

On the Credibility of Wikipedia: an Accessibility Perspective

Rui Lopes, Luís Carriço

{rlopes, lmc}@di.fc.ul.pt

HCI & Multimedia Group
LaSIGE/Universidade de Lisboa

Outline

- Background
- Credibility model of Wikipedia
- Experiment
- Discussion

Background

Web Architecture

- Web's architecture is decentralised.
 - **Good!** No central control. Everyone contributes independently with information and links.

Web Architecture

- **Credibility** and **verifiability** is leveraged through the Web's architecture:
 - Anyone can create a Web page with information, and
 - Link to external Web sites to corroborate the credibility of such information, build trust (Nielsen, J., 2000)

It's the user.

- The **end user** reaches the Web page and, when in doubt, follows external links to verify the credibility of its information.
- Information must be **accessible** (in *strict* and *broad* senses) in its UI, otherwise not verifiable and, consequently, reduced **credibility** (Muir & Morai, 1996).

Web Arch. (reprise)

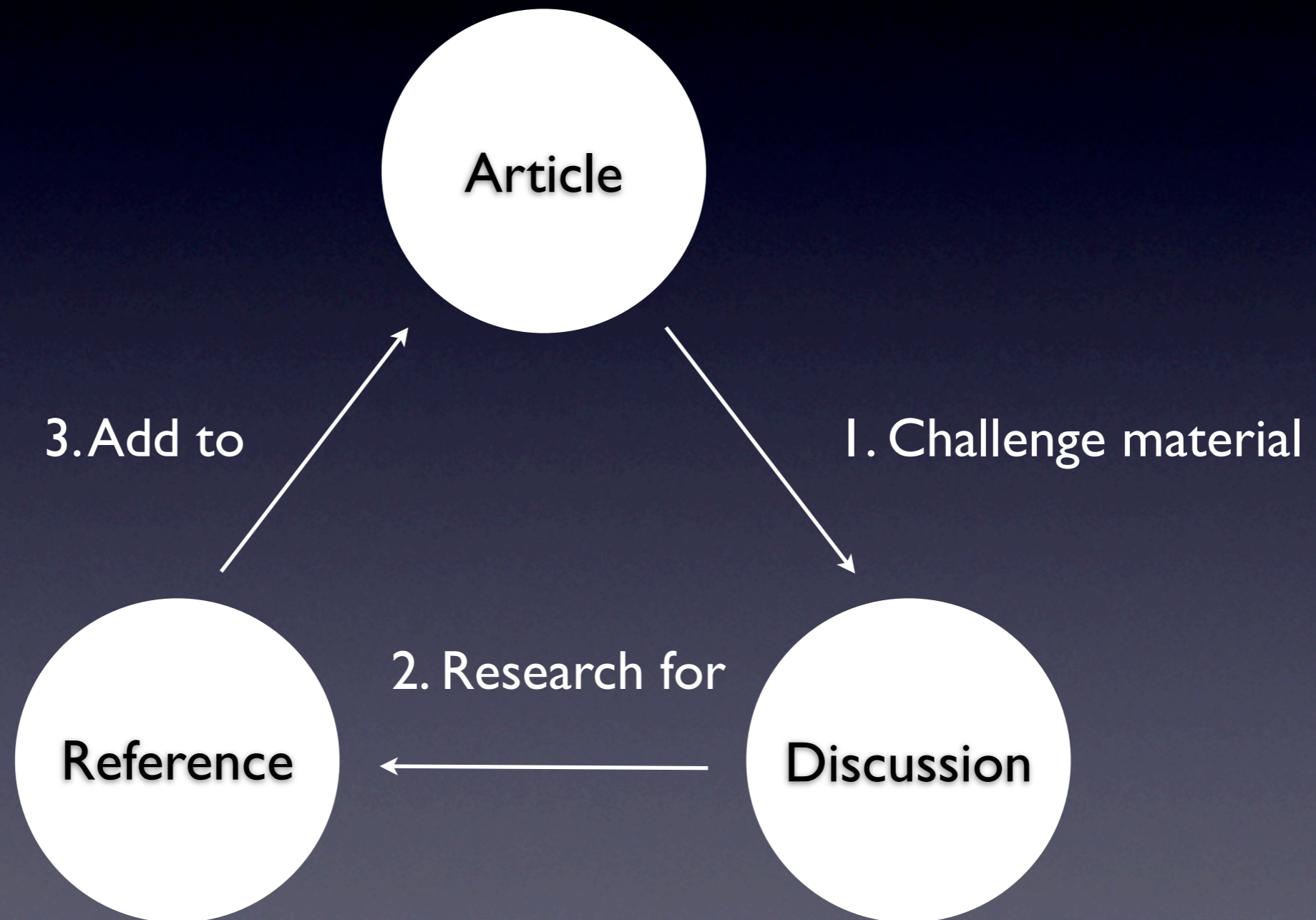
- Web's architecture is decentralised.
 - Good! No central control. Everyone contributes independently with information and links.
 - **Bad!** No central control. Changed information (and corresponding UI) without notice, broken links.

