

# Coding Protocol for Third-Party Annotators

Shabnam, Mehrdad, Daniel

April 2022

## 0.1 Protocol

1. SFCR includes information on General food requirements and guidance and sets out many requirements that apply across food commodities, along with food-specific requirements. Therefore, SFCR sets the overall context and framework for food safety. Reviewing SFCR, a set of codes and categories are induced. However, various regulations elaborate additional and more specific requirements for specific food products. A variety of food products are covered by the Food-specific requirements and guidance (FSRG), including dairy products, egg and processed egg products, fish, fruits and vegetables, meat and food animals, honey, icewine and maple syrup. Information in (FSRG) related to specific foods are not all covered by SFCR solely. However, there are numerous FSRGs and we are interested in the ones which refine or complement the already derived codes from SFCR. The principals are, being selective and achieving a high level of saturation. Therefore, based on the initial screening of documents we decided to exclude the last category from further investigation, on the basis of being a niche area that is country-specific and unlikely to be generalized. Also, based on the intuition that handling meat products and food animals or egg and processed egg products are more involved than the rest we limit the scope of open coding to these parts.

Moreover, like all legal texts, FSRG and SFCR have several cross-references that include intertwining content with other documents e.g., SFCR enabling acts e.g., Safe Food for Canadians Act (SFCA), Food and Drug Regulations (FDR) and Food and Drug Act (FDA). Cross-references we saw in those FSRGs and SFCRs prompted our investigation of FDR, FDA, and SFCR enabling acts.

2. You should not follow cross-references as a subject of processing although there are previous works that identify cross-references intent and the way to navigate them. There are provisions that require to comply with some cited provisions or there are exceptions to the previous provisions, in these cases the information content of the statement should imply whether a label will be assigned to a statement. For instance, “217 A prepackaged food must have a label that meets the requirements of these Regulations

applied or attached to it in the manner set out in these Regulations.” You don’t follow the content of “these regulations”.

3. We have a classification where requirements can be of types of data, measurement, time constraint, and convention.
4. According to our coding procedure, a unit of analysis for annotation purposes is sentenced level. These sentences are determined by the sentence-splitter and might not always be grammatically correct since the sentence-splitter (spacy) uses its dependency parser to divide texts into sentences. Moreover, you should look at context as in provisions with nested bullets, the preface may impact the sub-bullets. In other words, in the case of interrelated itemized sentences, you should check the context in order not to miss any potential label due to only considering the genuine meaning of the itemized sentence. Thus, you should read the preface and decide whether the sub-bullets add new data, measurement, time constraint, or convention.
5. Data: This label applies to statements of type data definition, data collection, or data retention.
6. Data definition: Characterizes the information i.e., the statement is labelled as data definition when it defines information. It defines what data, the labels must bear. In most cases where there are phrases e.g., “must be labeled/graded with”, “label must bear”, “must be shown on”, “label/grade must be attached/applied”, “are required to be labelled with”, “must be shown as”, “shall be located on” they elaborate on the expressions, mark or stamp that must or may be shown on the label, grade or on the container of a food product. A few examples are listed below:

“254 The label of prepackaged eggs that are graded in accordance with these Regulations must bear

  - (a) a declaration of net quantity; and
  - (b) in the case of eggs that are pasteurized in the shell, the words “Pasteurized” and “pasteurisé”, as well as the expressions “Graded Canada A Before Pasteurization” and “classifié Canada A avant pasteurisation” or the expressions “Graded Grade A Before Pasteurization” and “classifié catégorie A avant pasteurisation”, as the case may be.”

