December 2017

The Representation of the Canadian Government’s Warrantless Domestic Collection of Metadata in the Canadian Print News Media

Alan Del Pino
The University of Western Ontario

Supervisor
Dr. Jacquie Burkell
The University of Western Ontario

Co-Supervisor
Dr. Pam McKenzie
The University of Western Ontario

Graduate Program in Media Studies

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Arts

© Alan Del Pino 2017

Follow this and additional works at: http://ir.lib.uwo.ca/etd

Part of the Other Film and Media Studies Commons

Recommended Citation
http://ir.lib.uwo.ca/etd/5038

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.
Abstract

In January 2014, the former NSA contractor Edward Snowden revealed that Canada’s foreign intelligence agency CSEC was engaging in warrantless electronic surveillance of Canadians by monitoring communications metadata. Prior to these disclosures Canadians knew very little about metadata and about how the CSEC used information technology to collect electronic intelligence. Media outlets such as newspapers are important sources through which Canadians learn about issues such as warrantless surveillance of citizens. However, to date no research analyzes how Canada’s warrantless domestic collection of metadata has been represented in the Canadian new media. This thesis addresses this gap by analyzing the representation of the Canadian government’s domestic collection of metadata in three Canadian news publications, the Toronto Star, the Globe and Mail, and the National Post, from January 2013 to December 2016. This project performs a qualitative and quantitative content analysis of 91 articles on this topic.

The following factors were studied: how the CSEC and the OPC define metadata, how the articles define metadata, the topics observed in the introductory paragraphs of the articles, and how the topics develop over time. The study found a highly significant relationship between the year that the articles were published and the topics that were observed in the introductory paragraphs of the articles. Furthermore, across all news publications there was a fairly even distribution of articles that define metadata by including either the CSEC’s or the OPC’s definition of the term. This means that if Canadians learned about this issue by reading any of the three news publications, they would develop a range of perspectives regarding how metadata is defined. In addition, if Canadians only read the introductory paragraphs of the articles in any of the three publications they would be equally informed about how the coverage on this issue changes over time.
Keywords
Acknowledgments

Without the patient guidance of Dr. Jacquie Burkell this project would not have been possible. Dr. Pam McKenzie also played a critical role in developing this project and making sure that I saw it through to its end. In addition to my co-supervisors, I would like to thank all of the faculty and staff at FIMS who provided a positive support system throughout my time at Western. Matt Ward especially took the time out of his day on many occasions to listen to my academic woes and he regularly provided excellent advice.

To my fellow graduate students, I thank you for the countless hours of ironic and at times cynical conversations about life in academia and the uncertain future after it. You helped me make the best of this stressful and demanding graduate school experience. Yimin Chen and Zak Bronson also taught me the rules of the game when it came to quickly snapping up free food at social events as well as complementary drink tickets at the Grad Club. I am indebted to you both for this sage advice.

Lastly, I would like to thank my parents and all of the faculty at the University of Guelph-Humber who believed in my graduate school ambitions when I was still a young and aimless undergraduate student. My Dad, more than anyone else supported my aspirations to make it in grad school and constantly advised me to keep working hard when the going was tough. I would also like to thank Dr. Natalie Evans and Dr. Greg Kelley at the University of Guelph-Humber who provided outstanding academic recommendations that helped me get into Western’s competitive Master of Arts in Media Studies program with external funding.
# Table of Contents

Acknowledgments.................................................................................................................. iii  
Table of Contents ...................................................................................................................... iv  
List of Tables ............................................................................................................................ vi  
List of Figures ........................................................................................................................... vii  
List of Appendices .................................................................................................................... ix  
Chapter 1 ................................................................................................................................... 1  
  1 Introduction .......................................................................................................................... 1  
Chapter 2 ................................................................................................................................... 4  
  2 Background and Context ...................................................................................................... 4  
    2.1 Introduction ....................................................................................................................... 4  
    2.2 Metadata: Definitions and Use in National Security ....................................................... 4  
    2.3 Issues in Citizen Privacy and National Security .............................................................. 5  
      2.3.1 Oversight and Transparency of National Security ..................................................... 5  
      2.3.2 The Privacy-Security Trade-Off ................................................................................ 8  
      2.3.3 Concerns about the Warrantless Collection of Citizen Metadata for National Security ................................................................................................................................. 9  
      2.3.4 Importance of Communication around these Issues ................................................. 10  
    2.4 The Role of the Press in a Democratic Society ............................................................... 10  
    2.5 Research Questions ......................................................................................................... 12  
Chapter 3 ................................................................................................................................... 13  
  3 Methods .............................................................................................................................. 13  
    3.1 Sample selection ............................................................................................................. 13  
      3.1.1 Newspaper Readership and Demographics ............................................................... 13  
      3.1.2 How the Articles were sourced ................................................................................. 15  
    3.2 Data Analysis .................................................................................................................. 17
3.2.1 Chi-Square Test ................................................................. 21

3.3 Structure of the thesis .......................................................... 22

Chapter 4 ..................................................................................... 23

4 Findings ..................................................................................... 23

4.1 Overall Description of the Articles ......................................... 23

4.2 Metadata Definitions ............................................................. 26

4.2.1 How the Articles Define Metadata ....................................... 26

4.2.2 Definitions: What metadata IS; what metadata is NOT .......... 27

4.2.3 Evaluation: Metadata as Unimportant; Metadata as Meaningful ...... 34

4.2.4 How the OPC and CSEC Define Metadata ......................... 39

4.3 The Topics Observed in the Introductory Paragraphs of the Articles .......... 48

4.3.1 Transparency/Oversight ....................................................... 50

4.3.2 Scary/About You/ Spying ..................................................... 53

4.3.3 Breaking the Law ................................................................. 56

4.3.4 Terrorists/Terrorism ............................................................ 58

4.3.5 Other ................................................................................. 60

4.4 How the Topics Develop Over Time ....................................... 61

4.4.1 Lede Paragraphs Published by Year and by Topic Combining Years 2014-2015 ................................................................. 61

4.4.2 Topics by Publication (Percent per Publication) ................... 63

5 Discussion and Conclusion ....................................................... 66

6 Bibliography .............................................................................. 71

Curriculum Vitae ........................................................................ 89
List of Tables

Table 1 Initial Sample of Articles.......................................................... 16

Table 2 The Final Sample Size of the Articles Included in the Study....................... 17
List of Figures

Figure 1 Percentage of Articles Published by Publication.................................................. 23

Figure 2 Percentage of Articles Published by Year ............................................................... 23

Figure 3 Articles by Publication and Year (Percent Per year Per Publication) .................. 24

Figure 4 Percentage of Articles Per Publication Per Year, 2014 and 2015 Combined.......... 26

Figure 5 How the Two Definitional Categories are Represented from 2013-2016 by the Percentage of Articles per Year .................................................................................. 33

Figure 6 How the Two Evaluative Categories are Represented from 2013-2016 by the Percentage of Articles per Year .................................................................................. 38

Figure 7 The Percentage of Articles per Publication that Include a Definition of Metadata as Important .................................................................................................................. 39

Figure 8 The Percentage of All of the Articles that Directly Reference CSEC vs. Percentage of All the Articles that Directly Reference the OPC ......................................................... 42

Figure 9 The Percentage of All Articles that Include Elements of Only the CSEC Definition of Metadata, Only the OPC Definition, Both Definitions, or Neither Organization’s Definition of Metadata .................................................................. 43

Figure 10 The Percentage of Articles per Publication that Contain Statements Which Include Elements of Only the OPC Definition of Metadata, Only the CSEC definition, Both Definitions, or Neither Organization’s Definition of Metadata .................................................................. 44

Figure 11 The Percentage of Articles per Publication that Contain Passages which Directly Reference and Attribute Only the OPC Definition, Only the CSEC Definition, Both Definitions, or Neither Organization’s Definition of Metadata .................................................................. 46

Figure 12 The Percentage of Articles per Year that Include Passages which Directly Reference Only the OPC Definition, Only the CSEC Definition, Both Definitions, or Neither Organization’s Definition of Metadata .................................................................. 47
Figure 13 The Percentage of All Lede Paragraphs from Each of the Categories .................. 49

Figure 14 Percentage of Lede Paragraphs Published Per Year for Each Category Combining Years 2014-2015................................................................. 62

Figure 15 Topic by Publication (Percent per Publication)............................................. 63
List of Appendices

Appendix A: Newspaper Articles. .................................................................................................. 79
Chapter 1

1 Introduction

In order for national security agencies to conduct surveillance on Canadian citizens, they must normally obtain a legal warrant that is issued by a judge. However, in certain instances, national security agencies conduct surveillance without obtaining a warrant. This form of surveillance raises concerns because it removes the critical component of judicial or court ordered oversight. Furthermore, this type of surveillance is worrisome because of the large scope of contemporary electronic surveillance technologies.

Warrantless electronic surveillance is enabled by the collection and analysis of metadata. Metadata is the contextual information that surrounds an electronic message, the so-called ‘envelope information,’ and it can include phone numbers, the length of a phone call, IP addresses, email addresses, as well as other electronic information (“Metadata and Privacy”). When metadata is collected in large quantities, it can reveal a startlingly accurate depiction of an individual’s social relationships and behaviours (“Metadata and Privacy” 7).

The issue of warrantless electronic surveillance, specifically metadata collection, by the Canadian government has risen to the awareness of the Canadian public since 2013, when the former National Security Agency (NSA) contractor Edward Snowden leaked top-secret documents to journalist Glenn Greenwald concerning the warrantless surveillance related activities of America’s NSA (Lyon, “Surveillance, Snowden” 3). The Snowden disclosures alarmed Canadians because they also revealed that the Communications Security Establishment Canada (CSEC) was conducting warrantless bulk metadata collection on Canadian citizens as well as international travelers by setting up a surveillance mechanism that would allow the agency to track and locate travellers’ cell phones once they had signed into the free Wi-Fi at Toronto Pearson International Airport (“CSEC used Airport”; “Spy Agencies”).
The NSA leaks sparked a lively debate in the Canadian print news media regarding warrantless electronic surveillance and the significance of metadata (Berthiaume; Deibert; Mitrovica). Canadian Members of Parliament, the Privacy Commissioner of Canada, as well as journalists from the Canadian print news media began questioning CSEC’s surveillance activities ("Canada Needs Spying Debate"; Freeze, “How CSEC Became”).

In 2015, Snowden released disclosures about CSEC’s individual surveillance operations (Brean; Parsons; Freeze, “Canadian spy program”). In 2016, Canada’s domestic surveillance agency CSIS was found to have broken the law by retaining metadata on Canadians that were unrelated to national security threats (Freeze, “CSIS claims transparency”; Boutilier). CSIS failed to inform the Canadian courts that it was indefinitely retaining the metadata that it had collected from previously court-authorized surveillance operations (Freeze, “CSIS claims transparency”; Boutilier). In the same year, CSEC was found to have illegally shared Canadian metadata to members of the international intelligence community since 2013 (Freeze, “Privacy watchdog urges”). CSEC failed to remove information that could be used to pinpoint Canadians before sharing the metadata (Freeze, “Privacy watchdog urges”).

Groups fighting for citizen privacy and those who represent the national security agenda propose conflicting perspectives on whether citizens should be concerned about the government’s warrantless collection of metadata (Mitrovica; Freeze, “Segal says”). National security advocates suggest that metadata is required because it helps to identify security threats (Freeze, “MacKay Approved”). Privacy advocates argue that government metadata collection is an invasion into the private lives of citizens (Cavoukian and Levin).

News publications offer information to citizens about this important and controversial debate. This issue is complex and it is important to analyze how it has been represented in the Canadian print news media so that we can better understand what Canadians have been told about Canada’s collection of metadata.
This thesis analyzes the complex and at times contradictory conversation that plays out in the Canadian print news media regarding the representation of CSEC’s warrantless collection and analysis of citizen metadata. This study performs a qualitative and quantitative content analysis of articles that discuss the Canadian government’s warrantless domestic collection of metadata published in the *National Post*, the *Toronto Star*, and the *Globe and Mail* from 2013 to 2016.
Chapter 2

2 Background and Context

2.1 Introduction

This chapter introduces three important contextual considerations: the meaning of metadata as well as the type of information that it can reveal when it is used to conduct warrantless surveillance by national security agencies; issues of concern about the warrantless collection of citizen metadata for national security purposes; and the important role of the press in representing these issues to Canadian citizens.

2.2 Metadata: Definitions and Use in National Security

Electronic surveillance is assisted by the mass collection and analysis of metadata, which is a part of what national security agencies refer to as signals intelligence (SIGINT) (Rudner, “Signals Intelligence” 473). Signals intelligence can be defined as the collection of electronic communications through the use of “sophisticated, covert interception technologies capable of monitoring terrestrial, microwave, radio, Internet and satellite communications along with other electromagnetic emissions” (Rudner, “Signals Intelligence” 474).

The Office of the Privacy Commissioner of Canada (OPC) defines metadata quite broadly as “data that provides information about other data. It is information that is generated as you use technology, and lets you know the who, what, where, when, and how of a variety of activities” (Metadata and Privacy 1). The OPC additionally states that:

In the communications context, metadata provides certain details about the creation, transmission and distribution of a message. As such, metadata can, for example, include the date and time a phone call is made or the location from which an e-mail was accessed. (Metadata and Privacy 1)
Metadata is produced as a result of individuals communicating over electronic devices (Metadata and Privacy 3). These devices can include “landline telephones, mobile phones, desktop computers, laptops, tablets or other computing devices” (Forcese 129).

Metadata is often compared to the information on the outside of an envelope. At first blush, metadata collection may appear to reveal only a small amount of detail about individuals since metadata does not record the content of an electronic communication. However, the OPC states that the information on the outside of an envelope can often provide meaningful insight about an envelope’s content (Metadata and Privacy 3). When national security agencies collect metadata in bulk quantities, a vivid and striking picture emerges that demonstrates a startlingly accurate depiction of an individual’s habits, interests, and social connections which can be used to predict an individual’s future actions, whereabouts, and relationships with others (Metadata and Privacy 7). Metadata “can sometimes be more revealing than [a message’s] content” because it allows an agency to make inferences and predictions about how groups of individuals may behave in the future (Metadata and Privacy 1-4).