Credibility on Wikipedia

Wikipedia

- Free, editable-by-all online encyclopaedia
- Iterative, collaborative editing thrives for (perceived sense of) quality, to ensure its **credibility**
- Credibility policies defined through self-organisation of contributors

Credibility-building Lifecycle



I. Challenge material



This article does not [cite](#) any [references or sources](#).

Please help [improve this article](#) by adding citations to [reliable sources](#). [Unverifiable](#) material may be challenged and removed.



This article needs additional [citations for verification](#).

Please help [improve this article](#) by adding [reliable references](#). Unsourced material may be [challenged](#) and removed.

layouts of the keys^{[[citation needed](#)]},
of the keyboard.

counterparts^{[[verification needed](#)]},
they are legal). This stance is large

tator, who was banned from
policies.^{[[27](#)][[unreliable source?](#)]}

- `{{unreferenced}}`,
`{{refimprove}}`
- Discussion follows
- `{{fact}}`,
`{{verify_source}}`,
`{{verify_credibility}}`
- Discussion follows

2. Research for References

- Each contributor searches for a credible source (from her point of view)
 - Verifiable (and accepted) by contributors
 - Limited to their knowledge, experience, **abilities, disabilities**

3. Adding References

- Footnotes are created with `[[cite book]]`, `[[cite journal]]` and, specially, with `[[cite web]]`
- Article information is updated with new citations and references, with **navigable links**
- *(repeat the cycle)*

Does this lifecycle
delivers credible
information for
everyone?

Experiment

Method

- Accessibility **evaluation** framework
 - Subset of WCAG (16 automated *checkpoints* applied to HTML structures)
- Setup:
 - 100 (random) seed pages from Wikipedia
 - Extract all hyperlinks to external pages
 - Tidy (to detect problems in HTML)
 - Checkpoints yield accessibility quality (%
passed/total)

