

Thema International Steering Committee Minutes

*Wednesday 15th March 2017, 10:30-12:00 GMT (UTC+0),
The Dark Room, Olympia, London*

Attendees:

Luc Audrain (Hachette Livre)	Martin Lüning (MVB)
Graham Bell (EDItEUR)	Renek Mendruń (BN Poland)
Marie Bilde (Pruneau)	Ivan O'Brien (Publishing Ireland)
Maria Börman (Bokinfo)	Brian O'Leary (BISG)
Pierre Boudigues (Electre)	Dominique Parisis (CLIL)
Francesca Cacciapaglia, (Informazioni Editoriali)	Jesús Peraita (FGEE / DILVE), Chair
Francis Cave (EDItEUR consultant)	Christer Perslöv (Bokinfo)
Michel Cerveillin (BTLF)	Simonetta Pillon (IE-Online)
Virginie Clayssen (Editis)	Vincent Poulvelarie (Dilicom)
Tim Devenport (EDItEUR & ISNI-A)	Suzanne Rozario (EDItEUR)
Fride Fosseng (Bokbasen)	Sofie Salmon (CLIL)
Alain Fournier (BTLF)	Chris Saynor (EDItEUR)
Bente Franck-Sætervoll (Bokbasen)	Henan Sun (Ingenta China)
Noah Genner (BNC)	Howard Willows (Nielsen)
Alex Haffner (MVB)	Jon Windus (Nielsen)
Janne Koskinen (Bonnier) (obs)	Johanne Worsaae Petersen (Gyldendal)
Karina Luke (BIC)	

Summary of decisions in this meeting:

- Minutes of previous meeting approved
- National qualifiers proposed for addition in April approved
- Thema 1.3 scheduled for Q1 of 2018 (unless there is specific demand for an earlier date)

1. Welcome and introductions

JP welcomed delegates to the 8th meeting of the *Thema* International Steering Committee (TISC), and asked all present to introduce themselves.

2. Minutes of the previous meeting

JP introduced the minutes of the meeting held in Frankfurt in October 2016, and asked for any comments. In the absence of any issues, the minutes were approved.

3. Report on current activity

JP asked CS to present his report on current activities.

CS welcomed JK from Bonnier Finland as an observer, in preparation for the creation of a new Finnish national group. The Australian Publishers Association has also set up a group to participate, and will be actively promoting use of *Thema* in Australia. There has been further interest in *Thema* from Czechia (Czech Republic).

EDItEUR has created a half-day 'Introduction to *Thema*' training course and it has been delivered in New York and Seattle as part of the ongoing BISG training programme. GB mentioned that the Seattle event was significant in that it was attended by staff from Amazon, and he hoped that Amazon Europe would

announce that *Thema* would be its preferred subjects scheme (in Europe) during the remaining part of the London Fair.

Post-meeting update: Isabel Gouveia of Amazon duly announced this at the BIC Seminar. Amazon's 'browse by subject' in its European stores is primarily driven by Thema classifications, though it also makes use of keywords where some Thema subject categories are not detailed enough (this is especially the case in genre fiction).

CS updated the group on various translation progress. The Spanish translation was recently improved and updated for version 1.3, and was for the first time now available for download as an Excel spreadsheet and HTML file, as well as via the interactive browser (<http://editeur.dyndns.org/thema>). This brings the total to 16 languages. In addition, Greek, Mandarin, Canadian French also exist but are not yet available, Czech and Slovak translations are progressing, and a Korean translation has been completed but is not yet available to EDItEUR).

With regard to translation, CS reported that there were certain subject headings that have more or less consistently raised issues in translation, where the heading is idiomatic or unfamiliar, or where a literal translation does not capture the concept or meaning of the original. Since it is fundamentally important that each code conveys the same meaning in each language, so English headings are being annotated to ease the work of translators. He illustrated the French translation process: the English headings and notes were translated professionally, but the French national group has also set up a cross-industry working group to examine the translation to ensure it captures the right meanings and that the terms used are appropriate to the book industry. CS noted that translations should also – wherever possible – include a translation of the *Thema* Basic Instructions document too. He also appealed for national groups to suggest examples of well-classified titles (both simple and hard-to-classify titles).