2.3 Issues in Citizen Privacy and National Security

Three issues are important for this thesis because they are discussed in the Canadian print news media: the oversight and transparency of the national security management, the trade-off between citizen privacy and national security, and concerns about the warrantless collection of citizen metadata for national security.

2.3.1 Oversight and Transparency of National Security

The two Canadian agencies that are responsible for investigating national security threats are the Canadian Security Intelligence Service (CSIS) and the Communications Security Establishment Canada (CSEC). CSIS is Canada’s national domestic surveillance agency that has a rigorous oversight mechanism which requires the security agency to obtain a warrant whenever intrusive data collection occurs from Canadians (“Intelligence Collection”). CSEC proposes the following information about its role in Canada’s national security framework:
CSE is Canada's national cryptologic agency. Unique within Canada's security and intelligence community, CSE employs code-makers and code-breakers to provide the Government of Canada with information technology security (IT Security) and foreign signals intelligence (SIGINT) services. CSE also provides technical and operational assistance to federal law enforcement and security agencies. (“About us”)

CSEC is a member of the international Five Eyes intelligence network, which also includes security agencies from the U.S., the U.K., Australia, and New Zealand (“Partnerships”). In 2016 the Globe and Mail argued that CSEC has been illegally sharing metadata about Canadian citizens with the Five Eyes Intelligence network “for years” (Freeze, "Spy Agency”).

SIGINT technologies allow the Five Eyes network to monitor, store, and share the electronic communications of suspected terrorists and other individuals who are a threat to national security. Rudner states that the Five Eyes network allows these security agencies to pool their intelligence resources together so that collectively there is a “near-global SIGINT capability to collect and deliver real-time communications intelligence on foreign targets” for national security purposes (“Canada’s Communications Security” 479). In addition, Rudner states that as part of this Five Eyes partnership the allied security agencies are not allowed to “target one another or their respective nationals” (“Canada’s Communications Security” 479).

Walby and Anaïs reference the Auditor-General Sheila Fraser’s 2004 report on Canadian intelligence agencies, which criticizes both the lack of information and detail in the CSE commissioner’s reports and CSEC’s broad scope, resulting in the Commissioner only being able to review a small portion of CSEC’s operations (369).

National security measures require oversight and regulation in order to effectively assess the extent to which the privacy rights of citizens may be infringed upon (Solove 36-37). Historically, Canada’s national security agencies have had limited parliamentary oversight, and it was only in 1996 that the government appointed a CSE Commissioner, “a judicial office with a mandate to review and report upon the agency’s activities with
respect to compliance with the law” (Rudner, “Signals Intelligence” 485). The appointment of a CSEC commissioner in 1996 appears to demonstrate some form of accountability for the security agency yet problems persist regarding effective government oversight (Rudner, “Signals Intelligence” 485; Walby, Anais 369). Since 9/11 the heightened attention to national security has often favoured the implementation of national security measures that have limited oversight and regulation. Solove suggests that oversight and regulation do not have to come at the cost of weakening national security measures (Solove 2). Without effective oversight, there is no guarantee that the privacy rights of citizens are being protected from potential government overreaches. Government overreaches on citizen privacy can occur due to new developments in intrusive electronic national security measures.

In addition to limited parliamentary oversight, Canada’s national security agencies have a history of secrecy. CSEC in particular has been shrouded by government secrecy for many years, and its existence was only revealed in 2001 because of the passage of the “Anti-Terrorism Act” that would demarcate the agency’s official mandate (Rudner, “From Cold War” 474-475). Rudner states that for decades prior to 2001 CSEC operated without a legislative mandate, and was “arguably, also the most secretive component of the Government of Canada” (“From Cold War” 475; “Signals Intelligence” 97). Due to the historically secretive nature of CSEC, there is growing concern among citizens and privacy advocates that the agency is jeopardizing citizen privacy rights in the name of protecting national security (Rudner, “Signals Intelligence” 487).

Since 2005, concerns have been raised by the former privacy commissioner of Canada, Jennifer Stoddart, regarding the sweeping powers that are granted to “agencies involved in national security” as a result of the passing of Canadian counter-terrorism legislation (Couturier). In a 2005 news release issued by the Office of the Privacy Commissioner of Canada, Stoddart calls for “greater accountability, transparency and oversight” for the Canadian agencies that protect against national security threats (Caidi, Ross 667; Couturier). The current Privacy Commissioner of Canada Daniel Therrien recommends “expert, independent oversight” instead of an oversight body that is connected to the national security establishment “so that rights are effectively protected” (“Privacy and
Canada’s). The case that Canada’s national security agencies require increased powers is often presented by comparing individual privacy with national security concerns.

2.3.2 The Privacy-Security Trade-Off

In the wake of the September 11th terrorist attacks, national security agencies have frequently positioned increases in electronic surveillance as being necessary or required in order to protect citizens from the new looming terrorist threat. This new political climate that favours the implementation of increasingly more intrusive national security measures situates citizen privacy as being less important than national security. The national security perspective frequently downplays the privacy implications associated with CSEC’s metadata collection program.

The urgent need to prevent future terrorist attacks has created what David Lyon refers to as a “panic regime” (Surveillance After 35). Within this regime, citizens have become more willing to give up their civil liberties - such as their right to privacy - in exchange for the perceived protections that are enabled as a result of increases in electronic national security measures (Surveillance After 35). This trade-off suggests that in order for national security to be preserved citizens must expect a decrease in their right to privacy. Schneier (“Data and Goliath” 156) argues, “when the security versus privacy trade-off is framed as a life-and-death choice, all rational debate ends” (“Data and Goliath” 156).

Different stakeholders take very different positions on this issue. There is growing concern among academics and privacy advocates that increases in national security impose unnecessary restrictions on the civil liberties and privacy rights of everyday citizens (Caidi, Ross 663-664). Privacy advocates dismiss the idea proposed by CSEC that its mass collection of metadata does not infringe on the privacy rights of everyday citizens. In many cases, a variety of methods exist that allow citizen privacy rights to be protected while at the same time limiting the extent to which national security is compromised such as more rigorous independent oversight as well as increased parliamentary oversight (Schneier “Data and Goliath” 156). Government agencies responsible for protecting the privacy rights of citizens include the federal Office of the Privacy Commissioner of Canada (OPC) as well as the provincial Information and
Privacy Commissioner of Ontario; each Canadian province has a provincial Information and Privacy Commissioner as well.

On December 6th, 2016, the Privacy Commissioner of Canada Daniel Therrien and his provincial and territorial counterparts published a review of Canada’s national security framework (“Privacy and Canada’s”). In this review, when focusing on metadata and national security they argue “that the National Defence Act be amended to clarify that the CSE’s powers with respect to the collection, use and disclosure of personal information be accompanied by specific legal safeguards to protect the privacy of Canadians” (“Privacy and Canada’s”).

2.3.3 Concerns about the Warrantless Collection of Citizen Metadata for National Security

CSEC states that signals intelligence is normally directed at foreign entities and it involves “targeting and intercepting foreign communications, decrypting or decoding them, and analyzing their content to see what they reveal” (“Foreign Signals”). Canadian government officials who defend CSEC’s collection of metadata frequently repeat the response that the agency does not collect the content of citizen communications (“How does CSE”).

Critics argue that the warrantless mass collection of citizen metadata by government security agencies like CSEC results in severe intrusions into the privacy rights of everyday citizens (Bauman, et al. 127). Citizens may have no involvement in terrorist activities yet they can still be put under surveillance if they are socially associated with an individual who has connections with a terrorist network (84). The fact that mass metadata collection amounts to mass surveillance suggests that all citizens are considered to be potential threats to national security by government agencies like CSEC. This raises important questions regarding the right that citizens have to privacy in their use of electronic communication devices that are often critical in many forms of contemporary communication.
2.3.4 Importance of Communication around these Issues

Canadians across the nation should engage in a discussion that focuses on the meaning behind the electronic information that the government is collecting. In order to effectively participate in this discussion Canadians must consider the privacy-security trade off, effective oversight for our national security agencies, the transparency between CSEC and the public, and the meaning of metadata. In order for Canadians to become informed participants in this debate they need to understand how these critical issues have been represented in the Canadian press.

2.4 The Role of the Press in a Democratic Society

In a liberal democratic country like Canada, the press plays the fundamental role of communicating to citizens the information that is considered to be “socially important at any given time” (McNair 29). Without the press, citizens within a democratic nation would face difficulties in obtaining the latest information about how government officials and politicians are managing sensitive issues such as privacy rights. The press empowers a country’s citizens by enabling them to make more informed decisions regarding whether to support or protest against the actions of government representatives (McNair 1).

Since the 19th century, the news media have established the need for the freedom of the press from the influences of those who hold positions of power in society in order to ensure that the public can develop their own critical and objective stance on current events (Schultz 24). Freedom of the press helps to ensure that the positions discussed in the news media are not simply representative of the views of society’s elites such as those who work for the government (Hackett and Zhao 180-181).

The press therefore acts as a mediator between “political actors and the public” (McNair 105). The Canadian print news media are expected to critique the actions of government officials in order to ensure that there is a mechanism built into society that acts as a form of accountability for “those in positions of political, corporate, economic and social power” (Schultz 1). This function of the press as a watchdog is often adversarial to the
positions of government actors in conducting a set of checks and balances in order to weigh whether government officials are acting in the best interest of a country’s citizens (Hackett and Zhao 140; Schultz 2-3). Schultz states that “the process of finding, distilling, and analyzing the information that is the media’s commodity also ensures its political role, the core of its self-definition as the fourth estate” (2).

Scheufele and Tewskbury identify three ways in which news media can influence a society. First, the way an issue is framed in articles may influence how that issue is understood by audiences (Scheufele and Tewskbury 11). Second, a central function of the news media is to set the agenda: determining the main stories that the public should be informed about (McNair 29), and influencing the salience of topics “so that an issue becomes the focus of public attention, thought, and perhaps even action by the news media” (Bryant and Oliver, 1). A study by McCombs and Shaw suggests that the news media play a critical role in influencing how the general public perceives the importance of disparate issues. Third, the news media may prime an audience, suggesting “that they ought to use specific issues as benchmarks for evaluating the performance of leaders and governments” (Scheufele and Tewskbury, 11), for example the effectiveness of presidential candidates in achieving their goals (Iyengar and Kinder, 63).

In 2017, with the development of increasingly more interconnected networks of communication these three characteristics of framing, agenda setting, and priming can now occur at an accelerated speed. With the emergence of the 24 hour news cycle and the sharing of articles on popular social networking sites like Facebook and Twitter the media can now frame articles, set the agenda, and prime audiences at a much greater speed.

Newspaper articles are a gateway into understanding how the privacy implications that surround CSEC’s metadata collection program have been represented to Canadians across the nation. The 1981 Royal Commission on Newspapers reiterates the importance of the news media by stating “major daily newspapers … remain the primary source of public affairs information not only for the top decision-makers but also for the most politically attentive segment of the population at all levels” (Kent 137). By investigating
some of the largest Canadian daily newspapers this project will demonstrate how the press represents the privacy implications that surround CSEC’s collection and analysis of metadata.

2.5 Research Questions

The main research questions that this project aims to investigate is as follows:

1. *How has the complex issue of the Canadian government’s warrantless domestic collection of citizen metadata been represented to the general public in the National Post, the Toronto Star, and the Globe and Mail since the 2013 Snowden disclosures?*

2. *How has metadata been defined in the National Post, the Toronto Star, and the Globe and Mail?*

3. *What topics are observed in the lede paragraphs of the articles?*

4. *What is the relationship between the topic, publication, and year that the articles were published?*
Chapter 3

Methods

This chapter provides an overview of the selection of the sample of newspaper articles to be analyzed and describes the data analysis. The chapter concludes with the outline for the thesis.

3.1 Sample selection

3.1.1 Newspaper Readership and Demographics

The three Canadian newspapers that this project analyzes are the Toronto Star, the Globe and Mail, and the National Post. The Globe and Mail and the National Post were selected because they are the two most read national Canadian newspapers and the Toronto Star was chosen because it is the most read Canadian daily newspaper (“Newspaper Topline”).

Both the Globe and Mail and the National Post are considered to be national newspapers in Canada. Both publications release digital as well as print editions daily. A recent study demonstrates that across all four quarters of 2016 the Globe and Mail surpassed the National Post in readership with a combined digital and print audience of 2.23 million readers for an average weekday issue (Rody-Mantha). The National Post received 1.52 million readers in comparison within the same period (Rody-Mantha). These numbers suggest that the Globe and Mail has 31.84% more readers than the National Post throughout the weekdays. This data was collected from Canadians from the ages of 18 and up (Rody-Mantha).

The same study demonstrates that the Toronto Star is the most read newspaper in both Ontario and Toronto. In Ontario, the Toronto Star accumulated a readership of 1.91 million readers in 2016 when both print and digital readerships are combined (Rody-Mantha). The second Ontario newspaper that trails the Toronto Star’s readership is the Toronto Sun which only amassed a grand total of 839,000 readers in 2016 (Rody-Mantha). In the city of Toronto, the Toronto Star saw readership numbers of 1.27 million
readers when print and digital readerships are combined (Rody-Mantha). The second
most read newspaper in Toronto is the Metro but it only saw readership numbers of
725,000 (Rody-Mantha). The numbers listed demonstrate the 2016 readership of an
average weekday issue of these papers for individuals aged 18 and over (Rody-Mantha).

The three publications are separated by their political orientations. The National Post’s
political leaning is conservative (“World Newspapers”). The Toronto Star is more
liberally oriented as a publication (“World Newspapers”). Lastly, the Globe and Mail is
listed as a liberal publication (“National Canadian Newspapers and News Sites”).

