

Wikipedia and the Problem of Authorship. Aaron Swartz's Hypothesis

Krzysztof Gajewski
Institute of Literary Research
Polish Academy of Science

Oradea/Nagyvárad, 7-8 October 2016

Table of Contents

The Gang of 500 vs. The Anonymous Horde

Alternatives: Persistent Word View and Persistent Word Revision

Verification of Swartz's thesis and Problems Arising

- The Double Face of Wiki Text

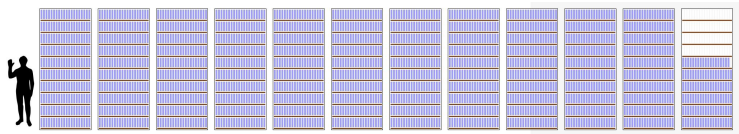
- Concise as a Wikipedia Article

- The Authors

- Copied Content Problem

Conclusions. Ritual vs. Conflict View of Communication

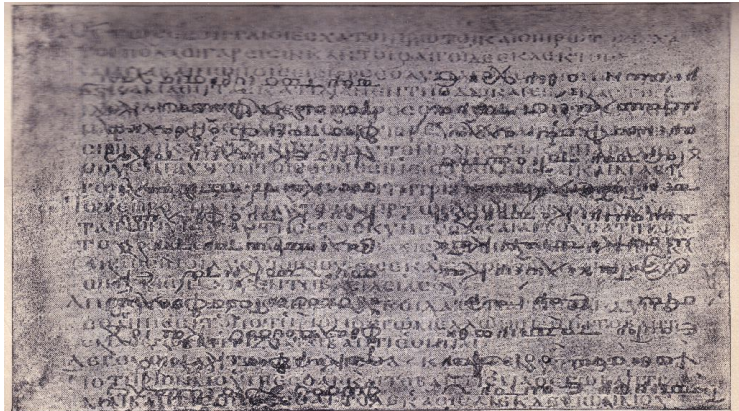
English Wikipedia Size in Volumes



Wikipedia pure text version in print (multimedia excluded)

- 12 stacks
- 2319 volumes (Britannica size)
- 5.2 million articles (as of July 2016)
- 3.068 billion words (as of July 2016)

Wikitext as a palimpsest



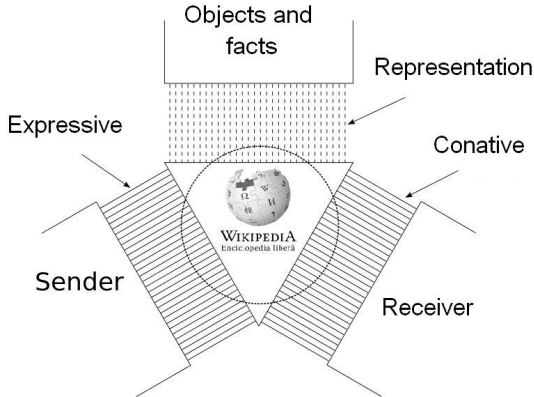
„you cannot change the content of Wikipedia, you can only supply a new one. ”

