

# **Formatting JATS**

**as easy as 1-2-3**

Tony Graham  
Mentea  
13 Kelly's Bay Beach  
Skerries, Co Dublin, Ireland  
info@mentea.net  
@MenteaXML  
<http://www.mentea.net>

Version 1.0 – 2 April 2014

© 2014 Mentea

All rights reserved. The author grants the U.S. National Library of Medicine permission to archive and post a copy of this paper on the Journal Article Tag Suite Conference proceedings website.





# Formatting JATS

as easy as 1-2-3

**JATS Preview stylesheets** 5

**Aside: GitHub** 11

**XSLT 1.0: Government body** 11

**XSLT 2.0: PLOS ONE** 15

**XSLT 3.0: xslt3testbed** 21

**References** 24

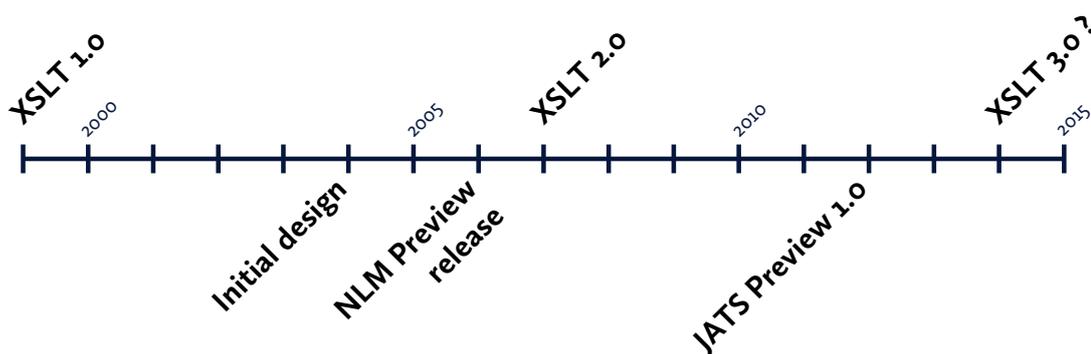
Appendix A – **About** 25





## Reconstructed timeline

4



Reconstructed from comments in code, downloads, and emails with Kim Tryka and Tommie Usdin.

## Why still XSLT 1.0 in 2012?

5

- XSLT 1.0 still dominant on some platforms
  - .NET
  - Linux/Unix
- Also tested with XSLT 2.0
- NLM stylesheets developed circa 2006/2007
  - One well-known XSLT 2.0 processor
  - Java only

## What does it do?

6

- Preprocessing
  - Convert OASIS tables to HTML tables
  - Massage citation format
  - Some require XSLT 2.0
- Formatting
  - XML to HTML
  - **XML to XSL-FO for formatting as PDF**
- Post-processing
  - HTML to XHTML for MathML

The only part that I've needed to use, and the only part being covered, is the transformation to XSL-FO and formatting to PDF.

## Customizability

7

“These stylesheets are **provided as a point of entry for JATS users** who may not have the resources to create them from scratch. Because there are many varied implementations of JATS, you should have no expectation that these stylesheets will create production ready files in any arbitrary system. Instead, the stylesheets should be customized for your particular needs.”

“Because **we view these stylesheets as a template for a customized solution, not the solution itself**, we will accept changes that fix an actual bug, but we will not merge in changes that we view as “customization”. For example, we will accept changes that fix a problem which otherwise leads to failure in creating a final output file, but we will not accept changes that focus on presentational aspects of the final output (such as font changes, margin changes, graphics sizing, etc).”

Statement about customisation from JATSPreviewStylesheets README with added emphasis.

## XSLT features supporting customizability

8

- Templates
- Modular stylesheets
- Named attribute sets

## Templates

9

- `match` matches a context in source XML
- Content of `xsl:template` instantiated when template is applied

```
<xsl:template match="td">
  <fo:table-cell xsl:use-attribute-sets="td">
    <xsl:call-template name="process-table-cell"/>
  </fo:table-cell>
</xsl:template>
```

Elements in the body of the template not in the XSLT namespace are copied to the result, and elements and attributes in the XSLT namespace are acted on by the XSLT processor.

## Modular stylesheets

10

```
<xsl:include
  href = uri-reference />
```

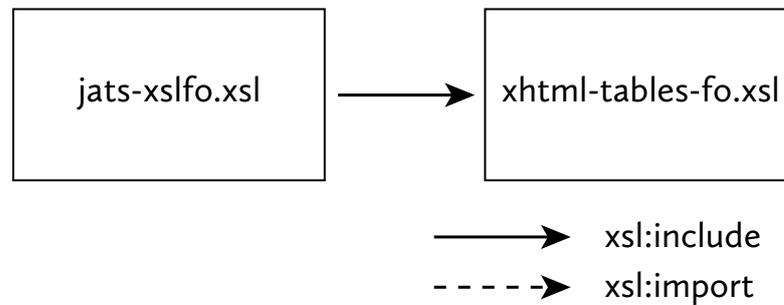
- href refers to other stylesheet
- Children of other `xsl:stylesheet` replace `xsl:include`

```
<xsl:import
  href = uri-reference />
```

- href refers to other stylesheet
- Imported definitions and template rules *not* part of importing stylesheet
- Have lower *import precedence*

## Imports in JATS XSL-FO preview

11



There are more interesting block diagrams later.

## Overriding templates

12

- Template in importing stylesheet overrides same context in imported
- Good when overriding complete function of template
- Extra overhead if you just want to change one little thing

## Attribute sets

13

- Named set of attribute definitions
- Use in multiple places
- Definitions evaluated in each context where used

```
<xsl:attribute-set name="fig">
  <xsl:attribute name="keep-together.within-page"
    >always</xsl:attribute>
  <xsl:attribute name="id">
    <xsl:value-of select="generate-id()" />
  </xsl:attribute>
</xsl:attribute-set>
```

Since attribute definitions in attribute sets are evaluated each time the attribute set is used, the value of the `id` attribute will be unique to each context.

