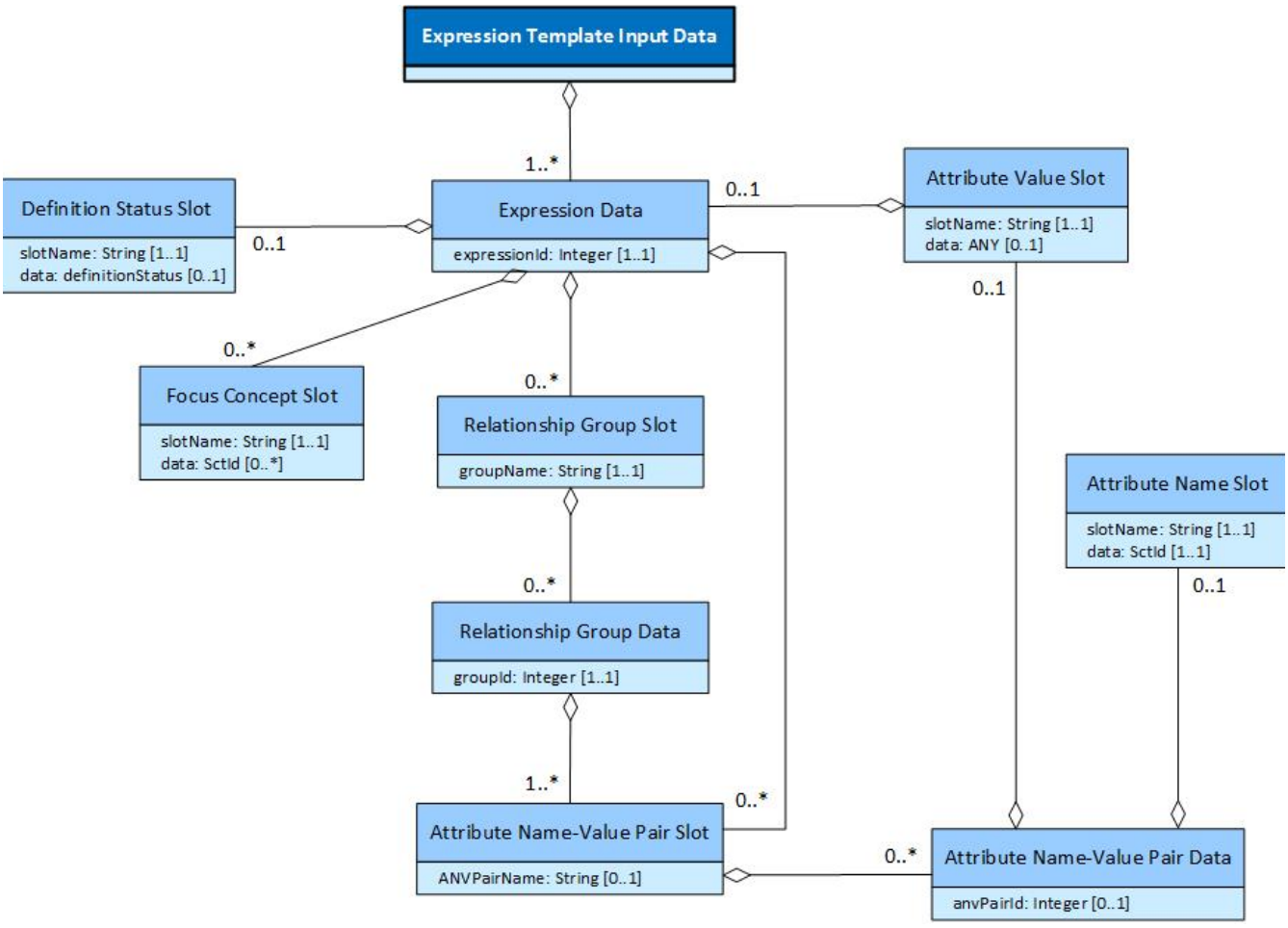
7.1. Preparing Input Data

Before a template can be processed, it is important that the input data is represented in a clear and unambiguous way. This is required to ensure that the template is processed in the expected manner, and the intended results are produced. In this section, we explain some of the considerations in representing and preparing the input data for processing.