Wikitext as a palimpsest

- (cur | prev) ○ 15:04, 6 April 2016 Ian Rose (talk | contribs) ... (41,576 bytes) (-131) ... *(Revert to the revision prior to revision 713917375 dated 20*
- (cur | prev) ○ 14:58, 6 April 2016 ScratchClubYoutube (talk | contribs) ... (41,707 bytes) (+131) ... (undo | thank) *(Tag: Visual edit)*
- (cur | prev) ○ 08:18, 25 March 2016 DrKay (talk | contribs) ... (41,576 bytes) (-15) ... *(repetition of year, the "four" is the dated info, not the individ*
- (cur | prev) ○ 23:12, 16 March 2016 FACBot (talk | contribs) ... (41,591 bytes) (+5) ... *(Promoting 'Thunder (mascot)' to Featured Article status) (ui*
- (cur | prev) ○ 08:58, 16 March 2016 Ian Rose (talk | contribs) m ... (41,586 bytes) (-3) ... *(Does that arrow serve a purpose? Not seen it in other ai*
- (cur | prev) ○ 08:56, 16 March 2016 Ian Rose (talk | contribs) m ... (41,589 bytes) (-34) ... *(Tweak/trim some wording, duplink, 2x whitespace) (unc*
- (cur | prev) ○ 03:30, 16 March 2016 Montanabw (talk | contribs) ... (41,623 bytes) (-28) ... *(removed Category:Individual horses using HotCat mas*
- (cur | prev) ○ 22:10, 12 March 2016 Montanabw (talk | contribs) ... (41,651 bytes) (-8) ... *(Not precisely "desensitized", more like "trusts his rider th*
- (cur | prev) ○ 00:24, 12 March 2016 Atsme (talk | contribs) ... (41,659 bytes) (-1) ... *(remove stray parenthetical close) (undo | thank)*
- (cur | prev) ○ 00:23, 12 March 2016 Atsme (talk | contribs) ... (41,660 bytes) (+26) ... *(fix captions) (undo | thank)*
- (cur | prev) ○ 15:58, 9 March 2016 Atsme (talk | contribs) ... (41,634 bytes) (-10) ... *(→Super Bowl appearances: trim excess words) (undo | thank;*
- (cur | prev) ○ 07:49, 9 March 2016 Montanabw (talk | contribs) ... (41,644 bytes) (-93) ... *(tighten prose per Biodef) (undo | thank)*
- (cur | prev) ○ 07:34, 9 March 2016 Montanabw (talk | contribs) ... (41,737 bytes) (-57) ... *(Ce per Biodef) (undo | thank)*

Revision history of a Wikipedia article

Model of Communication



- Receiver = we
- Sender = ?

*„the free encyclopedia that **anyone** can edit.”*

The Gang of 500 vs. The Anonymous Horde

Jimmy Wales: The Gang of 500

Jimmy Wales:

- 73.4% of all the edits — 2% of the users (1400 people)
- 50% of all the edits — 0.7% of the users (524 people)¹

¹Swartz 2006

What Is an Edit?

Two types of edits:

1. uploading textual content
2. „wikisation” of a content already uploaded

Top 10 contributors to „Alan Alda” entry

1. by edits: 7 registered vs. 3 anonymous
2. by letters added: 8 anonymous vs. 2 registered

Aaron Swartz: The Anonymous Horde

Apparent exceptions

1. translations
2. plagiarism

Alternatives: Persistent Word View and Persistent Word Revision

Persistent Word View (PWV) is based on

1. number of letters she input
2. the popularity of the content²

Top 10% most active editors generated 86% of Persistent Word View (Feb. 2006)

²Priedhorsky 2007

Persistent Word Revision (PWR) is

*The sum total of subsequent revisions persisted by the words in a revision.*³

³Research:Content persistence

Persistent Word View vs. Persistent Word Revision

- PWV stresses a role of a reader
- PWR — editors decide of the value of the content

Verification of Swartz's thesis and Problems Arising

Research Sample

30 random entries from Polish Wikipedia („Random article”):

1. Powstanie Kantonalistów
2. Gmina Czarnylas
3. Aurora (telenowela)
4. Podróż na Tajemniczą Wyspę
5. Oberwiera
6. Park Narodowy Ałtaj-Tawanbogd
7. Kościół św. Mikołaja w Wilnie
8. Klaffer am Hochficht
9. Anthaxia attenuata
10. Jerzy Panek (polityk)
11. Wojna z terroryzmem
12. FSO Polonez Analog
13. Strange Frontier
14. Praszywe (Dolina Łatana)
15. Chelmsley Wood
16. Droga krajowa B5 (Niemcy)
17. San Roque (Mariany Północne)
18. Krupy (powiat sokołowski)
19. Alex MacDowall
20. Scottish Premier League (2002/2003)
21. Bitwa pod Olszanicą
22. Rodrigo Oliveira de Bittencourt
23. Hrabstwo White (Tennessee)
24. Bradford (Ohio)
25. Most Królowej Jadwigi w Poznaniu
26. Walenty Forys
27. (Get A) Grip (On Yourself)
28. Andrzej Ekiert
29. Droga krajowa nr 471 (Węgry)
30. Wyganki

