

Thema International Steering Committee Minutes

*Wednesday 11th October 2017, 10:15-12:00 CEST (UTC+2),
Room Brillanz 4.2, Messe Frankfurt*

Attendees:

Luc Audrain (Hachette Livre)	Merve Okçuoğlu, (Turkish Publishers Association)
Detlef Bauer (Libri GmbH)	Brian O'Leary (BISG)
Graham Bell (EDItEUR)	Bruno Pacheco (APEL), guest
Pierre Boudigues (ELECTRE)	Dominique Parisi (EDITIS)
Francis Cave (EDItEUR consultant)	Tuula Pelkonen-Tiri (Kirjavalitys)
Francesca Cacciapaglia, (Informazioni Editoriali)	Lisbeth Hakansson Petré (Bokinfo)
Virginie Clayssen (EDITIS)	Jesús Peraita (FGEE / DILVE), chair
Ricardo Costa (Metabooks), guest	Christer Perslöv (Bokinfo)
Fride Fosseng (Bokbasen)	Simonetta Pillon (Informazioni Editoriali)
Alain Fournier (BTLF)	Vincent Poulvelarie (DILICOM)
Bente Franck-Sætervoll (Bokbasen)	Suzanne Rozario (EDItEUR)
Noah Genner (Booknet Canada)	Sophie Salmon (CLIL)
Simei Junior (Metabooks), guest	Chris Saynor (EDItEUR)
Janne Koskinen (Bonnier Finland)	Howard Willows (Nielsen)
Karina Luke (BIC)	Jon Windus (Nielsen)
Martin Lüning (MVB / VLB)	Johanne Worsaae Petersen (Gyldendal)

Summary of decisions in this meeting:

- Approval of new Finnish and CERLALC-affiliated ISBN Agencies groups
- Minutes of the previous meeting were approved
- Listings of *Thema* categories should incorporate sub-version numbers (eg v1.2.6 instead of just v1.2)
- Approval of the proposals for place qualifiers, including both national extensions to be added to v1.2 and (provisionally) the core qualifiers to be added to v1.3
- The *Thema* 1.3 working group should not consider the addition of detailed age range functionality that duplicates the feature of <AudienceRange> in ONIX

1. Welcome and introductions

JP welcomed delegates to the meeting of the *Thema* International Steering Committee (TISC), and noted the pleasing growth in the number of attendees. Delegates introduced themselves, and guests from Brazil gave a brief overview of their interest in ONIX.

2. Minutes of ISC Meeting held 15th April 2017, London Book Fair, and matters arising

JP introduced the minutes of the last meeting.

JP noted the two new groups for approval, the Finnish national group and the CERLALC-affiliated Latin-American ISBN Agencies group (a supra-national group). CS said that that representatives of the Finnish group were present, but that the CERLALC group is not present in person – he described the role of CERLALC within the book industry and its work to promote *Thema* (and ONIX). **The new Finnish and CERLALC-affiliated Latin-American/Caribbean ISBN agencies *Thema* groups were formally approved.**

PW said other matters arising from the minutes would be covered in the main agenda, and **the minutes were approved.**

3. Report on current *Thema* development work and a summary of activities since last meeting

CS introduced his activity report on EDItEUR's work since the last meeting. The full report is available separately, but CS mentioned a few of the key points.

Training: CS noted that he is working on updating the various documentation available for *Thema* via the EDItEUR website, and in particular beginning to build a couple of application notes (similar to those available for ONIX) – one to explain how *Thema* can work within the context of ONIX and the other ONIX metadata elements and the other to highlight various worked examples of real books and real subject classifications. Delegates were asked to let CS know of any interesting real-life examples they had come across that could be included.

Translation: *Thema* is currently available in 16 languages. No new translations have been added since the London Book Fair, but EDItEUR is awaiting confirmation that the Chinese, Korean and Green translations can be published. Work is also ongoing with a Canadian French localisation. CS emphasised that any lack of clarity encountered during translation should be reported to EDItEUR, as the English heading can then be modified during the update for *Thema* v1.3. CS also noted that translations can include explanatory notes that are not necessarily part of the English text. He asked that if national groups had completed translations of the Basic User Instructions, EDItEUR would like to link to them from its website (though it would also be happy to host the documents too).