## JATS Preview supporting customizability

14

- Global variables
- Attribute sets
- Named templates

## Example customization

15

- Add to attribute set from JATS stylesheets

```
<xsl:attribute-set name="td">
  <xsl:attribute name="line-stacking-strategy"
    >max-height</xsl:attribute>
</xsl:attribute-set>
```

- New attribute set reusing merged td attribute set

```
<xsl:attribute-set name="td-small"
  use-attribute-sets="td">
  <xsl:attribute name="line-height">10pt</xsl:attribute>
  <xsl:attribute name="border">none</xsl:attribute>
  <xsl:attribute name="padding-top">0pt</xsl:attribute>
  <xsl:attribute name="padding-bottom">0pt</xsl:attribute>
</xsl:attribute-set>
```

- Override JATS stylesheet in more-specific context

```
<xsl:template
  match="td[ancestor::table[@style = 'small']]">
  <fo:table-cell xsl:use-attribute-sets="td-small">
    <xsl:call-template name="process-table-cell"/>
  </fo:table-cell>
</xsl:template>
```

The `xsl:attribute-set` extends the 'td' defined in the JATS Preview stylesheet.

The new 'td-small' attribute set includes the attribute definitions from all declarations for the 'td' attribute set plus the definitions contained in its definition.

The template matches on a more-specific context than the general-purpose template for `td` in the JATS Preview stylesheets, so in those particular contexts, the XSLT processor uses this template, which adds a different set of attributes to the generated `fo:table-cell` but which still uses the 'process-table-cell' named template from the JATS Preview stylesheets as is used in the original template for `td`.

This illustrates in a nutshell how a customisation is able to extend, override, and reuse the constructs in the core JATS Preview stylesheets.

## Summary: JATS Preview

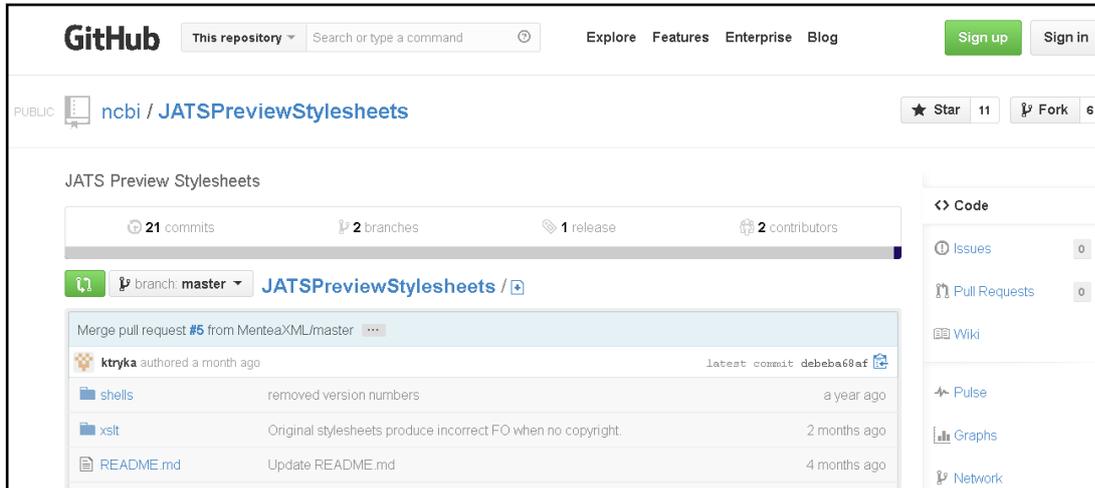
16

- XSLT 1.0
- Not accepting customisations into core
- Stylesheet structure facilitates customisations

## Aside: GitHub

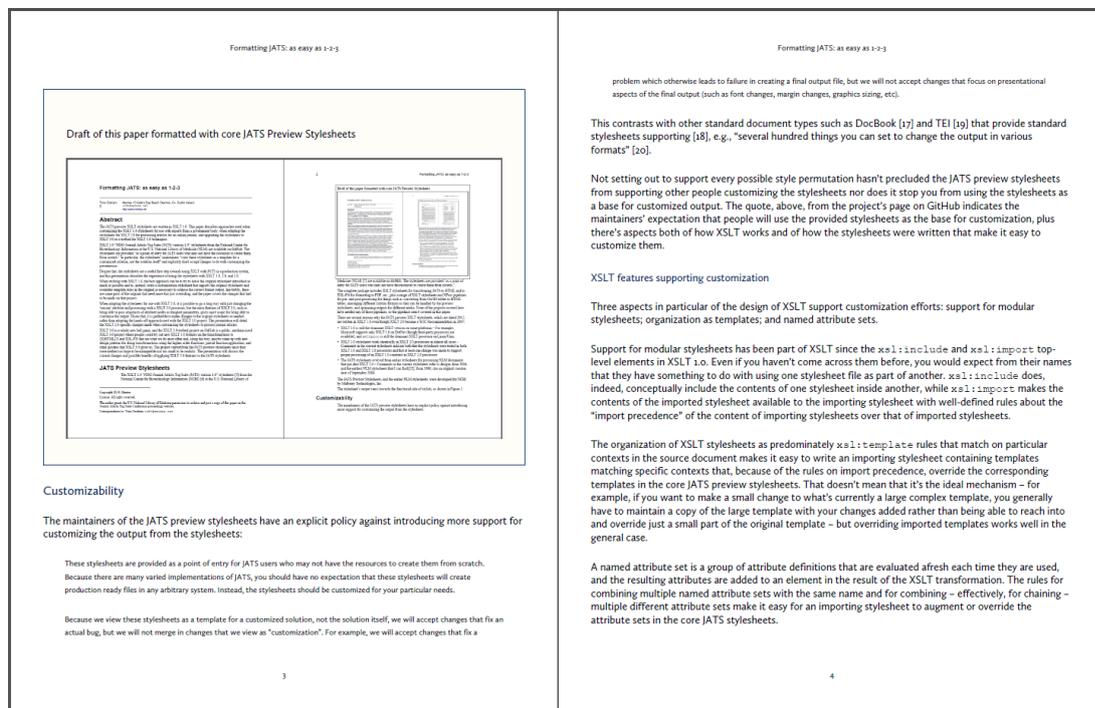
17

- “World’s largest open source community”
- Git distributed version control system
- Easy to “fork” – make your own version of projects
- Easy to “pull” merge requests from other projects