According to the Globe and Mail’s 2017 newspaper media kit, the demographics of its
average weekday readership is quite diverse (“Globe Newspaper”). The media kit states
that 35% of its readership is under 34, 19% between the ages of 35 and 49, 25% between
50 and 64, and lastly 21% over 65 (“Globe Newspaper”). The Globe and Mail’s
readership is also broken down according to the average income of its readers. The
majority of the readers of this publication (65%) make an income of less than $100,000
(“Globe Newspaper”). The second largest segment of readers earn $100,000 to $200,000
annually and 25% of the publication’s readers make up this grouping (“Globe
Newspaper”). Lastly only 9% of the Globe and Mail’s readers make over $200,000
annually (“Globe Newspaper”).

The Toronto Star’s 2017 media kit demonstrates the publication’s demographics by
focusing on the print and digital readership that was seen in the second quarter of 2016.
This media kit relies on research conducted by Vividata as well (“Toronto Star”). Within
this time period, when the print and digital readerships are combined, the Toronto Star’s
weekly audience consisted of 2,337,000 baby boomers (born 1945-65, “Glossary of
terms”), 2,558,000 readers between the ages of 24-54, 2,227,000 multicultural individuals
(people who were not born in Canada who also speak a foreign language), 978,000
mothers, 1,128,000 individuals who earn a household income of $125,000 and up, and
1,632,000 millennials (born between the beginning of the 1980s and the beginning of the
2000s, “Glossary of Terms”; “Toronto Star”). The media kit also claims that 30% of the
publication’s readers “were born outside of Canada”, 33% speak another language other
than English, mothers represent 19% of the publication’s readers, and 22% of the publication’s readers are “affluent Canadians,” with a household income of $125,000 or more (“Toronto Star”).

The National Post’s 2017 media kit quotes Vividata readership data from the fourth quarter of 2015 (“National Reader”). The data from this report suggests the “total weekly footprint” of the paper which reaches a variety of different audiences (“National Reader”). For adults aged 18 and over, this publication has a weekly following of 4,796,000 individuals (“National Reader”). In 2015, the National Post saw the following weekly readership numbers: the 18-24 age bracket consisted of 743,000 readers, the 25-34 range included 1,045,000 individuals, 35-49 was the largest segment of readers at 1,197,000, the 50-64 bracket was the second largest grouping at 1,123,000, and the seniors’ bracket of people aged 65 and over consisted of 688,000 readers (“National Reader”). The 18-24 age range represented 15.5% of the National Post’s weekly readers, the 25-34 segment 21.8%, 35-49 the largest grouping represented 25%, 50-64 was slightly lower with 23.4%, and people 65 and over represented only 14.4% of the weekly readership for this paper (“National Reader”).

3.1.2 How the Articles were sourced

The sample of articles was sourced from the Canadian Major Dailies online database. Articles were sorted by date, month, and year in ascending order (“Canadian Major”), and according to the publication. The articles were selected from January 2013 to December 2016 in order to analyze how the conversation on the Canadian government’s collection of metadata has developed in the Canadian print news media since the 2013 Snowden disclosures.

The first step in the data collection process was identifying articles that included the term “metadata” in the full text and then reviewing the articles to make sure that they were relevant to this study. The total initial sample size included 215 articles.

The second step in the data collection process was reading each article to remove any articles that did not reference the Canadian government’s collection of metadata.
articles only described the NSA’s activities in large-scale government metadata collection and did not describe Canadian metadata collection. 3 articles touched on the role of CSIS but did not specifically speak about CSEC, the agency responsible for Canadian signals intelligence. Finally, 6 duplicate articles were removed. This process resulted in a sample of 130 unique articles.

Table 1 Initial Sample of Articles

<table>
<thead>
<tr>
<th>News Publication</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>The National Post</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>The Globe and Mail</td>
<td>14</td>
<td>20</td>
<td>6</td>
<td>11</td>
<td>51</td>
</tr>
<tr>
<td>The Toronto Star</td>
<td>20</td>
<td>12</td>
<td>4</td>
<td>21</td>
<td>57</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>37</td>
<td>13</td>
<td>36</td>
<td>130</td>
</tr>
</tbody>
</table>

The sample size was reduced further during the data analysis process to focus the scope of the project. Articles were removed if they primarily discussed Canadian bills that were proposed or implemented from 2013-2016 concerning the government’s collection of metadata. These articles were removed because this project does not analyze how the legal environment in Canada has developed from 2013-2016 regarding the government’s collection of metadata. Instead, the project focuses primarily on how the Canadian government’s collection of metadata has been represented to the Canadian public in the Canadian print news media since the 2013 Snowden disclosures.

Articles were also removed from the sample if they focused primarily on CSEC collecting metadata and conducting surveillance operations on foreign countries. A small number of articles discussed CSEC being caught spying on Brazil and these articles were removed from the sample. The articles were removed because they are out of the scope of this project which deals primarily with how the Canadian government’s domestic collection of metadata has been represented to the Canadian public.

Articles were also removed if they only discussed telecoms or police services collecting metadata. These articles were removed to focus the scope of the project which primarily
analyzes how CSEC’s collection of metadata has been represented to the public in the Canadian print news media. The final sample size includes 91 articles in total which consist of articles, columns by in-house journalists, editorials, and opinion pieces.

Table 2 The Final Sample Size of the Articles Included in the Study

<table>
<thead>
<tr>
<th>News Publication</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>The National Post</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>The Globe and Mail</td>
<td>11</td>
<td>13</td>
<td>4</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>The Toronto Star</td>
<td>11</td>
<td>6</td>
<td>2</td>
<td>17</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>23</td>
<td>8</td>
<td>30</td>
<td>91</td>
</tr>
</tbody>
</table>

3.2 Data Analysis

This study first relied on conventional qualitative content analysis (Hsieh and Shannon 1279) as a research method in order to shed light on how the National Post, the Toronto Star, and the Globe and Mail represent the Canadian government’s warrantless collection of metadata since the Snowden 2013 disclosures. The qualitative data analysis software HyperRESEARCH was used to organize the data set, code the articles for revealing information, build reports that display the coded data, as well as to efficiently determine overarching themes (“Qualitative analysis”). This conventional qualitative content analysis analyzed the headline as well as the lede paragraph of the articles to search for overarching themes. All sections of the articles (the headline, lede paragraph, and body) were analyzed when researching how the articles define metadata.

Hsieh and Shannon state that “conventional content analysis … is usually appropriate when existing theory or research literature on a phenomenon is limited” (1279). Since the 2013 Snowden disclosures, little scholarly research has been written on the degree to which CSEC may infringe upon the privacy rights of citizens by conducting mass
taxpayer funded metadata collection. Furthermore, no studies investigate the debates around these issues in the Canadian print news media.

This qualitative content analysis relied on me selecting themes that emerged from the data. Themes were not selected by any preconceived notions that I may have developed prior to undertaking the study (Hsieh and Shannon 1279). Hsieh and Shannon state that in “conventional content analysis … researchers avoid using preconceived categories, instead allowing the categories and names for categories to flow from the data” (1279). The process of themes materializing from reading the data is referred to as “an inductive approach” to theme selection according to Ryan and Bernard (89).

I carefully read all of the articles to identify initial themes and patterns. These were coded both on paper and in HyperRESEARCH and discussed with the co-supervisors. The second stage of analysis was a more focused analysis of how the articles defined metadata within the body of the text. The articles were carefully read in HyperRESEARCH to code for any sentences or paragraphs that either defined metadata or provided an explanation of what metadata is, and then analyzed how these definitions compared to standard Canadian definitions (e.g., from CSEC and the OPS). The third stage of analysis returned to the first paragraphs of the articles to describe the themes readers would encounter first in an article.

One of the techniques that was used to identify themes was selecting ideas or concepts that were repeated numerous times throughout the data (Ryan and Bernard 89). In an attempt to spot recurring themes, I specifically focused on identifying and recording recurring “conversation topics, vocabulary, recurring activities, meanings, feelings, or folk sayings and proverbs” which could signal the existence of a topic that was represented numerous times across the data set (Taylor et al. 171-172). Ryan and Bernard state, “The more the same concept occurs in a text, the more likely it is a theme” (89). Metaphors were also coded if they were used to define or provide examples of metadata.

I first looked for recurring concepts that were represented in each individual publication. I then searched for the recurring concepts that were observed across all of the publications so that overarching themes could be found (Taylor et al. 171).
While reading the data on a publication-by-publication basis to uncover recurring themes I kept a log for recording “hunches, interpretations, and ideas” (Taylor et al. 171). Taylor et al. state that active recording is important when “developing themes and concepts” because it allows researchers to, later, look back on the ways in which the data left notable impressions (Taylor et al. 171). It was important to keep a detailed record of the salient issues, concepts, or ideas that emerged from a close reading of the data set (Taylor et al. 171).

By noting the similarities and differences between each group, I was able to select additional themes and subthemes that emerged from the data (Ryan and Bernard 91).

Prior to the process of interpretation and analysis I returned to the data set to categorize the statements, sentences, or paragraphs that demonstrated the overarching themes that were previously observed across all of the news publications. Each statement, sentence, or paragraph was organized according to the theme that the item represented. The introductory or lede paragraphs of the articles were also thematically categorized. This process of categorization further narrowed down the scope of the material being investigated in this research project.

Once the process of categorization had been completed I performed a close reading of the coded text organized according to their thematic groups. This focused close reading of the data was performed to tease out any new inferences that may not have been previously apparent. Close readings of the coded text occurred numerous times throughout the data analysis process. This process of revisiting the data in an attempt to discover additional insights coincides with the “interpretive research tradition” of qualitative research (Krippendorff 88). Krippendorff states that “acknowledging the holistic qualities of texts, these scholars feel justified in going back and revising earlier interpretations in light of later readings; they settle for nothing less than interpretations that do justice to a whole body of texts” (88). Any new insights or impressions were also recorded in order to assist in the analysis of the data.

Once the themes, subthemes, and categories had been selected and tagged in HyperRESEARCH, I then began the process of recording the data. The following themes
were found regarding how the articles define metadata: metadata defined by what it is, by what it is not, metadata as unimportant, and metadata as important. The themes observed in the lede paragraphs of the articles include: transparency/oversight, scary/about you/spying, breaking the law, terrorists/terrorism, and other. Krippendorff describes the recording process as “observers, readers, or analysts interpret[ing] what they see, read, or find and then stat[ing] their experiences in the formal terms of an analysis” (126). Elo and Kyngäs state that “there are no systematic rules for analysing data; the key feature of all content analysis is that the many words of the text are classified into much smaller content categories” (109).

At this point, I returned to the data set to begin the process of “[searching] for multiple interpretations by considering diverse voices (readers), alternative perspectives (from different ideological positions), oppositional readings (critiques), or varied uses of the texts examined (by different groups)” (Krippendorff 88). Unlike the more quantitative approaches to content analysis I first applied a predominantly qualitative process to data analysis to better understand the broader context of how CSEC’s collection of metadata had been represented in the Canadian print news media (Krippendorff 88). After the data had been analyzed through a qualitative lens, I then moved on to a quantitative approach to data analysis.

Quantitative content analysis was also employed to analyze the data. All of the articles included in the sample set were analyzed using HyperRESEARCH to count the number of times that specific themes and definitions of metadata were represented in the articles. Krippendorff states that “there is no point in counting unless the frequencies lead to inferences about the conditions surrounding what is counted” (28). Counting the frequency in which different themes and different definitions of metadata were represented in the articles demonstrated if there was an overrepresentation or an underrepresentation of different viewpoints. This quantitative information was presented using bar graphs to clearly demonstrate how each position had been represented in the Canadian print news media. Bar graphs were also used to demonstrate how the themes were represented within each of the three publications as well as to demonstrate how the themes were represented on an annual basis from 2013-2016. The findings reveal how the
Canadian government’s collection of metadata has been represented to the Canadian public in the *National Post*, the *Toronto Star*, and the *Globe and Mail* since the 2013 Snowden disclosures.

### 3.2.1 Chi-Square Test

Chi-square test was used to analyze the relationship between the different quantitative variables that were observed in the study. Vaughan describes the chi-square test as “an inferential statistical test that is used to examine relationships between two variables with nominal or ordinal data” (Vaughan 75). The test is also used to demonstrate the probability in which conclusions are correct “allowing us to state how certain we are of our conclusion[s]” (Vaughan 77). Vaughan explains that:

> The larger the chi-square score, the larger the discrepancy, and the more likely it is that the two variables being studied are related. Recall that the null hypothesis assumes no relationship between the two variables. Therefore, the larger the chi-square score, the smaller the probability for the null hypothesis to be true. When the probability (the p-value) is equal to or smaller than a pre-set value, usually 0.05, we will reject the null hypothesis and conclude that the relationship between the two variables is statistically significant; i.e., there is a real relationship in the population, not just a chance relationship in the sample. (81)

In addition, she states “what we really need to know is whether or not the p-value associated with this chi-square score is less than the pre-set level, usually 0.05” (81).

The test can vary based on the degrees of freedom of “a contingency table” (Vaughan 77-82). A contingency table is simply “the cross tabulation of two variables” (Vaughan 77). Vaughan describes degrees of freedom as “a statistical term associated with every statistical test … in the case of the chi-square test, it means the number of cells whose cell frequencies are free to change once the row marginal totals and the column marginal totals are fixed” (82). Calculating the degrees of freedom of a contingency table allowed me to accurately measure the critical values of chi-square linked to the significance level of 0.05 or 0.01 (Vaughan 83). Vaughan states “usually we use a significance level of...”
0.05; i.e., we will reject the null hypothesis if the probability for it to be true is equal to or less than 0.05” (83).

3.3 Structure of the thesis

Chapter 4: This chapter will discuss the findings from both the qualitative as well as the quantitative data analysis. The chapter will begin by discussing the overall description of the articles. The number of articles from each publication that focus on this issue will be reported and discussed, along with the number of articles published each year from 2013-2016.

