1. Hi, my name’s Tanja Vienonen, and I’m from the National Library of Finland. Today, I’m going to talk about our efforts to unify the metadata used in Finnish repositories.

2. This harmonising metadata work was part of TAJUA, the Open Scientific Publishing project. It’s a national project in Finland to promote Open Access principles, and improve the availability of Finnish research.

3. Over 60 organisations in Finland have a repository, there are 13 instances, some of them shared. Seven are hosted in the National Library. DSpace is most commonly used platform, together with qualified Dublin Core fields. All of this amounts to having over 300,000 OA full-text publications saved in the repositories.

4. Initially, the most challenging aspects were the kaleidoscopic metadata practices and non-standard DC produced from rushed, ad hoc solutions, which have also made aggregation problematic.

5. Here’s some examples of problem metadata, which, if you can see them, are mostly in Finnish.

6. Needless to say, partially outdated metadata guidelines only for graduate theses hasn’t been enough to unify the overall metadata. The muddled practices have made maintenance, development, migrations, and complying with aggregators’ requirements difficult.

7. Galloping to the rescue were the 80 some experts of the metadata working group. Of these, a handful of most interested people formed a smaller group to bear the brunt of the work. Experts from the National Library were involved through the whole process.

8. We approached the work with the end result in mind, as well as acknowledged that a single format isn’t enough to comply with all the requirements. We also concluded that a rich and accurate enough internal format would have to do. Therefore, semantic uniformity was prioritised over exactly “right” DC fields.

9. The work started with me exploring the most common formats, standards, and aggregators. And then compiling the most used metadata fields and qualifiers from the repositories, and based on these, making suggestions of more accurate metadata fields to be used when the ones in use shouldn’t really be used.

10. Still, the first draft of the guidelines was a monster of a spreadsheet.

11. To make access and editing easy, the spreadsheets were saved into Google Drive, all the meetings had remote access, and developments were communicated through email and the National Library’s online forum.

12. Then, all of a sudden, it was done.
13. The final version of the guidelines is available at one of the National Library’s public wiki pages. It consists of 62 DC based fields, and 11 extra ones were put in a separate namespace for use in the annual data collection of the Ministry of Education and Culture, if needed.

14. Here’s one result of the makeover. The dc.date qualifiers present here were cut to a sixth from before.

15. Six of the 62 metadata fields were labelled important, because we argue that these at least should be present in every metadata record.

16. In the future, there are yet more presentations being given, in order to receive a repository audience as wide as possible in Finland. The implementation of the metadata guidelines should start in the autumn.

17. This experience has taught me that uniform metadata and getting the repository community involved are essential for this kind of work to succeed. And the easier participation is, the better.

18. All I can say is, we did it! And got the Saimaa-Ringed-Seal-of-Approval to prove it.

19. Thanks for listening! Here’s a link to the guidelines and how to contact me.