## XSLT 1.0: Government body

18



The paper for this talk formatted using XSLT 1.0 stylesheets

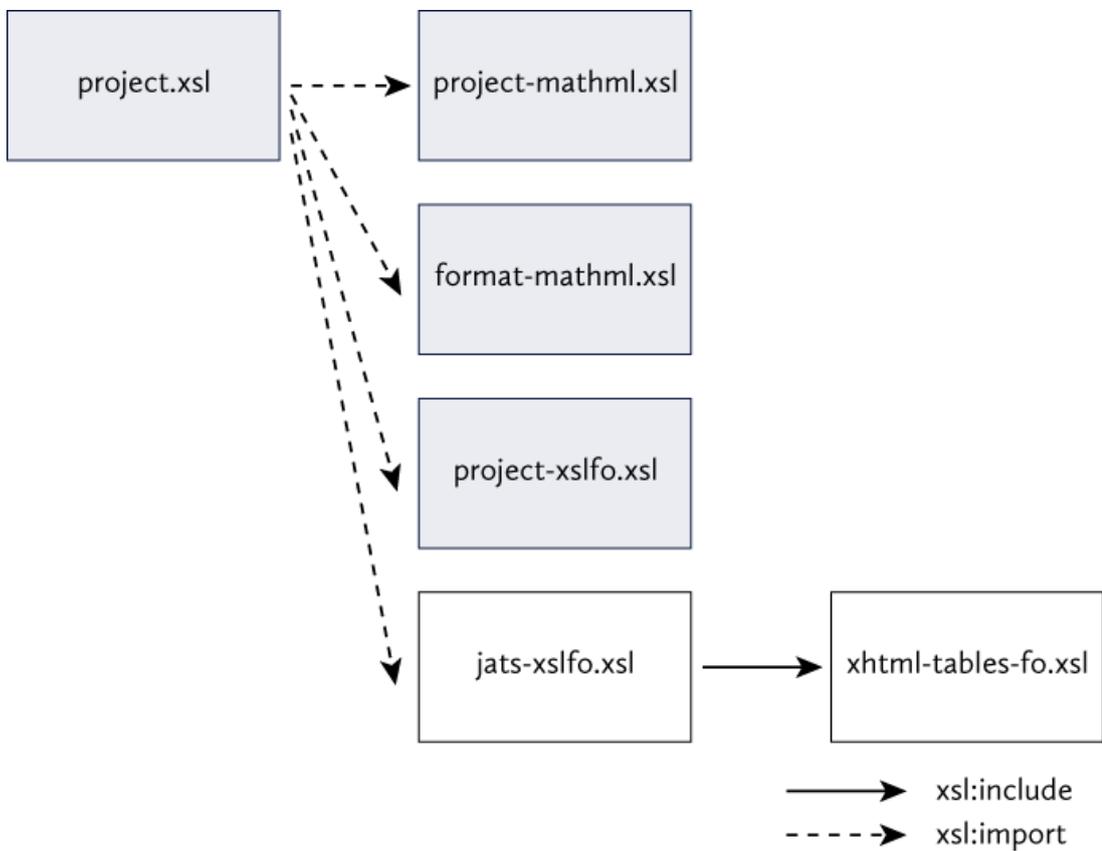
## Project details

19

- Source: variation on JATS Blue with custom metadata
- Result: similar page design to JATS preview stylesheets
- XSLT 1.0 because...
  - Client preference
  - Body and back content unchanged from JATS
  - Page design similar to JATS preview
- Customisation...
  - Changes in new modules
  - Import JATS Preview stylesheets

## Import structure

20



**MathML fix-up modules**

21

- Separate modules that can be dropped when problems solved
- `project-mathml.xsl` – add parentheses around display equation number
- `format-mathml.xsl` – workaround too-high accented characters

$$\frac{SE(\hat{p})/\hat{p}}{-\ln(\hat{p})} > .175 \text{ when } \hat{p} \leq .5$$

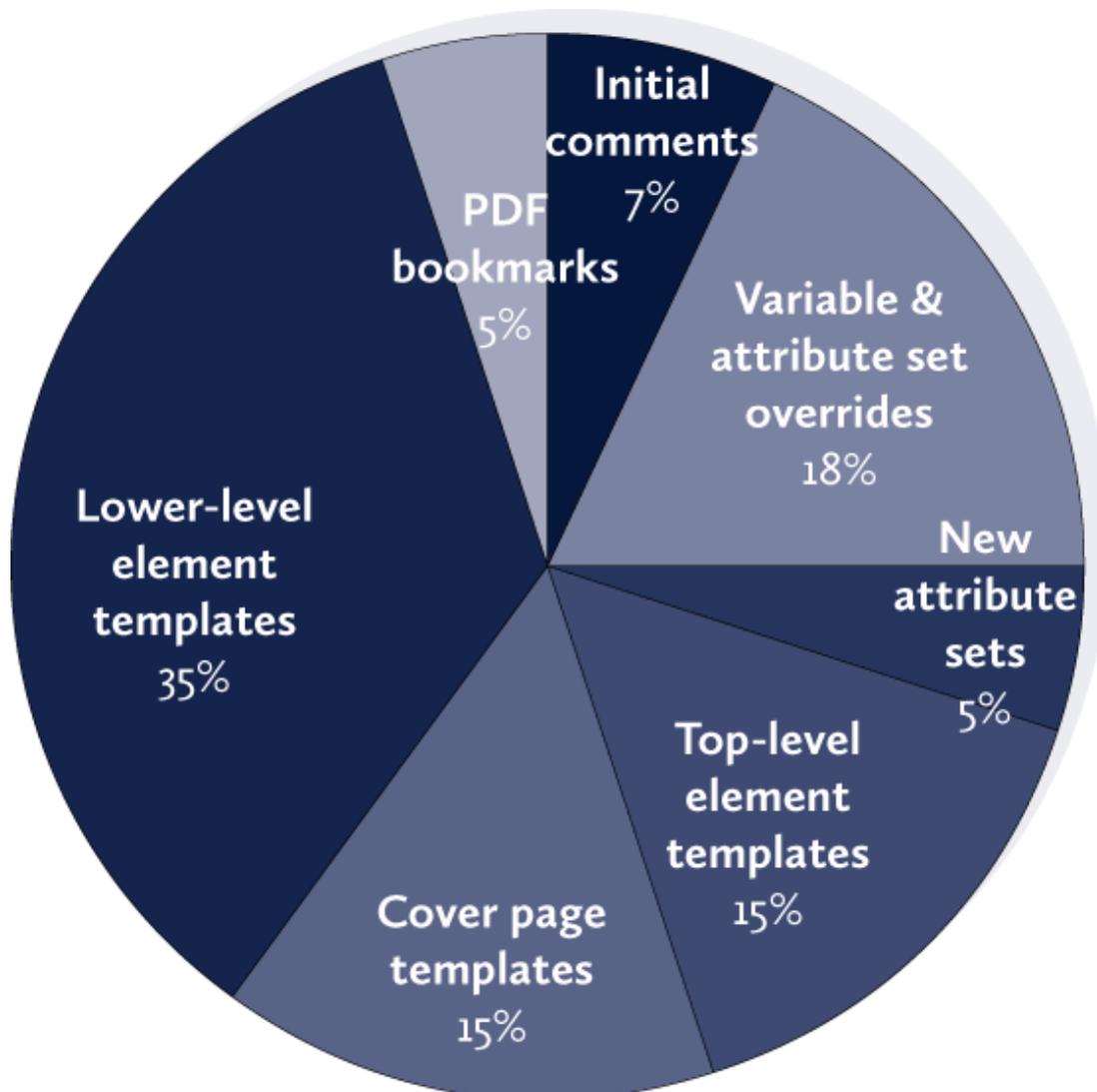
becomes

$$\frac{SE(\hat{p})/\hat{p}}{-\ln(\hat{p})} > .175 \text{ when } \hat{p} \leq .5$$

