RDFa Metadata in LaTeX*

Michael Kohlhase
Jacobs University, Bremen
http://kwarc.info/kohlhase

October 8, 2013

Abstract

The `rdfmeta` package allows mark up Ontology-based Metadata in LaTeX documents that can be harvested by automated tools or exported to PDF.

Contents

Experimental! do not use!

*Version v0.2 (last revised 2012/09/23)

rdfmeta.dtx 2371 2013-10-08 10:42:02Z kohlhase

1 Introduction

The `rdfmeta` package allows mark up extensible metadata in SᴛEX documents, so that that it can be harvested by automated tools or exported to PDF. It is also intended to support the new metadata infrastructure for the OMDoc format [Koh06] introduced in OMDoc1.3 [Koh10] (see [LK09] for the relevant ideas and and [KKL10] for an application).

Metadata are annotated as key value pairs in the semantic environments provided by SᴛEX. In most markup formats, the metadata vocabularies are fixed by the language designer. In SᴛEX, the `rdfmeta` package allows the user to extend the metadata vocabulary.

```
\importmodule[../ontologies/cert]{certification}
...
\section[id=userreq,hasState=$\statedocrd{\tuev}$]{User Requirements}
...
```

```
<imports from="../ontologies/cert.omdoc#certification"/>
...
<omgroup xml:id="userreq">
  <metadata>
    <link rel="../ontologies/cert.omdoc#certification:hasState">
      <dc:title>User Requirements</dc:title>
      <resource rel="../ontologies/cert.omdoc#certification/statedocrd"
                resource="../ontologies/cert.omdoc#certification/tuev"/>
    </link>
  </metadata>
...
</omgroup>
```

Example 1: Metadata for Certification

Take, for instance, the case where we want to use metadata for the certification status of document fragments. In Figure 1 we use the `hasState` key to say that a section has been approved by the TÜV, a specific certification agency. There are two concerns here. First, the `hasState` key has to be introduced and given a meaning, and same for the (complex) value `\statedocrd{\tuev}`. This meaning is given in the `certifiation` ontology which we imported via the `\importmodule` command. The ontology can be marked up in SᴛEX (see Figure 2), with the exception that we use the `\keydef` macro for the definition of the `hasState` relation so that it also defines the key. For the details of this see the next section.

2 User Interface

We now document the specifics of the environments and macros provided by the `rdfmeta` package from a user perspective.

2.1 Package Options

`showmeta` The `rdfmeta` package takes the option: `showmeta`. If this is set, then the metadata

Key Group Identifier	Macros	Package/Class
dcm@person	DCMPerson	dcm.sty
dcm@institution	DCMInstitution	dcm.sty
dcm@sect	section	dcm.sty
assig	assignment	hwexam.cls
inclassig	includeassignment	hwexam.cls
quizheading	quizheading	hwexam.cls
testheading	quizheading	hwexam.cls
module	module	modules.sty
termdef	termdef	modules.sty
view	view	modules.sty
omgroup	omgroup	omdoc.sty
ignore	ignore	omdoc.sty
omtext	omtext, definition, axiom, assertion, example, inlinedef	omtext.sty, statements.sty
phrase	phrase	omtext.sty
problem	problem	problem.sty
inclprob	includeproblem	problem.sty
req	requirement	reqdoc.sty
spf	sproof, spfcases, spfcase, spfstep, spfcomment	sproof.sty
termref	termref	statements.sty
symboldec	symboldec	statements.sty

Figure 1: Key Group Identifiers in sTeX

keys are shown (see [Koh12a] for details and customization options).

sectioning

The remaining options can be used to specify metadata upgrades of standard keys. The `sectioning` option upgrades the `\part`, `\chapter`, `\section`, `\subsection`, `\subsubsection`, `\paragraph` macros (and of course their starred variants).

2.2 Extending Macros and Environments by Metadata Keys

\keydef

The main user-visible feature of the `rdfmeta` package is the `keydef` macro. It takes two arguments, a "key group identifier" and a key name. In a nutshell, every sTeX command that takes metadata keys comes with a "key group identifier" that identifies the set of admissible keys; see [Koh12a] for details on this concept. Figure 1 gives an overview over the key groups and their identifiers in sTeX.