The Double Face of Wiki Text

Output Version of Wikipedia Article

Article [Talk](#)

Thunder (mascot)

From Wikipedia, the free encyclopedia

Thunder is the [stage name](#) for the [horse](#) who is the official live animal [mascot](#) for the [Denver Broncos](#) football team. Three purebred [horses](#) whose coats lightened with age until they turned completely white. [Sharon Magness-Blake](#) has owned all three horses, and trainer since 1999. As of 2016, Thunder has appeared in four [Super Bowls](#) with the team since 1998. The original Thunder performed at Super Bowl III. Thunder III appeared in [Super Bowl XLVIII](#) and [Super Bowl 50](#). Thunder III also made appearances in [Times Square](#) and on television during the pre-game promotion for Super Bowl XLVII. He routinely appears in parades, makes hospital and school visits, and attends various airplanes, ridden in elevators, and appeared indoors at press conferences and banquets.

Thunder's duties as mascot typically include leading the team onto the field at the start of every [home game](#), and a [gallop](#) down the field after a [touchdown](#) or [field goal](#). Thunder and his rider also interact with fans before the game; the horse is particularly popular with children. Thunder served as Thunder need to remain calm in situations that would normally frighten most horses, such as being in football stadiums during [pyrotechnics](#), cheerleaders waving [pom-poms](#), and other spectacles common to [National Football League](#) (NFL) games. Thunder also wears a [horse head mask](#) atop a Broncos uniform.

The original Thunder, later named "Thunder, Sr.", was described as bold and courageous. He was a [stallion](#) registered as JB Kob the Broncos from 1993 until his retirement in 2004. He continued making community appearances until his death in 2009. Thunder was a [gelding](#) registered as Winter Solstice. He had been the personal [pleasure riding](#) horse of Magness-Blake. Judge described his role as mascot but eventually grew into it. He retired from mascot duties in early 2014 but appeared in the 2016 Super Bowl 50 Broncos win over the [Carolina Panthers](#). "Thunder III", a gelding registered as Me N Myshadow, was the [understudy](#) to Thunder II, at age three when he was [started under saddle](#). He began performing at preseason games in 2013. Although Thunder II was still Thunder III was sent to Super Bowl XLVIII because he was younger and better able to handle air travel. He is described as laid-back and performing.

Contents [\[hide\]](#)

[1](#) History

Front-end: Output Version of Wikipedia Article (OVA)

Input Version of Wikipedia Article

'''Thunder''' is the [[stage name]] for the [[horse]] who is the official live animal [[mascot]] for the [[Denver Broncos]] [[Arabian horse|Arabians]] have held this role since 1993, all [[gray (horse)|gray horses]] whose coats lightened with age. Blake|Sharon Magness-Blake]] has owned all three horses, and Ann Judge has been their rider since 1998 and trainer since 1999. [[Super Bowl]]s with the team since 1998. The original Thunder performed in [[Super Bowl XXXII]] and [[Super Bowl XXXIII]] and [[Super Bowl 50]]. Thunder III also made appearances in [[Times Square]] and on television morning news shows in [[New York City|New York]]. Thunder IV routinely appears in parades, makes hospital and school visits, and attends various other public functions. Thunder V appeared indoors at press conferences and banquets.

