

Prompt Used to Generate the Report

You are an academic researcher producing a polished, professionally formatted, white paper analyzing a recent Government of Canada public consultation. The length of the report can be anywhere between 15-30 pages assuming single spacing and minimal whitespace between major sections of the paper (i.e. no page breaks between major sections).

I have providing two Government of Canada documents that form the core of the dataset used to perform your analysis: 1) A “What We Heard” report; and 2) an excel data file containing the raw data of more than 11,000 public responses the government of Canada gathered as part of their public consultation on the next version of Canada’s National AI Strategy. The “What We Heard” report is divided into two main analysis sections, the first section, which is titled “Online Consultation” (starting on page 7 of the report), is based on the public responses (contained in the excel file). The second section, titled “Task Force Reports” (starting on page 10 of the report) is based on reports that I have not provided, which were submitted by an “expert task force”. In this task you are to focus only on the content of the “What We Heard” report that focuses exclusively on the public response data, in other words focus only on statements and claims made about the public consultation data. Do not analyze information related to the “Task Force Reports”.

The excel file is formatted as follows: the questions that the Government of Canada posed to the public are contained in a header row, and the individual responses (just over 11,000 of them) are contained in the subsequent rows. In other words, each column contains all the individual answers to a specific question. Each blank cell in the table represents an answer that an individual did not provide for that question.

Generally speaking, you should perform the analysis using accepted methods in the social sciences for coding and analyzing quantitative and qualitative data such as is included in the raw data. You should also use sentiment analysis to better characterize the intent of the respondents when providing their responses.

Your report must contain the following sections and information in those sections as described:

Section 1: Demographic Breakdown

In this section, provide a detailed demographic analysis based on the first several columns in the excel file. Those include age, sector information (academic, business), and other information.

Read the article and background: <https://profjasonmillar.substack.com/p/the-government-of-canada-falls-short>

Section 2: Accuracy of the Government of Canada's Claims

In this section, I am trying to understand how accurately the government's analysis of the public response data represents that data. Your job is to:

- 1) Analyze both documents and assess the accuracy of the various claims and findings contained in the "What We Heard" report against the raw data (contained in the excel file) they used to produce the key findings, proposals, and other key takeaways.
 - a. Provide a deep analysis of the accuracy of the claims as they relate to the public response data. In doing so, indicate whether the claims are accurate, exaggerated, or understated based on your analysis of the frequency of supporting evidence in the raw data. Provide statistical information explaining your findings.
- 2) Extract and provide a list of "Top 10 Themes" that you feel accurately reflects the public response data. Provide statistical information explaining your findings.

Section 3: Accuracy of Specific Claims

In this section, I am trying to understand the accuracy of specific claims the government of Canada makes in the "What We Heard" report. The specific claims are:

- 1) On page 5 of the report, the government claims that their methodology was applied "ensuring accuracy and comprehensive representation of all perspectives."
 - a. Question for accuracy analysis:
 - i. Does their stated methodology provide enough evidence to support the claim that it ensured accuracy and comprehensive representation of all perspectives?
- 2) On page 5 of the report, the government claims that their methodology resulted in "thorough, unbiased" reporting.
 - a. Question for accuracy analysis:
 - i. Based on your results in Section 2, was their reporting "thorough and unbiased"?
- 3) On page 16 of the report, the government claims that "The methodology, data science and use of AI in this project are all aligned with and conform to the Treasury Board of Canada Secretariat's guidance on the responsible use of artificial intelligence in government". I have provided links to the full "guidance on the responsible use of artificial intelligence in government" below. Use it to answer the following questions.
 - a. Questions for accuracy analysis:

Read the article and background: <https://profjasonmillar.substack.com/p/the-government-of-canada-falls-short>

- i. Based on your reading of the “guidance on the responsible use of artificial intelligence in government”, how well does the government of Canada’s “What We Heard” report align with that guidance?
- ii. Where is it particularly well aligned and where does it seem to fall short?

Links to the full “guidance on the responsible use of artificial intelligence in government”:

Here is some critical background information on the Government of Canada’s official guidance on how to use generative AI when conducting government business. This document is critically important for the analysis because the Government of Canada claims that they completed the “What We Heard” report in compliance with their “Guide on the use of generative AI”.

- This document is the overarching “Guidance Document” on using generative AI when conducting government of Canada work. It is meant to set out guiding principles and help government employees understand how to categorize uses as relatively low risk versus relatively high risk. The document can be found here: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html>
- There are 12 Government of Canada principles that are meant to guide the effective and ethical use of AI in the federal government. They can be found in the list at the top of this page: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/principles.html>
- This government of Canada website spells out the various dos and don’ts on how to use generative AI at work. Of particular interest is the table listing principles, and corresponding dos and don’ts. The website is here: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/generative-ai-your-daily-work.html>

Section 4: Reverse Engineering the Government of Canada’s Analysis

In this section I would like you to reverse engineer the AI prompts that the Government of Canada might have used to generate the results that they report in the “What We Heard” report. They claim that they used generative AI to analyze the raw public response data, and they report on which models they used, but they don’t provide information about the AI prompts they used. Because they do not provide the prompts in their report, it is important to try to understand what prompts might have led to the output they report.

In preparing this section, iteratively test and retest a set of prompts, and apply them to the raw public response data in the excel file, until you get output that produces an analysis

Read the article and background: <https://profjasonmillar.substack.com/p/the-government-of-canada-falls-short>

supporting the various key findings and claims made in the “What We Heard” report (excluding, of course, the information related to the expert task force). This might take a while, but I really want you to get as close as you can. The goal of this section, in particular, is to try to determine what prompting biases could lead to any discrepancies you found in Section 2.

Report on the full set of resulting prompts. Also report in detail on the methodology you use to produce them.

Section 5: Methodology

In this section, provide a detailed explanation of your methodology, including any weaknesses contained in the analysis that arose due to the quality of the data. Include a detailed explanation of the AI model(s) you used in performing all of the work. Include detailed information about how you performed the sentiment analysis. Throughout this section, provide sample responses alongside brief descriptions of what your analysis told you about those particular responses (e.g. this response was typical of an answer supporting the call for regulations, or this response demonstrates such and such a category used as part of the sentiment analysis).

Section 6: World Cloud

Produce a nice-looking word cloud, similar in look and feel to that in the “What We Heard” report, which is representative of the excel document raw data according to your analysis of key themes and respondent intent.