Industry Analysis

January 2016

WMF Reading Team

Introduction

1. This deck is a collection of the Industry analysis performed by the Reading team over the last year.
2. It has been used for background to the Reading strategy work and will be also be used for the review of the Reach strategic approaches.

Table of Contents

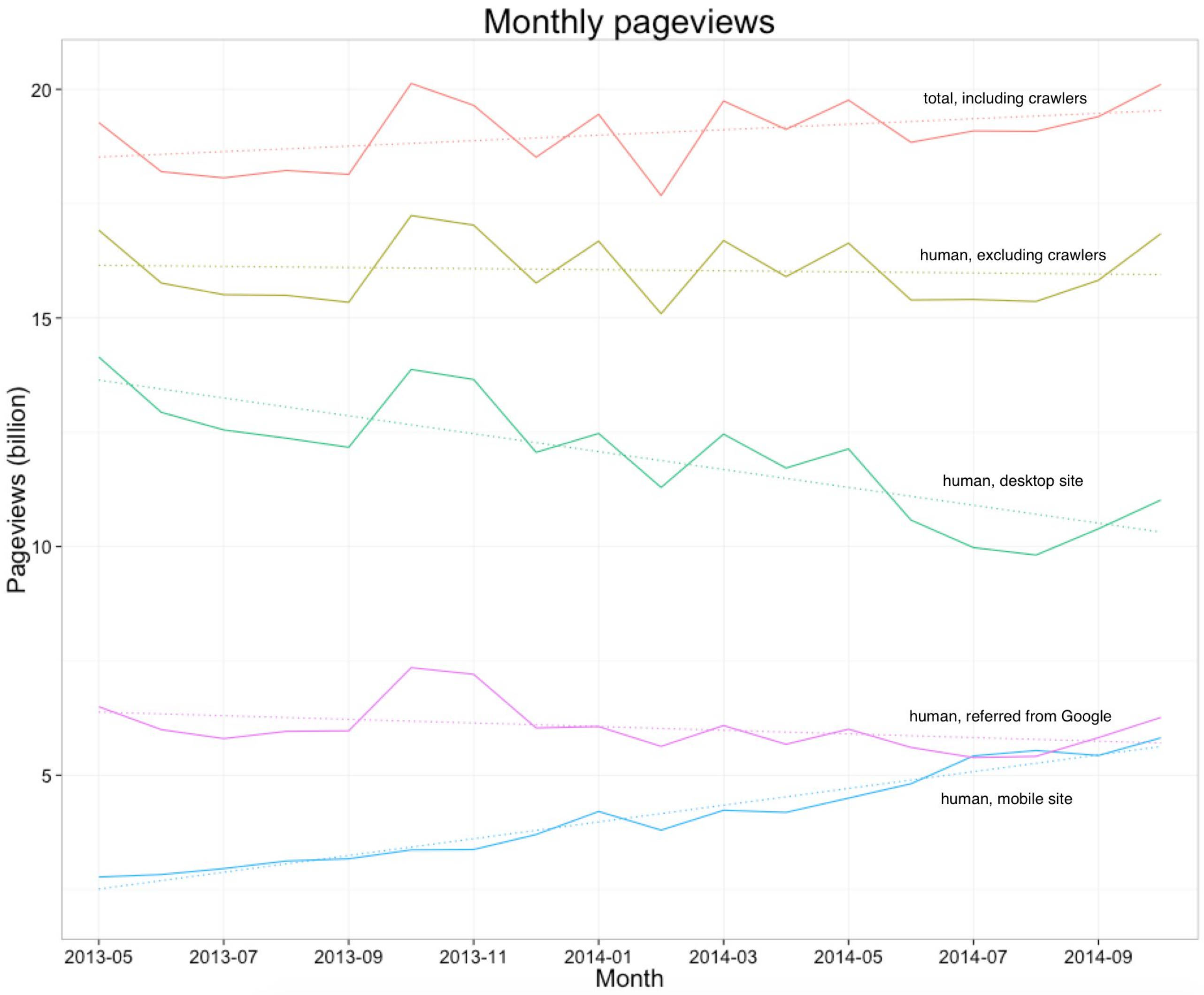
Traffic Reports

1. The last time we did an exhaustive review of traffic was late 2014. We are updating this work -- see posts to mobile-l.
2. Analytics is working on a unique devices metric (coming soon)

World

Globally, total human readership (excluding crawlers and automated traffic) is **flat**

This is the result of a **decline in desktop readership** entirely compensated by **mobile growth**.

