

LaTeX to MathML – archiving mathematics

Kaveh Bazargan

15 July 2004

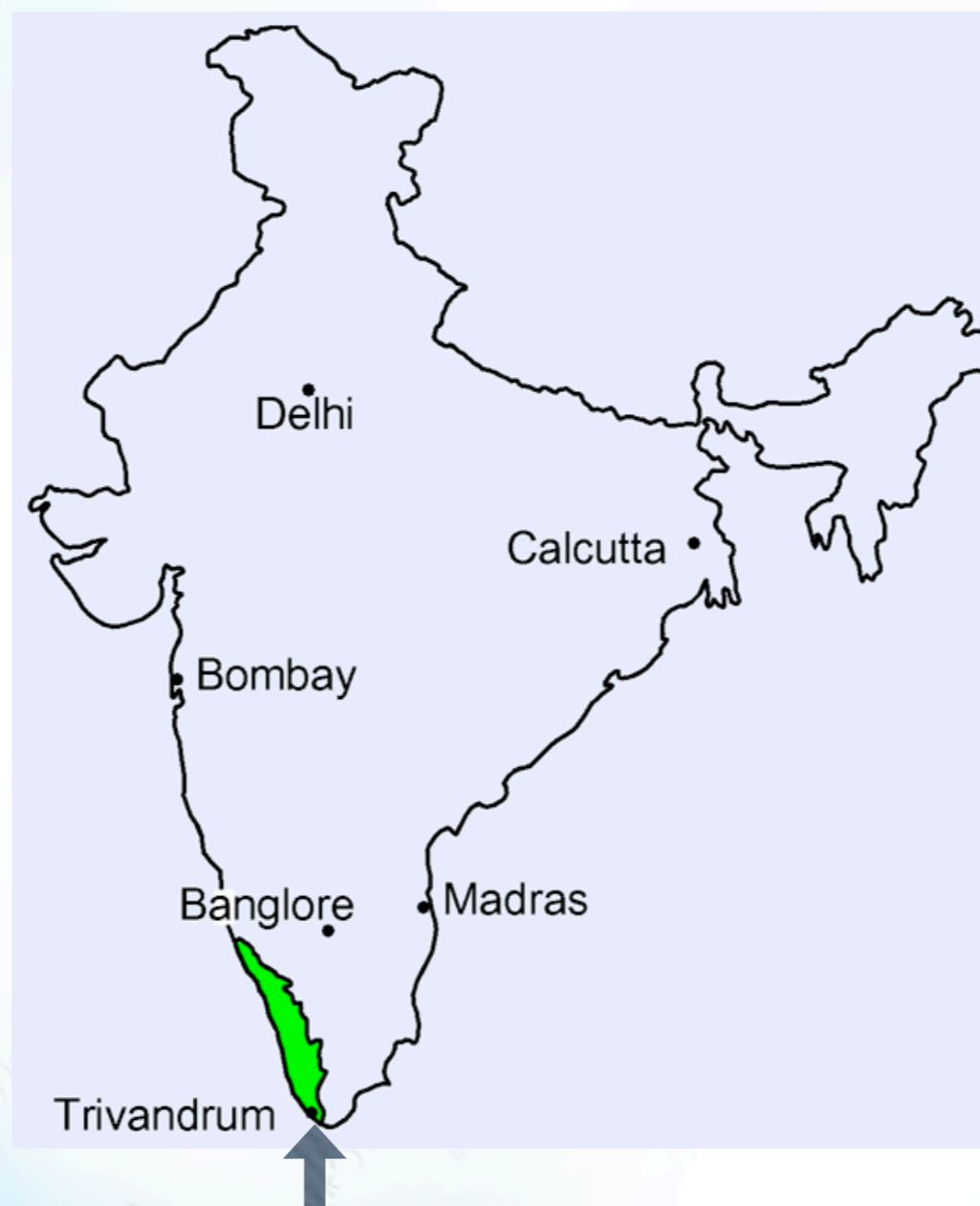
Who we are

- ◆ Typesetters (or content managers)
 - ◆ One of many!
- ◆ We do journals only
- ◆ Get manuscripts after acceptance
- ◆ Deliver PDF and XML

What we do

- ◆ Copy edit
- ◆ Graphics
- ◆ House style
- ◆ Cross references
- ◆ All deliverables must match

Where we are



$q_t(\theta^*(\pi), \pi) - b_T \sum_{i=1}^T \nabla q_t$
around π . Now,
 $\sup_{\pi' \in B(\pi, \epsilon)}$
 $\sum_{i=1}^T \epsilon_i (\pi - \pi')$



RIVER VALLEY
TECHNOLOGIES

Our staff



Our clients

- ◆ Academic publishers:
 - ◆ Institute of Physics
 - ◆ Elsevier
 - ◆ Blackwells
 - ◆ ...