(Latest formatter has rewritten MathML support)

**What's in project-xslfo.xsl?**

22

**Summary: XSLT 1.0**

23

- Customisation on top of JATS Preview stylesheets
- Preview stylesheets provided sufficient hooks

The image displays a grid of 12 sample PLOS ONE pages, arranged in two rows of six. Each page represents a different type of content commonly found in scientific journals:

- Page 1 (Top Left):** Title page for a paper titled "A Single Kernel-Based Approach to Extract Drug-Drug Interactions from Biomedical Literature". It includes the journal logo, authors, and abstract.
- Page 2 (Top Middle):** A figure showing three line graphs (A, B, C) representing different experimental conditions. The y-axis is labeled "Normalized Frequency" and the x-axis is "Time".
- Page 3 (Top Right):** A table with multiple columns and rows, likely containing statistical data or experimental results.
- Page 4 (Middle Left):** A table with multiple columns and rows, similar to the one on page 3.
- Page 5 (Middle Middle):** A diagram showing a flowchart or dependency graph with nodes and arrows.
- Page 6 (Middle Right):** A table with multiple columns and rows, similar to the ones on pages 3 and 4.
- Page 7 (Bottom Left):** A table with multiple columns and rows, similar to the ones on pages 3 and 4.
- Page 8 (Bottom Middle):** A diagram showing a flowchart or dependency graph, similar to the one on page 5.
- Page 9 (Bottom Right):** A table with multiple columns and rows, similar to the ones on pages 3 and 4.

Sample PLOS ONE pages.

## Project details

- Peer-reviewed, open-access, online publication
- Public Library of Science
- JATS/NLM markup
- Lights-out batch formatting with XSL-FO
- Previously produced use 3B2 and (presumably) manual fix-up
- XSLT 2.0 because...
  - Big differences in metadata, figure, table handling
  - Needed vendor extensions
- Customisation...
  - Modified version of `jats-xslfo.xsl`
  - Additional XSLT modules

## PONE “features”

26

- Figures and tables float to top (or bottom) of page
- Figures column-wide or page-wide
  - No size information in XML
- Figure graphic+caption can’t overflow page
- Tables column-wide, page-wide, or page-high
  - Page-high may be single column
  - May be multiple pages
  - No width indication in XML
  - No row spanning (thank goodness!)
- No figures or tables allowed after start of back matter

## XSLT/XSL-FO “features”

27

- Page-wide floats
  - Vendor extension for column-wide
- Floats don’t break
- Floats only at top of page
  - Bottom-float extension available but unused
- Graphic size not available to XSLT
- *Fire-and-forget* processing

## Table handling

28

- “Pre-format” tables in three widths on *long* pages
  - Column-wide, page-wide, (width of) page-high
  - Prefix table IDs with string indicating width
- Format to area tree XML
- Compare area trees for each table
- Use width with least area and no overflow
- Recreate as multiple `fo:float` if overflows page
  - Re-use table column widths from area tree to remain consistent

# Picking "Best" Tables

Table 1. Statistics of the DDI corpora.

Statistics	Training set	Dev set	Test set
Documents	241	126	126
Dependencies	1037	519	519
Relations	200	100	100
Types of dependencies	100	50	50

Table 2. Effectiveness of parameters.

$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	AUC	ACC
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51

Table 3. Performance of our approach in comparison with other approaches.

Approach	TP	FP	FN	FN	P	R	F	AUC	MCC	AUC
MB [3]	342	354	212	2917	0.05	0.19	0.07	0.19	0.15	-
Our approach	506	237	246	1973	0.11	0.22	0.15	0.22	0.14	0.24
LMB-FM [32]	532	375	223	1895	0.12	0.23	0.15	0.23	0.15	-
FRU-RLT [3]	529	377	226	1894	0.11	0.22	0.15	0.22	0.14	-
Struc2Vec [33]	502	376	225	1895	0.12	0.23	0.15	0.23	0.15	-
BNKLEL [34]	420	296	335	1625	0.12	0.23	0.15	0.23	0.15	-
Naive Bayes approach	503	376	225	1894	0.11	0.22	0.15	0.22	0.14	0.24
Our approach	502	281	274	1924	0.12	0.23	0.15	0.23	0.15	0.24

Table 1. Statistics of the DDI corpora.

Statistics	Training set	Dev set	Test set
Documents	241	126	126
Dependencies	1037	519	519
Relations	200	100	100
Types of dependencies	100	50	50

Table 2. Effectiveness of parameters.

$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	AUC	ACC
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51

Table 3. Performance of our approach in comparison with other approaches.

Approach	TP	FP	FN	FN	P	R	F	AUC	MCC	AUC
MB [3]	342	354	212	2917	0.05	0.19	0.07	0.19	0.15	-
Our approach	506	237	246	1973	0.11	0.22	0.15	0.22	0.14	0.24
LMB-FM [32]	532	375	223	1895	0.12	0.23	0.15	0.23	0.15	-
FRU-RLT [3]	529	377	226	1894	0.11	0.22	0.15	0.22	0.14	-
Struc2Vec [33]	502	376	225	1895	0.12	0.23	0.15	0.23	0.15	-
BNKLEL [34]	420	296	335	1625	0.12	0.23	0.15	0.23	0.15	-
Naive Bayes approach	503	376	225	1894	0.11	0.22	0.15	0.22	0.14	0.24
Our approach	502	281	274	1924	0.12	0.23	0.15	0.23	0.15	0.24

Table 1. Statistics of the DDI corpora.

Statistics	Training set	Dev set	Test set
Documents	241	126	126
Dependencies	1037	519	519
Relations	200	100	100
Types of dependencies	100	50	50

Table 2. Effectiveness of parameters.

