GSIM Sprint 2 - As it Happens

A daily blog by Thérèse Lalor on the progress of the Sprint 2 - 16 to 28 April 2012

Saturday 14 April - Sunday 15 April

The Sprint Organising Team spent Friday setting up the conference rooms at the Statistical Training Institute.

We have been enjoying the beautiful weather and lovely cherry blossoms in Daejeon. Much of the weekend was spent at a coffee shop which in front of the Airport bus stop waiting to greet participants as they arrived in Daejeon - the last pick up was 12.30AM Monday!

Monday 16 April

Today was the first day of Sprint 2. There was a mix of new and familiar faces in the room - with some Sprint 1 participants joining us again as well as new participants. Mr Woo, the Commissioner of KOSTAT (our hosts), joined us for some opening remarks and we were then given a presentation from our KOSTAT colleagues on their NARA system.

The focus for today was on context setting. We were joined by Steve Vale (via a pre-recorded video) for the HLG BAS context and had presentations on the feedback received for GSIM v0.3, how other industries present their information models, DDI and SDMX and the possible use cases to be looked at during Sprint 2. There were some interesting discussions as we started to touch on issues that will no doubt form part of the work plan in the coming days.

Tonight it is time for some Korean BBQ and birthday drink with one of our participants.

Tuesday 17 April

The focus of today was to review all the feedback received on GSIM v0.3. I think the participants found this interesting as it gave a good overview of the issues to be solved during Sprint 2. A document was created to note how each piece of feedback would be resolved and to provide basis to report back to the reviewers.

The feedback identified we need to make the scope and purpose clearer. Work started on reviewing what was agreed at Sprint 1 and considering how it could be better communicated. The organising and design principles were also identified as a key priority and a group has started work on enhancing those.

At the end of the day we had a meeting with the Practitioner group to update them on our progress. Although it is still early days, it was good to make the connection and we will talk to them again on Friday.

Before the Sprint each participant was asked to bring some sweets or biscuits from home to share. The idea was that each day we would be able to sample something from a different country during morning or afternoon tea. Now in the support room, we have a plethora of interesting snacks from across the world. Today we tried pumpkin sweets from Korea.

Wednesday 18 April

This morning saw the first "soap box" presentation of the Sprint. This gives participants a chance to briefly present an idea they want to share with the group. The topic this morning was a template for the defining characteristics of each information object.

Following this we did work plan for the rest of Sprint 2. We are hoping to have the first cut of the modelling done by Friday which will allow use case testing and defining relationships to other models.

Thursday 19 April

Today it was straight down to work!

We had a soap box to begin the day. This was about how GSIM plays a vital role in developing shared understanding of information management in official statistics. Currently there are lots of existing standards/models. Some have limited definitions and many use the same term to mean different things. With rich definitions, GSIM can fill the gap here.

A lot of modelling work was progressed throughout the day. As I went into the different discussion rooms to see the progress, it was clear that participants were looking at the existing standards and models to inform their work. Phrases such as "are you saying this is the same as in X?" or "I think we should look at how Y does this" were frequently heard.

Participants are also putting their knowledge of the statistical process to good use. In order to describe and critique, real world examples were used.

At the end of the day, the groups have done the first cut of modelling lower level objects such as variable, concept, classification and statistical activity. The groups also have started to document the relationship between GSIM and other standards and models.

There are lots of outputs to share and discuss during the meeting with other practitioners tomorrow.

Friday 20th April

A modelling day today. The participants split into 3 groups and everyone got involved in the modelling. It was interesting to compare the approach taken by the two modelling groups from yesterday. It is clear to the participants that there is a tension when doing this work. On the one hand, GSIM should be a communication tool but on the other it needs enough detail to map to other standards (e.g. DDI and SDMX). This was a discussion that came up with the practitioner group as well. I think the participants have hit a happy medium of producing diagrams that are abstracted up a little from the detail and going to lower detail only on some objects.

At the end of week 1, the participants are a bit tired but still full of enthusiasm - with many talking about continuing to model objects over the weekend. So far, they have reviewed concept, variable, classification, population, statistical unit, data source, statistical activity, dataset, production step and questionnaire

Tonight we had our group dinner at a lovely restaurant just outside Daejeon to celebrate a fruitful week.

Saturday 21 April - Sunday 22 April

The participants looked rather tired on Friday night after a long week. I think everyone was looking forward to a weekend break. Some went to Seoul on Saturday and KOSTAT kindly invited us to visit the Magoksa Temple on Sunday.