Input Data Representation

Template input data may be represented in a variety of forms, ranging from flat tabular structures to nested serializations. Irrespective of the format, however, it is important that there is no ambiguity as to how each piece of input data should be used to create the resulting expressions. This can be particularly challenging where repetition of relationship groups or attribute name-value pairs is required.

The UML diagram below illustrates the logical structure of expression template input data. Each set of *Expression Template Input Data* includes the data intended to be used to create one or more expressions. The data used to populate a single expression is referred to in this model as *Expression Data*. Each *Expression Data* (identified by an expression id), may include at most one *Definition Status Slot* (with a slot name and a definitionStatus value), zero or more *Focus Concept Slots* (each with a slotName and zero or more values), zero or more *Relationship Group Slots* (each with a group name), and zero or more ungrouped *Attribute Name-Value Pair Slots* (each with a name). Each Relationship Group Slot has zero or more *Relationship Group Data* instances in the input data (each identified by a group id). Each of these *Relationship Group Data* instances has input data for one to many *Attribute Name-Value Pair Slots*. And for each *Attribute Name-Value Pair Slot* within a *Relationship Group Data* instance, there are zero to many *Attribute Name-Value Pair* Data instances (identified by an anvPair id), each with at most one *Attribute Name Slot* (with name and value), and at most one *Attribute Value Slot* (with name and either a simple data value, or an *Expression Data* instance of its own).



Unknown macro: 'caption-figure'

Input Data Examples

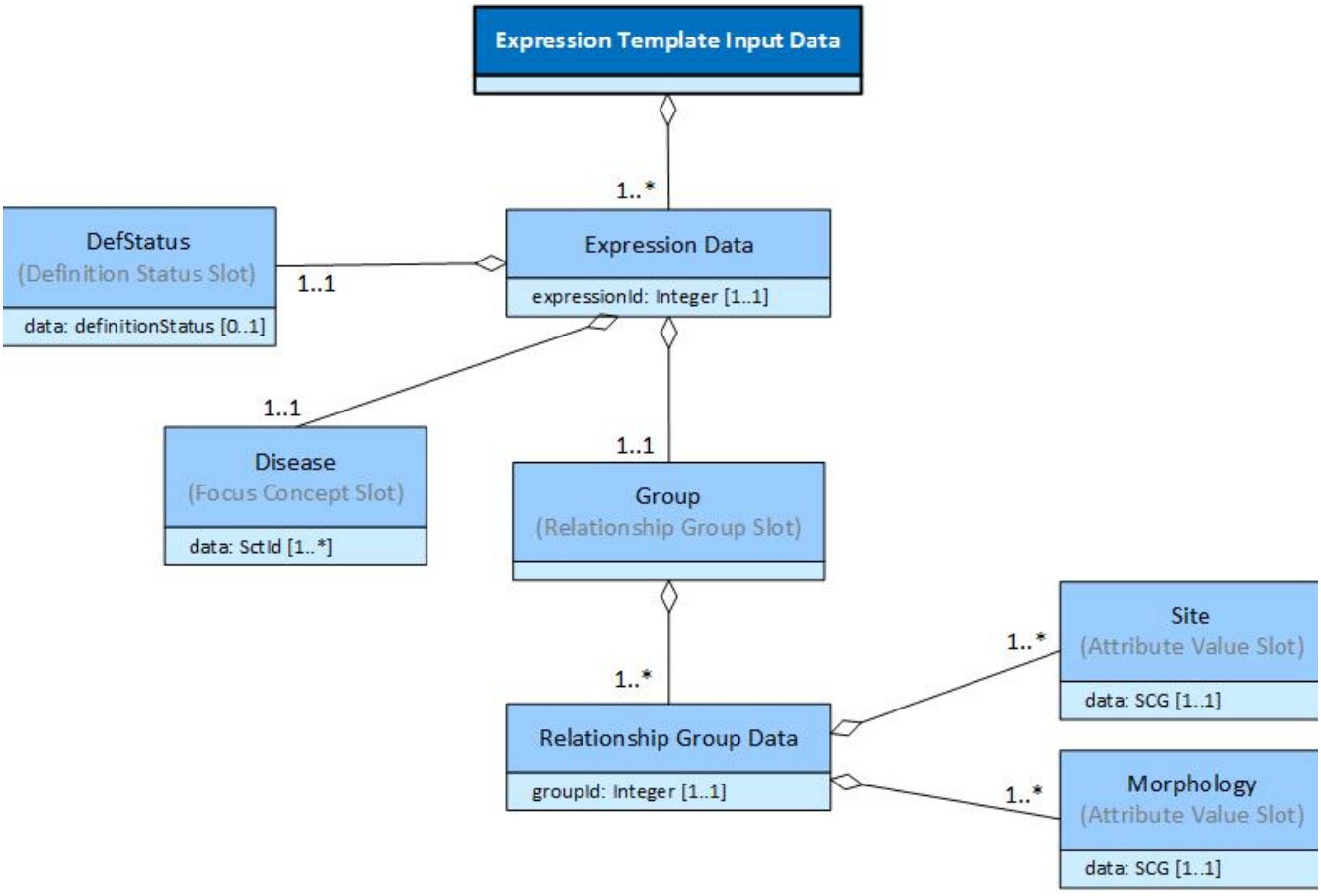
In this section, we provide some examples of unambiguous expression template input data, and discuss how this input data can be used to populate each expression.

