

# Adobe Portable Document → Format: Risks for long-term preservation

Johan van der Knijff  
Koninklijke Bibliotheek  
November 16, 2009

## Background on PDF research at the KB

- KB's mandate to ensure long-term preservation and accessibility of received materials
- PDF now prevalent format for digital (scientific) publications: almost 90% of all files in e-Depot is in PDF format.

## Questions

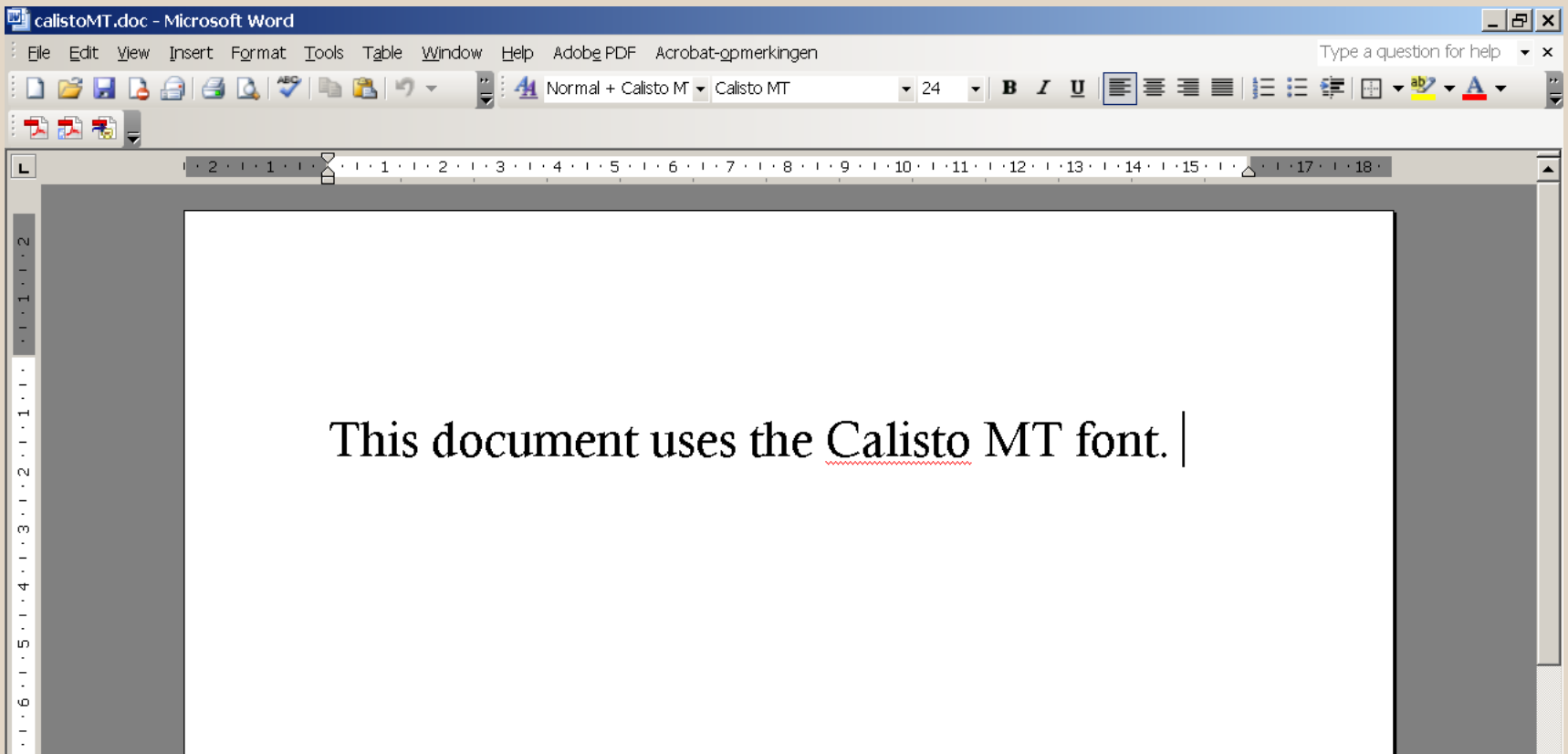
- So what are the long-term preservation/accessibility risks for materials in existing collection?
- How can we mitigate these risks?
- How can we minimise the risks for new materials?

