SECURITY INTELLIGENCE REVIEW COMMITTEE

TOP SECRET

SIRC REVIEW 2015-02 REVIEW OF CSIS'S USE OF DATA MANAGEMENT AND EXPLOITATION ACTIVITIES

SUMMARY

- This review examined CSIS's Data Acquisition Program (DAP) and its associated governance structures. The DAP was created in June 2011 to manage the task of integrating bulk datasets within CSIS's data exploitation program. The acquisition of these bulk datasets is meant to provide context and insight to CSIS investigations, particularly when combined with high value intelligence, much of which is collected under warrant.
- The review focussed on the collection of datasets for use by the DAP, and whether the collection is in compliance with the *CSIS Act*, in particular section12, which requires that collection be limited to information that is "strictly necessary."
- The review examined the bulk datasets in the DAP's holdings, which are broken down into two broad types. The first type is "referential," which means they are openly sourced and publicly-available and therefore not "collected" under the authority of section 12 of the *CSIS Act*. Though SIRC agrees with the principle that the acquisition of purely referential datasets would not constitute "collection" per se, SIRC does not agree with CSIS's assessment that its largest referential dataset is, in fact, referential.
- SIRC recommended that CSIS re-evaluate its referential bulk datasets to ensure that they should continue to be considered referential; those that are deemed not to be referential should be assessed against the "strictly necessary" threshold.
- CSIS also has "non-referential" datasets, which are "collected" under the authority of the CSIS Act, and so must meet the collection threshold of "strictly necessary." Despite this, SIRC found no evidence to indicate that CSIS has appropriately considered the threshold as required in the CSIS Act. SIRC therefore recommended that CSIS undertake a formal and documented assessment for each of its existing non-referential datasets to ensure the information was collected to the extent that was "strictly necessary."
- SIRC further recommended that CSIS halt its acquisition of bulk datasets until it has in place a formal process of assessment to confirm that the datasets meet the required threshold.

ATIP version FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET-

Table of Contents

1 INTRODUCTION	
2. METHODOLOGY	
2.1 Review Activity and Criteria	4
3 BACKGROUND: THE DATA ACQUISITION PROGRAM (DAP)	5
3.1 Governance	6
4 REFERENTIAL BULK DATASETS	8
4.1 The Case of Bulk Data	8
5 NON-REFERENTIAL BULK DATASETS	11
5.1 Addressing strictly necessary	11
5.2 The principle of proportionality	14
5.3 Assessing instances where bulk collection is required	15
5.4 Testing bulk data for strict necessity	16
6 CONCLUSION	18
ANNEX A - SUMMARY OF FINDINGS	19
ANNEX B - SUMMARY OF RECOMMENDATIONS	20

ATIP version FEB 2 0 2019

2

SIRC STUDY 2015-02

TOP SECRET-

1 INTRODUCTION

Data gathering and exploitation is increasingly regarded as essential in the fight against international terrorism and other threats to national security.

This review marks SIRC's first examination into CSIS's Data Acquisition Program (DAP) and its associated governance structures. The review focussed on the collection of datasets for use by the DAP, and whether the collection is in compliance with the *CSIS Act*, in particular section12, which requires that collection be limited to information that is "strictly necessary."

Though the acquisition of bulk datasets by CSIS began in earnest in 2011 with the creation of the DAP, the review found that there is currently no formal governance framework to set limits around the acquisition and use of bulk datasets and no guidance concerning how to apply the limitations of "strictly necessary" to bulk dataset collection. In addition, SIRC found that CSIS has not sought formal legal guidance with respect to bulk dataset acquisition. Accordingly, SIRC recommends that CSIS finalize and implement the governance framework for dataset acquisition no later than February 1, 2016.

SIRC examined the list of bulk datasets in the DAP's holdings, which are broken down into two broad types. The first type is "referential", meaning that they are used primarily to facilitate identity verification, such as matching a name with a phone number. According to CSIS, because referential datasets are openly sourced and publicly-available, they are not "collected" under the authority of section 12 of the *CSIS Act.* SIRC agrees with the principle that there are instances when the acquisition of purely referential datasets would not constitute "collection" per se. However, SIRC does not agree with CSIS's assessment that its largest referential dataset is, in fact, referential according to its own criteria. SIRC recommends that CSIS re-evaluate all referential bulk datasets against its criteria to ensure that they should continue to be considered referential; those that do not should be assessed against the strictly necessary threshold.

