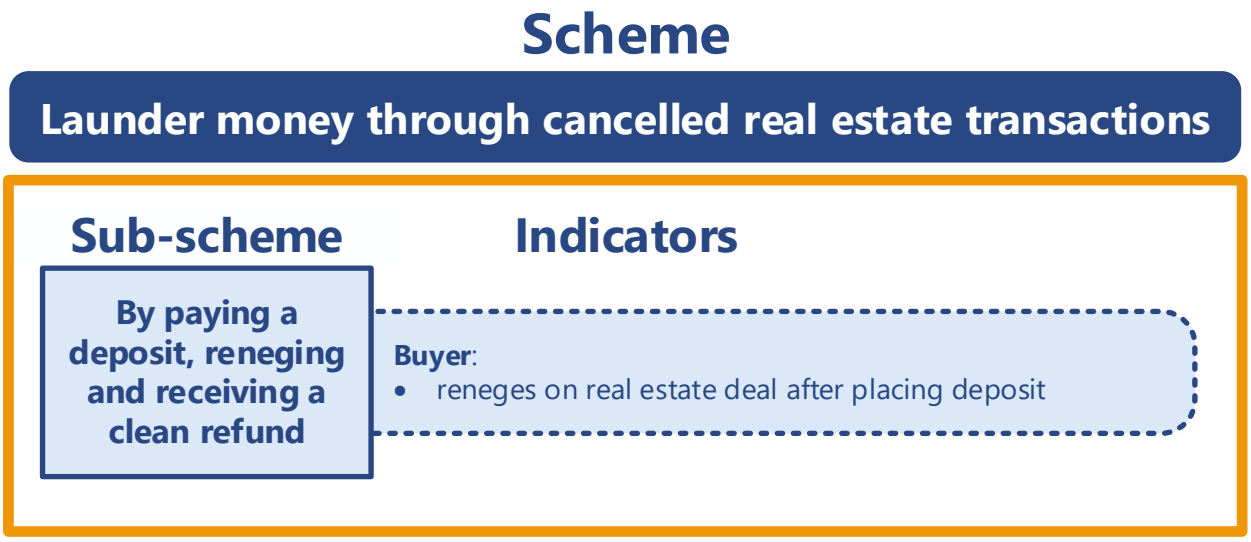
Just as with legitimate businesses, there is an incentive for criminal enterprises to maintain offices near their customers or clients. For drug dealers, this minimizes movement with the illegal substances while increasing the accessibility for their customers. Additionally, a city may be geographically divided between competing criminal organizations, confining activities to a prescribed area.

From an investment perspective, launderers may keep properties close by, so that they can more effectively be monitored and/or managed (Unger et al., 2011). Unger et al. also hypothesize that perpetrators may want to acquire property in particularly good neighborhoods, so as to reduce suspicions and increase the potential for an appreciation in value. An owner with properties in both very good and very bad areas may be an indication of such criminal enterprise.

4.3.7 Scheme: launder money through cancelled real estate transactions

The core motive of many money laundering schemes is to get rid of dirty money and receive clean money in return, such as a cheque from a lawyer or a real estate brokerage. Cancelling a real estate transaction enables the launderer to receive a signed lawyer's cheque as a refund, less a fee, in return. This is especially advantageous if a majority of the initial down payment and lawyer fees were paid for using criminal proceeds.

Figure 4-10: Summary of scheme: laundering money through cancelled real estate transactions



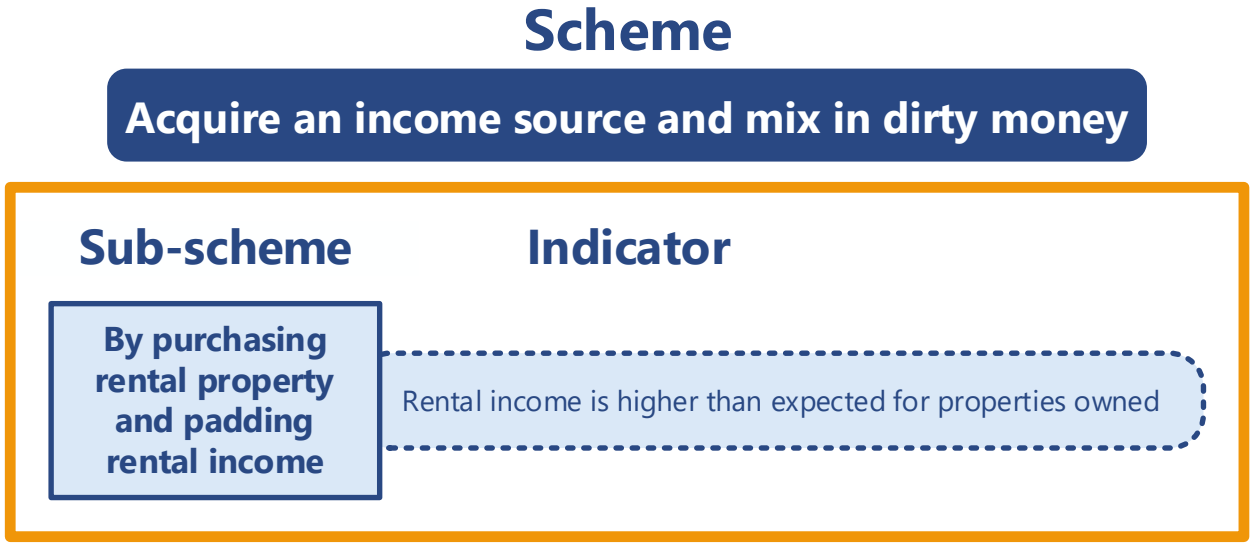
4.3.7.1 Sub-scheme: paying a deposit, reneging, and receiving a clean refund

Due to the high-priced nature of real estate transactions, a cancelled transaction and returned deposits can clean large amounts of money relatively quickly. Use of this scheme for cleaning cash is limited by the amount of cash a lawyer is willing to accept, but suffers from no such limitations for money acquired through fraud or when conspiring with a lawyer. Larger amounts could also be laundered by conspiring with a seller to utilize this scheme repeatedly on different lawyers or real estate brokers.

4.3.8 Scheme: acquire an income source and mix in dirty money

The ownership of cash-businesses is a method of laundering money that has stood the test of time. The term "money laundering" allegedly originates from Al Capone's use of laundrettes as a foil for prohibition-era alcohol sale profits (Unger, 2009). In theory, any legitimate business can serve this purpose, but some are more advantageous than others. In the residential real-estate sector, the relevant variant of this scheme relates to rental incomes.

Figure 4-11: Summary of scheme: acquiring an income source and mixing in dirty funds



4.3.8.1 Sub-scheme: purchasing rental property and padding rental income

Rental property can be employed both as housing and as a steady source of legitimate income which can be padded with the proceeds of crime. Relatively large amounts of cash may be readily accepted by the owner's bank if the building is in a low rent area where tenants are more likely to pay their rent in cash.

This sub-scheme is difficult to detect through the data that is currently available. Landlords are required to declare rental income on their tax returns, but renters in BC are excluded from the requirement to list their rent paid. As such, money launderers may claim significantly higher rental income than collected and make up the difference with dirty money while some individuals will declare lower earnings to evade taxes. This results in a data gap in the current structure and additional details are provided in the sections 5 and 6.

4.3.9 Miscellaneous indicators

Other indicators may be useful for detecting money laundering activity, but do not fall within the confines of the schemes detailed above. For example, individuals with a criminal record may be more likely to launder money, but such a record does not play a necessary part in any particular money laundering effort. It follows that these miscellaneous indicators may be helpful in detecting money laundering when combined with the scheme-derived indicators.

Figure 4-12: Summary of miscellaneous indicators

Miscellaneous Indicators

Company:

- Has frequent international transactions but no apparent business
- Is party to suspicious transactions

Owner:

- Has frequent international transactions
- Is party to suspicious transactions
- Has a criminal record
- Is a resident in a risk territory
- Is from a risk territory
- Has a spouse from a risk territory
- Has forfeited property in the past
- Has been prosecuted for money laundering in the past

5 WORKING GROUP DATA HOLDINGS

Working group members from the federal government of Canada and provincially in British Columbia hold an array of data that could supplement the data currently being used in anti-money laundering activities.

This chapter presents the results of a survey of working group members at the federal and BC levels. The inquiry covered the details of existing real estate- and AML-related data holdings and activities at surveyed institutions, the presence of data gaps and their current data sharing practices.

5.1 KEY POINTS

- Some data holdings are not leveraged for the purposes of AML because the holder lacks an AML mandate.
- Organizations involved in AML activities tend to use a case-by-case approach to detect money laundering, rather than a systematic data driven detection strategy or use data driven approaches that are limited by the data they can access.
- Recurring data gaps relate to information on beneficial ownership, property/financing legal arrangements, mortgage and wealth data, as well as relationship data among those transacting property transfers.
- Several data gaps could be filled with increased sharing of data between public institutions, subject to the Canadian legal framework.

5.2 SUITABILITY OF WORK STREAM DATA FOR AN AML DATA FRAMEWORK

Work Stream members have diverse data holdings that are pertinent to an AML data framework in real estate. The suitability of the data holdings for each institution is briefly addressed in this section, along five dimensions. These dimensions cover operational aspects such as the ease to share and link data. These dimensions are useful for identifying practical issues arising from implementing an AML data framework for real estate, but are not intended as a form of rigorous comparison between institutions.

Accessibility: This dimension considers the ease of employing the data within a data framework. Particularly, this relates to the difficulty of parsing and exporting the information. This will depend on the format and system holding the data, such as standardized databases, or a series of PDFs. It can also relate to the data itself, such as having fields that contain unstructured text, which may be difficult to parse.

Strength of identifiable information: Identifiable information in each dataset facilitates successful linking with other data sets. Identifiable information is any data that distinguishes individuals or businesses, such as names, birth dates, addresses or property descriptors. Consistent identifiers, or identification numbers such as property IDs or business numbers, are useful if they are unique and persistent across time, and across data sources. For example, a unique property ID is stronger than an address which can be written in multiple ways and is more susceptible to transcription error or manipulation.

Coverage of desired population: Money laundering is a relatively rare phenomenon, and as a result it can be hidden amongst the volume of activity in the marketplace. Having a comprehensive data set is a

prerequisite for a data driven approach to establish typical patterns and identify anomalies to transactional norms. For these reasons, data coverage is an important characteristic to consider to assess how data can be used for AML purposes. Examples of coverage limitations include tax data from the Canada Revenue Agency (CRA), which typically excludes data relating to individuals who do not pay taxes, are non-residents, or make incomplete declarations or, FINTRAC datasets, which exclude most international and cash transactions below certain thresholds or made outside established time periods. These exclusions are not necessarily problematic in the context of a data framework providing the data does not preclude the detection of ML schemes from the data. An example would be loan data which only includes mortgages from major banks. Money launderers may prefer smaller banks with smaller compliance departments in an effort to avoid detection or rely on non-banks or private lending.

Coverage may also be limited to certain time spans or geographical locations.

Fit between data dimension and analytical concept: With the exception of FINTRAC, data held by work stream members is generally not collected for the purposes of detecting money laundering in real estate. Furthermore, some aspects of what may be detected using data indicators are complex or broad. For example, while the terms of a mortgage reflects a person's obligation to pay, it is very difficult to assess a person's ability to pay. Income data within the CRA's data holdings could be used to bolster such an indicator and historical records may provide insight into overall wealth information. Similarly, the use of family, close friends, and business associates as nominal owners is a well-known ML scheme. Data from the CRA can be used to identify family members, but generally excludes extended family, friends, and business associates. In essence, the data does not capture much of the analytical concept of close ties. These limitations can complicate any repurposing of the data into an AML data framework for real estate.

The particular analytical concepts discussed in these sections will relate closely to the schemes and indicators which make use of the data points held by the given institution.

Accuracy of data recorded: This dimension measures the extent to which the data recorded correctly captures the phenomenon it was designed to measure. For example, tax data is intended to measure income from which a tax liability can be calculated. A property transaction record is intended to capture the identity of the new owner of the property and some associated transaction details, such as the date of the transfer, the price paid, etc. This dimension will mainly be reflective of the rigour of the collection effort, the validation procedures employed, and the incentives for providing accurate information.

5.3 Statistics Canada

Statistics Canada currently houses extensive data holdings collected solely for statistical and research purposes and cannot be used for enforcement purposes. The most relevant sources are the Canadian Housing Statistics Program (CHSP), the Business Register (BR) and the Uniform Crime Reporting (UCR) survey. Statistics Canada also has an extensive data sharing agreement with the Canada Revenue Agency (CRA). Section 5.4 will detail the CRA's data holdings.

5.3.1 Statistics Canada Data

5.3.1.1 Canadian Housing Statistics Program

The Canadian Housing Statistics Program (CHSP) is an administrative census of residential properties in Canada. The CHSP's holdings currently covers British Columbia, Ontario, Nova Scotia (for 2018, 2019, and 2020), and New Brunswick (2019 and 2020), with additional provinces and years being added as the project develops. The CHSP aims for full coverage of Canada, including 13 provinces and territories, 35 metropolitan areas and 5,000+ municipalities by 2022.

The CHSP data holdings are derived by linking administrative records from Statistics Canada, provincial and territorial title and land assessment records, and other government sources. For British Columbia, the main source of data is British Columbia Assessment (BCA).

The data contain a number of valuable characteristics on properties including their date of construction, assessment value, and location. Additional information on property owners is also available, such as residency and immigrant status, income, age, gender and legal type for properties owned by legal entities. Significant data pieces not currently available via the CHSP include mortgage information, vacancy status, rental rates charged, and details on renovations undertaken.

The data can be linked with other records through reliable identifiers for nearly all properties and owners. It could thus be readily utilized in an AML RE data framework for research purposes. These data cannot be used for enforcement.

5.3.1.2 Business Register

The Business Register (BR) contains a consistent, comprehensive listing of all businesses in Canada for the purposes of conducting scientific surveys of business activity and performance. It includes relevant data points such as revenues, expenses and number of employees, taken from surveys, tax records and profiling by dedicated staff.

Importantly, the BR also maintains a record of ownership linkages among companies operating in Canada, which can be used to identify the parent companies within complex legal structures. In large part, created through data collected under the authority of the Corporate Returns Act and schedule 9 filings of business taxes generally for small entities. Additionally, per CRA Schedule 50, private corporations are required to disclose shareholders holding 10% or more of common and preferred shares, up to a maximum of ten individuals. Not all ownership relations are reflected on the BR, however, in large part because ownership relations are not automatically mapped from submitted tax data, and because minority owner relations are not recorded in the mapping of ownership. Additionally, this information is only available for corporations who have had tax activity in Canada. It is possible for shell corporations which own property to declare no tax activity in Canada. Such corporations would appear on the BR, when incorporated in Canada, but would have very limited data. In theory, this could result in some money laundering-related ownerships having little associated data on the BR.

As with the CHSP, BR data are in a searchable format that facilitates analysis and can be linked with other records through reliable identifiers. As with the CHSP, these data can be used only for research purposes. These data cannot be used for enforcement.

5.3.1.3 Uniform Crime Reporting Survey

The Uniform Crime Reporting (UCR) Survey contains records of police-reported crime incidents from every police agency in Canada, dating back to 1962. During the 1980's the survey transitioned from an aggregate survey to collecting microdata, which includes detailed characteristics of incidents and the accused, and additionally victims if the offence involved violence.. To some extent, the microdata can be linked to other data sources. The UCR Survey captures incidents of money laundering as violation of the Criminal Code and does not have variables that can identify the type of money laundering, such as real estate.

The Uniform Crime Reporting Survey data reflects crimes that are reported to police or detected by police. The volume of ML incidents reported by police can be influenced by a number of factors. First, since ML schemes are formulated to avoid detection, they should appear as legitimate and aim to not be detected by law enforcement. Second, the extent to which police statistics include money laundering may be related to the allocation of resources to detecting and investigating this type of crime. Third, some instances of money laundering that come to the attention of police services may not be captured in the UCR data due to the collection method which is limited to the identification of the four most serious violations of an incident. . Due to the difficulty in prosecuting ML in Canada, the most serious charge on money launderers often relates to the predicate offence that created the funds to be laundered.

Due to these factors, reported instances of money laundering in police data cannot be used as a measure of the accuracy of other variables, or to estimate the prevalence of money laundering in the real estate market in Canada.

5.3.1.4 Integrated Criminal Courts Survey

The Integrated Criminal Court Survey (ICCS) seeks to maintain information on appearances, cases, and charges in provincial, territorial, and superior courts in Canada. The survey includes information on court activity, case characteristics, case processing information, case outcomes, and characteristics of the accused.

The data from the ICCS is presented at the case level. It follows that, if a case has multiple charges with differing decisions and sentencing outcomes, the most serious charge is selected to represent the case.

5.3.1.5 Other Resources

Additional work completed by Statistics Canada could be of interest in an AML framework. The Financial Performance Indicators for Canadian Business, a now discontinued program, generated financial performance indicators for small to medium Canadian businesses based on detailed tax records. This information was produced by geography and industry type. This program could be reinstated in an AML data framework to produce standards for Canadian businesses, thereby identifying outliers economy-wide. This could serve as an identifier since illicit businesses may over report their earnings to launder proceeds of crime or may lack other financial measures consistent with earned revenues such as payroll deduction expenses and other operating expenses.

5.3.2 Relevant Statistics Canada Data Points

5.3.2.1 Canadian Housing Statistics Program

Table	DP #	Description

Natural Persons	02	Name
Natural Persons	03	Address
Natural Persons	04	Date of birth
Natural Persons	05	Country of birth
Natural Persons	06	Properties owned, current, identifiers
Natural Persons	08	Tax filings / activity in Canada
Natural Persons	16	Date of immigration to Canada
Natural Persons	17	Immigration class
Legal Persons/Arrangements	02	Name
Legal Persons/Arrangements	03	Address
Legal Persons/Arrangements	06	Industry
Legal Persons/Arrangements	19	Properties owned, current
Properties	02	Address
Properties	03	Type
Properties	04	Owners, identifiers
Properties	06	Occupied by owner, flag
Properties	11	Construction date
Real Estate Transactions	06	Property assessed or expected value, time of sale
Miscellaneous	09	Property values in area

5.3.2.2 Business Register

Table	DP #	Description
Legal Persons/Arrangements	02	Name
Legal Persons/Arrangements	03	Address
Legal Persons/Arrangements	04	Date of Registration
Legal Persons/Arrangements	05	Country of Registration
Legal Persons/Arrangements	06	Industry
Legal Persons/Arrangements	07	Immediate owners, identifiers
Legal Persons/Arrangements	09	Ownership structure
Legal Persons/Arrangements	10	Tax filings / activity in Canada
Legal Persons/Arrangements	11	Number of employees
Legal Persons/Arrangements	17	Revenue
Legal Persons/Arrangements	23	Country of control
Legal Persons/Arrangements	24	Related and associated corporations, identifiers

5.3.2.3 Uniform Crime Reporting Survey

Table	DP #	Description
Miscellaneous	10	Crime rate in area

5.3.2.4 Integrated Criminal Courts Survey

Table	DP #	Description
Natural Persons	20	Has been a criminal record, flag
Natural Persons	22	Has been accused of money laundering, flag

5.3.3 Suitability of Statistics Canada data within an AML data framework

Accessibility: Statistics Canada data is already processed and fit for analysis. As such, it has high levels of accessibility.

Strength of identifiable information: The CHSP and BR data contain very strong identifiers that allow for data linking. For the CHSP data, a small fraction of identifiers may be misattributed due to the use of probabilistic linking methods for property owners.

For the UCR data, the identifiers are comparatively weak; in many incidents no name is available or only aliases are known. Where names are available, they are often obscured by the application of a SOUNDEX cipher. SOUNDEX is a phonetic algorithm applied to protect the identity of the individuals involved[23]. It follows that this hinders the linking process, if employed. Identifiers for ICCS are generally quite strong.

Coverage of desired populations: The coverage of the Statistics Canada data highlighted here is extensive but not complete. The CHSP covers only owners of residential properties, not commercial properties. Geographically, the CHSP is currently limited to properties and owners in Ontario, British Columbia, Nova Scotia and New Brunswick, but aims to have full coverage of residential property owners in Canada by 2022. Although the data is updated annually, the most recent data available is for 2019; the program began in 2017.

The Business Register covers all federally- and provincially- incorporated entities and businesses that pay Canadian income taxes. Corporations registered in a foreign country are included when they have taxable activity in Canada, which could exclude shell companies whose main activities involve holding real estate or providing a small number of loans as part of an ML scheme. Additionally, ownership structures are not complete for all businesses and exclude minority ownership relations. The Business Register also conducts profiles on companies to fill in details unavailable through survey or administrative (tax) programs. Profiling is conducted for larger corporations, and as such, shell companies are unlikely to be profiled

The UCR covers all reported-police incidents across Canada, and as such, does not reflect incidents of unreported crimes.

The ICCS covers all provinces and territories from 2005 onward, and smaller parts of Canada from 1994 onward. However, four jurisdictions – Prince Edward Island, Ontario, Manitoba and Saskatchewan – do not provide superior court data to the ICCS survey. This could cause an underestimation of serious cases, potentially including money laundering and organized crime, as they may be processed in superior courts.

Fit between data dimension and analytical concepts: The data held at Statistics Canada, and particularly the information produced by the CHSP, is employed extensively across the indicators, leading to the output of many analytical concepts. Many of these are fairly straightforward; the identity of the legal owner of the property, their age, the characteristics of the property including its type and

[23] SOUNDEX replaces a name with a letter and a series of numbers in a deterministic fashion. Different names will produce an identical SOUNDEX code, particularly where they sound somewhat similar. The result makes it more difficult to link a particular SOUNDEX codes to a particular person.

location, the dates at which the property was sold, etc. However, some of these concepts are more intricate and harder to match. For example, there is no direct information in the CHSP for the indicators which relate to ownership tied to countries with higher levels of corruption or bank secrecy. Approximations using immigration data or current address (sometimes available for non-resident property owners) can be used but are imperfect measures of source of funds.

The CHSP includes assessment values received from BC Assessment and other data providers which can be used as a proxy for the fair market value.

One advantage of the data held in the BR is that ownership structures and their changes are mapped. Complex structures may be an indication of attempted money laundering. However, the knowledge of ownership structures is bounded by the borders of Canada; direct foreign owners are known from schedule 50 tax filings, but further links in the ownership structure are unknown. In addition, the de facto owners may differ from the de jure owners available in the BR.

The UCR is employed in an indicator measuring the level of crime in an area. While there are known discrepancies between reported crimes and crime levels observed from victimization surveys, it should still be a fairly good measure of crime levels by geography area such as municipalities or neighbourhoods. A drawback is that not all police services report geo-locator information for criminal incidents, and there may be some instances when the locator represents the police station rather than the location of the criminal incident.

The ICCS data is employed for the indicator of whether a person has been prosecuted for money laundering in the past. As with the other indicators in the miscellaneous category, their utility lies in identifying cases of heightened suspicion, in this case related to significant money laundering-adjacent activity. The analytical fit is here quite high as there are not many ML prosecutions in Canada and most of these result in a conviction on a related predicate offence charge.

Accuracy of data recorded: In general, Statistics Canada data has a high level of accuracy as significant effort is expended to validate, verify, and clean the data. Some inaccuracies necessarily remain either as a result of probabilistic linking methods or as a result of inaccuracies in the data received by the agency or from non-reporting. However, for CHSP data, there is a strong incentive among data providers to achieve a high level of accuracy on the identity of the owner of a given property and on the assessed value of said property.

UCR data includes some geographic inaccuracies in the location of incidents. As noted, however, some police services do not provide geo-locator information for the incidents and there may be instances when the locator represents the police station rather than the location of the criminal incident. This may occur in some jurisdictions and results in lower geographic resolution of the data.

5.4 Financial Transactions Analysis Centre of Canada (FINTRAC)

The Financial Transactions Reports and Analysis Centre of Canada (FINTRAC) is Canada's Financial Intelligence Unit (FIU) and Anti-Money Laundering and Anti-Terrorist Financing (AML/ATF) regulator – an agency specifically established to detect, deter, and prevent money laundering and terrorist activity financing, as well as threats to the security of Canada. The Centre produces actionable financial intelligence in support of investigations of Canada's police, law enforcement and national security

agencies in relation to money laundering, terrorist activity financing and threats to national security. It also generates strategic financial intelligence, including research reports on trends and developments for the public, regime partners, policy decision-makers and international counterparts that shines a light on the nature and extent of money laundering and terrorist activity financing inside and outside Canada.

FINTRAC is an administrative-type of FIU,that receives financial transaction records from reporting entities and acts at arm's length from law enforcement. It operates according to the Canadian Charter of Rights and Freedoms, and the various privacy laws in Canada. FINTRAC has a large data set from which it conducts it analysis, and develops financial intelligence disclosures. This data set is the principal repository of suspicious financial transactions undertaken by individuals and businesses in Canada. Many real estate schemes rely on significant cash transactions or cross-border flows of funds, which are available in the data received by the agency.

Four of FINTRAC's data holdings are particularly relevant for the detection of money laundering in the real estate sector: Large Cash Transaction Reports, Electronic Funds Transfer Reports, Suspicious Transaction Reports, and the Money Services Business Registry.

5.4.1 FINTRAC data

5.4.1.1 Large Cash Transaction Reports (LCTRs)

Reporting entities must submit an LCTR to FINTRAC whenever cash is received in:

- an amount of $10,000 or more in the course of a single transaction; or
- two or more amounts of less than $10,000 each that total $10,000 or more. In these cases, an LCTR must be submitted if the transactions were made within 24 consecutive hours of each other by or on behalf of the same individual or entity.

These are stored into a database maintained and used by FINTRAC. Reporting entities are:

- Accountants and accounting firms (when carrying out certain activities on behalf of their clients)
- Agents of the Crown that sell money orders
- British Columbia notaries (when carrying out certain activities on behalf of their clients)
- Casinos
- Dealers in precious metals and stones
- Financial entities
- Life insurance companies, brokers, and agents
- Money services businesses
- Real estate developers, brokers, and sales representatives[24] (when carrying out certain activities)

[24] Under the PCMLTFA, an entity is considered to be a real estate broker or sale representative when it acts as an agent for the purchase or sale of real estate and are registered and licensed to do so by the province. This includes the buying or selling of land, houses, commercial buildings, etc. A real estate developer means an individual or entity other than a real estate broker or sales representative who, in any calendar year after 2007, has sold the following to the public:
-five or more new houses or condominium units;
-one or more new commercial or industrial buildings;
-one or more new multi-unit residential buildings each of which contains five or more residential units; or
-two or more new multi-unit residential buildings that together contain five or more residential units.

- Securities dealers

The vast majority of the reports originate from financial entities. Records are kept by FINTRAC for at least 10 years and up to 15 years, unless they were part of a disclosure and are then retained indefinitely**.** The data has identifiers for the parties involved in the transactions, including names, addresses, telephone numbers, client identification, dates of birth, and account numbers.

5.4.1.2 Electronic Funds Transfer Reports (EFTRs)

ETFRs record flows of funds into and out of Canada at the request of a client. Financial entities, money services businesses, and casinos are required to report all such transfers of $10,000 or more made within 24 consecutive hours. FINTRAC receives millions of such reports a year. The data has identifiers for the parties involved in the transactions, including names, addresses, telephone numbers, client identification, dates of birth, and account numbers.

5.4.1.3 Suspicious Transaction Reports (STRs)

Reporting entities are required to file a suspicious transaction report whenever they have reasonable grounds to suspect that a completed or attempted transaction is related to the commission of a money laundering or terrorist financing offence. The reporting entities are the same as with Large Cash Transaction Reports.

FINTRAC receives over 200,000 such reports a year from reporting entities (FINTRAC, 2019). The frequency and quality of reports vary significantly between reporting entities given the volume of transaction activities and the subjective nature of STRs. Very few reports are received from the real estate sector; in fact, FINTRAC received 62 such reports from BC between 2014 and 2018 (German, 2019, p. 59). Over the past few years, FINTRAC has engaged with real estate regulatory bodies, associations and businesses across the country to strengthen compliance in this sector including STR reporting obligations (FINTRAC, 2019). In addition, as part of the 2019 federal budget, the Centre has received funding to increase its outreach and examinations (including STR reporting obligations) in the real estate and casino sectors with a focus on the province of British Columbia (FINTRAC, 2019).

STRs are one of the most valuable report types submitted to FINTRAC. In addition to the prescribed data, the reports allow through a free-form field (i.e. Part G - Description of suspicious activity) for an expansion on the descriptive details surrounding transactions. That said, the information in this free-form field can be difficult to parse systematically.

