“Safe Data” – Statistics Canada’s Data Confidentiality Classification Tool

Tiered Access Workshop
Council of Professional Associations on Federal Statistics
January 30, 2020

Presented by:

Steven Thomas

Chief Methodologist for the Center for Confidentiality and Access

Statistics Canada

Tiered Access at Statistics Canada

- Statistics Canada recognizes that a one-size-fits-all access solution unnecessarily limits researcher access to valuable data resources
- Building on the ‘Five Safes Framework’, there are 5 recognized levers to ensure safe access solutions
- Recognizing that not all data has the same confidentiality concerns allows access under a variety of different access scenarios
- Measuring the confidentiality of datasets allows disclosure risks to be mitigated appropriately

CCT: Contextual Elements

Previous view on confidentiality

- Black and White view: Information was either confidential or it wasn't
- Information was made public or protected within 'safe settings' (Headquarters or Research data Centers)
- Greatly influenced how Statcan employees worked and how the agency's information holdings were managed

CCT: Contextual Elements

Current View

- User-centric
  - Users have the information/data they need, when they need it, in the way they want to access it, with the tools and knowledge to make full use of it.
- Modern Workforce
  - enable employees to work in a way that enhances productivity and collaboration

Challenge

Identify and understand the risks associated with data assets in order to take calculated risks with data access

CCT: Contextual Elements

- The Confidentiality Classification Tool (CCT)
  - Self-administered assessment tool to be aware of the confidential nature of our data holdings
- Operationally Feasible
  - Simple and lean tool - Intended for use by statistical program practitioners
- Risk Management
  - Understand the probability and severity of a breach in confidentiality associated with a particular dataset

CCT: Disclosure Risk and Sensitivity

- Probability considerations:
  - Disclosure risk assessment through four types of disclosure
  - Identity, Attribute, Inferential and Residual
- Severity considerations
  - Sensitivity assessment
  - Adapted from levels from the Office of the Information and Privacy Commissioner of Ontario
- These assessments are combined to arrive at a score that determines the level of confidentiality

CCT: Preview – Disclosure Risk

Disclosure Risk Self-Assessment Checklist			
Purpose:	Provide a self-assessment of the disclosure risk associated with the product under examination.		
Instructions:	Please answer all questions below and provide an explanation for your chosen ratings.		
	Please refer to the short Guide in Step1 for more detailed instructions and examples of how to fill out this questionnaire.		
Section 2.1: General Information			
Date completed:		Completed by (name):	
Class of product (GSBPM)		Product name:	
Division or FRC:		SDDS number: (where applicable)	
Section 2.2: Disclosure Risk Questionnaire			
Disclosure Type	**Question**	**Choose the rating that best applies**	**Comment/Explanation/Assumptions**
Identity Disclosure			
	Can confidential information be revealed directly through the information displayed by the product?	○ LOW risk as SIGNIFICANT effort is required ○ MEDIUM risk as SOME effort is required ○ HIGH risk as LITTLE effort is required	
Attribute Disclosure			
	Can confidential information be revealed by grouping attributes available in the product?	○ LOW risk as SIGNIFICANT effort is required ○ MEDIUM risk as SOME effort is required ○ HIGH risk as LITTLE effort is required	
Inferential Disclosure			
	Can confidential information be revealed through a probabilistic statement derived from the information contained in the product?	○ LOW risk as SIGNIFICANT effort is required ○ MEDIUM risk as SOME effort is required ○ HIGH risk as LITTLE effort is required	
Residual Disclosure			
	Can confidential information be revealed by combining the information contained in this product with that from other products?	○ LOW risk as SIGNIFICANT effort is required ○ MEDIUM risk as SOME effort is required ○ HIGH risk as LITTLE effort is required	

CCT: Preview - Sensitivity

Sensitivity Self-Assessment Checklist			
Purpose:	**To assess the level of sensitivity of the specified product based on the impact a breach would have on an individual, business or institution.**		
Instructions:	**Please complete all questions below.**		
	Please provide explanations or assumptions for your rating assessment.		
	For more detailed instructions, please refer to the Guide in Step1 of this Spreadsheet.		
Section 3.1: General Information			
Date completed:		Completed by: (name)	
Class of product (GSBPM)		Product name:	
Division or FRC:		SDDS number: (where applicable)	
Section 3.2: Sensitivity Question			
Please Assess the Impact on Respondents of any Disclosure		**Rating**	**Comment/Explanation/Assumptions**
Select the **highest** level of impact a breach would have on an individual, business, or institution from the Sensitivity Rating Legend below, and provide a justification for why you believe this is the case. The impact must reflect an outcome that can reasonably be expected to occur, rather than one based on an extreme scenario.		○ Severe (5) ○ High (4) ○ Medium (3) ○ Low (2) ○ Negligible (1)	
Sensitivity Rating Legend			
Would the breach of the information in an identifiable format:			
Severe (5)	...cause **severe harm** to an individual or business?		
High (4)	...cause **considerable harm** to an individual or business?		
Medium (3)	...cause **reputational damage or embarrassment** to an individual or to a business?		
Low (2)	...cause **minimal harm** to an individual or business?		
Negligible (1)	...**not cause any harm** as the information is considered to be publicly available?		