“(3) The relative firmness of the cheese must be identified by the following expressions:

  - (a) “Soft White Cheese” and “fromage à pâte fraîche” or “fromage frais”, if it has a moisture on fat-free basis content of 80% or more;
  - (b) “Soft Cheese” and “fromage à pâte molle”, if it has a moisture on fat-free basis content of more than 67% but less than 80%;

“321 Fresh fruits or vegetables that are sent or conveyed from one province to another or imported must be labelled with the applicable size designation that is set out in the Compendium, if any. The size designation must

- (a) be shown in close proximity to the grade name;
- (b) in the case of prepackaged fresh fruits or vegetables, other than consumer prepackaged fresh fruits or vegetables,
  - (i) if their container is a reusable plastic container, be shown in characters that are at least 1.6 mm in height, or
  - (ii) if their container is not a reusable plastic container, be shown in characters of at least the minimum character height that is set out in paragraph 320(1)(b) for the grade name”

As an example, food products must be labeled with the relevant size designation, and that is a data definition since we know what size designation means. However, in the following example, the “descriptive term” doesn’t define any data since we don’t know what information it characterizes. “(3) The descriptive term referred to in paragraph (1)(e) must be shown in close proximity to the common name and in characters that are at least the height that is the greater of

- (a) one-half the height of the characters in which the common name is shown, and
- (b) 1.6 mm.”

“233 (1) The metric units that must be shown in a declaration of net quantity of a consumer prepackaged food must be in

- (a) millilitres, if the net volume of the food is less than 1 000 mL;
- (b) litres, if the net volume of the food is 1 000 mL or more;
- (c) grams, if the net weight of the food is less than 1 000 g; and
- (d) kilograms, if the net weight of the food is 1 000 g or more.”

This example shows that the preface is used in conjunction with the bullet (a) to define the data, where the metric units must be shown in milliliters, but other bullets are simply different formats of this data. Also, in the below example, the preface, along with bullet (a), defines quantity, while bullet (b) refers to a different format (decimal system) of the data. “235 If the declaration of net quantity of a consumer prepackaged food is shown in metric units and the quantity is less than one metric unit, the quantity must be shown

- (a) in words; or
- (b) in the decimal system, with a zero preceding the decimal point.”

“231 The declaration of net quantity of a consumer prepackaged food must be shown

- (a) in the case of a consumer prepackaged food that is listed in the document entitled Units of Measurement for the Net Quantity Declaration of Certain Foods, prepared by the Agency and published on its website, as amended from time to time, by volume, weight or numerical count in accordance with that document; or
- (b) in the case of a consumer prepackaged food that is not listed in the document referred to in paragraph (a),
  - (i) if the food is liquid, gas or viscous, by food volume, or, if the food is solid, by weight, or
  - (ii) if the established trade practice in respect of the food is to show its net quantity in a manner that is different than what is required by subparagraph (i), in accordance with that established trade practice.”

In this example, the declaration of net quantity must be shown by volume, weight, numerical count, or in accordance with some cross-references. Data definition is the first part of defining net quantity; however, we do not consider the following bullets to be data definitions.

“(2) If prepackaged fresh fruits or vegetables that are labelled in accordance with this Part are placed inside of a second container and the resulting product is prepackaged fresh fruits or vegetables, other than consumer prepackaged fresh fruits or vegetables, the second container is not required to be labelled with the information referred to in subsection (1) if that information is readily discernible and legible without having to open the second container and that information is not obscured by the container.”

In this example, there is a cross-reference needed in order to understand what the “information” is. More importantly, the compliance here, “is not required to be labelled with” is permission, not a requirement.

In requirements such as the two mentioned below, data is not explicitly defined. To define data, we cannot use generic phrases, such as “that information” or “the statement”.

“(2) That information must be shown in characters that are in the case of a tray with an overwrap or an egg carton, on the top or side of the tray or egg carton, at least 1.5 mm in height;”

“(2) The statement must be shown on or in close proximity to the pictorial representation, if the representation is shown on the principal display panel;”

7. Data collection: The statement is labeled as Data collection when it needs a document to be recorded. For instance:

“(2) If a complaint is received, the operator must implement the procedure and prepare a document that sets out the details of the complaint, the

results of the investigation and the actions taken based on those results and keep it for two years after the day on which the actions are completed.”