CSIS also has "non-referential" datasets, which it categorizes as having been "collected" under the authority of the *CSIS Act*, and so must meet the collection threshold of "strictly necessary." Despite this, SIRC found no evidence to indicate that CSIS has appropriately considered the threshold as required in the *CSIS Act*. SIRC therefore recommends that CSIS undertake a formal and documented assessment for each of its existing non-referential datasets to ensure the information was collected to the extent that was "strictly necessary." Ultimately, SIRC recommends that CSIS halt its acquisition of bulk datasets until it has in place a formal process of assessment to confirm that the bulk datasets meet the required threshold.

SIRC STUDY 2015-02

TOP SECRET

2 METHODOLOGY

One of the central objectives of SIRC's first review into CSIS's data acquisition and exploitation activities was to establish a baseline of information into this activity. To that end, the review sought to establish the quantity and types of bulk¹ datasets used by CSIS's DAP and to explore the policy and/or legal considerations relied upon by CSIS to guide its activities in this new area of activity. The review also considered how CSIS used its bulk datasets by examining a sample of products generated through the exploitation of these bulk datasets. This included an examination of the taskings from operational branches for the products of data analytics.

The core review period for this study was January 1, 2013 to December 31, 2014, although information falling outside of this period was examined to make a full assessment.

2.1 Review Activity and Criteria

SIRC set out to assess CSIS's acquisition of bulk datasets against the requirements of the *CSIS Act*, in particular section 12, which requires that collection be limited to information that is "strictly necessary;"² this included an intent to examine CSIS's test for evaluating "strictly necessary" as it is applied to the acquisition of bulk datasets. SIRC also set out to examine the DAP's internal policies and procedures with respect to data exploitation and acquisition. To that end, SIRC sought policy or framework documents related to data acquisition and exploitation.

SIRC met with CSIS representatives to provide context to the issues under review. The discussions included multiple meetings with representatives

by SIRC benefited from several demonstrations provided by of its methods and tools associated with data exploitation.

ATIP version

¹ SIRC uses the term "bulk" to distinguish it from more targeted collection.

²Section 12 reads as follows: 12. (1) The Service shall collect, by investigation or otherwise, to the extent that it is strictly necessary, and analyse and retain information and intelligence respecting activities that may on reasonable grounds be suspected of constituting threats to the security of Canada and, in relation thereto, shall report to and advise the Government of Canada.

SIRC STUDY 2015-02

TOP SECRET

3 BACKGROUND: THE DATA ACQUISITION PROGRAM (DAP)

The DAP was created in June 2011 to manage the task of integrating bulk datasets within the Operational Data Analysis Centre (ODAC), formed in 2006 to develop a data exploitation program. The acquisition of these bulk datasets is meant to provide context and insight to CSIS investigations, particularly when combined with high value intelligence, much of which is collected under warrant.

ATIP version

FEB 2 0 2019

November 18, 2015

5

SIRC STUDY 2015-02

TOP SECRET

Generally speaking, there are three ways in which information from bulk datasets is engaged. First, datasets can be used to conduct indices checks by using information or a selector already connected to a potential threat, such as an address, phone number to search for "hits" in the data. Second, datasets can be

used to enhance knowledge of a target by searching the data for previously undetected trends, links or patterns between and among data points. In this way, the data can also be used to create a number of visual representations

Finally, CSIS uses the data to try to identify previously unknown individuals of interest by linking together types information which have mirrored threat behaviour,

The DAP also supports more focused inquiries. For instance, regional investigators regularly approach the DAP for such as "individuals who have not otherwise

come to CSIS's attention

3.1 Governance

As noted, the DAP has developed a procedure to address the handover of the data from its source to the point that the data is ready for exploitation by DAP analysts; however, beyond the more technical aspects of ingestion, **SIRC found that there is no governance framework guiding the collection, retention and use of bulk datasets.** This situation remains today despite the fact that SIRC has seen references in earlier CSIS documentation to the need to validate the authority to collect and to manage the risk of over-collection by confining collection to that which is "strictly necessary."¹⁰ Moreover, the mandate of the DAP includes the need to "identify critical acquisition issues and develop a framework for obtaining bulk data."¹¹To that end, SIRC was told that a formal governance framework was drafted two years ago, but it has yet to be finalized.