## Outline

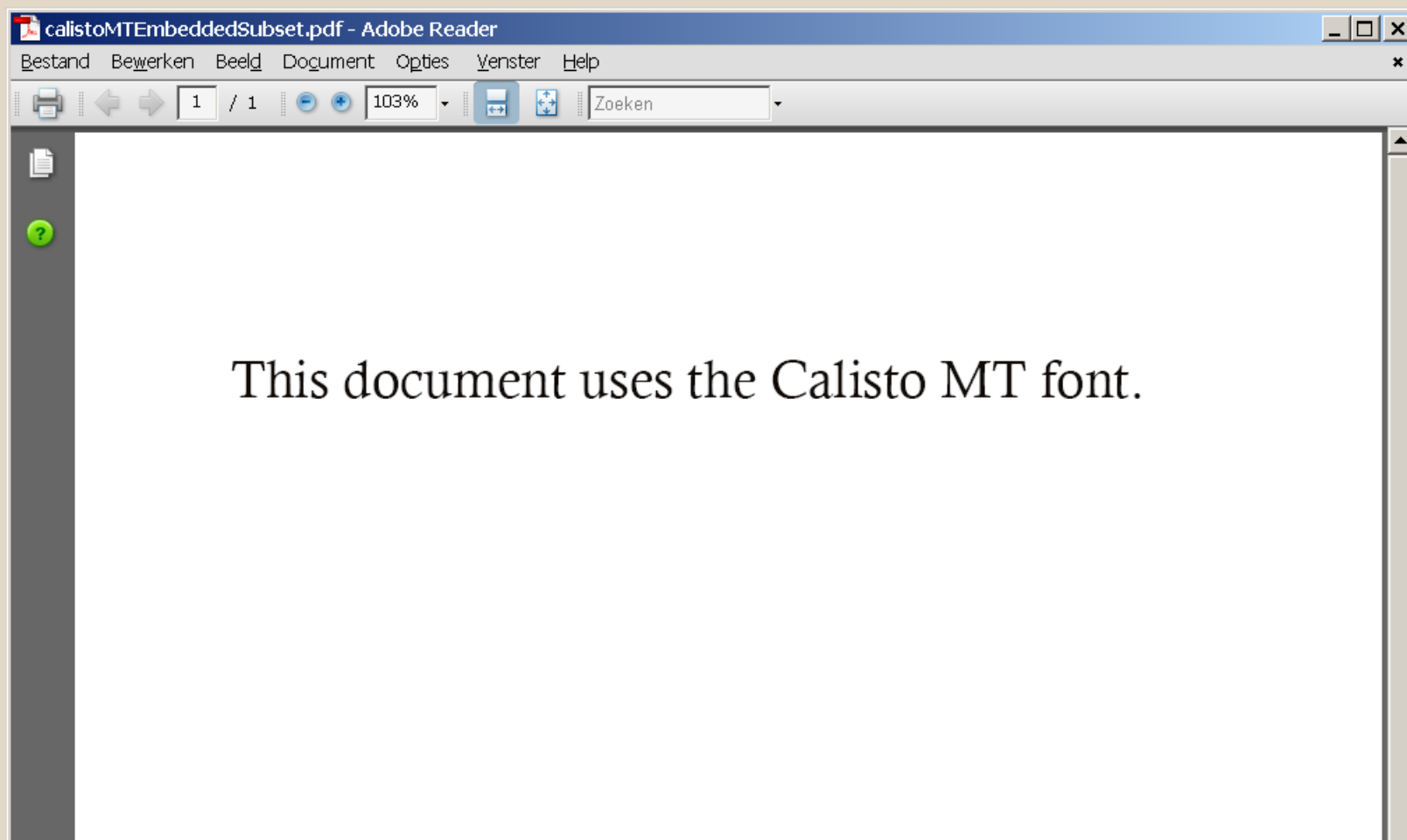
- Fonts
- External references
- Embedded content