“The grading certificate must be signed by the grader and include the following information:

- (a) the name and address of the producer;
- (b) the name of any person who is acting on behalf of the producer;”

“98 (1) Despite subsection 306(1), a licence holder may import ungraded eggs if they before the import, notify the Minister in writing of the quantity of ungraded eggs that are intended to be imported, the date of the import and the name of the licence holder and address of the establishment referred to in paragraph (c);”

“(2) Any person who sells a food at retail, other than a restaurant or other similar enterprise that sells the food as a meal or snack, must prepare and keep documents that include the information specified in paragraphs (1)(a), (c) and (d).”

8. Data Retention: A statement is labeled data retention when a document needs to be maintained, such as, “(3) The documents referred to in subsections (1) and (2) must be kept for two years after the day on which the food was provided to another person or sold at retail, and must be accessible in Canada.”
9. Sensing/Masurement: Measurements is the association of numbers with physical quantities, characteristics and phenomena such as temperature, dimension measures, weight, numerical count, color, firmness, PH level, fat content, water level, bacteria concentration, test, pressure, and humidity. You need to specify which measurement types you found. If a new measurement type is found in a statement, please write it in the column named “others”.

“The thawing procedures must prevent net gain in weight over the frozen weight, when poultry products are thawed for repackaging.” To prevent net weight gain in a specific phase, someone must assess and keep track of it.

“When water (including ice) is used for chilling and comes in contact with carcasses, carcass parts and giblets during chilling process, the water retention must be assessed.” The requirement explicitly asks for an assessment of water retention.

“The poultry product will be considered “frozen” and labelled accordingly when freezing becomes more extensive i.e. deeper than 4mm.” As an example, the product will be labelled based on its freezing depth.

“188 (1) The container of a consumer prepackaged food that is set out in column 1 of Table 1 of Schedule 3 must be of a size that corresponds to a net quantity by weight or by volume that is set out in column 2 or 3.”

It is necessary to measure the food product in this instance since its net quantity must match the container size.

“(2) Subject to subsection (3), if a lot contains the number of units set out in column 1 of Part 1 of Schedule 5, the inspector must collect from the lot at least the number of units set out in column 2 and the units collected constitute the sample referred to in subsection (1).” This example measures the number of units.

“(5) In the case of an inspection of a consumer prepackaged food that consists of a liquid, the net quantity of the food must be determined on the basis of the assumption that the liquid is at a temperature of 20°C.” This example implies that the net quantity must be adjusted and selected based on liquid temperature constraint; therefore, the liquid temperature must be measured as a condition.

“(6) In the case of an inspection of a consumer prepackaged food that consists of a frozen liquid food and that is normally sold and consumed in a frozen state, the net quantity of the food must be determined when the food is in a frozen state.”. When the food is in a particular state, the net quantity must be measured and determined.

“(3) For the purposes of subsection 6(1) of the Act, labelling a consumer prepackaged food with a declaration of net quantity does not constitute labelling a food in a manner that is false, misleading or deceptive if (b) the actual net quantity of the food is, subject to the tolerance provided under subsection (5), not less than its declared net quantity.” One can determine the falseness, misleadingness, and deception of the labelling by measuring the net quantity value.

“343 The percentage of the contents of a multi-ingredient food commodity that are organic products must be determined in accordance with CAN/CGSB-32.310.” In this example, “The percentage of the contents of a multi-ingredient food commodity” must be measured and determined.

“Product temperature requirements must be met at the time of shipping.”, “If the timeframe has elapsed, the FSIS inspector will be required to monitor the temperature of the load to verify compliance, as per CFR 590.950(b).”, and “48 (1) If a low-acid food is in a hermetically sealed package, an operator must apply the scheduled process referred to in subparagraph (3)(a)(i) and, if batch thermal treatment is applied, must use a temperature-sensitive indicator that visually indicates that the package has been thermally treated.” These requirements discuss temperature monitoring that requires measurements.