CS reported that new national extensions from the Spanish and UK groups were approved in January. The Spanish were primarily place qualifiers, though there were also the first national extensions to the language qualifiers. The UK added educational qualifiers for the Northern Irish education system.

The user interface for the online browser for *Thema* was updated slightly, and now includes links to the home pages of the organisations responsible for the translations. CS also noted that EDItEUR had begun to add synonyms to the online browser too. While the codes, headings and notes are all visible, synonyms or alternative labels are invisible but *are* included when searching. They can be used to ensure alternative spellings (even common misspellings) and alternative terms can be used in searches (*eg* a search for 'labor' should find a heading 'labour law', and a search for 'expressive therapy' should find a heading 'creative therapy'), or to improve search success in language variants (*eg* by including '*computadora*' in any (Spanish) category that includes '*ordinador*'). These alternatives can be used extensively for place names too, where sometimes both an Anglicised name and the local name of a place can be included. Originally, for Swiss and Belgian places, there were many place name variants listed. With the use of synonyms, only the usual Anglicised names are visible, while the others are hidden synonyms.

Synonyms can be added in any of the language translations to improve interactive search success. They should not replace the use of Class here and See also references.

CS reported on the progress with place qualifiers. There are large gaps in the geographical place qualifiers for countries where there is no national group – for example Greece has lots of books but no place qualifiers. The principles on which qualifiers can be added into the scheme *as if there were* a national group. CS proposed that a Working Group be formed to compile a list of places that are missing, but are needed to qualify large numbers of books published in countries which are using *Thema*. It is critical that this is not seen as an exhaustive gazeteer of place names, but it should include current cities and regions, and also cultural areas, particular points of interest, physiographic features and so on. Volunteers are sought for the WG, and suggestions are sought from *all* national groups (and other *Thema* users) with the

aim of having a list of proposals in time for Frankfurt. Various attendees indicated they would be prepared to work on the WG.

CS introduced the notes on cultural significance. He explained that such qualifiers are intended to have perennial significance – so a qualifier for a single anniversary would not be appropriate because bookshops may only dedicate a section of the store to the subject for a limited time. The proposal from the UK for a qualifier for ‘Shakespeare’ is good example of this. Bookshops will dedicate a permanent section of the store to all aspects of Shakespeare – plays, poems, criticism, biographies, adult or children’s, fiction or non-fiction, perhaps even tourist guides, textbooks and others. It is of extreme importance to the UK book trade. EDItEUR suggests that a ‘cultural significance’ qualifier could be incorporated by broadening code 5P ‘relating to specific groups and cultures’ to ‘relating to specific groups or social and cultural interests’, then adding a 5PK child ‘relating to specific and significant cultural interests’. This 5Px could then be extended with national extensions – eg 5PK-GB-S for ‘Shakespeare’. Such extensions would be for when there is a significant need in a national book supply chain that cannot be expressed using a core code and existing qualifiers.

He noted also that the introduction of 5Px would be a change introduced in version 1.3 at the earliest, but EDItEUR wanted to see how other national groups might use such an option – for exceptional needs, particularly cultural interests, within your supply chains that cannot be met by existing codes and qualifiers. He emphasised that the section would not be appropriate for temporary, faddish or fashionable interests (the UK group promises not to propose a Brexit code...), but for where there is a more-or-less permanent or perennial section reserved in many bookstores. BFS questioned the breadth of what could be added, and HW replied that potential extensions are not limited to ‘great authors’, but that the key would be to ensure that 5PK-xx- extensions are limited to *significant subjects of enduring interest which cannot otherwise be expressed in Thema or ONIX, and where bookshops often have dedicated sections or merchandising*. National groups should provide suggestions or examples of how such a section might be extended to EDItEUR over the next few months.

