

# Typesetting XML with ConTeXt

PKP 2019 Conference Barcelona

---

Denis Maier

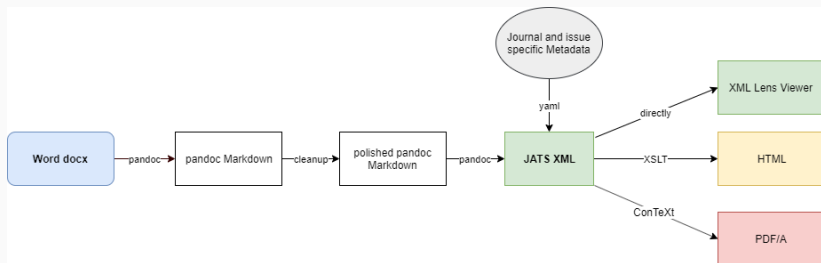
22. November 2019

University Library Bern

# Starting position

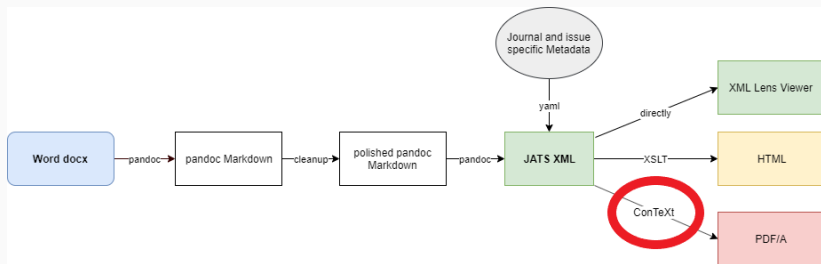
- Print journal “JUDAICA” => e-only OA journal on Bern University Publishing
- The Task:
  - Single source workflow for web view and typeset output (PDF, XML, HTML)
  - JATS XML input
  - PDF: high quality typography without manual typesetting
  - PDF/A
  - Open source
- Solutions:
  - Lens Viewer
  - XSLT for HTML
  - But how should we produce PDF from XML?

# Our Workflow



Source: Maier, Denis Christian; Stutzmann, Jan (21 November 2019). JATS for Judaica. Designing an XML-first Publication Workflow for a Humanities Journal. In: PKP 2019 Conference Barcelona. Universitat Autònoma de Barcelona. 20-22 Nov 2019. <https://doi.org/10.7892/boris.135828>

# Our Workflow



- TeX, markup based plain-text typesetting system, by Donald Knuth, 1977–1986
- LaTeX: Macropackage for TeX, initially by Leslie Lamport, 1980s–today
- ConTeXt: Macropackage for TeX, by PRAGMA ADE, 1990–today

## Why a TeX-based solution?

- High quality typesetting (Knuth-Plass linebreaking algorithm, Microtypography)
- Customizable
- Multiplattform
- Open Source

- Consistent interface, predictable commands
- PDF/A out of the box
- Native XML processing

## How does it look like

```
1 \setupbodyfontenvironment[default][em=italic] % use italic as em, not slanted
2
3 \definefontfamily[mainface][rm][Linux Libertine 0]
4
5 \setupbodyfont[mainface,11pt]
6
7 \setuppapersize[A5][A5]
8
9 \setuppagenumbering[location={footer,center}]
10
11 \setupindenting[yes,medium,next]
12
13 \setuphead [section, subsection] [number=yes]
14 \setuphead [section] [style=\bf]
15 \setuphead [subsection] [style=\em]
16
17 \starttext
18
19 \section{Introduction}
20 Lorem ipsum {\em dolor sit amet}, consetetur sadipscing elitr,
21 sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat,
22 sed diam voluptua.\footnote{This is a footnote.}
23
24 \subsection{Methods}
25 Lorem ipsum dolor sit amet, consetetur sadipscing elitr,
26 sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat,
27 sed diam voluptua.
28
29 \stoptext
```



- XML input
- ConText style setups
- Mappings from XML elements to ConTeXt

# ConTeXt XML processing (1): A minimal JATS XML article

```
4 <article xmlns:mml="http://www.w3.org/1998/Math/MathML" xmlns:xlink="http://www
5 <front>
30 <body>
31 <sec id="introduction">
32 <title>Introduction</title>
33 <p>Lorem ipsum <italic>dolor sit amet</italic>, consetetur sadipscing
34 elit, sed diam nonumy eirmod tempor invidunt ut labore et dolore
35 magna aliquyam erat, sed diam
36 voluptua.<xref ref-type="fn" rid="fn1">1</xref></p>
37 <list list-type="bullet">
38 <list-item>
39 | <p>A list item</p>
40 </list-item>
41 <list-item>
42 | <p>Another list item</p>
43 </list-item>
44 </list>
45 <p>Lorem ipsum <italic>dolor sit amet</italic>, consetetur sadipscing
46 elit, sed diam nonumy eirmod tempor invidunt ut labore et dolore
47 magna aliquyam erat, sed diam
48 voluptua.<xref ref-type="fn" rid="fn2">2</xref></p>
49 <disp-quote>
50 <p>Lorem ipsum dolor sit amet, consetetur sadipscing elit, sed diam
51 nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam
52 erat, sed diam voluptua.</p>
53 </disp-quote>
54 <p>Lorem ipsum dolor sit amet, consetetur sadipscing elit, sed diam
55 nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat,
56 sed diam voluptua.<xref ref-type="fn" rid="fn3">3</xref></p>
57 </sec>
58 <sec id="conclusions">
59 <title>Conclusions</title>
60 <p>Lorem ipsum dolor sit amet, consetetur sadipscing elit, sed diam
61 nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat,
62 sed diam voluptua.</p>
63 </sec>
64 </body>
65 <back>
78 </article>
```

```
\startxmlsetups xml:jatssetups
  % Discard everything
  \xmlsetsetup{#1}{*}{-}
  % Select main nodes
  \xmlsetsetup{#1}{article|front|body|back}{xml:*}
  % Formatting
  \xmlsetsetup{#1}{bold|italic}{xml:*}
\stopxmlsetups
\xmlregistersetup{xml:jatssetups}
```

```
\startxmlsetups xml:body  
  \directsetup{document:start}  
  \xmlflush{#1}  
\stopxmlsetups
```

```
\startxmlsetups xml:p
  \begingroup
    % Get language attribute if available
    \xmlfilter{#1}{.[@xml:lang]/command(xml:language)}
    % Flush content
    \xmlflush{#1}
  \endgroup
  % end paragraph
  \par
\stopxmlsetups
```

```
\startxmlsetups xml:italic  
  {\em \xmlflush{#1}}  
\stopxmlsetups
```

```
\startxmlsetups xml:bold  
  {\bf \xmlflush{#1}}  
\stopxmlsetups
```

## A very important article

*John Doe*

### 1 Introduction

Lorem ipsum *dolor sit amet*, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.<sup>1</sup>

- A list item
- Another list item

Lorem ipsum *dolor sit amet*, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.<sup>2</sup>

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.<sup>3</sup>

### 2 Conclusions

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.

---

<sup>1</sup> Footnote text.

<sup>2</sup> Footnote text.

<sup>3</sup> Another footnote.

## So? Should I use this?

MAYBE

- It works
- It meets our requirements

But:

- Yet another tool in the toolchain
- You need someone that is familiar with it



Email: [denis.maier@ub.unibe.ch](mailto:denis.maier@ub.unibe.ch)

Github: <https://github.com/denismaier/jats-demo-workflow>

Licensed under a Creative Commons Attribution 4.0 International License.