“The licence holder must submit to the veterinary inspector a written protocol for each all poultry products which will be crust frozen outlining the time period and location for the equilibrium of internal and external product temperatures such that an internal temperature is between 4°C and -2°C.”, “Use of the term “air chilled” or similar phrases must be restricted to carcasses or portions when licence holder can demonstrate

through a written PCP program and validation data that there is no net increase in the weight of the carcasses as a result of post evisceration washing, chilling and drainage as per “Poultry Retained Water Control Program”. As these two examples require measurement requirements to be written, they are also data collection requirements.

You should note that just detecting the measures or the units of measurements e.g., °C , mm, cm<sup>2</sup>, g/kg, mL/L, pink, white, red and so forth is not enough. For instance: “272 (1) The label of a prepackaged processed fruit or vegetable product must bear (f) the total percentage of sweetening ingredients added, if any, in the case of frozen fruits packaged in sugar, invert sugar, dextrose or glucose in dry form;” In this example, the data definition on the label may contain some metric units.

“(2) Despite paragraphs (1)(a) and (c), 500 mL may be shown as 0.5 L and 500 g may be shown as 0.5 kg.” and “(3) That information, other than the declaration of net quantity, may be shown in characters that are at least 0.8 mm in height if (a) the information that a label is required by Division 2 to bear is shown on the principal display panel” These statements outline the font and format of labels, which we later elaborate on as a convention.

10. Time constraint: Temporal restrictions or conditions. In this context, we have observed two main types of time constraints: interval/deadline and periodicity. Deadline: a point in time by which something must be done. Periodicity: the tendency, quality, or fact of recurring at regular intervals.

In the following samples, you’ll find examples of events that must or may occur in a fixed period of time or periodically.

Interval/deadline: “91 (1) Any person who has received a request from the Minister for a document referred to in section 90, or any part of such a document, must provide it to the Minister

- (a) within 24 hours after receipt of the request, or within,
  - (i) any shorter period that is specified by the Minister, if the Minister believes that it is necessary in order to identify or respond to a risk of injury to human health associated with a food commodity, or
  - (ii) any longer period that is specified by the Minister, if the Minister believes that the document is not necessary for a recall that is or may be ordered under subsection 19(1) of the Canadian Food Inspection Agency Act; and
- (b) if provided electronically, in a single file and in plain text that is capable of being imported into and manipulated by standard commercial software.”

In this example, bullet (a) has specifically mentioned the deadline as “24 hours after receipt of the request”.

“(3) The documents referred to in subsections (1) and (2) must be kept for two years after the day on which the food was provided to another person or sold at retail, and must be accessible in Canada.”

“138 (1) Within 24 hours before the slaughter of a food animal and in accordance with the document entitled Ante-mortem Examination and Presentation Procedures for Food Animals, prepared by the Agency and published on its website, as amended from time to time, a licence holder must conduct an ante-mortem examination of the food animal or of a sample from the shipment that the food animal is part of, which must include, in the case an equine or a bird other than one that is a game animal or an ostrich, a rhea or an emu, the examination of the documents referred to in subsection 165(1).”

“11(2) If a complaint is received, the operator must implement the procedure and prepare a document that sets out the details of the complaint, the results of the investigation and the actions taken based on those results and keep it for two years after the day on which the actions are completed.”

“The process must be monitored, by a designated plant employee, a minimum of every 2 hours for crust disappearance and internal product temperature.”

Periodicity:

“(2) The operator must, at least once every 12 months,

- (a) conduct a recall simulation, based on the recall procedure;
- (b) prepare a document that sets out the details of how the recall simulation was conducted and the results of the simulation, and keep that document for two years after the day on which the recall simulation is completed.”