Semantically, `\keydef{`⟨*keygroup*⟩`}{`⟨*key*⟩`}` defines a symbol just like the `\symdef` macro from the `modules` package [KGA12]. But it also extends the syntax of sTeX itself: it adds a key ⟨*key*⟩ to ⟨*keygroup*⟩, which allows to state the corresponding metadata as a key/value pair in the sTeX macro or environ-

ment. Following the ideas from [LK09], the metadata is transformed to RDFa metadata [Adi+10] in OMDoc, where the identifiers of relations are exactly the symbols introduced by the corresponding `\keydef`.

```
\documentclass{omdoc}
\usepackage{stex,rdfmeta,amstext}
\begin{document}
\begin{module}[id=certification]
% \metalanguage[../owl2onto/owl2]{OWL2}
  \keydef{omtext}{hasState}
  \keydef{omgroup}{hasState}
  \symdef{hasState}{\text{hasState}}
  \symdef{statedocrd}[1]{rd. #1}
  \symdef{tuev}{\text{T\"UV}}
  \begin{omgroup}[id=foo,hasState=test]{Definitions}
  \begin{definition}[for=hasState]
    A document {\definiendum[hasState]{has state}} $x$, iff
    the project manager decrees it so.
  \end{definition}
  \begin{definition}[for=statedocrd,hasState=$\statedocrd\tuev$]
    A document has state \definiendum[statedocrd]{rd. $x$},
    iff it has been submitted to $x$ for certification.
  \end{definition}
  \begin{definition}[for=tuev,hasState=$\statedocrd\tuev$]
    The $\tuev$ (Technischer \"Uberwachungs Verein) is a national
    certification agency in Germany.
  \end{definition}
\end{omgroup}
\end{module}
\end{document}
```

Example 2: A simple Ontology on Certification

In our example in Figure 2 we have defined a key `hasstate` in the `omtext` key group[1] and a symbol `hasstate` via `\addkey{omtext}{hasstate}`. Furthermore, we have defined the meaning of the relation expressed by the `hasstate` symbol informally and specified some possible objects for the relation (that could of course have been done in other modules as well). We have made use of this metadata ontology and the new key `hasState` in the example in Figure 1.

2.3 Redefinitions of Common LaTeX Macros and Environments

The `rdfmeta` package redefines common LaTeX commands (e. g. the sectioning macros) so that they include optional KeyVal arguments that can be extended by `\keydef` commands. With this extension, we can add RDFa metadata to any existing LaTeX document and generate linked data (XHTML+RDFa documents) via the LaTeXML translator.

[1] For the `\omtext` environment and key group see [Koh12d]

2.4 Extending Packages with rdfmeta

The rdfameta package also exposes its internal infrastructure for extending the redefinitions. Note that the upgrade macros can only be used in LaTeX packages, as the macro names contain @. Consequently, this section is only addressed at package developers who want to extend existing (i.e. not written by them) packages with flexible metadata functionality.

\rdfmeta@upgrade

\rdfmeta@upgrade is the basic upgrade macro. It takes an optional keyval argument an a command sequence ⟨*cseq*⟩ as a proper argument and (if that is defined), redefines \⟨*cseq*⟩ to take a keyval argument. There is a variant

\rdfmeta@upgrade*

\rdfmeta@upgrade* that has to be used to upgrade macros that have a starred form (e.g. \section and friends). Note that \rdfmeta@upgrade* upgrades both forms (e.g. \section and \section*).

\rdfmeta@upgrade uses four keys to specify the behavior in the case the the macro to be upgraded already has an optional argument. For concreteness, we introduce them using the \section macro from standard LaTeX as an example. \section has an optional argument for the "short title", which

optarg

will appear in the table of contents. The optarg key can be used to specify a key for the existing optional argument. Thus, after upgrading it via \rdfmeta@upgrade*[optarg=short]{section}, we can use the updated form \section[short=⟨*toctitle*⟩]{⟨*title*⟩} instead of the old \section[⟨*toctitle*⟩]{⟨*title*⟩}. Actually, this still has a problem: the \section* would also be given the short key and would be passed an optional argument (which it does not accept). To rem-

optargstar

edy this we can set the optargstar key to no. In summary, the correct upgrade command for \section and \section* would be

```
\rdfmeta@upgrade*[optarg=short,optargstar=no]{section}
```

The \rdfmeta@upgrade* macro also initializes a metadata key-group (a named set of keys and their handlers; see [Koh12b] for details) for the section macro with an id key for identification (see [Koh12e] for details). Often, the name of the key-group is the same as the command sequence, so we take this as the default, if we want to specify a different metadata key-group name, we can do so with the

keygroup

keygroup key in \rdfmeta@upgrade*.

idlabel

If idlabel is set to ⟨*prefix*⟩, then the LaTeX label is set to the value ⟨*prefix*⟩.⟨*id*⟩, where ⟨*id*⟩ is the value given in the RDFa id key. This allows to use the normal LaTeX referencing mechanism in addition to the semantic referencing mechanism provided by the sref package [Koh12f].

2.5 Limitations

In this section we document known limitations. If you want to help alleviate them, please feel free to contact the package author. Some of them are currently discussed in the sTeX TRAC [sTeX].

1. Currently the coverage of the redefinitions of standard commands in the rdfmeta package is minimal; we will extend this in the future.

2. The `\rdfmeta@upgrade` macro only works with single arguments, this should be easy to fix with `\case` for the argument string.
3. I am not sure `\rdfmeta@upgrade` works with environments.
4. it would be convenient, if we had a macro `\keydefs`, which takes a list of keygroups, so that we can define keys in multiple groups in one go, e.g. `\keydefs{omtext,omgroup}{hasState}` in Figure 2. But the obvious "solution"

```
\newcommand\keydefs[2]{\@for\@I:=#1\do{\keydef{#1}{#2}}}
```

   does not work for me.

3 The Implementation

The `sref` package generates two files: the LaTeX package (all the code between ⟨*package⟩ and ⟨/package⟩) and the LaTeXML bindings (between ⟨*ltxml⟩ and ⟨/ltxml⟩). We keep the corresponding code fragments together, since the documentation applies to both of them and to prevent them from getting out of sync.

We first set up header information for the LaTeXML binding file.

```
⟨*ltxml⟩
package LaTeXML::Package::Pool;
use strict;
use LaTeXML::Package;
⟨/ltxml⟩
```

3.1 Package Options

We declare some switches which will modify the behavior according to the package options. Generally, an option `xxx` will just set the appropriate switches to true (otherwise they stay false).[1]

EdN:1

```
⟨*package⟩
\DeclareOption{showmeta}{\PassOptionsToPackage{\CurrentOption}{metakeys}}
\newif\if@rdfmeta@sectioning\@rdfmeta@sectioningfalse
\DeclareOption{sectioning}{\@rdfmeta@sectioningtrue}
\ProcessOptions
⟨/package⟩
⟨*ltxml⟩
DeclareOption('showmeta','');
DeclareOption('sectioning','');
⟨/ltxml⟩
```

The first measure is to ensure that the right packages are loaded. From the from STeX collection, we need the `sref` package (see [Koh12f]) for handling keys, the `modules` package for exporting the `\keydef` (see [KGA12]).

```
⟨*package⟩
\RequirePackage{sref,modules}
```

[1] EdNote: need an implementation for LaTeXML

\rdfmeta@loaded and we define a macro `rdfmeta@loaded` just for the purpose of determining whether the `rdfmeta` package is loaded.

\rdfmeta@loaded

```
\newcommand\rdfmeta@loaded{yes}
```

\rdfmeta@sectioning And another macro `rdfmeta@sectioning` to determine wether the sectioning macros have been redefined.

\rdfmeta@loaded

```
\if@rdfmeta@sectioning\newcommand\rdfmeta@sectioning{yes}\fi
⟨/package⟩
```

3.2 Key Definitions

\keydef The `\keydef` macro is rather simple, we just add a key to the respective environment and extend the export token register for the current module by an `\addmetakey` instruction.

```
⟨*package⟩
\newcommand\keydef[2]{\addmetakey{#1}{#2}%
\expandafter\g@addto@macro\this@module{\addmetakey{#1}{#2}}}
⟨/package⟩
⟨*ltxml⟩
DefConstructor('\keydef','');
⟨/ltxml⟩
```

\listkeydef The `\listkeydef` macro is analogous, but uses `\addmetalistkey` instead. instruction.

```
⟨*package⟩
\newcommand\listkeydef[2]{\addmetalistkey{#1}{#2}%
\expandafter\g@addto@macro\this@module{\addmetalistkey{#1}{#2}}}
⟨/package⟩
⟨*ltxml⟩
DefConstructor('\listkeydef','');
⟨/ltxml⟩
```

3.3 RDFa upgrade Facilities

We first define the keys for the `\rdfmeta@upgrade` macro.

```
⟨*package⟩
\def\@yes@{yes}
\addmetakey*{upgrade}{idlabel}
\addmetakey*{upgrade}{optarg}
\addmetakey*[yes]{upgrade}{optargstar}
\addmetakey*{upgrade}{keygroup}
```

\rdfmeta@upgrade This upgrade macro gives extended functionality according to the optional keys. The top-level invocation just differentiates on whether a star is following:

```
\def\rdfmeta@upgrade{\@ifstar\rdfmeta@upgrade@star\rdfmeta@upgrade@nostar}
```

Both cases are almost the same, they only differ in the third line where they call `\rdfmeta@upgrade@base` or `\rdfmeta@upgrade@base@star` defined above. In particular, both take the arguments originally intended for `\rdfmeta@upgrade`.

```
\newcommand\rdfmeta@upgrade@nostar[2][]{\metasetkeys{upgrade}{#1}%
\ifx\upgrade@keygroup\@empty\def\@@group{#2}\else\def\@@group{\upgrade@keygroup}\fi
\rdfmeta@upgrade@base{#2}{\@nameuse{\@@group @\upgrade@optarg}}}
```

They set the metakeys from the second argument, then set `\@@group` to be the intended group (if the `keygroup` key was specified, it takes precedence over the default `#2`).

```
\newcommand\rdfmeta@upgrade@star[2][]{\metasetkeys{upgrade}{#1}%
\ifx\upgrade@keygroup\@empty\def\@@group{#2}\else\def\@@group{\upgrade@keygroup}\fi
\rdfmeta@upgrade@base@star{#2}{\@nameuse{\@@group @\upgrade@optarg}}}
⟨/package⟩
⟨*ltxml⟩
⟨/ltxml⟩
```

`\rdfmeta@upgrade@base` This auxiliary macro and is invoked as `\rdfmeta@upgrade@base{`⟨*cseq*⟩`}{`⟨*optarg*⟩`}`, where ⟨*cseq*⟩ is a command sequence name. It checks if `\`⟨*cseq*⟩ is defined (if not it does nothing), saves the old behavior of `\`⟨*cseq*⟩ as `\rdfmeta@`⟨*cseq*⟩`@old`, and then redefines `\`⟨*cseq*⟩ to take a keyval argument and passes ⟨*optarg*⟩ as the optional argument.

```
⟨*package⟩
\newcommand\rdfmeta@upgrade@base[2]{\@ifundefined{#1}{}%
{\message{redefining macro #1,}
\ifx\upgrade@idlabel\@empty\srefaddidkey{#1}\else\srefaddidkey[prefix=\upgrade@idlabel]{#1}\fi%
\expandafter\let\csname rdfmeta@#1@old\expandafter\endcsname\csname #1\endcsname%
\expandafter\renewcommand\csname #1\endcsname[2][]%
{\metasetkeys{#1}{##1}\@nameuse{rdfmeta@#1@old}[#2]{##2}}
\addmetakey*\@@group{\upgrade@optarg}}}
```

`\rdfmeta@upgrade@base@star` This is a variant of `\rdfmeta@upgrade@base`, which also takes care of the starred variants of a macro.

```
\newcommand\rdfmeta@upgrade@base@star[2]{\@ifundefined{#1}{}%
{\message{redefining macros #1 and #1*,}
\ifx\upgrade@idlabel\@empty\srefaddidkey{#1}\else\srefaddidkey[prefix=\upgrade@idlabel]{#1}\fi%
\expandafter\let\csname rdfmeta@#1@old\expandafter\endcsname\csname #1\endcsname%
```

In this case, we cannot just use `\newcommand` for dealing with the optional argument because the star is between the command sequence and the arguments. So we make a case distinction on the presence of the star. `\rdfmeta@`⟨*cseq*⟩`@old`.

```
\expandafter\renewcommand\csname #1\endcsname%
{\@ifstar{\@nameuse{rdfmeta@#1@star}}{\@nameuse{rdfmeta@#1@nostar}}}%
```

the macros `\rdfmeta@`⟨*cseq*⟩`@star` and `\rdfmeta@`⟨*cseq*⟩`@nostar` that are defined in terms of `\rdfmeta@`⟨*cseq*⟩`@old` handle the necessary cases. The second one is simple:

```
\expandafter\newcommand\csname rdfmeta@#1@nostar\endcsname[2][]%
{\metasetkeys{#1}{##1}\edef\@test{#2}%
```

```
\ifx\@test\@empty\@nameuse{rdfmeta@#1@old}{##2}%
\else\@nameuse{rdfmeta@#1@old}[#2]{##2}\fi}%
```

For `\rdfmeta@`⟨*cseq*⟩`@star` we have to take care of the optional argument of the old macro: if the `optargstar` key was set, then we pass the second argument of `\rdfmeta@upgrade@base` as an optional argument to it as above.

```
\ifx\upgrade@optargstar\@yes@%
\expandafter\newcommand\csname rdfmeta@#1@star\endcsname[2][]%
{\metasetkeys{#1}{##1}\@nameuse{rdfmeta@#1@old}*[#2]{##2}}%
\else%
\expandafter\newcommand\csname rdfmeta@#1@star\endcsname[2][]%
{\metasetkeys{#1}{##1}\@nameuse{rdfmeta@#1@old}*{##2}}%
\fi%
\addmetakey*\@@group{\upgrade@optarg}}}
⟨/package⟩
```

3.4 Redefinitions

If the `sectioning` macro is set, we redefine the respective commands

```
⟨*package⟩
\if@rdfmeta@sectioning
\message{redefining sectioning commands!}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{part}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{chapter}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{section}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{subsection}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{subsubsection}
\rdfmeta@upgrade*[optarg=short,optargstar=no]{paragraph}
\fi
⟨/package⟩
⟨*ltxml⟩
⟨/ltxml⟩
```

3.5 Finale

Finally, we need to terminate the file with a success mark for perl.

```
⟨ltxml⟩1;
```

References

[Adi+10] Ben Adida et al. *RDFa Core 1.1. Syntax and processing rules for embedding RDF through attributes*. W3C Working Draft. World Wide Web Consortium (W3C), Aug. 3, 2010. URL: `http://www.w3.org/TR/2010/WD-rdfa-core-20100803/`.

[KGA12] Michael Kohlhase, Deyan Ginev, and Rares Ambrus. *modules.sty: Semantic Macros and Module Scoping in sTeX*. Self-documenting LaTeX package. Comprehensive TeX Archive Network (CTAN), 2012. URL: http://www.ctan.org/get/macros/latex/contrib/stex/modules/modules.pdf.

[KKL10] Andrea Kohlhase, Michael Kohlhase, and Christoph Lange. "sTeX – A System for Flexible Formalization of Linked Data". In: *Proceedings of the 6th International Conference on Semantic Systems (I-Semantics) and the 5th International Conference on Pragmatic Web*. Ed. by Adrian Paschke et al. ACM, 2010. ISBN: 978-1-4503-0014-8. DOI: 10.1145/1839707.1839712. arXiv: 1006.4474v1 [cs.SE]. URL: http://portal.acm.org/citation.cfm?id=1839707.

[Koh06] Michael Kohlhase. OMDoc – *An open markup format for mathematical documents [Version 1.2]*. LNAI 4180. Springer Verlag, Aug. 2006. URL: http://omdoc.org/pubs/omdoc1.2.pdf.

[Koh10] Michael Kohlhase. "An Open Markup Format for Mathematical Documents OMDoc [Version 1.3]". Draft Specification. 2010. URL: https://svn.omdoc.org/repos/omdoc/branches/omdoc-1.3/doc/spec/main.pdf.

[Koh12a] Michael Kohlhase. *metakeys.sty: A generic framework for extensible Metadata in LaTeX*. Self-documenting LaTeX package. Comprehensive TeX Archive Network (CTAN), 2012. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/metakeys/metakeys.pdf.

[Koh12b] Michael Kohlhase. *metakeys.sty: A generic framework for extensible Metadata in LaTeX*. Self-documenting LaTeX package. 2012. URL: https://svn.kwarc.info/repos/stex/trunk/sty/metakeys/metakeys.pdf.

[Koh12c] Michael Kohlhase. *metakeys.sty: A generic framework for extensible Metadata in LaTeX*. Self-documenting LaTeX package. 2012.

[Koh12d] Michael Kohlhase. *omtext: Semantic Markup for Mathematical Text Fragments in LaTeX*. Self-documenting LaTeX package. Comprehensive TeX Archive Network (CTAN), 2012. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/omtext/omtext.pdf.

[Koh12e] Michael Kohlhase. *sref.sty: Semantic Crossreferencing in LaTeX*. Self-documenting LaTeX package. 2012. URL: https://svn.kwarc.info/repos/stex/trunk/sty/sref/sref.pdf.

[Koh12f] Michael Kohlhase. *sref.sty: Semantic Crossreferencing in LaTeX*. Self-documenting LaTeX package. Comprehensive TeX Archive Network (CTAN), 2012. URL: http://www.ctan.org/tex-archive/macros/latex/contrib/stex/sref/sref.pdf.

[Koh12g] Michael Kohlhase. *sref.sty: Semantic Crossreferencing in LaTeX*. Self-documenting LaTeX package. 2012.

[LK09] Christoph Lange and Michael Kohlhase. "A Mathematical Approach to Ontology Authoring and Documentation". In: *MKM/Calculemus Proceedings*. Ed. by Jacques Carette et al. LNAI 5625. Springer Verlag, July 2009, pp. 389–404. ISBN: 978-3-642-02613-3. URL: `http://kwarc.info/kohlhase/papers/mkm09-omdoc4onto.pdf`.

[sTeX] *Semantic Markup for LaTeX*. Project Homepage. URL: `http://trac.kwarc.info/sTeX/`.