



日本のメタデータ出版の現状

Current State of Meta-data Publishing in Japan

Nov. 2, 2007
Kazuhiro Hayashi
Journal Manager of CSJ
(Steering Committee member of
SPARC/JAPAN)

SPARC/JAPAN Seminar 2007

A Case Study of CSJ

SPARC/JAPAN Seminar 2007

In 1990-1996(-2001)

- CSJ-SGML project
- Collaborated with TOPPAN printing company
- Developing went well, learning a lot
- Management went wrong – irreversible process in the key work-flow
- They did not improve the DTD of SGML only ver. 1.0
- Costly

SPARC/JAPAN Seminar 2007

In 1997

- Visiting ACS (Columbus)
- SGML – Xyvison system
- Established work-flow

1997-2002

- Seeking an alternative
- Changing the system - it's really hard!

SPARC/JAPAN Seminar 2007

A Practical Solution

- TeX—3B2 system (from 2002-2007)
- Modifying SGML(XML)—3B2 system
- TeX is just like SGML or XML
- It does not have any information about the layout

At that time

- Publisher must have had their own DTD for XML
- It was flexible but one did not fit the other
- There was no standard

SPARC/JAPAN Seminar 2007

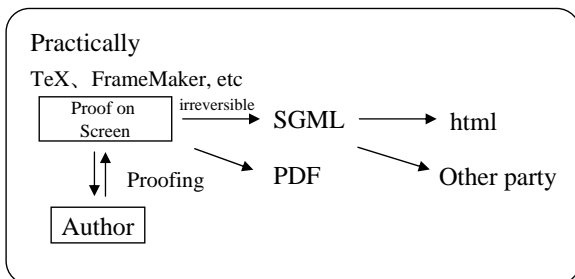
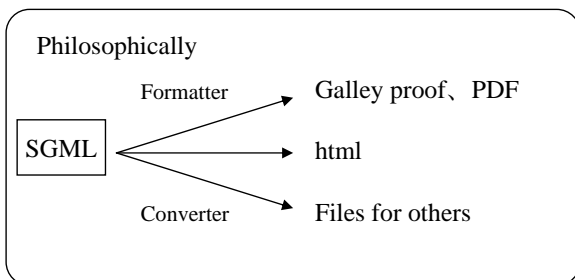


Fig.1 gap between ideal and reality of SGML publishing

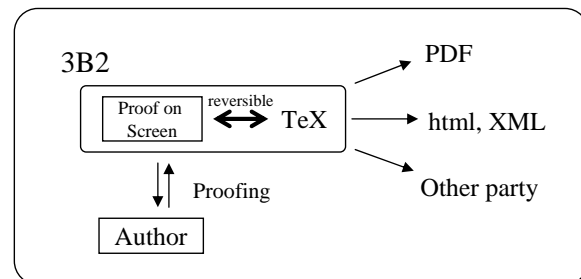


Fig.2 TeX-3B2 System

林、情報の科学と技術2002

SPARC/JAPAN Seminar 2007

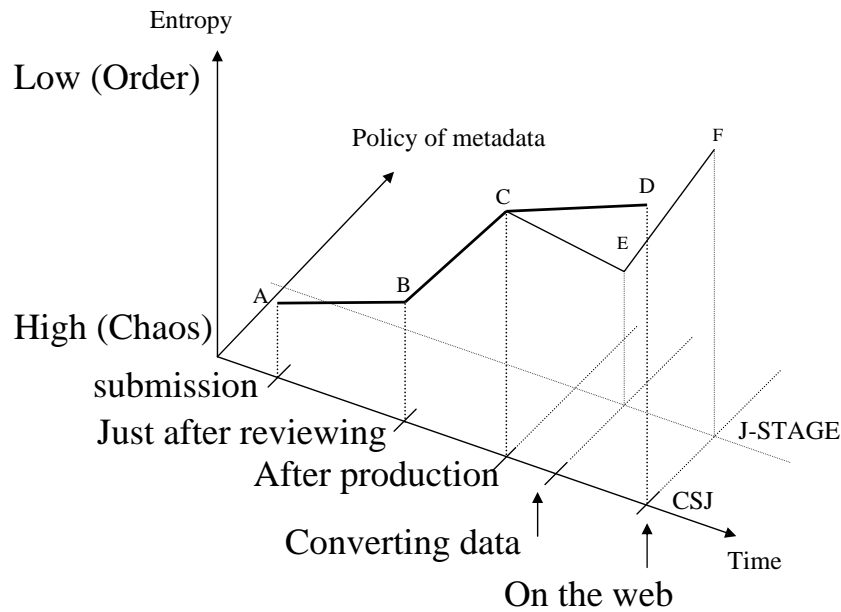


Fig.3 Concept of the entropy of metadata

林、情報の科学と技術2002(一部改変)

SPARC/JAPAN Seminar 2007

NLM-DTD

- Standard for scholarly publishing
- At least we must have a way to produce NLM-DTD based XML files from its original metadata
- Especially critical for Biology and Medicinal Journals (NLM)
- Then basically XML-based publishing can solve all the problems effectively

SPARC/JAPAN Seminar 2007

CSJ New Aspects

- To be continued

SPARC/JAPAN Seminar 2007

How about other societies?

SPARC/JAPAN Seminar 2007

Why is Meta-data Publishing not so popular in Japan?

- **Conflict with publishing in Japanese**
It must be compatible with Japanese publishing because of the cost of business investment
- **The way to upload metadata to J-STAGE**
They can use J-STAGE with a simple bibliographic text data which can be made by cut-and-paste from PDF files.

SPARC/JAPAN Seminar 2007

Severe Situation of J-STAGE

- **It works well as a point of E-journal service**
They provide full text PDF, ref-linking and so on.
- **But the meta-data is a kind of text database**
- **It can not provide xhtml online journals from their internal data now**
- **Is this just because of J-STAGE?**

SPARC/JAPAN Seminar 2007

Severe Situation of Printing Company

- It works well as a point of E-journal service
- For now?

- It cannot release online prior to print
- Again it cannot or it is hard to produce html or xhtml
- They say they are capable, but it costs much because the process is not so efficient

SPARC/JAPAN Seminar 2007

How about TeX?

We are comfy with TeX!

SPARC/JAPAN Seminar 2007

Advantage and Disadvantage of TeX

- Yes, TeX is excellent! I agree
- Skilled authors can do anything
- It is still inevitable for math-equations

However,

- It is “noisy” from a view point of meta-data
- It needs compiling to see the print-view
- Terrible authors’ macro
- Not popular except for in some areas (math)
- Incompatible with MS-WORD

Again NLM-DTD

- Standard for scholarly publishing
- At least we must have a way to produce NLM-DTD based XML files from its original metadata
- Especially for Biology and Medicinal Journals (NLM)
- Basically XML-based publishing can solve all the problems

Not Pessimistic

- It is a good time to introduce the settled meta-data publishing
- Standard for scholarly publishing
- Cost effective

To improve the situation of scholarly publishing in Japan