Rebuilding define.pdf by reading Word and XFDF files

David Garbutt
BIOP AG
Basel

Where do the fields come from?

LaTeX example input

LaTeX example output

LaTeX example result is PDF

MS Word XML format

The original table

WORD XML structure

Result!

Where has all the structure gone?

Let me count the ways...

What is XFDF?

Reading XFDF with a XML Map

Read with SAS and reshape

Split into page and variable info

Example output as csv

Ways of building define PDF

Proposal

Further reading

Rebuilding define.pdf by reading Word and XFDF files

David Garbutt
BIOP AG
Basel

Where do the fields come from?

LaTeX example input

LaTeX example output

LaTeX example result is PDF

MS Word XML format

The original table

WORD XML structure

Result!

Where has all the structure gone?

Let me count the ways...

What is XFDF?

Reading XFDF with a XML Map

Read with SAS and reshape

Split into page and variable info

Example output as csv

Ways of building define PDF

Proposal

Further reading