SIRC therefore recommends that the governance framework for bulk dataset acquisition be finalized and implemented no later than February 1, 2016. SIRC believes that this framework should, among other issues, set parameters around collection based on the statutory requirement that collection be limited to that which is "strictly necessary." Ongoing dataset management issues would need to be addressed

ATIP version

Gated: _____ FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

as well, to ensure that the datasets being used by the DAP continue to be relevant and those datasets that are not, are deleted. As CSIS finalizes its governance framework, SIRC also suggests that some consideration be given to the experience of

As will be discussed in the next two sections, SIRC is of the opinion that CSIS's lack of an overarching framework In

particular, CSIS has not set out, in explicit terms, how the datasets meet the test of strict necessity. Neither does CSIS have a formalized process that would require senior level officials to approve and challenge the decision to ingest new datasets based on those assessments.¹⁴ Finally, the few existing procedures documenting data acquisition do not contain a process to determine whether or not it was "strictly necessary" to acquire the data in question.

ATIP version

FEB 2 0 2019

¹⁴UK Intelligence Services Commissioner, 2014 Annual Report.

SECURITY INTELLIGENCE REVIEW COMMITTEE

SIRC STUDY 2015-02

TOP SECRET

4 REFERENTIAL BULK DATASETS

CSIS's "data-catalogue" contains two broad types of datasets:

"referential." CSIS has noted that referential datasets are not acquired for the express purpose of exploitation, meaning that they do not have a stand-alone intelligence value. Instead, these datasets are meant to "enrich" or provide context into intelligence collected by CSIS in more conventional ways. Examples of Canadian referential datasets are

According to CSIS, referential datasets share two criteria: these datasets are openly acquired, meaning no tradecraft or special CSIS access was employed in their acquisition and they are available publicly, meaning that anyone would be able to acquire them. The result, as CSIS argued, is that "the acquisition of referential material, in whatever format, that is openly and publicly available is not acquired under any specific authority because it is not considered by DMEX to be collection pursuant to section 12 of the *CSIS Act*."¹⁵ Instead, the collection of referential data was likened to acquiring a book for the CSIS library or acquiring a digital copy of the phone book.

4.1 The Case of Bulk Datasets

For the past few years, CSIS has participated in

initiative

¹⁵CSIS response to SIRC memo.

ATIP version

Gated: FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

SIRC's review of CSIS documentation suggests that datasets do not meet the criteria laid out for referential datasets, as the information is not acquired openly nor is it publicly available.

First, CSIS documentation states that During a briefing, SIRC was told that

Moreover, CSIS documentation strongly suggests that,

In other words, this is a unique collection of data that could not be assembled through other means.

Second, SIRC noted that the type of information contained in datasets is , well beyond publicly-available information.

In addition, CSIS has expressed an expectation that will do more than simply provide identity verification" or "enrich" other collected data.

SIRC recognizes that open source reference material can be very useful in the conduct of lawful investigations into threats to national security.

SIRC accepts that CSIS may acquire such referential data. However, SIRC found that datasets did not meet CSIS's own criteria regarding referential data, and as such, constituted "collection" as understood in the CSIS Act. Accordingly, SIRC recommends that CSIS assess the acquisition of datasets against the "strictly necessary" threshold.

ATIP version

SIRC STUDY 2015-02

TOP SECRET

Furthermore, SIRC recommends that CSIS re-evaluate all other referential bulk dataset acquisitions to ensure that they meet the criteria for inclusion in the referential category. Those that do not should be assessed to ensure the information was collected to the extent that was "strictly necessary."