5.4.1.4 Money Services Business registry

Money services businesses (MSBs) in Canada are required to register with FINTRAC. The Centre compiles and publishes information on these registered MSBs, including the names, addresses, and activities of their branches and agents. The data also contains some additional identifying details, such as their incorporation date and jurisdiction, incorporation number, business license number, and MSB registration number. Since some money launderers have been known to set up MSBs to launder proceeds of crime, linking suspicious properties or their owners to such MSBs could prove valuable in AML efforts. The data are available in bulk in XML format, making it suitable for processing and analysis. The entries contain several unique identifiers as well as names and addresses, allowing to link the data to other sources.

5.4.1.5 Other data sources

FINTRAC also receives casino disbursement reports, terrorist property reports, and voluntary information records. For operational purposes, terrorist property reports are submitted by paper to FINTRAC. However, the reports' data points are inputted into FINTRAC's database, and captured electronically. Voluntary information records about suspicions of money laundering or about the financing of terrorist activities and FIU queries are similar to STRs but are submitted, respectively, by persons or entities that are not legally required to report, and other FIUs. FINTRAC houses Cross-Border Currency Reports and Customs Seizure Reports, which it receives from the Canadian Border Services Agency (CBSA).

FINTRAC also has a Memorandum of Understanding with the Real Estate Council of British Columbia in accordance with which the RECBC shares data with FINTRAC.

Finally, FINTRAC also keeps a record of reporting entities and all administrative penalties it has issued.

5.4.2 Relevant FINTRAC data points

Table	DP #	Description
Natural Persons	18	Financial transactions, identifiers
Natural Persons	19	Person is party to suspicious transactions, flag
Legal Persons/Arrangements	15	Financial transactions, identifiers
Legal Persons/Arrangements	16	Entity is party to suspicious transactions, flag
Financial Transactions	02	Date
Financial Transactions	03	Amount
Financial Transactions	04	Sender, identifier
Financial Transactions	05	Recipient, identifier
Financial Transactions	06	Country of origin (for EFTRs only)
Financial Transactions	07	Destination country (for EFTRs only)

5.4.3 Suitability of FINTRAC data within an AML data framework

Accessibility: FINTRAC data is digital, housed in exportable relational databases. Some fields contain unstructured text, which may be difficult to parse in an automated or systematized way.

Strength of identifiable information: FINTRAC data has identifiers such as first and last name, date of birth, address, phone number, account number, and (for LCTRs and STRs) client number for natural persons. They receive corporation, trust or other entity name, address, phone number, account number, and (for LCTRs and STRs) incorporation number for legal entities and arrangements.

All prescribed data received by FINTRAC is either legally considered as mandatory, mandatory where applicable, or require reasonable efforts to be provided by reporting entities. As such, these data points are not always present within each data set (LCTRs, EFTRs, and STRs) or available for all parties involved in transactions.

Coverage of desired populations: Coverage of financial transactions is not complete. LCTRs cover receipt of cash transactions above $10,000, while EFTRs cover transactions above $10,000 that are entering or leaving the country at the request of a client. The movement of money outside of the

country is not covered, nor are large domestic transfers. As such, some money laundering related transactions are not included in the database.

Coverage of STRs is difficult to assess. The level of STR reporting varies by industry given the volume of transaction activities, the subjective nature of STRs, and the industries' level of knowledge and awareness of their obligations under the Proceeds of Crime (Money Laundering) and Terrorist Financing Act (PCMLTFA). To address this, FINTRAC engages reporting entities in both general and targeted education and awareness activities. The Centre also provides extensive guidance through publications on its website and uses a risk-based approach to select reporting entities that are examined each year to assess, among other things, their compliance with STR obligations. While lawyers in Canada and Quebec notaries are identified as reporting entities under the PCMLTFA, they are exempt from PCMLTFA obligations, including reporting suspicious transactions. This is as a result of a Supreme Court decision. Additional details are provided in Box 3.1.

Information on both the sender and recipient is available in all EFTRs, but not necessarily in LCTRs and STRs, in part because these can have a different number or arrangement of involved parties.

Fit between data dimension and analytical concepts: The indicators which rely on FINTRAC data either use the data to identify financial ties with other countries, or to simply identify whether a person or entity has engaged in suspicious transactions.

With regard to financial ties, FINTRAC reports are generated from completed or attempted transactions, so the presence of a given transaction report is a good reflection of the existence of a financial tie. However, although EFTs reflect transactions received from abroad or sent to abroad, these transactions are not directly linked to a given real estate transaction or mortgage payment. Furthermore, the absence of transaction reports to other countries does not necessarily indicate the absence of a financial tie to other countries. The transactions may simply fall below the reporting threshold.

With regard to STRs, there is a strong fit between the data collected and its use for AML. STRs are submitted by reporting entities when they encounter ML suspicions, so the existence of an STR directly reflects at least some expert threshold of suspicion. However, reporting entities employ their own judgement and internal rating systems when determining when and how to submit a suspicious transaction report. Additionally, reporting entities are restrained by the limited quantity of information they can share with each other, complicating their efforts to coordinate on improving their systems.

These limitations make the level of noise and density of information on money laundering in the STRs unknown. Due to the subjective nature of STRs, in which it is submitted to FINTRAC following a financial transaction that occurs or is attempted, and for which there are reasonable grounds to suspect that the transaction is related to the commission or attempted commission of a money laundering or terrorist activity financing offence, there are limitations in identifying over or under reporting. However, in FINTRAC's role as a regulator, it conducts assessments and reviews of the quality, timing, and volume of reports it receives, including STRs. If FINTRAC identifies any deficiencies in a reporting entities' STR obligations, it will inform the reporting entity, and depending on the severity of deficiency, may issue an administrative monetary penalty (AMP).

In addition, neither the fitness of the data for the detection of ML, nor the efficacy or success of FINTRAC's operations can be determined by equating the number of financial transaction reports

FINTRAC receives to the number of financial intelligence disclosures it provides to law enforcement agencies and other disclosure recipients. While other countries and FIUs may use such metrics, they are often a judicial or law enforcement type of FIU, and have different legal frameworks than Canada. This often results in foreign FIUs being able to access reports directly from reporting entities, rather than having a substantial amount of reports being received by them. In Canada, one of FINTRAC's financial intelligence disclosures may contain over 100 financial transaction records. Alternatively, a foreign FIU's financial intelligence disclosure may only contain one or two financial transaction records.

Accuracy of data recorded: As with any type of data, the accuracy of information has limitations. Fields within forms may contain falsified or mistaken information (e.g. spelling errors or information placed in wrong part of form). While the accuracy is unknown and undisclosed for all reports, it is presumably high for LCTRs and EFTRs, as both of these are threshold-based reports, which can be identified by reporting entities in a systematic manner and reported to FINTRAC.

STR forms have upfront validation to ensure that mandatory fields in the report are completed, that data is provided in the appropriate format and, wherever possible, that the content of the fields is accurate.

BOX 5.1
THE REGULATION OF LAWYERS AND THE FEDERATION OF LAW SOCIETIES OF CANADA

The 2016 Financial Action Task Force (FATF) Mutual Evaluation Report (MER) for Canada noted that "lawyers in Canada are frequently involved in financial transactions, often related to high-risk sectors, such as real estate, as well as in the formation of trust and companies" (FATF, 2016a, pp. 80-1), and that "the legal profession in Canada is especially vulnerable to misuse for money laundering and terrorist financing risk" (FATF, 2016a, p. 15).

Lawyers were highlighted by the FATF because the AML/ATF provisions in federal regulations towards legal professionals are not in force and lawyers are not reporting suspicious activities of their clients to FINTRAC.

CANADA (ATTORNEY GENERAL) V FEDERATION OF LAW SOCIETIES OF CANADA

Lawyers were originally included in the PCMLTFA in 2001, where the legislation required them to collect information about their clients and, among other things, report suspicious transactions to FINTRAC. Additionally, the PCMLTFA allowed FINTRAC warrantless access to law offices and documents subject to compliance examinations. Following constitutional challenges launched by several law societies via the Federation of Law Societies of Canada (FLSC), the Attorney General of Canada agreed to exclude lawyers from the Act in 2002 pending the outcome of litigation.

In 2015, the case was escalated to the Supreme Court of Canada (SCC). The FLSC argued that the requirements under the PCMLTFA violated the *Charter of Rights and Freedoms* (the Charter). Additionally, the FLSC claimed that the provisions were unconstitutional because they breached solicitor-client privilege.

The SCC found that provisions requiring lawyers to collect client information, keep records and have lawyers' offices undergo compliance searches violated section 7 (right to life, liberty and security of the person) and section 8 (protection against unreasonable interference with a reasonable expectation of privacy) of the Charter. However, the SCC acknowledged the important public purpose of Canada's AML/ATF Regime and did not preclude the possibility that Parliament could impose AML/ATF obligations on the legal profession that are within constitutional boundaries.

As it stands, the SCC's ruling has left the regulation of the conduct of lawyers to provincial and territorial law societies, which are self-regulatory organizations. The law societies are responsible for regulating, monitoring, and sanctioning the conduct of their members in the public interest, including their potential involvement in money laundering schemes.

In practice, the FLSC provides model rules which are adopted by the provincial and territorial bodies. The FATF MER noted that the representatives of the FLSC "did not demonstrate a proper understanding of ML/TF risks of the legal profession" (FATF, 2016a, p. 81) and that the SCC ruling "raises serious concerns" (FATF, 2016a, p. 15).

In response, AML experts have recommended the establishment of legislation requiring lawyers to report suspicious transactions to their law society (Maloney et al., 2019) or for the Canadian government to be more accountable for the regulation of lawyers (Bromwich, 2018).

5.5 Canada Revenue Agency

The Canada Revenue Agency (CRA) processes tax filings from individuals, legal persons, and other entities in Canada and as such, collects a significant amount of data. In addition, the CRA has the capacity to include data from external databases, such as law enforcement data, for the purpose of assessing tax compliance. However, this data are not accessible to the divisions devoted to tax administration. Identifying ML has not been a specific focus for the CRA since money launderers may be tax compliant.

Tax data collected through BC provincial taxes are covered in section 5.9.4, BC Taxpayer Administration Compliance & Services.

5.5.1 Canada Revenue Agency data

5.5.1.1 Federal Tax Records (T1, T2, T3)

The CRA houses the current and historical tax records for natural persons, legal persons, and legal arrangements, such as trusts. The data extends back to the 1980s and includes many details on the activity levels of these entities.

The T2 tax records filed for legal persons include information on the Canadian beneficial owners of the company, as well as the proportion of the company owned by foreign entities. However, legal persons may list nominees as the entity's owners, complicating the analysis of the data for AML purposes.

Individual taxes, recorded on T1 records, also contain a wealth of information on the characteristics of individuals, including income and its source, dependants, and family members. Since nominees are often close family members, utilizing such links is useful in an AML context.

T3 records are filed by trustees of Trusts whenever a resident beneficiary receives income from the trust, or when a beneficiary's interest in the trust changes.

Individuals owning and renting real estate or other property are required to complete a Statement of Real Estate Rentals (Schedule T776). They must list the addresses of all rental properties, the number of units rented, as well as the gross rental income received.

CRA data are housed in a Common Agency Data Warehouse, contains reliable identifiers, and is currently used in systematic analyses to detect risk cases.

5.5.1.2 Other data sources at Canada Revenue Agency

The enforcement side of the CRA accesses many external sources on a case-by-case basis. For example, the CRA has the authority to obtain mortgage data from a credit rating agency, property ownership data from the BC Land Title Survey Authority, or transaction data from financial institutions in order to conduct investigations. The CRA can obtain information about persons or entities in another country (Foreign-based Requirements) on a case–by-case basis. Furthermore, an "Unnamed Persons Requirement" can be issued to companies or individuals domestically, requesting information on third

parties. This has been used to obtain information from property developers on any pre-construction assignment sales made on the properties under construction.

At a more systematic level, the CRA receives EFTR data from reporting entities. Additionally, specialty tax data from the provinces is also collected, such as data from the implementation of BC's Speculation and Vacancy Tax, Land Owner Transparency Act, and the BC Condo and Strata Assignment Integrity Register (CSAIR).

The CRA also acquires bulk data on building, demolition, and renovation permits.

5.5.2 Relevant CRA Data Points

5.5.2.1 T1 tax filings

Table	DP #	Description
Natural Persons	02	Name
Natural Persons	03	Address
Natural Persons	04	Date of birth
Natural Persons	08	Tax filings / activity in Canada
Natural Persons	09	Income, time of purchase
Natural Persons	10	Income, historic
Natural Persons	13	State of health (i.e. incapacitated / healthy)
Natural Persons	23	Spouse, identifiers
Natural Persons	24	Relatives, identifiers
Natural Persons	25	Income from rent

5.5.2.2 T2 tax filings

Table	DP #	Description
Legal Persons/Arrangements	02	Name
Legal Persons/Arrangements	03	Address
Legal Persons/Arrangements	06	Industry
Legal Persons/Arrangements	09	Ownership structure
Legal Persons/Arrangements	10	Tax filings / activity in Canada
Legal Persons/Arrangements	11	Number of employees
Legal Persons/Arrangements	17	Revenue
Legal Persons/Arrangements	18	Revenue from rent
Legal Persons/Arrangements	23	Country of control

There may be additional data points present in Statistics Canada's Business Register (BR) derived from tax data provided by the CRA.

5.5.3 Suitability of CRA data within an AML data framework

Accessibility: Tax data is digitized.

Strength of identifiable information: Tax data has very strong identifiable information including name, gender, address and SIN for people and name, address, and business number for legal persons and arrangements.

Coverage of desired populations: The coverage is very high; nearly complete for both individuals and companies. Some individuals who do not pay taxes, such as some students and unemployed, do not appear in the data.

For businesses, the coverage is similarly high, although some variables, such as the number of employees, stem from schedules which do not apply to all businesses.

Non-residents are generally not covered. This results in a data gap for foreign money launderers who launder their gains in Canada, and for many Canadians who have spent years abroad; estimating their wealth based on historical income would be less feasible for this group.

The CRA investigates ML occurring through sales, including house flipping, as owners must pay taxes on capital gains received through the sale of a secondary property. However, the agency is not currently able to record property transactions accomplishedthrough transfers of corporate ownership, in large part because such transfers are not documented in a central database. Similarly, unfinished condominium assignments are currently not recorded[25] alongside transfers of finished properties, so flipping such properties would not be automatically identifiable in the tax data.

Another coverage gap within the CRA's data holdings is information on tenant payments to landlords. Landlords must report rental income in their T1 fillings, and fill a T776. However, tenants are generally not obligated to report the amount of rent paid, so the landlord data cannot be corroborated with tenant data. This means that ML schemes relating to underutilized properties cannot be systematically detected without additional information.

Fit between data dimension and analytical concepts: In general, the fit between the data points and their use within an AML data framework is moderate. Income reported corresponds to most legitimate income, which may be used as a proxy for an individual's legitimate ability to afford their property. Some money launderers, however, may also report income from illegitimate sources as a part of the laundering process. This results in some reported income being derived from illegitimate sources, and as such, reported income may not necessarily correspond to legitimate disposable income. The CRA also has wealth indicators, such as value of assets relative to yearly fillings, which could be an additional indicator of ability to pay.

For identifying nominees, the fit is also moderate. Tax data only covers close relatives filing their taxes together; limiting it to mostly children and legal partners. The historical coverage is moderate as well; going back to the 1980s but no further. The parents of adults from that time would not be identifiable. A larger conceptual challenge is, as noted, that money launderers that may have confidants within their extended family or outside their family, as these will not be part of relationship structures mapped through tax data.

[25] The CRA does, however, issue Unnamed Persons Requirements to selected developers to obtain this information.

The indicators which rely on T2 tax records attempt to identify shell corporations used for owning property or facilitating their purchase. In general there is a good fit between tax-recorded activity in Canada and real activity in Canada. Tax authorities strive to record all taxable business activity, which closely mirrors economic activity by legal persons.

Accuracy of data recorded: Accuracy for most dimensions is likely to be high. However, the use of nominal owners and/or directors for legal persons and arrangements could be prevalent among money launderers. The accuracy of beneficial ownership data may therefore be in doubt. Finally, money laundering businesses may also be over-reporting their revenues, in an attempt to create some justification for their illicitly obtained funds.

5.6 The Bank of Canada

The Bank of Canada tracks several housing-related indicators, due to their relevance for financial stability. The Bank does not track ML activities or attempt to estimate their prevalence, but does analyze activities that may be related to ML activities, such as the prevalence of house flipping and the activity of real estate investors and private and/or unregulated lenders.

5.6.1 Data sources from the Bank of Canada

The Bank of Canada does not itself produce data sources relevant to AML. The Bank has however access to very detailed information on the housing market, particularly for Ontario. For federally regulated financial institutions (FRFIs), the Bank receives anonymized loan-level data on the characteristics of individual mortgage originations and renewals. The Bank also receives accompanying household income information at the time of origination. It receives similar information on all insured mortgages and all mortgages securitized through government securitization programs, but does not receive similar data on uninsured mortgages for non-FRFIs. The Bank of Canada has acquired land registry data for Ontario from Teranet, including loan sizes, property values, and interest rates covering all lenders, including private lenders. Finally, The Bank of Canada receives loan payments data from TransUnion.

5.6.2 Relevant Bank of Canada Data Points

There are currently no data points in the data framework that use data from the Bank of Canada. However, there are a number of loan and wealth data points that may or may not have equivalents within Bank of Canada data or amongst their data providers.

Table	DP #	Description
Loans/Mortgages	06	Size
Loans/Mortgages	07	Interest rate
Loans/Mortgages	08	Term length
Loans/Mortgages	09	Opening date (first payment date)
Loans/Mortgages	10	Closing date (last payment date)
Loans/Mortgages	11	Date of transfer
Loans/Mortgages	12	Current balance
Loans/Mortgages	13	Payments (adjusted to reflect monthly payments), identifiers
Loans/Mortgages	14	First date of being 90 days in arrears
Loans/Mortgages	15	Collateral type
Loans/Mortgages	16	Collateral value

Loan Payments	02	Loan, identifier
Loan Payments	03	Payment date
Loan Payments	04	Payment amount
Properties	15	Mortgages on property, current, identifiers
Properties	16	Mortgages on property, historic, identifiers
Miscellaneous	01	List of regulated mortgage issuers in Canada

5.6.3 Suitability of Bank of Canada data within an AML data framework

Accessibility: Bank of Canada data is maintained in a digital and exportable format as much of is it received from external data providers.

Strength of identifiable information: For their mortgage-related data, the identifiers have been removed due to privacy legislation. Mortgages on a property in Ontario can be followed over time in Teranet data, but the holders are not identifiable. While private lender data is available for Ontario, the identities of these lenders is not included in the data.

Coverage of desired populations: The data covers mortgage originations and renewals, both insured and uninsured, for FRFIs. Another set of data covers transactionally - and portfolio-insured originations from all financial institutions that access government-sponsored securitization programs. For Ontario, the data in addition covers mortgages provided by less regulated and unregulated lenders. However, as the private lenders are not identified and the data is overall not linkable, schemes in which the identity of the lender are central – such as loan-back schemes – will not be detectable with this data.

Fit between data dimension and analytical concepts: The loan and mortgage related data points capture important characteristics of the financing of the real estate purchase, such as mortgage amount, interest rate, and term length. Some indicators rely directly on these properties of a mortgage and as such, there is a strong fit between the data gathered, and its use within an AML data framework.

One of the key analytical concepts of interest is measuring a person's ability to afford a given property or mortgage. The details of the mortgage are essential for such an assessment, but must be weighed against accurate income and wealth data for the mortgage holder. Without such linking, the data has little ability to measure this analytical concept.

Accuracy of data recorded: The data received by the Bank of Canada is not a regulatory requirement and so it is not standardized and can be prone to error. However, the Bank makes every effort to clean, organize and validate the data received. As with the CRA data, there may be systematic misreporting among money launderers.

5.7 Royal Canadian Mounted Police (RCMP)

The RCMP investigates money laundering offences and can access many data sources on a case-by-case basis during the course of their investigations.

In Canada, when criminal groups hide their assets in real estate, they typically do not register these purchases in their own names. Elaborate structures set up through law firms enable criminals to disguise their controlling interests in companies and assets through the use of nominees or offshore trusts. As a

result, real estate is often the safest asset that is disassociated from the proceeds of crime, and hidden away from any potential asset confiscation by law enforcement.

Money laundering investigations stem from predicate offence crimes, and unless a property was directly implicated in the course of an investigation, such as a grow op, or an illegal gambling house; the properties would not be a focus of the case. The Proceeds of Crime investigation attempts to identify all of the assets associated to the criminals implicated in the predicate offence investigations, but unless a clear path of evidence draws them to properties in their names, or the names of a closely known associates; law enforcement would not be aware of these assets, or able to seize them.

The RCMP submits incidents to the UCR survey administered by Statistics Canada on a regular basis, and case information is provided to prosecutors and other law enforcement agencies when required.

5.7.1 Relevant RCMP Data Points

Table	DP #	Description
Natural Persons	20	Criminal record, flag
Natural Persons	22	Has been prosecuted for money laundering, flag
Miscellaneous	10	Crime rate in area

5.7.2 Suitability of RCMP data within an AML data framework

Accessibility: The RCMP uses separate databases for incidence reports and for investigations. The incidence reports databases are shared with Statistics Canada and are therefore in practice exportable. Criminal Records are housed in the National Repository of Criminal Records. Records and fingerprints have been digitalized over the previous decade.

Strength of identifiable information: Identifiers are strong for criminal records where the full names and dates of birth are used. Fingerprints are also used in the National Repository of Criminal Records, but those are not feasible to use in an AML data framework. Police reports for crime incidences use a Soundex-encrypted name and, where available, date of birth do identify suspects, which combine to form rather weak identifiers. In addition, incidence records may contain no names or pseudonyms.

Coverage of desired populations: The coverage is weak as many money launderers go undetected. The data covers apprehended, detected, and/or reported launderers. Coverage for crime rates data should be very high, spanning all of Canada.

Fit between data dimension and analytical concepts: The indicators which rely on RCMP data on criminal records and charges of ML use these to identify an ML and/or other criminal history. In general, the RCMP data strives to record these same aspects, so these data points would have a good fit within an AML data framework.

For ML charges, these are often dropped by prosecutors, but many related charges are sustained in such cases, indicating the presence of money laundering. A conviction for ML or a related offence is similarly a strong signal, although the money laundering may not involve a real estate transaction.

For crime-rate data, there are two analytical concepts of interests. The first is are areas with a perceived danger to law-abiding citizens. This should fit closely with reported crime levels. The second is areas in

which criminal enterprise is undertaken. This should generally correspond to the crime rate in an area, although there is no data to affirm this.

Accuracy of data recorded: This dimension is difficult to assess. It is presumed that many money launderers have no criminal record, thus lowering the strength of this signal. The presence of a criminal record is, in most cases, an accurate signal, but the absence of it may not be. A similar problem exists for those accused of, or charged with, money laundering.

5.8 Canada Mortgage and Housing Corporation (CMHC)

The CMHC is a Crown Corporation that provides mortgage loan insurance, assist in the development of affordable housing initiatives, and provides research on housing to the Canadian government. As such, the CMHC collects information on mortgages and the housing industry in Canada.

The CMHC is a Work Stream member and has participated to the inquiry. However, CMHC mortgage insurance activities are considered out of scope as these are commercial and are more akin to private sector organizations financial institutions, rather than the public sector.

The CMHC has also engaged in work on developing a Market Integrity Index using acquired data to produce a number of money laundering indicators, and aggregating these into a measure of the extent of perceived bad actors in the housing market (Abdallah, Koulouris, & Pereboom, 2020). This work has been developed on Quebec data, but may be expanded to other provinces, including British Columbia.

5.9 British Columbia Public Institutions

British Columbia collects mortgage- and housing-related data in a number of distinct institutions and databases.

5.9.1 BC Assessment

British Columbia Assessment estimates property values in British Columbia for the purposes of collecting property taxes. While assessed property values will often differ from the actual market value of a house, it is the closest proxy amongst public holdings. The discrepancy between income and housing value, or between expected and actual sales prices are central to several schemes, making this an important data source. Many other provinces have similar assessment agencies. BCA data are used in the production of the Canadian Housing Statistics Program data at Statistics Canada.

5.9.1.1 Relevant BC Assessment Data Points

Table	DP #	Description
Properties	02	Address
Properties	03	Type
Properties	04	Owners, identifiers
Properties	05	Sales, historic, identifiers
Properties	08	Number of units
Properties	11	Construction date
Real Estate Transactions	05	Date of transaction being finalized
Real Estate Transactions	06	Property assessed or expected value, time of sale
Real Estate Transactions	07	Sales price

Miscellaneous	09	Property values in area

5.9.1.2 Suitability of BC Assessment data within an AML data framework

Accessibility: The data is regularly shared with Statistics Canada, and as such, has a high level of accessibility.

Strength of identifiable information: Very strong, as this data is currently linked with other datasets at Statistics Canada, leading to the production of official statistics. Each property has an individual and unique property identifier.

Coverage of desired populations: The coverage is complete as it covers all properties. And some large (>3) number of previous sales.

Fit between data dimension and analytical concepts: Data points such as property address, type, and date of sale are straightforward and therefore have a mostly uncomplicated use within an AML data framework. Indicators which rely on sales prices often rely on both the recorded sales price and the expected sales price for detecting under/over-pricing and the financing requirements of the transaction. Sales price data held by BC Assessment is a direct measure of the recorded sales price, but the use of assessment value as a proxy for expected sales price is less straightforward. While assessors do attempt to mirror market value, and update the assessment roll yearly, there is still lag which may result in a fit that may not be ideal for AML purposes, particularly during very active market periods.

For assessment data coming from other provinces, there may be larger lags between assessments and other particularities which increase differences between assessment values and market prices.

Some provinces also have assessment value appreciation restrictions, which may create a wider gulf between assessment value and the market value.

Accuracy of data recorded: The accuracy should be high for all data points. Typically, the data lists only nominal owners of property.

5.9.2 BC Land Title and Survey Authority

The British Columbia Land Title and Survey Authority (BC LTSA) is responsible for the official records related to land titles and mortgages in BC. As such, nearly all changes of ownership, mortgages with claims on ownership, and related agreements are registered with the agency. For example, their Form B contains valuable information on a mortgage agreement's lender and borrower, as well as details on the specifics of the transaction, including the principal amount, the interest period, and the interest rate.

BC LTSA faces operational constraints to analysis and alternate use of their data. Much of the data collected by BC LTSA is not stored or structured for systematic analysis. Individual records are retrievable but exist mainly in digital image format. Records are kept in an easily readable format for a year, after which they are replaced by their digital image counterpart. This is not conducive to historical or comprehensive data analysis requirements. BC LTSA operates on a cost-recovery basis to fulfill a well-defined land-title function. Expanded funding and an expanded mandate would be required to make changes necessary to make the data useable for AML purposes.

5.9.2.1 Relevant BC LTSA Data Points

Table	DP #	Description
Natural Persons	02	Name
Natural Persons	03	Address
Natural Persons	06	Properties owned, current, identifiers
Natural Persons	07	Properties owned, historic, identifiers
Natural Persons	15	Mortgages held, identifiers
Legal Persons/Arrangements	02	Name
Legal Persons/Arrangements	03	Address
Legal Persons/Arrangements	19	Properties owned, current, identifiers
Legal Persons/Arrangements	20	Properties owned, historic, identifiers
Legal Persons/Arrangements	21	Mortgages held, identifiers
Properties	02	Address
Properties	04	Owners, identifiers
Properties	05	Sales, historic, identifiers
Properties	15	Mortgages on property, current , identifiers
Properties	16	Mortgages on property, historic, identifiers
Loans/Mortgages	02	Property mortgaged, identifier
Loans/Mortgages	03	Holders, identifiers
Loans/Mortgages	04	Previous holders, identifiers
Loans/Mortgages	05	Lenders, identifiers
Loans/Mortgages	06	Size
Loans/Mortgages	07	Interest rate
Loans/Mortgages	08	Term length
Loans/Mortgages	09	Opening date (first payment date)
Loans/Mortgages	10	Closing date (last payment date)
Loans/Mortgages	11	Date of transfer
Real Estate Transactions	02	Property, identifier
Real Estate Transactions	03	Buyers, identifiers
Real Estate Transactions	04	Sellers, identifiers
Real Estate Transactions	05	Date of transaction being finalized
Real Estate Transactions	18	Involved notaries, identifiers
Miscellaneous	07	Interest rates on comparable mortgages

5.9.2.2 Suitability of BC LTSA data within an AML data framework

Accessibility: The data is held in an SQL database with a proprietary software layer enabling interaction with the information. This makes the land-title data accessible through a web-interface, but it is not clear whether the data can be separated from the software. It may be tied inextricably to the software itself. In addition, much of the archived data is held in scanned PDF and TIFF files, which are both very difficult to parse.