Example 1

The expression template below is used to create expressions that represent a | Disease| with one or more | Finding site| and | Associated morphology| . When using expression templates, such as this one, in which attribute name-value pairs and relationship groups may be repeated, the input data should be explicit about which data values are used to populate each slot, and how these values are grouped into relationship groups.

```
[[+tok (=== <<<) @DefStatus]] [[+id (<< 64572001 |Disease| ) @Disease]] :
   [[@Group]] { 363698007 |Finding site| = [[+ (<< 272673000 |Bone structure| ) @Site]],
                116676008 |Associated morphology| = [[+ (<< 72704001 |Fracture| ) @Morphology]] }
```

To support the creation of input data for this expression template, the logical model in Unknown macro: 'caption-ref' above can be specialized by replacing the 'Slot' classes (e.g. 'Relationship Group Slot' and 'Attribute Value Slot') with the name of the respective slots in the template, and simplifying where possible. The resulting logical model of input data for the above expression template is shown below in Unknown macro: 'caption-ref'. Please note that this model has been simplified by removing unnamed logical classes, which have a cardinality of 1..1 and no data attribute. For example, attribute name-value slots are not required in this example. In general, attribute name-value slots are only required where both the attribute name and the attribute value are represented using a slot.



Unknown macro: 'caption-figure'

By populating this logical model with input data, as shown below in Unknown macro: 'caption-ref', the expression template can be processed to generate completed expressions. Please note that the first column in the table below is used to group together the input data intended to populate each expression. Subsequent columns are named according to the associated slot in the expression template. Relationship group slots are used to group the data that is intended to populate a single relationship group. Attribute name-value slots are not required in this example. They are are only required where both the attribute name and attribute value use a slot.

Unknown macro: 'caption-table'

Expression Data	DefStatus	Disease	Group	Site	Morphology
1	===	46866001 |Fracture of lower limb|	1	12611008 |Bone structure of tibia|	72704001 |Fracture|
2	<<<	92196005 |Benign neoplasm of lung|	1	39607008 |Lung structure|	3898006 |Neoplasm, benign|
		92038006 |Benign neoplasm of bronchus|	2	955009 |Bronchial structure|	3898006 |Neoplasm, benign|
3	<<<	60667009 |Closed fracture of rib|	1	113197003 |Bone structure of rib|	34305007 |Fracture, multiple, closed|
		36991002 |Closed fracture of upper limb|		371195002 |Bone structure of upper limb|	
4	===	16119006 |Abscess of jaw|	1	70925003 |Bone structure of maxilla|	44132006 |Abscess|
		109327001 |Abscess of facial bone|			
		128234004 |Disorder of maxilla|			

Using input data shown in Unknown macro: 'caption-ref' to populate the given expression template will result in following four expressions.