ATIP version

Gated: _____FEB 2 0 2019

SECURITY INTELLIGENCE REVIEW COMMITTEE

SIRC STUDY 2015-02

TOP SECRET

5 NON-REFERENTIAL BULK DATASETS

The DAP also collects non-referential datasets, which CSIS categorizes as having been "collected" under the regime of section 12 of the *CSIS Act*, and therefore having to meet the collection threshold of "strictly necessary."

datasets vary: some datasets are provided by

SIRC set out to assess the datasets against the relevant legislative threshold, expecting to see that CSIS considers the limitations of strictly necessary in its acquisition process.²² In fact, there is no standing requirement to document a justification for the collection of the datasets against the threshold of strict necessity. SIRC also expected that CSIS would have sought legal guidance to address the boundaries and applications of "strictly necessary" to bulk collection. Despite references to the need for a legal opinion in documentation in 2005 when ODAC was being created, ²³SIRC found that there has been no formal legal guidance from CSIS Legal Services regarding whether or not the acquisition of bulk datasets was consistent with CSIS's mandate as outlined in section 12 of the CSIS Act.

What follows is an assessment by SIRC of what should constitute the principles against which adherence to "strictly necessary" are measured in this context, as well as examples to illustrate those principles.

5.1 Addressing strictly necessary

During the parliamentary debates surrounding the passage of the *CSIS Act*, the then Solicitor General noted that "the mandate [of CSIS] as worded in the Bill before you, limits all security investigations to those that are "strictly necessary," in the interests of national security. This is a clear signal that the mandate is to be interpreted narrowly.²⁴ The Solicitor General would subsequently add that "[w]e do not want [CSIS] to collect any more than is strictly necessary, because it is the collection that is the potential violation of people's privacy and rights."²⁵

ATIP version FEB 2 0 2019

²²SIRC requested copies of the a blank version of which appeared in the documentation, for select datasets. These forms should contain the justification for datasets as they are acquired. SIRC was told that a draft, to be finalized along with the overarching policy framework.

²³Data Exploitation Task Force Report, 2005.

²⁴Debates of the House of Commons, February 10, 1984, p.1274.

²⁵ Parliamentary Committee on Justice and Judicial Matters, May 24, 1984, p.28:52.

SIRC STUDY 2015-02

TOP SECRET

"Strictly necessary" means more than merely useful or expedient, two criteria which tacitly underpinned much of the explanation provided to SIRC by CSIS personnel over the course of this review. SIRC holds that when determining whether information collected by CSIS is "strictly necessary," and thus within the limits of CSIS's lawful mandate, three principles come into play:

- 1) The *CSIS Act* places meaningful parameters around, and limits upon, collection. It was the clear intent of Parliament to do so and there nothing to indicate to this day that the law should be interpreted otherwise.
- The onus is on CSIS to proactively assess whether any collection activity meets the strictly necessary threshold, and hence whether collection in each case is justified.
- 3) If there is any uncertainty left at the end of the assessment, it is the limitation that should prevail, and collection should not take place.

These principles, which represent a fair and honest interpretation of the powers provided to CSIS by its mandate, must explicitly frame all CSIS's collection activities. SIRC notes in passing that a review of practices in other jurisdictions suggests that the concept of "necessity" is not uncommonly used to animate discussion of intelligence activities. For example, the UK's Independent Reviewer of Terrorism Legislation explains in a discussion on bulk collection that state interference with a human right must satisfy several tests, including that the interference is "necessary in pursuit of a legitimate aim." The Independent Reviewer states that, though the courts are generally willing to find that a legitimate aim is being pursued, "*necessary*' means less than *'indispensable*', but more than merely *'admissible*' or *'useful.*'²⁶In the context of Canada, section 12 of the *CSIS Act* does not stop at "necessary," but rather refers to "strictly necessary," which is a higher threshold.

Given the high threshold to which collection under "strictly necessary" must adhere, SIRC was concerned by the absence of an overarching policy or governance framework and by CSIS's failure to seek legal guidance. Moreover, in validating whether or not the collection of bulk datasets adheres to "strictly necessary," CSIS employs a set of procedures that contain no instructions or guidelines on how the threshold should be interpreted. Although the documented procedures include questions regarding dataset

dated:

²⁶Anderson's discussion in Chapter 5, Legal Constraints. This quote refers to standards derived from the European Convention on Human Rights, which have been imported into UK law.