$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	AUC	ACC
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51
0.01	0.01	0.01	0.01	0.01	0.71	0.51

Table 3. Performance of our approach in comparison with other approaches.

Approach	TP	FP	FN	FN	P	R	F	AUC	MCC	AUC
MB [3]	342	354	212	2917	0.05	0.19	0.07	0.19	0.15	-
Our approach	506	237	246	1973	0.11	0.22	0.15	0.22	0.14	0.24
LMB-FM [32]	532	375	223	1895	0.12	0.23	0.15	0.23	0.15	-
FRU-RLT [3]	529	377	226	1894	0.11	0.22	0.15	0.22	0.14	-
Struc2Vec [33]	502	376	225	1895	0.12	0.23	0.15	0.23	0.15	-
BNKLEL [34]	420	296	335	1625	0.12	0.23	0.15	0.23	0.15	-
Naive Bayes approach	503	376	225	1894	0.11	0.22	0.15	0.22	0.14	0.24
Our approach	502	281	274	1924	0.12	0.23	0.15	0.23	0.15	0.24

Page-wide

Column-wide

Page-high

Three tables formatted in each of three widths, with preferred versions highlighted.

# Sized and placed tables

DDI Extraction from Biomedical Literature

1. Negation  
171. **Atorvastatin** should not be administered concomitantly with ampicillin, ciprofloxacin, gentamicin, netilmicin, or tobramycin.

2. Need more context  
172. **Alcohol** increases bioavailability by 50%, cimetidine, and valproates.

3. Posing errors  
173. **Acetazolamide** may prevent the urinary antiseptic effect of methenamine.

4. Association errors  
174. **Amiodarone** may suppress certain CYP450 enzymes, including CYP2A6, CYP2C8, CYP2C9, and CYP3A4.

(a) False positives

1. Coarctator structures and appositions  
FN1. There is little if any clinically significant interaction between **Acetaminophen** and **metformin**.

FN2. There were transient increases in liver ALT and AST when **CANCIDAS** and **oxytocin** were co-administered.

2. Drugs in different subphases class  
FN3. Based on total **ergosterol** concentrations, **griseofulvin** increased the AUC by 25% and reduced the plasma and renal clearance by 20% and 30%, respectively.

3. Need more context  
FN4. Patients taking **warfarin** or other **coumarin-derivative** anticoagulants should be monitored regularly for changes in prothrombin time or INR.

FN5. Because there are no data on the compatibility of **Novolog** and **crystalline zinc insulin preparations**, **Novolog** should not be mixed with these preparations.

(b) False negatives

doi:10.1371/journal.pone.0049011.g005

Table 2. Effectiveness of parameters.

$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	AUC	ACC
0.02	0.05	0.46	0.97	21	0.83	0.18
0.04	0.04	0.10	0.10	0.01	0.17	-
0.06	0.03	0.08	0.10	0.01	0.17	-
0.08	0.01	0.12	0.17	0.01	0.15	-
0.10	0.07	0.06	0.02	0.18	0.18	0.14
0.12	0.12	0.03	0.02	0.07	0.19	-
0.14	0.14	0.02	0.01	0.02	0.12	-

FN1. **Acetaminophen**, **R**, **FN2**, **FN3**, **FN4**, **FN5**, **FN6**, **FN7**, **FN8**, **FN9**, **FN10**, **FN11**, **FN12**, **FN13**, **FN14**, **FN15**, **FN16**, **FN17**, **FN18**, **FN19**, **FN20**, **FN21**, **FN22**, **FN23**, **FN24**, **FN25**, **FN26**, **FN27**, **FN28**, **FN29**, **FN30**, **FN31**, **FN32**, **FN33**, **FN34**, **FN35**, **FN36**, **FN37**, **FN38**, **FN39**, **FN40**, **FN41**, **FN42**, **FN43**, **FN44**, **FN45**, **FN46**, **FN47**, **FN48**, **FN49**, **FN50**, **FN51**, **FN52**, **FN53**, **FN54**, **FN55**, **FN56**, **FN57**, **FN58**, **FN59**, **FN60**, **FN61**, **FN62**, **FN63**, **FN64**, **FN65**, **FN66**, **FN67**, **FN68**, **FN69**, **FN70**, **FN71**, **FN72**, **FN73**, **FN74**, **FN75**, **FN76**, **FN77**, **FN78**, **FN79**, **FN80**, **FN81**, **FN82**, **FN83**, **FN84**, **FN85**, **FN86**, **FN87**, **FN88**, **FN89**, **FN90**, **FN91**, **FN92**, **FN93**, **FN94**, **FN95**, **FN96**, **FN97**, **FN98**, **FN99**, **FN100**.

FN1. **Acetaminophen**, **R**, **FN2**, **FN3**, **FN4**, **FN5**, **FN6**, **FN7**, **FN8**, **FN9**, **FN10**, **FN11**, **FN12**, **FN13**, **FN14**, **FN15**, **FN16**, **FN17**, **FN18**, **FN19**, **FN20**, **FN21**, **FN22**, **FN23**, **FN24**, **FN25**, **FN26**, **FN27**, **FN28**, **FN29**, **FN30**, **FN31**, **FN32**, **FN33**, **FN34**, **FN35**, **FN36**, **FN37**, **FN38**, **FN39**, **FN40**, **FN41**, **FN42**, **FN43**, **FN44**, **FN45**, **FN46**, **FN47**, **FN48**, **FN49**, **FN50**, **FN51**, **FN52**, **FN53**, **FN54**, **FN55**, **FN56**, **FN57**, **FN58**, **FN59**, **FN60**, **FN61**, **FN62**, **FN63**, **FN64**, **FN65**, **FN66**, **FN67**, **FN68**, **FN69**, **FN70**, **FN71**, **FN72**, **FN73**, **FN74**, **FN75**, **FN76**, **FN77**, **FN78**, **FN79**, **FN80**, **FN81**, **FN82**, **FN83**, **FN84**, **FN85**, **FN86**, **FN87**, **FN88**, **FN89**, **FN90**, **FN91**, **FN92**, **FN93**, **FN94**, **FN95**, **FN96**, **FN97**, **FN98**, **FN99**, **FN100**.

DDI Extraction from Biomedical Literature

Table 3. Performance of our approach in comparison with other approaches.