Strength of identifiable information: The strength of identifiers is moderate. Land title data includes name, occupation, and address, but no stronger identifiers such as a SIN or a BC tax identifier. However, the Property Transfer Tax form, collected by the BC Ministry of Finance, does contain identifiers which

could be used to combine BC LTSA data with stronger identifiers. These include the date of birth, individual tax number, SIN, country of citizenship, and where applicable, business number for the buyer, and name and address for the seller. Both the land title data and the transfer tax data include, where possible, the parcel identifier number (PID), so the two form can be linked via this PID and the date of the transaction.

Mortgage data includes the name and address of the borrower(s), the name, address, and occupation of the lender(s), and the parcel identifier number for the property.

Where forms are submitted by an agent, the name address and phone number of the agent is provided.

Coverage of desired populations: The coverage for current ownership is high; all land transfers must be registered with the BC LTSA. Historic data for ownership dates back to the mid-1800s. Digital ownership records are maintained from 1982 forward, although there are more data entry errors and concatenated fields in the earlier data.

Interests in a property can be recorded with LTSA, so enforceable loans against the property are registered, most commonly as mortgages, and occasionally as builder's liens.

Fit between data dimension and analytical concepts: The fit between the data collected and its use within an AML data framework is fairly high for the land title data. Ownership in the registry corresponds to legal ownership, while a sale corresponds to a legal sale.

Recorded sales prices and mortgage terms are important concepts in an AML data framework to indicate a buyer's capacity to make a purchase or the frequency of which a property is flipped. Mortgage data recorded in BC LTSA files can be used to confirm the existence of a mortgage and the parties thereof, but cannot be relied on for other financing details. Either other sources of mortgage data must therefore be found, or expanded detail must be collected by the BC LTSA to account for mortgage data points listed in the data framework.

Accuracy of data recorded: For ownership data, the accuracy is presumed to be high, except for the use of person nominees. For sales prices, a similar precision is expected since this amount is subject to transfer taxation. For mortgages, however, it is known that the registered BC LTSA mortgage information is not sufficiently precise to determine actual amounts borrowed or the monthly payment amount. Lenders typically provide upper limit principal values in order to reduce their paperwork burden. In addition, the details of the mortgage quickly become outdated as there is often no subsequent filings until the claim (debt) is cleared or a new party issues a loan secured by the property.

5.9.3 BC Land Owner Transparency Registry

In 2018 the government of BC passed the Land Owner Transparency Act (LOTA), which obligates owners to file a transparency declaration, and for indirect owners to file a transparency report specifying who the beneficial (natural person) owners are. For corporations and trusts which own properties, this will increase transparency regarding land ownership. This data will be placed in a Land Owner Transparency Registry (LOTR).

It is as yet unclear how reliable the registry will be, as the declaration is self-reported, but it does add an additional hurdle for would-be launderers or tax evaders, and there are also significant penalties

associated with filing false or misleading information. The registry will be housed at the BC Land Title Survey Authority.

5.9.3.1 Relevant BC Land Owner Transparency Registry Data Points

Table	DP #	Description
Natural Persons	06	Properties owned, current, identifiers
Legal Persons/Arrangements	08	Beneficial owners, identifiers
Legal Persons/Arrangements	09	Ownership structure
Legal Persons/Arrangements	19	Properties owned, current, identifiers
Properties	04	Owners, identifiers

5.9.3.2 Suitability of BC Land Owner Transparency Registry data within an AML data framework

Accessibility: It is currently unclear how the data will be stored, but large portions of it, without the strongest identifiers, will be accessible to the public in some manner. The future exportability of the data as a whole is also unknown.

Strength of identifiability: Information on reported beneficial interest holders has strong identifiers including name, date of birth, last known address, SIN or tax number, and residency status. Information on the reporting corporation, trust, or partnership includes their business name, business address, and relevant identifier (e.g. business number, incorporation number, trust reference number, etc.).

Coverage of desired populations: The coverage will initially be sparse but will eventually roll out to have more extensive coverage. A LOTA submission is required for any change of title. For most changes of title, this will simply identify the purchaser as a non-reporting body (e.g. an individual or publicly traded corporation). For all others, the land transfer will not go through until the accompanying transparency report has been submitted.

Existing owners who are required to file a transparency report must also do so within a prescribed period, although it remains to be seen how many will do so. Presumably, money launderers who are hiding behind a shell corporation might neglect this requirement. Updated transparency reports are also required whenever there is a change in beneficial ownership, such as when a corporation which owns a property itself changes hands. However, there is currently no mechanisms to identify and follow up on such cases, so these may also not be covered in the LOTR data.

The LOTR covers formal and informal beneficial ownerships through nominees such as family members who have legal title to the lands when these arrangements are conducted through trusts, partnerships, corporations, as well as unwritten contracts and arrangements which mirror these. However, it is likely that informal nominee owners will not be captured when these are engaging in money laundering. When such owners circumvent formal legal arrangements in order to increase the obscurity of the beneficial ownership, they may defend this obscurity by submitting an inaccurate transparency declaration.

Fit between data dimension and analytical concepts: In many cases we are interested in the identity of the natural person who has beneficial ownership of a given property. This is exactly what the LOTR seeks to record, so there is a high degree of fit between the data dimensions being recorded and the analytical concept of interest.

Accuracy of data recorded: Data submitted in the transparency reports will not be verified so the accuracy may be low, particularly for money launderers who have an incentive to falsify their reports. There is, however, a risk of being audited, and the resultant threat of significant penalties from submitting false statements.

5.9.4 BC Taxpayer Administration Compliance & Services

The Taxpayer Administration Compliance & Services (TACS) is the data system employed by the BC Ministry of Finance. It houses data stemming from tax-related operations, including efforts to address specific housing market issues, such as the Speculation and Vacancy Tax (SVT). This tax is levied only in designated taxable regions, namely municipalities within the Capital Regional District and the Metro Vancouver District, and will exempt properties that are principal residences or rented for at least six months per year. While this tax entails a self-declaration, its administration requires data on the owners of property, their income sources and the use of their owned properties. Specifically, the tax is higher for foreign owners and satellite families, which are families whose Canadian income is less than half of their total household income. TACS also houses data related to the administration of the Property Transfer Tax. This tax is greater for foreign national, foreign corporations and taxable trusts, and as such, contains information on the nature of property purchasers.

In order to reduce anonymous land ownership and tax evasion, the provincial government of BC introduced the Condo and Strata Assignment Integrity Register (CSAIR), effective January 2019. On a quarterly basis, developers must report all condo assignments in both pre-construction and completed apartments. Data from this registry will greatly decrease the opaqueness of the condominium market and allow for better identification of this money laundering sub-scheme.

5.9.4.1 Relevant BC TACS Data Points

Table	DP #	Description
Natural Persons	02	Name
Natural Persons	03	Address
Natural Persons	04	Date of birth
Natural Persons	06	Properties owned, current, identifiers
Natural Persons	06	Properties owned, historic, identifiers
Natural Persons	09	Income, time of purchase
Natural Persons	10	Income, historic
Natural Persons	16	Date of immigration to Canada
Natural Persons	23	Spouse, identifiers
Natural Persons	26	Is a trustee, flag
Legal Persons/Arrangements	02	Name
Legal Persons/Arrangements	03	Address
Legal Persons/Arrangements	04	Date of registration
Legal Persons/Arrangements	05	Country of incorporation
Legal Persons/Arrangements	06	Industry
Legal Persons/Arrangements	07	Immediate owners, identifiers
Legal Persons/Arrangements	08	Beneficial owners, identifiers
Legal Persons/Arrangements	18	Revenue from rent
Legal Persons/Arrangements	19	Properties owned, current, identifiers

Legal Persons/Arrangements	19	Properties owned, historic, identifiers
Legal Persons/Arrangements	22	Lender regulatory regime
Properties	02	Address
Properties	03	Type
Properties	04	Owners, identifiers
Properties	05	Sales, historic, identifiers
Properties	06	Occupied by owner, flag
Properties	07	Vacant, flag
Properties	09	Number of vacant units
Properties	12	Renovations, extent
Loans/Mortgages	05	Lenders, identifiers
Loans/Mortgages	06	Size
Real Estate Transactions	02	Property, Identifier
Real Estate Transactions	03	Buyers, identifiers
Real Estate Transactions	04	Sellers, identifiers
Real Estate Transactions	05	Date of transaction being finalized
Real Estate Transactions	07	Sales price
Real Estate Transactions	19	Amount of down payment
Real Estate Transactions	22	Sale is an assignment (pre-construction)

5.9.4.2 Suitability of BC TACS data within an AML data framework

Accessibility: Data from the various real estate related tax measures are used for the assessment of provincial taxes and shared with the CRA, implying that the data is in a digital format and is exportable.

Strength of identifiable information: These are strong and include a person's SIN, date of birth, and/or Individual Tax Number (ITN) for individuals, and business number, incorporation number, and/or trust number for legal persons and arrangements.

Coverage of desired populations: Very high but not complete. The data available from the Property Transfer Tax includes properties that were transferred from 2012 onwards. As such, the coverage of properties owned by persons may be incomplete and historical coverage will be sparse.

Additionally, there are geographic carve-outs for some of the taxes, which could be utilized by money launderers. The SVT only covers buildings with less than 4 rental units, so wealthier money launderers could be excluded. Finally, the data only covers BC.

Fit between data dimension and analytical concepts: One analytical concept of interest is whether the owned property is underutilized, for example by being left vacant. The BC Speculation and Vacancy tax intends to collect this same information, but is largely not applied to multi-unit rental properties. For covered population of properties with fewer than four rental units, it may be difficult to determine whether vacancy is a result of underutilization due to money laundering, or simple inability to consistently find tenants.

For condo assignments, the fit is high as the tax data seeks to identify the same type of record-dodging behavior utilized by preproduction condo ML schemes. As such, the fit between this data and its use within an AML data framework is moderate.

For data relating to the Property Transfer Tax, the fit is high as the information is collected during the legal sale or transfer of ownership of a property. Limited historical data can diminish the fit for the assessment of properties owned by natural and legal persons. Additionally, the financial information collected through the Property Transfer Tax can be leveraged for a few indicators, such as whether the property is acquired without a mortgage. The lack of additional financing details limits the use of this data for additional mortgage-related indicators.

Accuracy of data recorded: The accuracy of the data is unknown. Vacancy rates are self-reported and only verified in an audit. Satellite family status is also self-reported, and in some cases nearly impossible to verify as Canadian institutions do not have a reliable way of knowing income received and kept abroad.

Enforcement of the Speculation and Vacancy tax may be limited by lack of related data points. For example, BC Finance does not have access to international transfer data for BC property owners, nor up-to-date data on PR status changes. Thus, the accuracy of the self-reported data is difficult to test and improve.

5.9.5 BC Financial Services Authority

The BC Financial Services Authority (BC FSA) (formerly Financial Institutions Commission (FICOM)), regulates credit unions, insurance and trust companies, pensions, and mortgage brokers. The Real Estate Council of BC (RECBC) and the Office of the Superintendent of Real Estate (OSRE) will be integrated within the BC FSA, and as such, the organisation will also regulate real estate agents and brokers.

The BC FSA houses the Mortgage Broker Tracking System (MBTS), the agency's most relevant data source for AML purposes. This system is a registrar of all mortgage brokers conducting business in the province, and includes personal information such as their date of birth, and their business and home addresses. Mortgage brokers can play a role in money laundering in real estate due to the nature of their work and as such, ML schemes could be detected with high quality broker information. The records extend back to the 1980s, but currently, there is no way to link mortgage brokers to properties without conducting an individual investigation. It follows that the utility of this data for AML purposes is presently limited.

Real estate agents may also facilitate money laundering due to the nature of their work. As such, linking agent information to transactions could further bolster an AML data framework. The RECBC collects information on real estate agents and brokers including personal information, such as their date of birth, practice number, licensing history, and brokerage. Since every agent works within a brokerage, the RECBC also collects the incorporation number, legal name, and brokerage activity. The data also includes a criminal history indicator (no record, minor, serious), as well as a discipline history and enforcement information when a complaint has been filed. The RECBC has limited transaction data based on its ongoing audits of brokerage firms.

The BC FSA also houses a database named Watchdog, which is used to organize data related to examinations and investigations into mortgage-, insurance-, and real estate brokers, on a case-by-case basis. This has constrained uses for AML purposes due to its limited coverage.

5.9.5.1 Relevant BC FSA Data Points

Table	DP #	Description
Real Estate Transactions	17	Involved mortgage and real estate brokers/agents, identifiers

5.9.5.2 Suitability of BC FSA data within an AML data framework

Accessibility: The MBTS data is housed in an SQL database with a custom ASPX/Vb.net layer over top, making the data parsed and exportable. Data on investigations is housed in a system built on Lotus Notes and consists mostly of links to files (such as PDF files and Word documents) organized in various structures, depending on the case. This data would be difficult to use systematically due to the irregular structure and mix of file formats and internal file structures. Most of the historical case data has been stored in physical files.

The data collected by the REBC is held within an extractable SQL database housed in Oasis, a semi-customized application. As such, the accessibility should be high.

Strength of identifiable information: Both mortgage broker and real estate agent identifiers are strong, and include date of birth, number from a government-issued ID, business address, Mortgage Broker Registration Number and known aliases. Corporate registrations include a corporate registry number as well. There is currently no way to link brokers or real estate agents to property transactions without manual intervention to link brokers and agents to clients and their properties.

Coverage of desired populations: The coverage of the MBTS is high but not complete; all registered mortgage brokers are in the database. Registrations must be renewed every two years. However, sales may still be illegally brokered by non-registered persons, undertaken without the use of any third parties, or by using by using a registered broker as a front to submit to lenders.

The coverage is also high for real estate agents and brokers. All licensed real estate agents are in the database and compliance audits on real estate brokerages are performed every four years. It is illegal to purchase a property through a non-licensed agent, and the Office of the Superintendent of Real Estate (OSRE) investigates such activities.

Fit between data dimension and analytical concepts: The analytical concept of interest is facilitators of money laundering, namely mortgage brokers and real estate agents. The information is not currently linked to real estate transactions. The fit between this data and its use within an AML data framework is expected to be high if broker and real estate agent information can be accurately linked to real estate transactions.

Accuracy of data recorded: The data consists of mandatory professional registrations and as such should be very accurate. There are no known systematic inaccuracies in this data.

6 DATA AVAILABILITY

Chapter 4 provided an overview of methods of money laundering in the real estate sector, while Chapter 5 provided a summary of data held by data work stream members and other relevant public institutions. This chapter builds upon these two chapters to provide a high level overview of which ML sub-schemes are detectable with data currently held by the relevant public institutions cited.

6.1 KEY POINTS

- Many of the indicators for ML schemes and sub-schemes described in the real estate framework cannot be constructed from data currently held by Work Stream member institutions and other relevant public institutions that were part of this review.
- Notable data gaps exist for indicators relating to mortgages, wealth, and relationships of owners.
- ML schemes that are detectable via data would require more data sharing and linking for further analysis.

6.2 DATA AVAILABILITY BY SUB-SCHEME

This section presents the results of a qualitative assessment conducted to measure the ability to construct each of the indicators identified for money laundering in real estate, based on the availability of the required data points within the current data holdings of public institutions.

It looks at the existence of data and the accessibility or readiness to use data (e.g., is it in digital format), but not the feasibility of acquiring, sharing or linking the data, nor the intended data use (strategic or tactical). Strategic and tactical use of data are considered later.

Recall that the data framework is based on four levels of data granularity, from highest (most general) to lowest levels (most specific) starting with schemes, sub-schemes, indicators and finally data-points.

The following analysis summarizes the availability of data to calculate indicators at the sub-scheme level. This aggregation may obscures some details on the availability of the data for the component indicators. Additional details, and a full list of availability by indicators, is available in Appendix 3. The summarization uses the following approximated categories:

> **Fair to Good data availability:** data required as an input to the calculation of the ML indicators for the sub-scheme are entirely or mostly held within public data holdings. It should be possible to construct indicators intended to detection of ML using this sub-scheme.
>
> **Poor data availability:** data is available for some of the indicators, but not all of the indicators in the sub-scheme. It may be possible to construct ML indicators for this sub-schemes, but some indicators lack required data, and greater operational difficulties are expected.
>
> **Mostly Unavailable:** Most or all of the required data to calculate the indicators for the sub-scheme does not exist, is not held in a digital format, or is not available within public data holdings. It is likely not possible to construct indicators for detect ML using this sub-scheme based on the data's limited availability.

Data for six of the 22 sub-schemes are held in public institutions and rated as Fair to Good:

1. Using a shell company or trust for obscuring beneficial ownership: most of the associated data points are available within the Business Register maintained at Statistics Canada and through tax data. Foreign corporations may not exist on the register, and the validity of beneficial ownership information on the register has not been sufficiently tested.
2. Lending the money to oneself: the data points for this sub-scheme are required for the lender and most of them are available within the Business Register maintained at Statistics Canada and through tax data.
3. Undervaluing a property purchase and paying the difference under the table: baselines from which to find outliers can be made from BCA property assessment and sales data.
4. Buying a property and selling it as soon as possible: all sales in BC are recorded at the LTSA and BC TACS, so it is possible to spot properties with frequent activity.
5. Purchasing a property from a co-conspirator and selling it back to them: current and former owners of a property is part of the land title record maintained by the LTSA and BC TACS.
6. Purchasing a pre-construction condominium and selling it before it is completed: these sales are recorded in the BC Condo and Strata Assignment Integrity Register (CSAIR)[26].

Table 6-1: Overview of data availability by money laundering scheme and sub-scheme

Scheme	Sub-scheme	Data Availability
To unlawfully obscure beneficial ownership	…by using a shell company or trust.	Fair to Good
	…by using a formal or informal nominee owner.	Mostly Unavailable
	…by having an unclear address.	Poor
To funnel cash/money through mortgages	…by acquiring many mortgages.	Poor
	…by repeatedly mortgaging a property.	Mostly Unavailable
	…by acquiring an outsized mortgage.	Poor
	…by lending the money to oneself.	Fair to Good
	…by acquiring very valuable property with a large mortgage.	Poor
	…by taking over existing mortgages.	Mostly Unavailable
	…by making mortgage payments in cash.	Mostly Unavailable
	…by undervaluing and paying the difference under the table.	Fair to Good

[26] As of January 1st 2019, CSAIR requires developers of condos and strata in BC to collect and report all residential purchase agreements, including both pre-sale and completed lots. This database is held within the BC Ministry of Finance.

To quickly convert dirty funds into a real estate asset	…by buying the house outright with dirty or mixed funds.	Poor
	…by defaulting on a loan to one-self or an associate.	Mostly Unavailable
To flip houses to turn dirty funds into clean funds	…by buying a property and selling it as soon as possible.	Fair to Good
	…by conspiring with the owner and selling it back to them.	Fair to Good
	…by renovating a property with cash/dirty funds.	Mostly Unavailable
	…by purchasing a pre-construction condo and reselling it before the completion date.	Fair to Good
	…by selling an existing property on assignment	Mostly Unavailable
To hide capital in other jurisdictions	…by purchasing property in a country with strong property rights.	Poor
To purchase property for criminal use	…by purchasing property in an area of interest.	Poor
To launder money through cancelled real estate transactions	…by paying a deposit, reneging, and receiving a clean refund.	Mostly Unavailable
To acquire an income source and mix in dirty money	…by purchasing rental property and padding rental income.	Mostly Unavailable

6.3 DATA SHARING EFFECTS FOR STRATEGIC AND ENFORCEMENT FRAMEWORKS

Prevailing legislation clearly distinguishes enforcement efforts from strategic ones, and as such, these are considered independently in the following section.

6.3.1 Data sharing for strategic insights

The data framework can enhance strategic insights to produce informative and actionable guidance, bolstering Canada's AML regime. The objective would be to produce estimates and metrics of money laundering in real estate at the aggregate level, to be disseminated to authorized governmental entities and policymakers.

For example, new indicators could allow for the analysis of the relationship between policy initiatives and outcomes. Additionally, a strategic framework could produce aggregate figures on geographical "hot spots" for ML in the real estate sector, as well as aggregate information on the characteristics of launderers, such as sex, age, and declared income. The magnitude of ML in the real estate sector could

be assessed quantitatively, including the number and value of the properties that may be implicated in ML, through a red flag analysis of the schemes and sub-schemes.

Research could also focus evaluating the effectiveness of the typologies presented, identifying which indicators are most relevant, and which data may be most gainfully analyzed for enforcement purposes. Feedback on indicators and sub-schemes from participating institutions would supplement the framework to further validate what is relevant and narrowing the scope of relevant data points could also serve to mitigate invasion of privacy concerns. Further work could generate additional typologies, as ML evolves with new technologies and schemes.

Table 6-2 presents data sharing option scenarios and how these may contribute to calculation of ML indicators. This is meant as a template of detectable sub-schemes with current data holdings while increasing data sharing among public institutions that focus on research. The ability to detect schemes and sub-schemes will change as data gaps are filled and additional schemes and sub-schemes are added to the framework.

As indicated in Table 6-2, limiting the extent of data sharing sharply restricts the sub-schemes which can be detected. The scenarios examined are as follows:

1. Analysis of Statistics Canada's data holdings, including tax data held by Statistics Canada.
2. Analysis of BCA and BC LTSA's data holdings.
3. Analysis of data within Statistic Canada's data holdings, with the addition of BCA and BC LTSA data.
4. Centralization of all data held at public institutions or sharing between all identified institutions. This scenario is identical to the one underlying **Error! Reference source not found.**.

Utilizing the data currently available within Statistics Canada would allow for the detection of one sub-schemes, with an additional six sub-schemes receiving a "Poor" rating. Similarly, accessing BCA and BC LTSA data would be sufficient to produce indicators for three sub-schemes, and possibly one additional sub-scheme (which has a "Poor" data availability).

Analyzing Statistics Canada data, supplemented with BCA and BC LTSA data, increases the number of detectable sub-schemes to four, with an additional seven sub-schemes having partly missing data for some of their indicators. A centralized model with data collected from all identified public institutions would allow for the generation of strategic insight on five sub-schemes.

Table 6-2: Data availability by money laundering scheme and sub-scheme and data sharing scenario focused on generating strategic insights

		Data Accessed			
		Statistics Canada data only	**BCA and BC LTSA data**	**BCA, BC LTSA and Statistics Canada**	**All data held at public institutions**
Scheme	**Sub-scheme**	**Data availability**			

To unlawfully obscure beneficial ownership	…by using a shell company or trust.	Fair to Good	Mostly Unavailable	Fair to Good	Fair to Good
	…by using a formal or informal nominee owner.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by having an unclear address.	Poor	Poor	Poor	Poor
To funnel cash/money through mortgages	…by acquiring many mortgages.	Mostly Unavailable	Poor	Poor	Poor
	…by repeatedly mortgaging a property.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by acquiring an outsized mortgage.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Poor
	…by lending the money to oneself.	Poor	Mostly Unavailable	Poor	Fair to Good
	…by acquiring very valuable property with a large mortgage.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Poor
	…by taking over existing mortgages.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by making mortgage payments in cash.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To quickly convert dirty funds into a real estate asset	…by undervaluing and paying the difference under the table.	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good
	…by buying the house outright with dirty or mixed funds.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Poor
	…by defaulting on a loan to one-self or an associate.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To flip houses to turn dirty funds into clean funds	…by buying a property and selling it as soon as possible.	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good
	…by conspiring with the owner	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good

	and selling it back to them.				
	…by renovating a property with cash/dirty funds.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by purchasing a pre-construction condo and reselling it before the completion date.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Fair to Good
	…by selling an existing property on assignment	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To hide capital in other jurisdictions	…by purchasing property in a country with strong property rights.	Poor	Mostly Unavailable	Poor	Poor
To purchase property for criminal use	…by purchasing property in an area of interest.	Poor	Mostly Unavailable	Poor	Poor
To launder money through cancelled real estate transactions	…by paying a deposit, reneging, and receiving a clean refund.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To acquire an income source and mix in dirty money	…by purchasing rental property and padding rental income.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable

6.3.2 Data sharing for enforcement purposes

The indicators can also be used to pinpoint individual cases of the known ML schemes, leading to an investigation by law enforcement. AML regime partners, such as the RCMP, FINTRAC, the enforcement arm of the CRA, or a new unit, could leverage this data framework for the identification of the sub-schemes identified.

As with the generation of strategic insights, the ability to produce indicators for the known sub-schemes depends on the data available. Nearly all sub-schemes require additional data, not currently available to AML regime partners.

Table 6-3 summarizes the ability to construct indicators for known money-laundering sub-schemes using various combinations of data. As data available from newer initiatives in BC, such as the Land Owner Transparency Registry (LOTR), are in development, they have been excluded from this analysis. The examined scenarios are:

1. Analysis of data held by FINTRAC or the RCMP. The data availability assessment is identical for both organizations.
2. Analysis of the data held by FINTRAC or the RCMP, combined with housing data from BC Assessment and BC LTSA. The data availability assessment is identical for both organizations.
3. Analysis of the data held by all identified BC institutions. These include all the taxation data housed in BC TACS, and data from BC Assessment, BC LTSA, and BC FSA.
4. Centralization of all publicly held data or sharing between all identified institutions with the exception of Statistics Canada, given the explicit legislative restrictions for using Statistics Canada data to identifying individuals to law enforcement.

The first scheme, "to unlawfully obscure beneficial ownership", warrants further clarification. British Columbia has introduced legal requirements to declare beneficial ownership when registering property. The two main sub-schemes of this scheme, "by using a shell company or trust" and "by using a formal or informal nominee" provide effective obfuscation of ownership. The indicators focus on ways of detecting the use of this scheme when the beneficial owner is not known, or when documents declaring beneficial ownership have been falsified. Detection of undeclared nominee owners is particularly difficult: indicators for the detection of this sub-scheme rely on unavailable data points concerning an owner's wealth, relationships, and mental faculties. Other indicators for detecting this sub-scheme rely on details of the housing transaction which remain unrecorded, such as name changes partway through the process, or the source of the deposit payment. The requirements to disclose beneficial ownership in BC may directly measure the use of this scheme once fully implemented.

The lack of housing-related data limits FINTRAC and the RCMP's ability to detect sub-schemes in the real estate data framework. Supplementing LEA data with housing data from BC Assessment and the LTSA of BC would expand the number of detectable sub-schemes to three, with an additional two sub-schemes having partly missing data.

Indicators capable of detecting five sub-schemes would be possible if data were combined from all BC institutions. An additional six sub-schemes receive a Poor data availability rating in this scenario, implying that some indicators may be possible to construct for these sub-schemes as well. There is significant overlap between the scenarios, in part because few sub-schemes are detectable with currently held data, and in part because much of the detection power comes from the housing data which is included in most of the scenarios.

The scenarios shows that, in general, more sub-schemes are detectable with greater combinations of data. It may, however, be possible to produce indicators at separate institutions with differential access to data and combine the results/flagged cases to gain somewhat more coverage. The number of sub-schemes detectable through such an approach may become greater should some of the data gaps discussed below be filled in.

Table 6-3: Data availability by money laundering scheme and sub-scheme and data sharing scenario focused on generating insights for enforcement purposes

Scheme	Sub-scheme	FINTRAC data only / RCMP data only	BCA, BC LTSA, and FINTRAC / RCMP data	BCA, BC LTSA, BC TACS, BC FSA	All data held at public institutions*
To unlawfully obscure beneficial ownership	…by using a shell company or trust.	Mostly Unavailable	Mostly Unavailable	Fair to Good	Fair to Good
	…by using a formal or informal nominee owner.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by having an unclear address.	Mostly Unavailable	Poor	Poor	Poor
To funnel cash/money through mortgages	…by acquiring many mortgages.	Mostly Unavailable	Poor	Poor	Poor
	…by repeatedly mortgaging a property.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by acquiring an outsized mortgage.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by lending the money to oneself.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Poor
	…by acquiring very valuable property with a large mortgage.	Mostly Unavailable	Mostly Unavailable	Poor	Poor
	…by taking over existing mortgages.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by making mortgage payments in cash.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To quickly convert dirty funds into a real estate asset	…by undervaluing and paying the difference under the table.	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good
	…by buying the house outright with dirty or mixed funds.	Mostly Unavailable	Mostly Unavailable	Poor	Poor
	…by defaulting on a loan to one-self or an associate.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To flip houses to turn dirty funds into clean funds	…by buying a property and selling it as soon as possible.	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good
	…by conspiring with the owner and selling it back to them.	Mostly Unavailable	Fair to Good	Fair to Good	Fair to Good

	…by renovating a property with cash/dirty funds.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
	…by purchasing a pre-construction condo and reselling it before the completion date.	Mostly Unavailable	Mostly Unavailable	Fair to Good	Fair to Good
	…by selling an existing property on assignment	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To hide capital in other jurisdictions	…by purchasing property in a country with strong property rights.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Poor
To purchase property for criminal use	…by purchasing property in an area of interest.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To launder money through cancelled real estate transactions	…by paying a deposit, reneging, and receiving a clean refund.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable
To acquire an income source and mix in dirty money	…by purchasing rental property and padding rental income.	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable	Mostly Unavailable

6.4 DATA LIMITATIONS: DATA NOT AVAILABLE FROM PUBLIC INSTITUTIONS

The most expansive data sharing and access scenario examined above leaves many sub-schemes with "Poor" and "Mostly Unavailable" data availability ratings. This is a result of several key data gaps within the data holdings of Canadian public institutions.