National Extensions: following the London Book Fair, a number of national extension qualifiers were added to *Thema* v1.2. In April, we added national extensions for France, Francophone Canada, Ireland and the UK, mostly place qualifiers, with a few education and time period qualifiers for Quebec and Ireland. In August, there were new national extensions for Germany and Spain – educational qualifiers for German, and place qualifiers for Spain. DB explained that there are a large number of German education qualifiers because the *lande* administer schools, exams and curricula differently. The additions have been agreed with the major German education publishers. JP said that *Thema* 1.2 has added extensions in several sub-versions, and suggested that the filenaming of the documents such as the spreadsheets should indicate the sub-version. CS agreed EDItEUR would begin to do this.

Online browser: GB reported on minor updates to the *Thema* online browser. In particular there are more tooltips, and there is a level of SKOS-based RDF support for potential use of *Thema* in Linked Data. (As an aside, GB noted that with the online browser, if a term is not translated, the fallback to English is automatic. As a result, on the translation template, it is best NOT to copy the English term into the blue area for any terms which you have not translated. Simply leave the blue cell blank, and the browser will use the English anyway. The English should only be copied into the blue area for genuine 'loanwords'. JP raised the issue that this might be problematic because the translated list of categories would have blanks. GB said we'll perhaps need an extra column to identify the difference between a translation or loanword, and a fallback.) Finally, the browser has also changed address, to <http://ns.editeur.org/thema>, and it also now supports https (secure HTTP) for privacy and security reasons.

Synonyms: CS has added a range of synonyms to the browser, often gathered from anonymous logs of user searches. These synonyms are not visible via the browser itself. But if someone does a search for a word that is not part of the visible *Thema* text, it may still find the correct category. National groups are welcome to add their own synonyms – which might be translations of the English synonyms or entirely independent. CS gave the example that the category heading "Food & drink: spirits, liqueurs & cocktails" was accompanied by the synonyms 'Gin, whisky, vodka, rum'. AF noted that these were not really synonyms – CS agreed, saying they were more properly just extra indexing terms, and potential candidates for sub-categories in the future. But there are also genuine synonyms, and common misspellings that use the same mechanism in the browser. CS also mentioned that searches for 'Champagne' link to the French

region, but also – because of an added synonym – to the category for Wines. ML warned that the synonyms could make the spreadsheets very complex, so should not be included in the public spreadsheets, and CS agreed.

Place qualifiers: CS introduced the list of more than 600 proposed place qualifiers compiled by the working groups (most are national extensions, some are core qualifiers). The working group sought to create a list of places required by current *Thema* users in those countries that did not have their own *Thema* representatives. There were well over 1200 original suggestions, which the working group winnowed down by checking that there were sufficient books featuring each place. Reaching out to friends and colleagues in other countries also informed the list (*eg* for the Netherlands, Hungary, Greece, Brazil), and CS expressed his appreciation for this feedback. A few countries were avoided because of ongoing local work (*eg* Japan, Finland, Turkey, Portugal, Czechia, New Zealand). CERLALC also delivered suggestions for most Latin-American countries. These countries may make further place proposals in the future.

CS explained that the proposals will be added to *Thema* in two stages – if approved, the national extensions can be added to *Thema* v1.2 immediately, but there are some proposals for core qualifiers that will be delayed until v1.3 is ready.

CS thanked the members of the working group, Rikke Skov, Lada Kris, Adam Cholewka, for their hard work on this proposal, and asked for comments on the proposal. GB noted that while the proposals were extensive, they were not necessarily comprehensive, and there will likely be further place qualifiers added in future.

JP asked for approval of the additions and modifications, and there were no objections.

4. Update and discussion of development of *Thema* 1.3

CS noted that the working group on *Thema* 1.3 had also been meeting and had been working through the very large number of proposals for a significant revision of the core subject codes and qualifiers. He expected a full list of recommendations from the working group should be available around the end of the year (December or January). The candidate list would then need to be carefully considered by the full ISC before ratification and publication of *Thema* v1.3 in early 2018.