Approach	TP	FP	FN	FN	P	R	F	AUC	MCC	AUC
MB [3]	342	354	212	2917	0.05	0.19	0.07	0.19	0.15	-
Our approach	506	237	246	1973	0.11	0.22	0.15	0.22	0.14	0.24
LMB-FM [32]	532	375	223	1895	0.12	0.23	0.15	0.23	0.15	-
FRU-RLT [3]	529	377	226	1894	0.11	0.22	0.15	0.22	0.14	-
Struc2Vec [33]	502	376	225	1895	0.12	0.23	0.15	0.23	0.15	-
BNKLEL [34]	420	296	335	1625	0.12	0.23	0.15	0.23	0.15	-
Naive Bayes approach	503	376	225	1894	0.11	0.22	0.15	0.22	0.14	0.24
Our approach	502	281	274	1924	0.12	0.23	0.15	0.23	0.15	0.24

F1 score: 0.15, precision: 0.11, accuracy: 0.15, MCC: 0.14, Matthews correlation coefficient: 0.14

doi:10.1371/journal.pone.0049011.t003

Table 3. Performance of our approach in comparison with other approaches.

Approach

TP

FP

FN

FN

P

R

F

AUC

MCC

AUC

MB [3]

342

354

212

2917

0.05

0.19

0.07

0.19

0.15

-

Our approach

506

237

246

1973

0.11

0.22

0.15

0.22

0.14

0.24

LMB-FM [32]

532

375

223

1895

0.12

0.23

0.15

0.23

0.15

-

FRU-RLT [3]

529

377

226

1894

0.11

0.22

0.15

0.22

0.14

-

Struc2Vec [33]

502

376

225

1895

0.12

0.23

0.15

0.23

0.15

-

BNKLEL [34]

420

296

335

1625

0.12

0.23

0.15

0.23

0.15

-

Naive Bayes approach

503

376

225

1894

0.11

0.22

0.15

0.22

0.14

0.24

Our approach

502

281

274

1924

0.12

0.23

0.15

0.23

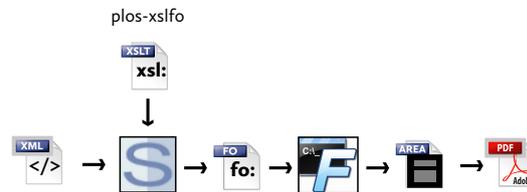
0.15

0.24

Column-wide and page-wide tables placed on pages.

## Usual processing model

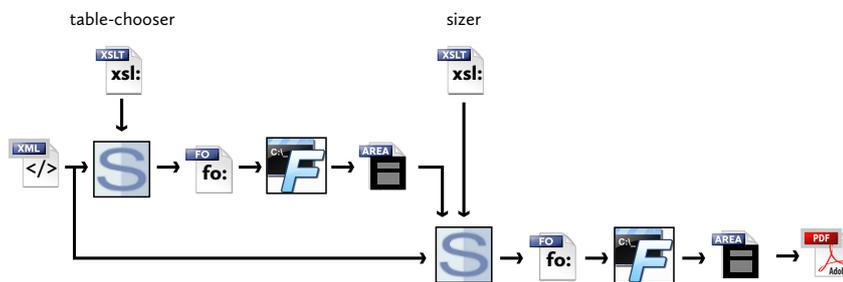
31



The conventional XSLT-XSL-FO processing model.

## Table-handling processing model

32



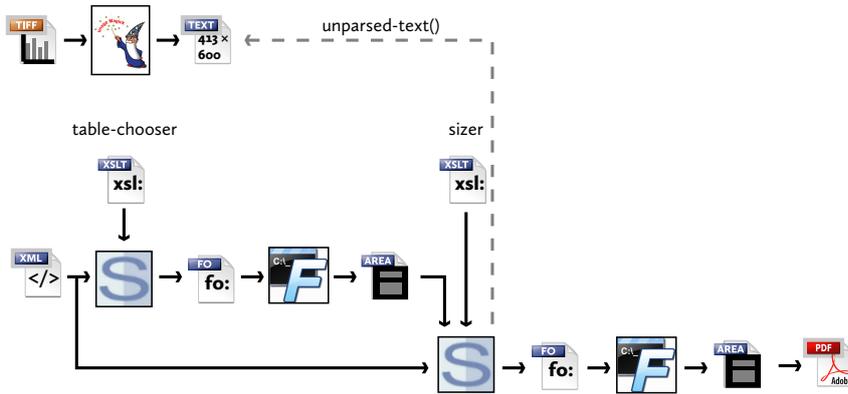
The processing model including preprocessing tables to generate an area tree from which to determine the preferred width for each table.

## Graphics handling

33

- Get TIFF graphics
- ImageMagick `identify` gives graphic size and resolution
- “Pre-format” caption at both widths to get exact size
- Choose best width
- (Possibly) scale down graphic so caption also fits on page

# Figure-handling processing model



Processing model when graphics handling added.

# Floats after back matter

**Figure 4.** PAb-induced apoptosis in *trypsin* cell lines. *Flow cytometry* analysis revealed the proportion of sub-G1 phase cells (apoptotic cells) in the 72h (ND), 96h (D), and 120h (PAB). The experiments were repeated at least three times.

**Author Contributions**  
 Read and approved the manuscript: SM, HC, JY, VS, DC. Conceived and designed the experiments: SM, HC, VS, DC. Performed the experiments: SM, HC, VS, DC. Analyzed the data: SM, VS, DC. Wrote the paper: SM.

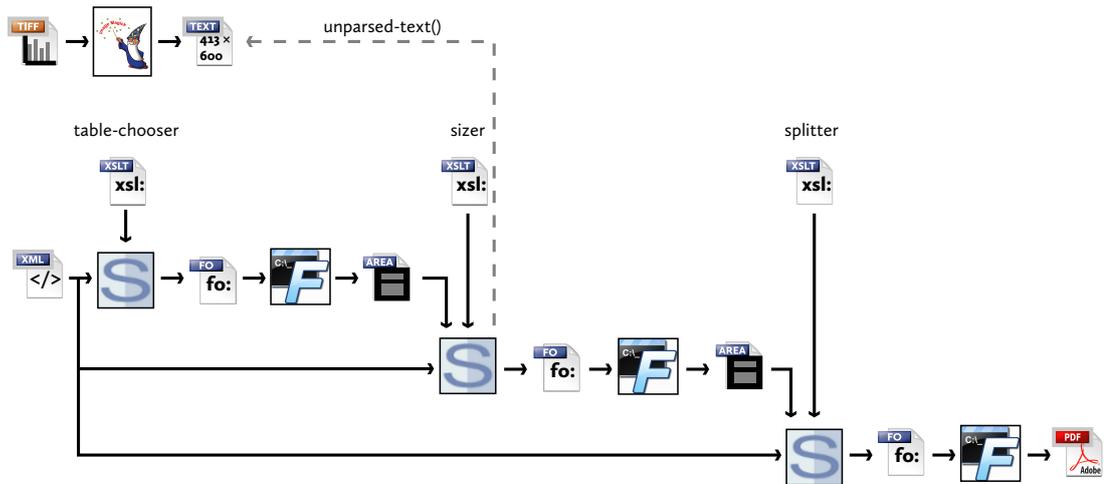
**Figure 5.** Inhibitory effect of PABs on tumor growth in xenograft SCIP<sup>+</sup> mouse models. (A) A significant difference in tumor volume ( $P < 0.05$ ) was observed between PAB-treated mice and control groups. (B) Representative photos of tumor-bearing mice. (C) A significant increase in survival was observed in PAB-treated mice compared with other treatment groups ( $P < 0.05$ ).