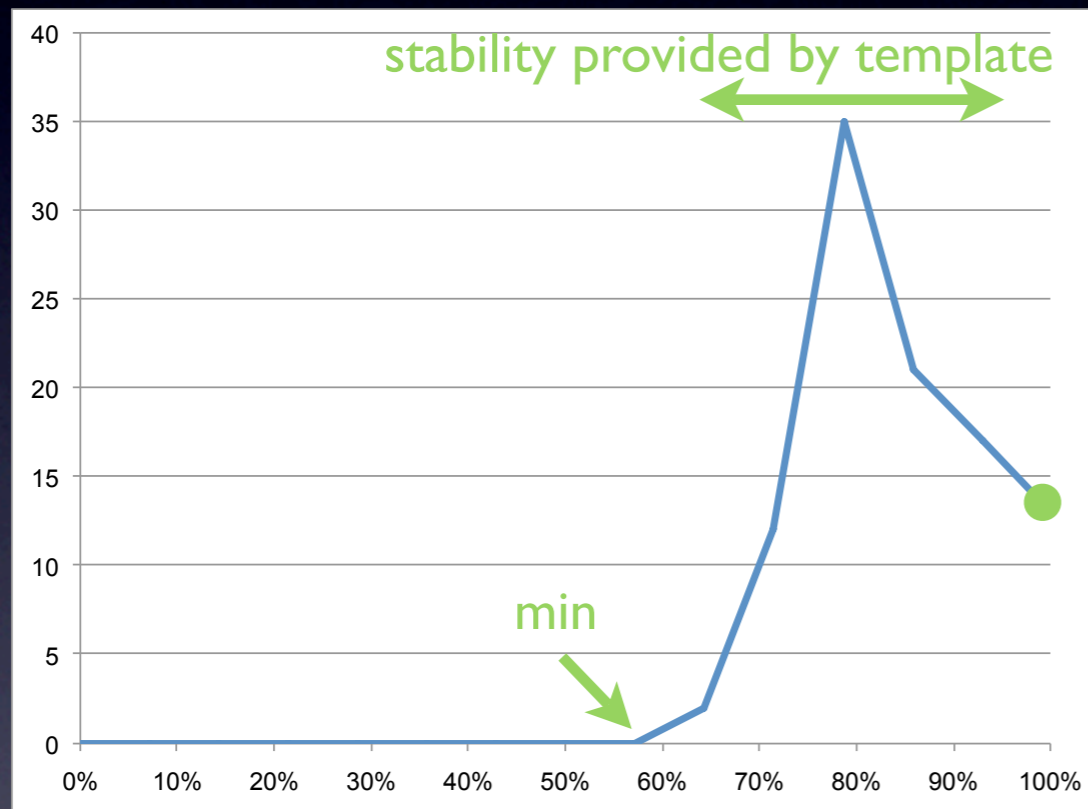
Results

- ~2.5x ratio of external pages per article
- 13% seed reach 100% quality vs 2% for external (world **wild** web!)

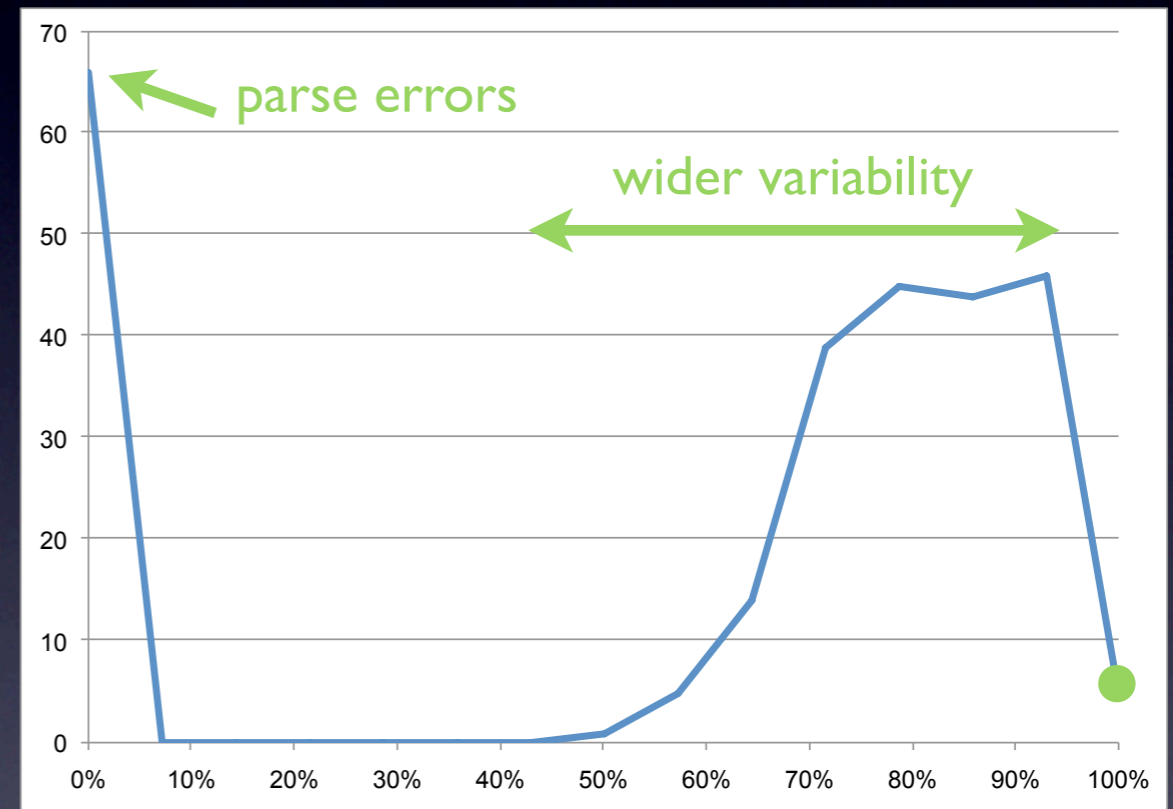
	Seed	External
Avg	84%	60%
Max	100%	100%
Min	64%	0%
Stddev	9%	36%

