Base Units for Multimodal Documents

Everything which can be seen on each page of a multimodal document should be captured as a **base unit**.

continuous homogeneous text block
sentence fragments initiating a list
headings, titles, headlines
photos, drawings, diagrams, figures (without caption)
captions of photos, drawings, diagrams, tables
text in photos, drawings, diagrams
icons
table cells
list headers list items
list labels (itemizers)
items in a menu
page numbers
footnotes (without footnote label)
footnote labels
running heads
emphasized text
horizontal or vertical lines which function as delimiters between columns or rows
lines, arrows, polylines which connect other base units

Everything on the page should belong to just **one** base unit. Sentences divided by a page or column break should be marked as two base units. The caption of figures, tables etc. is always marked as an extra base unit. Horizontal or vertical lines are marked as separate base units if they cannot be viewed as part of another layout element.

In certain very restricted cases, we allow base units inside base units:

- visual layout

    - emphasized text portions in a sentence/heading
    - icons or similar pictorial signs in a sentence
    - text pieces in a diagram or picture
    - arrows and other graphical signs in a diagram or picture
    - menu items in an interactive pop up menu

- other function

    - orthographic sentences
    - document deictic expressions occurring in a sentence