fhiso

Requirement to Separate Evidence and Conclusion

Contents

1. Abstract

This is a functional requirement for the clear separation of evidence and conclusion in the Data Model. Evidence is objective and must be recorded "as is". Conclusions are subjective and there may be different interpretations of the available evidence.

2. Requirement

If we do not distinguish these two types of data then it limits our ability to share accurately with other researchers – researchers who may have a different take on those past events.

The same case applies to online content such as family trees. Most of these already do not make the distinction and are widely criticised as a result. By contrast, a newspaper archive must record what was printed. While it would be permissible to highlight a personal name, a place reference, or a date reference, it would be wholly wrong to provide a subjective interpretation.

This issue affects the use of mark-up in narrative text, and the separate proposal entitled Narrative Text discusses this in more detail. It also applies to properties extracted from a given source such as age and occupation. For instance, an occupation of "charrer" might be someone who burns wood in the making of barrels, or it could be a misspelled version of "charer", as in charwoman. Knowledge of the sex, age, and previous occupations of the person may hazard an educated guess but that is no longer evidence.

It also affects the Data Model's representation of a person, and how the associated attributes (e.g. parentage, date of birth, place of birth, etc.) were derived.

3. Use Cases

A written date may be ambiguous, or may have uncertain characters, or it may rely on context outside of that section of text. For instance, a reference of "Last Friday" could be given an interpretation, and an equivalent machine-readable value, but only if someone includes some reasoning and some context from elsewhere. It's therefore a conclusion and not a verbatim reflection of the written date.

A "Last Friday" reference based on a newspaper publication date, or a letterhead, is less subjective than, say, a person reference, or a place reference for that matter. An article that simply says "John Smith was fined 10s for loitering" is insufficient, in itself, to identify the John Smith. The name is objective - the identification is subjective. The two types of mark-up required are not equivalent.

A date in, say, a Hindu calendar must be recorded "as is". Converting it to the Gregorian calendar is not a true reflection of the evidence. There is no agreed "epoch" between many calendars and so a clean, algorithmic conversion is not always possible. This is explained in the FHISO submission on 'Calendars'.

4. References

The Narrative Text proposal to FHISO distinguishes semantic mark-up that support shallow (objective) or deep (subjective) semantics. Cfps 18, http://fhiso.org/files/cfp/cfps18.pdf.

The Calendars proposal to FHISO explains that applying date conversions from other world calendars to the Gregorian one is a true reflection of the evidence. Cfps 13, http://fhiso.org/files/cfp/cfps13.pdf.

A discussion of the relevance of this separation to mark-up on the historical-schema.org blog site. http://historical-data-schema.blogspot.ie/2012/02/examples-please.html#!/2012/02/examples-please.html.