

Questions and Answers

Question	Response/Reaction
1. What is DIP?	DIP stands for Data Inventory Project. It refers to a project being carried out by the Policy Research Data Group [PRDG] to collect information about the extent and nature of data holdings within government departments. It will also be used to provide input for future endeavors and information for the Treasury Board Secretariat (TBS) data catalogue and their Open Data Pilot Project.
2. What is the purpose of DIP?	<p>The DIP is the critical first step in the development of a government-wide inventory of data holdings. The overall objective is to develop an inventory of datasets which may be used to support policy and program development and the work of Statistics Canada and the TBS.</p> <p>Although the government of Canada produces thousands of data sets on a broad variety of topics including environmental, scientific, geospatial, statistical and financial data, there is no single inventory of these data sets that can be used to determine which data are available to support policies and programs and what data gaps may exist.</p>
3. How will information from DIP be used by the Policy Research Data Group (PRDG) and others?	The metadata from DIP will be used to determine the availability of data to support policy and program development needs. Comparing data needs to data available from departments will provide the basis for identifying data needs and eventually data gaps for policy purposes. The current initiative focuses solely on the development of a data inventory.
4. How will the DIP information be collected?	Statistics Canada's Special Surveys Division will be adapting its standard electronic data collection tool to capture the inventory information from departments. It will be deployed to departments in much the same way as the Public Service Employees' Opinion Survey was collected in the past.
5. What is a dataset?	A dataset is a single file containing information that is produced by a program or project, or is obtained through a survey or administrative source. Thus, a dataset is any organized collection of data or information that has a common theme. A dataset is a single file saved on a memory storage device such as a server or hard drive. A single dataset can be broken into multiple parts, but if it's within the same file then it's a single dataset, otherwise it would be a master-file and

	sub-sets.
6. What is the difference between a dataset and a database?	A dataset is a single file containing information that is produced by a program or project, or is obtained through a survey or administrative source. A database is a collection of data organized according to a conceptual structure describing the characteristics of these data and the relationships among their corresponding entities. Databases containing information related to key policy issues are in the scope of DIP.
7. How would the questionnaire be completed for a database as opposed to a dataset.	Each case will be different. Some departments may decide that reporting a database as a single dataset makes sense while another database in another department may be better reported as a group of datasets. Please contact the focal point (coordinator) from your department if you need assistance.
8. Which datasets should departments include in the DIP?	DIP seeks to identify all datasets related to the departments' programs and activities. Datasets that are considered in-scope are those datasets that are being used to address or can be used to address your key departmental objectives. Even if a dataset is obtained from another government department (for example: survey data collected by Statistics Canada), your department should still report on it.
9. Which datasets should be excluded?	Databases related to the department's personnel and financial activities or their facilities should not be included. Similarly catalogues describing their library's holdings and commercial databases should not be included unless these catalogs specifically referred to the data holdings of the department the address key policy issues.
10. Should datasets from Statistics Canada that are held by departments be included?	Yes. ALL data holdings within a department are in scope if they relate to one of the key policy issues. The reason for doing so is to enable the Deputy Heads to know which data are held within the Department. This should include public use microdata files, share files and special tabulations.
11. How long is the DIP questionnaire?	The DIP questionnaire contains a maximum of 57 questions. However, most data sets can be described using 25 to 30 questions.