SIRC STUDY 2015-02

TOP SECRET-

SIRC attempted to address this ongoing gap by requesting that CSIS generate a proper assessment for acquired bulk datasets selected for review.²⁹ The assessments provided were highly general, with few details or specific examples of the kind that SIRC would expect to see to justify the collection of bulk datasets.

the main thrust was that the information contained in the bulk datasets was in some way connected to an ongoing, mandated investigation.

Missing, in SIRC's view, is the level of specificity required to judge whether any of these cases, "strictly necessary."

SIRC agrees with CSIS that establishing a connection between the collection of a bulk dataset and a legitimate investigation or threat as defined by section 2 of the *CSIS Act* is an essential component of any consideration of strict necessity. On its own, however, SIRC finds that this is not sufficient to meet the threshold of "strictly necessary," which imposes meaningful limitations on CSIS's collection activities. It follows that a greater degree of detail is required to make the case of strict necessity.

ATIP version FEB220020049

SECURITY INTELLIGENCE REVIEW COMMITTEE

²⁹ The datasets selected are:

SIRC STUDY 2015-02

TOP SECRET

5.2 The principle of proportionality

One of CSIS's core operating principles asserts that "the least intrusive techniques must be used first, except in emergency situations or where less intrusive investigative techniques would not be proportionate to the gravity and imminence of the threat."³² To be proportionate, the expected intelligence must be sufficient to justify the intrusion and there must be an assessment of whether the information could not reasonably be obtained by less intrusive means. SIRC is concerned that the current direction of CSIS's bulk collection and exploitation activities, which favours expanding CSIS's direct access to bulk data, are at odds with the principle of proportionality.

ATIP version

³²The proportionality principle is well established in other jurisdictions, including in the UK, where the Intelligence Services Commissioner stated that intrusions into privacy must be proportionate, meaning that a less intrusive means could not have been used. "Report of the UK Intelligence Services Commissioner for 2014". Similarly, for the Australian Inspector General, the factors to be considered in the proportionality test include the privacy impacts of proposed investigative activity, the public interest in the proposed investigative activity, and the availability and effectiveness of less intrusive investigative techniques. IG of Intelligence and Security, Submission to the "Inquiry into potential reforms of National Security Legislation," August 2012.

SIRC STUDY 2015-02

TOP SECRET

Indeed, one of the potential long-term implications of bulk data acquisition is the accumulation of a significant volume of information on individuals, including Canadian citizens, with no connection to a threat.

For SIRC, cases are examples of what may constitute the difference between collection that is "strictly necessary" and collection that merely holds the unknown potential to be useful. In both cases, threat related information

was

already being collected by CSIS. Therefore the strict necessity of the additional information such as was provided through the sharing in bulk of information becomes, for SIRC, harder to justify.

5.3 Assessing instances where bulk collection is required

SIRC recognizes that some situations will yield no alternatives to bulk collection.

ATIP version

dated: _____FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

fill an intelligence gap

is a

relatively poor indicator of a threat.

Without an objective assessment of why this bulk information is strictly necessary, it is harder for SIRC to distinguish this type of relatively unfocused bulk collection from "fishing" or "profiling." Unfortunately, CSIS has not provided an assessment of this data, including how

through other means

CSIS also did not assess the likelihood of the bulk datasets successfully identifying an "unknown, unknown" through the use of examples of this method being successful, for example, or some other type of objective assessment. As a general principle, the more unfocused the collection, the more stringently should CSIS apply the test of strict necessity.

5.4 Testing bulk data for strict necessity

Overall, SIRC found no evidence to indicate that CSIS's data acquisition program has satisfactorily considered the threshold of "strictly necessary" as required in the CSIS Act. SIRC therefore recommends that CSIS undertake a formal and documented assessment for each of its existing non-referential bulk datasets to ensure the information was collected to the extent that was "strictly necessary;" this practice should extend to all future bulk dataset acquisitions.