# Fonts

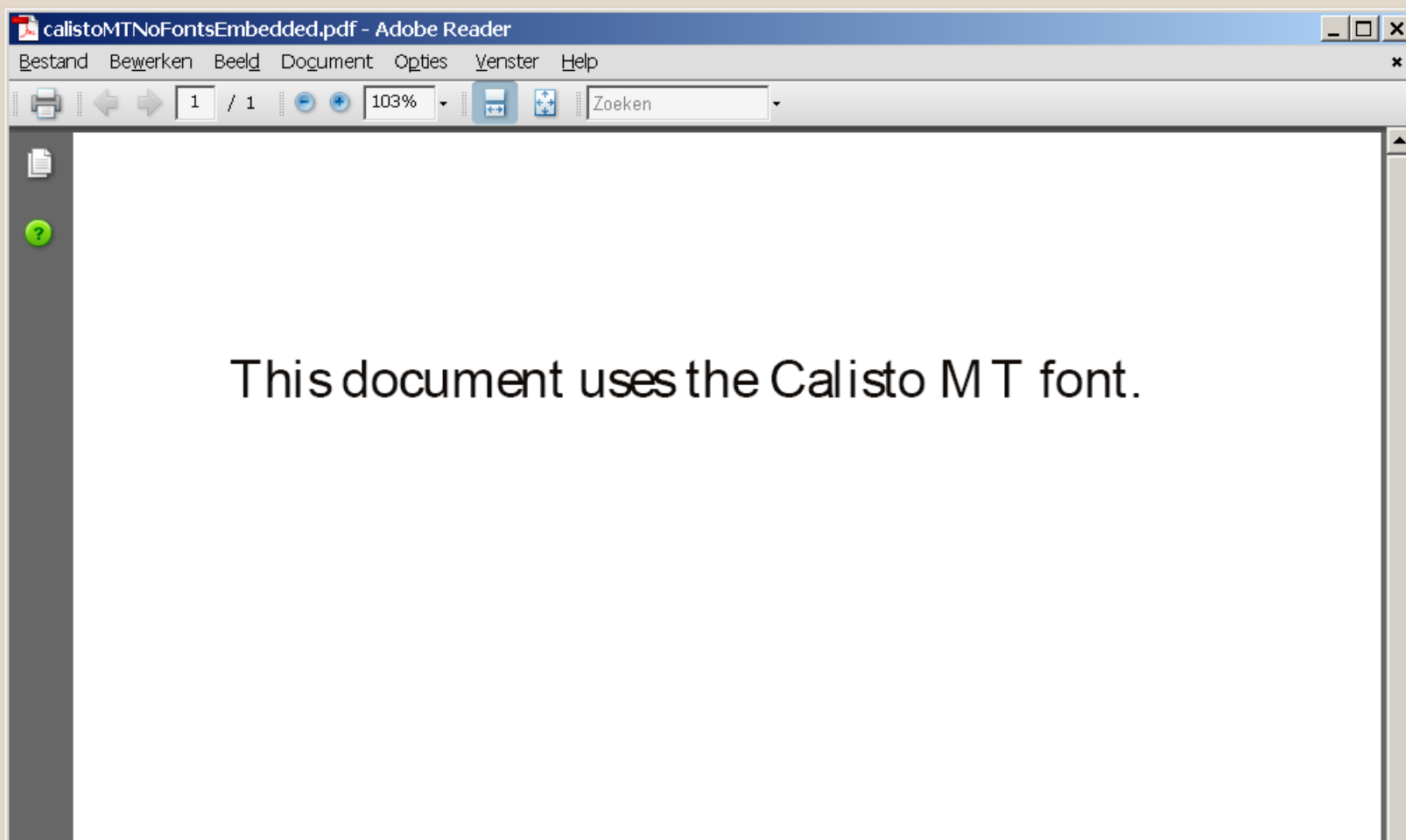
Export the following Word document to PDF without embedding any fonts:



## Appearance on creator's system:



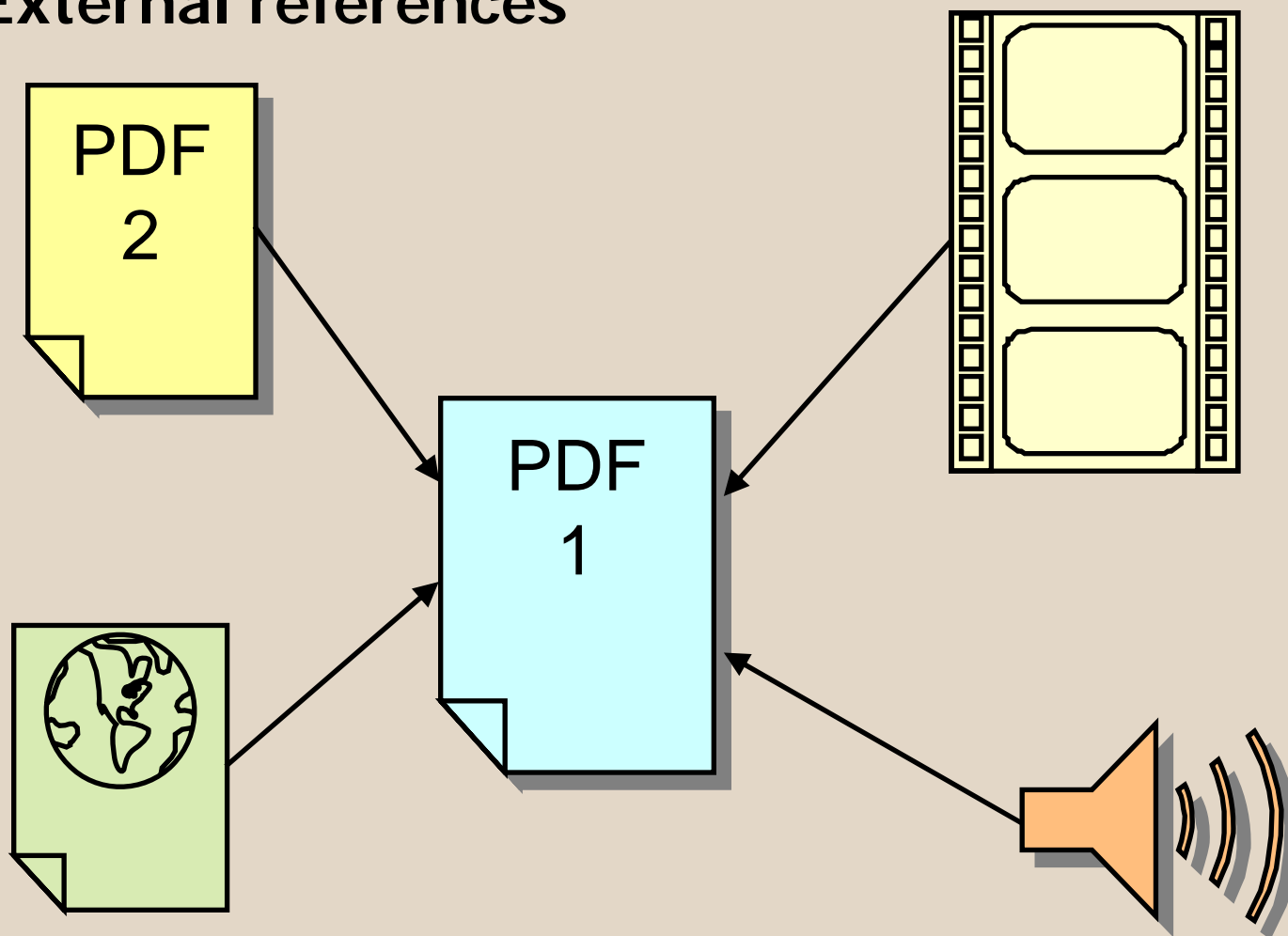
# Appearance on some user's system in 30 years time?



## Risks

- Appearance of documents with non-embedded fonts can be very different from what the creator had in mind (worst case: illegible!)
  - ‘Standard’ fonts on current operating systems may become exotic in future environments
- Use tools such as JHOVE to establish whether all fonts in a file are embedded

