



# SGML & Markup

Stephen Miller

AAC / ÖAW

*Texts & Files* SS 2007

Universität Wien



# Steven DeRose

- Coombs, James H., Allen H. Renear, and Steven J. DeRose. 1987. "[Markup Systems and the Future of Scholarly Text Processing](#)." *Communications of the ACM* 30 (11): 933-947.
- DeRose Steven J., David G. Durand, Elli Mylonas, and Allen H. Renear. 1990. "What is Text, Really?" *Journal of Computing in Higher Education* 1 (2): 3-26.



# What is Markup?

Stephen Miller

AAC / ÖAW

*Texts & Files* SS 2007

Universität Wien



# Concepts of Markup

- Presentational
- Procedural
- Descriptive
  - James H. Coombs, Allen H. Renear, Steven De Rose, “Markup Systems and the Future of Scholarly Text Processing,” *Communications of the ACM*, 30 (1987): 933-47.
  - <http://www.oasis-open.org/cover/coombs.html>



# Types of Markup

1. Presentational
2. Procedural
3. Descriptive



# Print

## SEASONS AND MONTHS

### *The New Year*

I. Manx, Yn Vlein Noa. *The New Year*. January 1st.

“... anciently the first day of November was the first day of the year: and the Mummings on the eve of All Saints' Day still begin their petition with these remarkable words, 'To-night is New Year's night. Og-u-naa. The Moon shines fair and bright. Tro-la-la.'” Kelly (*circa* 1790), s.v. Baal-sauin. *See* November 1st.



# Typescript

## Seasons and Months

### The New Year

I. Manx, Yn Vlein Noa. The New Year. January 1st. ''... anciently the first day of November was the first day of the year: and the Mummings on the eve of All Saints' Day still begin their petition with these remarkable words, 'To-night is New Year's night. Og-u-naa. The Moon shines fair and bright. Tro-la-la.''' Kelly (circa 1790), s.v. Baal-sauin. See November 1st.



# Presentational

Season Header1 : uppercase

Header2 : italic / centred

Use Adobe Garamond  
+ Old Style figures

I. Manx, Yn Vlein Noa. The New Year.  
January 1st. "... anciently the first  
day of November was the first  
year: and the Mummers on the  
Saints' Day still begin their  
with these remarkable words,  
is New Year's night. Og-u-na  
shines fair and bright. Tro-  
Kelly (circa 1790), s.v. Ba  
November 1st.

Body text: 10/12

Header1 12/14

Header2 10/12

Footnotes 9/10



# Procedural

```
\input macros.tex
\centerline{{\title{}}Seasons and Months}}
\medskip
\centerline{{\subtitle{}}The New Year}}
\smallskip
```

```
{\bf{}}I.} Manx, Yn Vlein Noa. {\it{}}The New Year}.
January 1st. ''... anciently the first day of
November was the first day of the year: and the
Mummers on the eve of All Saints' Day still begin
their petition with these remarkable words, 'To-
night is New Year's night. Og-u-naa. The Moon
shines fair and bright. Tro-la-la.''' Kelly
({\it{}}circa} 1790), s.v. Baal-sauin. {\it{}}See}
November 1st.
```

(Plain TeX)



# Descriptive

```
<div1 type="chapter">
<head type="main">Seasons and Months</head>
<div2 type="section">
<head type="sub">The New Year</head>
<p>
<hi rend="strong">I.</hi> Manx, Yn Vlein Noa. <hi rend
="emphasis">The New Year</hi>. January 1st.
&ldquo;&hellip; anciently the first day of November
was the first day of the year: and the Mummings on the
eve of All Saints&apos; Day still begin their
petition with these remarkable words, &lsquo;To-night
is New Year's night. Og-u-naa. The Moon shines fair
and bright. Tro-la-la.&rsquo;&rdquo; Kelly (<hi
rend="emphasis">circa</hi> 1790), s.v. Baal-sauin.
<hi rend="emphasis">See</hi> November 1st.
</p>
</div2>
```

(TEILite DTD)



# SGML

Stephen Miller

AAC / ÖAW

*Texts & Files* SS 2007

Universität Wien



# SGML

- Standard Generalized Markup Language ISO 8879 (1986)



# SGML, HyTime, DSSSL

- SGML ISO 8879:1986
  - Document Structure
- HyTime ISO/IEC 10744:1992
  - Hypermedia
- DSSSL ISO/IEC 10179:1996
  - Document Style Semantics



# SGML

- “SGML is a metalanguage for defining markup languages, and HTML is an instance, specified as a document-type definition (DTD), of such a language.”
  - Larry Press, *CACM* 38/3 (1995): 21.



# Before SGML... GML

- C. F. Goldfarb, “A Generalized Approach to Document Markup.” *Proceedings of the ACM SIGPLAN SIGOA Symposium on Text Manipulation*. New York: ACM, 1981. 68-73.
  - Adapted as "Annex A. Introduction to Generalized Markup" in ISO 8879 (1986)



# Before GML... GM

- *A Brief History of the Development of SGML*
  - [www.sgmlsource.com/history/sgmlhist.htm](http://www.sgmlsource.com/history/sgmlhist.htm)



## And now, XML...

- “What has been will be again, what has been done will be done again; there is nothing new under the sun.”

Ecclesiastes 1:9 NIV



# SGML

- “SGML is a **metalanguage** for defining **markup languages**, and HTML is an **instance**, specified as a **document-type definition (DTD)**, of such a language.”
  - Larry Press, *CACM* 38/3 (1995): 21.



# Keywords

1. Metalanguage
2. Markup Language
3. Document Type Definition



# Structured Documents & SGML

1. Documents are **generic**
  - concept of **Document Type**
2. Content can be **modelled**
  - concept of **Content Model**



# Document Type Definition (DTD)

- Markup required by a Document Type to present the Content Model formally declared in a Document Type Definition (DTD)
- The DTD is written in SGML and
  - defines the **allowable** markup
  - the **rules** for the use of that markup



# Validation

- How do you know that you have followed the tags and rules for their usage set out in the DTD enforced?
- By validating your document instance against the Document Type Definition using an SGML Parser
- If the file passes without error then you have created a valid SGML instance