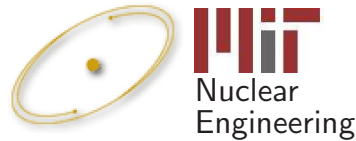


Web Publishing Mathematics With MathML

Ian Hutchinson



Professor and Head,

Department of Nuclear Science and Engineering, MIT

[Author: TtH (TeX to HTML), TtM (TeX to MathML)]

Web Publishing means serving Readers



By “publishing” I mean making (\sim static) readable pages available on the web.

There are other crucial potential MathML interactive uses, which are more dynamic. There are also uses as document exchange format that don't require the materials to be displayed or read.

MathML has in common with HTML that it is not a page-description medium.

- Author does not have precise control over what reader sees.
- Layout is partly determined by document structure, partly by browser.
- Neither HTML nor MathML are appropriate if you want the reader to see exactly what you see.
- Contrast with PDF (or PS) which IS a precise page-description medium, but is much clumsier for browsing.

What Types of Web Publishing is MathML good for? (in principle)



- Islands of mathematics inside browsable material. [e.g. problem sets]
- Documents that will be read predominantly on-screen. [increasing proportion]
- Conveniently navigable and searchable documents. [e.g. user manuals]
- Small documents that should load and display quickly. [e.g. announcements]

Notice that the compelling reasons for HTML/MathML are to do with speed and convenience in the context of a browser. MathML cannot compete with PDF as printable medium. These reasons are negated if MathML is awkward.

Authors need readers

MathML's biggest handicap over its entire (~ 7 year) history is simply that it has NOT been handled by the common web browsers. Until recently it was impossible to put up a self-contained MathML page that could be read by both IE and Mozilla/Netscape. Even now, they (Mozilla or IE) need plugins or fonts to be specially downloaded, and can't print MathML.

Insert rant about standards vs usability and poor choices ...

How do you author MathML?



You Don't!

MathML is far too verbose and intolerant a syntax to author directly (unlike HTML).
... But you don't need to.

TeX and Word (etc) both have powerful ways of translating into MathML.

By choice I don't know much about MSWord but MathType can output MathML.

There are two powerful TeX/LaTeX translators. I wrote one of them: TtM (TeX to MathML). It has been available for several years (free for linux).

You write your documents using the standard tool (TeX) that you are used to using. Then you can produce high-quality printable forms (e.g. PS or PDF using dvips or pdftex) **and** HTML (and/or) MathML (using TtH and/or TtM) from the same source. When changes are needed you change the source (tex) and retranslate.

