

# Data Entry Specs 1.0.1 Overview

Wolfgang Schmidle, Klaus Thoden, Malcolm D. Hyman  
Max Planck Institute for the History of Science, Berlin, Germany

6th October 2008

## 2 General Markup

### 2.1 Pages

Type the entire content of one page, then go on to the next page. Do not mix the content of different pages.

#### 2.1.1 Page breaks, Page Numbers and Running Heads

Page breaks are marked by `<pb>`. If the page has a page number, type it within the `<pb>` tag. Type the page number exactly as it appears in the book. If there is a running head on the page, it is marked by `<rh>` and `</rh>`. Type the running head immediately after the `<pb>` tag.

#### 2.1.2 Catchwords and Signatures

Do not type catchwords and signatures.

### 2.2 Text Blocks

#### 2.2.1 Headings

Headings are marked by `<h>` and `</h>`.

#### 2.2.2 Paragraphs

Paragraphs are marked by `<p>` and `</p>`.

#### 2.2.3 Block Quotations

A block quotation is marked by `<q>` and `</q>`. Do not type repeating quotation symbols.

## **2.3 Columns and Tables**

### **2.3.1 Columns**

Columns are marked by `<col>` and `</col>`. The columns on a page are numbered from left to right.

### **2.3.2 Tables**

A table is marked by `<tb>` and `</tb>`. Use # as field separators. Do not type horizontal or vertical lines.

## **2.4 Notes**

### **2.4.1 Marginal Notes**

A marginal note on the left is marked by `<mg1>` `</mg1>`, and a marginal note on the right is marked by `<mgr>` `</mgr>`. Type the marginal note in separate lines, starting after the line it is the closest to.

### **2.4.2 Footnotes**

Footnotes are marked by `<fn>` and `</fn>`. Type the footnote where it appears on the page. Insert the footnote symbol in the `<fn>` tag. The corresponding footnote symbol in the main text is marked by `<n>`. Insert the footnote symbol in the `<n>` tag.

## **2.5 Printed Images**

### **2.5.1 Figures**

Where a figure occurs in the text, put a `<fig>` tag in a separate line.

### **2.5.2 Drop Caps**

Type a drop cap as an ordinary capital letter. Do not use the `<fig>` tag for ornamented drop caps.

## **2.6 Handwritten Notes**

Handwritten notes are marked by `<hd>`. Do not type the note itself.

## **2.7 Unreadable Text**

### **2.7.1 Characters You are Unsure About**

If you are not sure about a character, type `<?>` after it. If you are unsure about a whole paragraph, type `<?>` directly after the `<p>` tag, i.e. `<p><?>`. A completely unreadable character is typed as `@`. If it is unclear how many characters are unreadable, use `<gap>` instead of `@`.

### **2.7.2 Unknown Characters**

If there is an unknown character in the text, add it to the list of unknown characters. From then on, type its number whenever it occurs in the text.

## **3 Latin Alphabet**

### **3.1 General**

#### **3.1.1 Punctuation**

Type a space after the punctuation marks `. , : ; ! ?` Type a space before the opening brackets `( [` and after the closing brackets `) ]`

#### **3.1.2 Hyphens**

If the line ends with a hyphen, type it.

### **3.2 Characters**

#### **3.2.1 Characters to be Typed Directly**

If a character with diacritics can be easily typed and encoded as Unicode character, type it directly.

#### **3.2.2 Other diacritics**

If a character cannot be typed directly but is a combination of a known character and some known diacritic, type `\`, then the diacritic, then the character.

### 3.3 Type Styles

#### 3.3.1 Italics

Use underscores `_ _` to mark the beginning and end of words or whole lines in italics. Encode only up to a few lines of text in this way. A whole paragraph in italics is marked by `it` in the `<p>` tag, i.e. `<p it>`.

#### 3.3.3 Small Caps

If you can identify capital letters as small caps, use `<sc>` `</sc>`.

#### 3.3.4 Subscript and Superscript

Use `<_>` and `</_>` for subscript. Use `<^>` and `</^>` for superscript.

### 3.4 Latin Ligatures

Resolve easy ligatures silently. Resolve difficult ligatures, too, but type `{ and }` around them.

## 4 Greek Alphabet

### 4.1 Characters

Type greek characters directly as Unicode characters.

### 4.2 Punctuation

The rules for Latin punctuation apply. In addition, type the mid-dot `·` directly.

### 4.3 Greek Ligatures

Resolve letter variations silently. Resolve all ligatures and type `{ and }` around them. If a ligature contains some diacritics, type them.

## 5 Mathematical Symbols

### 5.1 Fractions

Type fractions in one line. Use `{ / }` to mark beginning, fraction line and ending.