Which parts of an article are actually being read?

Tilman Bayer
Wikimania 2018, Cape Town, July 22, 2018
Beyond pageviews

What knowledge do people access on Wikimedia projects, and how often? Pageview numbers have been our main source for answering this question. However, they only tell us how often a page is opened - not how much of it is being read.
This talk

- Tour through some new data sources that became available in recent years
- These come from academic research or from data WMF collected during work on software features (products). But they yield insights for editors too.
- Myself: Analyst in the WMF Audiences (product) department, Wikipedia editor since 2003
Data on scroll actions

On the Android Wikipedia app, 

Around 32% of pageviews do not involve scrolling down - i.e. reader looks only at the top of the page.
Scrolling data

On the Android Wikipedia app,

- The bottom of the page is only reached on about 5% of pageviews
- Only 25% of users ever reached the end of any page (during one week)
Clickstream data
Public dataset by WMF, counting how often wikilinks are followed from one page to the other.

Likewise indicates that the top of an article is read much more often that its end.

Clickthrough rate of wikilinks vs. their position on the source page, 0=top, 1=bottom (from Paranjabe et al.)
Pages may not be read linearly.

Here, the reader moved through several Wikipedia articles, jumping up and down, while looking for a specific fact about diabetes. (Mechanical Turk experiment by Halfaker et al., 2013)

Grey area = position of the view pane, i.e. the visible area of the screen.
Eyetracking studies

Expensive lab equipment allows to find out which part of the screen a user looks at.

WMF has never used this so far, but several academic researchers (from Germany and Scotland) have applied it to (desktop) Wikipedia.
Figure 1. Gaze plots example from Mark Zuckerberg article. The dots are fixations, larger dots mean greater fixation durations and the lines between the fixations are saccades.
Insights from eyetracking studies

- Most readers look first at TOC, then infobox.
- In "lookup" tasks, >45% of time is spent scanning TOC and lists.
- In "learn" tasks, it's <10%.
- Images provide visual orientation, too:
  - "39% of all [eye] contact points lead from image to image, in mutual directions (previous or next)."

(sources)
Section expansions

On the mobile web version of Wikipedia, article sections are collapsed until the reader taps the heading.

We instrumented this for an A/B test, and now have some data about how often each section is accessed.

(click for animated screencast)
On 60% of mobile pageviews, not a single section is opened, i.e. the reader only looks at the introduction.
Much more interest for personal aspects (“Early life and career”) than for political work (“Legislative career”, “presidency”)

Almost no-one views “external links”
Some takeways from these and other pages

- Mobile web readers seem very uninterested in the “external links” section.
- “References” is rarely opened as well (but footnotes can also accessed from the text).
- Lower section less popular, but heading/topic seems to matter.
Page previews data

- **New feature** (2017/18) on desktop Wikipedia
- Reader hovers over a link to see an excerpt of the linked data
- Heatmap of hovers shows which parts of the page were read, but also which topics(links) readers are interested in looking up briefly from the source page
Barack Obama

From Wikipedia, the free encyclopedia

Barack Hussein Obama II (born August 4, 1961) is an American politician who served as the 44th President of the United States from 2009 to 2017. The first African American to assume the presidency, he was previously the junior United States Senator from Illinois from 2004 to 2008. Obama was born in 1961 in Honolulu, Hawaii, two years after the territory was admitted to the Union as the 50th state. Raised largely in Hawaii, Obama also spent one year of his childhood in Indonesia. After graduating from Occidental College in Los Angeles, he moved to New York City in 1983, where he worked as a community organizer in Chicago. In 1988 Obama enrolled in Harvard Law School, where he was the first black president of the Harvard Law Review. After graduation, he became a civil rights attorney and professor, and taught constitutional law at the University of Chicago Law School from 1992 to 2004. Obama represented Illinois in the United States Senate from 1997 to 2004.

Obama received national attention in 2004 with his unsuccessful run for the Democratic Party nomination and his landslide November election to the Senate. In 2008, Obama was nominated for president by his campaign against other Democratic primary candidates Hillary Clinton and John Edwards. He was elected over Republican John McCain and was inaugurated on January 20, 2009. Nine months later, Obama was the 2009 Nobel Peace Prize laureate, accepting the award with the caveat that he felt there were others "far more deserving of this honor than me."

During his first two years in office, Obama signed many landmark bills into law. The main reforms were the Affordable Care Act, the Dodd–Frank Wall Street Reform and Consumer Protection Act, and the Don't Ask, Don't Tell Repeal Act of 2010. The President's health care and financial regulations served as economic stimulus and to the Great Recession. After a lengthy debate over the national debt limit, Obama signed the Budget Control Act and the American Taxpayer Relief Act. In foreign policy, Obama increased U.S. troop levels in Afghanistan, reduced nuclear weapons with the United States–Russia New START treaty, and ended military involvement in Iraq. He ordered military intervention in Libya in opposition to Muammar Gaddafi. Gaddafi was killed by NATO-assisted forces, and he also ordered the military operation that resulted in the 2011 attack by the United States and its allies.

After winning re-election by defeating Republican opponent Mitt Romney, Obama signed into law for a second term in 2013. During his second term, Obama promoted awareness for LGBRTQ Americans. His administration faced legal that the Supreme Court's decision to strike down civil marriage bans in California lost. The Court took up the question of gay marriage in the United States v. Windsor. However, according to the 1954 Supreme Court decision in Brown v. Board of Education, the lawsuit was brought by the NAACP, which represented a coalition of civil rights groups. Obama also advocated for gun control in response to the Sandy Hook Elementary School shooting, and issued several executive actions concerning climate change and immigration. In foreign policy, Obama stated...
Thank you!

Questions?

Tbayer at wikimedia dot org

[meta:Research:Which parts of an article do readers read]