In this example, the first sentence (preface in combination with bullet (a)) mention “every 12 months” which is consistent with periodicity time constraint. Keeping records for a period of time is required by bullet (b).

However, you can see samples that include time (even its metric unit) but does not imply time constraint, such as, “The delayed evisceration time must be incorporated into total chilling time for the affected product.”, “The license holder can validate alternate delayed evisceration time (other than 30 minutes).”, “When developing the monitoring procedure following must be followed: Select time for examination.” or “225 The label of a prepackaged food must be applied or attached in such a manner that the label is still applied or attached at the time it is sold.”

11. Convention: There are a number of requirements for the legibility/usability of labels. Most commonly, these include standards for the font (dimension measure) and format ( colors, metric units, languages, and order) of the



expressions/characters/marks/information. We don't have control over measuring or deciding them, for example:

“(2) That information must be shown in characters that are

- (a) in the case of a tray with an overwrap or an egg carton, on the top or side of the tray or egg carton, at least 1.5 mm in height; and
- (b) in the case of a container other than a tray with an overwrap or an egg carton, at least 6 mm in height.”

“(3) The geographic origin of a food must, subject to the requirements of any other federal or provincial law, be shown

- (a) in close proximity to the name and principal place of business of the person by or for whom the food was manufactured, processed or produced; and
- (b) in characters of at least the same height as those in which the information referred to in paragraph (a) is shown.”

“(2) That information must be shown in characters that are

- (a) in the case of a tray with an overwrap or an egg carton, on the top or side of the tray or egg carton, at least 1.5 mm in height; and
- (b) in the case of a container other than a tray with an overwrap or an egg carton, immediately below the common name, at least 13 mm in height.”

In this example height, (generally format) is a given and it is not something we have control over.

12. As we went through the provisions of the regulation we noticed some words that might imply vagueness such as hazard, contamination, necessary, close to, sufficient, appropriate, adequate, distinct, discernible, as much as, promptly, quickly, nearly, immediately, clearly, etc. The important point here is that you need to investigate whether the sentence has any content which implies a label in spite of these vague words. For instance, “(3) The geographic origin of a food must, subject to the requirements of any other federal or provincial law, be shown in close proximity to the name and principal place of business of the person by or for whom the food was manufactured, processed or produced.” In this example, although determining the proximity is vague, the “geographic origin must be shown on the food product” is a data definition requirement.

“230 The declaration of net quantity that is shown on the label of a consumer prepackaged food must

- (a) be in distinct contrast to any other information or pictorial representation on the label; and
- (b) show the numerical quantity in boldface type.”

In this example, “the declaration of net quantity” must be shown, is the data definition, and “distinct contrast” is vague.

“If such a word or expression is shown on the label, it must be shown in close proximity to the common name.”, “208 Any information that a label is required by these Regulations to bear must be clearly and prominently shown and readily discernible and legible to the purchaser under the customary conditions of purchase and use.” Leaving aside vague words like close proximity, clearly, prominently, and discernible, “any information” or “such word or expression” are too generic to be considered data definitions.

“11 (1) Any food that is imported must have been manufactured, prepared, stored, packaged and labelled in a manner and under conditions that provide at least the same level of protection as that provided by sections 47 to 81.” In this example, the same level of protection is not known unless you take a look at the cross-reference. For example, it might mean sanitizing hands or washing tankers. The statement does not have enough content to be considered as a measurement. As a result, it is impossible to determine the amount of “same level of protection” just by reading this statement.

”(d) be equipped with instruments to control, indicate and record any parameters that are necessary to prevent contamination of the food;” In this example, we don’t know all the necessary “parameters” to prevent contamination (it is too generic). There are a number of parameters that can be referred to, such as temperature, time monitoring, space requirements, and sanitation rules, but not all may require measurement.

“Where conventional tank chilling is used, care must be taken to ensure that: sufficient overflow of water is provided to ensure the removal of extraneous matter prior to final icing.” Although “sufficiency” is a vague term, it is clear that measuring the water level to reach a specific level is essential.

