The Reliability of Wikipedia

Zsuzsanna Müller
PhD student
University of Debrecen, Hungary

Internet

- the access to information has become easier

- sheer number of information sources has been increasing

- there can be found more and more anonym texts in the web containing incomplete, inexact and incorrect information

the **evaluation of reliability** has become one of the most important issues of **information acquisition**

The last few years **evaluation criteria** can be applied efficiently in the processes of information seeking and evaluation were determined by:

- studies from e.g. Rieh and Belkin; Frieder, Kantor, Cool and Belkin; Amento
- librarian aids e.g. of UC Berkely Library, Cornell University Library
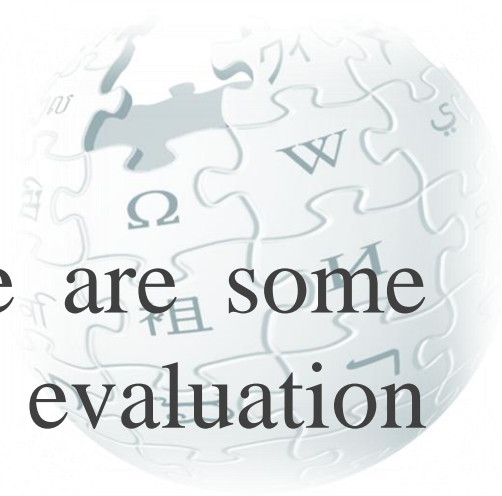
In spite of these efforts there are some **shortcomings** in the field of evaluation of online information sources

- earlier studies focus on the examination of quality and not specifically of reliability

- previous researches on the topic are too specific, while librarian aids are too general
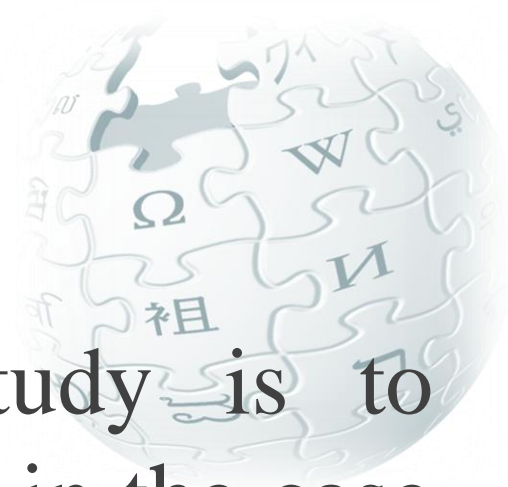
The **aim** of the current study is to develop criteria for evaluation in the case of specific information sources

Wikipedia

- its encyclopedic nature
- it is being applied more and more in the field of education

Wikipedia

- free, online, multilingual encyclopedia
- contains 3,626,934 articles
- has 14,465,628 registered users
- can be edited collaboratively by anyone who has internet access

    → its content is often disputed

"what is contributed is more important than the expertise or qualification of the contributor" - can be read in Wikipedia

This does not mean that the contributions are totally uncontrolled.

There are more and more **tools** inside Wikipedia which **have been created to control the editing process**:

- Five pillars - the principles through which Wikipedia operates

- Policies - demonstrate standards that all users should follow

- Guidelines - contain practices for following those policies

The **success of this quality control mechanism** has been proved by several studies published in this field:

Stvilia, Smith, Gasser and Twidale:

- most of contributors take the issue of quality very seriously which results in continuing development in the collaborative work, therefore also in the quality improvement process

Denise, Smith and Williamson:

- the registered users edit more often than the unregistered

- the more often a user contributes, the higher reliability he/she has

- anonym unregistered volunteers contribute on a higher level, than those registered

<u>Wilkinson and Huberman</u>:

- strong correlation between the number of edits, number of distinct contributors and the article quality

<u>Voss</u>:

- the more people read an article the more errors are amended - the popularity is important factor in the quality of content

CREATING CRITERIA SYSTEM FOR RELIABILITY EVALUATION OF WIKIPEDIA

Evaluation criteria of printed documents - starting point to determine criteria for reliability evaluation of Wikipedia

- Patrick Wilson's cognitive authority theory

- Robert S. Taylor's quality model

Patrick Wilson's cognitive authority theory

Criteria	Factors
cognitive authority (=reliability)	authorship
	publishing data
	document type
	intrinsic plausability

Robert S. Taylor's quality model

Categories	Criteria
ease of use	
noise reduction	
quality	accuracy
	comprehensiveness
	currency
	reliability
	validity
adaptability	
time-consuming	
cost-consuming	