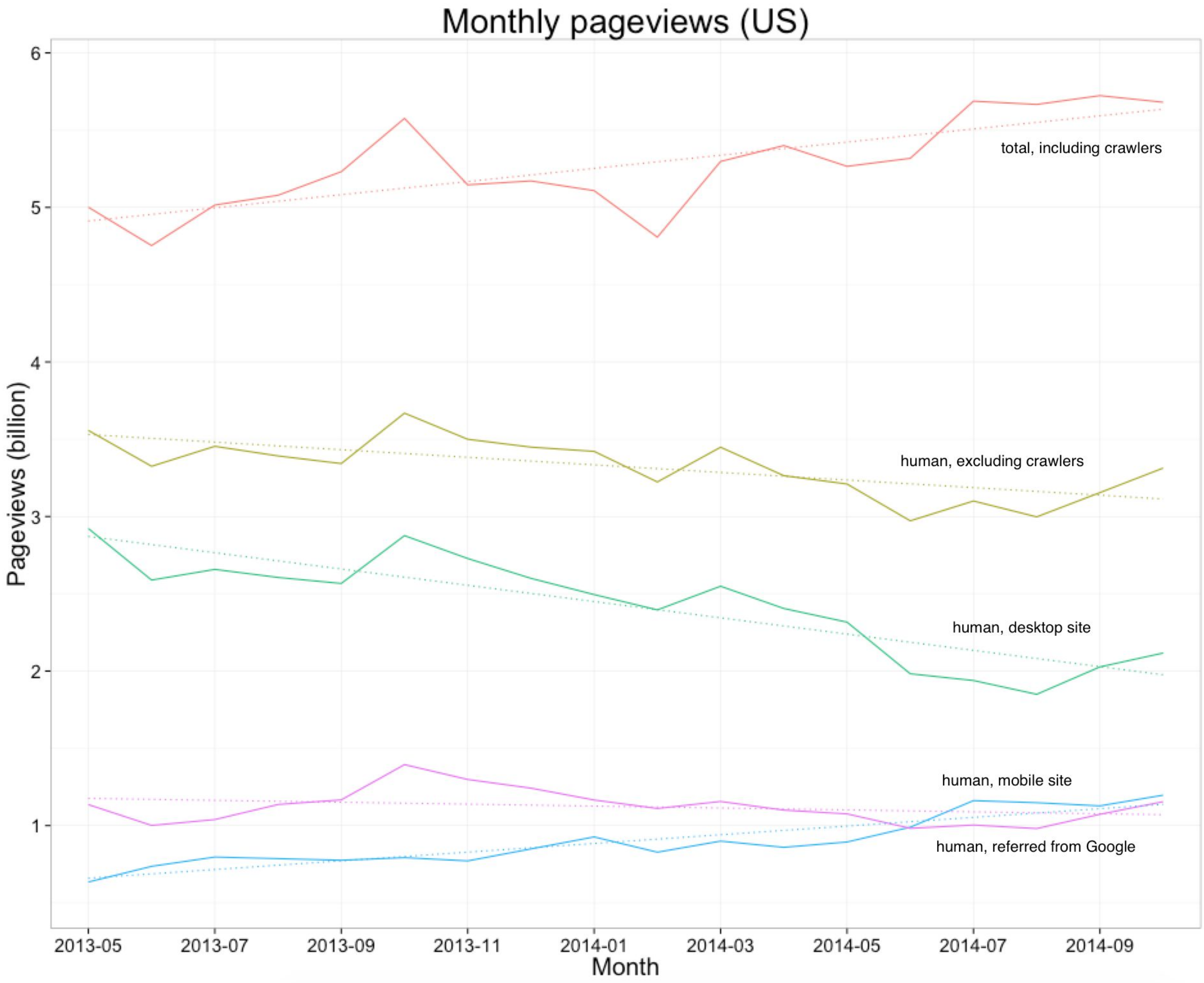


US

The US is the country with the combined **largest readership**.