CS noted that the ISC had been presented with a more strategic question for *Thema* 1.3, regarding the addition of ‘audience’ detail – in particular age ranging (From... To...). Should *Thema* incorporate a stronger specification of the intended audience or readership for a book (using one of two possible methods), or should suggestions about audience be ‘parked’ and not considered for 1.3?

JP asked whether the results sent to EDItEUR so far were clear? CS noted that only a handful of replies had been received from people not present at the meeting, and opinions were widely spread.

FC asked for clarification of the age ranging, whether both lower and upper bounds would be mandatory. CS confirmed that it would not be mandatory – both methods allowed for simple From... ages as well as From... To... ranges. GB summarised that there were three options:

1. Do nothing (the level of readership age detail in *Thema* is about right for the moment)
2. Or, add more detail
 - a. allowing free specification of age ranges
 - b. allowing guided specification of age ranges

If choosing 1., the ISC should also consider whether *Thema* should include age ranging at all (*ie* should we plan for reducing or removing audience age codes altogether in some future update?) GB also noted retailers who use age bands strongly prefer guided specification of narrow bands, as wide bands are not useful in retail.

DB expressed the view that it should not be in both ONIX and *Thema* – we should not create more redundancy. LA and BFS agreed with this concern, both warning that having it in both could introduce internal contradiction. BO'L noted that the BISG subject committee's view was that age ranging should not be added to *Thema*. HW had surveyed the UK group, but the result was not clear cut – some thought age ranges should not be in *Thema* at all, but a slight majority supported option b). He also noted that there were other elements of audience or 'treatment' that were an important part of the *Thema* approach. CS noted the Irish and Japanese group supported dealing with age ranging only in ONIX (and thus did not want more audience detail in *Thema*). HW noted that not everyone used ONIX, and JWP stated this was the case in Denmark, for example. JP pointed out there were already overlaps in school grades too, and while this may be undesirable, there are arguments in favour of having age ranges in both ONIX and *Thema*. The Spanish group favoured option b) if additions were made.

Other speakers (FC, AF, NG) supported the 'do nothing' option. RC strongly supported a clear distinction, rather than having age ranging in both ONIX and *Thema*.

GB noted that the consensus of the ISC so far supported option 1 – that the current level of detail was sufficient, and potentially even the possibility of a reduction in Audience information at the next non-compatible update of *Thema* (ie in some future *Thema* version 2.0). But he noted this implied a problem for *Thema* users not using ONIX, eg in Denmark and proposed another option – that enhanced *Thema* age ranges could be used within the ONIX <Audience> composite. BFS and others asked what benefit would this have? Denmark could introduce ONIX!

JW asked whether there were any issues with the way ONIX supported age ranging? GB noted that ONIX was like option 2.a, allowing freeform From... To... ranges (for interest age, school grade, calendar age) with the attendant possibility of contradictions like "From 9 to 6" (though an advanced schema may police that possibility). BFS asked whether there were any reported problems with this? GB replied no.

SP supported the option of having some more detailed audience age ranging in *Thema*, in part simply because it is important for a 'classification scheme' to be able to deal with audience age (particularly for children) at least at a simple level, and in part because not everyone uses ONIX. The problem of overlap is simply about being careful. HW agreed, suggesting that it might be important for mapping, and JP also agreed – while we might not want to add more age detail to *Thema*, the existing functionality should not be removed. AF made the point that *Thema* was not a 'classification scheme' but a 'subject category scheme', and that audience was separate. DB supported the position that *Thema* should retain at least some audience age functionality.

MGS felt unable to take a strong position, but was conscious that *Thema* was a pragmatic scheme where some level of age detail would be useful (even if not academically perfect). MO noted the low use of ONIX within Turkey, which would again suggest at least some age ranging be retained in *Thema*. JWP noted that the Danish preference would be for option 2.a.

LA – not entirely seriously – suggested that age ranging could be developed further in *Thema* and removed from ONIX, but as BFS pointed out, there are more users of ONIX than *Thema*, and this would simply cause more problems in ONIX.