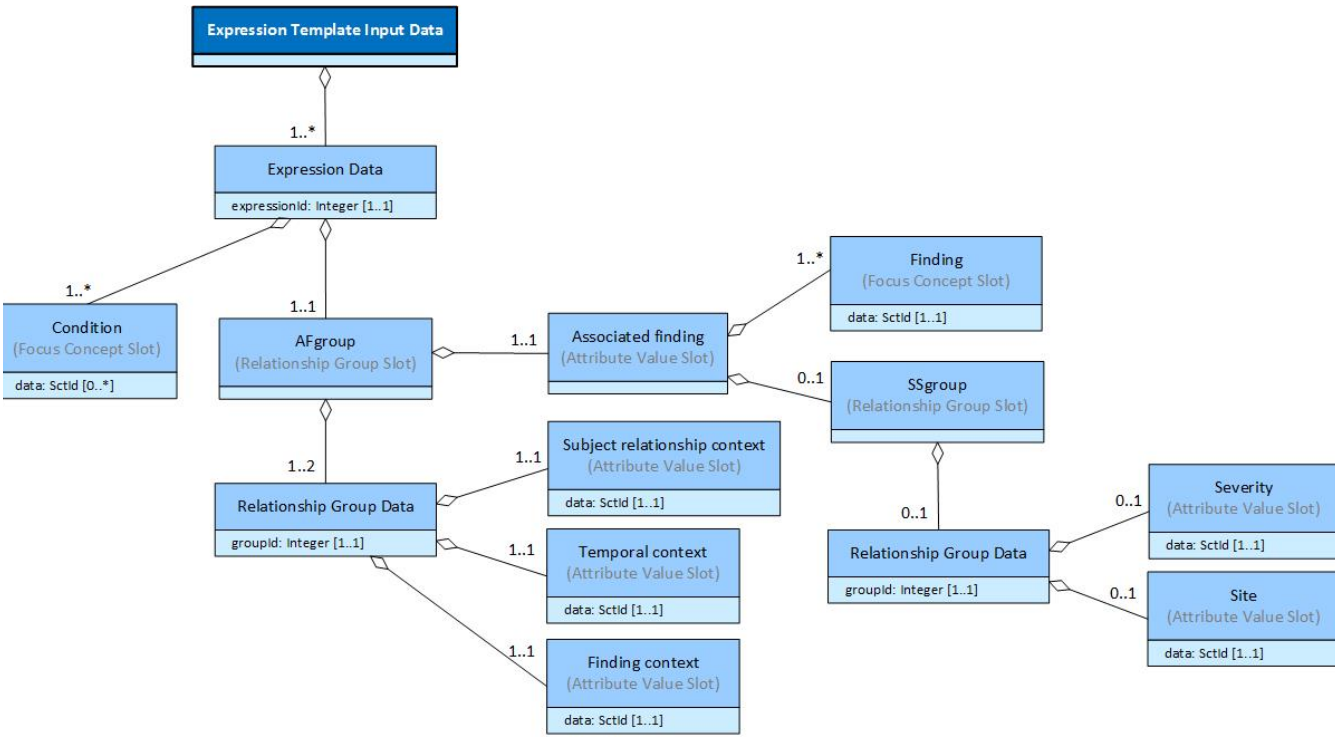
	Expression
1	=== 46866001 |Fracture of lower limb| : { 363698007 |Finding site| = 12611008 |Bone structure of tibia| , 116676008 |Associated morphology| = 72704001 |Fracture| }
2	<<< 92196005 |Benign neoplasm of lung| + 92038006 |Benign neoplasm of bronchus| : { 363698007 |Finding site| = 39607008 |Lung structure| , 116676008 |Associated morphology| = 3898006 |Neoplasm, benign| }, { 363698007 |Finding site| = 955009 |Bronchial structure| , 116676008 |Associated morphology| = 3898006 |Neoplasm, benign| }
3	<<< 60667009 |Closed fracture of rib| + 36991002 |Closed fracture of upper limb| : { 363698007 |Finding site| = 113197003 |Bone structure of rib| , 363698007 |Finding site| = 371195002 |Bone structure of upper limb| , 116676008 |Associated morphology| = 34305007 |Fracture, multiple, closed| }
4	=== 16119006 |Abscess of jaw| + 109327001 |Abscess of facial bone| + 128234004 |Disorder of maxilla| : { 363698007 |Finding site| = 70925003 |Bone structure of maxilla| , 116676008 |Associated morphology| = 44132006 |Abscess| }