Criteria system for evaluation of Wikipedia

Primary Criteria	Secondary Criteria	Tertiary Criteria
quality	*accuracy*	
	comprehensiveness	
	currency	
	reliability	author
		content
		entry qualifications
		verifiability
	validity	

As the four criteria are of uneven significance, they were <u>weighing</u> with the method of direct estimation:

Main criteria	Secondary criteria	Scores to be given
author (2)	qualification, accomplishment	0-negative 1-neutral 2-administrator
verifiability (2)	references, other bibliographical data	0-not satisfactory 1-satisfactory
entry qualification (1)		0-negative 1-neutral 2-featured
content (1)		0-not satisfactory 1-satisfactory

EVALUATION AND ITS METHOD

- <u>Examination I</u>. conventional, content-based evaluation
- <u>Examination II</u>. evaluation based on the specific criteria system
- <u>Examination III</u>. comparison based on the results of earlier researches

Samples

SAMPLE I. ten articles from a special field of history (Greek stepmother mythology)

SAMPLE II. ten randomly selected featured articles from the field of history

Results – Examination I.

- basic information can be found in Wikipedia articles
- Wikipedia articles usually contain more complementary information

----

- information to be found in Sample I. was judged as reliable

Results – Examination II.

- reputation, acknowledgement and qualification of authors can usually be determined as neutral
- verifiability is often unsatisfactory

- only <u>three articles </u>out of the ten can be qualified as reliable in Wikipedia
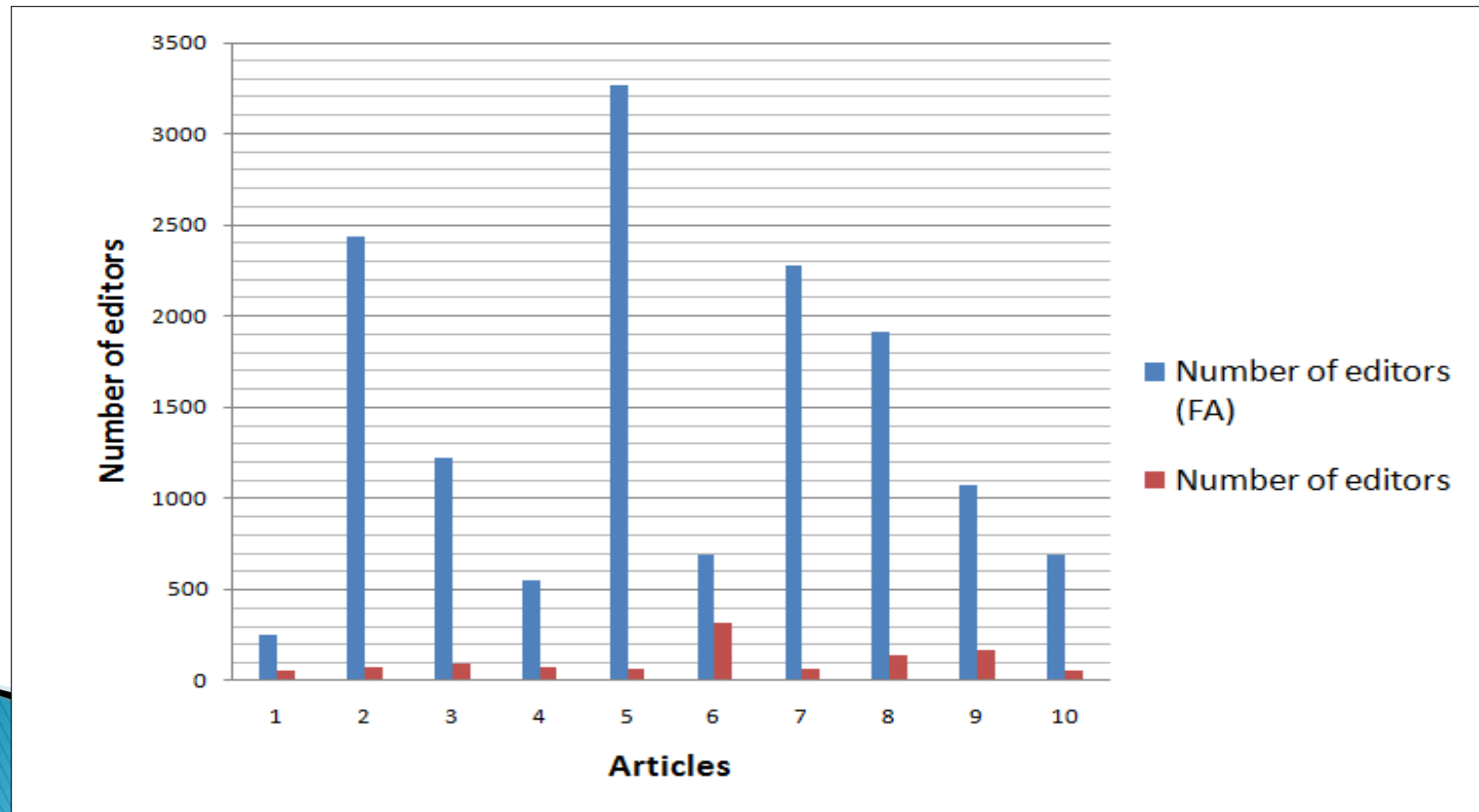
Results – Examination III.