<p>12. To whom will the questionnaires be sent?</p>	<p>Each department will have a designated co-coordinator who will be in contact with the survey team. The coordinator will be responsible for forwarding questionnaires to respondents in the departments and relaying questions to the help desk or survey team.</p>
<p>13. If you are providing us with an Excel based collection tool, can we just populate that table and send it in instead of completing the questionnaire?</p>	<p>No. The Excel collection tool is not meant to replace the questionnaire. The collection tool was created in order to help responding departments begin consolidating the information prior to collection.</p>
<p>14. Can we simply send you existing spreadsheets of the information rather than manually entering in the questionnaire?</p>	<p>No. However any inventories that already exist in your department can be used to populate the Excel based collection tool, and complete the e-questionnaire directly.</p>
<p>15. The collection period seems short for the amount of information the project is looking for. Is there any flexibility in the collection deadline?</p>	<p>We understand that departments may have very large data holdings and that completing the questionnaire will be time consuming. Collection is planned from the beginning of April until the end of May. We will be providing all respondents with an excel based collection tool that will allow them to begin recording the information on their data holdings prior to collection.</p>
<p>16. Are the Data Inventory Project (DIP) and the Open Data Project the same?</p>	<p>No. DIP is a separate project being undertaken by the PRDG and the Open Data Initiative is a Treasury Board Secretariat project. Certainly there are similarities between the two projects; however, they are separate undertakings.</p>
<p>17. How do we group datasets?</p>	<p>In general they should be grouped according to a logical theme (i.e. same data over several years, same data for different geographic areas, data that is all collected for the same project with the same goal in mind). The subject matter experts in your department would be in the best position to determine if the grouping is logical or not.</p>

<p>18. How will the information contained in the inventory be protected?</p>	<p>One of the last questions of the questionnaire, asks “Who should be allowed to access the information you provided in this questionnaire”. Once all the data has been collected, a list of each dataset and the response to this question will be established. The deputy minister for each department will be asked to vet this list which will eventually determine the availability of the inventory to other departments and the general public.</p>
<p>19. What is the reference period for the datasets</p>	<p>The reference period of the dataset or database is the time period for which the information applies. If a database maintains records back to 1975 and is still being updated then the reference period is 1975 to present.</p>
<p>20. Are old datasets which are no longer used considered in scope?</p>	<p>Yes, all datasets should be considered in-scope. If the datasets were collected and used to address key policy issues then for historical purposes they are in-scope for the survey. If your department has determined that they are important enough to keep and save, then they are probably important enough to be reported on. Priority should be given to new and current datasets, but the older datasets where ever possible/feasible should also be included.</p>
<p>21. Our department creates reports based on datasets. Are these reports in scope?</p>	<p>Even if such reports were considered by a department to fall under the definition of a dataset, these reports would be grouped with the datasets they were derived from. If the reports are of importance to the project, program or survey more detail about the types of files available can be included in the abstract question. However, if the reports are only administrative, just as a list of names of the people who have accessed the dataset, then they can be excluded.</p>
<p>22. Should data holdings held by departmental libraries be included in DIP?</p>	<p>No. Library holdings of commercial databases should not be included. If the library catalogues include information on departmental data holdings then they could be useful in helping you complete the survey.</p>
<p>23. Can I use acronyms and abbreviations that are common to my department?</p>	<p>As this inventory is meant to be utilized by several departments it is best to avoid acronyms and abbreviations and instead write out the full definition/title.</p>
<p>24. Can we report on an incomplete dataset?</p>	<p>Yes, just enter in all the relevant information that is available.</p>

<p>25. What if I encounter technical difficulties with the Internet-based survey?</p>	<p>If a respondent has any technical questions or requires assistance, they should access the on-line help pages or contact Statistics Canada at the following email address sos@statcan.gc.ca or call the assistance line at 1-877-949-9492.</p>
<p>26. Where can I find out more information about the survey?</p>	<p>You can find more information about the survey at www.gcpedia.gc.ca/wiki/Data_Inventory_Project and at http://stdsimdb/cgi-bin/imdb_reports/p3Instr.pl?Function=assembleInstr&Item_Id=122361&lang=en&db=imdb&adm=0&dis=0</p>
<p>27. What is the security of the information that we are sending to you over the internet?</p>	<p>For more information on the security of Statistics Canada electronic questionnaires please visit : https://www68.statcan.gc.ca/eq-ge/en/login-connexion</p>
<p>28. This questionnaire is different from the one we were shown in February and early March.</p>	<p>Yes it is. In the interest of reducing the response burden of departments and increasing the quality and quantity of information provided in the questionnaire we have reviewed the content and removed and reordered several questions.</p>