Example 2

The expression template below is used as a pattern for family history expressions. It contains a nested relationship group (i.e. SSgroup) inside the outer relationship group (i.e. AFgroup). To populate this expression template, the input data must be clear as to where each value should be used, and how these values should be grouped into relationship groups and expressions.

```
[[+id (<<  413350009 |Finding with explicit context|  ) @Condition]]:
   [[ 1..2 @AFgroup ]] { [[1..1]]  246090004 |Associated finding|  = ([[+id (<<  404684003 |Clinical finding|  ) @Finding]]:
      [[0..1 @SSgroup]] { [[0..1]]  246112005 |Severity|  = [[+id (<  272141005 |Severities|  ) @Severity]],
                  [[0..1]]  363698007 |Finding site|  = [[+id (<  91723000 |Anatomical structure|  ) @Site]] }),
            [[1..1]]  408732007 |Subject relationship context|  = [[+id (<  444148008 |Person in family of subject|  ) @Relationship]],
            [[1..1]]  408731000 |Temporal context|  = [[+id (<  410510008 |Temporal context value|  ) @Time]],
            [[1..1]]  408729009 |Finding context|  = [[+id (<  410514004 |Finding context value|  ) @Context]] }
```

To support the creation of input data for this expression template, the logical model in Unknown macro: 'caption-ref' above can be specialized as shown below in Unknown macro: 'caption-ref'. Please note that this model has been simplified by removing unnamed logical classes, which have a cardinality of 1..1 and no data attribute.



Unknown macro: 'caption-figure'

This logical model can be populated with input data, as shown below in Unknown macro: 'caption-ref'.

Unknown macro: 'caption-table'

Expression Data	Condition	AFgroup	Finding	SSgroup	Severity	Site	Relationship	Time	Context
1	266898002 |Family history: Respiratory disease|	1	195967001 |Asthma|	1	24484000 | Severe|		444301002 |Mother of subject|	410589000 |All times past|	410515003 | Known present|
2	161077003 |Father smokes|	1	77176002 |Smoker|	1	255604002 | Mild|		444295003 |Father of subject|	15240007 |Current|	410515003 | Known present|

	161078008 |Mother smokes|	2	77176002 |Smoker|	1	24484000 |Severe|		444301002 |Mother of subject|	15240007 |Current|	410515003 |Known present|
3	160288009 |Family history: neoplasm of skin|	1	372130007 |Malignant neoplasm of skin|	1	6736007 |Moderate|	113179006 |Skin structure of nose|	444304005 |Sister of subject|	410511007 |Current or past (actual)|	410515003 |Known present|
	275937001 |Family history of cancer|			2	255604002 |Mild|	88089004 |Skin structure of lip|			

Using the input data shown above to process the given expression template will result in following three expressions.