## External references



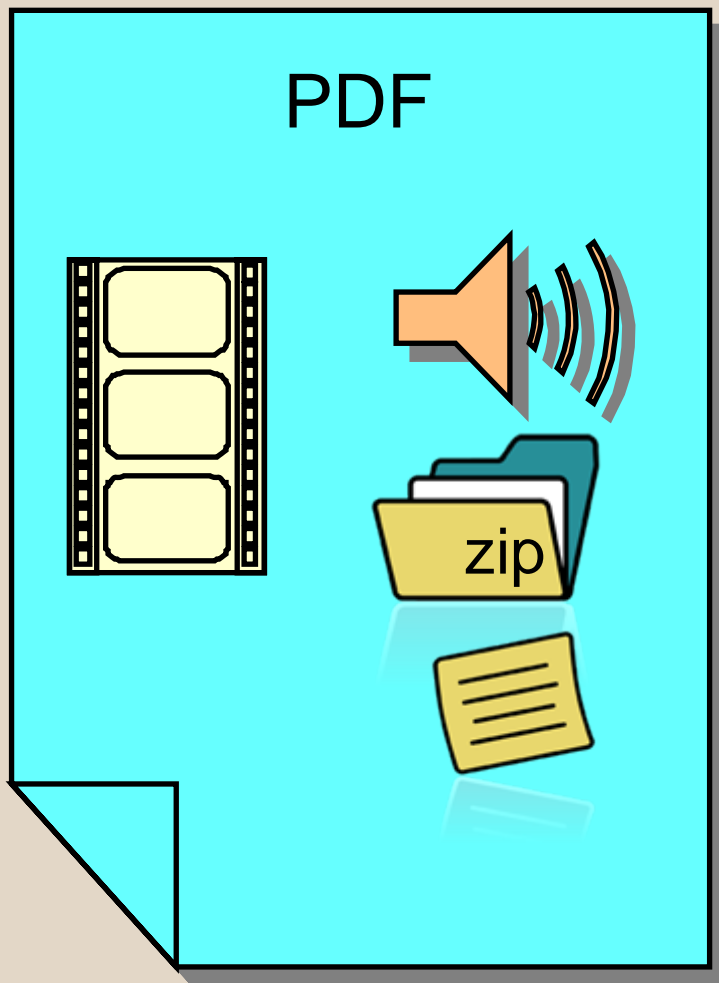
## Risks

- External objects may become inaccessible
- Display of external objects can rely on external applications that may not be available in the future

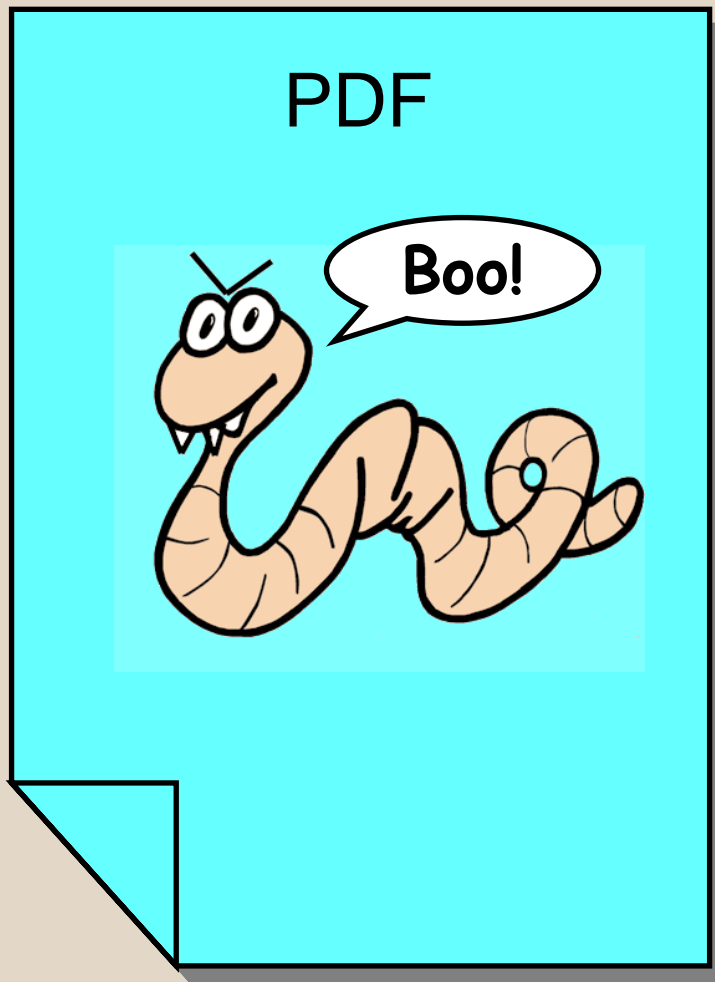
Example: embedded video in some obsolete format

- This will affect document appearance (and possibly actual contents)

# Embedded content



This can be embedded as well



## Risks

- Display of embedded objects can rely on external applications that may not be available in the future

Example: embedded video in some obsolete format

- *Anything* can be embedded/attached, so embedding of viruses, worms and other malicious code is possible as well

## Conclusions

- PDF format includes several features that compromise the long-term accessibility of PDF documents
- For new materials the risks can be minimised by following KB PDF Guidelines (next slide)
- The KB will use tools (when available) to monitor the risks

## KB PDF Guidelines

Downloadable here:

[www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/PDF\\_Guidelines.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/PDF_Guidelines.pdf)

Note: Guidelines are updated on a regular basis