Figures and tables are required to not appear after the start of the back matter.

## Splitting at back matter 36

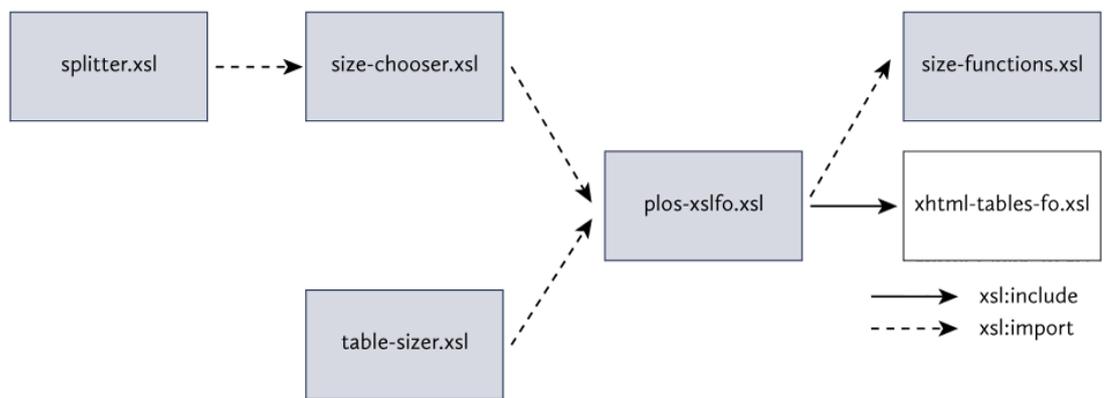
- Format “final” FO with right-width tables and figures to area tree
- Compare positions of first “back” content and last float
  - back plus bits from front, body
- Generate new FO with either one or two fo:page-sequence
- If second fo:sequence, it contains only back matter so floats in first appear before back matter

## Putting It All Together 37



The full processing model.

## Import structure 38



All the top-level stylesheets use plos-xslfo.xml for basic formatting.  
 splitter.xml does everything size-chooser.xml does, and more, so it imports that file rather than importing plos-xslfo.xml directly.

## Summary: XSLT 2.0

39

- It shouldn't be this hard
- Column-wide floats require vendor extension
- Navigating area tree isn't easy
- No standard for area tree XML made it harder and even less portable
- Creating new FO and reprocessing easier than rewriting area tree
- EXPath Binary Module (and a TIFF-handling library!) could avoid using ImageMagick
  - Or use vendor extension

## XSLT 3.0: xslt3testbed

40

<https://github.com/MenteaXML/xslt3testbed>

- Trying out new XSLT 3.0 features
- Converting existing JATS stylesheets to XSLT 3.0

## Why?

41

“...the design process does not include enough feedback; by the time people start reporting their usability experiences, the decisions are difficult to change.”

- Early start on patterns and idioms to help adoption
- Find infelicities in spec (and implementations)
- The time is right
  - Project started November 2013
  - XSLT 3.0 Last Call WD – 12 December 2013

Quote from Micheal Kay, editor of XSLT 3.0 spec: <http://www.biglist.com/lists/lists.mulberrytech.com/xsl-list/archives/201403/msg00332.html>

Motivation comes from looking for a better way to get people using the new version:

- 1997: Wanted to discuss DSSSL so started DSSSList
- 1998: XSL-List started – people tried every new XSL feature as it came out
- 2004–2007++: People had working XSLT 1.0 systems and there weren't many XSLT 2.0 processors, so adoption slow
- 2013–2014: Looking for a quicker win than mailing lists, and people now used to working with GitHub projects

## W3C Process

42

- End game for a W3C spec:
  - Last Call
  - Candidate Recommendation
  - Proposed Recommendation
  - Recommendation
- Changes after “Last Call” require more documentation and substantiation

## Why JATS?

43

- Simpler than, e.g., DocBook or TEI
- Not a toy
- Potentially useful to authors and archives
- Existing XSLT stylesheets available

## Why JATSPreviewStylesheets?

44

<https://github.com/NCBITools/JATSPreviewStylesheets>

- XSLT 1.0
  - Easy for new contributors to add XSLT 2.0-isms
- Public domain
  - No copyright issues
  - XSLT 3.0 stylesheets also public domain
- Explicitly not supporting gazillion customisation parameters, PIs, etc.
  - Simpler processing
  - Fewer user expectations

## xslt3testbed goals

45

- Trial different techniques
- Open for dipping into to try random ideas
- Develop patterns and idioms
- Develop XSLT 3.0 package for XHTML tables
  - `xsl:package` new in XSLT 3.0
  - XHTML tables used in many document types

## xslt3testbed non-goals

46

- Single best way of doing anything
  - Multiple ways to solve the same problem are okay
- Definitive XSLT 3.0 testbed
  - It's easy to fork and make your own version
- Complete stylesheet for all of JATS
  - Existing stylesheets don't cover everything yet either

## Results so far

47

- Trying out maps, anonymous functions, and `xsl:iterate`
- Small advances in multiple areas
- Both XSL-FO and XHTML stylesheets
- More details in XML Prague 2014 talk  
<http://www.mentea.net/resources/xslt30testbed-slides.pdf>

## 6 W3C Bugzilla bu<sup>H</sup>Htickets so far

48

ID	Product	Comp	Assignee	Status	Resolution	Summary	Changed
24207	XPath /	XSLT 3.0	mike	NEW	---	XPath-level element and attribute constructors for use in anonymous functions	Mon 11:14
24199	XPath /	XPath 3.	jonathan.robie	ASSI	---	[XP30] No 'FunctionBody' production in body of spec?	2014-01-31
24200	XPath /	XPath 3.	jonathan.robie	RESO	WONT	"as" SequenceType vs 'TypeDeclaration' in XPath/XQuery 3.0?	2014-01-07
23118	XPath /	Function	mike	RESO	FIXE	'V' in fn:id	2013-09-01
23944	XPath /	XSLT 3.0	mike	RESO	FIXE	xsl:package/xsl:expose position	2014-01-28
23932	XPath /	XSLT 3.0	mike	CLOS	FIXE	Attribute sets provide attribute instructions	2013-12-10

6 bugs found.