Disclosure Risk Matrix

Sensitivity (Impact scale)											
	5	0	2	4	5	5	6	7	7	8	9	9
	4	0	2	3	4	5	5	6	6	7	8	8
	3	0	1	2	3	4	5	5	6	6	6	**7**
	2	0	1	1	2	3	4	4	4	5	5	5
	1	0	1	1	1	2	2	2	3	3	3	4
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
		Potential Disclosure Risk										

CCT: Governance

Data Custodians:

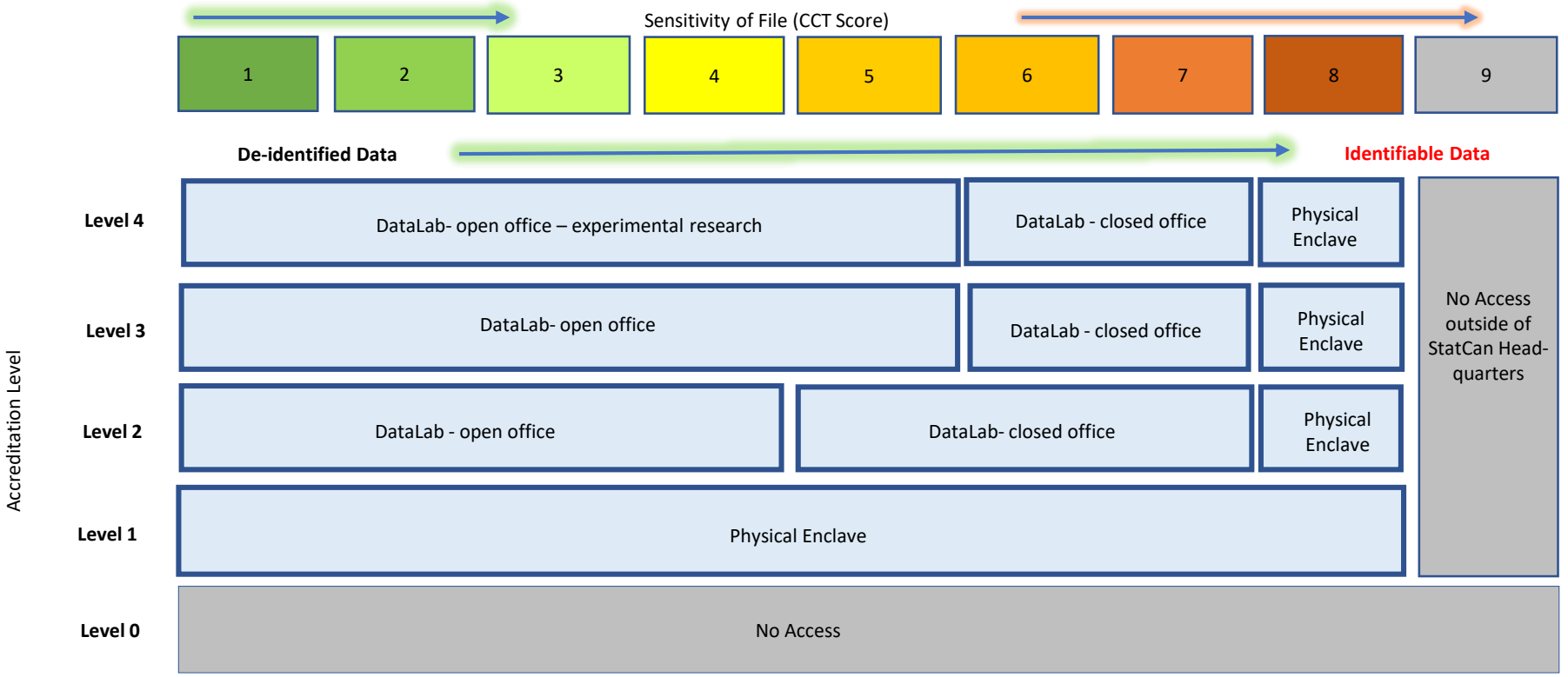
- Perform assessment
- Take steps to reduce risks if possible / necessary

Directors:

- Sign-off on final assessment
- Sign-off on access options

- Review committee:
  - Monitor scores obtained
  - Standardize practices
  - Hear appeals / special requests

Prototype Access Framework – Concordance Table



The highest level of CCT score of the datasets associated with a project dictate the mode and location of access

Location:
Open office = cubicle environment
Closed office = personal office or conference room
Physical Enclave = Designated Certified Room (DataLab/RDC/FRDC)

Questions or Comments

To download the CCT:

https://open.canada.ca/ckan/en/dataset/2c910c37-c684-561e-9e0b-1d5bb6ca5fb9

Contact:

Steven.Thomas@canada.ca