GB stressed that this 5PK proposal will most likely be introduced with version 1.3, and that because it opens the way for national extensions, it might be used not only for Shakespeare, but also for other cultural interests in other countries. At this stage, EDItEUR is simply exploring how it might be used.

Finally, CS introduced the question of interest age ranges. At present, *Thema* can express open ended interest age ranges such as ‘from about 5 years’. There are no upper bounds. Upper bounds could be introduced, either by having separate codes for ‘to about 7 years’, or by creating children of the ‘from 5’ code to mean ‘from 5 to 6’, from ‘5 to 7’, from ‘5 to 8’ and so on. The former would require two codes for each book to specify a range, the latter only one. Upper bounds would be optional, and if a single code ‘from 5 to 8’ were specified, a recipient could still if necessary truncate it to create an open-ended age range. GB noted that retailers prefer relatively narrow age ranges. BFS commented that publishers will not necessarily agree on particular age bands, which could prove problematic for booksellers. HW recounted the experience with BIC’s open ended ranges, which were originally proposed as non-overlapping bands, 1–3, 4–6, 7–10 and so on. IOB commented that some books are indeed suitable for a narrow range (eg 4–6) whereas others are suitable for a wider range (eg 4–10). He also noted that a reluctant reader audience may have a different (higher) range, but GB noted that this is *interest age* (which would not be higher for a reluctant reader). AF suggested it would be useful to align publishers and libraries on a common set of age ranges.

IOB and BFS suggested that a set of ‘upper bound’ codes would be inherently more flexible than a pre-defined set of bands. The disadvantage, as GB pointed out, is that publishers would tend to use inappropriately wide bands, which booksellers do not want.

LA questioned whether audience age ranging is in scope for *Thema*, when it can already be expressed in ONIX <AudienceRange>. GB noted this was a good question, as the group had previously taken the view

that *Thema* does not need to express things that can be carried in ONIX. But there may be a need to express audience ranges in *Thema* for where *Thema* is used separately from ONIX. CS stressed that this was a request for the national groups to provide comments.

4. Proposals for French, Irish, Canadian French and British national extensions

JP asked CS to cover the various national extensions proposed for introduction into *Thema* in April, which were separately documented.

CS highlighted the new French administrative regions, which have required changes to the normal parent-child relationships (like Tibet in a previous update). GB noted this has a technical effect for those who have implemented a *Thema* hierarchy in software – you cannot discover a code's parent by removing a letter. Systems will have to maintain a separate record of each code's parent (*eg* in an additional column in a database table). This 'parent' column is included in all available downloadable spreadsheets.

The new Irish national group proposed a range of place, education and interest group qualifiers.

Post-meeting update: County Longford was inadvertently omitted from the original Irish proposals, but was added subsequently.

The Canadian French proposals include new and updated (in French) place names, language national extensions, time periods and new codes for the Quebec educational system. The UK qualifiers include a small number of national extensions for places and a couple of new interest extensions. CS also highlighted two existing codes for Pennines and Cotswolds, which were originally added at 'European level'. Equivalent codes have been proposed within England (since that is where the Pennines and Cotswolds are), but the older codes are equivalent and not (at this stage) deprecated. The Tibet process is not a suitable fix for this.

The various proposals were approved without objection.

LA asked whether *Thema* should contain semantic information about whether places are cities, towns, regions, or are capitals. JP suggested this should be considered for 1.3. GB noted a limited amount of this information is present in some cases – *eg* 1DSE-ES-AA and 1DSE-ES-AAA are Almeria (Province) and Almeria (City) – but this has been added for disambiguation rather than to provide semantic value.

5. Timetable for development of version 1.3

JP moved the meeting on to the potential schedule for version 1.3. CS posed the question whether a version 1.3 should be prepared for autumn of 2017, or spring 2018? 1.2 was released May 2016, and *Thema* aims for an 18–24 month cadence (but always subject to demand).