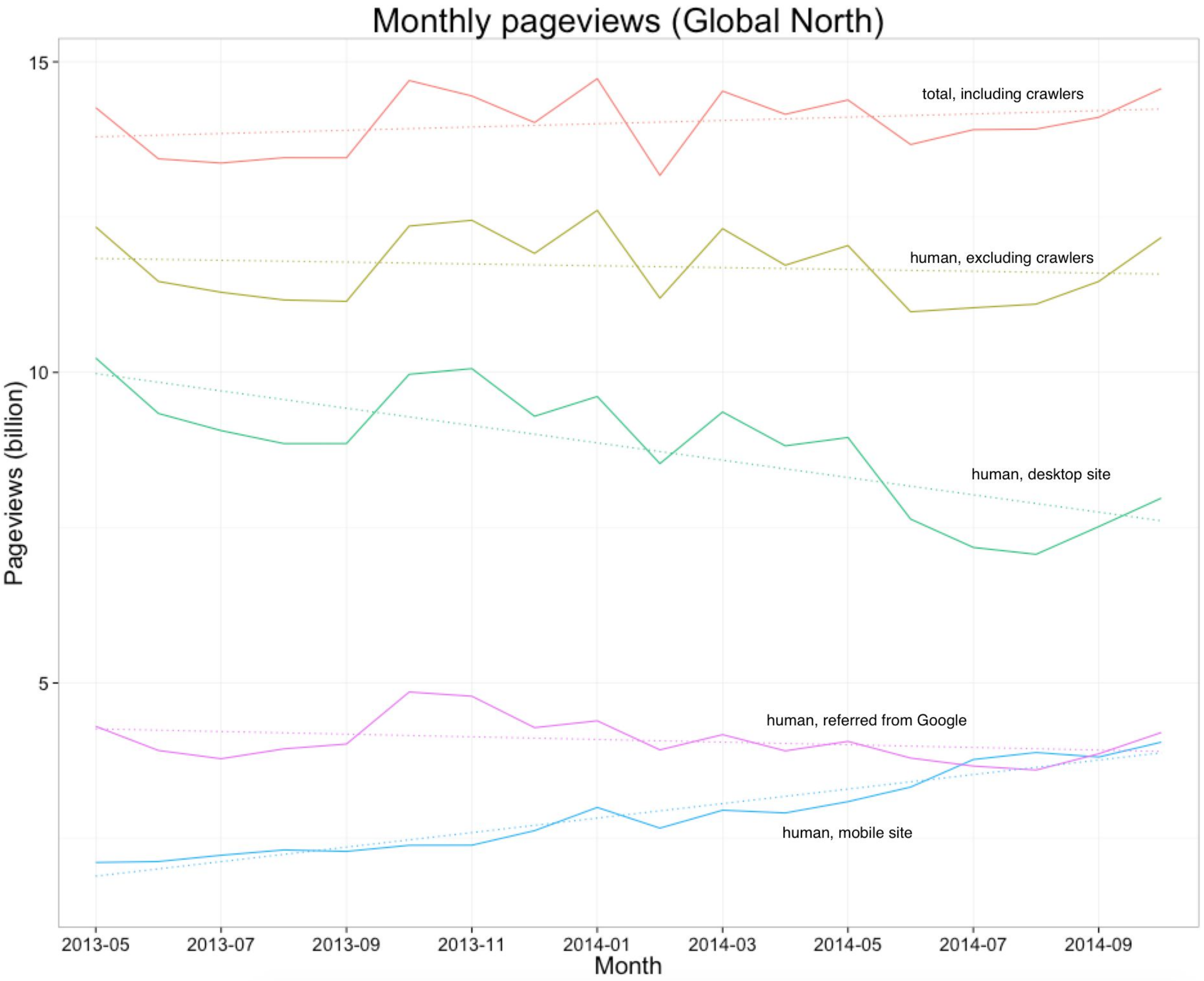
Total human readership from the US is **declining**: **mobile growth in the US is lagging** behind global mobile growth trends and is not offsetting the **decline in desktop traffic**.