Both the qualitative and quantitative results concerning how the articles define metadata in the body of the text will then be reported. The definitions have been sorted into categories in order to demonstrate the similarities or differences between how metadata is defined in the articles. The percentage of articles per year, and per publication that define metadata according to the different categories will be discussed as well.

After the articles have been introduced, the qualitative findings concerning how CSEC and the OPC define metadata will be reported. Following these findings, the percentage of all the articles that include the OPC definition or the CSEC definition of metadata will be discussed and analyzed. I will also review the percentage of all the articles that directly reference CSEC or the OPC when metadata is defined. The percentage of articles per publication, and the percentage of articles per year will also be analyzed.

The themes observed in the introductory paragraphs of the articles will now be reviewed and analyzed. Quotes from the articles will be used to demonstrate examples of how the themes are represented in the body of the text. The percentage of articles from each thematic category will be presented and discussed. I will show how each theme develops over time as well as how each thematic category is represented in the three publications.

Chapter 5: This chapter will review the significant findings from the data analysis. This chapter will also include the conclusion as well as insights for future research.
Chapter 4

4 Findings

4.1 Overall Description of the Articles

Figure 1 Percentage of Articles Published by Publication

Figure 1 introduces the total percentage of articles per publication which discuss the Canadian government’s domestic collection of metadata. Out of the total sample of 91, the Toronto Star and the Globe and Mail represent an almost equal portion of the articles. The articles from the Globe and Mail represent 41.8% of the total articles. The articles from the Toronto Star amount to 39.6% of the articles, while the articles in the National Post make up only 18.7% of the sample.

Figure 2 Percentage of Articles Published by Year
Figure 2 documents the distribution of the articles across years, reporting the percentage of the 91 articles that discuss this issue from 2013-2016. Several notable differences are observed over this period. In both 2013 and 2016 the highest percentage of articles were published across all publications. These two years correspond with significant events that occurred concerning the Canadian government’s collection of metadata as observed in the introduction to this project. Year 2014 saw a decline in coverage compared to years 2013 and 2016. In 2014, articles discussed CSEC’s collection of metadata from travelers who passed through a large Canadian airport. Compared to the other years, 2015 saw a sharp decline in coverage. Unlike the other years, no critical events were observed across all of the publications concerning the Canadian government’s collection of metadata in 2015.

Figure 3 Articles by Publication and Year (Percent Per year Per Publication)

Figure 3 illustrates the percentage of the articles that each publication published from 2013 to 2016 on the Canadian government’s collection of metadata. Each of the publications shows different trends in reporting as the year’s progress. For the *Globe and Mail* the coverage begins quite high in 2013 with 28.94% of the publication’s articles published in this year. In 2014, the reporting increases slightly to 34.21% of the publication’s articles. Then in 2015 there is a drop in coverage where only 10.52% of *The Globe and Mail’s* articles are published. Finally, in 2016 there is a jump where 26.31% of the publication’s articles are published in that year.
The *Toronto Star* also saw a large amount of coverage in 2013 with 30.55% of its articles being published in that year (see fig. 3). Unlike the *Globe and Mail* there is quite a large dip in coverage for the *Toronto Star* in 2014 where only 16.66% of its total articles are published. A similar trend is seen in 2015 where the *Toronto Star* publishes only 5.55% of its articles in this year. However, in 2016 the *Toronto Star* publishes the highest percentage of its articles in comparison to any other publication for the year. 47.22% of the *Toronto Star*’s total articles were published in 2016.

The *National Post*’s coverage peaks in 2013 and slowly decreases until 2016 (see fig. 3). A small exception to this trend is seen in 2015 where the *National Post* publishes the lowest percentage of its total articles in this year. In 2013, the *National Post* published 47.05% of its total reporting on this issue. Like the *Toronto Star*, in 2014 a dip in coverage is seen by the *National Post* where 25.27% of all of its articles are published in this year. In 2015, the *National Post* published only 11.76% of its articles on this issue. However, this is a higher percentage than any other publication for this year; the *Globe and Mail* published 10.52% of its total articles in 2015 and it was the second highest for the year. Lastly, in 2016 the *National Post* reports the lowest percentage of its articles on this issue by publishing only 17.64% of its articles in this year.

The observable trends across all publications show that a large amount of reporting occurred in 2013 which is when the Snowden revelations were first disclosed (see fig. 3). Two out of the three publications, the *Toronto Star* and the *National Post*, see a dip in reporting in 2014 while the *Globe and Mail* sees a slight increase in articles published for the year. All publications demonstrate a lower amount of coverage in 2015 compared to other years. Then in 2016 both the *Globe and Mail* and the *Toronto Star* see a drastic jump in coverage. The *National Post* also sees an increased in coverage in 2016 but by a smaller amount than the *Toronto Star* or the *Globe and Mail*.

Figure 4 shows that the focus on the issue changed from 2013-2016 across all publications.
Figure 4 Percentage of Articles Per Publication Per Year, 2014 and 2015 Combined.

Figure 4 introduces how the publications focus their reporting when years 2014 and 2015 are combined due to the low percentage of articles that were published in 2015. Only 8.79% of the total articles were published in 2015. Furthermore, these years were combined because in 2015 no major events or disclosures were discussed across all three publications concerning this topic. With the years 2014 and 2015 combined, a chi-square test of the relationship between the year of publishing and the publication makes the case for a marginally significant relationship between these variables ($\chi^2(4) = 7.973$, $p<.1$) (Vaughan 87). This relationship demonstrates that the National Post is most likely to focus its reporting on this issue in the early years of this study (2013-2014), the Globe and Mail in the intermediate years (2014-2015), and the Toronto Star in the last year (2016).

4.2 Metadata Definitions

4.2.1 How the Articles Define Metadata

In order to understand how the Canadian government’s warrantless collection of citizen metadata has been represented in the Canadian print news media, it is important to analyze how metadata has been defined in those same media. Of the 91 articles, 61 articles (67.03% of the total) included at least one definition or explanation of metadata,
and 30 articles did not include any definition of metadata. It is very important to note that a single article could include more than one definition of metadata.

The definitions and explanations of metadata were identified for analysis as samples of text that provide examples of metadata or terms that are used to represent metadata, examples of the information that metadata does and does not record, and examples of metaphors that represent metadata. These were coded to identify thematically related categories using HyperRESEARCH (a qualitative analysis program). All passages in each article which provided a definition or explanation of metadata were identified and analyzed.

Across all of the passages there were 4 distinct categories. Two categories focus on the definition of metadata: first, the type of information that is included in the category; second, the type of information that is not included. The third and fourth categories are evaluative in nature: the third category focuses on the notion of metadata as trivial or unimportant, revealing information only about communications, and not about people, while the fourth focuses on the idea that metadata is important, because it actually reveals a great deal about people.

4.2.2 Definitions: What metadata IS; what metadata is NOT

Statements published in the articles define metadata by what it is and by what it is not. It is important to note that a single article could include multiple statements which define metadata differently. For instance some articles could include passages that define metadata by what it is as well as other passages that define metadata by what it is not. Therefore a single article could include different statements which define metadata in different ways. Inclusive definitions use both general and specific terms. Exclusive definitions compare metadata to the information that it is not.

What Metadata IS:

The first category describes representational patterns that emerge from the data that make claims concerning what metadata is. This kind of definition is present in 61 of the 91 articles (67% of the total, 100% of the articles that defined metadata included statements
that described what metadata is). Often in this category, metadata is described as phone logs or IP addresses. The following quote from the *Globe and Mail* demonstrates this description of metadata. Articles are identified in the text by a reference number: see “Appendix A: Newspaper Articles” for full references.

The crux of the issue is ‘metadata’ - or logs of phone and Internet communications. (33)

This category includes many general and vague statements which described metadata by arguing that it is a record of the communications people share on the Internet: transmission, or traffic data. The following quotes from the *Globe and Mail* define metadata as data that is about transmission patterns as well as about online traffic.

Mr. Forster said airport metadata monitoring was about identifying data transmission patterns and never spying on people. (28)

In June, The *Globe and Mail* unearthed records showing that CSEC has been collecting some Canadian ‘metadata’ - telephone and Internet traffic records - in the course of scouring global telecommunications trails for investigative leads. (34)

At other times metadata is described vaguely, e.g., through statements such as “data that details the circumstances around electronic communications” or as “communications-related information” (Toronto Star 14; National Post 16).

In other statements metadata is described by more specific terms that make claims concerning the electronic information that metadata records. Some examples of this type of information include: the length of phone conversations, the numbers that are dialed from cell phones, as well as the sender and recipient of emails and text messages. The following statement from the *Toronto Star* demonstrates this representation of metadata.

This class of information can include the destination and duration of phone calls, emails, and text messages. (6)
Another common generic description of metadata, which is stated numerous times throughout the articles, is that it is “data about data.” The following quote from the *Globe and Mail* demonstrates this representation of metadata.

‘It is data about data, so it is well within the parameters,’ Mr. Rigby told the committee. Mr. Forster… explained that CSEC wanted to build a picture of the data signatures of public WiFi spots, such as Internet cafes, hotels and airports. (28)

The data signatures term is important because it further disconnects metadata from electronic communications. If a passage claims that metadata only records “data signatures,” this suggests that metadata does not record information that is about communications (Globe and Mail 28).

Lastly, some passages in this grouping define metadata as information that records geographic locations. Some of these definitional passages claim that metadata records the location of individuals and other passages argue that metadata records the location of electronic devices. These passages that describe metadata appear in multiple articles but the following quotes from the *Toronto Star* and the *National Post* make the case that metadata is information that is connected to a person’s location.

The information being tracked - known as ‘metadata’ - includes the duration and location of the calls, the numbers and location of both parties and other identifying information. (30)

The operation involved the processing of at least two weeks of identifying information associated with our mobile devices, their location in time and space, primarily in Canada (beginning at a major international airport). (10)

This category introduced the information that is considered to be metadata according to statements that were made within the articles. Metadata is defined in these statements as a record of phone logs and a record of the online communications that people share over the Internet. Several general and vague descriptions of metadata are introduced as well in statements that were made within the articles, which describe metadata using blanket
terms that are nonspecific such as by arguing that it is information about transmissions or about communications traffic. Metadata is also broadly claimed to be information that is about information. More specific descriptions of metadata are introduced as well to make the case that metadata can record information that is connected to the location of individuals. Other specific descriptions of metadata claim that it records detailed information about telephone, text, and e-mail communications such as: who people are communicating with, how long they communicate for, as well as the phone numbers and emails that are used to communicate with others.

What Metadata is NOT

The second category defines metadata by discussing the information that it does not record. Only 33 (36.3% of the total articles, 54% of those providing a definition) include statements that define metadata in this fashion. In this category, metadata is defined by its relationship to a message’s content: specifically, the focus is on the notion that metadata is not the content of an electronic message, or not “private communications.” This quote from the Globe and Mail shows how metadata is defined in one of the passages as not being the content of a shared communication.

…. metadata are not ‘private communications’; that is, not the content of communications, just data associated with those communications. (23)

In addition to defining metadata by contending that it is not content, brief, general, and broad descriptions of metadata are stated in several of the articles. Similar to the last category, the majority of these statements refer to metadata using nonspecific terms. These quotes from the Toronto Star show this form of definition.

It is the context but not the content of a communication. Context, not content. (11)

‘…metadata’ - the what, where, and how of emails, texts or cellphone calls - not the actual content of conversations, written exchanges, or photos. (27)

A smaller number of definitional passages point out that metadata does not capture the communications that are exchanged as part of a conversation. These descriptions are
slightly more specific than those that simply state that metadata does not record content. The following quotes from the *Globe and Mail* and the *Toronto Star* show this form:

> Metadata is information about an electronic communication excluding the spoken or typed words. (16)

> That means the who, where and when of telecommunications exchanges, but not the what - the actual content of what is said or written. (28)

An envelope metaphor is introduced in several statements that are made in the articles to make the case that metadata does not record the personal and at times private information that people share when they communicate over electronic devices. The following quote from the *Globe and Mail* demonstrates this description of metadata:

> ‘Metadata is the envelope information,’ surveillance czar John Adams would tell his counterparts during briefings, sources say. Sometimes, the major-general-turned-mandarin would even wave a paper envelope to make his point that the metadata program is far from the digital equivalent of steaming open letters. Citizens' communications contents were, are and would forever be sacrosanct inside the envelope, and off-limits. CSEC merely wanted a better glimpse at the address, return address and other routing information. (20)

This metaphor defines metadata by arguing that it is only the surface level information that is connected to electronic communications. Therefore, this metaphor reinforces the same message which argues that the information, or communicative content, that is stored within electronic messages is not considered to be metadata.

Lastly, another metaphor is used to reinforce the argument that metadata is not content. This metaphor represents metadata not as the picture that is taken when an image is photographed, but rather as the contextual information that is created when a photograph is captured. An example of this description of metadata is found in the following quote from the *National Post*:
In a rare public appearance that follows unprecedented scrutiny of the ultra-secretive spy agency, Mr. Forster denied CSEC had been monitoring the private communications of Canadians as it vacuumed up metadata, or ‘data about data.’ Comparing communications to a photograph, he said ‘the picture is the content. But what comes with that picture is other bits of data around the date, the time, the focal length, the aperture, the pixels, so it's data about it, but it's not the picture.’ (2)

Definitional passages in this category used many different techniques to reinforce the argument that metadata does not record the content that is shared when people communicate using electronic devices. In some cases, the passages explicitly argued that metadata does not record content by stating that metadata does not record the verbal words, or written messages that people exchange over electronic devices. In other cases, metaphors were used to argue the same message by defining metadata as not being the letters contained in mailed envelopes or the pictures that are taken when photographs are captured. Similar to the last category, metadata was defined by statements that use broad and general language which makes the case that metadata is only contextual information.

Overall, articles are almost twice as likely to contain a definition of metadata in terms of what it is (61 out of 91 articles, 67%) by including statements that define metadata in this fashion than by discussing what metadata is not (33 out of 91 articles, 36.3%) include statements that define metadata by what it is not.
Figure 5 documents the percentage of articles per year that include statements from only one of the definitional categories of metadata, both of the definitional categories, or neither of the two categories. The 2014-2015 year segments are combined into one category since only a small number of articles are published on this topic in 2015. Only 7 of the 91 articles were published in 2015.

