Base

GSIM Review

Issues to resolve related to Base group:

- What goes in the Base Group? All objects are identifiable and administered, so it makes sense to have Identifiable Artefact and Administrative Details in the Base group. **Should Agent, Individual and Organisation be moved to the Business group?** Some more examples of use and a better reason behind the grouping in the Base Group than that currently provided in GSIM Specification IV Technical Information paragraphs 109, 110, 121, 113 would be appreciated.

- Identifiable Artefact inconsistencies:
  - All identifiable objects inherit all attributes from Identifiable Artefact. **What constitutes an identifiable artefact?**
  - Some information objects in GSIM have an identifiable artefact tag. There is a note in the UML to say that ProcessInput is not an identifiable artefact.
  - Some information objects have a direct association with Identifiable Artefact. The items directly associated with Identifiable Artefact currently are not saying they are Identifiable Artefacts but that they can be associated with (any) Identifiable Artefact in a particular way.

  - In GSIM, the name and description attributes have a cardinality of 0..1 and a value type of Text. LIM proposes that these attributes should have cardinality of 1..1 and value type MultilingualText. **What cardinality should be used?**

- Definition of Agent is currently: "An actor that performs a role in relation to the statistical Business Process." Propose to change the definition to align with PROV-O: "An agent is something that bears some form of responsibility for an activity taking place, for the existence of an entity, or for another agent’s activity”

- Add in Change Event, Change Event Tuple, Role and Agent in Role (the last two replace Agent Role)
The Base Group in GSIM is described as:

109. These objects can be seen as the fundamental building blocks that support many of the other objects and relationships in the model. They form the nucleus for the application of GSIM objects. They provide features which are reusable by other objects to support functionality such as identity, versioning etc.

110. The GSIM Base Group consists of two sets of objects
1. Those that give identity and administrative details that are re-usable by other information objects.
2. Those that model the organizations and individual that may provide or consume data and referential metadata.

A: Identity and Administrative Details

111. The only base artefact in GSIM that gives underlying identity is the Identifiable Artefact. It can be inherited by any class in GSIM for which identity is required.
112. There is no attempt in GSIM to model the administration of items in repositories such as the maintenance agency, versioning, repository functions. However, the Identifiable Artefact does have a link to Administrative Details where such details can be added using the GSIM extension methodology.

B. Information Providers, Information Consumers, Organizations, and Individuals

113. Information Providers and Information Consumers are the respective sources for and the targets of data and referential metadata collection and dissemination. Each Agent can play the role of Information Provider or Information Consumer in a particular context of collection or dissemination. The same Agent may play the role of Information Provider in one context and the role of Information Consumer in another context. For any one Agent Role there must be a single Agent that plays the role: this is actual Organization or Individual that is the Information Provider or Information Consumer.

114. If the Agent is an Organization then it is possible to specify the structure of the Organization in terms of sub Organizations or Individuals.
A. Base Package in LIM

Figure 1: Base Group in LIM

Change Event

1. The Change Event was introduced to have a general way to manage changes in the states of information objects.

2. LIM Definition: A Change Event captures that a change has occurred. It identifies the objects that have been affected, and the new objects that have been created due to the change.

3. An association to Agent In Role was introduced in order to identify all Agents, with a specific Role, involved in the Change Event.

4. The Change Event has a changeDate attribute which is used to convey the event time and a changeType attribute that has a controlled vocabulary to accurately map to a recognized object lifecycle.

5. The Change Event identifier attribute indicates the change that is applied to Identifiable Artefacts.

Change Event Tuple

6. LIM Definition: A Change Event Tuple records which Identifiable Artefacts were changed by a Change Event. It keeps track of the source Identifiable Artefact(s) to which the Change Event was applied and the resulting target Identifiable Artefact(s).

Role and Agent In Role

7. Role and Agent In Role were introduced in order to have a more flexible and extensible way of adding Roles to Agents without having to change existing information objects. The introduction of these objects provides a way to relate the Agent, Role and Administrative Details objects and supersedes the GSIM information object AgentRole.

8. LIM Definition for Role: The responsible function involved in the statistical Business Process.
9. LIM Definition for Agent In Role: An Agent acting in a specific Role.

10. An Agent In Role may apply to either type of Agent - an Organization or Individual. The object is intended to reflect a single Agent acting in a single Role and as such is a very unambiguous representation. A common example would be to identify which Individuals or departments within an Organization provide administrative data.

Identifiable Artefact
11. An important decision was made that all attributes should be mandatory in Identifiable Artefact. The only exception to this rule is Local ID. This is an identifier that uniquely references an information object in a local context. The consequence of this decision was a change in the cardinality of some attributes and the moving of some attributes to other information objects.

12. After mapping to DDI, three new attributes were added (local ID, versionDate and versionRationale).

13. Three new associations were added i.e. one (applies to) from the new object Change Event and two (source, target) from the Change Event Tuple.

14. The cardinality of the relationship to Administrative Details changed from 0 ... * to 0 ... 1. This both simplifies the model and provides focus for all non-mandatory attributes to be carried.

Agent
15. There has been one new usage association relationship introduced, from Agent In Role to Agent. Cardinality 0 ... * at Agent In Role and 1 at Agent. This supersedes the relationship directly to Agent Role in GSIM v1.1

Administrative Details
16. A documentation attribute was added to Administrative Details to refer to an official document that has been published by the organization.
Proposals related to expanding Statistical Program and links to it:

- The Business part is abstract and can be more detailed for the objects Statistical Support Program, Business Process, Statistical Program, Statistical program cycle. These objects need to handle new ways of producing statistics, like one Statistical Program handling collection and microdata editing and several other Statistical programs using the result.

- GSIM needs to ensure there are adequate information objects that will enable flow of information through corporate and statistical activities. Note: The corporate information objects are likely to be available in other models – examples would include Customer Information, and do not necessarily need to be created in GSIM.

- A statistical program is in a context of a Subject Area/Domain. Is the Subject Field the same as a Subject Domain like labor market or prices? Add a relationship between Statistical Program and Subject Domain/Field.

- When designing a Statistical Program it can be made from a subject matter domain perspective. The total design needs to take into account several Statistical Program that are “linked” to each other through the information flow/dependencies. All Statistical Program doesn’t result in a Product. Propose to add a relationship between Statistical Program and Product. A Statistical Program can result in zero or many Products. A Product can be created from one and only one Statistical Program. The recursive relation to Statistical Program and the attribute "Subject Matter Domain" already handle the relationship between different Statistical Programs within the same subject matter domain.

Proposals related to removing/simplifying Statistical Program and links to it:

- Consider re-naming Statistical Program and related objects - This is because in many statistical offices, the term "statistical programme" is used to refer to an annual or multi-annual planning document setting out all of the statistical work the office will do i.e. a list of all surveys and/or products. As currently defined in GSIM, "Statistical Program" is not so much a programme ("program" is usually used in the context of a computer program), but a specific activity within a wider programme. Therefore the name of this object can cause confusion for users. Perhaps better alternatives could be "Statistical Activity" or "Statistical Workflow". A more radical step would be to simplify this part of GSIM, and merge "Statistical Program" with "Business Process", to give "Statistical Business Process" - which would demonstrate better alignment with GSBPM.