6.4.1 Mortgage data

Financial arrangements are a key modus for money launderers, hence mortgages, being a central real estate financing tool, are critical to the AML data framework: mortgage data could be used in the identification of eight of the sub-schemes and 34 indicators in the data framework.

The Bank of Canada is only the work stream member to access mortgage information in a systematic manner, although the data they use is de-identified. The main data providers of this data are credit rating agencies and land title authorities, which may include private entities such as Teranet and credit bureaus. Credit rating agencies would also typically have little reliable information on mortgage terms aside from the identifiers and mortgage balance information.

Most loan-back schemes, however, are likely to be conducted through non-regulated lenders, and these are generally not included in credit rating agency data.

BC LTSA collects mortgage information within British Columbia, but key pieces of data for several sub-schemes, such as the precise amount borrowed, the interest rate, and the amount outstanding, are missing. Other land title authorities may collect similar information in other provinces.

The CMHC is studying the benefits of establishing data standards in the Canadian mortgage industry (CMHC, 2020). Should such standards be adopted, it may be easier to collect comparable data from various private sector actors involved in the mortgage industry.

6.4.2 Beneficial ownership data

Corporations' ability to act as third parties while being controlled by the launderer lends itself to a multitude of money laundering schemes. In the real estate sector, opaque legal persons form the core of the most well-known money-laundering sub-schemes, including simple ownership obscured through a corporation, and the lending of money to oneself through a series of corporations. Without accurate data on who owns which corporations, it is almost impossible to know who owns which properties, who lends to whom, and who sells to whom.

Information on corporate ownership is available within the CRA, Statistics Canada, and on a case by case basis to law enforcement. However, there are concerns relating to the quality of this data, as the corporations themselves may have nominee owners and/or directors. Moreover, there are no central records of changes in corporate ownership.

Beneficial ownership information can be obtained via production order from financial entities (e.g. banks, credit unions). There are sources available, though not all corporations are captured, and these can only be accessed on a case-by-case basis.

Beneficial ownership registries are valuable to AML efforts. Adding transparency to corporate ownership is key to obtain more information on property owners and creates obstacles to money laundering through real estate. To be effective, a registry should include all corporations, whether federally or provincially incorporated.

Ensuring the accuracy of the information in a beneficial ownership registry is also a challenge. The UK has a public beneficial ownership registry, and this has allowed civil society organizations to evaluate the quality of the data therein. Many discrepancies were discovered, in part because the UK Companies House, responsible for maintaining the registry, does limited verification of the data supplied. For instance, over 10,000 companies declared a foreign company as the beneficial owner – with 72% of these owner companies located in tax havens (Global Witness, 2018). Without higher quality data, the registry becomes much less useful for AML and fraud-related purposes (BEIS, 2019).

Vetting by public-interest vetting is one of the key benefits of making a beneficial ownership registry publicly accessible. A public registry could also provide more timely access and ait in investigating cross-border financial crimes (ISED 2020). However, a public registry also entails greater privacy concerns and potential risks for identity theft, fraud, and harassment (ISED 2020). It is also possible public access to information on the beneficial owners of Canadian corporations would displace licit and illicit activity into other structures with potentially lower reporting requirements, such as trusts or foreign corporations.

British Columbia is taking steps to collect beneficial ownership data. The LOTA will produce a Land Owner Transparency Registry (LOTR) for legal entities and arrangements with an interest in BC land. There is the risk, of course, of corporations or persons not declaring the true owners behind a nominee,

but the threat of criminal sanction arising from prosecution could deter this. Furthermore, BC is has been conducting public consultations on establishing a central registry of beneficial ownership for privately held BC corporations, irrespective of their ownership of land.

6.4.3 Relationship data

In an analysis of convicted money laundering cases in Canada, Schneider (2004) found the use of nominee owners in 61% of cases. Nominee owners can be relatives, friends, business associates, lawyers, and shell companies. Friends and business associates are hard to identify without relatively invasive approaches, but knowledge of an owner's spouse and close relatives would go a long way toward uncovering nominee ownership arrangements.

Tax data is likely the most promising way to create family links by identifying individuals who submit their taxes together, thereby creating "tax families". Current tax data dates back to the 1980s, allowing the identification of most parent-child and spousal relationships in the country. However, other familial relationships may not be captured by the data. The use of tax data for such links also faces limitations with regard to immigrant, transnational families and non-tax-filers. In spite of these difficulties, the CRA is employing some approaches to map relationships between family members, associated businesses, trusts, partnerships, and friends.

Commercially available PEP (politically exposed persons) databases, and court or police incident records, could be used to identify corrupt sources of funds and potential criminal associates.

6.4.4 Individual wealth data

A couple of the indicators rely on identifying a discrepancy between an owner's available income /accumulation of wealth and their acquisition of property. The UK has enshrined this approach in law, empowering courts to issue Unexplained Wealth Orders, obliging property owners to account for the origin of their wealth or have their assets seized.

Wealth information at the level of the individual is not available in any known public data source. Statistics Canada produces estimates of household wealth, but these are in aggregate based on surveys of a small sample of the population and are thus unsuitable for determining the extent of this phenomenon.

A proxy for wealth could be estimated from historical tax records and pension contributions, but this would not directly measure assets and liabilities. Given widely different levels of savings between individuals, any such "wealth approximation" effort would require a significant amount of work and methodological rigour to be of use in an AML framework.

Data on liabilities is not currently available to working group members, but it is to a large extent collected by the credit rating agencies. Such data would be necessary to create a good wealth estimate at the individual level.

6.4.5 Rental income and rental payments data

An ML scheme that is particularly difficult to detect involves purchasing rental property and padding ostensible rental income with the proceeds of crime. This can be done either by declaring rent for unoccupied units or by under-charging on rent and making up the difference with the proceeds of crime.

Companies and individuals are currently required to report their rental incomes, but there is very little corresponding reporting coming from tenants[27]. Furthermore, where multiple properties are owned, rental income is aggregated across all properties. It is very difficult to estimate when reported rental income is fraudulent without more granular data.

6.4.6 Data on non-residents

The data held by Canadian institutions cover mostly residents of Canada. This limits the detection of money laundering which can be transnational. Non-resident owners of Canadian properties have unknown incomes, unknown spouses, and those spouses have unknown incomes. Financial transactions between non-residents property holders or lenders in other countries are unrecorded.

Canada does have tax information exchange agreements with various foreign jurisdictions. This exchange is typically conducted in regard to a specified situation identified by the CRA, or for identified individuals.

6.5 Effect of filling data gaps with further public institutions data

Table 6-4 performs the same summary of data availability as presented in **Error! Reference source not found.** under the assumption that high quality structured data is available for the key data gaps. The scenario assumes these are filled for the more key mortgage data points, as well as data points for beneficial ownership, relationships, and rental payments[28]. Data gaps estimated to be more difficult to fill, such as wealth, itemized mortgage payments, or interactions with real estate agents, were left unfilled.

Not all sub-schemes reach a Fair to Good level of data availability under these assumptions, but significantly more do.

Table 6-4: Data availability with key data gaps filled

Scheme	Sub-scheme	Data Availability
To unlawfully obscure beneficial ownership	…by using a shell company or trust.	Fair to Good
	…by using a formal or informal nominee owner.	Mostly Unavailable
	…by having an unclear address.	Poor
To funnel cash/money through mortgages	…by acquiring many mortgages.	Poor
	…by repeatedly mortgaging a property.	Fair to Good
	…by acquiring an outsized mortgage.	Poor

[27] Tenants in Quebec must declare their rental expenses on their annual tax returns. Additionally, tenants in Ontario qualifying for the Trillium tax credit may also declare their rental expenses.

[28] Specifically the data points: M_01, M_04, M_06, M_07, M_08, M_11, M_12, M_14, M_17, LPO_08, LPL_08, NPO_24,NPS_03, P_10, and X_06. See Appendix 1 for descriptions.

	…by lending the money to oneself.	Fair to Good
	…by acquiring very valuable property with a large mortgage.	Fair to Good
	…by taking over existing mortgages.	Fair to Good
	…by making mortgage payments in cash.	Mostly Unavailable
To quickly convert dirty funds into a real estate asset	…by undervaluing and paying the difference under the table.	Fair to Good
	…by buying the house outright with dirty or mixed funds.	Poor
	…by defaulting on a loan to one-self or an associate.	Poor
To flip houses to turn dirty funds into clean funds	…by buying a property and selling it as soon as possible.	Fair to Good
	…by conspiring with the owner and selling it back to them.	Fair to Good
	…by renovating a property with cash/dirty funds.	Mostly Unavailable
	…by purchasing a pre-construction condo and reselling it before the completion date.	Fair to Good
	…by selling an existing property on assignment	Mostly Unavailable
To hide capital in other jurisdictions	…by purchasing property in a country with strong property rights.	Poor
To purchase property for criminal use	…by purchasing property in an area of interest.	Poor
To launder money through cancelled real estate transactions	…by paying a deposit, reneging, and receiving a clean refund.	Mostly Unavailable
To acquire an income source and mix in dirty money	…by purchasing rental property and padding rental income.	Fair to Good

7 DATA MODELS FOR STORAGE AND ANALYSIS

Data and information relating to ML is currently being held by various regime partners and is being shared on a limited basis. These limitations reflect legislative and privacy restrictions based on various organization's operating procedures, their authorizing Acts as well as the Canadian Charter of Rights and Freedoms. As such, organizations are limited in the full use of the real estate data framework for analytical and enforcement purposes.

Three models could be considered to facilitate greater access and sharing of data. The models are intended to take into account the institutional landscape that currently exists. Key to the implementation of these models is the development or availability of new data to close data gaps in the data framework presented.

1. A "Distributed" model where data is held by respective oversight organizations and enhanced access is authorized (or data is shared) among regime partners.

2. A "Centralized" model where data is consolidated and made accessible to the regime partners.

3. A "Hybrid" model where data is organized by separate custodial and analytical functions undertaken by existing or new partners or units. A data custodian would be responsible for collecting, processing, and housing data. A separate coordinating organization or unit would lead access and analysis of the data for AML purposes.

There are pros and cons to each model, the main factors being practical operational considerations and privacy impacts.

7.1 THE DISTRIBUTED MODEL

The distributed model most closely aligns with Canada's current AML regime. Reflecting the mandates, resources available, and practices of each individual organization, data can be shared to specified organizations, typically law enforcement agencies, when there are reasonable grounds for suspicion of money laundering. While this would likely satisfy privacy considerations in the prevailing legislation, it does limits the ability to share and leverage data to its full extent for AML purposes.

The distributed model currently used could be made more effective with increased data sharing which would be contingent on regime partners' having an enhanced capacity and authority to share data. This would entail separate data-sharing arrangements, strong and consistent data management practices amongst all involved parties, backed by the legal authority to do so. A distributed model, even with enhanced sharing and data management practices, introduces an element of duplication by the organizations involved; any shared administrative data would often have to be linked before being analyzed, and this can be a difficult, costly and time-consuming process. Institutions wishing to analyze shared data would each have to independently link the required data together.

7.2 THE CENTRALIZED MODEL

A centralized model would enable the systematic analysis of combined data for the detection of money laundering, with all the data held and managed within one institution. Organizational and operational efficiencies to standardize and avoid duplication of work are key advantages of this model. A centralized model ensures that the processing, cleaning, quality, and documentation of the data is consistent throughout the various sources, and that they are properly linked, leading to reliable and accurate analysis results. This model is ideal for systematic analyses, both for tactical and strategic insights, as all the available data is held within one institution which specializes in AML.

There are also potential operational inefficiencies with a centralized model which could affect the non-AML-operations for institutions contributing data to the centralized data holder. If existing data holders were to cede some or all of the control over their processes of collecting and storing data – in an effort to create consistency across institutions – this may negatively affect their ability to adjust their practices in order to better serve their non-AML mandates. A balance would need to be established between the ability of the centralized data holder to clean the data it receives, and changes to the data-collection processes aimed at ensuring sufficient data quality and coverage.

Privacy concerns have been raised by advocacy groups about the consolidation of data sources such that a centralized model would potentially entail greater privacy and security concerns (Nielsen, 2015, in a report for the Privacy Commissioner of Canada).

7.3 THE HYBRID MODEL

A hybrid model combines aspects of the distributed and centralized models by processing and storing subsets of data within designated data custodian institutions separate from AML responsibilities. A separate organization or unit would perform a coordinating role between the AML regime partners and data custodians to enable access, linking, and analysis of the data, as well as ensure that consistent data management practices are implemented.

The hybrid model would leverage existing public data holders to create key sources of data on subjects related to the AML data framework for real estate, such as real estate transactions, taxes, or professional/criminal relationships and their linkages within newly AML empowered custodial institutions. Custodial institutions would be responsible for data cleaning and processing, as well as maintaining the quality of the data and accompanying documentation. The custodial functions could entail limited and prescribed responsibilities for the analysis of this data but not for the detection of ML. Designated custodial institutions may require additional resourcing to undertake their expanded roles.

The coordinating unit as the locus of AML expertise could be responsible for leading data linkage and access functions, particularly with respect to non-regime partners, as well as supporting analysis being undertaken by regime partners. This unit could also be charged with the responsibility to ensure appropriate scopes and safeguards for any analyses performed. It could lead in the maintenance and development of the typology data framework and for red flag analysis arising with shared data, leading to a better assessment and understanding of ML as practised in Canada. This unit could also lead the development of metrics aiming to measure the effectiveness of Canada's AML regime and of relevant policy interventions

The location of the coordinating unit would need to be decided based on best fit. It could potentially reside within an existing regime partner such as FINTRAC, be a stand-alone unit, or consortium of regime partner members along the lines of the ACE Fusions group. A consortium would be in line with the need for enhanced co-operation and co-ordination of efforts to combat ML as recommended by academic experts such as Brigitte Unger and the report on Combatting Money Laundering in BC (Maloney et al., 2019).

There are several advantages to the hybrid approach, chief of these is its potential to maintain privacy and security. Sensitive data would be held and maintained separately by respective custodial institutions and AML regime partners compared to the centralized model. The coordinating unit could conduct its analysis of the various custodial databases designated for investigative purposes, and share suspicious transactions with law enforcement, in keeping with data sharing practices currently in use among regime partners. Second, data gaps can be filled at the custodial institution where subject-matter expertise and strong data management practices are already established. New data can be efficiently incorporated into existing analyses, leveraging procedures for accessing and linking data from the custodial institution. Third, the additional data collected by the custodial institution could be used for non-AML research and policy purposes, where appropriate. Finally, the use of privacy enhancing technologies can be explored within the hybrid approach in order to strike the right balance between AML and privacy concerns. See Appendix 4 for a brief note on Privacy Enhancing Technologies.

Several details must be considered for the implementation of a hybrid model such as the development of information sharing mechanisms, the designation of data custodians and a review of existing legislative authorities. Storage, processing and maintenance among custodians should be consistent, and developing standards would ensure uniformity among participating institutions. Full or limited access may be granted to regime partners based on their responsibilities and mandate.

In its ability to leverage data for the detection of ML in real estate, the hybrid model is the closest approximation to the centralized model. More analysis would be required to assess the extent to which this model would mitigate privacy and security concerns.

8 INTERNATIONAL AML DATA MODELS

A data model is a system for collecting, storing, accessing, and analyzing data in the production of a specific output. The collection and use of data is central to modern AML regimes which rely on various data models for producing both actionable intelligence and a better understanding of money laundering activities.

The types of data collected, the ways in which it is analyzed and how and to whom data is shared, and differ in small and large ways between countries. The following provides an overview of the AML data practices in Canada, Australia, the UK, the US, and the Netherlands. These countries were chosen for their cultural and institutional similarities with Canada, and for their leadership in AML practices.

8.1 KEY POINTS

- Canada is evolving its public-private partnership strategy to enable a more structured approach to the exchange of information on particular cases or threats between law enforcement and financial institutions. This approach is similar to established programs in the UK, US, Netherlands and Australia.
- The sharing of data between public institutions is most systematic in the Netherlands where iCOV automates authorized information requests between institutions.
- Large scale data analysis programs which combine and analyze encompassing data sets are being developed in the UK and Australia.
- Access to beneficial ownership information on legal persons and arrangements is most developed in the UK where a public registry has been made available. The EU's 5th AML directive has committed its members to establish beneficial ownership registries.
- Law enforcement agencies (LEAs) can access a database of suspicious/unusual transactions or suspicious activities/matters in all of the foreign jurisdictions studied, excluding Canada. Canada's anti-money laundering and anti-terrorist financing regime was designed to ensure that the rights enshrined in the Canadian Charter of Rights and Freedoms are not compromised.

8.2 OVERVIEW OF INTERNATIONAL AML DATA MODELS

	CANADA	AUSTRALIA	UNITED KINGDOM	UNITED STATES	THE NETHERLANDS
FINANCIAL TRANSACTION REPORTS	Suspicious and threshold transactions. Disclosures are provided to relevant LEAs and other agencies when legal grounds are met.	Suspicious matters and threshold transactions. Accessed by LEAs and other institutions.	Mainly suspicious activities. Accessed by LEAs	Mainly suspicious activities. Accessed by LEAs. Merged with local data at some LEAs	Mainly unusual and suspicious transactions. Accessed by LEAs and other institutions. Part of iCOV.
PUBLIC-PUBLIC DATA COOPERATION	Sharing of some data from CBSA to FINTRAC. FINTRAC receives Voluntary Information Records (VIRs) from LEAs and other agencies.	Fintel Alliance coordinates and facilitates data exchange with 20+ agencies on particular cases and threats.	Sharing of some HMRC data on borders and trade with the NCA.	Extensive array of coordinating institutions and groups.	iCOV infrastructure for automated data sharing between public institutions. Other initiatives.
PUBLIC-PRIVATE DATA COOPERATION	Partnerships that coordinate specific data disclosure exchange on particular predicate offences. Cooperation on development of indicators.	Fintel Alliance coordinates and facilitates data exchange on particular cases and threats.	JMLIT coordinates and facilitates data exchange on particular cases and threats.	FINCEN Exchange coordinates information sharing among LEAs and financial institutions	SCTF coordinates and facilitates data exchange on particular cases and threats. Automated access to some banking data by LEAs.
LARGE SCALE ANALYSIS INITIATIVES	Bulk analysis of other organizations' data at FINTRAC.	Merging of AUSTRAC data with data at other agencies. Pilot project for joint analysis of banking data.	Development of "industrial scale data analytics" at NCA.	Merging and analysis of FinCEN data with LEA data at multiple agencies.	Limited thematic analyses at iCOV on non-merged data. Pilot project to combine data sets at Dutch FIO and FIOD. Development of AMLC suite of data tools.

BENEFICIAL OWNERSHIP REGISTRIES	Registry of property owners being deployed in BC. Work underway to increase beneficial ownership transparency at the federal level	No registry of beneficial owners.	Publicly accessible beneficial ownership registry.	No registry of beneficial owners.	Plans for a beneficial ownership registry.

8.3 CANADA AML DATA MODELS

The Financial Transactions and Reports Analysis Centre (FINTRAC), Canada's Financial Intelligence Unit, uses data to detect, prevent, and deter money laundering and terrorist financing. FINTRAC analyzes data received from reporting to identify transactions and patterns that are suspected of relating to money laundering, terrorist activity financing or other threats to the security of Canada. They may proactively issue disclosures to enforcement agencies when a money laundering or terrorist financing offence, or a threat to the security of Canada is suspected, given that legislative thresholds are met. The legislative threshold specifically requires reasonable grounds to suspect that a transaction is relevant to a money laundering or terrorist financing offence. FINTRAC's analysis of reports is enhanced by two mechanisms. First, law enforcement agencies can inform FINTRAC through voluntary information records about ongoing investigations with a perceived ML or terrorist financing nexus which can heighten the scrutiny on pertinent transactions. Second, focused public-private partnership projects around particular predicate offences leads to tagged Suspicious Transaction Reports (STRs) which help guide analysis.

8.3.1 Financial transaction reporting

FINTRAC's primary source of data is mandatory reports received from a large list of reporting entities. These entities must monitor and report suspicious transactions through suspicious transaction reports (STRs). In addition, reporting entities are required to submit a large cash transaction report (LCTR) when a reporting entity receives $10,000 or more in cash in the course of a single transaction, or when it receives two or more cash amounts totaling $10,000 or more made within 24 consecutive hours by or on behalf of the same individual or entity. Certain reporting entities are also required to submit an electronic funds transfer report (EFTR) upon a transmission of instructions from a client for the transfer of $10,000 or more out of or into Canada in a single transaction or in two or more transactions totaling $10,000 or more made within 24 consecutive hours by or on behalf of the same individual or entity. Furthermore, FINTRAC also receives other types of reports such as Terrorist Property Reports (TPRs), Casino Disbursement Reports (CDRs) and Cross Border Currency Reports (CBCRs). FINTRAC received over 235,000 STRs; 10 million LCTRs; 17 million EFTRs; 200,000 CDRs, and 60,000 CBCRs in 2018-19 (FINTRAC, 2019).

The data collected by FINTRAC is not shared with other institutions, given Charter and legislative considerations. Instead, it is analyzed solely by FINTRAC and a disclosure is made to law enforcement or

other relevant agencies when the legislative thresholds are met.[29] The majority of disclosures are released to the RCMP. Disclosures are also often sent to multiple enforcement agencies when relevant legal thresholds are met. Provincial and municipal police forces, and the Canadian Security Intelligence Service also received several hundred disclosures annually, with a few hundred going to the Canada Border Services Agency (CBSA) and the Canada Revenue Agency (CRA) as well as provincial securities agencies and the Competition Bureau (FINTRAC, 2019).

Law enforcement and other agencies, as well as the public, can send FINTRAC Voluntary Information Records (VIRs) on individuals or entities who may be suspected of activity related to money laundering or terrorist financing. This aids FINTRAC's efforts in analyzing the data and may lead to subsequent disclosures to the law enforcement agency that originally submitted the information, or to other agencies, as prescribed by the legislation. Over 2,700 such records were provided in 2018-19.

8.3.2 Public-public data cooperation

There are currently no formal public-public data cooperation initiatives in Canada. However, Canadian public institutions do collaborate and share information when legally required to do so. For example, the CBSA sends data on cross-border movements and seizures of currency or monetary instruments of $10,000 or more to FINTRAC.

8.3.3 Public-private data cooperation

FINTRAC has engaged in a number of public-private partnership (PPP) projects with reporting entities (e.g. financial institutions) in Canada centered on particular predicate offences. In general, these projects are led by a single private sector institution in cooperation with other institutions from the same or other sectors. With the cooperation of members, FINTRAC develops and disseminates financial indicators, in the form of an Operational Alert, for the given predicate offence, and facilitates the tagging of any STRs generated by reporting entities, as a result of these indicators. Project Protect was the first of its kind, focusing on ML related to human trafficking. As a result of these efforts, nearly 2,000 STRs were submitted under this project in 2016, and many more since. Comparatively, less than 600 STRs relating to human trafficking were received in the previous year. Similarly, in 2015-16, FINTRAC issued 19 disclosures related to human trafficking, but issued 120 such disclosures in 2016-17 and 142 in 2017-18 (FINTRAC, 2018). This growth in disclosures shows the success of the partnership in producing actionable intelligence. In addition, the majority of these disclosures relate to people or businesses which were not flagged to FINTRAC through voluntary information records, and as such they represent new leads for Canadian LEAs.

The success of Project Protect spawned similar projects focused on fentanyl trafficking (Project Guardian), romance fraud (Project Chameleon), and suspicious casino-related transactions (Project Athena). One of the latest projects, Project Athena, began with a focus on casinos and underground banking, but has since expanded and added subgroups focusing on money laundering through the use of luxury goods, luxury vehicles, and real estate. The project has since evolved to focus on lawful information sharing between law enforcement, FINTRAC, financial institutions and relevant regulators. However, certain data sharing barriers exist due to legislative constrains.

[29] In 2018-19, there were around 2,300 disclosures (FINTRAC, 2019). The number of disclosures is small compared to the number of reports received for numerous reasons. For instance, each disclosure can contain as many (even over 100) individual transactions.

8.3.4 Large scale analysis initiatives

FINTRAC combines its individual data sets in order to analyze them in aggregate to detect money laundering.

The CRA also performs targeted aggregate analysis to detect tax fraud and tax evasion, although those activities are not done with the explicit purpose of detecting money laundering.

8.3.5 Beneficial ownership registries

British Columbia has recently implemented a beneficial ownership registry for residential properties. There is also work being done on a parallel registry for beneficial owners of companies incorporated in British Columbia. There is no centralized beneficial ownership registry for federally incorporated companies, nor for companies incorporated in the other Canadian jurisdictions. Consequently, ownership information is often of little use to law enforcement because money launderers often do not maintain direct ownership of the properties and companies they utilize in their illegal enterprise.

8.4 AUSTRALIA AML DATA MODELS

Use of data in the Australian AML regime mainly takes two forms. First, the Australian Transaction Reports Analysis Centre (AUSTRAC), the Australian FIU, collects transaction data from reporting entities using both suspicion and prescribed thresholds. AUSTRAC analyzes these internally and provides access to the data to outside organizations and partners, as appropriate. Second, the Fintel Alliance is an initiative that brings together many public and private institutions, and their differential access to data, for joint efforts to target particular predicate crimes or laundering techniques.

8.4.1 Financial transaction reporting

The Australian Transaction Reports and Analysis Centre (AUSTRAC – the Australian FIU) receives transaction reports from reporting entities, analyzes them, and passes them on to law enforcement if there is evidence of potential criminal activity.

AUSTRAC receives a number of reports from reporting entities. As with all FATF members, these include Suspicious Matter Reports (SMRs), which are submitted whenever a reporting entity suspects that a person or transaction is linked to crime. As in the Canadian AML regime, there are also a number of threshold transaction reports submitted. These are triggered whenever a cash amount over $10,000 is exchanged, whenever cash over $10,000 is transferred across the border, or whenever any amount is electronically transferred into or out of Australia.

Once collected, the data is analyzed by AUSTRAC and other governmental agencies. Within AUSTRAC, algorithms have been developed to “manage inconsistencies between data sets and recognize linkages between them to identify persons of interest.” (AUSTRAC, 2019, p. 16). External entities, including the Australian Federal Police and the Australian Criminal Intelligence Commission have full access to the database. They can query and export data points and conduct their own analyses. Private institutions, such as banks, that have designation through the FINTEL Alliance can also access the data, but only as it pertains to a crime type the group is investigating in a FINTEL Alliance project. Over 2 million searches were performed in 2018-19 by parties external to AUSTRAC (AUSTRAC, 2019, p. *v*). It is unclear how complex of an analysis can be done with these searches.

8.4.2 Public-public and public-private data cooperation

FINTEL Alliance is a partnership between a large number of public and private-sector institutions in Australia. It sets itself apart from other PPPs in part by having physical offices (in Sydney and Melbourne) to which member organizations second staff on a full-time or part-time basis. The shared physical facilities promote knowledge sharing among members through personal interactions and limited, monitored access to AUSTRAC's database. This leads to a deeper understanding of capabilities and constraints and is conducive to conceptualizing achievable projects for the Alliance (Chadderton & Norton, 2019, p. 27).

The Fintel Alliance have completed a number of projects since they were established and have published some metrics on the number of "detections" the collaborations have led to. For example, their 2018-19 annual report claims a six-fold increase in the number of child exploitation detections and a four-fold increase in outlaw motorcycle gang detections (Fintel Alliance, 2019).

The project process begins by when a target is proposed by one or more members and approved by the Strategic Advisory Board. At this point, AUSTRAC issues compulsory production orders to relevant institutions – both public and private – which have the required data. The compulsory nature of the orders legally shields private sector organizations who would otherwise be more cautious or face legal liabilities when sharing data on their clients (Chadderton & Norton, 2019, p. 17).

The assembled data is analyzed and is either directly acted upon or is used to generate a typology report which is then sent out to reporting entities. Reporting entities can then submit SMRs referencing the particular suspicion, such as Child Sexual Exploitation, covered in the report.

The Fintel Alliance has commenced or completed 49 projects since it was established in 2017. Most of these projects list only public agencies and institutions as the "main partners". (Fintel Alliance, 2019). Other partners will be brought in on an ad-hoc basis, depending on the targets and type of project. This is especially the case if the project is after a more niche area or crime type.

