Why a list of comets is one of Wikipedia's longest pages

March 10 2016

Keep scrollin', scrollin', scrollin', scrollin'

The top 10 longest pages on Wikipedia

---

1. List of law clerks of the Supreme Court of the United States

2. List of United States counties and county equivalents

3. List of compositions by Johann Sebastian Bach

4. Opinion polling for the 2015 United Kingdom general election

5. 1919 Birthday Honours

6. 1918 Birthday Honours

7. List of comets by type

8. List of Dutch inventions and discoveries

9. List of Amtrak stations

10. Constituency election results in the United Kingdom general election, 1929

Correct as of March 9 2016

Why are some pages on Wikipedia so much longer than others?

A new study has found it is not only how famous or popular the subject is, but more if it generates what the researchers call 'cumulative growth effect'.

Out of the 38 million pages on the website the [longest on the English Wikipedia](#) is a 'List of law clerks of the Supreme Court of the United States', while '1919 Birthday Honours' and 'Lists of Comets by type' are also in the top 10.

The biggest page that isn't a list is 'Opinion polling for the 2015 United Kingdom general election', at number three according to Wikipedia, and 'Constituency election results in the United Kingdom general election, 1929' at 10.

None of these subjects are likely to capture the general public's imagination, but new research has found the way Wikipedia works produces a snowball of editors onto a page – known as a 'cumulative growth effect'.

The [open source platform](#) relies on a host of editors – or Wikipedians – to contribute, organise and edit the site and researchers found that articles tend to snowball, with the number of contributions to an article increasing as it gets longer and longer.

Without this cumulative growth effect, articles would have been up to 45 per cent shorter according to Aleksi Aaltonen, of Warwick Business School.

Dr Aaltonen said: "We found the way in which Wikipedia's [content](#) production is organised is inherently motivating for users. The gradual nature of content development encourages and inspires users and editors - it snowballs.

"This motivational mechanism may emerge from the fact that once the content is there subsequent users are able to build on it bit by bit and see the results immediately rather than having to contribute an entire article.

"As a consequence, articles that are edited heavily and therefore grow in length will continue to be edited more."

In the article Cumulative Growth in User-Generated Content Production, published in *Management Science*, Dr Aaltonen and Stephan Seiler, of Stanford University, used eight years of data looking at articles on the Roman Empire on the English Wikipedia.

The data set contains the full text of every version of all articles from the beginning of the online encyclopaedia in January 2001 to January 2010. This allowed the researchers to track the evolution of the content of each article on the Roman Empire.

"Looking at this data set, we found surprisingly strong evidence for a cumulative growth effect," said Dr Aaltonen. "More specifically, when we controlled for the popularity of individual articles and the overall editing activity at different times as well as for a number of other factors, we still found that the current length of the article has a positive impact on the amount of editing it receives."

The Roman Empire section, which comprises 1,310 unique articles, was chosen by the researchers as the knowledge on the topic could be assumed to undergo relatively little change during the sample period. This helped to control for possible new information on the subject coming to light that could otherwise affect the evolution of the content.

Dr Aaltonen, who teaches Business Systems Analysis and Strategic Information Management on Warwick Business School's Undergraduate programme, said: "The main lesson we can draw from the findings is

that any action that increases content can trigger further contributions.

"Two ways to achieve an increase in content and grow the length of a page is to incentivise users to contribute content or, even more directly, to pre-populate articles with content. These interventions can lead to a magnified effect as more editors are attracted to the content and as they contribute so it snowballs.

"Because many organisations such as Sony, Xerox, Disney, Microsoft and Intel harness wikis—server programs that allows users to collaborate in forming the content of a website—often using the same software and a similar page layout, these findings are likely to carry over to those related platforms.

"Importantly, we also found the additional activity induced by the cumulative growth effect leads to an increase in content quality.

"Our findings are therefore relevant for the design of other open content production platforms as well. Importantly, the platform provider has some degree of control over the source of the motivational mechanism we identify - that is the amount of content."

Provided by University of Warwick