There was general consensus – voiced by BFS – that option 1. is preferred. So for CS, **the *Thema* 1.3 working group should not consider age ranging** (while there are other age-related issues that will be considered). It was noted that this does not meet the needs of the Danish group.

JP drew the discussion to a close, noting that the arguments and opinions had been thoroughly aired, and while 1.3 would not tackle age ranging, it was not ruled out for later updates.

5. Updates from National Groups

JP asked if any of the national groups had any updates, as time was limited.

BP – the Portuguese group was revising the translation of *Thema* 1.2, and working on proposals for some national qualifiers

KL – BIC is revising its product data accreditation scheme, and *Thema* will be recommended for all levels of accreditation and compulsory at the highest levels. It has also announced publicly that the old BIC subject scheme is frozen and users are recommended to transition to *Thema* (the scheme has been frozen for some years, but this public announcement is new). HW noted the new Nielsen system has gone live which enables Nielsen subject classifiers to use *Thema*. An upcoming BIC training on *Thema* is sold out. HW also repeated the news that Amazon EU has declared *Thema* to be its preferred subject scheme.

JWP – Denmark is trying to make *Thema* mandatory for its books in print database.

DB – all the big publishing groups now using *Thema* in Germany. There is now a special interest group with around 40 active members from across the supply chain, which is planning to produce best practice advice on metadata.

LHP – in Sweden, *Thema* has been in use for a couple of years and is mandatory (just one retailer is still using BIC and it will switch at the end of 2017). Bokinfo runs *Thema* education once or twice a year.

FF – in Norway, classification is carried out by Bokbasen and uses *Thema*. The first retailer will be implementing *Thema* on its website very soon. Bokbasen also focuses on training, with combined ONIX / *Thema* courses twice a year.

MO – in Turkey, resources are limited to a single translator, who is working on improving the 1.2 translation. We hope in the future to build a strong metadata study group.

TPT – in Finland in November we are going to have a publisher's seminar on *Thema* and some ONIX training. JK – it is very important to share knowledge across the Finnish book sector.

AF – building on official endorsement of *Thema* by the publishers, distributors, booksellers in Quebec, we are ready to go ahead with a formal national group: the first meeting is in November.

NG – the Canadian group is focused on proposals for national extensions (eg for indigenous cultures). The automated BISAC to *Thema* translator will be updated with the forthcoming BISAC 2017 to *Thema* 1.2 mapping when it is released (by BISG).

BOL – the BISAC 2017 are about to be approved, and the updated mapping will follow.

SS – we intend to continue to work on both CLIL domestically, but with *Thema* for international trade, and there will be a mapping between them. We have had discussions with French Canada about translation

FC – in Italy, we are revising the *Thema* v1.2 translation, and there will likely be some updates to national extensions for place qualifiers. We are not yet using *Thema* widely, but hope to progress as soon as possible: we are currently working on a trial with an Italian publisher.

MGS – in Australia, we have in principle agreement from the publishers and booksellers [to adopt *Thema*], but the major activity has been the rebuilding of the main metadata aggregation system TitlePage, which has been brought back in house. MGS expressed his surprise at the lack of awareness of the importance of subject classification among retailers in Australia, but he is intent on progressing *Thema*.

RC – at Metabooks, BISAC is a mandatory field, but the system will show a mapped *Thema* code as a first step towards adoption. The biggest book chain in Brazil, Saraiva, has already agreed to use *Thema*, and it will implement *Thema* in its bookstore before the end of 2017.

JP – DILVE now accepts *Thema*, and the need is to have automatic mapping between the [transitional] IBIC and *Thema*. There will be a little loss of precision, but Spanish publishers and retailers prefer this to changing all the systems quickly.

6. Any other business

JP called for any other important business, but there was none.

CS reminded delegates that they should be members of the *Thema* ISC Google Group.

7. Date and location of next meeting

JP advised that the next *Thema* ISC face-to-face meeting will be during the London Book Fair, venue and time to be confirmed, and closed the meeting.

Chris Saynor / Graham Bell

EDItEUR

11th October 2017