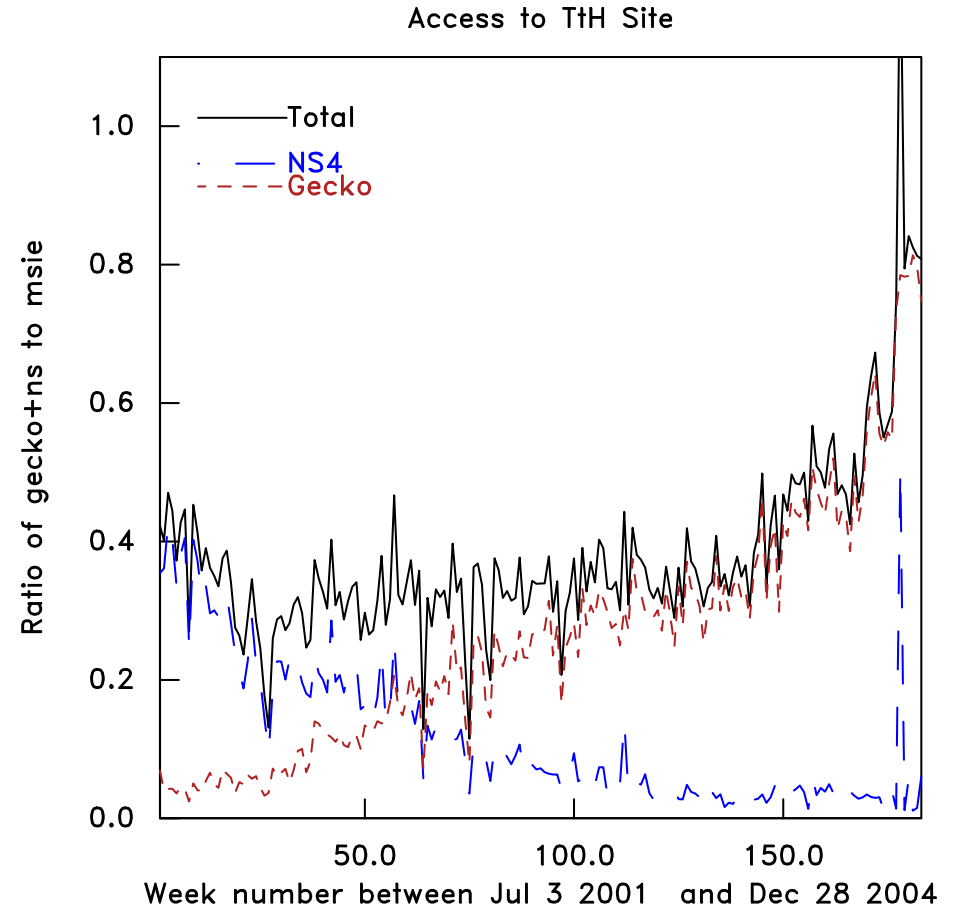
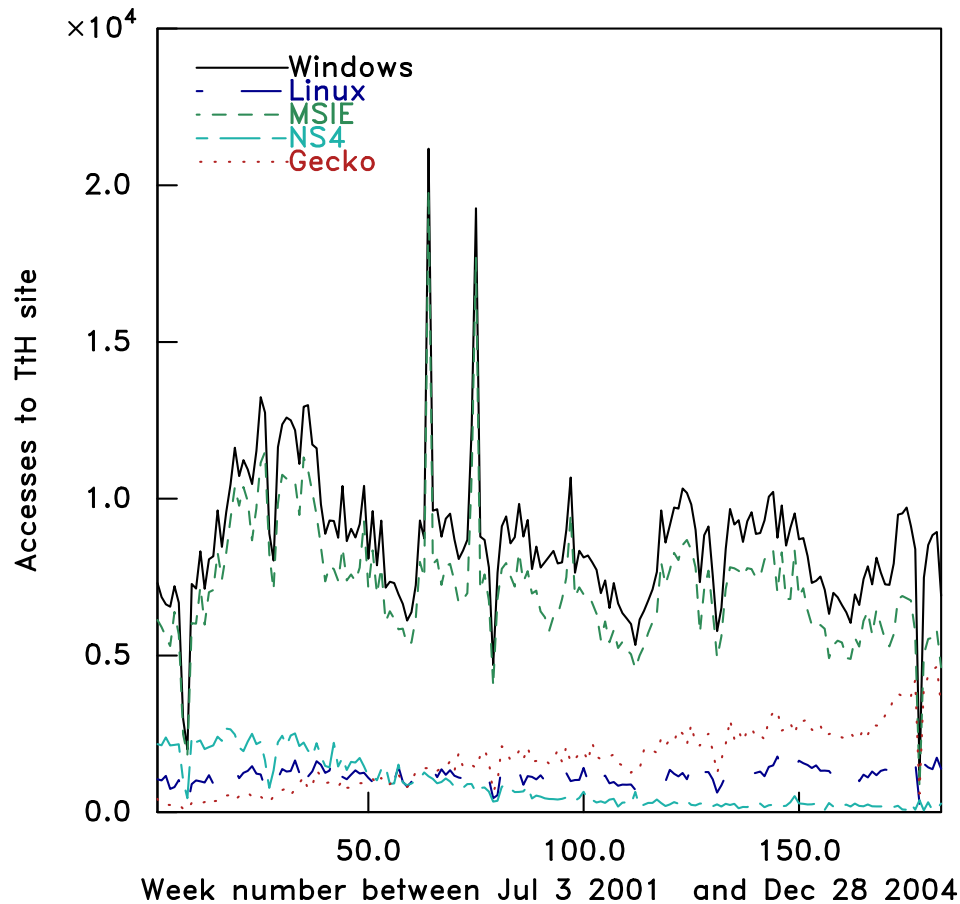
Adoption of MathML has *not* been waiting for authoring tools. They exist.

It has been (and to some extent still is) waiting for widespread rendering ability.

Some history data on browsers



Gathered for past 3.5 years at the TtH web site, which I take to be a good indicator of people who are interested in putting mathematics on the web.



Windows $\sim 80\%$. Linux $\sim 15\%$.

Slow, small linux increase.

MathML rendering is required for both Mozilla and IE.

Mozilla (Gecko-based) browsers risen to nearly 50%. Mostly in last 6 months.

Large Userbase for HTML publishing of TeX

Of order 100,000 based on download statistics of TtH and TtM

But that still means it is just a niche.

Typically 15 each TtH (HTML), TtM (MathML) **linux** executables downloaded per week. Implies perhaps comparable interest.

Typically 200 **total** TtH* (HTML) downloads per week
(including windows executables and C source).

There have been very roughly 50,000 downloads of TtH, in past 5 years.

There is substantial interest in translating TeX to HTML!

Predominantly this comes from students, academics, and researchers worldwide.

Numbers of TtM downloads are too small to make a believable assessment of user base that prefers MathML.

[* TtM is not available for free download for windows.]

Conclusions



MathML has a number of weaknesses. But it's close to being the best we've got for browsable web publishing of mathematics.

Producing MathML from the dominant mathematics authoring systems is rather straightforward.

Reading MathML is currently still not as effortless as it needs to be.

Browsers need improvements. Progress is being made.

Will MathML “take off” and take over web mathematics publishing?

My guess is that it won't.

There will remain a place for formats with perfect layout control (PS, PDF etc.)

MathML will have to establish its own niche.