FF asked that the range of suggestions should be sent to the Committee. BFS asked what volume requests had been received, and GB reported the list of suggestions was still relatively short – maybe 30 potential new categories – with nothing viewed as urgent. JW suggested that a certain level of demand is healthy, and CS agreed that current proposals would be made available via the Committee's Google Group.

There was general agreement that 1.3 would be schedule for early 2018, unless committee members indicated particular urgent updates had arisen.

GB reminded the Committee members that they should all be members of the Google Group, and if they are not, contact EDItEUR.

6. Updates from national groups

JP asked for significant news from any of the national groups.

GB noted the first *Thema* training sessions held in the US, and BOL mentioned that the BISAC to *Thema* mapping had been updated to align with BISAC 2016 and *Thema* 1.2. NH also confirmed that the

automated mapping tool provided by BookNet Canada has been updated to incorporate the new version of the mapping.

RM reported a new member of staff (at Biblioteka Narodowa?) was working on geographical extensions for Poland, and the group was waiting on the National Library integrating *Thema* into the library systems.

The Danish group (represented by MB) has completed an evaluation of the implementation of *Thema* across the Danish book trade, but the Board of the Danish PA has not yet approved the report.

BFS confirmed that *Thema* is fully implemented in Norway, with all new books being categorised.

JW and HW reported that in the UK, the transition is progressing. Nielsen has seen an uptick in supply of *Thema* from publishers, and will be able to use *Thema* as the ‘master scheme’ (potentially mapping to BIC and BISAC where required) later in 2017. This relies on the recently updated BIC to *Thema* mapping and the new ‘reverse mapping’ from *Thema* to BIC. BIC (the organisation) has strengthened its advice that publishers and retailer should transition from BIC (the scheme) to *Thema*, confirming that the BIC scheme will not be further developed. KL noted that the annual BIC seminar at the London Book Fair includes a session on *Thema* and there is training (from HW) planned for later in the year.

In Germany, ML reported that 1.7m records – over 90% of the total – now have good *Thema* codes. However, he reported that some publishers assigned categories that are overly broad, so there will need to be some quality process to ensure the categories match the contents, and possibly some education to ensure publishers use the scheme correctly.

AF reported that in French Canada, the trade wishes to adopt *Thema* as the primary scheme, rather than the current *Thèmes Electre*, and that the work on the Canadian French translation is ongoing. The possibility of a mapping to BISAC is appealing.

In Italy, while IE-Online has started implementing and promoting *Thema*, adoption by Italian publishers is limited so far.

Sweden has begun a series of training courses for publishers and booksellers.

JK reported that Bonnier Finland is aiming to create a national group for the Finnish book industry, and to participate in future development of *Thema*.

NG confirmed that most Canadian publishers were supplying *Thema* category information to BookNet Canada’s Catalist system, but there remained an issue with US publishers who supply data with only BISAC categories. The update of the BISAC to *Thema* mapping is helpful. No significant Canadian retailers were yet using *Thema*, but all had plans to do so in the near future.

In France, there are plans to update the mapping from CLIL to *Thema*, and the key Dilicom database now includes *Thema* subject categories (for both new and existing books).

Finally, JP reported that DILVE can now accept *Thema* codes uploaded by publishers, alongside the existing Spanish IBIC codes. Work is nearly complete on a tool to automate mapping from IBIC to *Thema* (which is different from the UK-derived BIC to *Thema* mapping). GB asked whether that mapping tool might also be valuable in the UK and Italy (where CCE is very similar to the BIC scheme). JP agreed it might, and that he would discuss with the UK and Italian stakeholders.

7. Any other business

There was no other business.

8. Date and location of next meeting

JP advised that the next *Thema* ISC face-to-face meeting will be during the Frankfurt Book Fair on 11th October 2017, 10am, room Brillanz to be confirmed. JP closed the meeting.

Graham Bell
EDItEUR
15th March 2017