8.4.3 Large scale analysis initiatives

AUSTRAC's data has been combined with the datasets of other government agencies so that they can perform their own analyses on the combined data, In addition, AUSTRAC devotes resources to ensuring adequate record linking in these situations. For example, AUSTRAC worked with the Department of Human Services (DHS) to link DHS data with transactions data in order to detect welfare fraud (AUSTRAC, 2019, p. 24). Similar linking with tax data aided in 4,500 Australian Tax Office cases, raising ATO's recovered funds by over AUD 200 million (AUSTRAC, 2019, p. 24).

The Innovations Hub of Fintel Alliance is also working on a future approach to large-scale data monitoring and use. One central project involves developing ways to analyze disparate and distributed datasets for risks which only become visible when all sets are joined, but to do so in a privacy-preserving manner using Privacy Enhancing Technologies (PETs). The goal is to "enable real-time detection and alerting of criminal activities, using a secure system that supports privacy-preserving data matching and machine learning" (AUSTRAC, 2019, p.54). If successful, the Fintel Alliance will effectively be able to combine multiple data sets from private entities without centralizing the actual data or leaking information between data holders. See Appendix 4 for a brief overview of PETs.

8.4.4 Ownership registries

Records identifying the immediate owners of most corporations are maintained at the Australian Securities and Investment Commission, although many are also maintained at the State or Territorial levels where they were incorporated. However, no registry of beneficial ownership information is currently maintained in Australia.

8.5 UNITED KINGDOM AML DATA MODELS

The use of data in the UK AML regime is distinguished in four ways. First, there is widespread bulk access to the United Kingdom Financial Intelligence Unit (UKFIU)'s Suspicious Activity Report (SAR) database, which at least 77 organizations can use. Second, financial institution data is often contributed to individual cases through the Joint Money Laundering Intelligence Taskforce (JMLIT), the UK's public-private partnership initiative. The National Crime Agency (NCA) is also developing a future capability for "industrial scale data analytics", but details are at this stage sparse. Finally, the UK has built one of the world's first public register of beneficial owners for businesses and legal entities.

8.5.1 Financial transaction reporting

A large number of reporting entities are required to submit Suspicious Activity Reports (SARs) to the UK FIU, which is housed within the NCA. SARs have a lower reporting threshold than STRs in Canada, and the information received through SARs is more limited. The UK FIU then makes these SARs available in bulk to many national and regional LEAs, such as the London Police and the Serious Fraud Office, and to non-LEAs such as the Department for Environment, Food, and Rural Affairs (DEFRA). In 2018, the SAR database was available to nearly 5,000 accredited end-users in 77 different organizations (FATF, 2018, p. 43).

Analysis of SAR data is mostly done at these law enforcement and other institutions, who query the data for ongoing cases and reports relevant information to their jurisdiction. DEFRA regularly does a query for its acronym, which would bring up SARs related to both DEFRA employees and DEFRA financing. DEFRA also searches the SAR database for any links to ongoing investigations they are undertaking in connection with their regulatory duties (NCA, 2019e). Thousands of financial investigative personnel have received financial intelligence training from the NCA, enabling the routine use of STRs in their investigations (FATF, 2018, p. 45). The Mutual Evaluation Report on the UK notes that "even the smaller police forces have specialist financial investigators which enhances their ability to use financial intelligence in investigations." (FATF, 2018, p. 6). Some bulk analysis is also undertaken by police (NCA, 2019c).

The UKFIU largely devotes its resources to processing two specialist types of SARs: Defence Against Money Laundering (DAML) SARs and Fast-Track SARs.

A DAML SAR[30] provides some legal protection to the reporting entity. In return, the reporting entity must delay the activity or transfer until a consent is received from the UKFIU, which may take up to 7 business days. A DAML SAR will often trigger a brief investigation, which may result in charges being placed against the involved parties. (FATF, 2018, p. 49; NCA, 2019d; 2019f).

[30] There is also a parallel Defense Against Terrorist Financing (DATF) SAR.

A SAR is fast-tracked to law enforcement when there is suspicion of an immediate or ongoing crime involving vulnerable persons, such as when an elderly person is in the process of being defrauded, or when there is a suspicion of human trafficking (NCA, 2019b; 2019c).

The UKFIU received 443,000 regular SARs and 34,000 DAML SARs in 2018-2019. Around 3,700 SARs were fast-tracked (NCA, 2019b).

To improve SAR reporting among professionals, the UK has established an Office for Professional Body Anti Money Laundering Supervision (OPBAS) (FATF, 2018, p. 6)[31]. Its objective is to ensure that the existing private supervisors for the various professional bodies meet the standards required by the AML regulations.

8.5.2 Public-public data cooperation

The NCA regularly accesses Her Majesty's Revenue and Customs' (HMRC) cross-border cash reports and imports and exports database. However, it is unclear exactly how this is used, nor whether it is cross-checked against STRs (FATF, 2018, p. 52).

A newly established National Economic Crime Centre (NECC), housed within the NCA, is intended to lead and coordinate efforts on combating economic crime in the UK. The NECC acts as a centre point for information-sharing across multiple agencies using what is known as the "Section 7 Gateway". Section 7 is a section of the Crime and Courts Act which states that "A person may disclose information to the NCA if the disclosure is made for the purposes of the exercise of any NCA function." It also allows the NCA to use this information for any other of its functions, and to disclose the information to third parties (Crime and Courts Act, 2013).

8.5.3 Public-private data cooperation

The Joint Money Laundering Intelligence Taskforce (JMLIT) first became operational in 2015 and, as such, is a pioneer in public-private partnerships for law enforcement in the AML space. It gathers state and private actors, particularly police and banks, to cooperate on combating ML and related crimes.

In practice, JMLIT allows for law enforcement to gain timely access to private party data, such as banking information, by keeping an open line of communication between law enforcement and reporting entities. These private parties may then elect to submit relevant information and data from their own data holdings to the NCA through the "Section 7 Gateway" and/or through a SAR to the UK FIU. The use of this gateway is facilitated by having a member of the NCA present during JMLIT meetings, to which participants are allowed to discuss otherwise restricted information.

Sharing sensitive information through JMLIT is generally employed on individual cases, which are brought to what is called the Operations Group. The long-standing cooperation has given NCA agents a good understanding of the data holdings of the member financial institutions. In 2018-19 JMLIT supported over 140 investigations (NCA, 2019b).

There is also a number of Strategic Working Groups within JMLIT in which strategic financial intelligence is shared for the purposes of threat trend assessment and typology development. These groups issue

[31] Of note, OPBAS oversees professional bodies such as the Chartered Institute of Taxation, the Association of Chartered Certified Accountants, as well as the Law Society / Solicitors Regulation Authority, the Law Society of Northern England, and the Law Society Scotland.

alerts to banks with corresponding codes which the banks are to include in any related SARs that they submit.

8.5.4 Large scale analysis initiatives

The UKFIU focuses its analytical efforts on processing DAML and Fast-Track SARs, and as such, conducts limited large scale analysis of submitted STRS. This is a point of critique raised in the most recent FATF Mutual Evaluation Report (FATF, 2018, p. 6).

The NCA has received additional funding to develop a National Data Exploitation Capability (NDEC). It intends to provide "industrial scale data analytics" and "be the foundation for a suite of data driven products and services that will underpin access to and understanding of intelligence relating to national security." (NCA, 2019a). There are numerous potential uses of this service for anti-money laundering, though as it is still in development, it is too early to determine its impact in practice.

As noted, law enforcement agencies have access to SAR data and can perform bulk analyses on these. A number of agencies use this access to produce strategic reports and analyses (FATF, 2018, p. 55).

8.5.5 Beneficial ownership registries

In 2016, the UK created the register of People with Significant Control, one of the world's first public beneficial ownership registers. It makes it easy for law enforcement and others to look up beneficial ownership data on UK companies, limited liability partnerships and similar entities. The submitted data is neither validated nor verified (BEIS, 2019), so both unintended errors and fraudulent entries could be a significant issue. The public nature of the register, however, does allow third parties to identify and report discrepancies. For example, the Non-Governmental Organization (NGO) Global Witness analyzed the data and issued a report highlighting various inconsistencies implying faulty or fraudulent reported data (Global Witness, 2018).

8.6 UNITED STATES AML DATA MODELS

The United States financial system plays a significant role within the world economy. Specifically, its size, complexity, international reach and plethora of innovative financial products are subject to infiltration by money launderers (FATF, 2016b) who also seek to wash their gains abroad corrupting other marketplaces.

The United States has put emphasis on their anti-money laundering and anti-terrorism financing regime since the September 11 2001 attacks. The Financial Crimes Enforcement Network (FinCEN,) the United States FIU, has modernized its data framework to ensure that government agencies and LEAs have timely access to reports received by the intelligence unit. The United States has focused on public-public partnerships and inter-agency collaboration with initiatives such as the SARs Review Teams, the Financial Crime Task Forces, and the Fusion Centres. Additionally, the FinCEN Exchange was recently established with the aim of enhancing information sharing between public and financial institutions. The United States has also leveraged Geographic Targeting Orders to fill in data gaps within their framework.

8.6.1 Financial transaction reporting

FinCEN, housed within the United States Department of Treasury, is the country's financial intelligence unit responsible for combatting ML and other financial crimes. Reporting entities include financial institutions, casinos, and money services businesses but are missing key sectors such as lawyers, accountants, and real estate agents (FATF, 2016b). Reporting entities are mandated to report suspicious transactions to FinCEN through SARs and other reports such as Currency Transaction Reports (CTRs).

The main pieces of legislation on money laundering in the United States are the Bank Secrecy Act (BSA) and the USA PATRIOT Act. The BSA mandates designated entities to report cash transactions exceeding $10,000 USD and suspicious activity to the Financial Crimes Enforcement Network (FinCEN), regardless of their Federal or State registration (FATF, 2016b). This information is housed within FinCEN's database of the Bank Secrecy Act. Additionally, financial institutions must establish and maintain information on the beneficial owners of legal entities (FATF, 2020).

The USA PATRIOT ACT of 2001 was enacted to strengthen the investigatory tools available to law enforcement as well as implement measures to detect and prosecute money laundering (FinCEN, 2001). Of note, FinCEN can issue section 314(a) requests which obligates financial institutions to search their records for matches to the provided persons or companies and report back on their holdings. FinCEN initiates these requests on behalf of domestic and foreign LEAs (USA PATRIOT Act of 2001). Section 314(b) allows financial institutions to share information with each other in cases of suspected money laundering or terrorist financing for the purposes of reporting the findings to the federal government (USA PATRIOT Act of 2001).

Direct access and use of the BSA database is available to authorities at the local, State and Federal levels (FATF, 2016b). This allows for continued operational and strategic analysis by law enforcement at various levels of government. Approximately 10,000 users from over 100 agencies have self-service access to the database through a FinCEN portal, where they are able to query and analyze the data acquired through the BSA. Additionally, nine law-enforcement and supervisory agencies have access to bulk data, enabling them to merge FinCEN data with their own databases (FATF, 2016b, p. 55). Large Federal law enforcement agencies also have liaison staff within FinCEN, allowing for an open line of communication with the FIU.

FinCEN may link its BSA database to external sources, namely open source and commercial databases such as state corporation records, property records, court records, and vehicle registrations. Furthermore, the agency has indirect access to law enforcement data through partnerships with various agencies such as the Drug Enforcement Administration (DEA), the Federal Bureau of Investigation (FBI), and the United States Secret Service (USSS).

A recent emphasis on Information Technology (IT) modernization has led to SARs being available in the BSA data within one day of being received by FinCEN. Furthermore, FinCEN may request records from financial institutions on individuals or legal entities without requiring a court order or subpoena (USA PATRIOT Act of 2001).

8.6.2 Public-public data cooperation

One of the strengths of the United States AML regime, as stated by FATF (2016b), is the inter-agency cooperation and collaboration. This is accomplished through a variety of public-public initiatives, namely the SAR Review Teams, the Financial Crime Task Forces, and the Fusion Centers.

SAR Review Teams are found in all Federal districts in the country and leverage the task force environment to proactively review SARs received within that region (FATF, 2016b). These teams typically meet on a monthly basis and are comprised of representatives from various governmental agencies, such as the Internal Revenue Service – Criminal Investigation (IRS-CI), the Immigration and Custom Enforcement (ICE), the DEA, as well as the FBI (Clark, 2010). The goal of SAR Review Teams is to reduce the strain on the resources available to law enforcement and FinCEN by outsourcing the analysis of SARs within a geographical location to an independent team. This is possible due to FinCEN's ability to share the data collected under the BSA with other governmental agencies.

Financial Crime Task Forces combine local, State and Federal levels of government to cooperate and exchange information centered on a specific threat or geographical location. There are over 55 task forces across the country, led by IRS-CI, who leverage BSA data as well as IRS tax information to conduct investigations (FATF, 2016b). For example, the Organized Crime Drug Enforcement Task Force (OCDETF) is a nationwide task force aiming to coordinate investigations into serious drug trafficking threats and their financial infrastructure (FATF, 2016b). The El Dorado Task Force (EDTF) focuses on financial crime within the New York and New Jersey areas, regrouping 250 staff members, investigative teams of analysts, law enforcement agencies as well as prosecutors (FATF, 2016b).

Fusion Centers are information sharing and analysis hubs maintained by State or local governments. They are responsible for receiving, analysing, as well as sharing threat information and actionable intelligence. Fusion Centers leverage financial information, national intelligence as well as local and State information (FATF, 2016b). They develop extensive analytical products such as special event threat assessments and long-term trend analysis on specific homeland security threats (Homeland Security, 2017) while seeking to share information regarding threats and actionable intelligence (FATF, 2016b). For example, the OCDETF Fusion Center (OFC) is a data centre focused on analyzing drug and drug related financial data by merging information from FinCEN and its partner agencies, including the DEA, FBI, IRS, and ICE (FATF, 2016b).

8.6.3 Public-private data cooperation

Financial institutions, including banks, casinos, MSBs, and Brokers or Dealers in Securities, are encouraged to share information amongst themselves after sending notices to FinCEN, as per section 314(b) of the USA Patriot Act. They may share information on individuals and organizations through secure channels, with the aim of identifying and reporting money laundering or terrorist financing. This allows FinCEN to obtain a more comprehensive overview of suspicious transactions.

FinCEN Exchange is a voluntary program launched in 2017, with the objective of convening LEAs, FinCEN, and financial institutions to share contextual information on money laundering, terrorist financing, organized crime, and financial crime investigations. Section 314(b) of the USA PATRIOT Act enables financial institutions to assist in FinCEN Exchange briefings, when invited to participate. This initiative seeks to provide additional information to financial institutions on specific cases, thereby enabling them

to better fulfill their BSA reporting requirements. In turn, this increased reporting aids FinCEN and LEAs pursue their investigations (FinCEN, 2020).

8.6.4 Large scale analysis initiatives

Large-scale monitoring in the United States is mainly done through Financial Crime Task Forces and Fusion Centers. These initiatives seem to be mostly focused on an issue or geographical location. Additionally, law enforcement and various levels of government may access FinCEN data and merge this information to their own databases. As such, they may leverage this capability to perform scans relevant to their mandate.

8.6.5 Beneficial ownership registries

Similarly to Canada's AML regime, financial institutions in the US are required to verify the beneficial ownership of their customers, but there is no centralized registry of beneficial ownership.

8.6.6 Other initiatives

In recent years, Geographic Targeting Orders (GTOs) have been leveraged by FinCEN to gather data on corporate all-cash purchasers of luxury real estate in specified areas. In general, GTOs are a tool listed in United States law (Bank Secrecy Act of 1970) requiring financial institutions to report on specified transactions within a geographical location. The real estate GTOs were first introduced in January 2016, and legally mandated title insurance companies to report the beneficial owners of limited liability companies (LLCs) purchasing properties valued above a certain threshold without a bank loan. This order aimed to increase the transparency around LLCs who purchase luxury real estate without any external financing, thereby circumventing ML reporting regulations. The initial GTO was ordered for the counties of Manhattan and Miami-Dade and was subsequently amended to include additional high-risk counties in California, Florida, New York and Texas. The GTOs are valid for 180 days and have been consistently renewed.

A quantitative analysis of all-cash real estate purchases by LLCs before and after the GTO was conducted to evaluate the impact of the policy (Hundofte & Rantala, 2018). Before the GTO was issued, purchases of properties by corporations with no financing constituted approximately 10% of the dollar volume in the sample studied. After the policy change, all-cash purchases by LLCs fell by approximately 70%. Furthermore, the analysis showed a nationwide drop in corporate purchases into real estate and estimated an annualized drop in volume of $45 billion (Hundofte & Rantala, 2018). These results highlight the impact of policy changes that aim to decrease the opacity inherent to the real estate market.

8.7 NETHERLANDS AML DATA MODELS

The use of data in the Dutch AML regime is the most developed among the countries studied. The Dutch have created sophisticated automated data sharing and information exchange mechanisms which ensures that a plethora of analytical and enforcement groups have quick access to large volumes of data. Chief among these is iCOV - the Infobox for Criminal and Unaccountable Assets. iCOV aggregates public data sets and automatically services information requests on the data.

The Netherlands also have a series of cooperative groups which access these data to varying degrees for the purposes of knowledge creation and/or enforcement.

8.7.1 Financial transaction reporting

As in the other AML regimes, a number of reporting entities who interact with potential money launderers are required to submit reports to the national FIU, called Unusual Transaction Reports (UTRs) in the Netherlands. UTRs are only accessible to the FIU, but are made available to enforcement agencies if they are deemed to arise sufficient suspicion, thus turning them into Suspicious Transaction Reports (STRs). In 2018-2019, the FIU received almost 400,000 UTRs[32] and released nearly 58,000 as STRs organized in 8500 files (FIU-the Netherlands, 2019).

The conversion to an STR may occur through three main pathways:

1. **Analysis and conclusion of suspicion by the FIU**. This step can include the use of external data sources such as requests for additional information from the reporting entities and requests of related iCOV reports. At the time of the 2018 annual report, low computational capacity limited the complexity of analytical algorithms (some based on artificial intelligence) which could be leveraged in this step. This path accounted for around 30% of STRs in 2018 (FIU-the Netherlands, 2019).
2. **Automated matching on the Reference Index of Criminal Investigations and Subjects** – a database provided by the police to the FIU. This matching accounted for more than half (55%) of STRs in 2018 (FIU-the Netherlands, 2019).
3. **Requests from the National Public Prosecutor**. The prosecutor can forward requests for investigation from a large number of Dutch institutions to the FIU, upon which the FIU attempts to find matches in their UTR database. If these are found, the UTRs turn into STRs and are released.

Released STRs go into both the police data warehouse and the iCOV database.

8.7.2 Public-public data cooperation

The Dutch AML regime has several institutions whose purpose includes bringing disparate parties together to coordinate knowledge sharing and leverage differential data access for knowledge generation and enforcement actions.

The Financial Expertise Centre (FEC) is a public-public partnership between legal, tax, and law enforcement authorities aimed at sharing information and developing knowledge on financial crimes, with the goal of strengthening the integrity of the financial sector. Partners use this forum to request information from each other and complex information exchanges are facilitated through the use of "datarooms". The FEC was created in 1988 and thus precedes iCOV as a way to exchange information.

The Anti-Money Laundering Centre (AMLC) is housed within the Tax and Custom Administration's enforcement arm and touts its "unique access to data", namely tax information, as one of its strongest suits. The AMLC leverages this data access and its partnerships with academia to explore emerging and complex ways money is laundered – going beyond traditional methods like the use of cash intensive

[32] This number excludes over 350,000 UTRs which were received as a result of a recent change mandating a UTR filing for any transaction related to a list of high-risk countries. FIU-the Netherlands also makes this exclusion in most of their reporting as these UTRs are considered to carry very little information.

businesses - in order to better understand the phenomena and how it is changing to develop up-to-date signals and or indicators.

The Dutch have also established 10 Regional Information and Expertise Centres (RIECs), and one national centre (the LIEC) for information sharing and coordination at local levels. The RIECs generate integrated "images" which "shows what the manifestations of undermining crime are in a specific area, as perceived and/or suspected by RIEC partners." (RIEC-LIEC, 2019, p. 10, auto-translated). There were 127 local and regional "static" undermining images produced in 2018, implying each RIEC produces several. They are working on setting up "Joint Data Labs" between some of the RIECs, which would be used to develop "dynamic" images. It is not clear what data is used in either case, although localized knowledge and experience seems to be a factor.

At a more aggregate level, The Netherlands has implemented a novel and efficient way to handle information requests between different parties. A large share of such requests are completed automatically, enabling investigators to gain actionable intelligence quickly without large amounts of paperwork or delays.

The Infobox for Criminal and Unaccountable Assets (iCOV) aggregates data from suspicious transaction reports, tax administration data, tax fraud data, police data, central banks data, criminal convictions, real estate data, and customs data. They then prepare automated and custom reports for member organizations based on their access to data. Each report contains data from sources these organizations would otherwise have access to if they requested it formally from the data holder. iCOV serves to speed up and centralize the process.

iCOV produces three types of reports.

- **iRVI:** Capital and income reports, including all details on income, wealth, and revenue.
- **iRR:** Relations Reports, including all formal (financial) relations with whom a person has come into contact, and their contacts as well (up to 3 layers deep).
- **iRT:** Reports on thematic threats. More analytical pieces which involve complex queries across datasets. For example, iCOV created indicators on ownership of real estate, how real estate is financed and what real estate transactions flows look like.

One of the key organizing principles of iCOV is that the datasets it holds are kept separate. While the same query may be run across several datasets, these datasets are not combined to enable even more complex queries. In addition, partner organizations do not have direct access to the data; reports on individuals or clusters must be requested – so bulk analysis is mostly done by investigators at iCOV.

8.7.3 Public-private data cooperation

A centralized database similar to iCOV exists for privately held banking data. The Banking Data Referral Portal (BDRF) combines data from banks and payment service providers on account holders, ultimate beneficial owners, and account opening and closing dates. Account balances and transactions data are also being added. The Dutch investigative services and tax authorities employ this tool to aid in their investigations.

The Serious Crime Task Force (SCTF) is a pilot-project started in 2019 and housed within the FEC. It combines FEC members with private sector parties in order to facilitate information exchange and knowledge development. For example, a typology is developed either collaboratively or by a public-

sector partner, and are shared with the SCTF's private sector partners. These partners then test the typology on their own data and report back the results to SCTF. Eventually, the result can be instituted as standardized algorithms within these partner organizations, and suspected matches can be further investigated.

The SCTF is also used to exchange information on particular cases. Efforts are focused on professional money launderers and other key facilitators of ML, and there is a standing commitment that any cases submitted by private-sector partners will be followed up on by public-sector partners.

8.7.4 Large scale analysis initiatives

The Dutch AML regime incorporates widespread access to data from various regime partners. However, such data is mainly provided on case-by-case basis, limiting the extent of analyses done on aggregated data. The thematic reports produced by iCOV are perhaps the closest equivalent to a systematic data-driven approach to AML among partners studied, but the data employed is never gathered for analysis in a single database.

There are, however, some pilot projects aimed at better combining select data sets to enable the analysis of aggregated datasets. The Dutch FIU has an information sharing project with the Fiscal Information and Investigation Service (FIOD) which aims to "digitize the sharing of information between both organizations". This project may draw similarities to iCOV or the Banking Data Referral Portal, but it employs the FIOD's Financial Criminal Investigation network (FCI-net) application which can "match data sources locally and anonymously". They claim this can eventually be used to match with data from FIUs outside of Europe (FIU-the Netherlands, 2019, p. 10).

Similarly, the AMLC has built a set of tools named the AMLC suite, including the "AMCL Browser". This tool can be used to visually and interactively search through large amounts of data to find relationships between, for example, subjects, companies, and addresses. It also includes the "AMLC matcher" which finds relationships between different data set, allowing them to be combined, and a "Data Framework" which allows analysts to perform various analyses on the data, such as network analysis algorithms (AMLC, 2020).

8.7.5 Beneficial ownership registries

The European Union's 5th AML Directive has committed members to establish public registries of corporate beneficial ownership, which should include the name, the month and year of birth and the country of residence and nationality of the beneficial owner as well as the nature and extent of the beneficial interest. The Netherlands has committed to creating a beneficial ownership register, some parts of which will be public. Existing companies will have an 18-month grace period to update their registration so it will not be complete until 2021 at the earliest. iCOV also maintains records of financial linkages between companies, including known ownership relations.

9 DATA MANAGEMENT

Data is a valuable asset and, when properly managed, leads to countless analytical possibilities and insights. The following chapter aims to provide data management recommendations and guidelines, leading to databases suitable for statistical analysis. Each section tackles an aspect of data management that should be considered when establishing or enhancing a data framework. These are generic guidelines on data management based on practices by Canada's national statistics agency, Statistics Canada, a world-leading statistical agency, provided as base guidelines for the consideration of those managing and using data.

The sections of data management are:

1. Privacy, Necessity, and Proportionality;
2. Information Management;
3. Data Quality;
4. Creating Identifiers and Linking Data Tables; and
5. Data Editing, Validation and Imputation.

This chapter is based on resources publicly and internally available from Statistics Canada. It includes information from the agency's Data Quality Guidelines (Statistics Canada, 2019), Data Quality Assurance Framework (Statistics Canada, 2017), as well as Guidelines on the Management of Statistical Microdata Files, amongst others.

9.1 KEY POINTS

- The benefits gained from data collection endeavours should be proportional to the privacy impacts of a specified data strategy.
- Strong information management should be implemented to provide effective control and protection of the data, while ensuring accurate data interpretation.
- Data should be structured in a machine-readable, standardized format that is well suited to its intended and potential use.
- Data quality should be assessed to determine the data's ability to accurately fulfill its intended purpose.
- Data editing, validation, and imputation methods should be used to ensure that the data is of high quality and fit for analysis.

9.2 PRIVACY, NECESSITY, AND PROPORTIONALITY

Any collection of personal information impacts the privacy of those to whom it pertains. As such, the benefits gained from data collection endeavours must be proportional to the privacy impacts it entails. Proper data management practices can turn data into a valuable asset, but prior consideration must first be given to the necessity of the data collection and the proportionality of its impacts.

In this context, necessity refers to the specification of the data strategy and the data required to achieve an analytical outcome. Proportionality is the equilibrium at which the needs of the data strategy are fulfilled while ensuring that privacy is adequately protected. Excess data collection, a data strategy

which is unnecessary to achieve a desired outcome, or a data collection where the risks outweigh the benefits incurred by the analytical outcome, would not be proportional.

The Office of The Privacy Commissioner of Canada encourages government institutions to assess the following questions as part of a four-part test, aimed at determining the necessity and proportionality in collecting personal information (OPC, 2019; 2020):

- Is the measure demonstrably necessary to meet a specific need?
    - Is it rationally connected to a public goal that is pressing and substantial?
    - Is there empirical evidence in support of the initiative?
- Is it likely to be effective in meeting that need?
    - Was it carefully designed to achieve the objective in question?
- Is the loss of privacy proportional to the need?
    - The more severe the impact on privacy, the more clear and important the goal should be.
- Is there a less privacy-invasive way of achieving the same end?
    - Is there empirical evidence that other means will not achieve the objective? Have reasonable steps been taken to ensure that the minimum amount of personal information required to achieve the objective has been collected?

Statistics Canada has incorporated these questions into a Necessity and Proportionality framework, now applied agency-wide. This information is shared internally at Statistics Canada. The framework aids in reaching an equilibrium where the necessity of the data strategy is fulfilled and privacy is protected. As such, this framework can be used as a template by other institutions when collecting personal information, helping to ensure that privacy is being protected.

9.2.1 Necessity

Under the Necessity and Proportionality framework, the necessity of the data collection is examined at three levels. The first level clarifies what is being achieved through the data collection. The questions to be answered include which insights are to be generated, why they are required, and who they are for.

The second level clarifies what has been done previously and elsewhere to generate the same insights. What are the outcomes of those efforts compared to the expected outcome of the proposed new data collection? What are the limitations of these efforts that would be overcome by the collection?

The third level clarifies in detail the level of granularity, quality, and scope required to generate the insights highlighted at the first level. It clarifies the exact pieces of data required and by doing so, also clarifies which data would not be necessary.

9.2.2 Proportionality

The counterparts to the necessity of the data collection are the sensitivity of the data and the ethical issues raised by its collection, storage, and use. The sensitivity of the data relates to how personal it is. Statistics Canada is developing a sensitivity scale, and the work is still ongoing.

Assessing the ethical issues raised by the data collection requires transparency around the efforts and consultations with relevant stakeholders and ethics experts.