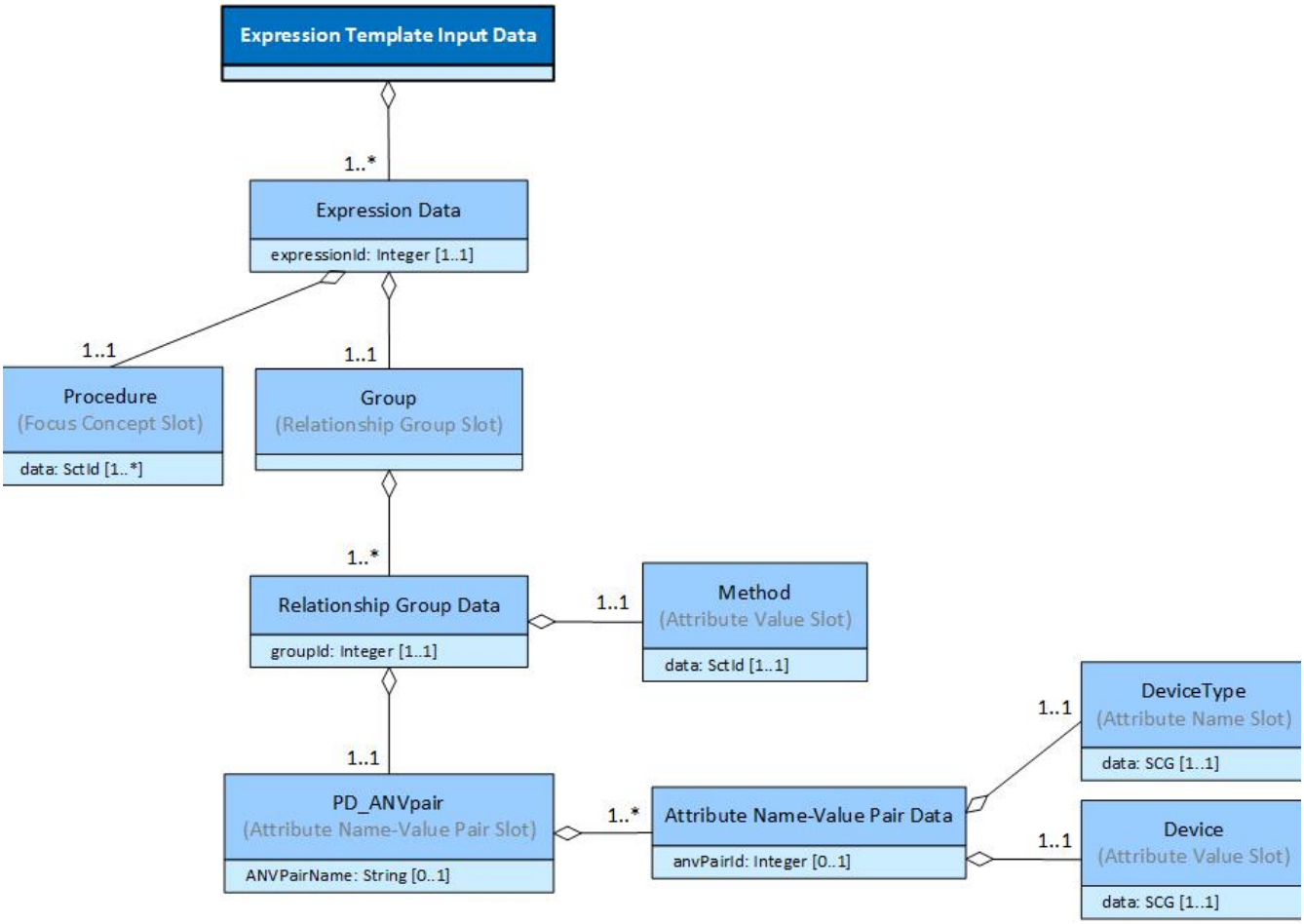
	Expression
1	266898002 |Family history: Respiratory disease| : { 246090004 |Associated finding| = (195967001 |Asthma| : { 246112005 |Severity| = 24484000 |Severe| }), 408732007 |Subject relationship context| = 444301002 |Mother of subject| , 408731000 |Temporal context| = 410511007 |Current or past (actual)| , 408729009 |Finding context| = 410515003 |Known present| }
2	161077003 |Father smokes| + 161078008 |Mother smokes| : { 246090004 |Associated finding| = (77176002 |Smoker| : { 246112005 |Severity| = 24484000 |Severe| }), 408732007 |Subject relationship context| = 444295003 |Father of subject| , 408731000 |Temporal context| = 15240007 |Current| , 408729009 |Finding context| = 410515003 |Known present| }, { 246090004 |Associated finding| = (77176002 |Smoker| : { 246112005 |Severity| = 255604002 |Mild| }), 408732007 |Subject relationship context| = 444301002 |Mother of subject| , 408731000 |Temporal context| = 15240007 |Current| , 408729009 |Finding context| = 410515003 |Known present| }
3	160288009 |Family history: neoplasm of skin| + 275937001 |Family history of cancer| : { 246090004 |Associated finding| = (372130007 |Malignant neoplasm of skin| : { 246112005 |Severity| = 6736007 |Moderate| , 363698007 |Finding site| = 113179006 |Skin structure of nose| }, { 246112005 |Severity| = 255604002 |Mild| , 363698007 |Finding site| = 88089004 |Skin structure of lip| }), 408732007 |Subject relationship context| = 444304005 |Sister of subject| , 408731000 |Temporal context| = 410511007 |Current or past (actual)| , 408729009 |Finding context| = 410515003 |Known present| }