The vast majority of **automated traffic** (crawlers) comes from the US.



Global North

The Global North is responsible for **72% of total human readership**, trends are similar to those observed globally.

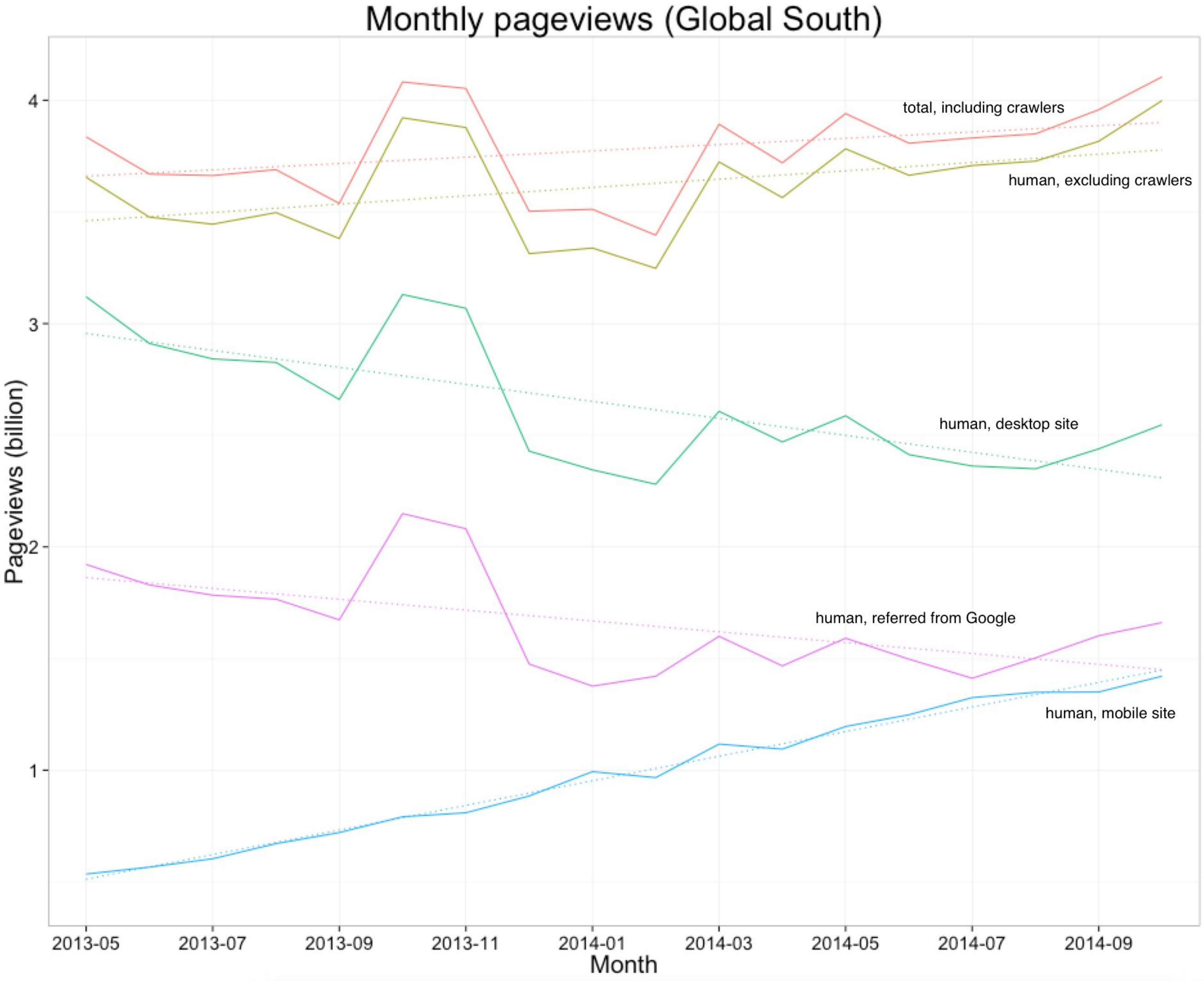


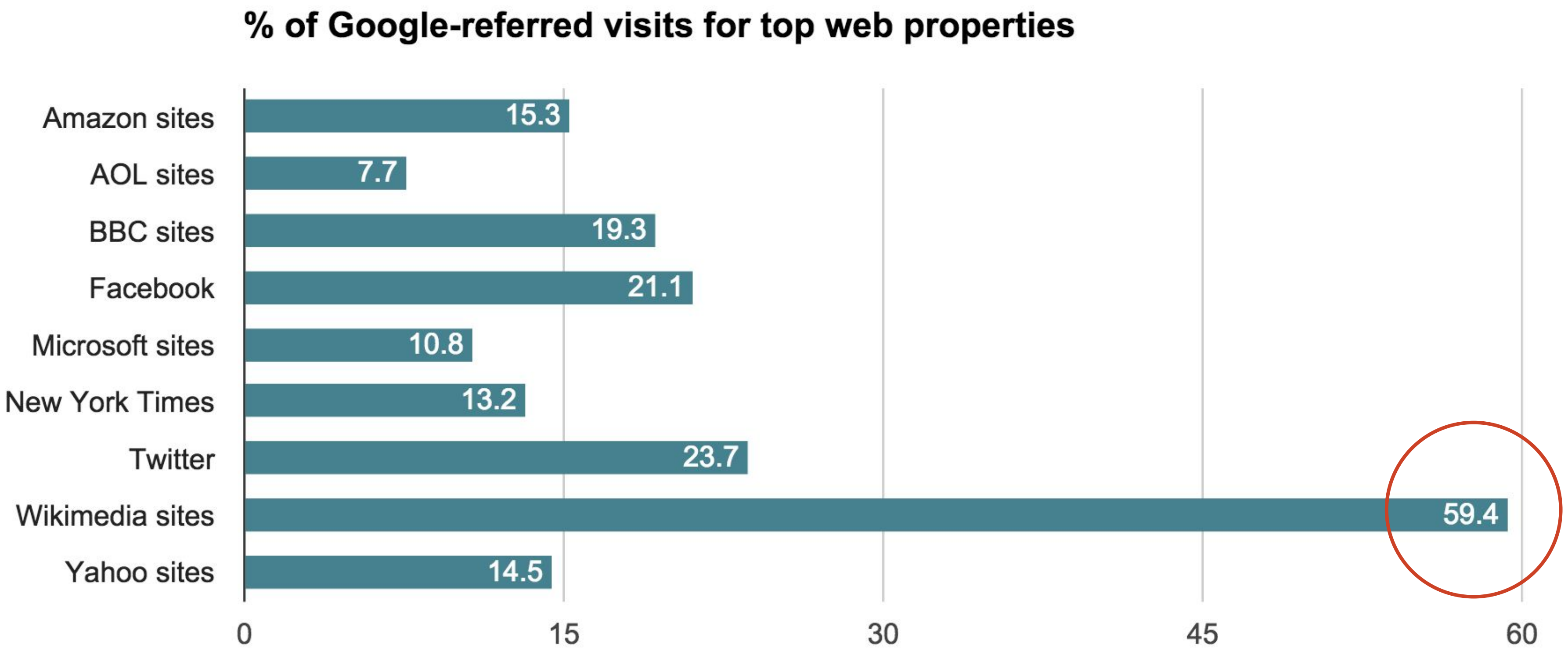
Global South

The Global South is responsible for **24% of total human readership**.

It's seeing the **fastest mobile growth**, although the majority of mobile traffic still comes from the Global North.

Automated traffic from the Global South is a negligible fraction.