When the percentage of articles per year from 2013-2016 which include statements that define metadata by the two definitional categories are analyzed several important findings are observed. When a chi-square test is conducted comparing articles that include statements from neither definitional category, only one of the two categories, or both, a marginally significant relationship is observed between these variables ($X^2(4)=8.529$, $p<.1$). This data suggests that the articles are more likely to include more comprehensive statements that define metadata in 2013 than in later years. In 2013, 53.3% of the articles for the year include statements from both definitional categories. In the later years we see a large decrease in the percentage of articles per year which include statements from both definitional categories: 25.8% in the 2014-2015-year segment and 26.7% in 2016.
When the two definitional categories are analyzed across the three publications no significant relationship is seen between the likelihood that a publication will define metadata by either of the two definitional categories. When comparing the first definitional category across all publications the following chi-square score is observed \( (X^2(2) = 3.609, \text{ n.s.}) \). When comparing the second definitional category across all publications a chi-square score of \( (X^2(2) = 1.077, \text{ n.s.}) \) is observed. These scores make the case that there is no significant relationship across publications regarding the likelihood that a definition will be provided that defines metadata by what it is or what it is not.

### 4.2.3 Evaluation: Metadata as Unimportant; Metadata as Meaningful

These next two categories are evaluative. The first category argues that metadata is *unimportant* because it is information that is about communications instead of information that is about individuals. This category uses minimizing language to reinforce the claim that metadata is not important. Also, metadata is familiarized to present the argument that this information is inconsequential. The second category argues the alternative position that metadata is *meaningful* because it discloses a large amount of information about people’s lives.

**Metadata as Unimportant**

Overall, 45.1% of the total articles include statements which argue that metadata is information that is unimportant. It is important to note however that the majority of the articles that define metadata as unimportant include statements or references that are proposed from individuals who are a part of the national security framework; primarily individuals working for CSEC propose this definition of metadata. Metadata is argued to be inconsequential through three strategies: first through the use of minimizing language; second, by ‘familiarizing’ metadata; and third, by emphasizing the claim that metadata is information about *communications*, and not about people.

The first strategy describes metadata by using minimizing language to make the case that metadata is not important. This argument is observed in the following quote from the *Toronto Star*.
In any case, he went on, it was only ‘metadata’ that the spies were monitoring - such as whom the travelers were contacting. Agents didn't listen in on any actual conversations. (35)

The phrase “only metadata” makes the case that metadata is information that is insignificant or inconsequential.

Another example of the minimizing language that is used to define metadata in this category is when it is argued that metadata is simply “just data.” A quote from the Globe and Mail, which was previously cited in the second category, defines metadata by arguing that it is “just data associated with … communications” (23). When metadata is represented as being “just data” that is related to communications this makes the case that metadata does not contain information that is confidential or private.

Similarly, minimizing language is used in statements made in the articles when it is argued that metadata is mainly numerical information. When metadata is claimed to be mostly numerical data this creates the impression that metadata is disconnected from the meaning that is embedded in the written or verbal messages that people exchange using electronic communication devices (National Post 3). This quote from an article published in the National Post in 2016 introduces this argument by discussing the minimizing language that is used to describe metadata.

But the agency had been indefinitely keeping the metadata, using it to gain ‘insight otherwise impossible to glean,’ Noel wrote in his decision. He acknowledged the raw metadata ‘consists mostly of numbers,’ ‘may only have limited privacy impacts’ and had yielded useful intelligence, creating new investigative leads. (3)

The use of the envelope metaphor emphasizes the familiar nature of metadata, and thus minimizes any concerns about metadata collection. Here, metadata is argued to be the information that appears on the outside of envelopes that are delivered through the postage system. In this argument, metadata is explicitly linked to the information that we have always disclosed when letters are mailed, thus implying that individuals should not
be concerned about the collection of this information by national security agencies. This example from the *Globe and Mail* exemplifies the envelope metaphor that appears frequently in statements made within the articles to make the case for this idea.

‘Metadata is the envelope information,’ surveillance czar John Adams would tell his counterparts during briefings, sources say. Sometimes, the major-general-turned-mandarin would even wave a paper envelope to make his point that the metadata program is far from the digital equivalent of steaming open letters.

Citizens' communications contents were, are and would forever be sacrosanct inside the envelope, and off-limits. CSEC merely wanted a better glimpse at the address, return address and other routing information. In this alluring metaphor, CSEC was nothing more than a third party who happens to see the outside of a mailed letter while it's in transit. Metadata could be observed, mapped and shared without running afoul of laws precluding domestic surveillance. (20)

Most of the descriptions of metadata argue that it is information that is about communications instead of being about people. This third strategy diminishes the importance of metadata by blurring the connection between metadata and the actions of individuals. When passages claim that metadata is information that is about communications instead of about individuals this makes the case that metadata is harmless because it is only related to patterns about communications instead of patterns about people. The examples listed previously argue this idea including when metadata is described in passages as “just data associated with … communications” as claimed in a statement made in an article published in the *Globe and Mail* (23). The envelope metaphor, published in the *Globe and Mail*, also develops this argument by describing metadata as “the outside of a mailed letter while it's in transit” (20).

Definitional passages in this grouping use three strategies to develop the argument that metadata is information that is unimportant. The first strategy used minimizing terms to argue that people should be unconcerned about government metadata collection. Minimizing language was used, in one instance, to make the claim that metadata is inconsequential because it does not record the meaning that is stored in the messages that
people exchange electronically. The second strategy familiarized metadata to make the case that metadata is information that is irrelevant. Here it is argued that metadata is simply information that has always been disclosed when mailing letters. If metadata is information that has been freely disclosed in the past this makes the case that people should not be concerned about national security agencies collecting this info. Lastly, the third strategy contends that metadata is disconnected from individuals by claiming that metadata is information that is about communications instead of being about people.

**Metadata as Important**

This final category argues that metadata is information that is about people, as information that people should value as significant, and as information that people should be concerned about. A much greater percentage of the total articles (65.9%) include passages or statements that define metadata in this fashion. This grouping argues the opposing perspective to the last category which claimed that metadata is meaningless. Conversely this grouping makes the case that people should be worried about the government’s collection of metadata because several passages argue that metadata provides revealing information about individuals. Similar to the former evaluative category the majority of the articles which include statements that define metadata as important either quote or reference an individual that has been interviewed by the publication, a court ruling, or a statement made by the privacy commissioner of Canada. The following quote from the *National Post* makes the case that metadata can tell us a lot about people’s lives.

> Such metadata has the power to show our movements and associations through an airport, across town, or across the country. Our mobile devices and the associated metadata leave behind a digital trail that can reveal where you live, work, travel, what you purchase online, who you associate with, even what time you are likely to go to bed, wake up and leave home. (10)

In this description of metadata, a trail metaphor is introduced which claims that metadata can be used for tracking purposes by national security agencies. Some definitional passages in this category include specific descriptions of the information that metadata is
claimed to record. These passages are very different than the general descriptions of metadata that appear in the former categories. In this grouping, clear descriptions of metadata are provided which make the case that this information is connected to the actions of individuals and it is argued that it can be used to record very revealing details about people’s lives.

**Figure 6 How the Two Evaluative Categories are Represented from 2013-2016 by the Percentage of Articles per Year**

Figure 6 documents the percentage of articles per year that include statements from only one of the evaluative categories of metadata, both of the evaluative categories, or neither of the two categories. When the percentage of articles per year that included statements that define metadata as information that is unimportant is compared with the year that the articles were published, no significant relationship is seen between these two variables ($X^2_{(2)}=0.967$, n.s.).

Conversely, a chi-square test shows that there is a statistically significant relationship between the year that articles are published and the percentage of articles per year that include passages that define metadata as information that is important ($X^2_{(2)}=6.462$, $p<.05$). This relationship demonstrates that as the year’s progress, the publications are more likely to publish articles that contain passages which define metadata as information that is important.
Figure 7 The Percentage of Articles per Publication that Include a Definition of Metadata as Important

Figure 7 documents the percentage of articles for each of the individual publications that include passages that define metadata as information that is important. This data allows the reader to observe the variations across the three publications.

There is also a significant relationship between publication and the likelihood that an article includes passages that define metadata as information that is important ($X^2(2)=0.001$, p<.01). When analyzing the three publications, it is clear that *Globe and Mail* articles were most likely (86.8%) to include passages that define metadata as important. The *Toronto Star* is the second publication that is most likely to include a passage that defines metadata in this way with 55.6% of its articles containing a passage that defines metadata in this fashion. Lastly, *National Post* articles are least likely to include a passage that defines metadata as important with only 41.2% of its total articles including a passage of this type.

4.2.4 How the OPC and CSEC Define Metadata

The last section discussed some of the arguments that were introduced in the passages of the articles concerning how metadata is defined. This section begins by introducing two competing perspectives concerning the definition of metadata which were developed by the Communications Security Establishment Canada and the Office of the Privacy Commissioner of Canada. It is important for Canadians to consider the differences between how these two agencies define metadata because they have conflicting
perspectives regarding the privacy/security trade-off. When defining metadata, many of the articles included passages with characteristics of both of these definitions and some of the articles included passages that directly cite one or both of the organizations’ definitions of metadata.

The Office of the Privacy Commissioner of Canada’s (OPC) definition of metadata is explained and developed in 3 different sources: the OPC’s 2013-2014 Annual Report to Parliament, a 2014 report commissioned by the OPC titled “Metadata and Privacy a Technical and Legal Overview,” and lastly a 2006 fact sheet released by the OPC named “The Risks of Metadata.” CSEC’s definition of metadata comes from the Communication Security Establishment Commissioner’s 2013-2014 Annual Report as well as from CSEC’s official website which describes metadata in a section titled “Metadata and our Mandate.” Many of the articles in this study include passages that reference either the CSEC or OPC definition of metadata.

The OPC definition describes metadata as “data about data or information about information” (The Risks). These phrases appear frequently in statements made in the articles. In addition to these two phrases, the OPC defines metadata as information that “sometimes can be more revealing than the actual content of a communication” (“Annual Report”). Computer scientist Daniel Weitnzer is quoted by the OPC to develop the argument that:

Metadata [is] arguably more revealing [than content] because it’s actually much easier to analyze the patterns in a large universe of metadata and correlate them with real-world events than it is to go through a semantic analysis of all of someone’s email and all of someone’s telephone calls. (“Metadata and Privacy”)

The OPC also states that metadata “may not be the contents of our communications, but [it] … can paint a profoundly detailed picture of our lives” (“Metadata and Privacy”). The 2013-2014 annual report published by CSEC argues that metadata can be used to “compile a detailed profile of an individual” and it is claimed in a report published by the OPC that metadata leaves behind a “digital … personal trace” that can “identify individuals” (“Annual Report”; “Metadata and Privacy”). The OPC states that metadata is
created as individuals use technological devices (“Metadata and Privacy”). Metadata is also “a hidden level of extra information that is automatically created and embedded in a computer file” (The Risks). Lastly the OPC critiques the idea proposed by the former chief of CSEC John Adams that metadata is similar to the information that appears on the outside of mailed letters. The OPC argues that “the size, shape or colour of an envelope can sometimes be quite revealing as to what message it contains” (“Metadata and Privacy”).

CSEC defines metadata as routing information that is “associated with a communication” (“Communications Security”). CSEC also states that “metadata excludes the content of a communication” and this is one of the key differences between how CSEC defines metadata in comparison to the OPC (“Communications Security”). One of the CSEC sources states that “while metadata reveals a certain amount of information about devices, users and transmissions, it is contextual and does not expose the content of emails, phone calls or text messages” (“Metadata and Our”). The idea that metadata is not content appears numerous times in statements made throughout the articles.

Of the 91 articles, 48.4% include passages that contain certain elements of the OPC’s definition of metadata. For instance, some articles include statements that make the case that metadata is similar to leaving behind an electronic trail that can be followed. Other articles include passages which contend that while metadata is not content, it can still provide very revealing details about a person’s life. In other instances, the articles may only include statements that describe metadata using the “data about data” phrase. It is important to note that only a small portion of these articles (7.7%) include passages that directly attribute the OPC when an OPC definition of metadata is introduced or referenced. This suggests that readers are rarely informed that the OPC definition of metadata is being referenced when metadata is defined in the passages of the articles.

Out of the 91 articles, 53.85% include a definition of metadata that is similar or related to how CSEC defines the term. Most of these passages argue that metadata is not content. Another core concept that reflects the CSEC definition is the claim that metadata is information that is about communications rather than about people. As with the OPC
definition, only a fraction (25.3% of the 91 articles) include statements that directly reference or attribute CSEC when defining metadata. This suggests that the majority of the articles do not include statements that attribute the security agency when its definition of metadata is included.

**Figure 8 The Percentage of All of the Articles that Directly Reference CSEC vs. Percentage of All the Articles that Directly Reference the OPC**

Figure 8 provides the reader with a brief overview of the percentage of the 91 articles that include statements that directly reference or attribute CSEC’s or the OPC’s definitions of metadata. When including a definition of metadata, the passages of the articles are 3.29 times more likely to reference and attribute CSEC’s definition of metadata than the OPC’s definition. This suggests that readers are more likely to be aware of how CSEC defines metadata in comparison to how the OPC defines the term. It is important to note that *this graph only records the percentage of the total articles that include statements that directly attribute CSEC or the OPC when a definition of metadata is cited, introduced, or described.*
Figure 9 The Percentage of All Articles that Include Elements of Only the CSEC Definition of Metadata, Only the OPC Definition, Both Definitions, or Neither Organization’s Definition of Metadata

Figure 9 shows the percentage of the 91 articles that include characteristics or elements of only the CSEC definition, only the OPC definition, both definitions, or neither definition. The neither category represents the percentage of the 91 articles that do not include a definition of metadata as well as the articles that include passages that may define metadata in different way than the OPC or the CSEC definitions.