Example 3

The expression template below represents a procedure with a single method and one or more procedure devices. Please note that in the first attribute name-value pair, both the attribute name and the attribute value use a slot. Because this name-value pair is repeatable, the input data needs to include an attribute name-value pair slot to ensure that the corresponding attribute name and attribute value stays connected.

```
[[+id (<< 71388002 |Procedure| ) @Procedure]]:
   [[1..1 @Group]]
      { [[1..* @PD_ANVpair]] [[+id (< 405815000 |Procedure device| ) @DeviceType]] = [[+ (< 260787004 |Physical object| ) @Device]],
        [[1..1]] 260686004 |Method| = [[+ (< 129264002 |Action (qualifier value)| ) @Method]] }
```

To support the creation of input data for this expression template, the logical model in Unknown macro: 'caption-ref' above can be specialized as shown below in Unknown macro: 'caption-ref'. Please note that this model has been simplified by removing unnamed logical classes, which have a cardinality of 1..1 and no data attribute.



Unknown macro: 'caption-figure'

This logical model can be populated with input data, as shown below in Unknown macro: 'caption-ref'. Please note that because the first attribute name-value pair is repeatable and uses a replacement slot for both the attribute name and attribute value, the input data needs to include the attribute name-value pair slot to ensure that the corresponding attribute name and attribute value stays connected.

Unknown macro: 'caption-table'

Expression Data	Procedure	Group	PD_ANVpair	DeviceType	Device	Method
1	387713003 |Surgical procedure|	1	1	363699004 |Direct device|	2282003 |Breast prosthesis, device|	257867005 |Insertion - action|
2	71388002 |Procedure|	1	1	363699004 |Direct device|	313025003 |Hearing aid battery|	282089006 |Replacement - action|
			2	363710007 |Indirect device|	6012004 |Hearing aid, device|	

Using the input data shown in Unknown macro: 'caption-ref' to process the given expression template will result in the following two expressions.