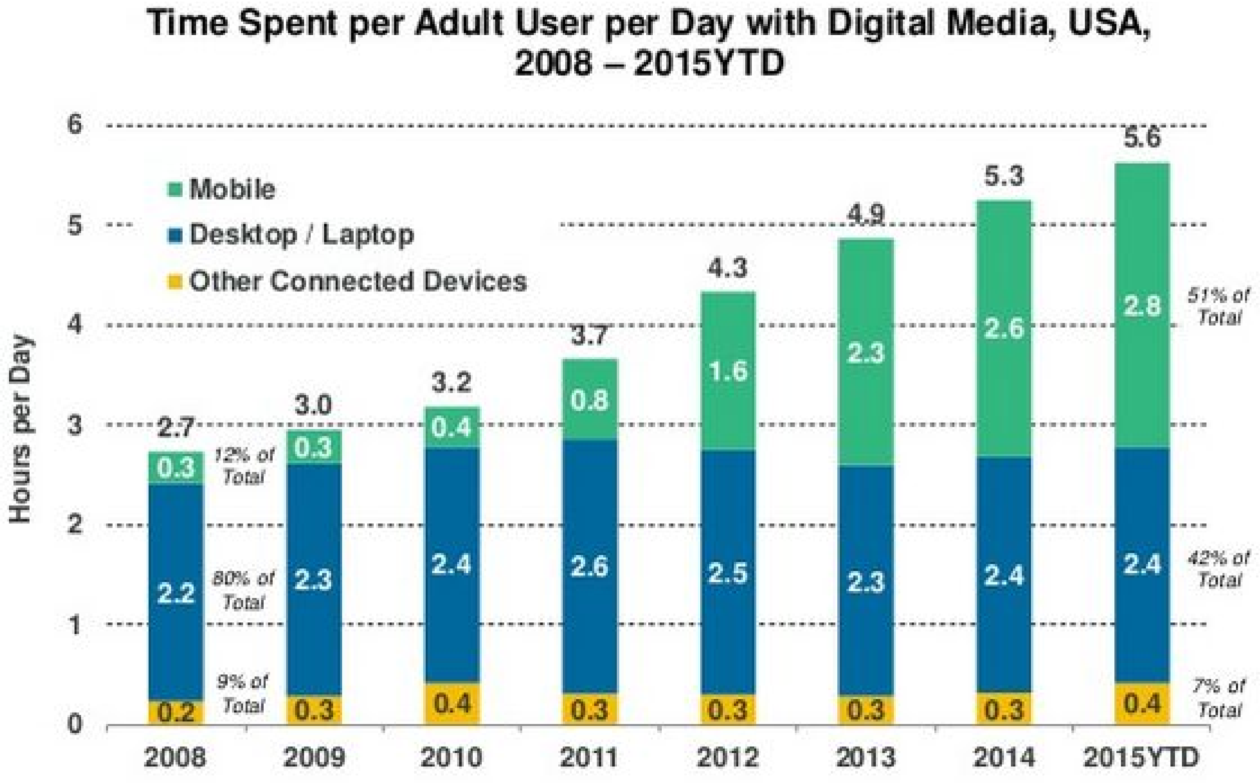
We rely on Google referrals up to **5x more** than any other top web property or media company.

(Source: comScore Source/Loss report July 2015)

2015 US Traffic Trend Updates

- We put together a model for forecasting page views for December 2015
  - A linear regression forecast a 5% drop in page views from the US to enwiki
  - However, there is some nuance around the use of this model and a rise in page views is also a possibility
- Current page views from Readership report indicate that page views are “roughly” flat year over year, but some caveats about comparisons apply.

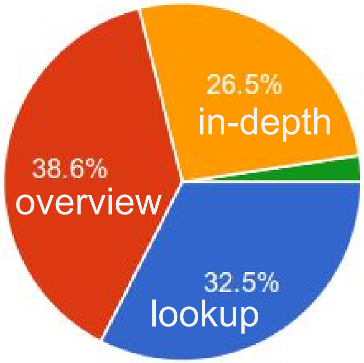
Internet *Usage* (Engagement) Growth Solid
+11% Y/Y = Mobile @ 3 Hours / Day per User vs. <1 Five Years Ago, USA



