# Traceability Modeling for the Engineering of Heterogeneous Systems

Report on Nasser Mustafa’s thesis, by Daniel Amyot

## Methodology

Literature review and survey with professionals to collect requirements, and “demonstration by construction” (against requirements) and case study for validation. Wasn’t there a chance to either have other people use the taxonomy of traceability model and get their feedback?

## Threats

Chapter 3 on threats (**located in an unorthodox place in the thesis**) seems to confuse threat mitigation with remaining threats. Very few remaining threats are actually discussed, which suggest that most threats were eliminated (instead of mitigated). This is actually not the case. Mitigations are good, but this does not prevent the remaining threats from being identified**. Overall, what are the remaining threats?**

3.3: No researcher bias because of the survey? You still ask the questions, select the participants, and collect and analyze the results… just like you were the only one to select and analyze the papers in the literature survey. There is plenty of bias threats remaining here.

**4.1: How was Google Scholar used? It does not support the complexity of this query nor the ability to search keywords or abstracts…**

4: **Where is the list of the 328/330 selected papers**, and the analysis of the criteria? How can we link your synthesis results to the evidence?

P38: **only 4 papers (which ones?) that discussed modeling heterogeneous artefacts… Sounds awfully small. What is the boundary between heterogenous and homogenous exactly?** Are SysML and i\* heterogeneous? Are UML and Java? You have more than this in Table 11. Sounds like a threat to the validity of this work.

5: **I am not sure how to read Table 7**… What are the columns (and why are contributions spread, and evolutionary relations missing)? Is it entirely aligned with the text (no!)? Not sure how to read Table 8 either!

6: Link metamodels from commercial tools such a s DOORS?

7: If you had to redo the survey again, what would be different? **Many questions such as Q8-Q9-Q11 in the survey should have allowed more than one answer to be selected.** Questions on phases/levels are unclear. In terms of results, some frequencies (e.g., of tools used) would have been appreciated.

7: I disagree with some claimed limitations of DOORS. But more importantly, are you proposing a better tool? You should have asked explicitly about tool shortcomings!

7: p.74 many of the concerns identified from the survey are actually **hard to relate to questions and answers…** For example, how do you conclude “2- The need for recording historical information about artifacts and trace links.”?

8: p. 76: Req1 and Req2 are the same! Also: how are the requirements traced to the previous chapters (and not just textual interpretations of part of the traceability metamodel)?

8: p. 79, Fig. 13: **Why have traces** (is **ordered** necessary?), and why is each TraceElement part of exactly one trace? **What is the URI of a keyword in a line of code**? Not tied to Ecore, great; how will it be implemented?

**8: Fig. 13: How can I constraint the use of a taxonomy or a set of predefined trace types for a traceability model? How can I ensure all tests are traced to a requirement? In other words, how can I implement a TIM? Constraints are on individual elements are insufficient, need them on TraceabilityRoot too.**

**8**. Fig.13: Homomorphic to common meta-metamodels! Are you rediscovering MOF/Ecore? Model, Element, Class, Association, Attribute, Constraint, Package,…!!! Predefined data types include taxonomies!

8: The more flexible the metamodel (Fig. 13), the more difficult it becomes to analyze and exploit! Is it too flexible?

9: **Taxonomies are great, but what does this mean in practice? If you are renaming one of the 4 SysML types of links, will you change their standard**? Will you create yet another standard with mappings? Where is the full taxonomy by the way (online)?

9: What if I want to use **French names for my links**? What is a link type is not there (Hurt in GRL)?

9.7: Looks more like compliance than validation (as people are not involved). Has anyone used it and provided feedback?

**p. 99 Fig. 18: What are the differences between Containment, Has-A, and Part-Of? Does this violate your Consistency criterion (#3)?** Isn’t there a risk that for a same project, in practice, different links will be used for the same concept, and the same link for different concepts (like coding with IDC10 in medicine)? Users may not understand the subtleties of your taxonomy.

10: Table 17: test 14 does not exist... But more importantly, none of thee tests seem to require the use of a **Trace** (the sequenced group). Useful?

11: good coverage, but avoids URIs, RDF, elements inside models, tools, people with different usages, …

11: Again, no coverage of the **Trace** class from Fig. 13!!!

11: With the presence of URIs, **I can cover this in DOORS too**!

**Is a generic solution something desirable**? What is wrong with domain- or problem-specific solutions (p. 60)? Can your generic solution actually replace all the approaches discussed in your literature review (after all, linking heterogeneous systems surely support linking homogeneous systems)? If not, then where does it break? Usefulness is also related to the analyzability of the traces.

Was the traceability model used by others?

Tools: OSLC is great, but complicated and not perfect. It also needs the individual tools to talk OSLC and provide sufficient services. Have you used that technology?

**Traceability requirements of industrial partner** (CAE). Which are (beyond heterogeneous artefacts)…? Were they satisfied?

## Presentation

* The table of content is missing! The heading for the List of Tables is also missing. I also strongly recommend to insert a list of acronyms (as there are many in the thesis).
* References 6 and 22 are the same!
* References 7, 9, and 11 are the same!
* References 12 and 24 are the same!
* References 31 and 139 are the same!
* References 40 and 136 are the same (even if they have different dates)!
* References 50 and 70 are the same!
* References 79 and 123 are the same!
* References 110 and 125 are the same!
* Many references have missing information on page numbers, publishers, etc. Many also have extra dots for some reason.
* There is no logic in the sorting of the references (not alphabetical, not temporal)