





# SGML


Stephen Miller  
AAC / ÖAW  
*Texts & Files* WS 2005–06  
Universität Wien



## 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.


2 Texts & Files WS 2005–06 / SGML © AAC / ÖAW 2005



## Keywords

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


3 Texts & Files WS 2005–06 / SGML © AAC / ÖAW 2005



## Structured Documents & SGML

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


4 Texts & Files WS 2005–06 / SGML © AAC / ÖAW 2005



## 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

5 Texts & Files WS 2005–06 / SGML © AAC / ÖAW 2005



## 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 files passes without error then you have created a valid SGML instance

6 Texts & Files WS 2005–06 / SGML © AAC / ÖAW 2005