



WikiHist.html:

English Wikipedia's Full Revision History in HTML Format

Blagoj Mitrevski, Tiziano Piccardi, Robert West



A faint, light-gray network diagram serves as the background for the central text. It consists of numerous thin, gray lines forming a complex web of triangles and polygons, resembling a molecular or network structure.

**Can we rely on Wikitext to get
the links on a Wikipedia page?**

Implement.

Problem

"**Niue**" (`{{lang-niu|Niu^-e}}`) is an [[island country]].

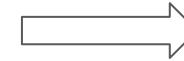
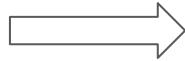


Modules not available in the dump (LUA)

`Niue (Niuean: <i lang="niu">Niu^-e</i>) is an island country.`

Multiple Docker instances

Wikitext Full history

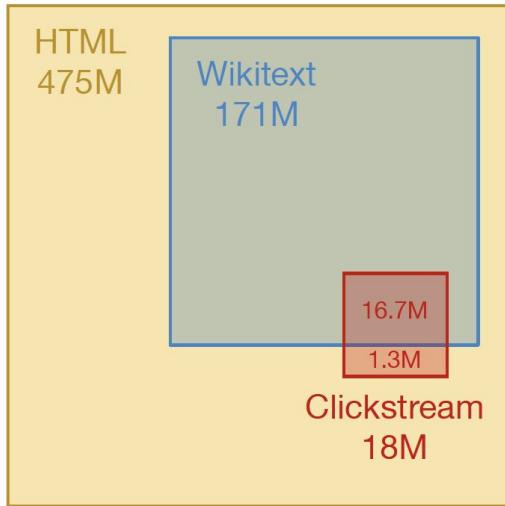


HTML Full history

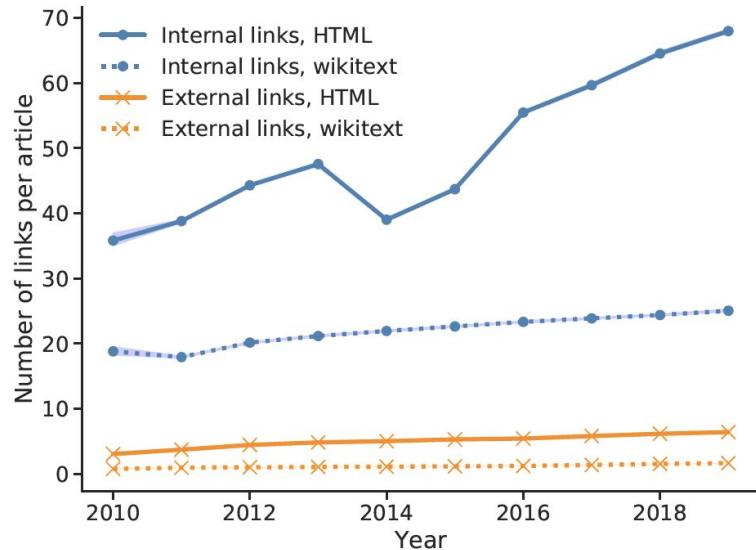
Page rendered by using the correct template revision

What did we discover?

Clickstream shows transitions that are not possible on the Wikitext links network



In average in 2019 a page contains 3 times the number of links visible in Wikitext



Full dataset (7TB) available at:
<https://zenodo.org/record/3605388>