Proportionality is achieved when the data needs of the data strategy are fulfilled and privacy is protected. This equilibrium will often require flexibility on the necessity of the data collection, particularly regarding the third level specifying the granularity, quality, and scope of the data collected. Greater data sensitivity or efforts with greater ethical issues indicate that the intended level of data collection should be adjusted, where applicable. Where this is not possible, proportionality can be achieved by taking other approaches to mitigate the risks raised by the ethical considerations. For example, increasing security measures around the collection, storage, and processing of sensitive information can lead to a reduction in risks.

Achieving proportionality is an iterative process, with each adjustment to necessity or risk mitigation efforts triggering a reassessment of sensitivity and ethical issues.

9.2.3 Alternatives

As part of the framework, it is important to document alternatives to the proposed collection which have been dismissed and the reasons for their dismissal. Some may have been mentioned in level two of the necessity assessment, but other approaches may not have been covered. This process further increases clarity around the necessity of the proposed collection effort.

9.2.4 Testing and implementation

As part of the Necessity and Proportionality framework process, collection efforts having reached equilibrium should be tested at a more constrained scale to provide greater proof that the intended outcomes are achievable through the proposed data strategy.

9.3 Information Management

Information Management (IM) is an integral aspect of the organization of data, as it serves to document the data management cycle and inform data stewards on retention practices. It ensures adequate and effective control of the data acquired, while enabling analysts to deduce useful and accurate insights. Proper implementation of an information management framework also allows the data to retain value for potential future analyses. Inaccurate or incomplete information management can result in flawed analytical conclusions and vulnerability to staff churn.

In recent years, the federal government has pushed departments and agencies to adopt more digital data applications and organization, spearheaded by the Data Strategy Roadmap for the Federal Public Service (Government of Canada, 2018). This plan outlines the target outcomes as a user-centred, collaborative and digitally-centred government, enabled by themes such as a better integration between IT and IM as well as the development of relevant data strategies at the agency and department level.

The following section begins by noting IM Principles, abridged from Statistics Canada's Information Management Vision, relevant to the content of this report. The creation and organization of documentation and retention practices are discussed, which will permit data users to better understand, analyze and interpret the information available while adhering to privacy considerations. This information has been referenced from the Guidelines on the Management of Statistical Microdata Files and Aggregate Statistics, used internally at Statistics Canada.

9.3.1 IM principles

Information is a strategic asset, which should be easily accessible to the right person in a timely manner. Adhering to the IM Principles outlined below will lead to strong information management systems for organizations, enabling for the use of data at its full potential.

Of note, organizations should strive to have a strong governance of the information they hold, by setting up policies and ensuring that employees have the knowledge and tools to adhere to these plans. These policies should be implemented consistently throughout the organization, with clear directives for new systems or projects. Additionally, strong information management practices avoid duplication, which ensures that the information is secure and private while allowing all users to see modifications and updates. The organization should be aware of the information it holds, manage this information using a single source of authority, as well as monitor access to the information.

The current and future needs of data users should be considered during the planning and implementation stages of an information management system. This system should be continually updated and improved based on user feedback. The required information should be easily accessible to all internal and external users. Furthermore, access, use, sharing and collaboration of the information by the users should be maximized while adhering to information control to ensure that security and privacy is maintained.

Technological advances, such as Cloud-native and open solutions, should be leveraged to meet these information management principles.

Information Management Principles	
Governed	**Designed with users**
• Appropriate governance of information through adequate processes and tools • IM responsibilities are well-defined and communicated to staff, who are knowledgeable and can apply the procedures	• The needs of internal and external users are well-understood and considered during the development of IM policies and processes • There is consistent evaluation and improvements of policies and processes based on user feedback
Avoids duplication	**Discoverable**
• The same information is stored only once within an organization • Data and documents are not unnecessarily copied or duplicated	• The relevant information is easily accessible through the implementation of tools and resources
Interoperable	**Trusted**
• Implementation of organization-wide standards and processes, including well-defined IM requirements for new systems	• Knowledge of the information held by the organization, where it is, and who has access to it • Information is managed using a single source of authoritative information
Safeguarded	**Open and transparent**

• Data retention procedures follow operational needs, legislative policies, or historical needs, where appropriate • Information is properly managed to ensure its readability, reliability, and authenticity	• Users have open access to information • Access, use, sharing and collaboration is maximized, while respecting security, privacy and confidentiality
Secure and private	**Technology-enabled**
• Adequate information control and monitoring is executed to protect privacy and confidentiality • Only users with a "need to know" have access to information, and all rights to access are appropriately managed	• Information is generated and managed with modern digital tools to meet both user and IM requirements

9.3.2 Documentation

Proper and up-to-date documentation allows employees to efficiently identify and interpret data files, effectively manages information over time as well as supplements data quality by providing data users with information which will lead them to draw accurate conclusions.

In general, data documentation may be divided in four categories:

1) Internal management documentation;
2) General description of the data;
3) More specific information about the data elements, commonly referred as a data dictionary; and

User manual meant for external data users.

Organizations may select the categories relevant to their dataset or framework. The amount of documentation appropriate for an organization depends on the goal, usage and the length of time the data will be retained. Additional documentation may be necessary for large and complex databases, such as information about the data structure and framework.

The table below outlines examples of information that may be included within the documentation of each category.

Information that may be contained within the documentation categories	
Internal management	**General description of data**

• The intended use or purpose of the data • Whether the data has specific or widespread usage • The number of data users and the type of users (such as other departmental agencies) • Legal limitations to data sharing (such as disclosing personal information) • Whether the data is part of a repeated data collection or a one-time event • The length of time the information is to be retained	• The name of the program or activity • The reference period • The computer file format (e.g. Excel, SAS, CSV, Oracle, SQL) • The version identifier, if applicable • The retention period • The source data (including location and version number), if applicable • The description of the data file • The data location • The documentation location
Data dictionary	**User manual**
• The identification and description of every data element • Information about each variable (such as the variable name, the type, the position within the file, the label or description, the valid values and labels, the variable width) • The logic for derived variables • Important note for users, if applicable	• Information on the purpose, objectives, subject matter or content of the data • A description of the data population • The time frame or reference period of the data • Quality reports • Methodology information such editing and coding, imputation, and the creation of derived variables • Key concepts, variables (or characteristics) and classifications used • Important note for users, if applicable

Internal management documentation aims to inform the organization of their data holdings and keeps track of the external data users with whom the data has been shared. This ensures that data sharing privacy considerations are met. This information should be shared with internal data users and as such, it should be stored on an internal server. Therefore, it may also include information that is distinctively meant for internal work, as it will not be shared outside the organization.

Documentation containing a general description of the data allows the file to be read and interpreted as intended without requiring additional input from the data steward or the file's creator. It should include information about the format, context and content of the file. This document should be shared with all internal data users, and may be shared externally if it contains information not available within the user manual.

A data dictionary is the most important documentation when it relates to data. It informs users on the different variables available and their definitions, as variable names in databases are often not informative or may contain abbreviations. The data dictionary should also list the possible values and the different categories for categorical data. Data users typically repeatedly refer to the data dictionary, and as such, it should be widely available to internal and external users. It may also inform potential users as to whether the data would satisfy their needs.

Finally, a user manual should be created if the data is shared externally, or analyzed independently from the data collection and processing phase. It should include information on the purpose of the data, appropriate uses, as well as supplemental information on the quality of the data. Any methodological changes should be noted, especially those aiming to correct problems in the original data. This document should be widely shared to external data users, along with a data dictionary.

This information should be stored in a standardized format, easily accessible to authorized data users. The content should be updated as required, such that the information is up to date. This can be accomplished by storing the documentation on a shared drive and establishing consistent file naming conventions for quick reference. If the data requires restricted access, it should be placed in separate folders from the documentation, which has a more open access. Links should be provided to easily access both the data and the documentation.

9.3.3 Data structuring

It is important to maintain data in a machine-readable, standardized format that is well suited to its intended and potential use. In general, the data should be housed in a non-proprietary database structure with the ability to export and query the information. This ensures that data analysis and accessibility are not inhibited by the constraints of the interface provided by the particular software solution employed.

The following outlines the main differences between the Structure Query Language (SQL) and NoSQL data structuring systems. In general, SQL databases are older, more established and organized, while NoSQL structures provide additional flexibility, but may come with the caveat of an increased difficulty in analysis.

9.3.3.1 SQL

The most common data structuring solution is the Relational Database Management System (RDBMS). This is commonly referred to as a Structured Query Language (SQL) database, due to the Structured Query Language used to interact with the database. Examples include PostgreSQL and MySQL.

An SQL database organizes the data into one or more tables of columns and rows, with a unique identifier for each row. Typically, a table stores information from an entity or data dimension while each row represents an instance of that entity. For example, one table may cover properties with every row representing a unique property. The columns will store characteristics of that entity, such as assessment value or number of bedrooms. Multiple tables are related to each other via correspondence tables. A simple example is illustrated below, consisting tables of natural persons, properties, and their correspondence table.

Table 1: Natural Persons

ID	FirstName	LastName	SIN	
NP000000001	Jane	Smith	950999999	
NP000000002	John	Doe	850888888	

Table 2: Properties

ID	Street	Number	PostalCode	Value

P000000001	Wilboard Ave	22	H4Y 2P5	350000
P000000002	Pine Crescent	4	L3S 4K9	470000

Correspondence table: Natural persons and properties

ID	NaturalPersonID	PropertyID	OwnershipShare
OR00000001	NP000000001	P000000001	1
OR00000002	NP000000001	P000000002	0.5
OR00000003	NP000000002	P000000002	0.5

As depicted, the correspondence table serves to denote which natural persons, identified by their NaturalPersonID, own which property, or unique PropertyID. The corresponding is usually not one-to-one, substantiating the need for a reliable concordance table.

SQL databases are an older and more established technology relative to NoSQL databases. As such, expertise on structuring and implementation of the framework is widely available. The SQL language is employed to manage and interact with the data held in the RDBMS. This language has a relatively straightforward syntax structure and is a programming skill that is widely available. Furthermore, the highly structured organization of SQL databases is ideal for analysis. Statistical and other analytical projects are typically run on structured data tables. Additionally, the SQL language effortlessly allows the merging of various tables, thereby creating customized tables suitable for analysis.

The main drawback of SQL is that it is comparatively inflexible once implemented. As a result, it requires a thorough plan of potential future uses during the establishment process. The various tables and columns should be mapped out, as each row within a table has exactly the same number of columns. Whenever new information must be incorporated, the column structure and the database code may require significant modifications. For example, some property owners may be minors and information about their legal guardians must be now be included. That additional information may be recorded in additional columns within the Natural Persons table, thereby increasing the width of the table for all rows, or an additional correspondence table should be created.

9.3.3.2 NoSQL

NoSQL databases are much more unstructured as they do not possess a pre-defined tabular structure. Any structure can be implemented on a NoSQL system, making this format highly flexible and scalable while potentially introducing additional difficulties while querying and analysing the data. Additionally, noSQL has no widespread programming language such as SQL for RDBMS systems. The following section focuses on document-oriented NoSQL frameworks such as MongoDB and Apache CouchDB.[33]

[33] A document-oriented database is designed to retrieve one or more entire entries – such as a person's complete file – and can therefore be useful when this is the most common use case. By contrast, SQL databases excel at retrieving only specified columns from each row – limiting the amount of data transmitted when many rows are retrieved.

Other possible NoSQL structures include an otherwise unstructured key-value stores, a wide-column stores, and graph (network) stores.

In general, a document-based NoSQL database is composed of one or more collections, with each collection containing one or more entries (or documents) of similar types. Entries are composed of key-value pairs. For example, the information in the tables above can be presented in two collections.

Collection 1: Natural Persons

```
 [{
  "_id": "NP000000001",
  "FirstName": "Jane",
  "LastName": "Smith",
  "Properties": [{
    "id": "P000000001",
    "share": 1.0
  },{
    "id": "P000000002",
    "share": 0.5

  }]
},
{
  "_id": " NP000000002",
  "FirstName": "John",
  "LastName": "Doe",
  "Properties": [{
    "id": "P000000002",
    "share": 0.5
  }]
}]
```

Collection 2: Properties

```
 [{
  "_id": "P000000001",
  "Street": "Wilboard Ave",
  "Number": "22",
  "Postal Code": "H4Y 2P5",
  "Value": 350000,
  "Owners": [{
    "id": "NP000000001",
    "share": 1.0
  }]
},
{
  "_id": " P000000002",
  "Street": " Pine Crescent ",
  "Number": "4",
  "Postal Code": "L3S 4K9",
  "Value": 470000
  "Owners": [{
    "id": "NP000000001",
    "share": 0.5
  },{
    "id": "NP000000002",
    "share": 0.5
  }]
}]
```

In this example, the first entry in the Natural Persons collection, following the identification number, has a key for “FirstName”, with an associated value of “Jane”.

NoSQL also differs from SQL in the sense that entries, corresponding to rows in an SQL table, can have different structures, even when they are housed within the same collection. For example, a collection of property owners could contain both natural persons and legal persons. These two distinct entry types would have different keys within them – a corporation may have a key for its business number, or its legal type, whereas a natural person would contain a key for its SIN number or gender. The benefit of this framework is that all owners are stored in a single collection, without requiring additional columns for predominately empty or null rows. However, due to the inconsistent organization of the data, additional verification should be implemented on data retrieved from the database.

Another advantage of a NoSQL framework is that the data within an entry may be nested. Key-value pairs may contain entries comprised of text, numbers, timestamps, as well as a nested set of other key-value pairs, or lists containing any of the previous types. For example, the “Properties” key in the example above has a value containing a list of entries for each property. In turn, each of these entries have two key-value pairs; one for the identification number of the property and one for the property owner’s share of ownership. This use of nested structures and lists obviates the need for a separate correspondence table. However, it may also complicate the revising of information, as data is duplicated across collections. In the example above, a change of ownership would require a correction in both

collections, whereas in the SQL example, this change would entail a simple update of the correspondence table.

Due to the objects having no prescribed structure, it is relatively simple to add new information to the entities in a collection, as well as to entities of different types. As such, NoSQL databases may easily evolve as business requirements shift. Thus, this type of data storage is especially popular among start-ups. The relative lack of structure, however, leads to databases that may be more difficult to analyze. Analysts must first reshape the data into one or more tables containing the information they wish to analyze, with varying levels of difficulty depending on the structuring of the NoSQL framework. Since the data is semi-structured and contains a finite number of variations, the structure may still be conductive to data mining and analysis after it has been reshaped.

Finally, the NoSQL framework scales more efficiently than SQL, as NoSQL may be scaled across multiple computers or servers. Alternatively, SQL must usually be scaled within a single computer. Large technology companies often use the NoSQL structure due to this advantage.

All in all, many requirements should be considered when selecting a database system. The following table addresses the most relevant considerations and how they affect the choice between a relational (SQL) and non-relational (NoSQL) database system.

SQL vs NoSQL overview	
SQL: • Is more established • Is easier to analyze • Is less flexible • Is more difficult to set up and change	NoSQL: • Is less established • Is more difficult to analyze • Is more flexible • Is easier to set up and change

9.3.4 Data retention

In general, it is recommended that data be retained for as long as is allowed by prevailing legislation and privacy concerns. Data is often costly to collect and relatively inexpensive to store. While the use of stored data may not be immediately apparent, historical data can contain valuable strategic insights.

If possible, data which is not being actively used should be archived with more security and use precautions than is data which is current and being actively used, in order to lower the risk of data breaches. This is particularly true for data containing sensitive information or protected personal data.

9.4 DATA QUALITY

In general, data quality is an important metric that serves to assess the data's ability to accurately complete its intended purpose. High quality data will yield valid statistics and lead to correct interpretations of the data. As a result, quality must be considered at every step of the data collection, processing, analysis and, if applicable, dissemination stages. Additionally, data quality must be assessed and assured across source datasets to warrant their inclusion within a data framework. The quality of the data must also be disclosed to the data user to ensure adequate use of the data and correct inferences.

Statistics Canada publishes the Quality Assurance Framework (Statistics Canada, 2017) and the Statistics Canada Quality Guidelines (Statistics Canada, 2019) as tools for data quality management within the agency. Within this framework, Statistics Canada has established six dimensions of data quality. These data quality measures were created for the production of official statistics. However, they may be leveraged for the establishment and maintenance of federal and provincial databases. Data quality measures must be transparent to data users.

Statistics Canada's six data quality dimensions are

1. Relevance;
2. Accuracy;
3. Timeliness;
4. Accessibility;
5. Coherence and Comparability; and
6. Interpretability.

9.4.1 Relevance

Relevance consists of the degree to which the data produced satisfies the data user's needs. Data is relevant when it contains the information required by its users and allows them to extract that information in a straightforward manner. As such, it is important for the government entity to determine the data users and collect the data they require in a format suitable to analysis.

When considering the implementation of a data integration initiative, collecting data that is highly relevant for the goals of the framework will be a priority. It is important to consider what data is required for a strong framework, while considering the costs of collecting additional data, from a privacy, price, and resource standpoint. A cost to benefit analysis may be required, while respecting legislative constraints. For example, many indicators require data points relating to mortgage information which is scarcely available in a suitable format in the current governmental holdings. This information is highly relevant and as such, a cost to benefit analysis as well as a review of the legal considerations could lead to potential solutions for the acquisition of this data.

THE MAIN GUIDELINES FOR DATA RELEVANCE ARE:

- Create and maintain a list of current and potential users of the data
- State the goals and objectives for the use of the data by its users
- Determine what data should be collected and how it should be stored, depending on user needs
- Establish the optimal balance between the need for data and the privacy of individuals

9.4.2 Accuracy

Data accuracy is defined as the degree to which the data correctly measures the information it was designed to measure. It can be divided into two metrics; bias, representing the systematic errors in data and variance, the random error.

Bias is when the available data is not representative of the underlying population. For example, self-reported surveys can produce biased data, particularly in instances where respondent self-select to

participate (e.g. when the survey is not mandatory). Alternatively, bias also occurs when the data does not include variables that appropriately measure the outcome of interest or when changes to the data generating system or features are unaccounted for. On the other hand, variance includes errors occurring on an observation basis, or as random errors. As such, it will have a relatively smaller impact on analysis when compared to bias.

Associated with this data quality dimension is reliability, reflecting the data accuracy over time. This is especially important for repeated data collection, such as annual assessments or monthly measurements. Other important metrics include the validity of the variables within the database as well as the consistency of the data collected.

A focus on data accuracy is important during every stage of the data management process, in order to correctly extract information from the data. For example, allowing only closed-ended rather than open-ended responses on forms can be used to minimize data processing and ensure accurate and reliable data. Establishing procedures to detect errors in the stages listed will yield a high level of data accuracy.

THE MAIN GUIDELINES FOR DATA ACCURACY ARE:

- Use quality control methods and quality assurance processes to detect and control potential errors in the various phases of the production process
- Ensure that responses on forms are designed to allow accurate and reliable data collection
- Determine the quality of record linkage from different files
- Assess the representativeness of the data
- Minimize and evaluate processing errors, such as during data capture, edit and imputation, and file manipulation
- Use other sources to compare, evaluate and validate the data, where applicable
- Use new technologies to automate procedures as much as possible, thus minimizing errors from manipulation and data inconsistencies

9.4.3 Timeliness

Timeliness refers to the delay between the data collection period and the date where the information becomes available. Collecting and processing data in a timely fashion is often necessary to ensure data relevance. A parallel concept is punctuality, referring to the difference between the planned and actual availability of data.

In some cases, a compromise may be required between timeliness and the other data quality dimensions. Investing in innovation and resource management to streamline the data collection and processing stages may be necessary. These advances can bring sustained gains in timeliness and punctuality with limited impact on the other dimensions. Additionally, collecting data in a way that promotes minimal data processing may reduce errors and allow for prompt data availability.

THE MAIN GUIDELINES FOR DATA TIMELINESS AND PUNCTUALITY ARE:

- Plan and develop a schedule for the data collection and processing phase and ensure they are fulfilled with regular follow-ups
- Be transparent to data users and inform them about timeliness constraints
- Invest in resource management and new technologies to ensure timeliness, if applicable
- Assess the importance of data users' needs relative to the length of time and resources required to meet them

.

9.4.4 Accessibility

Data is accessible when users can obtain and use the required information with ease. This is especially important if the data users are not part of the data collection and processing phases, as an accessible dataset leads to lower chances of errors. This will ensure accurate data analysis and interpretation by the data users. This data dimension also refers to providing access to authorized users. The guidelines provided in the information management section should be leveraged to ensure high data accessibility.

Related to accessibility is clarity, referring to the inclusion of metadata and other information relating to the data produced. Once again, this is especially valuable for external data users, as it will inform them of the data they are analysing and of any processing peculiarities.

The principal standard of this dimension in the context of a data framework is the availability of data in an adequate structure and format. It is essential that data users receive data in a format that meets their needs, especially if this data has previously been processed. Structured data will also ensure that users can execute the current and potential future analyses required.

THE MAIN GUIDELINES FOR DATA ACCESSIBILITY ARE:

- Establish a system for documenting and archiving data
- Identify the data users and to determine their data format needs
- Create procedures for sharing restricted-use data and files and ensuring that these procedures are followed
- Make data accessible for current and future analysis
- Ensure that the data access and products comply with the organization's requirements
- Periodically update the metadata and note any processing adjustments
- Share the metadata with the data users and keep them informed of any collection or processing changes

9.4.5 Coherence and comparability

Coherence and comparability of data is defined as the ability to produce consistent statistical information that is comparable over time and can be used concurrently with other data sources. Coherent datasets follow established definitions and measurement protocols, ensuring their reliability

and their ability to be combined with other data sources. Comparability allows for the measurement of changes over time due to the change in true underlying values, as opposed to a change in definitions or measurement conventions.

As a result, it is recommended to use standard definitions and concepts throughout all datasets, as well as established and consistent data collection and processing methodologies. For data collection, it is advised to keep the fields consistent over time unless a change is warranted.

For data frameworks, coherence and comparability will ensure an accurate detection of indicators through data. Some indicators measure changes in data points over time, requiring strong data comparability. For example, determining a person's origins from their country of residence as opposed to their country of birth could lead to much different results. Keeping the method stable over time ensures comparability. Moreover, coherence of various data sources will guarantee the correct measurement of indicators while merging different sources.

THE MAIN GUIDELINES FOR DATA COHERENCE AND COMPARABILITY ARE:

- Keep concepts, definitions, classifications and methodologies up to date
- Ensure the fields during data collection are kept consistent, where applicable
- Document and regularly update the concepts, definitions and methodologies employed
- Keep variable names, definitions and processing steps as consistent as possible
- Specify when concepts and methods used at different times vary for the same variable, where applicable
- Assess and document the impact of using different methods at different times
- Evaluate how methodologies differ in other jurisdictions and how that may affect data comparability, where applicable

9.4.6 Interpretability

Data interpretability consists of the ease with which data users can understand, properly analyze and draw accurate conclusions from the data provided. This is achieved through the availability of supplementary information and metadata required for the appropriate use of the data. These additional statistical tools ensure that valid inferences are extracted from the data and that these interpretations are free from data user biases.

Metadata is defined as "information about statistical data and the statistical business process" (Statistics Canada, 2019). As such, providing metadata is an important factor in ensuring adequate data interpretability.

Statistics Canada has established three main components required by data users to establish meaningful data interpretability;

1. Concepts, variables and classifications

2. Methodology of data collection and processing
3. Data quality measures

Precise data interpretability is of great importance in a data framework, especially if this framework is housed outside of the data generating organization. Analysts must be kept informed of data definitions, methodologies and quality in order to accurately include relevant data in the framework. Additionally, any changes to these features must be disclosed to ensure the continued analysis of the data.

THE MAIN GUIDELINES FOR DATA INTERPRETABILITY ARE:

- Ensure that the organization's information management practices are up to date
- Provide metadata and information access to the data users
- Regularly update the data users on any changes to the three main components required for accurate data interpretability
- Ensure that quality indicators are available so that users are fully informed and can understand the strengths and limitations of the data

9.5 CREATING IDENTIFIERS AND LINKING DATA TABLES

Identifiers in data tables serve to uniquely distinguish observations within a data dimension. This leads to an organized database that allows the linking of information from different tables together. Identifiers may also replace identifying information, ensuring adequate privacy standards.

For example, a table containing data from a property dimension may list unique properties on every row. The characteristics may be recorded in the columns and contain information relevant to the property, such as land size, number of bedrooms and geographical location. A unique identifier for every property in this table may be created, with every identifier appearing only once for a given time period.

Another table may contain information on assessment values for properties. This table may display unique properties on every row, with their characteristics such as assessment value and assessment date as variables in the columns.

Linking both tables is necessary to gather more information, such as the median assessment value of properties within a certain number of bedrooms. This is done by determining which properties may be linked together from both datasets, while employing a linkage strategy. Challenges can arise when some values, such as street name and city, are missing, contain typological errors or produce links that are not one-to-one.

In general, generating identifiers and successfully linking them to related information in other tables is a complex task and must be taken into consideration when creating a database. The following section provides recommendations for the successful creation of identifiers and linking practices.

9.5.1 Creating identifiers

In general, identifiers are unique across a group of observations that require distinctive identification, as demonstrated in the example above. It is important to first establish the current objectives for the data and foresee any future potential uses. Identify which data dimensions are available within the data and which dimensions will require analysis. Create the identifiers according to the dimensions outlined, such as generating a property ID for every property or a citizen ID for individuals.

It may also be valuable to create identifiers within other identifiers, depending on the data structure and format. For example, one identifier may uniquely represent households while a second identifier distinctively characterises individuals within that household. These identifiers may be consolidated to create a unique identifier for every individual within these households, or they may be kept as two separate flags.

Identifiers may be numerical or contain characters. One way to generate these flags is to simply start at one and count. If using this technique, it is important to ensure that the data is not sorted in a meaningful way before producing the identifiers. For example, if the data is sorted by age before generating the identifiers, one may recognize that smaller IDs are younger, leading to privacy concerns. It is also possible to create identifiers while incorporating other information from the data. For example, an identifier may begin with “ON” or a province ID for individuals living in Ontario. This may be useful to quickly identify Ontario residents, while being broad enough to prevent the identification of an individual through their identifier.

It is recommended to keep the identifier constant across the same observation over time and across datasets, if possible. This greatly facilitates linkage between files.

All in all, identifiers are created to organize the data and uniquely identify observations without giving away any identifiable or protected information about the data itself.

THE MAIN GUIDELINES FOR CREATING IDENTIFIERS ARE:

- Determine the current and future uses of the data
- Establish the data dimensions available in the data and which dimensions will require analysis
- Generate a unique identifier for every distinct observation within a group and time period
- Ensure that the identifiers are uncorrelated to the underlying data
- Keep identifiers constant over time and over datasets, if possible

9.5.2 Linkage

Linking distinct data tables together may be necessary for institutions gathering large amounts of data from varying data dimensions, such as linking a data table with owner characteristics to their corresponding property. This process allows for the analysis of more information but can easily become a complex task as data from different datasets may contain errors, missing information and produce

erroneous links, among other issues. Statistics Canada has published a Record Linkage Project Process Model, which has been adapted for this section (Sanmartin, C. et al., 2017).

The first step consists of specifying the needs for linkage aiming to understand the concepts of interest and the data limitations that are present. Due to the complexity of a record linkage process, it is recommended to assess its necessity and its future use before undertaking the task. This may involve consultations with subject matter experts who are aware of the data gaps and analysis needs. Additionally, alternative existing data sets should also be explored to determine whether these meet the requirements for the project and any restrictions on the data must be considered. If existing data cannot satisfy the requirements outlined, record linkage may be a viable option.

The feasibility of record linkage must also be assessed. The data sets of interest must be identified as well as the unit of linkage, for example individuals, businesses or properties. The quality and availability of the linkage variables, that is the common variables across datasets to be linked, must be evaluated. These variables may be names, demographic variables such as gender and date of birth, geographic variables such as address and postal codes, and other unique identifiers such as social insurance and business numbers. Data custodians may be consulted if the data is housed externally. The feasibility of the linkage process based on the data and linking variables must be assessed by record linkage specialists.

Once record linkage is deemed appropriate, the next step involves the preparation of the data. The linkage variables must be identified. It may be necessary to subset only the selected variables to comply with privacy regulations. Additionally, linkage keys may be created by concatenating the values within the linkage variables. Once the variables are selected, their structure, format and code set must be standardized across datasets. The variables must be assessed for their accuracy, reliability and completeness. This involves analysing the rate of missing, incorrect or invalid values. More complex methods may also be employed to assess the discriminatory power of the linking variables. Finally, extensive data cleaning may be required to adequately prepare the data for record linkage.

With the data preparation completed, the linking process can begin. The linkage process is simple when different files have a common identifier or group of identifiers for the same observation. A link between the observations will be made if these variables agree. This is known as deterministic linking; it is essentially a rule-based linking strategy, following logical conditions. In general, deterministic linking is easier to implement and interpret, and comes with a lower computational burden. However, this strategy requires high quality data to generate accurate links (Dusetzina, S. et al., 2014).