Thunder's duties as mascot typically include leading the team onto the field at the start of every [[Home (sports)|home game field whenever the team scores a [[touchdown]] or [[field goal (American and Canadian football)|field goal]]. Thunder and the horse is particularly popular with children, who are allowed to pet him. The horses who have served as Thunder need to frighten most horses, such as being in football stadiums with thousands of cheering fans, exploding [[pyrotechnics]], cheer, compete to [[National Football League]] (NFL) games. Thunder shares mascot duties with [[Miles (mascot)|Miles]], a human who

The original Thunder, later named "Thunder, Sr.", was described as bold and courageous. He was a [[station]] [[bred regi|show|show horse]], who was team mascot for the Broncos from 1993 until his retirement in 2004. He continued making commu|Sr. was succeeded in 2004 by "Thunder II", an Arabian [[gelding]] registered as Winter Solstice. He had been the personal described him as being somewhat timid when he first began his role as mascot but eventually grew into it. He retired from : Super Bowl 50 victory parade in downtown [[Denver, Colorado|Denver]] following the Broncos win over the [[Carolina Panthers|Myshadow, who was the [[understudy]] to Thunder II, and trained specifically for mascot duties beginning at age three when he w|performing at preseason games in 2013. Although Thunder II was still active as team mascot during the 2013-14 season, Thun|younger and better able to handle air travel. He is described as laid-back, preferring to doze off during games when not be

```
==History==
```

[[File:ThunderIII TD509.ogg|thumb|upright=1.3|left|Video clip of Thunder III (He N Myshadow) making a run in 2014 after [[P

The original "Thunder", JB Kobask, was loaned to the Denver Broncos by Sharon Magness-Blake of Magness Arabians after she rode had a "white" horse. <ref name="NWork"/><ref name="Parkinson"/> (efn|Gray horses have a hair coat that depicts with age "flea-bitten"; white with small darker spots.<ref name="Animal Genetics"/>)<cite web|title=Gray Coat Color / Melanoma|url=http://Gray.asp|publisher=Animal Genetics Inc|accessdate=February 20, 2016|year=2014|></ref> Thunder's debut as team mascot was over the [[San Diego Chargers]].<ref name="Pomeroy"> Angela Moore, his rider at the time,<ref name="CSU"/> galloped him down name="Pomeroy"/> Thunder's mascot duties were expanded to include delivering the [[Football (ball)|American and Canadian football|refere]]s at the start of each home game. Spectators were allowed to interact with Thunder and pet him!<!--R. m Thunder II moved into a stall inside the stadium, don't want confusing detail--> prior to the game.<ref name="Parkinson"/>(<ci Horse: JB Kobask (Thunder)|url=http://www.arabianhorseworld.com/live/library/2012/10/33-kobask/|website=Arabian Horse World </ref> Magness said, "Thunder is friendly, and Broncos fans think it's good luck for the opposing team to pet our mascot."<!-- Wilson|title=Pet Tales: Watch for Broncos' horse on Sunday|url=http://www.post-gazette.com/life/pet-stories/2014/02/01/Watcl /201402010037|work=Pittsburgh Post-Gazette|accessdate=January 31, 2014|date=January 31, 2014|> Elway]] was particularly p T-shirts, and was referred to as the "second most popular Bronco" after former quarterback [[John Elway]]. Thunder appeared name="Parkinson"/> When team uniforms were [[Denver Broncos1997-present|redesigned in 1997]], Thunder was the inspiration to logo on the team's helmets.<ref name="Pomeroy"/>(<cite web|last=Pomeroy|first=L.&L.|title=Thunder: Denver's Favorite Bronco |u

Back-end: Input Version of Wikipedia Article (IVA)

- IVA is easy to determinate, calculated automatically, but includes WikiMarkUp
- OVA is what we actually see, what counts for a reader, but is difficult to precise

Methodological problems

- Wiki MarkuUp Language
- Non-textual content: links, bibliography, tables, sounds, images, graphics, diagrams, schemes etc.
- Styles: headings, lists, tables etc.
- Tranclusion (templates)
- Photos and other multimedia (WikiCommons)

Concise as a Wikipedia Article

The size of articles

The size of articles

developed	more than 10 lines	6	20%
somehow developed	4-10 lines	6	40%
undeveloped	less than 4 lines	12	40%