Results

(# pages vs. quality)



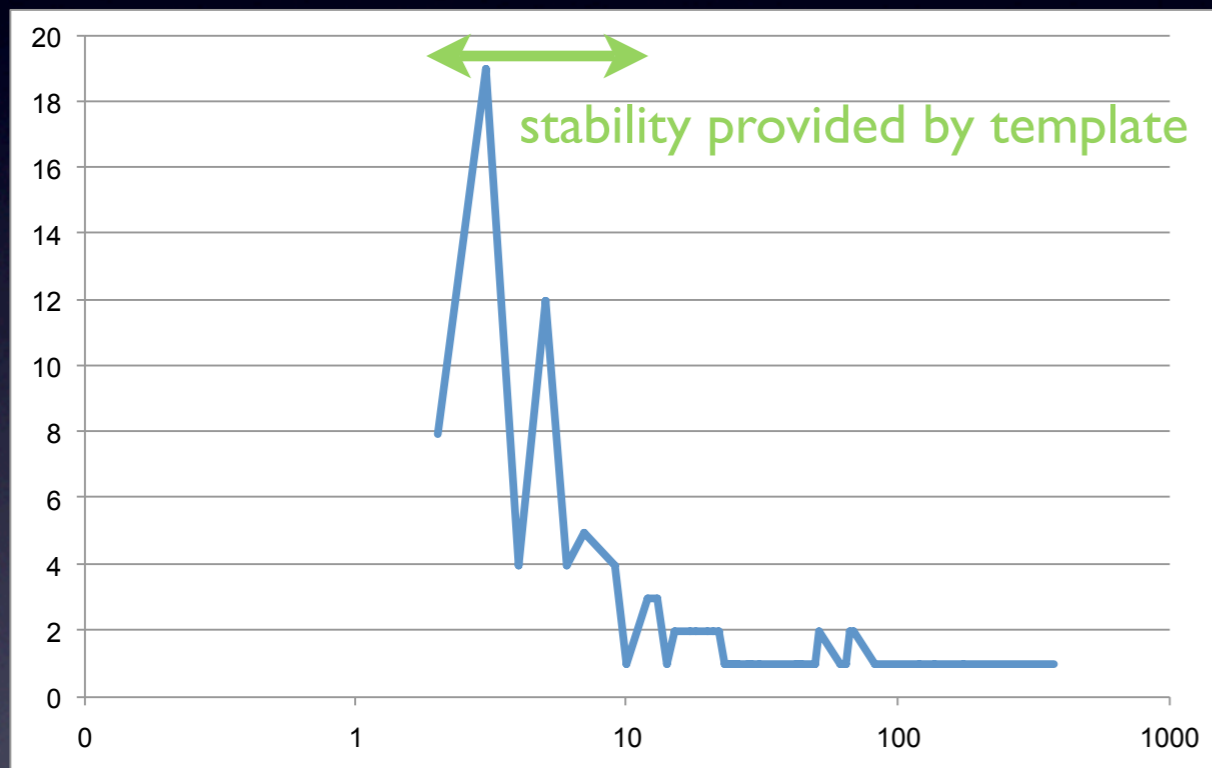
Seed pages



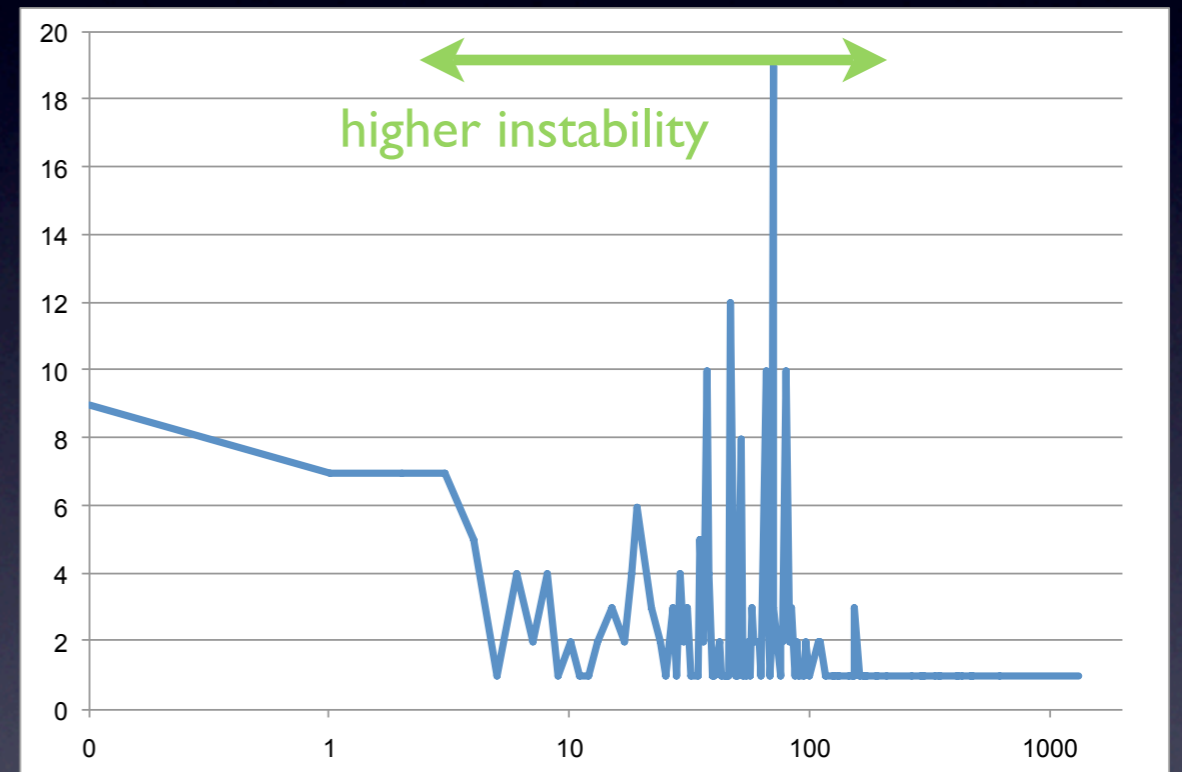
External pages

Results

(# pages vs. parse warnings)



Seed pages



External pages

Discussion

Contributors' Unawareness

- External pages linked by Wikipedia have **decreased accessibility** and **highly unstable quality**
- Dead links can be tracked with `{{dead link}}`, but seldom used mechanism
- Wiki markup provides highly accessible HTML, but no guidance on link creation: **contributors are on their own!**

Design Implications

- Important not to disrupt contributors' policies and processes, apply **incremental improvements** to the lifecycle:
 - *Challenge material*: `{{refinaccessible}}`, `{{verify_accessibility}}`
 - *Research for references*: availability and guidelines for automated accessibility evaluation tools (ETs)
 - *Adding references*: integrate ETs into Mediawiki, bots, apply policies

Limitations

- **Unique** external information sources (better one than none)
- **Ranking** of external pages not taken into account (is *higher rank* better than *more accessible*?)
- Only **automated** evaluation tests (though expert analysis follows the same pattern)

Summary & Ongoing work

- Instability and lower accessibility quality of citations on Wikipedia hurts its credibility
- Further analyse **correlation** between each article and its own external links
- Expand analysis to **other** authoritative information sources
- Compare results with **user studies** (and with different types of disabilities)

Thank you.

rlopes@di.fc.ul.pt