Probabilistic linkage consists of assigning a weight to every linking variable in a record pair, while accounting for the discriminatory power of the variables. For example, SINs would have a higher discriminatory power than names, as names are more prone to being duplicated, changed, misspelled, or vary across datasets. Higher weights imply a higher level of agreement, with a total linkage weight calculated for each record pair. A record pair with the same date of birth, city and name would be given a higher weight than a record pair with the same date of birth and a similar name. This weight is then compared to thresholds to determine whether the record pair represents a match, a non-match or requires further analysis. The parameters to be estimated from the data includes the weights and the thresholds. This method, while more complex, is recommended when the data is of lower quality. Probabilistic methods have been shown to outperform deterministic linking in this case, and as such, warrants the additional investment in time and resources (Dusetzina, S. et al., 2014).

The final step consists of validating the linking process. The internal validation serves to assess the quality of the matched records. This may involve the analysis of the overall linkage rates, as well as comparing the linkage rate across sub-groups to ensure an absence of bias. It is also recommended to note if any trends can be discerned in unlinked records. Simulations or manual verification of a subset of records can also be undertaken to generate linking error estimates such as the rate of false positive and false negative links.

As per privacy regulations and subject to the intended use of the linked data, identifiable information may be removed from the dataset, replaced by unique identifiers for distinct observations. The linkage keys should be stored within an anonymized linkage key file which contains only the unique identifiers and linking variables employed in the process. Access to this file should be restricted to individuals involved in the linking process.

THE MAIN GUIDELINES FOR LINKING DATA TABLES ARE:

- Outline the needs and data gaps filled by the data linking process
- Ensure that existing data cannot satisfy the requirements needed by linking datasets
- Assess the feasibility of the data linking process
- Standardize the linking variables across datasets
- Select an adequate linking strategy based on the data quality and desired outcomes
- Validate the data linking process
- Comply with privacy regulations, where applicable

9.6 DATA EDITING, VALIDATION AND IMPUTATION

Data editing, validation and imputation ensures adequate data quality and as such, are an integral aspect of data management. Strong data editing and validation practices will identify data entry errors, and may even identify fraud or other inconsistencies within the information collected. High quality and accurate data is especially important in a potential AML framework where small changes or abnormalities in the data may indicate use of a ML scheme.

Data editing is a distinct process from data validation as it is intended to clean and verify the data at the micro-data level, while data validation involves a verification at the macro-data level. Finally, imputation consists of assigning values for missing or invalid data that cannot be properly edited. Its use is dependent on the user needs as imputation is not always applicable.

The following section has been adapted from Statistics Canada Quality Guidelines (Statistics Canada, 2019) as well as the Guidelines for the Validation of Statistical Outputs, available internally at Statistics Canada.

9.6.1 Data editing

Data editing ensures that accurate and consistent data is produced. The goal of this phase is to investigate data that is missing, invalid or inconsistent. It may also lead to the identification of erroneous data records or processes. These checks can be logical in nature, resulting from definitions and subject-

matter expertise, or may involve quantitative methods such as statistical outlier analysis. External sources may also be leveraged in this phase.

Data editing should be conducted on an ongoing basis, that is, during the data collection, migration, and processing, as well as analysis stages of the data management process. There should be additional emphasis on editing when methodological or processing changes are implemented. Additional editing efforts should be undertaken on data that is self-reported, such as verifying and substantiating entries on random subsets, as well as deterministic subsets containing higher-risk entities.

Simple checks include verifying that the data entry is in concordance with the original data, where applicable. Additionally, the identification of extreme or improbable values can be accomplished by a simple scan of summary statistics. The consistency of a variable within a given record may also be explored. Simple univariate methods, such as the one proposed by Hidiroglou and Berthelot (1986), or more advanced machine learning procedures, such as isolation forests (Liu, Ting, & Zhou, 2008), can be leveraged for outlier detection. The presence of many outliers signals potential errors within the data. An assessment of missing values may ensure that bias is not introduced in the data, especially for central variables. This involves analysing whether missing values have the same characteristics as other variables.

In general, editing aims to detect and correct true errors within the data without generating bias or succumbing to over-editing. A point may be reached where just as many errors are be introduced as are corrected, which is counter-productive. As such, it is valuable to adequately determine the natural conclusion of edits and to reapply checks to values that were edited to ensure that no additional errors have been introduced. It is also important to ensure that edits are made consistently across the data as to not be self-contradictory.

The final step is to produce documentation regarding the editing process and follow-up on failed edits. Studying the process can serve as a way to improve data collection and definitions, evaluate the quality of the data and provide information on potential future enhancements. Analysing the rate of edits, that is, how many edits were required, directly reflects the quality of the underlying data. These rates should be reported to the data users, especially for key variables. The magnitude of the changes due to edits in macro-level summary statistics should also be assessed and recorded.

THE MAIN GUIDELINES FOR DATA EDITING ARE:

- Verify that the data entry matches the original data after migration
- Authenticate data that is self-reported
- Identify extreme or improbable values through an outlier detection process
- Ensure that the edits implemented replace true errors in the data
- Emphasize the reduction of errors in earlier phases, such as data collection, rather than error correction
- Keep all edits internally consistent, that is, not self-contradictory
- Analyze and record the rate of edits required and the magnitude of changes due to edits

9.6.2 Data validation

Data validation takes place at the end of the processing phase. The goal is to ensure that the collection of data is of high quality before it is disseminated, analyzed and shared, as in the case of a data framework. Subject matter expertise and diagnostic tools are leveraged to challenge the statistics and estimates produced by the data.

This phase is distinct from data editing as it is a later process, examining aggregates at the macro-level as opposed to values within records at the micro-level. However, validation checks may lead to editing at the micro-level, and as such, these two processes may run in parallel or iteratively.

Totals, averages, or percentiles should be explored during this phase, and challenged by subject matter expertise. Correlations and relationships between variables should also be considered. These values should be evaluated overall as well as within important sub-groupings. An analysis of changes over time can be conducted for time-series data, as well as a verification of seasonally-adjusted estimates where applicable. Cross-tabulations can assist in validating relationships through univariate tables, or frequencies, bivariate or two-way tables, as well as multivariate tables. It is important that unexpected or inexplicable results lead to further investigation and appropriate editing.

Validation should also include a review of the production process, such as the collection, coding, and editing phases, as well other procedures such as imputation, identifier creation and linkage. Increased scrutiny is required when conceptual, operational or methodological changes are implemented, or when the analysis work is done independently from the production work.

THE MAIN GUIDELINES FOR DATA VALIDATION ARE:

- Conduct data validation at the end of the processing phase
- Challenge the statistics and estimates produced by the data
- Verify cumulative and aggregated information, and ensure that these values as reasonable and explicable
- Evaluate values overall and within important sub-groupings
- Investigate and edit, if applicable, unexpected and inexplicable results
- Review the production process, especially when changes are implemented

9.6.3 Data imputation

Data imputation consists of replacing missing, invalid or inconsistent data that cannot be properly edited. This phase should only be implemented when these missing or invalid values impact the production of estimates or statistical analyses, or introduce significant bias within the data. As such, the usefulness and limitations of imputation must be assessed and deemed appropriate before its implementation.

Once imputation is deemed appropriate, available auxiliary information is leveraged to approximate the missing or inconsistent values as accurately as possible. As per the Fellegi-Holt principle, the fewest number of values are imputed so as to ensure that all records successfully pass the editing process, that is, require no further editing (Fellegi & Holt, 1976). This ensures that the original data is preserved as much as possible. Furthermore, the frequency distribution of the data should be maintained where possible (Fellegi & Holt, 1976). Adequate implementation of the imputation process leads to reduced bias and variance.

Auxiliary variables should have strong relationships with the variables to be imputed, and can be selected with statistical methods, such as modelling techniques including regressions, or with subject matter expertise. Auxiliary variables may also be found in alternative databases, contingent on the quality and suitability of that data. Imputation methods, either deterministic or statistical, are then selected based on the type of missing data and the goal of imputation. An imputation model is adequately selected if the data's first moments (the mean and the variance) hold, and will be predictive if the auxiliary variables have strong relationships with the imputed variables. All in all, imputation constitutes a serious modelling effort so as to adequately select auxiliary variables and an appropriate strategy.

The identification of values that have been imputed should be visible to the data users through documentation detailing the variables impacted and the methods selected. Additionally, it is recommended to include indicator variables within the data to denote the values that have been imputed. Both the imputed and unimputed values should be retained and the impact of the imputation process should be assessed. Data validation should be done on the imputed data to ensure that the distributions and estimates are plausible.

THE MAIN GUIDELINES FOR DATA IMPUTATION ARE:

- Assess if imputation is suitable and required for the data
- Impute the fewest number of values possible while ensuring that all records pass the editing process
- Select auxiliary variables that have strong associations with the variables to be imputed
- Determine an adequate imputation method based on the type of missing data and on the goal of imputation
- Document the imputation process and inform data users
- Assess the impact of the imputation process on estimates and analyses

10 CONCLUDING REMARKS

This report has provided an overview of the ways in which enhanced data applications can be used to combat money laundering in the real estate markets of British Columbia with the view of providing an understanding of the opportunities and challenges arising from using data for those in the AML policy, regulation and enforcement communities in Canada. The report outlined a data framework related to money laundering schemes that could be used to detect ML, provided a baseline overview of the relevant data currently held by public institutions (both within and outside the AML regime) and assessed the quality and availability of that data.

The report's contribution remains largely in the realm of concept, theory and description with the goal of encouraging discussion among AML practitioners about the role of data for the purpose of increasing the effectiveness of Canada's AML regime. The framework itself has to be tested empirically. It remains to be seen how effective the data-based indicators identified will be at detecting money laundering in real estate. Some will likely be more useful than others.

The implementation of a framework for money laundering in real estate in BC and Canada may require additional legal and privacy reviews, including the application of a necessity and proportionality process to weigh the privacy impacts with the benefits to be realized. It remains to be seen how much data would ultimately be needed to operationalise the enhanced data framework envisioned however. It is possible that the data needs for the most effective indicators are in the public domain already and/or can be accessed through established mechanisms such that the data needed prove to be less invasive than those of the framework as a whole. An assessment of the prevailing legislation, particularly with respect to data sharing, will be an important consideration to determine the feasibility of implementing the full data framework. Separate assessments would likely be required for models related to the generation of strategic and tactical insights.

The assessment should also take into account the particular data sharing/access model adopted. The adoption of a hybrid model would necessitate reviewing how existing public institutions or new units could serve in the roles of data custodians and data co-ordinators. Other models would require similar decisions on where to most effectively house the data and how to provide access to relevant institutions taking privacy impact into account.

It should be noted that, even without implementing new data sharing and access models, regime partners could bring more data to bear on their AML activities. Augmenting an organization's existing data holdings with publicly accessible data could advance the production of actionable information in a timely and cost advantageous manner. It could also be a precursor to the subsequent adoption of more comprehensive initiatives to implement the AML data framework to further enhance the regime's effectiveness. Additional data sources beyond that of public institutions, such as data collected by the private sector and law societies, would be important also and ought to be considered.

Although this report focused on ML related to residential real estate in British Columbia, the consideration of ML typologies as applied to other high value target assets or activities such as automobiles, art, gambling and false invoicing, in different geographies, will also yield valuable insights.

In the end, experimentation and empirical verification will ultimately be necessary to determine the benefits, utility and implications from implementing the data collection and linking described by the

AML framework for real estate. ML is a dynamic business service used by criminals and criminal syndicates that seeks to corrupt Canadians, their institutions and economy.

The consequences of ML do not stop at the Canadian border. International movement of funds, to and from Canada, also deserve attention, in co-operation with international partners fighting ML in their jurisdictions. Consideration ought also to be given on the applicability of the data approaches presented to fighting tax evasion, the proceeds of which are also subject to money laundering. The value of deterring or terminating these activities may never be fully known, but the arrest and contradiction of these activities will contribute to the economic and social well-being, security and stability of Canada and our international partners.

11 REFERENCES

Abdallah, W, Koulouris, E, & Pereboom, B. (2020). *Market Integrity Index (MII): Methodology and Application to Quebec Housing Market*. Draft Report, September 2020.

AMLC. (2020). *Data & Analytics*. Anti Money Laundering Centre. https://www.amlc.nl/data-analytics.

AUSTRAC. (2019). *2018-19 AUSTRAC Annual Report*. Australian Transaction Reports Analysis Centre.

Bank Secrecy Act of 1970, 31 U.S.C § 5326(a).

BEIS. (2019). *Review of the Implementation of the PSC Register* (BEIS Research Paper Number 2019/005). London: Department for Business, Energy & Industrial Strategy.

Bonawitz, K., Eichner, H., Grieskamp, W., Huba,D., Ingerman, A., Ivanov, V., Kiddon, C., Konecny, J., Mazzocchi, S., McMahan, H., Overveldt, T., Petrou, D., Ramage, D., & Roselander, J. (2019). Towards Federated Learning at Scale: System Design. *Proceedings of the 2nd SysML Conference,* Palo Alto, CA, USA.

Bromwich, R. (2018). (Where is) the Tipping Point for Governmental Regulation of Canadian Lawyers? Perhaps it is in Paradise: Critically Assessing Regulation of Lawyer Involvement with Money Laundering after Canada (Attorney General) v Federation of Law Societies of Canada. *Manitoba Law Journal, 41*(4), 1-26.

Canada v FLSC. (2015). *Canada (AG) v Federation of Law Societies of Canada*, 2015 SCC 7, [2015] 1 SCR 401

CanDIG. (2020). *National, distributed, analysis of locally-controlled private genomic data*. www.distributedgenomics.ca

Chadderton, P., & S. Norton. (2019). *Public-Private Partnerships to Disrupt Financial Crime: An Exploratory Study of Australia's Fintel Alliance* (SWIFT Institute Working Paper Series, no. 2017-003 (May)). doi:10.2139/ssrn.3392268.

Canadian Charter of Rights and Freedoms, Part 1 of the *Constitutional Act*, 1982, being Schedule B to the *Canada Act 1982* (UK), 1982, c 11.

Clark, M. (2010). *SAR Review Program – A National Perspective* [presentation]. https://www.ice.gov/doclib/aml/pdf/2010/clark.pdf

CMHC. (2020). *Mortgage Data Standards. Technology, business and policy joining forces to advance the industry.* Canadian Mortgage and Housing Corporation.

Cribb, R., & Oved, M. C. (2017, January 25). Hiding Cash in Canada; A Global Crackdown on Offshore Tax Havens Leads Cheats and Criminals to a Surprising New Mecca for Money Laundering and Tax Evasion. *The Toronto Star*.

Crime and Courts Act. (2013). c. 22 § 7. http://www.legislation.gov.uk/ukpga/2013/22/section/7.

Criminal Code, RSC. (1985). c C-46.

Dusetzina, S., Tyree, S., Meyey, A., Green, L. & Carpenter, W. (2014). *Linking Data for Health Services Research: A Framework and Instructional Guide.* United States: Agency for Healthcare Research and Quality.

van Duyne, P. C. (2013). Crime-Money and Financial Conduct. In B. Unger & D. van der Linde (Eds.), *Research Handbook on Money Laundering* (pp. 232-52). Edward Elgar Publishing.

Dwork, C. & Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science, 9*(3–4), 211–407.

FATF. (2007). *Money Laundering & Terrorist Financing through the Real Estate Sector*. Paris: Financial Action Task Force.

FATF. (2016a). *Anti-Money Laundering and Counter-Terrorist Financing Measures - Canada* (Fourth Round Mutual Evaluation Report). Paris: Financial Action Task Force.

FATF. (2016b). *Anti-Money Laundering and Counter-Terrorist Financing Measures – United States* (Mutual Evaluation Report). Paris: Financial Action Task Force.

FATF. (2018). *Anti-money laundering and counter-terrorist financing measures – United Kingdom* (Fourth Round Mutual Evaluation Report). Financial Action Task Force.

FATF. (2020). *Anti-Money Laundering and Counter-Terrorist Financing Measures – United States* (3rd Enhanced Follow-up Report & Technical Compliance Re-Rating). Paris: Financial Action Task Force.

Fellegi, I. & Holt, D. (1976). A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association, 71.*

FinCEN. (2001). *USA Patriot Act.* Washington: Financial Crime Enforcement Network, U.S. Treasury. https://www.fincen.gov/resources/statutes-regulations/usa-patriot-act

FinCEN. (2020). *FinCEN Exchange Questions and Answers.* Washington: Financial Crime Enforcement Network, U.S. Treasury. https://www.fincen.gov/resources/fin-exchange/fincen-exchange-frequently-asked-questions

Fintel Alliance. (2019). *Annual Report 2018-19*. Fintel Alliance.

FINTRAC. (2016). *Operational Brief: Indicators of Money Laundering in Financial Transactions Related to Real Estate* (FINTRAC-2016-OB001). Ottawa: Financial Transactions and Reports Analysis Centre of Canada.

FINTRAC. (2018). *Annual Report 2017-18*. Ottawa: Financial Transactions and Reports Analysis Centre of Canada.

FINTRAC. (2019). *Annual Report 2018-19*. Ottawa: Financial Transactions and Reports Analysis Centre of Canada.

FIU-the Netherlands. (2019). *FIU-the Netherlands Annual report 2018.* FIU-the Netherlands.

Gentry, C. (2009). *A Fully Homomorphic Encryption Scheme*. PhD thesis, Stanford University, http://crypto.stanford.edu/craig

German, P. (2019). *Dirty Money – Part 2. Turning the Tide - An Independent Review of Money Laundering in B.C. Real Estate, Luxury Vehicle Sales & Horse Racing*. Vancouver: Peter German & Associates Inc.

Global Witness. (2018). *The Companies we Keep. What the UK's Open Data Register Actually Tells us About Company Ownership.* London: Global Witness.

Government of Canada. (2018). *Report to the Clerk of the Privy Council: A Data Strategy Roadmap for the Federal Public Service.* https://www.canada.ca/en/privy-council/corporate/clerk/publications/data-strategy.html

Hidiroglou, M. & Berthelot, J.-M. (1986). Statistical editing and imputation for periodic business surveys. *Survey Methodology, 12,* 73-83.

Homeland Security. (2017). *Advancing the Homeland Security Information Sharing Environment: A Review of the National Network of Fusion Centers*. House Homeland Security Committee. Majority Staff Report.

Hundofte, S., and Rantala, V. (2018). Anonymous Capital Flows and U.S. Housing Markets (University of Miami Business School Research Paper No. 18-3). doi:10.2139/ssrn.3186634.

IMF. (2011). *Anti-Money Laundering/Combating the Financing of Terrorism (AML/CFT) – Topics.* Washington, D.C.: International Monetary Fund. https://www.imf.org/external/np/leg/amlcft/eng/aml1.htm.

Irwin, A. S. M., Choo, K.-k. R., and Liu, L. (2011). An Analysis of Money Laundering and Terrorism Financing Typologies. *Journal of Money Laundering Control, 15*(1), 85–111. doi:10.1108/13685201211194745.

ISED. (2020). *Strengthening Corporate Beneficial Ownership Transparency in Canada.* Inovation, Science and Economic Development Canada.

Keen, M., & Smith, S. (2006). VAT Fraud and Evasion: What Do We Know and What Can Be Done? *National Tax Journal*, *59*(4), 861–87.

van Koningsveld, J. (2013). Money Laundering – 'You don't see it, until you understand it': rethinking the stages of the money laundering process to make enforcement more effective. In B. Unger & D. van der Linde (Eds.), *Research Handbook on Money Laundering* (pp. 435-51). Edward Elgar Publishing.

Langdale, J. (2017, November). Impact of Chinese Transnational Crime on Australia: Intelligence Perspectives. Presentation delivered at the *New South Wales Police Force iFocus Conference*.

Lauter, K., Naehrig, M., & Vaikuntanathan, V. (2011). Can Homomorphic Encryption be Practical? *Proceedings of the 3rd ACM Cloud Computing Security Workshop (CCSW'11),* 113-124.

Levi, M. (2013). *Money Laundering Typologies: a Review of their Fitness for Purpose.* Government of Canada.

Liu, F., Ting, K. & Zhou, Z. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining,* 413-422. doi:10.1109/ICDM.2008.17.

Maloney, M., Somerville, T., & Unger, B. (2019). *Combatting Money Laundering in BC Real Estate*. Victoria: Expert Panel on Money Laundering in BC Real Estate.

Maxwell, N. (2020). Innovation and discussion paper: Case studies of the use of privacy preserving analysis to tackle financial crime. *Future of Financial Intelligence Sharing (FFIS) research programme.* Version 1.0 (29 June 2020).

NCA. (2019a). Deputy Director – National Data Exploitation Capability (NDEC). Job Posting by *Gatenby Sanderson*. https://www.gatenbysanderson.com/job/GSe54089/deputy-director-national-data-exploitation-capability-ndec/.

NCA. (2019b). *NCA UK Annual Report 2018-19* (HC 2379). National Crime Agency.

NCA. (2019c). *SARs In Action* (Issue 1). National Crime Agency.

NCA. (2019d). *SARs In Action* (Issue 2). National Crime Agency.

NCA. (2019e). *SARs In Action* (Issue 3). National Crime Agency.

NCA. (2019f). *SARs Regime Good Practice. Frequently Asked Questions. Defence Against Money Laundering (DAML)*. National Crime Agency.

Nielsen. (2015). *Exploring the Privacy Concerns and Priorities of Canadians.* Report prepared for The Office of the Pricacy Commissioner of Canada. The Nielsen Company.

OECD & The World Bank. (2018). *Improving Co-operation between Tax Authorities and Anti-Corruption Authorities in Combating Tax Crime and Corruption.* Organization for Economic Cooperation and Development & The World Bank.

OPC. (2017). *Privacy Enhancing Technologies – A Review of Tools and Techniques.* Office of the Privacy Commissioner of Canada.

OPC. (2019). *Statistics Canada: Invasive data initiatives should be redesigned with privacy in mind. Complaint under the Privacy Act*. Office of the Privacy Commissioner of Canada.

OPC. (2020*). Expectations: OPC's Guide to the Privacy Impact Assessment Process (Revised: March 2020)*. Office of the Privacy Commissioner of Canada.

PCMLTFA. (2000). *Proceeds of Crime (Money Laundering) and Terrorist Financing Act,* S.C. 2000, c 17.

RIEC-LIEC. (2019). *RIEC-LIEC Jaarsverslag 2018.* RIEC-LIEC.

Sanmartin, C., Trudeau, R., Trainor, C., Dasylva, A., Dosman, D., Evra, R., Grimes, C., McCarthy, I., McLeish, S., Nadeau, C., Qiu, T., Rollin, A., Rosa, J. & Saïdi, A. (2017). *Record Linkage Process Project Model*. Ottawa: Statistics Canada. Catalogue no. 12-605-X.

Schneider, S. (2004). Organized Crime, Money Laundering, and the Real Estate Market in Canada. *Journal of Property Research, 21*(2), 99–118. doi:10.1080/0959991042000328801.

Statistics Act, RSC. (1985). c. S-19.

Statistics Canada. (2017). *Statistics Canada's Quality Assurance Framework* (Third Edition). Ottawa: Statistics Canada. Catalogue no. 12-586-X

Statistics Canada (2019). *Statistics Canada Quality Guidelines* (Sixth Edition). Ottawa: Statistics Canada. Catalogue no. 12-539-X

Strizl, D. (2019). *Privacy Preserving Matching Using Bloom Filters: an Analysis and an Encrypted Variant.* Master's Thesis. University of Twente.

Tambay, J.-L., I. Carrillo-Garcia and S. Kanagarajah. (2015). *A New Approach for the Development of a Public Use Microdata File for Canada's 2011 National Household Survey*. Ottawa: Statistics Canada. Catalogue no. 99-137-XWE2015001.

Tax Justice Network (2018). *The Financial Secrecy Index*. Buckinghamshire: Tax Justice Network. https://financialsecrecyindex.com.

The Royal Society. (2019). *Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis.* The Royal Society.

Tomlinson, K. (2015, September 8). Banks Helping Wealthy Asian Investors Bend Chinese Rules to Bring Money into Canada. *The Globe and Mail*.

Transparency International Canada (2019). *Opacity: Why Criminals Love Canadian Real Estate (And How to Fix It).* Toronto: Transparency International Canada.

Unger, B. (2009). Money Laundering - A Newly Emerging Topic on the International Agenda. *International Review of Law and Economics,* 5(2), 807-19. doi:10.2202/1555-5879.1417.

Unger, B., Ferwerda, J., Nelen, H., and Ritzen, L. (2011). *Money Laundering in the Real Estate Sector: Suspicious Properties*. Cheltenham, UK: Edward Elgar.

USA PATRIOT Act of 2001, Pub. L. No. 107-56 § 314, 115 Stat. 272 (2001).

WEF. (2019). *The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value*. World Economic Forum.

12 APPENDIX 1: DATA POINTS AND POTENTIAL SOURCES

The data conceptual model highlights the need for data on buyers, sellers, the properties themselves, and the details on the financing of the property or property transaction. The table below organizes the required data points for all of the indicators into 8 categories. Buyers who are natural persons require different data points than do buyers who are corporate entities. Similarly, the details on the financing of a property can be divided into details on the characteristics of the loan, the payments of the loan, the provider of the deposit and/or down payment, etc. These are here structured into separate tables according to the type of data point, rather than the underlying category of the real estate arrangement they are measuring.