Only small variations are observed when comparing the percentage of articles that contain passages which only include elements of the CSEC definition of metadata in comparison to the percentage of all articles that contain passages which only include elements of the OPC definition (see fig. 9). Overall, there is only a 2.2% variation between the articles that include passages which only include characteristics of the OPC definition vs the articles that include passages which only include elements of the CSEC definition.

However, it is important to note that this graph does not reflect the percentage of all of the articles that include statements which directly reference the OPC or CSEC when defining metadata. 28.57% of the total articles contain passages that do not include a definition of metadata that includes characteristics of how the two organizations define
the term. However, this does not suggest that the passages in these articles do not include a definition of metadata within the body of the text.

A total of 34 articles (38.46% of the sample) include passages that define metadata which reflect the views of both CSEC and the OPC. This is double the number of articles that contain passages which only discuss one of the two organization’s definitions of metadata.

The articles that do not include either the CSEC or the OPC definition at times include passages that define metadata through the use of lists or examples. 27 (29.67% of all the articles) include passages that simply defined metadata by providing examples of the type of information that is included within metadata. Some examples include information such as: IP addresses, e-mails, phone logs, and geo-location information.

Figure 9 as a whole suggests that readers are almost equally likely to be exposed to either the OPC or CSEC definition of metadata. Figure 9 also suggests that while readers may be almost equally exposed to the CSEC or OPC definitions of metadata, readers may be less aware of how the two organizations differ in how they define metadata.

**Figure 10 The Percentage of Articles per Publication that Contain Statements Which Include Elements of Only the OPC Definition of Metadata, Only the CSEC definition, Both Definitions, or Neither Organization’s Definition of Metadata**
Figure 10 reports the percentage of articles by publication that contain passages which include characteristics of only the CSEC definition, only the OPC definition, or both definitions of metadata. Between the three publications several interesting variations are observed. The Globe and Mail is the publication that is most likely to include both the CSEC definition of metadata as well as the OPC’s definition with 57.89% of the publication’s articles containing passages which include characteristics of both definitions. A small variation of 2.63% of the articles for this publication contain passages that only include characteristics of how CSEC defines metadata in comparison to the OPC.

The Toronto Star is less than half as likely as the Globe and Mail to include characteristics of both definitions of metadata. Only 27.78% of the articles for this publication contain passages that include elements of both definitions and the same percentage is observed for the articles that contain statements which only include the OPC definition. The Toronto Star also sees a large variation where it is far more likely to only include the OPC’s definition of metadata in comparison only including CSEC’s. The Toronto Star is 2.5 times more likely to only include the OPC definition than CSEC’s.

The National Post is the publication that is least likely to include both definitions of metadata in statements that are made within the articles. The National Post also sees a large variation in reporting where only the CSEC definition of metadata is discussed in over twice as many passages than the OPC definition.
Figure 11 The Percentage of Articles per Publication that Contain Passages which Directly Reference and Attribute Only the OPC Definition, Only the CSEC Definition, Both Definitions, or Neither Organization’s Definition of Metadata

Figure 11 compares the three news publications to determine the percentage of articles per publication that contain passages which directly reference or attribute only the CSEC definition of metadata, only the OPC’s definition, or both definitions in the body of the text. Notable variations between the three publications are observed when comparing the data in this manner.

While the National Post demonstrates an even distribution of articles that include statements that reference and attribute CSEC and the OPC when metadata is defined in the passages, the other two publications rely more heavily on one definition than the other. Both the Globe and Mail and the Toronto Star reference and attribute only CSEC’s definition of metadata much more than the OPC definition.

The Globe and Mail shows the largest variation where the articles are 6.5 times more likely to contain passages that define metadata by referencing only the CSEC definition. It is important to note that while the CSEC definition of metadata may be referenced and attributed more frequently these numbers do not demonstrate whether the articles contain passages that critique or support how CSEC defines metadata. These numbers do suggest, however, that the Globe and Mail is far more likely to mention and attribute CSEC instead of the OPC when the passages define metadata.
A similar trend in reporting is viewed in the *Toronto Star* where the articles are 6 times more likely to include statements that attribute CSEC instead of the OPC when the statements define metadata. However, in this publication, the security agency is referenced and attributed in a much smaller percentage of the publication’s articles than in the *Globe and Mail*. The *Toronto Star* references and attributes the CSEC definition of metadata slightly more than half the percentage of times as the *Globe and Mail*. The *Toronto Star* also references and attributes the OPC definition of metadata slightly more than half the percentage of times as the *Globe and Mail*.

**Figure 12 The Percentage of Articles per Year that Include Passages which Directly Reference Only the OPC Definition, Only the CSEC Definition, Both Definitions, or Neither Organization’s Definition of Metadata**

Figure 12 introduces the percentage of articles per year that include statements which directly reference or attribute only the CSEC definition, only the OPC definition, or both definitions of metadata. Several large variations in reporting are observed when comparing the percentage of articles per year that include passages that only reference CSEC in comparison to the OPC.

In 2013, 10% of the articles contain passages that define metadata by only referencing how CSEC defines the term. In 2014, a large jump in coverage is seen where almost four times the percentage of articles for the year contain statements that define metadata by
referencing only the CSEC definition in comparison to year 2013. In 2014, 39.13% of the articles for the year contain passages that define metadata by only referencing CSEC’s definition. A large dip in coverage is observed in 2015 where 25% of the articles for the year include statements that only reference the CSEC definition. Lastly in the final year, another dip in coverage is observed where 20% of the articles in that year contain passages that only reference CSEC when defining metadata.

The percentage of articles for the year that contain statements which only reference the OPC when defining metadata is much lower than the overall percentage of articles that include statements which only reference CSEC. However, a similar ascending pattern in coverage is observed across the years for articles that include passages which only reference the OPC definition of metadata. In 2014, the same percentage of articles for the year is observed between the articles that contain passages which only reference the OPC’s definition and the articles that contain passages which include both definitions; 4.35% of the articles in 2014 contain passages that include both definitions of metadata or only reference the OPC definition. A dip in coverage is observed in 2015 where no passages exclusively reference the OPC’s definition. Lastly, year 2016 sees the highest percentage of articles than any other year which contain statements that only reference the OPC’s definition of metadata. 10% of the articles for the year in 2016 contain statements that exclusively reference the OPC.

Across all years, the percentage of articles per year that include passages that reference both definitions remains almost the same, ranging from 3.33% to 4.35%. A slight exception is observed in year 2015 where no passages reference both CSEC and the OPC’s definitions of metadata.

### 4.3 The TopicsObserved in the Introductory Paragraphs of the Articles

This next section analyzes the introductory or lede paragraphs of the articles included in this study. Lede paragraphs are important because they signal to the reader the main point or message that the author intends to develop. Lede paragraphs often distill the main
theme of the article into one small fact-filled segment which is placed at the beginning of the article. Lede paragraphs may also be the only piece of an article that a reader may quickly browse in addition to an article’s headline. Analyzing the topics in the lede paragraphs will shed light on the primary messages that journalists have attempted to convey to the Canadian public concerning this issue. The themes observed in these lede paragraphs were inductively determined by selecting the categories that were seen across all of the publications.

Five major categories were identified (see fig. 13). In the first category, the concepts of transparency and oversight work together to argue that Canada’s national security agencies operate with only marginal oversight and with very little transparency. The lede paragraphs that focus on transparency were combined with ledes that discuss oversight because these two topics often overlap. The ‘scary/about you/spying’ category consists of lede paragraphs which suggest that Canadians should be concerned about government metadata collection because it is argued that this amounts to intrusive government surveillance. The ‘breaking the law’ category contends that Canada’s foreign or domestic intelligence agencies have committed illegal acts by collecting metadata from Canadians in an unlawful way. The ‘terrorists/terrorism’ category contains lede paragraphs which claim that Canada’s intelligence agencies need metadata to protect against national security threats. Lastly the ‘other’ category groups together a collection of lede paragraphs which did not fit into the former categories. This category consists of three smaller subcategories.

**Figure 13 The Percentage of All Lede Paragraphs from Each of the Categories**
The largest number and percentage of lede paragraphs fall under the ‘transparency/oversight’ category; 35 lede paragraphs (38.5% of the total) discuss this topic. The second largest category of ledes are those within the ‘scary/about you/spying’ grouping with 30 lede paragraphs (33% of the total) making up this grouping. The ‘breaking the law’ category is the third largest grouping but there is a steep decline in the percentage of lede paragraphs that make up this category. Only 14 lede paragraphs discuss the ‘breaking the law’ category (15.4% of all of the total). The ‘terrorists/terrorism’ grouping is the fourth largest category consisting of 8 lede paragraphs or 8.8% of all the ledes. Lastly, the ‘other’ topic is the smallest category with only 4 lede paragraphs (4.4% of the total).

4.3.1 Transparency/Oversight

This is the largest grouping of lede paragraphs which includes 35 (38.45% of the total) ledes. Many articles begin by claiming that the oversight mechanism for CSEC is very thin. These articles argue in the introductory paragraphs that CSEC’s current oversight mechanism is ineffective because it is claimed that only one judge oversees the agency’s operations as a form of accountability. Some of the articles begin by arguing in the lede paragraphs that CSEC has grown to such a large size that more adequate accountability measures may be required to effectively oversee the security agency. An example of this theme can be seen in the lede paragraph of an article published in the National Post in 2013:

Opposition parties are calling for greater oversight of Canada's spy agencies as questions continued to swirl about the size and scope of super-secret U.S. and Canadian surveillance programs … CSEC is overseen by an independent commissioner, in this case a retired judge, who reviews its activities to ensure they comply with the law, while the Canadian Security & Intelligence Service is overseen by the Security Intelligence Review Committee, but otherwise no parliamentarians are involved. (7)

Another important issue is raised in this lede paragraph from the National Post that shows a consistent theme that is observed in this grouping. The lede suggests that
parliamentarians are not included in CSEC’s current oversight mechanism. Instead, several of the articles emphasize in their lede paragraphs that only a single judge is responsible for overseeing CSEC’s surveillance operations.

Other articles argue in the lede paragraphs that there is a lack of transparency between Canadian national security agencies and Parliament. One article from the *Globe and Mail* begins by arguing in the lede that Members of Parliament who are responsible for assessing national security laws are not cleared to view the secretive activities of Canada’s national security agencies. The article argues in its lede paragraph that individuals working for CSEC or CSIS would not be able to share information with parliamentarians due to a lack of security clearances (29). This quote from the lede paragraph of the article makes the case that there appears to be a culture of secrecy that surrounds Canada’s national security agencies:

> The chair of a parliamentary anti-terrorism committee says he was unaware of a federal spying program that allows for the collection of Canadians’ data trails. Conservative Senator Hugh Segal, who vets security laws as chair of the special Senate committee on anti-terrorism, said in an interview that he and other parliamentarians learned of the program’s existence only when they read about it in The Globe and Mail this week. In Canada, MPs and senators are not looped into the mechanics of surveillance programs. ‘We do not now have a mechanism to do that,’ Mr. Segal said, explaining that parliamentarians lack security clearances. ‘The people whom we would ask, who run the agencies, would be prohibited from giving us any of the details.’ (29)

A different article written by Ann Cavoukian, the former Information and Privacy Commissioner of Ontario, begins by arguing in the lede that Canada is even more secretive than the United States in its surveillance activities. In the lede paragraph, Cavoukian discusses the 2014 reforms to the NSA that were announced by Obama. Cavoukian then contrasts America’s reforms with Canada’s apparent lack of communication concerning the surveillance activities of CSEC (Globe and Mail 11). The following lede paragraph from the *Globe and Mail* argues that Canada has been more
reluctant than the United States in engaging in a dialogue with the public concerning the activities of its electronic surveillance agency.

Technology allows our every move to be tracked, collected and catalogued by our governments. U.S. President Barack Obama's announcement of reforms to the National Security Agency (NSA) demonstrates that free and open societies need a candid discourse on the surveillance powers of intelligence agencies. Yet, while our U.S. neighbours are debating the future of phone and Internet surveillance programs, our government is maintaining a wall of silence around the activities of the Communications Security Establishment Canada (CSEC). This silence is putting our freedoms at risk. (11)

Other articles begin by arguing in the lede paragraphs that CSEC operates under a veil of secrecy which is far too opaque. They claim that the agency lacks both public accountability as well as oversight. One article published in 2013 in the *Globe and Mail* contends in its introductory paragraph that, if Canada is investing close to half a billion dollars into CSEC’s annual budget, the public has a right to know more about how the agency operates. The following lede makes the case that Canada has invested a significant amount of money in the past into developing its surveillance capabilities while limiting accountability measures.

Communications Security Establishment Canada has a global reach for its surveillance and a budget that has ballooned to almost a half-billion dollars. But, as Colin Freeze reports, it lacks public accountability or oversight - allowing for a level of secrecy even some of its key architects say needs to change. It is known as ‘Camelot,’ and it is believed to be among the most expensive government buildings Canada has ever built. Next year, the analysts, hackers and linguists who form the heart of Communications Security Establishment Canada are expected to move from their crumbling old campus in Ottawa to a gleaming new, $1-billion headquarters. (21)
An article published in the Toronto Star in 2016 begins by arguing in its lede paragraph that CSIS collected metadata from Canadians who were not considered to be a threat to national security over a 10-year period without informing the courts. The following lede paragraph from the Toronto Star makes the case for this argument.

Michel Coulombe, Canada's top spy, is in deep trouble with the courts and his political boss, Public Safety Minister Ralph Goodale, over revelations CSIS kept a decade's worth of data on Canadians who are no threat to national security. (29)

In 2016, several articles begin by stating in the lede paragraphs that Canada illegally shared information about Canadian citizens with members of the Five Eyes partnership. The following article from the Toronto Star makes the case in its lede that there is a lack of transparency between the former Conservative government and the public concerning Canada’s surveillance activities.