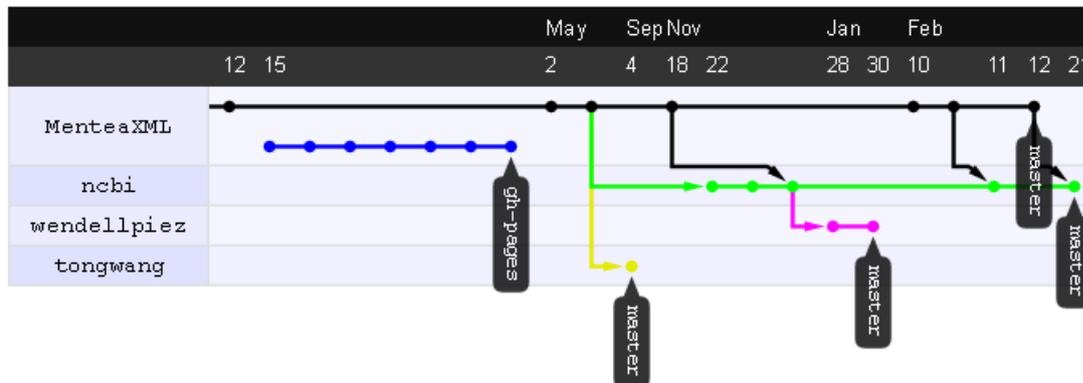
## 5 JATSPreviewStylesheets patches so far

49

### The JATSPreviewStylesheets network graph

All branches in the network using MenteaXML/JATSPreviewStylesheets as the reference point.

Show Help



## Other results

50

- One XSLT processor bug
- One change to Wendell Piez's JATS Oxygen plug-in
- Technique for hosting Oxygen plugins on GitHub

Pre-release

v0.0.2  
e7d142a

### Release test 2

tkg released this an hour ago · 1 commit to master since this release

Source code (zip)

Source code (tar.gz)

## Summary: XSLT 3.0

51

<https://github.com/MenteaXML/xslt3testbed>

- The time is right
- Useful in multiple arenas
- Results summarised on project wiki and <http://inasmuch.as/>
- Well suited for trying things out
- Go fork and multiply

## Conclusion

52

- JATS Preview stylesheets:
  - Explicitly don't support customisation
  - Good basis for your own customization
- Customise by:
  - Layer on top of existing styleheets
  - Modify your copy of the stylesheets
- Usable with XSLT 1.0, 2.0, or 3.0

## References

53

- slide 41 – Micheal Kay  
<http://www.biglist.com/lists/lists.mulberrytech.com/xsl-list/archives/201403/msg00332.html>
- slide 42 – W3C Process Document  
<http://www.w3.org/2005/10/Process-20051014/tr.html>
- slide 48 – Bugs so far  
[https://www.w3.org/Bugs/Public/buglist.cgi?email1=tgraham%40mentea.net&emailreporter1=1&emailtype1=substring&product=XPath%20%2F%20XQuery%20%2F%20XSLT&query\\_format=advanced](https://www.w3.org/Bugs/Public/buglist.cgi?email1=tgraham%40mentea.net&emailreporter1=1&emailtype1=substring&product=XPath%20%2F%20XQuery%20%2F%20XSLT&query_format=advanced)

## Appendix A

### About

Tony Graham 25

Mentea 25

#### Tony Graham

Tony Graham has been working with markup since 1991, with XML since 1996, and with XSLT/XSL-FO since 1998. He is Chair of the Print and Page Layout Community Group at the W3C and previously an invited expert on the W3C XML Print and Page Layout Working Group (XPPL) defining the XSL-FO specification, as well as an acknowledged expert in XSLT, developer of the open source xmlroff XSL formatter, a committer to both the XSpec and Juxy XSLT testing frameworks, the author of “Unicode: A Primer”, a member of the XML Guild, and a qualified trainer.

Tony’s career in XML and SGML spans Japan, USA, UK, and Ireland, working with data in English, Chinese, Japanese, and Korean, and with academic, automotive, publishing, software, and telecommunications applications. He has also spoken about XML, XSLT, XSL-FO, EPUB, and related technologies to clients and conferences in North America, Europe, and Australia.

#### Mentea

Mentea specialises in consulting and training in XML, XSL-FO, & XSLT. We are available for on-site meetings and classes, worldwide, but as well as on-site meetings and classes, we routinely keep in touch with clients through email, Skype, instant messaging, and telephone and through a secure, per-client or per-project wiki, revision-control, and issue-tracking system.

Our staff have been working with markup since 1991, with XML since 1996, and with XSLT/XSL-FO since 1998. Based in Dublin, Ireland, Mentea has a global reach: in recent projects, we have helped companies and organisations in the USA, Ireland, England, and France with their XSLT, XSL, and XML, including:

- Writing Schematron for a professional body
- Augmenting a XSLT-based automated schema documentation system that produces both HTML and PDF
- Extending FOP for a software company
- Training in XML, oXygen, DocBook, XSLT 2.0, and XSL-FO
- Formatting JATS to PDF for a scientific journal
- Writing XSLT stylesheets to convert non-XML into XML then into EPUB
- Writing XSLT to convert Excel into XML for a commercial bank

Mentea presents a unique range of skills extending beyond XML and XSL-FO/XSLT into Unicode, SGML, DSSSL, and programming in C, Java, Perl, Lisp, and other languages.

We understand how markup works. Our staff has worked with markup in Japan, USA, UK, and Ireland as user, consultant, and developer, with data in English, French, Chinese, Japanese, and Korean, with academic, automotive, publishing, software, and telecommunications applications, and in the Web Services and document processing arenas.

We are also interested in applying the tools for ensuring software quality – unit testing, code coverage, profiling, and other tools – to XML and XSLT/XSL-FO processing.

Through our associations and affiliations with other consultants around the world, we can call on extra help for large or specialised projects.



**MENTEA**