However, in these two examples: “133 A licence holder must provide a food animal with sufficient space to prevent the suffering of, injury to or death of the animal.” or “134 A licence holder must provide a food animal with sufficient ventilation to prevent the suffering of, injury to or death of the animal.” “sufficient space” and “ventilation” to prevent suffering or injury seem not to be measurable and actionable.

It appears that the three following requirements mention measuring temperature, humidity, PH, and pressure in order to reach a point according to the rules, ignoring the vagueness of the terms “appropriate level” and “adequate quantity”.

- (a) “65 (1) The temperature and humidity level in a facility or conveyance where a food is manufactured, prepared, stored, packaged or labelled or where a food animal is slaughtered must be maintained

at levels that are appropriate for the food or the food animal, as the case may be, and for the activity being conducted.”

- (b) “71 (1) An establishment must be supplied, as appropriate for the food or the food animal that is intended to be slaughtered, as the case may be, and for the activity being conducted, with water that is adequate in quantity, temperature, PH and pressure to meet the needs of the establishment;”
- (c) ”(c) must be capable of maintaining the temperature and humidity at levels that are appropriate for the food and, if necessary to prevent contamination of the food, be equipped with instruments that control, indicate and record those levels.”

“(4) The container of a consumer prepackaged food that is set out in column 1 of items 5 to 10 of Table 2 of Schedule 3 may have any volume capacity that is set out in Table 7 of that Schedule, in the case of metric containers, or Table 8 of that Schedule, in the case of imperial containers.” Although, you don’t know what are the values for “volume capacity” in the schedule, but the statement explicitly elaborates on the measurement of this dimension measure, where the container may have the same volume capacity of referred values.

“(2) The declaration of net quantity of a serving must be shown in accordance with the requirements of sections 231 and 233 to 237 respecting the declaration of net quantity of the food;” Data definition is present in this example, but the cited sections are not examined.

In some cases it cannot be determined what label will be assigned to the statement without reading the cited sections, e.g., “(3) Paragraphs (1)(b) and (c) do not apply in respect of any processed fruit or vegetable product that does not meet the requirements of the Regulations with respect to grade and that is sent or conveyed from one province to another, if it is labelled with the words “Substandard” or “sous-régulier”.”

“(2) Subsection (1) does not apply in respect of

- (a) a food additive;
- (b) a beverage that contains more than 0.5% absolute ethyl alcohol by volume; or
- (c) a food that is set out in Schedule 1 and that
  - (i) is unprocessed and is intended to be manufactured, processed or treated for use as a grain, oil, pulse, sugar or beverage,
  - (ii) has a label applied or attached to it, or accompanying it, that bears the expression “For Further Preparation Only” or “pour conditionnement ultérieur seulement”, and
  - (iii) is not a consumer prepackaged food.”

In this example, bullet (b) has measurement, but not other bullets.

13. Generally, we do not classify human actions, such as imports, exports, sales, conveys, deliveries, grading, removal, rejections, etc., because we are not sure if they are automated by software systems. For instance: “201 A food, whether prepackaged or not, that is sent or conveyed from one province to another or that is imported or exported, and whose label bears a common name printed in boldface type, but not in italics, in the Standards of Identity Document must meet any standard that applies in respect of that common name.” In this example, the label bears an expression, is not data definition and it is stated that some standards must be met.

Additionally, there are cases where a statement contains more than one decision. In these cases, if a decision asks for a label, the statement must have it. “333 (1) Ungraded eggs that are received at an establishment where eggs are graded by a licence holder must be graded and labelled with the applicable grade name that is set out in the Compendium or if they do not meet the requirements in respect of any grade that is set out in these Regulations, they must be rejected.” In this example, the food product must be graded and labelled with “the applicable grade name” or it must be rejected.