To learn that our digital surveillance agency broke privacy laws by revealing information about Canadian citizens to our allies is one thing. To learn that the Conservative government of the day, when apprised of this security breach, withheld the information from Canadians, is quite another. But that is where we are today, after learning of a major invasion of Canadian privacy more than two years after the fact. … This despite an effort Thursday to get ahead of this story with the first-ever background briefing for journalists from an official with the Canadian Security Establishment (CSE) - only 26 months after a software glitch was discovered that was sending metadata on Canadians to our Five Eyes allies without the proper scrubbing to hide identities. (25)

4.3.2 Scary/About You/ Spying

This grouping consists of 30 ledes, which represent 32.97% of the total lede paragraphs. Ledes are included in this category if the primary message observed in the lede paragraph states, suggests, or implies that national security agencies are conducting surveillance activities that are unnerving or worrisome. Many of these ledes argue that government surveillance is occurring on domestic Canadian citizens, and that much of this
surveillance is being conducted by security agencies that are supposed to target foreign threats. In other instances, lede paragraphs argue that the government is collecting too much information from citizens.

Ledes in this category focus heavily on developing the argument that national security agencies are collecting metadata from Canadians and they are therefore spying on citizens. Some of the lede paragraphs begin by explaining that Canada’s metadata collection program was previously suspended due to domestic surveillance concerns and was reinstated by the defense minister several years later (Globe and Mail 25). The following lede from the Globe and Mail makes the case that Canada’s metadata collection program was reinstated despite the domestic-surveillance concerns that were raised by a Canadian “federal watchdog agency” (Globe and Mail 25).

Defence Minister Peter MacKay approved a secret electronic eavesdropping program that scour global telephone records and Internet data trails - including those of Canadians - for patterns of suspicious activity. Mr. MacKay signed a ministerial directive formally renewing the government's ‘metadata’ surveillance program on Nov. 21, 2011, according to records obtained by The Globe and Mail. The program had been placed on a lengthy hiatus, according to the documents, after a federal watchdog agency raised concerns that it could lead to warrantless surveillance of Canadians. (25)

Some of the lede paragraphs in this grouping argue that the devices of travelers who passed through a Canadian airport were tracked by the government by collecting metadata from the devices. One lede, published in 2014 in the National Post begins by making this case and then continues to develop the argument that metadata can record revealing details about individuals without the person being aware that the information is being disclosed.

The operation involved the processing of at least two weeks of identifying information associated with our mobile devices, their location in time and space, primarily in Canada (beginning at a major international airport) … Our mobile devices and the associated metadata leave behind a digital trail that can reveal
where you live, work, travel, what you purchase online, who you associate with, even what time you are likely to go to bed, wake up and leave home. (10)

Other ledes begin by contending that Canadians are unintentionally swept into the Canadian government’s surveillance activities despite oppositional claims that are proposed by CSEC. The following 2014 National Post lede paragraph contends the security agency’s argument that Canadians are not targeted by its surveillance.

Canada has a spy problem. Over the past year and a half, Canadians have learned a great deal about the activities of the Communications Security Establishment (CSE). CSE is responsible for spying on communications abroad, protecting some government systems, and helping other federal departments spy on Canadians. CSE and the federal government alike insist that Canadians are not ‘targeted’ by our spies and assert that claims to the contrary are inaccurate or wrong. But CSE’s own rebuttals don’t hold water. (16)

Several other ledes in this category begin by critiquing CSEC’s statement that it “incidentally” conducts surveillance on Canadians. The following introductory paragraph from the National Post makes this case:

Canada's foreign electronic intelligence agency admits it ‘incidentally’ spies on Canadians, but wants to reassure the public it protects the privacy of that information. ‘In the course of targeting foreign entities outside Canada in an interconnected and highly networked world, it is possible that we may incidentally intercept Canadian communications or information,’ the Communications Security Establishment Canada (CSEC) said in a new statement posted on its website. It is the first time the country's ultra-secret signals intelligence agency has strayed from its standard assurance that it does not ‘target’ the electronic communications of Canadians. (14)

Several lede paragraphs begin by discussing NSA surveillance and then progress to discuss how this influences Canadians (National Post 13). Another common trend is ledes beginning by discussing NSA surveillance and then progressing to discuss how Canada’s
national security agencies conduct their surveillance (National Post 8; Toronto Star 22). Other ledes reverse this trend and begin by discussing Canada’s domestic surveillance by comparing it to America’s surveillance. The following lede paragraph from the *Globe and Mail* demonstrates this third trend in reporting.

The revelation that Canadians' phone calls and Internet activity are being monitored by government officials in much the same fashion as in the United States is disturbing and unacceptable. Even more troubling is the fact that the authority to carry out this surveillance came via ministerial directive, leaving Parliament out of the loop. The secret program should be halted until, at the very least, it can be debated in the House of Commons. (2)

4.3.3 Breaking the Law

This category consists of 14 texts which represent 15.38% of the total lede paragraphs. These ledes demonstrate that there are instances where Canada’s domestic or foreign national security agencies have broken Canadian laws by unlawfully collecting metadata on Canadians.

Many of the ledes published in 2016 begin by arguing that CSIS secretly collected metadata from Canadians who are considered to be “unrelated to national security threats” (National Post 3). These lede paragraphs begin by focusing on a Canadian federal court ruling that CSIS illegally retained Canadian metadata from court-authorized domestic surveillance operations. The following lede paragraphs from the *National Post*, the *Globe and Mail*, and the *Toronto Star*, respectively, make the case that a Canadian surveillance agency broke the law by keeping data that it was not supposed to save.

A previously unknown unit of Canada's intelligence service has been illegally keeping data unrelated to national security threats, the Federal Court disclosed Thursday. (3)

The Federal Court of Canada has faulted Canada's domestic spy agency for unlawfully retaining data and for not being truthful with judges who authorize its intelligence programs. (19)
Canada's spies for almost a decade illegally kept and analyzed data on people who posed no threat to national security, a federal court judge has ruled. In a scathing ruling, Justice Simon Noel said the Canadian Security Intelligence Service had illegally retained an unknown amount of data on ‘third party’ and ‘non-threat’ individuals since 2006. (13)

Other ledes begin by arguing that a Canadian intelligence agency illegally shared Canadian metadata with the Five Eyes intelligence network. One lede from the *Globe and Mail* starts by arguing that the information had been accidentally shared for many years:

A federal spy agency inadvertently shared logs of Canadians’ phone calls and Internet exchanges with intelligence allies such as the United States for years, a newly disclosed report says. (30)

A different lede paragraph from the *Toronto Star* argues that both of Canada’s intelligence agencies broke laws by conducting surveillance on Canadians without obtaining a warrant:

Both of the nation’s spy agencies were outed by their official Ottawa watchdogs this past week for breaching Canadians’ privacy rights, and for snooping on taxpayers without warrants. (5)

Another lede from the *Toronto Star* claims in its introduction that one of Canada’s intelligence agencies was aware of its illicit surveillance activities since 2013.

Canada's secretive electronic spying agency realized in 2013 it was breaking domestic privacy rules by transferring Canadians’ data to allied countries … (16)

A lede paragraph from the *Globe and Mail* begins by arguing that Canada’s foreign intelligence agency discontinued its surveillance operation at one point due to domestic surveillance concerns. The article contends in its lede that the surveillance operation was initiated as a response to foreign threats.

Persistent foreign spying threats prompted Canada's electronic-eavesdropping agency to embark on a counterespionage campaign so aggressive that its former
chief says he ‘shut the place down’ before it could be exposed to allegations of wrongful domestic surveillance. (26)

Another lede from the Globe and Mail argues that CSEC illegally conducted domestic surveillance by tracking the devices of travelers who accessed the free WIFI at a Canadian airport:

It seems that Canadians have likely been the subject of digital surveillance by the Communications Security Establishment Canada, our own cyber-spies, according to files obtained by the CBC from the U.S. whistleblower Edward Snowden. Apparently CSEC tracked the wireless devices of passengers emerging from Canadian airports for days. CSEC is supposed to monitor only foreign data, so if there were Canadians at this airport - not much of a stretch - these activities ‘constitute a clear violation of CSEC's mandates and almost certainly of the Charter’ … (37)

4.3.4 Terrorists/Terrorism

Eight ledes (8.79% of all the lede paragraphs) make up this grouping. The ledes of most of the articles in this category suggest that the Canadian government’s collection of metadata by its national security agencies is necessary to detect terrorist threats. Many of the lede paragraphs, including this one from the Globe and Mail, present government agency arguments that without metadata it wouldn’t be possible for CSEC to locate national security threats:

Canada's top security and spy-agency officials have given the first detailed public defense of secret government surveillance programs that collect telecommunications ‘metadata.’ We wouldn't be able to find or locate our targets without it, John Forster, chief of the Communications Security Establishment Canada, told a Parliamentary committee. The head of the foreign-intelligence electronic-eavesdropping agency, Mr. Forster said snooping on metadata is fundamental for the Canadian government to pick out foreign terrorists and other
targets ‘in a sea of billion and billions of communications traversing the globe.’

(28)

Several of the lede paragraphs argue that the government’s collection of metadata is concerning because domestic citizen information is also collected. The lede below, which is also taken from the *Globe and Mail*, argues that Canadians are also affected by national security agencies searching for terrorist threats.

Spy agencies in Canada, the United States and elsewhere have been caught harvesting huge amounts of potentially private data from the laptops, tablets and cellphones of millions of people, including their citizens. They say this is necessary because of the changed world of security threats, such as terrorism … in the years after the Sept. 11, 2001 attacks, the snooping came home to roost. Government agencies started to see a utility in collecting telecommunications data from citizens and non-citizens alike, so as to better pinpoint threats that could now arrive from anywhere - including from suspects who move between a dozen portable devices over the course of a day. (23)

Other lede paragraphs in this category argue that the government’s collection of metadata does not infringe on citizen privacy. Several of the articles claim in their lede paragraphs that there is tension between CSEC representatives and the courts or between CSEC and the public. Some of these ledes claim that CSEC does not view its collection of metadata from Canadians to be spying. This lede paragraph from the *National Post* makes a case for this line of reasoning where a high ranking CSEC official defends the agency’s domestic collection of metadata which they argue is helpful in locating foreign threats.

The head of Communications Security Establishment Canada defended the collection of ‘metadata’ on Monday, saying it helped identify foreign adversaries without snooping on the private communications of Canadians. Testifying before the Standing Senate Committee on National Security and Defense, John Forster shot back against allegations of overzealous government electronic surveillance that have arisen as a result of leaks by Edward Snowden. (2)
Another article from the *Toronto Star* argues in its lede paragraph that CSEC does not conduct surveillance on Canadians because Canadians are not targeted by the agency. The following introductory paragraph makes the case that Canadians are never targeted by CSEC. Instead, the lede paragraph begins by arguing that CSEC only conducts surveillance on foreign individuals.

Canada's top national security officials loudly defended the actions of the country's ultra-secretive intelligence operations and denied breaching the privacy of Canadians, saying only foreigners are ever targeted. (27)

### 4.3.5 Other

This category is quite small and it contains a compilation of three separate subcategories. The three subcategories were compressed into one category because of the low number of lede paragraphs in each of the former categories. The subcategories are titled: ‘not breaking laws,’ ‘new surveillance laws,’ and ‘increased national security powers.’ In total, only 4 ledes make up this ‘other’ category which amount to 4.4% of the total lede paragraphs. Compressing these categories allow the larger trends observed in this study to be more evident.

Only one of the lede paragraphs begins by arguing that CSEC has not broken the law by tracking the devices of travelers passing through a Canadian airport. This lede paragraph claims that the watchdog for the security agency found that CSEC was not guilty of conducting domestic surveillance because it only collected metadata. The following lede makes the case for this idea and it was published in 2014 in the *National Post*.

The independent watchdog who monitors the Communications Security Establishment Canada said Thursday the electronic spy agency had not snooped illegally on Canadians when it collected metadata at airports. CSE Commissioner Jean-Pierre Plouffe said in a statement he had looked into allegations about airport surveillance that surfaced after Edward Snowden leaked a document about the project to the CBC, but he had found no wrongdoing. (4).
The second subcategory that makes up this grouping consists of two lede paragraphs which argue that national security agencies require increased powers to conduct their surveillance. The following lede was published in the *Globe and Mail* and it makes a case for this argument.

The stakes are considerable, which is why the folks who run the national security apparatus have quietly and not-so-quietly been laying down markers as Ottawa reviews their powers. The argument goes they need more tools, and more leeway, to do their important work. (8)

The third subcategory deals with ‘new surveillance laws’ and it includes only one lede paragraph, published in 2016 in the *Globe and Mail*. This lede claims that new legislation is needed that can set clear limits on how national security agencies collect metadata:

Canada’s privacy czar is calling on the Liberals to fulfill a promise to pass laws constraining the federal spies who are allowed to capture records of Canadians' phone and Internet activities. The Communications Security Establishment needs new legislation because it has not been careful enough in handling such material, says Daniel Therrien, the Privacy Commissioner of Canada. (27)

### 4.4 How the Topics Develop Over Time

**4.4.1 Lede Paragraphs Published by Year and by Topic Combining Years 2014-2015**

Figure 14 indicates the percentage of the lede paragraphs per year that were published for each of the 5 categories. Similar to previous figures, figure 14 combines years 2014 and 2015.
When years 2014 and 2015 are combined, a highly significant relationship is seen between the year in which lede paragraphs are published and the topics that are observed in the ledes ($X^2(8)=28.511, p<.001$). As stated previously, years 2014-2015 were combined due to the low percentage of lede paragraphs which discussed this issue in 2015. Furthermore, these years were combined because no major events occurred across all publications in 2015 concerning this topic.

However, it is important to note that the results listed above are impacted by the small number of lede paragraphs that make up each individual topic. When the chi-square score was conducted using SPSS the software indicated that 9 of the cells have an expected count of less than 5. SPSS also indicated that only 1.34 cells should contain a count of less than 5. Years 2014-2015 were combined in an attempt to reduce the number of cells that contained small numbers. Yet, even with these years combined a slightly large number of cells, 9 in total, contain values that are less than 5. According to Vaughan, this suggests that “the chi-square score may be exaggerated by [the] small expected frequencies” (89).