- number of total revisions are higher in SAMPLE II. than in SAMPLE I.
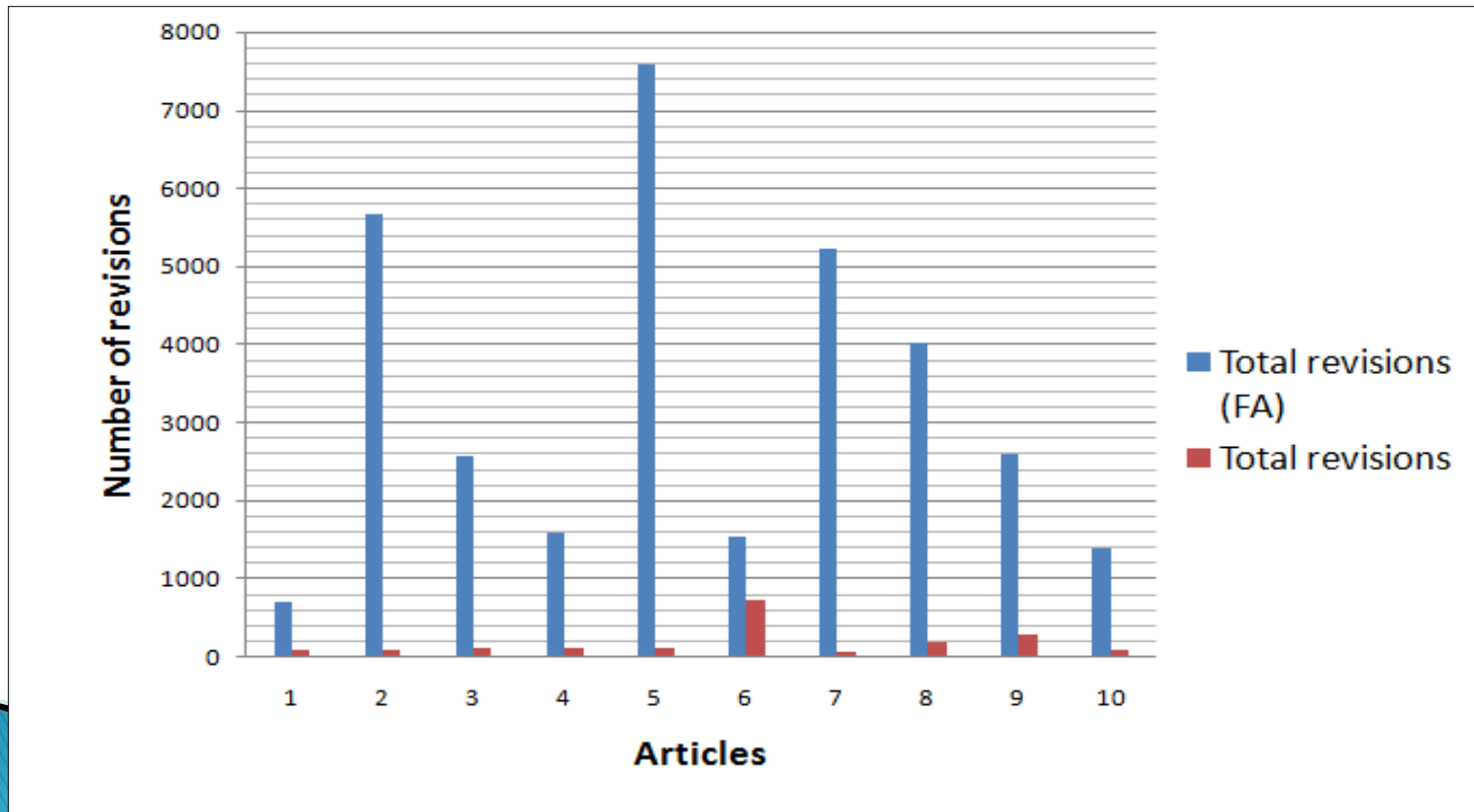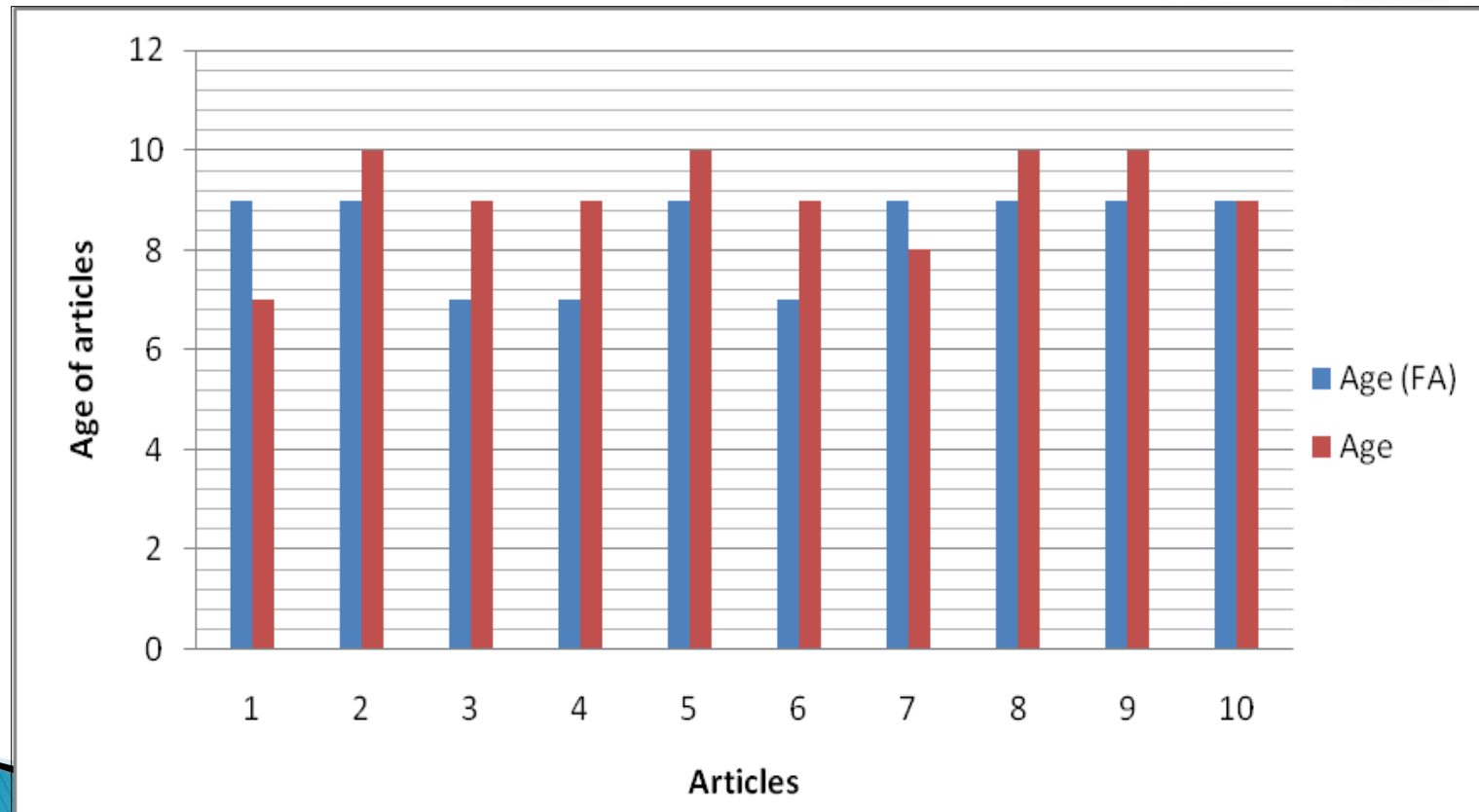
Results – Examination III.

- number of editors and are higher in SAMPLE II. than in SAMPLE I.

Results – Examination III.

- the age of articles are nearly equal

Results – Examination III.

▪ This means that according to the quality policy of Wikipedia the articles of SAMPLE I. have less authority than the articles of SAMPLE II., i.e. they are really less reliable.

THANK YOU FOR YOUR ATTENTION!