Subject	#	Data Point	Potential Source(s)
Natural persons	01	Identifier	-
	02	Name	BC TACS, BC LTSA, CHSP
Prefixes:	03	Address	BC TACS, BC LTSA, T1, CHSP
NPO – owners/ buyers/ sellers	04	Date of birth	BC TACS, T1, CHSP
NPL – lenders	05	Country of birth	CHSP
NPS – spouses of owners	06	Properties owned, current, identifiers	CHSP, BC LTSA, BC LOTR, BC TACS
NPR – relatives of owners	07	Properties owned, historic, identifiers	BC LTSA, BC TACS
	08	Tax filings / activity in Canada	T1
	09	Income, time of purchase	T1, CHSP, BC TACS
	10	Income, historic	T1, BC TACS
	11	Wealth	
	12	Legal guardians, identifiers	
	13	State of health (i.e. incapacitated / healthy)	T1
	14	Institutionalized, flag	
	15	Mortgages held, identifiers	BC LTSA
	16	Date of immigration to Canada	CHSP, BC TACS
	17	Immigration class	CHSP
	18	Financial transactions, identifiers	FINTRAC
	19	Person is party to suspicious transactions, flag	FINTRAC
	20	Criminal record, flag	RCMP, ICCS
	21	Has forfeited a real estate property in the past, flag	BC CFO
	22	Has been prosecuted for money laundering, flag	ICCS
	23	Spouse, identifiers	T1, BC TACS
	24	Relatives, identifiers	
	25	Income from rent	T1
	26	Is a trustee	BC TACS
Legal persons / arrangements	01	Identifier	-
	02	Name	BC TACS, BC LTSA, T2, BR
	03	Address	BC TACS, BC LTSA, T2, BR
Prefixes:	04	Date of registration	BC TACS, BR
LPO – owners / buyers / sellers	05	Country of registration	BC TACS, BR, T2
LPL - lenders	06	Industry	BC TACS, BR, T2
	07	Immediate owners, identifiers	BC TACS, BR
	08	Beneficial owners, identifiers	BC TACS, BC LOTR
	09	Ownership structure	BC LOTR, BR, T2

	10	Tax filings / activity in Canada	BR, T2
	11	Number of employees	BR, T2
	12	Phone number responsiveness	Manual Check
	13	Web presence, flag	Manual Check
	14	Logo, flag	Manual Check
	15	Financial transactions, identifiers	FINTRAC
	16	Entity is party to suspicious transactions, flag	FINTRAC
	17	Revenue	T2, T3, BR
	18	Revenue from rent	BC TACS, T2
	19	Properties owned, current, identifiers	BC LTSA, CHSP, BC TACS
	20	Properties owned, historic, identifiers	BC LTSA, BC TACS
	21	Mortgages held, identifiers	BC LTSA
	22	Lender regulatory regime	Manual Check
	23	Country of control	T2, BR
	24	Related and associated corporations, identifiers	BR
Properties	01	Identifier	-
Prefix: P	02	Address	BC TACS, BC LTSA, BCA, CHSP
	03	Type	BC TACS, BCA, CHSP
	04	Owners, identifiers	BC TACS, BC LTSA , BC LOTR, CHSP
	05	Sales, historic, identifiers	BC TACS, BCA, BC LTSA
	06	Occupied by owner, flag	BC TACS, CHSP
	07	Vacant, flag	BC TACS
	08	Number of units	BCA
	09	Number of vacant units	BC TACS
	10	Rent charged in units	
	11	Construction date	BCA
	12	Renovations, extent	BC TACS
	13	Renovations, date	
	14	Renovations, medium of payment	
	15	Mortgages on property, current, identifiers	BC LTSA
	16	Mortgages on property, historic, identifiers	BC LTSA
Loans / mortgages	01	Identifier	-
	02	Property mortgaged, identifier	BC LTSA
	03	Holders, identifiers	BC LTSA
Prefix: M	04	Previous holders, identifiers	
	05	Lenders, identifiers	BC TACS, BC LTSA
	06	Size	BC TACS
	07	Interest rate	
	08	Term length	
	09	Opening date (first payment date)	BC LTSA
	10	Closing date (last payment date)	BC LTSA
	11	Date of transfer	
	12	Current balance	
	13	Payments (adjusted to reflect monthly payments), identifiers	

	14	First date of being 90 days in arrears	
	15	Collateral type	
	16	Collateral value	
Loan payments *Prefix: MP*	01	Identifier	-
	02	Loan, identifier	
	03	Payment date	
	04	Payment amount	
	05	Payment medium	
	06	Transaction, identifier	FINTRAC
Real estate transactions *(incl. incomplete transactions)* *Prefixes:* *PT – property transactions* *PTU – unrealized property transactions*	01	Identifier	
	02	Property, identifier	BC TACS, BC LTSA
	03	Buyers, identifiers	BC TACS, BC LTSA
	04	Sellers, identifiers	BC TACS, BC LTSA
	05	Date of transaction being finalized	BC TACS, BC LTSA, BCA
	06	Property assessed or expected value, time of sale	BCA, CHSP
	07	Sales price	BC TACS, BCA, BC LTSA
	08	Name of buyer when first contacting broker	
	09	Name of buyer at time of sale	
	10	Date of first contracting broker	
	11	Date of name change	
	12	Date of initial deposit	
	13	Medium of initial deposit	
	14	Provider of Initial deposit, identifier	
	15	Medium of transaction documents (i.e. in-person, mail, fax)	
	16	Buyers insistence on medium of transaction documents	
	17	Involved mortgage and real estate brokers/agents, identifiers	BC FSA, RECBC
	18	Involved notaries, identifiers	BC LTSA
	19	Amount of down payment	BC TACS
	20	Medium of down payment	
	21	Date of transaction being cancelled	
	22	Sale is an assignment (pre-construction)	BC TACS
	23	Sale is an assignment (existing property)	
Financial transactions *Prefix: FT*	01	Identifier	-
	02	Date	FINTRAC, CRA
	03	Amount	FINTRAC, CRA
	04	Sender, identifier	FINTRAC, CRA
	05	Recipient, identifier	FINTRAC, CRA
	06	Country of origin	FINTRAC, CRA
	07	Destination country	FINTRAC, CRA
Miscellaneous *Prefix: X*	01	List of regulated mortgage issuers in Canada	Manual Check
	02	List of PEPs	
	03	Ranking of countries by capital controls	IMF
	04	Ranking of countries by corruption level	Transparency International
	05	Ranking of countries by secrecy laws / strength of AML regime	FATF, Tax Justice Network
	06	Rent charged in comparable units	

07	Interest rates on comparable mortgages	BC LTSA
08	Average length of time to complete a mortgage negotiation/transaction	
09	Property values in area	CHSP, BCA
10	Crime rate in area	UCR
11	Addresses of lawyer's offices	Manual Check
12	Addresses of corporate service providers	Manual Check

13 APPENDIX 2: COMPLETE MAPPING OF THE DATA FRAMEWORK

The Data Points in the mapping below employ prefixes and numbers corresponding to entries in the Data Points table in Appendix 1.

Scheme	*Sub-Scheme*	*Indicator*	*Data Points*
Unlawfully obscure beneficial ownership	By using a shell company or trust	- Owner is a company or trust	LPO_02, P_04, P_03
		- Owner is a numbered company	LPO_02
		- Owner company is incorporated abroad	LPO_05
		- Owner company is foreign-controlled	LPO_23
		- Owner company has foreign address	LPO_03
		- Owner company has public address (PO box or lawyer's office)	LPO_03, X_11, X_12
		- Owner company is in a state with strong secrecy laws	LPO_03, LPO_05, X_05
		- Owner company has no tax activity	LPO_10
		- Owner company has no employees	LPO_11
		- Owner company has no address	LPO_03
		- Owner company has no web presence / logo/ phone number/etc.	LPO_02, LPO_12, LPO_13, LPO_14
		- Owner company has complex ownership structure	LPO_09
		- Owner company was just established	LPO_04, P_05, PT_05
		- Owner company purchases residential property but its industry is not real estate	LPO_06, P_03
		- Owner company has few resources relative to property value	LPO_17, P_05, PT_06, PT_07
		- Owner company is a foundation, cultural or leisure association or non-profit entity	LPO_06, P_03
	By using a formal or informal nominee owner	- Owner and spouse have insufficient income/wealth for property	P_05, PT_05, PT_06, PT_07, NPO_01, NPO_09, NPO_10, NPO_11, NPO_23, NPS_06, NPS_09, NPS_10, NPS_11, P_15, M_06, M_17
		- Owner is related to other owners with insufficient income/wealth for their properties	NPO_24, NPR_06, P_05, PT_07, NPR_09, NPR_10, NPR_11
		- Owner is related to PEP	NPO_24, X_02
		- Owner is a minor	NPO_04, P_05, PT_05
		- Owner is incapacitated	NPO_13
		- Owner is institutionalized	NPO_14
		- Owner has a legal guardian	NPO_12
		- Buying client changes purchasing party's name shortly before closing the purchase	P_05, PT_08, PT_09, PT_10, PT_11
		- Buyer pays initial deposit with cheque from unrelated third party	PT_13, PT_14, NPO_23, NPO_24
		- Deposit is made through a lawyer's or notary's trust account	PT_13, PT_14
		- Buyer insists on providing signatures via fax only	PT_15, PT_16

		- Owner is a trustee	NPO_26
	By having an unclear address	- Owner has a PO box as their listed address	NPO_03
		- Owner has a lawyer or notary as their listed address	NPO_03, X_11
		- Owner has a Corporate Service Provider as their listed address	NPO_03, X_12
Funnel cash/money through mortgages	By acquiring many mortgages	- Owner holds multiple mortgages	NPO_06, NPO_15
		- Owner has insufficient income to support said mortgages	NPO_06, NPO_15, M_06, M_07, M_08, P_05, PT_05, PT_07, NPO_09, NPO_10, NPO_11, NPO_23, NPS_09, NPS_10, NPS_11
		- Owner acquired multiple properties around the same time	NPO_06, P_05, PT_05
		- Owner is linked to others who acquired multiple mortgages around the same time	NPO_06, NPO_15, NPO_24, NPR_06, NPR_09, NPR_10, NPR_11, P_05, PT_05, PT_07
	By repeatedly mortgaging a property	- Property is repeatedly mortgaged to the same owner	P_15, P_16, M_03, P_05, PT_05, LPO_07, LPO_08
		- Mortgage is paid back quickly	P_15, P_16, M_06, M_09, M_10, M_12
	By acquiring an outsized mortgage	- Mortgage has unusual loan-to-value	P_05, PT_06, PT_07, P_15, M_06
		- Mortgage has unusual interest rate	M_07, X_07
		- Mortgage is with an unregulated lender	LPL_02, LPL_03, LPL_06, LPL_22, NPL_02, NPL_03, X_01
	By lending the money to oneself	- Lender company is owned by borrower	LPL_07, LPL_08
		- Lender and borrower are same person	NPO_02, NPL_02
		- Lender company is incorporated abroad	LPL_03, LPL_05
		- Lender company is foreign-controlled	LPL_23
		- Lender company is in a state with strong secrecy laws	LPL_02, LPL_03, LPL_05, X_05
		- Lender company has no tax activity	LPL_10
		- Lender company has no employees	LPL_11
		- Lender company has no address	LPL_03
		- Lender company has no web presence / logo/ phone number/etc.	LPL_02, LPL_12, LPL_13, LPL_14
		- Lender company has complex ownership structure	LPL_09
		- Lender company was just established	M_09, LPL_04
		- Lender company is in non-financial sector	LPL_06
		- Lender is a natural person	NPL_02
		- Loan is syndicated between multiple unregulated companies or individuals	M_05, LPL_02, LPL_03, LPL_05, LPL_22, NPL_01, NPL_02, NPL_03
		- Loan has no collateral	M_15, M_16

		- Loan has collateral unrelated to the property	M_15, M_16
		- Loan has very low value collateral	M_15, M_16
	By acquiring very valuable property with a large mortgage	- Rental or business property remains unused	P_03, P_05, PT_07, P_08, P_09
		- Rent in rental property is under market value	P_03, P_05, PT_07, P_08, P_10, X_06
	By taking over existing mortgages	- Mortgage is transferred to non-owner	P_15, P_16, M_03, M_04, M_11
		- Seller retains existing mortgage	P_15, P_16, PT_04, PT_05, M_03
		- Existing mortgage is transferred along with property instead of writing new mortgage	P_15, P_16, M_03, M_04, M_11, P_05, PT_04, PT_05
	By making mortgage payments in cash	- Mortgage payments are paid in cash or other negotiable instruments	M_13, MP_03, MP_04, MP_05
Quickly convert dirty funds into a real estate asset	By undervaluing and paying the difference under the table	- Sales price is below expected sales price	P_05, PT_06, PT_07
	By buying the house outright with dirty or mixed funds	- Property acquired without a mortgage	PT_07, PT_19
		- First property acquired without a mortgage	NPO_06, NPO_07, PT_07, PT_19
		- Buyer rushes to complete transaction without good cause	P_05, PT_05, PT_10, X_08
		- Sale is completed without using a real estate broker or sales agent	PT_17
		- Deposit is paid with cash or other negotiable instrument	PT_13
		- Down payment is paid with cash or other negotiable instrument	PT_20
	By defaulting on a loan to one-self or an accomplice	- No interest or principal is repaid on mortgage	P_05, PT_05, P_15, P_16, M_14, M_09, M_10, M_13, MP_03, MP_04
		- Owner is an LLC which defaults on the mortgage	P_04, P_15, P_16, M_14, M_09, M_10, M_13, MP_03, MP_04
		- Other indicators from “By using a shell company or trust” (above)	P_04, M_14
		- Other indicators from “By lending the money to oneself” (above)	P_04, M_14, M_05
Flip houses to turn dirty funds into clean funds	By buying a property and selling it as soon as possible	- Owner buys properties frequently without the number of owned properties rising	NPO_06, NPO_07, P_05, PT_05
	By conspiring with the owner and selling it back to them	- Owner recently sold the same property	P_05, PT_04, PT_05
		- Property is frequently bought and sold	P_05, PT_05, P_03, P_11
		- Purchase and quick sale of property with significant increase/decrease in price	P_05, PT_05, PT_07
	By renovating a property with cash/dirty funds	- Property received extensive renovations paid for in cash	P_12, P_13, P_14
	By purchasing a pre-construction condo and reselling it before the completion date	- Condo sales date is before construction date, but not by developer	P_04, P_03, P_05, PT_04, PT_05, PT_22

	By selling an existing property on assignment	- Sale of existing property is an assignment	PT_23
		- Property is sold repeatedly on assignment before closing	PT_23
		- Uncompleted sale has very long closing date	PTU_05, PTU_12
Hide capital in other jurisdictions	By purchasing property in a country with strong property rights	- Owner address is in another country	NPO_03
		- Owner is from a country with strict capital controls	NPO_03, NPO_05, X_03, NPO_16
		- Owner address is in a country with strong bank secrecy laws / weak AML laws	NPO_03, X_05
		- Owner is from a country with high corruption	NPO_03, NPO_05, X_04, NPO_16, NPO_17
		- Owner is a PEP	NPO_02, X_02
		- Owner is related to a PEP	NPO_24, X_02
		- Owner does not have other economic activity in Canada	NPO_08
Purchase property for criminal use	By purchasing property in an area of interest	- Owner owns property in both very high crime and very low crime neighborhoods	NPO_06, P_02, X_09, X_10
		- Owner buys property above market price in a high crime neighborhood	NPO_06, P_02, P_05, PT_05, PT_06, PT_07, X_09, X_10
		- Buyer shows considerable interest in transactions related to buildings in a particular area	P_02, PTU_02
Launder money through cancelled real estate transactions	By paying a deposit, reneging, and receiving a clean refund	- Buyer reneges on real estate deal after placing deposit	PTU_12, PTU_13, PTU_21
Acquire an income source and mix in dirty money	By purchasing rental property and padding rental income	- Rental income is higher than expected for properties owned	LPO_06, LPO_17, LPO_18, NPO_25, X_06
Miscellaneous	Miscellaneous	- Company has frequent international transactions but no apparent business	LPO_15, FT_01, FT_02, FT_03, FT_04, FT_05, FT_06, FT_07
		- Company is party to suspicious transactions	LPO_16
		- Owner has frequent international transactions	NPO_18, FT_01, FT_02, FT_03, FT_04, FT_05, FT_06, FT_07
		- Owner is party to suspicious transactions	NPO_19
		- Owner has a criminal record	NPO_20
		- Owner is a resident in risk territory	NPO_03, X_04, X_05
		- Owner is from a risk territory	NPO_03, NPO_05, NPO_17, X_04, X_05
		- Owner's spouse is a resident in risk territory	NPO_23, NPS_03, X_04, X_05
		- Owner's spouse is from a risk territory	NPO_23, NPS_05, NPS_17, X_04, X_05
		-Owner has forfeited a property in the past	NPO_21
		-Owner has been prosecuted for ML in the past	NPO_22

14 APPENDIX 3: DATA AVAILABILITY BY MONEY LAUNDERING INDICATOR

14.1 METHODOLOGY

Error! Reference source not found. provides a summary of information at the data point level which is brought up to an aggregate estimation at the sub-scheme level.

14.1.1 Data point level

The availability of the data for each data point is assigned a value on a four point scale:

3. Data are held within a public institution, or are publicly available, and are stored in a format suitable for analysis, such as in structured and cleaned data sets.
2. Data are held within a public institution, or are publicly available, but are not stored in a format suitable for analysis, such as in an administrative database.
1. Data exists but would require significant work to compile or use; example include paper records, or a list of lawyer's offices.
0. Data does not exist or is not held by public institutions and is not publicly available.

14.1.2 Indicator Level

The score for the least available data point necessary for detecting the indicator is chosen. See Section 14.2, below, for a list of scores by indicator.

14.1.3 Sub-scheme Level

The scores of all the indicators of the given sub-scheme are averaged. The following simplification is then applied:

Less than one: Mostly Unavailable
Less than two: Poor
Two or greater: Fair to Good

Sub-schemes rated Fair to Good underwent and additional manual assessment of the estimated ability to construct the required indicators.

See Section **Error! Reference source not found.** for a list of assessments by sub-scheme.

14.2 DATA AVAILABILITY BY INDICATOR

Scheme	*Sub-Scheme*	*Indicator*	*Data Availability*
Unlawfully obscure beneficial ownership	By using a shell company or trust	- Owner is a company or trust	3
		- Owner is a numbered company	3
		- Owner company is incorporated abroad	3
		- Owner company is foreign-controlled	3
		- Owner company has foreign address	3
		- Owner company has public address (PO box or lawyer's office)	1
		- Owner company is in a state with strong secrecy laws	3
		- Owner company has no tax activity	3
		- Owner company has no employees	3
		- Owner company has no address	3
		- Owner company has no web presence / logo/ phone number/etc.	1
		- Owner company has complex ownership structure	3
		- Owner company was just established	2
		- Owner company purchases residential property but its industry is not real estate	3
		- Owner company has few resources relative to property value	3
		- Owner company is a foundation, cultural or leisure association or non-profit entity	3
	By using a formal or informal nominee owner	- Owner and spouse have insufficient income/wealth for property	0
		- Owner is related to other owners with insufficient income/wealth for their properties	0
		- Owner is related to PEP	0
		- Owner is a minor	2
		- Owner is incapacitated	3
		- Owner is institutionalized	0
		- Owner has a legal guardian	0
		- Buying client changes purchasing party's name shortly before closing the purchase	0
		- Buyer pays initial deposit with cheque from unrelated third party	0
		- Deposit is made through a lawyer's or notary's trust account	0
		- Buyer insists on providing signatures via fax only	0
		- Owner is a trustee	3
	By having an unclear address	- Owner has a PO box as their listed address	3
		- Owner has a lawyer or notary as their listed address	1
		- Owner has a Corporate Service Provider as their listed address	1
Funnel cash/money through mortgages	By acquiring many mortgages	- Owner holds multiple mortgages	2
		- Owner has insufficient income to support said mortgages	0
		- Owner acquired multiple properties in a short period of time	2
		- Owner is linked to others who acquired multiple mortgages in a short period of time	0
		- Property is repeatedly mortgaged to the same owner	0

	By repeatedly mortgaging a property	- Mortgage is paid back quickly	0
	By acquiring an outsized mortgage	- Mortgage has unusual loan-to-value	2
		- Mortgage has unusual interest rate	0
		- Mortgage is with an unregulated lender	1
	By lending the money to oneself	- Lender company is owned by borrower	0
		- Lender and borrower are same person	3
		- Lender company is incorporated abroad	3
		- Lender company is foreign-controlled	3
		- Lender company is in a state with strong secrecy laws	3
		- Lender company has no tax activity	3
		- Lender company has no employees	3
		- Lender company has no address	3
		- Lender company has no web presence / logo/ phone number/etc.	1
		- Lender company has complex ownership structure	3
		- Lender company was just established	2
		- Lender company is in non-financial sector	3
		- Lender is a natural person	3
		- Loan is syndicated between multiple unregulated companies or individuals	1
		- Loan has no collateral	0
		- Loan has collateral unrelated to the property	0
		- Loan has very low value collateral	0
	By acquiring very valuable property with a large mortgage	- Rental or business property remains unused	2
		- Rent in rental property is under market value	0
	By taking over existing mortgages	- Mortgage is transferred to non-owner	0
		- Seller retains existing mortgage	2
		- Existing mortgage is transferred along with property instead of writing new mortgage	0
	By making mortgage payments in cash	- Mortgage payments are paid in cash or other negotiable instruments	0
Quickly convert dirty funds into a real estate asset	By undervaluing and paying the difference under the table	- Sales price is below expected sales price	3
	By buying the house outright with dirty or mixed funds	- Property acquired without a mortgage	3
		- First property acquired without a mortgage	2
		- Buyer rushes to complete transaction without good cause	0
		- Sale is completed without using a real estate broker or sales agent	2
		- Deposit is paid with cash or other negotiable instrument	0
		- Down payment is paid with cash or other negotiable instrument	0
	By defaulting on a loan to one-self or an accomplice	- No interest or principal is repaid on mortgage	0
		- Owner is an LLC which defaults on the mortgage	0
		- Other indicators from “By using a shell company or trust” (above)	0
		- Other indicators from “By lending the money to oneself” (above)	0

Flip houses to turn dirty funds into clean funds	By buying a property and selling it as soon as possible	- Owner buys properties frequently without the number of owned properties rising	2
	By conspiring with the owner and selling it back to them	- Owner recently sold the same property	2
		- Property is frequently bought and sold	2
		- Purchase and quick sale of property with significant increase/decrease in price	2
	By renovating a property with cash/dirty funds	- Property received extensive renovations paid for in cash	0
	By purchasing a pre-construction condo and reselling it before the completion date	- Condo sales date is before construction date, but not by developer	2
	By selling an existing property on assignment	- Sale of existing property is an assignment	0
		- Property is sold repeatedly on assignment before closing	0
		- Uncompleted sale has very long closing date	0
Hide capital in other jurisdictions	By purchasing property in a country with strong property rights	- Owner address is in another country	3
		- Owner is from a country with strict capital controls	3
		- Owner address is in a country with strong bank secrecy laws / weak AML laws	3
		- Owner is from a country with high corruption	3
		- Owner is a PEP	0
		- Owner is related to a PEP	0
		- Owner does not have other economic activity in Canada	3
Purchase property for criminal use	By purchasing property in an area of interest	- Owner owns property in both very high crime and very low crime neighborhoods	3
		- Owner buys property above market price in a high crime neighborhood	2
		- Buyer shows considerable interest in transactions related to buildings in a particular area	0
Launder money through cancelled real estate transactions	By paying a deposit, reneging, and receiving a clean refund	- Buyer reneges on real estate deal after placing deposit	0
Acquire an income source and mix in dirty money	By purchasing rental property and padding rental income	- Rental income is higher than expected for properties owned	0
Miscellaneous	Miscellaneous	- Company has frequent international transactions but no apparent business	3
		- Company is party to suspicious transactions	3
		- Owner has frequent international transactions	3
		- Owner is party to suspicious transactions	3
		- Owner has a criminal record	3
		- Owner is a resident in risk territory	3
		- Owner is from a risk territory	3
		- Owner's spouse is a resident in risk territory	0
		- Owner's spouse is from a risk territory	3
		-Owner has forfeited a property in the past	2
		-Owner has been prosecuted for ML in the past	3

15 APPENDIX 4: PRIVACY ENHANCING TECHNOLOGIES

Organizations, whether public or private, are subject to legislative requirements that relate to privacy and information sharing. Furthermore, private organizations may not be in a position to share information that may affect market competition. As such, PETs are tools that are used to overcome barriers to data sharing. Privacy Enhancing Technologies (PETs) are set of techniques which are used to overcome barriers to data sharing by facilitating the sharing or analysis of data while protecting personal or business information and enhancing privacy protections.

The approaches themselves are not all new, but their use is an emerging phenomenon as requirements for data insights and concerns about data breaches have grown. This growing interest has led to recent reports by the World Economic Forum (WEF, 2019) and The Royal Society (2019), which highlight a variety of emerging techniques and their applications.

In Canada, PETs are used in some specialized projects, but their use does not appear to be widespread. Nor were any clear statements on the ways these technologies interact with existing privacy legislation readily available.[34]

Four distinct PETs are presented below to provide an illustration of the possibilities available with their use. These were chosen due to their applicability to preserving privacy in data sharing and dissemination scenarios.[35]

15.1.1 Differential Privacy

Differential privacy is a class of algorithms that adds noise to data such that information about an individual or entity is protected while allowing data users to extract and analyzormation about the population (Dwork & Roth, 2014). Analysis on datasets with suitable differential privacy algorithms will reach the same general conclusions as the original data, but will protect the privacy of individuals within the data by disturbing the original records. Such approaches are leveraged at Statistics Canada when producing Public Use Microdata Files (PUMFs). For example, the release of the 2011 National Household Survey employed a number of data perturbation methods including adding random noise to continuous variables, swapping responses to certain variables between cases, and the application of specific changes to a random selection of cases. Various controls were taken when applying these methods to ensure that the changes would not affect aggregate analytical conclusions across a range of dimensions (Tambay, Carillo-Garcia, & Kanagarajah, 2015).

It must be noted that this class of algorithms create a trade-off between accuracy and privacy as information on individual cases is actively lost and scrambled in the process (WEF, 2019). As such, differential privacy should be employed when data is shared with the goal of evaluating aggregate population or market trends as opposed to detecting anomalies on specific records.

[34] The office of the Privacy Commissioner has conducted a review of privacy enhancing technologies, but the review employed a different used of the term, referring to technologies which individuals employ to protect their personal data, such as secure messaging and anti-tracking tools for online browsing (OPC, 2017).

[35] Techniques relevant to data-sharing not explored here included Trusted Execution Environments (The Royal Society, 2019) and Secure Multiparty Computation (WEF, 2019).

15.1.2 Federated Analysis

Federated analysis seeks to share analytical insights from multiple independent datasets while abiding by data sharing restrictions. Independent analyses are conducted on the individual datasets, followed by a sharing of the insights received from these analyses (WEF, 2019). The results from the individual analyses are shared to create an aggregate model without sharing the underlying data.

This technique is especially useful when there is a large number of disparate datasets (WEF, 2019). For example, it is employed at Google to run models on users' phone. The analytical insights are then shared back to the company without personal information ever leaving the devices (Bonawitz et al., 2019). This protects the privacy of the users while aggregating the results received from all devices. One caveat is that the model obtained by employing federated analysis will typically be inferior to one that merged all distinct datasets (WEF, 2019).

In Canada, Federated Analysis is employed in the CanDIG project, which allows for the analysis of genetic data held at separate health care institutions in Canada (CanDIG, 2020).

15.1.3 Homomorphic Encryption

This PET encrypts data and allows for analysis to be conducted on the encrypted data (Lauter, Naehrig, & Vaikuntanatha, 2011). As a very simple example, encrypted values can be added together to produce an accurate encrypted sum, but the party performing the processing does not know what this sum is, only the party with the decryption key would be able to decrypt the sum. A more complex example would be matching an encrypted transaction pattern to rows in an encrypted database of transactions and returning the encrypted dissimilarity scores. The data holder can then decrypt the dissimilarity scores to see which rows require attention. The underlying information is never actually readable by the party processing or analyzing the data.

In theory, homomorphic encryption overcomes the weakness of federated analysis by allowing disparate datasets to be combined and analyzed as a whole, without exposing their content. In practice, current encryption algorithms limit the number and types of mathematical operations which can be performed on encrypted data, and the computational costs of operations are much greater than those of analysing unencrypted data (WEF, 2019). Developments have overcome some challenges by constructing a fully homomorphic encryption scheme, allowing many computations to be performed on encrypted data (Gentry, 2009).

In Australia, the Innovations Hub in Fintel Alliance is pursuing this technology for a future real-time alerting project. The implementation uses homomorphic encryption to create a secure multiparty computation environment in which Fintel Alliance will be able to run an analysis across data held at multiple institutions in order to detect patterns which are only visible when the data is analyzed as a whole. However, at no point can the party performing the analysis read the underlying data; any flagged sets of transactions must then be manually followed up on to learn more about them (Maxwell, 2020).

15.1.4 Zero-Knowledge Proofs

Zero-Knowledge Proofs (ZKPs) are used to confirm information without disclosing specific data about an individual or entity. For example, using ZKPs, a landlord could verify that a potential tenant meets a

certain minimum level of income, without revealing the tenant's exact income level. ZKPs can also be used to confirm a match for a record in a data set, without revealing details on the record being matched. For example, Bank A may have a dataset of fraudulent customers. Bank B may want to check new customers against this dataset without revealing to Bank A who their new customers are. ZKPs enable this kind of information sharing among financial institutions.

The Dutch Fiscal Information and Investigation Service (FIOD) and UK's HM Revenue and Customs have jointly created a European financial crimes information-sharing platform, FCI-net, which enables participating institutions to query of the information held by participating institutions without revealing the contents of the query or the contents of the data. Once a match is found, traditional channels can be used to obtain the data of interest. FCI-net uses Bloom Filters, an older technology which does not entirely preserve privacy, but the process is similar to the ZKP matching systems (OECD & The World Bank, 2018, p. 61-2; Stritzl, 2019).

16 APPENDIX 5: WORK STREAM MEMBERS

BC-Canada Working Group on Real Estate

Work Stream 1: Data Collection and Sharing

Co-chairs:
Jonathan Baron, A/Executive Director, Data, British Columbia Ministry of Finance
Haig McCarrell, Director, Investment Science and Technology Division, Statistics Canada

Work stream members:
Wahid Abdallah, Specialist, Policy Analysis, Canada Mortgage and Housing Corporation
Ellen Bekkering, Chief, Canadian Housing Statistics Program, Statistics Canada
Olga Bilyk, Senior Economist, Financial Stability Department, Bank of Canada
Luc Chartrand, Team Leader, Data Exploitation Lab, FINTRAC
Melanie Chauvin, Team Leader, Financial Crime Analysis Unit, RCMP
Jean-Philippe Deschamps-Laporte, Chief, Canadian Housing Statistics Program, Statistics Canada
Daniel-Alexandre Gagnon, Analyst, Financial Systems Division, Finance Canada
Annik Gougeon, Analyst, Canadian Housing Statistics Program, Statistics Canada
Philip Greenwood, Manager Compliance & Examinations, BC Financial Services Authority
Jason Jensen, Analyst, Canadian Housing Statistics Program, Statistics Canada
Élisabeth Koulouris, Senior Specialist, Housing Research, Canada Mortgage and Housing Corporation
Ian McDonald, Policy Analyst, Serious and Organized Crime, Federal Policing Strategic Direction, RCMP
Debbie Morreau, Real Estate Council of British Columbia
Bert Pereboom, Senior Manager, Housing Market Policy, Canada Mortgage and Housing Corporation
Raminder Pooni, Assistant Director, Criminal Investigations, Western Region, Canada Revenue Agency
Andrew Richardson, Chief of Compliance, Corporate Finance, BC Securities Commission
Amanda Taschereau, FINTRAC
Bruce Wallace, FINTRAC