It was quickly clear that the participants are now so involved in thinking about GSIM that the discussions continued the whole weekend. Many a wish for a successful GSIM Sprint was heard at Magoksa and there was a competition running to find a slogan for GSIM. Here are a few of the suggestions:

Great Statistics in the Making
Industrial strength statistics
Information for innovation

Monday 23 April

Today we had a number of participants leave and arrive. We said farewell to two of the Standards experts and welcomed a representative from UNESCAP.

There is a focus this week on how to package the information up and communicate it. Some of the participants worked on integrating all the modelling work into a coherent whole and others worked on how to communicate the information to the different audiences.

There is a lot of work to be done this week. Early in the day it was suggested that today we should work late to make significant progress on the work. It is now almost 11pm and the groups are still working hard!

Tuesday 24 April

A few tired faces around the table today after working late last night.

The theme of today was how to communicate the model. After a lot of robust discussions in the morning, the group came together in the afternoon to sort out some of the terminology that needed to be clarified. Do we have levels or layers? What are they called? Do we need to show both statistical unit and population at the highest level or just one? What about variable and concept? How do we show that the model is aimed at both traditional and new statistical processes? A lot of decisions were taken.

Another issue tackled today was some objects that the feedback on v0.3 suggested were missing from the model. One of the groups looked into Quality, Reference Metadata and Methodology.

Wednesday 25 April

I am writing the blog to the tune of objects, relationships and their cardinality "from product service to dataset, contains, 1 to many" as one group works on the lowest level of the model. It is a surprisingly rhythmic sound to type to.

Today a lot of work was undertaken on the documentation. We need to make sure that the groupings decided at the highest level flow down through all the levels of the model, as well as the definitions and descriptions.

The work on the use cases progressed further. The group is looking at use cases on environmental investments from the Netherlands and Web Scraping from IMF.
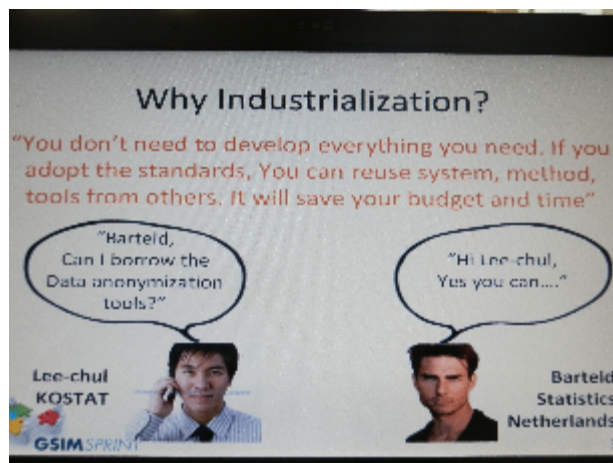
At the end of the day we had a web conference with the practitioner group and the HLG-BAS Secretariat. We gave them a dry run of the presentation to HLG BAS and gathered their comments on the work so far. I should give a special mention to our colleagues in the US and Canada who have participated in 3 webconferences at 5am over the last two week. Your comments/questions/insights have been invaluable to our process.

Thursday 26 April

After a morning spent finalising the model and use case testing, we visited the KOSTAT office to give a presentation. As there was a planned fire drill, we needed to get there quite early.

We were given a warm welcome by Mr Woo and the KOSTAT staff. It was good for the group to see the reactions to what we have been working on for the last two weeks. The presentation appeared to be well received. There were some interesting questions on modernisation in the Asia-Pacific region as well as a suggestion to ensure that people from developing countries are involved in the continutation of the work in some way.

Friday 27 April

Last day of Sprint 2.

This morning has been spent peer reviewing the documentation that has been created during the Sprint. This afternoon we reviewed the progress during Sprint 2 and reflected on Sprint process as a way to accelerate the development of GSIM. It was acknolwledged that the rapid development GSIM has had since the start of the year would not have been possible without the Sprints. In fact it would have take many more months (or years!) to achieve the same progress.

The final output of the Sprint was to present the outcomes of the Sprint to the HLG BAS group who were meeting in Geneva today. We received positive feedback from them on our progress and ended the 2 weeks on a high note.

We went out for dinner to celebrate the end of the Sprint and are now chatting and celebrating in the lounge of the 9th floor of the Statistical Training Institute (which has been our home for the last two weeks). Some of us are also being taught to play Go (a pleasure we have not time to take advantage of until now).

Thank you to all the participants of Sprint 2 - it was an intense, tiring but ultimately productive and enjoyable time that we spent together in Daejeon.