- only 13% of articles had more than 50 revisions.
- average OVA in the research sample: 1961 chars
- average IVA in the research sample: 3960 chars
- average IVA in the Polish Wikipedia: 2718 chars (Feb., 2014)

The Authors

Presumptions

- In most of the cases articles — up to 2 authors.
- The First Author — the biggest contribution in terms of OVA
- The Second Author — the second biggest contribution in terms of OVA

The First Author

- 11 of 30 articles — only one human author (9 of 11 — a bot contribution)
- The whole group of the First Authors consists of 29 members.
- A Wikipedist Cojan (ranked as 21st prolific editor in Wikipedia) happened to be the First Author of two random articles.
- 2 of 29 First Authors happened to be a bot: Tsca.bot and MalarzBOT.
- **Only 2 of 29 First Authors were anonymous**

The Second Author

- 8 of 30 articles — the Second Author added some facts or notes, 22 of 30 — redaction & wikisation only
- 2 of 8 factual intervention was made by anonymous, 6 of 8 by registered users.
- A Wikipedist Lowdown (one of the most prolific editors in Wikipedia) happened to be the Second Author of 2 random articles.

Authors

The characteristics of 1st and 2nd authors of articles

	all	registered	anonymous	bot
1st authors	29	25 (86%)	2 (7%)	2 (7%)
2nd authors	21	14 (67%)	4 (19%)	3 (14%)

Copied Content Problem

A Case of „Wojna z terroryzmem” („War on terrorism”)

- At 10:34 am, on 5 August 2005, an anonymous user A added a sentence:

„Such understood „war on terrorism” caused 25 000 civilian deaths”

- At 6:01pm, on 29 August 2005 another anonymous user B removed this sentence and wrote instead a new one:

„They [opponents of USA policy] also point out the numerous civilian casualties (several thousand) due to military operations.”

- Result?
 - The user B was granted the authorship of the sentence.
 - The user A hasn't been counted as an author of the final version of the entry at all.

Conclusions. Ritual vs. Conflict View of Communication

OVA Contribution

	chars	%
registered	40940	90,78%
anonymous	3353	7,43%
bots	806	1,79%

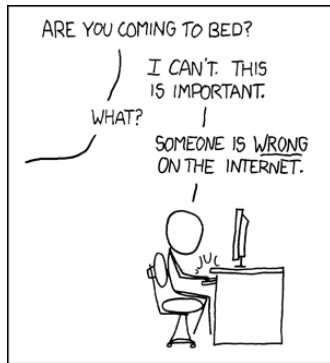
Pareto Principle vs. Long Tail

- Gang of 500 Theory → Pareto principle
- Anonymous Horde → Long Tail (Anderson 2006)

*„In a ritual definition, communication is linked to terms such as 'sharing,' 'participation,' 'association,' 'fellowship,' and 'the possession of a common faith.' This definition exploits the ancient identity and common roots of the terms 'commonness,' 'communion,' 'community,' and 'communication.' A ritual view of communication is directed not toward the extension of messages in space, but toward **the maintenance of society in time; not the act of imparting information, but the representation of shared beliefs.**”⁴*

⁴Carey 2009: 15

Conflict View of Communication



Conflict View of Communication⁵

⁵Conf. Jemielniak 2006: 124

References

1. Anderson Ch. 2006, The Long Tail: Why the Future of Business Is Selling Less of More. New York: Hyperion.
2. Carey J. W. 2009. Communication as Culture: Essays on Media and Society. New York: Routledge.
3. Jemielniak, D. 2013. Życie wirtualnych dzikich Netnografia Wikipedii, największego projektu współtworzonego przez ludzi, Warsaw: Poltext.
4. Swartz, A. 2006. Who Writes Wikipedia. (n.d.)
<http://www.aaronsw.com/weblog/whowriteswikipedia>
5. Swartz, A. nd. "Who Writes Wikipedia?" (Swartz 2006),
<http://www.aaronsw.com/2002/whowriteswikipedia/swartz2006>