When the relationship between the year that the lede paragraphs are published and the topics observed in the ledes are analyzed, several important findings are observed. In
2013, the highest percentage of lede paragraphs for the year focus on the ‘scary/about you/spying’ category; 60% of the ledes in 2013 begin by discussing this topic. Another important observation is that the second largest topic for this year is ‘transparency/oversight.’ 30% of the lede paragraphs published in 2013 discuss ‘transparency/oversight’ related issues.

In years 2014-2015 the largest percentage of lede paragraphs in this year segment (38.7%) begin by discussing the ‘transparency/oversight’ topic. In this year segment, the ‘scary/about you/spying’ topic only represents 29% of the ledes published from 2014-2015 which is a large decrease than what was observed in 2013.

Lastly, in 2016, another increase in reporting is observed where the majority of the lede paragraphs discuss the ‘transparency/oversight’ topic. 46.7% of the ledes in 2016 begin by discussing the ‘transparency/oversight’ category. This suggests that as the years progress the lede paragraphs are more likely to focus on the ‘transparency/oversight’ category. The ‘scary/about you/spying’ topic is only published in 10% of the lede paragraphs in 2016 which indicates that as the years progress the ledes are less likely to focusing on this topic.

4.4.2 Topics by Publication (Percent per Publication)

Figure 15 displays the percentage of lede paragraphs per publication that fell into the five topic categories. The most frequent categories include ‘transparency/oversight’ and ‘scary/about you/spying’ (see fig. 15).

**Figure 15 Topic by Publication (Percent per Publication)**
When the chi-square test is conducted comparing the topics found in the introductory paragraphs of the articles and the publication that the lede paragraphs were published within no significant relationship is observed ($X^2_{(8)}=7.684$, n.s.). A p-value of 0.465 is seen which is much higher than 0.05. Since the p-value is much higher than 0.05 we conclude that the null hypothesis between these variables is confirmed (Vaughan 77). This suggests that there are no significant differences across the publications concerning the percentage of lede paragraphs for each publication that begin by discussing the five topics.

One clearly observable trend is that the two categories ‘scary/about you/spying’ and ‘transparency/oversight’ were featured prominently. Both the National Post and the Globe and Mail published most of their lede paragraphs on the ‘scary/about you/spying’ topic with the former publishing 41% and the later 34% of their ledes on this topic respectfully. The ‘transparency/oversight’ topic was the second most published grouping for the National Post and the Globe and Mail as well. The National Post published 24% of its lede paragraphs on this topic, and the Globe and Mail 32%. The Toronto Star saw a different trend where most of its lede paragraphs began by discussing ‘transparency/oversight’ (53%) while the ‘scary/about you/spying’ grouping was the second highest category at 28%. However, it is important to note that these differences in percentages per publication are not significant.

The ‘breaking the law’ category was the third most discussed grouping across all publications. The Globe and Mail published 18% of its lede paragraphs on this topic, the Toronto Star 11%, and the National Post 18%. Both the Globe and Mail and the National Post discussed this topic an equal amount with each publication focusing 18% of its coverage on this grouping. Again, only nonsignificant variations are observed.

Surprisingly the ‘terrorists/terrorism’ grouping was the focus of only a small proportion of lede paragraphs across all publications. The Globe and Mail and the Toronto Star each published 8% of their ledes on this topic, and the National Post published only slightly more with 12% in total. This variation is nonsignificant.
Lastly, the ‘other’ category was the least discussed topic across all publications. The *Globe and Mail* published only 8% of its lede paragraphs on this topic and the *National Post* 6%. The *Toronto Star* published no lede paragraphs on this topic at all. Similar to the previous categories, this variation is not significant.

In addition, when this chi-square test was conducted using SPSS the software indicated that 7 of the cells had an expected count of less than 5 which is larger than the minimum expected count of 0.75. Since 7 of the cells included values lower than 5 this indicates that the results of the chi-square test may be skewed. As stated previously, Vaughan argues that the “chi-square score is sensitive to the effect of small expected frequencies” therefore the chi-square score of \((X^2_{(8)}=7.684, \text{n.s.})\) may be exaggerated (88). Often in these cases the researcher will attempt to combine categories to remedy this problem (Vaughan 90). However, in this case the categories could not be merged because this would result in too large of an overlap between the 5 unique categories that were observed in lede paragraphs (Vaughan 90).
5 Discussion and Conclusion

The themes found in the lede paragraphs serve an agenda-setting function as discussed by McNair and Bryant and Oliver. McNair argues that the media performs the critical role of determining and disseminating the main stories that the public should be concerned about (29). Bryant and Oliver claim that the media set the agenda “so that an issue becomes the focus of public attention, thought, and perhaps even action” (1). In addition, Scheufele and Tewksbury define agenda-setting as “a strong correlation between the emphasis that mass media place on certain issues (e.g., based on relative placement or amount of coverage) and the importance attributed to these issues by mass audiences” (11). It can be argued that the themes observed in the lede paragraphs of the articles set the agenda regarding the main topics that Canadians should view as important concerning CSEC’s warrantless domestic collection of metadata.

The concept of agenda setting is particularly important and relevant to this study due to the finding that there is a significant relationship between the year that articles were published and the topics observed in their ledes. This finding suggests that over time the Canadian print news media set the agenda by emphasizing different issues concerning CSEC’s warrantless domestic surveillance activities. Therefore, if a reader were to follow this issue in the *National Post, the Globe and Mail*, or the *Toronto Star* they may view different topics as important over time from 2013-2016.

The themes observed in the lede paragraphs also serve a priming function as discussed by Scheufele and Tewksbury (11). These authors argue that priming occurs when the media “use specific issues as benchmarks for evaluating the performance of leaders and governments. It is often understood as an extension of agenda setting” (11). In addition, these authors claim that “mass media can also shape the considerations that people take into account when making judgements about political candidates or issues” by priming audiences (Scheufele and Tewksbury 11). The concept of priming is important to this study because it develops the argument that a reader could follow this issue in the Canadian print news media and develop benchmarks for evaluating CSEC’s surveillance activities by simply reading the headlines and the lede paragraphs of the articles. Since a
significant relationship was found between the years that lede paragraphs were published and the topics observed in the ledes, it can be argued that over time the Canadian print news media use different issues as benchmarks for evaluating CSEC’s warrantless domestic surveillance from 2013-2016.

If a Canadian reader were to follow this issue in any of the three publications they would read that the discussion surrounding the Canadian government’s warrantless domestic collection of metadata is focused on four core themes. The two most prominent themes concerning transparency and oversight related issues as well as government surveillance that is scary, about you, or relates to spying. The breaking the law theme as well as the theme concerning terrorists/terrorism are represented in a much smaller percentage of the total lede paragraphs. Based on these percentages it can be argued that the Canadian print news media has set the agenda on this issue by primarily discussing both the transparency and oversight of CSEC as well as CSEC’s surveillance activities which the ledes claim are concerning, directed at citizens, and as a result equates to spying (see fig. 13). The average reader may be less concerned that CSEC had broken privacy laws by conducting its metadata enabled surveillance and even less concerned about terrorist or terrorism related issues due to the low percentage of the 91 articles which discuss these themes in the lede paragraphs (see fig. 13). In relation to agenda setting, this suggests that the Canadian print news media viewed CSEC breaking the law as well as terrorist/terrorism related issues as being less important than the former categories.

In 2013 the Canadian print news media set the agenda on this issue by focusing the majority of its reporting on transparency/oversight related concerns as well as government surveillance that is scary, about you, and consists of spying. If a reader were to follow this issue in 2013 they may be concerned about CSEC’s warrantless domestic surveillance due to the high percentage of lede paragraphs for the year that began by discussing the scary/about you/spying theme (see fig. 14); 60% of the ledes in 2013 discuss the scary/about you/spying theme (see fig. 14). Concurrently readers may also have been alarmed about transparency/oversight related concerns in 2013 due to 30% of the lede paragraphs for the year beginning by discussing this theme (see fig. 14). However, readers may view the transparency and oversight of CSEC as being less
important than the agency’s surveillance activates, in this year, since many articles argued in the lede paragraphs that the surveillance was scary, about citizens, and could be viewed as spying.

However as time progresses the focus on this issue changes as seen in the 2014-2015 year segment where the largest percentage of lede paragraphs discuss the transparency/oversight theme; 38.7% of the ledes for the year discuss this theme (see fig. 14). Most importantly, the scary/about you/spying theme is only represented in 29% of the lede paragraphs in 2014-2015 (see fig. 14). In relation to agenda setting, this suggests that in this year segment readers may become more concerned with the apparent lack of transparency and oversight of CSEC and less concerned that CSEC’s surveillance which is scary/about you/or results in spying on Canadian citizens. Based on these findings, it can be argued that from 2014-2015 the Canadian print news media set the agenda on this issue by viewing the transparency and oversight of CSEC as being more important than any other topic. This is a valuable finding because it suggests that over time the Canadian print news media slowly began to become more concerned about the secrecy that surrounds CSEC and its scarce oversight mechanism.

Lastly in 2016, an interesting observation is seen where the highest percentage of lede paragraphs for the year begin by discussing transparency/oversight related issues with 46.7% of the ledes discussing this theme (see fig. 14). Interestingly, the second largest thematic grouping for lede paragraphs published in 2016 consists of ledes which discuss government security agencies breaking the law (30% of the ledes for the year discuss this theme) (see fig. 14). In relation to agenda setting, this suggests that readers who follow this issue in 2016 may be more concerned with CSEC’s transparency/oversight related problems as well as the idea that our national security agencies have broken laws.

It is also important to note that the scary/about you/spying thematic category is only represented in 10% of the lede paragraphs in 2016 (see fig. 14). This shows that as the years progress the ledes set the agenda by focusing more on transparency/oversight related problems concerning CSEC and less on CSEC’s surveillance that was claimed to be scary, about citizens, and can be viewed as spying.
Furthermore, in 2016 the breaking the law theme is represented in a far greater percentage of lede paragraphs than any other year (see fig. 14). This may be an indication that in 2016, the Canadian print news media set the agenda on this issue by zeroing in on the illegal acts that the lede paragraphs claimed Canada’s national security agencies had committed. Since the Canadian print news media focused such a large percentage of its lede paragraphs in 2016 on CSEC breaking the law, this suggests that readers who followed this issue may view CSEC’s illegal activities as being more important in 2016 than any other year.

Several critical issues were introduced in the lede paragraphs of the articles from 2013-2016 that could be used as benchmarks for evaluating CSEC’s warrantless domestic surveillance activities (Scheufele and Tewskbury 11). The 2013 Snowden disclosures ignited the discussion in the Canadian print news media which questioned CSEC’s involvement in the NSA’s international surveillance activities. Another important issue was introduced in 2013 where one article argued in its lede paragraph that taxpayers fund CSEC’s annual budget, of close to 400 million dollars, therefore the lede claimed that the public has a right to know more about how CSEC conducts its surveillance. In 2014, many of the articles argued in their lede paragraphs that CSEC was found to have collected information from Canadians as well as from other travelers who accessed the Wi-Fi at Pearson International airport. Furthermore, the articles also argued in their lede paragraphs that the devices of travelers were tracked over a two week period by CSEC. Lastly, in 2016 many of the articles began by arguing in the lede paragraphs that CSEC had collected and shared Canadian metadata with the Five Eyes network for many years before informing the public. These concerns that were introduced in the lede paragraphs of the articles primed audiences by suggesting annual benchmarks that could be used to measure CSEC’s performance as a security agency from 2013-2016.

A reader could follow this issue in any of the three publications and be exposed to the same agenda setting and priming influences that were previously discussed. This is due to the fact that no significant relationship was found between the topics observed in the lede paragraphs of the articles and the publication in which lede paragraphs were published.
This analysis of articles in the Canadian print news media that address metadata collection and use demonstrates that Canadians will encounter a range of perspectives concerning the coverage of metadata regardless of the news publication that they read. When analyzing how the Canadian government’s warrantless collection of citizen metadata has been represented in the Canadian print news media it is clear that this issue is incredibly complex. This thesis acts as one of the first steps in shedding light on what Canadians are told about Canada’s warrantless domestic collection of metadata since the 2013 Snowden disclosures. This thesis combines both qualitative and quantitative forms of content analysis as a research method to analyze this topic, yet additional techniques could have been used to glean further insights from the data such as discourse analysis, linguistic analysis, or more complex forms of statistical analysis.

One limitation of this project is that it does not distinguish between the different types of articles that were included in this study such as editorials, opinion pieces, news reports, or columns. Another limitation is that only the lede paragraphs of the articles were thematically organized instead of also analyzing the entire body of the articles. This project also did not analyze the relationship between the authors of the articles and the thematic categories that were observed in the lede paragraphs. Lastly the final sample size of the articles was small and could have been increased to include other Canadian print news publications. Future research could address these problems by analyzing articles by their individual types, searching the entire body of the text for thematic categories, and looking for relationships between the author and the themes that are inductively observed.
6 Bibliography


Bryant, Jennings, and Mary B. Oliver. Media effects: advances in theory and research, Routledge, 2009.


Appendix A: Newspaper Articles.

The Globe and Mail


The Toronto Star


The National Post


Curriculum Vitae

Name: Alan Del Pino

Post-secondary Education and Degrees:

University of Western Ontario
London, Ontario, Canada
2015-2017 M.A

University of Guelph-Humber
Etobicoke, Ontario, Canada
2010-2015 B.A.

Humber College
Etobicoke, Ontario, Canada
2013-2015 Media Communications Diploma

Humber College
Etobicoke, Ontario, Canada
2010-2013 Public Relations Advanced Diploma

Honours and Awards:

Province of Ontario Graduate Scholarship
2015-2016

University of Guelph-Humber Merit Scholarship
2014

Related Work Experience

Teaching Assistant
The University of Western Ontario
2015-2017

Research Assistant
The University of Western Ontario
December 2016

Publications: