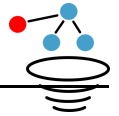


SciKGT_EX



Scientific Knowledge Graph TeX (SciKGT_EX) makes it possible to annotate specific **research contributions** directly in the LaTeX source code. The idea of extracting research contributions from scholarly publications is derived from the [Open Research Knowledge Graph \(ORKG\)](#).

With SciKGT_EX, authors of scientific publications can enrich their documents with structured, reduced and machine-readable information which represents the key points of the content they want to communicate. The production of this additional information improves electronic archiving of the information for the future and boosts discoverability in search engines and recommender systems. The contribution data is embedded into the PDF's XMP metadata where it can be retrieved by anyone who obtains the PDF document and persists for the lifetime of the document. Additionally, the contributions can easily be added automatically to scientific knowledge graphs like the ORKG.

Table of Contents

- [Installation](#)
 - [Template on Overleaf](#)
- [Using the LaTeX Package](#)
 - [Minimal Example](#)
 - [Optional Properties](#)
 - [Contribution Numbering](#)
 - [Invisible Markup](#)
 - [Referring to Entities](#)
 - [Defining Custom Properties](#)
- [Compatibility](#)
- [Testing](#)
- [License](#)

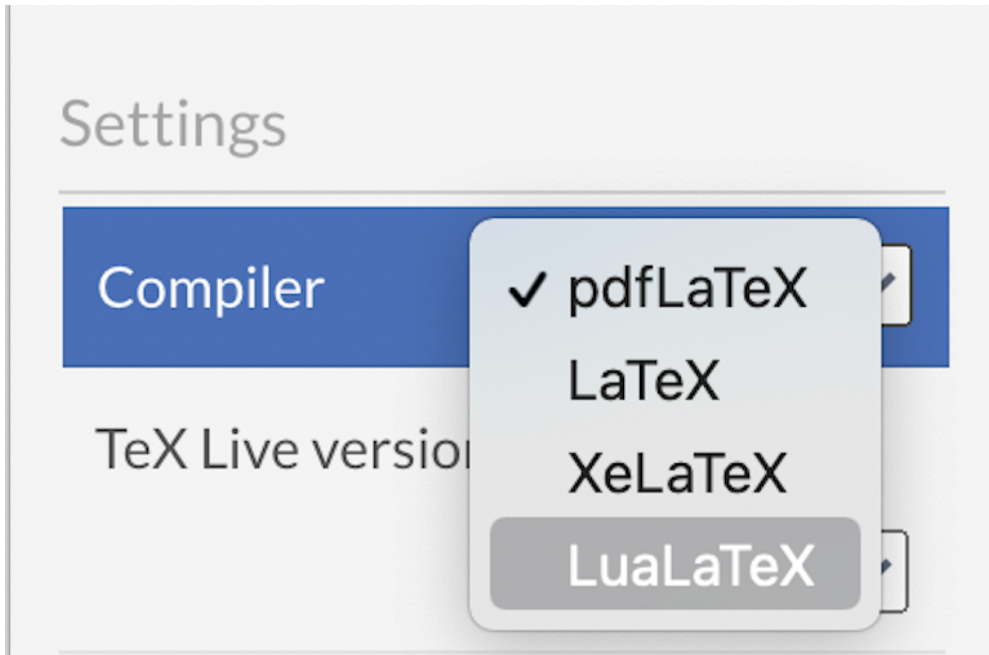
Installation

1. Copy the files `scikgtex.lua` and `scikgtex.sty` from the latest stable release to your latex project
 - latest stable release: <https://github.com/Christof93/SciKGT_EX/releases/tag/v2.2.0>
2. Set `\usepackage{scikgtex}` in your document preamble to use the package

Template on Overleaf

You can use this link: <https://www.overleaf.com/latex/examples/scikgtex-example/wrhmyrwfgrgw> to view a working example in Overleaf with some explanation. Open the template and start making your own metadata annotations right away!

It is necessary to compile your LaTeX source with LuaLaTeX for the package to work. This is typically straightforward with most modern LaTeX environments. In Overleaf it can be configured like this for example:



Using the LaTeX Package

Minimal Example

To create a minimal contribution we have to assign one of the 5 standard contribution properties to sentences or statements in the text:

- *research problem*
- *objective*
- *method*
- *result*
- *conclusion*

For each of these properties there exists a corresponding LaTeX command:

- `\researchproblem{..}`
- `\objective{..}`
- `\method{..}`
- `\result{..}`
- `\conclusion{..}`

Additionally we can mark title, authors and research field of the document with these three commands:

- `\metatitle{..}`
- `\metaauthor{..}`
- `\researchfield{..}`

We can surround expressions and passages in the text with these commands as illustrated by this example of a summarized research article by Wald, Ellen R., David Nash, and Jens Eickhoff on "Effectiveness of Amoxicillin/Clavulanate Potassium in the Treatment of Acute Bacterial Sinusitis in Children." *Pediatrics*, vol. 124, no. 1, 2009, pp. 9-15.:

```
\documentclass{article}
\usepackage{scikgtex}

\title{\metatitle{Effectiveness of Amoxicillin/Clavulanate Potassium in
the Treatment of Acute Bacterial Sinusitis in Children.}}
\author{\metaauthor{Ellen R. Wald} \and \metaauthor{David Nash}
\metaauthor{Jens Eickhoff}}
\researchfield*{antibiotic therapy}

\begin{document}
\maketitle

\begin{abstract}
The role of \researchproblem{antibiotic therapy in managing acute
bacterial sinusitis (ABS) in children} is controversial.
The purpose of this study was to \objective{determine the effectiveness of
high-dose amoxicillin/potassium clavulanate in the treatment of children
diagnosed with ABS}.

This was a \method{randomized, double-blind, placebo-controlled study}.
Children 1 to 10 years of age with a clinical presentation compatible with
ABS were eligible for participation.
\method{Patients were stratified according to age (<6 or ≥6 years) and
clinical severity and randomly assigned to receive either amoxicillin (90
mg/kg) with potassium clavulanate (6.4 mg/kg) or placebo}.
A symptom survey was performed on days 0, 1, 2, 3, 5, 7, 10, 20, and 30.
Patients were examined on day 14.
Children's conditions were rated as cured, improved, or failed according
to scoring rules.

Two thousand one hundred thirty-five children with respiratory complaints
were screened for enrollment; 139 (6.5\%) had ABS.
Fifty-eight patients were enrolled, and 56 were randomly assigned. The
mean age was 6630 months.
Fifty (89\%) patients presented with persistent symptoms, and 6 (11\%)
presented with nonpersistent symptoms.
In 24 (43\%) children, the illness was classified as mild, whereas in the
remaining 32 (57\%) children it was severe.
Of the 28 children who received the antibiotic, 14 (50\%) were cured, 4
(14\%) were improved, 4(14\%) experienced treatment failure, and 6 (21\%)
withdrew.
Of the 28children who received placebo, 4 (14\%) were cured, 5 (18\%)
improved, and 19 (68\%) experienced treatment failure.
\result{Children receiving the antibiotic were more likely to be cured
(50\% vs 14\%) and less likely to have treatment failure (14\% vs 68\%)
than children receiving the placebo}.
ABS is a common complication of viral upper respiratory infections.
\conclusion{Amoxicillin/potassium clavulanate results in significantly
more cures and fewer failures than placebo}, according to parental report
of time to resolution."
\end{abstract}
\end{document}
```

The produced document will then look like this:

Effectiveness of Amoxicillin/Clavulanate Potassium in the Treatment of Acute Bacterial Sinusitis in Children.

Ellen R. Wald David Nash Jens Eickhoff

December 19, 2021

Abstract

The role of antibiotic therapy in managing acute bacterial sinusitis (ABS) in children is controversial. The purpose of this study was to determine the effectiveness of high-dose amoxicillin/potassium clavulanate in the treatment of children diagnosed with ABS.

This was a randomized, double-blind, placebo-controlled study. Children 1 to 10 years of age with a clinical presentation compatible with ABS were eligible for participation. Patients were stratified according to age (<6 or ≥6 years) and clinical severity and randomly assigned to receive either amoxicillin (90 mg/kg) with potassium clavulanate (6.4 mg/kg) or placebo. A symptom survey was performed on days 0, 1, 2, 3, 5, 7, 10, 20, and 30. Patients were examined on day 14. Children's conditions were rated as cured, improved, or failed according to scoring rules.

Two thousand one hundred thirty-five children with respiratory complaints were screened for enrollment; 139 (6.5%) had ABS. Fifty-eight patients were enrolled, and 56 were randomly assigned. The mean age was 6630 months. Fifty (89%) patients presented with persistent symptoms, and 6 (11%) presented with nonpersistent symptoms. In 24 (43%) children, the illness was classified as mild, whereas in the remaining 32 (57%) children, the illness was classified as moderate to severe. Of the 56 children who received the antibiotic

As can be seen in the rendered pdf the marked properties can not be distinguished from the other sentences in the text. The annotations can be inspected in the file `xmp_metadata.xml`. The XMP file can be used as an inspection possibility for the user but it is not necessary to distribute it since the whole content is also directly embedded into the produced PDF file in the creation process. For our example the content of the metadata will look as such:

```
<?xpacket begin="?" id="48fdc517-5814-4d0c-cd03-0c296941c6"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/">
<rdf:RDF
  xmlns:orkg="http://orkg.org/core#"
  xmlns:orkg_property="http://orkg.org/property/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdf:Description rdf:about="https://www.orkg.org/orkg/paper/48fdc517-
5814-4d0c-cd03-0c296941c6">
    <rdf:type rdf:resource="http://orkg.org/core#Paper"/>
    <orkg:hasTitle>Effectiveness of Amoxicillin/Clavulanate Potassium in
the Treatment of Acute Bacterial Sinusitis in Children.</orkg:hasTitle>
    <orkg:hasAuthor>Ellen R. Wald</orkg:hasAuthor>
    <orkg:hasAuthor>David Nash</orkg:hasAuthor>
    <orkg:hasAuthor>Jens Eickhoff</orkg:hasAuthor>
    <orkg:hasResearchField>antibiotic therapy</orkg:hasResearchField>
```

```

<orkg:hasResearchContribution>
  <orkg:ResearchContribution
rdf:about="https://www.orkg.org/orkg/paper/48fdc517-5814-4d0c-cd03-
0c296941c6/contribution_ORKG_default">
  <orkg_property:researchproblem>antibiotic therapy in managing
acute bacterial sinusitis (ABS) in
children</orkg_property:researchproblem>
  <orkg_property:objective>effectiveness of high-dose
amoxicillin/potassium clavulanate in the treatment of children diagnosed
with ABS</orkg_property:objective>
  <orkg_property:method>randomized, double-blind, placebo-
controlled study</orkg_property:method>
  <orkg_property:method>Patients were stratified according to age
(&lt;6 or ≥6 years) and clinical severity and randomly assigned to receive
either amoxicillin (90 mg/kg) with potassium clavulanate (6.4 mg/kg) or
placebo</orkg_property:method>
  <orkg_property:result>Children receiving the antibiotic were
more likely to be cured (50% vs 14%) and less likely to have treatment
failure (14% vs 68%) than children receiving the
placebo</orkg_property:result>
  <orkg_property:conclusion>Amoxicillin/potassium clavulanate
results in significantly more cures and fewer failures than
placebo</orkg_property:conclusion>
  </orkg:ResearchContribution>
</orkg:hasResearchContribution>
</rdf:Description>
</rdf:RDF>
</x:xmpmeta>
<?xpacket end="r"?>

```

Optional Properties

Additional to the 5 standard ones, there are a big number of more specific properties which are optional and are generally used in a specific domain of science. For example properties of *p-value* or *accuracy* are useful for studies that include statistical examinations and can be attached to a contribution with `\contribution{p-value}{0.05}` and `\contribution{accuracy}{0.876}`. It is far more valuable to use properties which were also used by other people. Consult the [ORKG web portal](#) to find properties which are already used on papers related to your work.

Contribution Numbering

A scientific paper typically has a small number of distinct contributions. If we want to distinguish more than one contribution, we can number the contributions in the arguments of the marked properties. For example, an annotation of `\researchproblem[1]{..}` and `\researchproblem[2]{..}` adds two contributions with the respective research problems. These two contributions can have their own objectives, methods and results which must be numbered accordingly.

If two problems have a property in common (e.g. the same objective, or methods), we can assign the same property to two contributions using a comma between the arguments. Example: `\method[1,2]{..}`.

Invisible Markup

At some point, especially for advanced modeling, it will be desirable to add annotations to the metadata which are not explicitly rendered in text. This can be achieved using the starred variant of the defined or self-defined LaTeX commands.

For example, in the sentence 'the p-value was 0.01% higher than in the earlier experiment', we may want to report the actual p-value in the metadata since it is more clean. There might be many more subtle examples where for some reason you want the information in the metadata to look slightly different than in text. In such a case we can mark the p-value like this:

```
... the p-value was 0.01\% higher \contribution*{p-value}{0.06} than in
the earlier experiment ...
```

In the rendered sentence the content of the annotation (0.06) will be invisible. Since it is not rendered, it does not matter where in the document we put the command. This means that we can define a whole block of custom metadata independent of the text if needed. Let's say we have a result which did not make it into the final paper but can still be added to the metadata.

```
We tested the same experiment with different parameters and could not
reproduce the result.
\conclusion*[2]{No improvements}
\contribution*[2]{number of iterations}{10000}
\result*[2]{52\% accuracy \contribution*[2]{accuracy}{0.52}}
\method*[2]{Support Vector Machine}
```

Referring to Entities

Instead of using natural language to represent objects, we usually prefer URIs which uniquely identify resources in the Semantic Web. If we want to assign a URI as a property, we can use the `\uri{}` command inside an annotation.

```
\documentclass{article}
\usepackage{scikgtex}

\begin{document}
% adds a link to the URI of an entity as researchproblem
The role of
\researchproblem{\uri{https://www.orkg.org/orkg/resource/R12259}
{antibiotic therapy}} in managing acute bacterial sinusitis (ABS) in
children is controversial...
\end{document}
```

The `\uri` command takes the URI of an entity as a first argument and an optional label as the second. If a label is given, it is rendered as a hyperlink to the URI as can be seen in the picture.

1 Our research

The role of antibiotic therapy in managing acute bacterial sinusitis (ABS) in children is controversial.

<https://www.orkg.org/orkg/resource/R12259>

In the XMP metadata file this will result in a new node either with or without a label.

```
...
<orkg_property:researchproblem>
  <rdf:Description rdf:about="https://www.orkg.org/orkg/resource/R12259">
    <rdfs:label>antibiotic therapy</rdfs:label>
  </rdf:Description>
</orkg_property:researchproblem>
...
```

To add the entity without the hyperlink we can specify the entity with invisible markup like this for example:

```
The role of \researchproblem*
{\uri{https://www.orkg.org/orkg/resource/R12259}} antibiotic therapy in
managing acute bacterial sinusitis (ABS) in children is controversial...
```

Defining Custom Properties

It is possible to declare new properties (with custom namespaces) in your document preamble with `\addmetaproperty`. The namespace specification must be added as a comma-separated pair of prefix and uri in the optional argument. If no namespace is defined the self-defined properties will be added to the namespace `http://orkg.org/property` (which is also the case if you just use them with out defining).

However, it might occur that you want to use a property which is already defined semantically in an ontology on the web. In this case you can add the namespace to the annotation schema and use the property in you text. For example, suppose we want to use a property of an already existing ontology like the [argument model ontology](#). We can use their property `has_claim` by defining it in our preamble with the `\addmetaproperty` command like this:

```
\documentclass{article}
\usepackage{scikgtex}
\addmetaproperty[amo, http://purl.org/spar/amo#]{has_claim}
\begin{document}
We make the claim that \contribution{has_claim}{the earth is round}.
...
```

If there are properties with the same name but different meaning depending on the context we can specify the namespace in the actual annotation as well. For example there is a meaning of claim as a legal term denoting the extent of a patent application. Using two properties with the same name in a document can be handled like this:

```
\documentclass{article}
\usepackage{scikgtex}
\addmetaproperty[amo, http://purl.org/spar/amo#]{has_claim}
\addmetaproperty[patent, https://other.type/of/]{has_claim}
\begin{document}
\contribution{amo:has_claim}{The earth is round}.
Our patent has the following claim: \contribution{patent:has_claim}{An
apparatus for catching mice,
said apparatus comprising a base, a spring member coupled to the base, and
a cage}.
...
```

The metadata will list the custom namespace and correctly apply it to the annotations of the property.

Compatibility

By default SciKGT_EX overwrites the PDF metadata catalog and is not fully compatible with other metadata specification packages such as hyperxmp or pdfx. On top of that, the generated metadata currently can not comply with the PDF/A standard because of the user-defined properties. When submitting to conferences or other subsequent processors data written in the standard PDF metadata field is often erased. To prevent this use the SciKGT_EX compatibility mode.

The compatibility mode can be activated by importing the package with the compatibility parameter:

```
\usepackage[compatibility]{scikgtex}
```

In compatibility mode the metadata stream is linked to a new key in the PDF catalog called **SciKGMetad_{ata}**. Still, the metadata can be easily retrieved from the file by locating and extracting the uncompressed metadata stream.

This way the metadata can exist besides the standard metadata objects. Also, it will usually not be erased by any PDF to PDF/A conversion tools.

An example of how to retrieve SciKGT_EX metadata written in PDF/A compatibility mode can be found [here](#).

Testing

A number of integration tests can be run with:

```
sh test/run.sh
```


If desired individual test can be run with:

```
sh test/run.sh <directory_name_of_test>
```

To create a new test, make a new directory starting with the word test and copy the `run_test.sh` script into it. Then, add your LaTeX file called `test.tex` and a file with the expected metadata you want to test against (`xmp_metadata_expected.xml`). Make changes to `run_test.sh` to change the integration test as you see fit.

License

This work is published under [MIT License](#).

Copyright (C) 2022 by Christof Bless

Email: christofbless@gmail.com

Github: <https://github.com/Christof93>