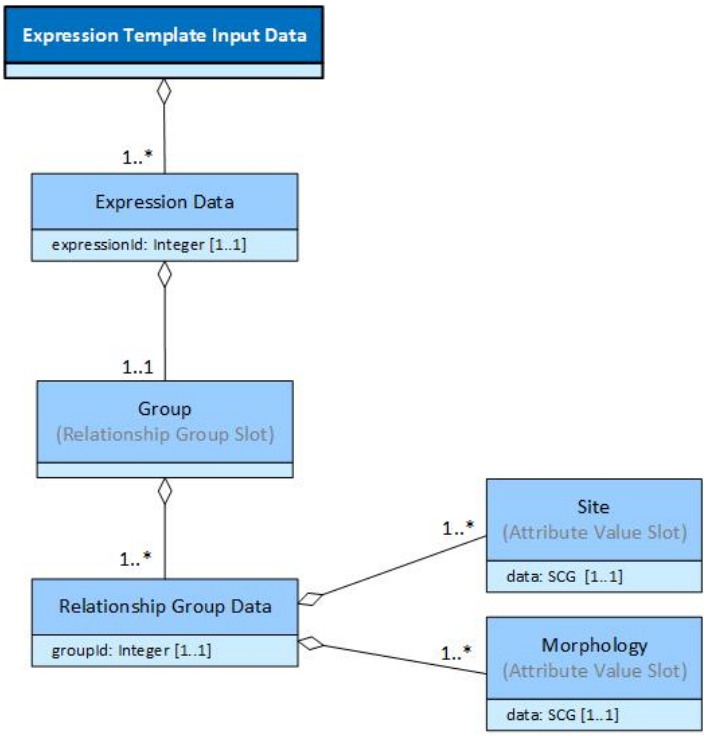
	Expression
1	387713003 |Surgical procedure| : { 363699004 |Direct device| = 2282003 |Breast prosthesis, device| , 260686004 |Method| = 257867005 |Insertion - action| }
2	384728007 |Replacement of device| : { 363699004 |Direct device| = 313025003 |Hearing aid battery| , 363710007 |Indirect device| = 6012004 |Hearing aid, device| , 260686004 |Method| = 282089006 |Replacement - action| }

Example 4

The expression template below represents a | Disease| with one or more values for | Finding site| and | Associated morphology| , grouped into one or more relationship groups.

```
64572001 |Disease| :
  [[@Group]] { 363698007 |Finding site| = [[+ (<< 272673000 |Bone structure| ) @Site]],
               116676008 |Associated morphology| = [[+ (<< 72704001 |Fracture| ) @Morphology]] }
```

To support the creation of input data for this expression template, the logical model in Unknown macro: 'caption-ref' above can be specialized as shown below in Unknown macro: 'caption-ref'. Please note that this model has been simplified by removing unnamed logical classes, which have a cardinality of 1..1 and no data attribute.



Unknown macro: 'caption-figure'

Unknown macro: 'caption-ref' below shows some example input data for the above template represented using the tabular format used in the previous examples. Unknown macro: 'caption-table'

Expression Data	Group	Site	Morphology
1	1	312763008 | Bone structure of trunk|	72704001 | Fracture|
	2	84667006 | Bone structure of cervical vertebra|	72704001 | Fracture|
2	1	71341001 | Bone structure of femur|	72704001 | Fracture|
3	1	12611008 | Bone structure of tibia|	72704001 | Fracture|

In addition to this tabular representation, there are a wide variety of other possible formats for representing template input data, including json, xml, tsv, csv etc. The exact format used will depend on the format required by the template processor. For example, the above input data can be represented in JSON as shown below.

Example JSON Representation of Input Data

```
{"Expression Data": [
        { "Group": [
                {  "Site":"312763008 |Bone structure of trunk|",
                    "Morphology":"72704001 |Fracture|" },
                {  "Site": "84667006 |Bone structure of cervical vertebra|",
                    "Morphology": "72704001 |Fracture|" } ] },
        { "Group": [
                {  "Site":"71341001 |Bone structure of femur|",
                    "Morphology": "72704001 |Fracture|" } ] },
        { "Group":[
                {  "Site":"12611008 | Bone structure of tibia|",
                    "Morphology": "72704001 |Fracture|" } ] } ] }
```

Simplification of Data Representation

While it is important that there is no ambiguity as to how each piece of input data should be used in processing the associated expression template, there are often opportunities to make the input data much simpler than is represented in the full logical model above. In particular:

- When the maximum cardinality of a relationship group is 1, there is no need to include the *relationship group* slot in the input data to group the attributes it contains;
- When the maximum cardinality of an attribute name-value pair is 1, there is no need to include the *attribute name-value pair slot* in the input data to group the corresponding name and value pairs;
- When there is either an attribute name slot with a fixed attribute value, or a fixed attribute name with an attribute value slot, there is no need to include the *attribute name-value pair slot* in the input data to group the corresponding name and value pairs.

With this in mind, the examples in 8. Expression Template Examples simplify the input data, where appropriate, using these assumptions and an implicit association with the logical model described above.