Open source software

- ◆ Can customize source code
- ◆ Minimized dependency on third parties
- ◆ Better support than commercial software
- ◆ Usually zero cost

Partial list of Open (free) software

TeX

PDFTeX

GhostScript

DVIPS

ImageMagick

SX

Linux

Apache

FTE

Emacs

VI

Perl

Tomcat

Java

GCC

Postgres

nsgmls

xmllint

GNU tools: make, grep, etc

Mozilla

Pine

PHP

Saxon

XSLTProc

OpenOffice

xPDF

Partial list of proprietary software

Mac OS X

Adobe Acrobat

Adobe Distiller

Adobe PhotoShop

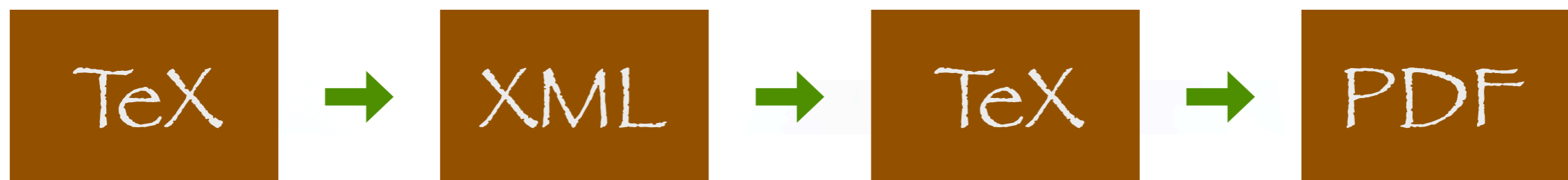
Adobe Illustrator

PitStop

What should be the workflow?

- ◆ Requirements:
 - ◆ Use author TeX files as input
 - ◆ XML/SGML must be definitive files
 - ◆ High degree of automation
 - ◆ Produce highest quality typesetting

Resultant automated workflow



Problem of TeX to XML

$$y = x^2$$

$$y = x^2$$

```
\def\power{x^2}
y = \power
```

```
\def\power#1{x^#1}
y = \power 2
```

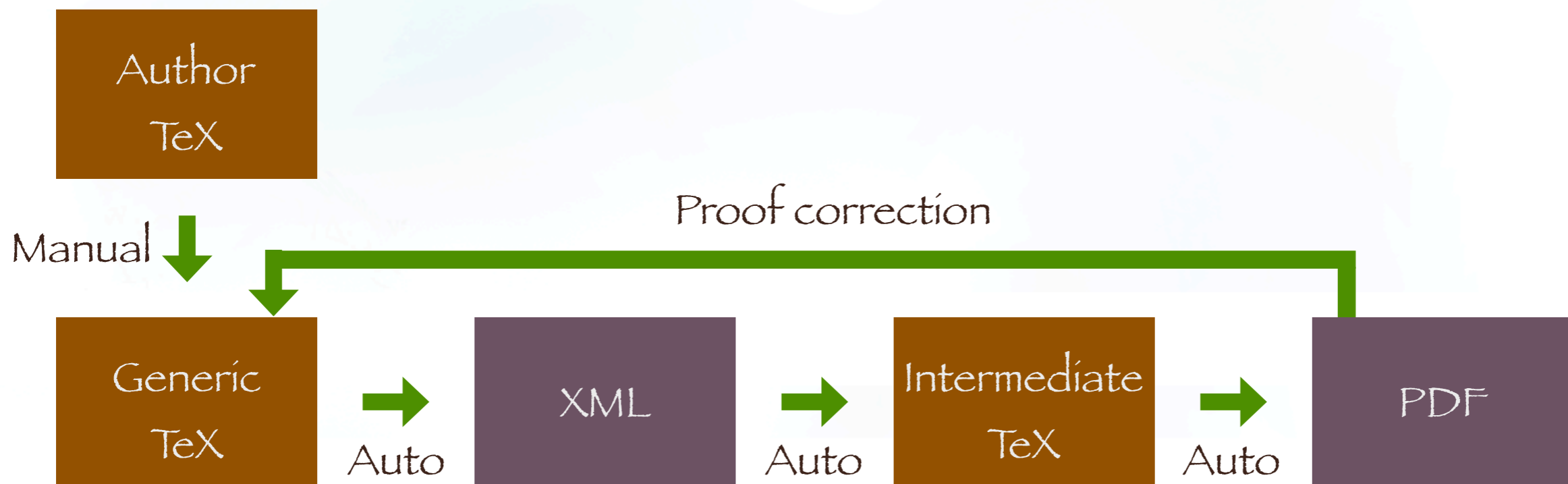
```
\def\power#1#2{#1^#2}
y = \power x 2
```

```
<mml:math>
<mml:mi>y</mml:mi>
<mml:mo>=</mml:mo>
<mml:msup>
<mml:mrow>
<mml:mi>x</mml:mi>
</mml:mrow>
<mml:mrow>
<mml:mn>2</mml:mn>
</mml:mrow>
</mml:msup>
</mml:math>
```

Use TeX as translator

- ◆ By definition, will give the same output
- ◆ But only gives DVI file for print
- ◆ ... use hidden messages (or `\specials`)
- ◆ TeX 'thinks' it is typesetting
- ◆ At the end, keep messages, discard DVI

Detailed workflow



Why are we in business?

- ◆ Each author works differently
- ◆ Language always needs improving
- ◆ Definitive XML/MathML is needed

How to put us out of business

- ◆ Create 'generic LaTeX'
- ◆ Better: submit XML/MathML
- ◆ Write translation filters