To assist in that task, SIRC has developed the following guidelines, each of which is meant to promote conformity to the threshold of "strictly necessary:"

- For any bulk information, a clear connection to a threat to the security of Canada as defined in section 2 of the *CSIS Act* must be established.
- It must be established that less intrusive means that would satisfy the intelligence requirements are not available as an alternative to bulk collection, consistent with the principle of proportionality.
- Finally, if there is no reasonable alternative to bulk collection, CSIS needs to objectively assess how closely connected the bulk information is to intelligence of value, with the goal of being increasingly strict the more broad the intended collection.

Given the absence of the proper application of guidelines such as those presented above in order to evaluate whether CSIS's bulk dataset collection has adhered to the threshold of "strictly necessary" as required in the *CSIS Act*, given the absence of a

ATIP version dated: FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

formalized governance framework in which these acquisitions take place, and given the lack of formal legal guidance to inform such acquisitions, SIRC makes the following recommendations:

First, SIRC recommends that CSIS cease to acquire any additional bulk datasets until it has put in place a formal and documented process to assess whether such acquisitions are collected to the extent that is "strictly necessary," taking into account the considerations raised by SIRC. This process should be put in place no later than February 1, 2016.

Second, Should this process not be in place by February 1, 2016, SIRC recommends that CSIS discontinue use of bulk datasets acquired under the DAP program until such time as it is.

ATIP version

FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

6 CONCLUSION

CSIS's current approach to the acquisition of bulk datasets runs the risk of overcollection. Though the metaphors of the "needle and the haystack" and "connecting the dots" may be persuasive in the abstract, they cannot on their own justify the collection of volumes of bulk data.

Overall, the Committee has found that the case remains to be made that the acquisition of these bulk datasets has met the threshold of "strictly necessary." Indeed, SIRC has seen no evidence to suggest that CSIS has systematically taken the threshold of "strictly necessary" into consideration; CSIS's lack of process, governance, and legal guidance around the acquisition and management of bulk dataset is lacking, and is not consistent with the practice of close allies.

SIRC believes that these elements will become even more critical moving forward,

As conveyed in this review, SIRC expects the principles made above in regard to "strictly necessary" will be reflected in all future acquisition assessments, and expects proper documentation of the assessment process. SIRC will follow-up in a timely manner in regards to our recommendations; going forward, SIRC will also ask for a full list of the bulk datasets annually as part of its annual certification of the Director's Report to the Minister of Public Safety. The Director's Report should reflect the status of CSIS's bulk dataset acquisition program.

> ATIP version FEB 2 0 2019

SIRC STUDY 2015-02

TOP SECRET

ANNEX A - SUMMARY OF FINDINGS

- SIRC found that there is no governance framework guiding the collection, retention and use of bulk data.
- SIRC found that bulk datasets did not meet CSIS's own criteria regarding referential data, and as such, constituted "collection" as understood in section 12 of the CSIS Act.
- SIRC found that there has been no formal legal guidance from CSIS Legal Services regarding whether or not the acquisition of bulk datasets was consistent with CSIS's mandate as outlined in section 12 of the *CSIS Act*.
- SIRC found no evidence to indicate that CSIS's data acquisition program has appropriately considered the threshold of "strictly necessary" as required in the *CSIS Act*.

ATIP version

SIRC STUDY 2015-02

TOP SECRET.

ANNEX B - SUMMARY OF RECOMMENDATIONS

- SIRC recommends that the governance framework for bulk dataset acquisition be finalized and implemented no later than February 1, 2016.
- SIRC recommends that CSIS cease to acquire any additional bulk datasets until it has put in place a formal and documented process to assess whether such acquisitions are collected to the extent that is "strictly necessary," taking into account the considerations raised by SIRC. This process should be put in place no later than February 1, 2016.
- Should this process not be in place by February 1, 2016, SIRC recommends that CSIS discontinue use of bulk datasets acquired under the DAP program until such time as it is.
- SIRC recommends that CSIS undertake a formal and documented assessment for each of its existing non-referential bulk datasets.
- SIRC recommends that CSIS assess the acquisition of bulk datasets against the "strictly necessary" threshold.
- SIRC recommends that CSIS re-evaluate all referential bulk dataset acquisitions to ensure that they meet the criteria for inclusion in the referential category. Those that do not should be assessed to ensure the information was collected to the extent that was "strictly necessary".

ATIP version

dated: _____FEB 2 0 2019