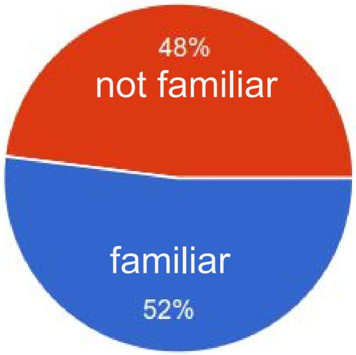
Source: eMarketer 9/14 (2008-2010), eMarketer 4/15 (2011-2015). Note: Other connected devices include OTT and game consoles. Mobile includes smartphone and tablet. Usage includes both home and work. Ages 18+ ; time spent with each medium includes all time spent with that medium, regardless of multitasking.

Preliminary Results from User Segmentation work

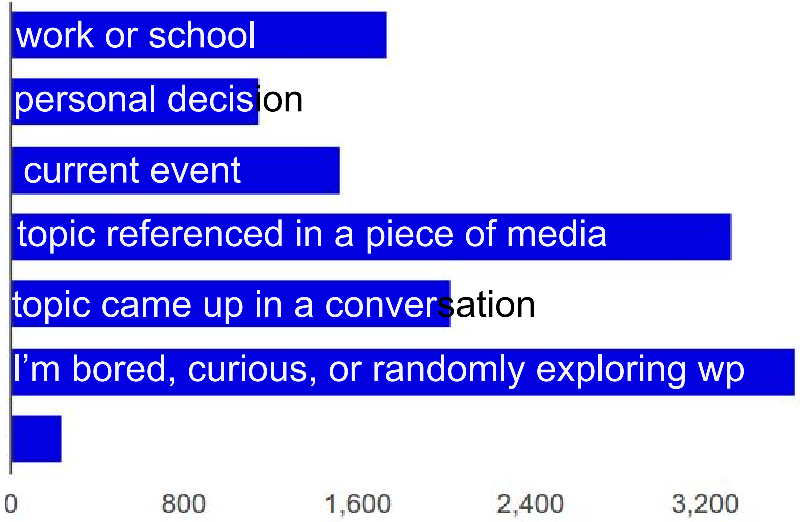
Q1. I'm reading this article to:



Q2. Prior to visiting this article:



Q3. I'm reading this article because:



* Survey data was collected during the Thanksgiving week with the goal of identifying missing labels. We do not recommend reading too much into the distributions from this survey.

Market Segments

1. We need a more granular understanding of our users to be able to target understandings and new features appropriately.
2. We use a 5 dimensional model:
    a. Consumer
    b. Channel
    c. Product or Service
    d. Geography
    e. Stage of Production

Customer

- Age
- Education
- Frequency
- Interests
- Occupation
- Gender

- School Age
- Young People
- Midlife
- Seniors

- Daily Internet Use
- Weekly
- < Weekly

- Men
- Women
- Genderqueer

- Communications
- News
- Social
- Culture
- Bored
- Studying

- Pre-Literate
- Literate
- Primary
- Secondary
- College +

- Student
- Education
- Service
- Manufacturing
- Agriculture

Channel

Platform

- Android
- iOS
- Mobile Web
- Desktop

Connectivity [1]

- Insufficient (none)
- Basic (2G)
- Good (3G)
- Excellent (> 3G)

Distribution

- Preload
- Owned & Operated Sites
- Syndication

[1] https://drive.google.com/a/wikimedia.org/file/d/0B9y-9JVBK7_5Y29LT2pYWm9Lak0/view

Product or Service

Experience

Attributes

Content

Lookup

Browse

Learning

Updated (to current expectations)

Accessible

Interactive

Low-bandwidth

Rich Media

Social

Unchanged

Topical

Evergreen

Niche

Definitions

Definitions

Structured

Long Form

Video

Image

Geography

North America
South America
Europe
Africa
Asia
Australia

Geo

Top 10 Wikis by PVs [1]

English
Spanish
German
Japanese
French
Chinese
Italian
Polish
Portuguese
Dutch

Language

[1] http://stats.wikimedia.org/EN/TablesPageViewsMonthlyCombined.htm

Stages of Production

Editing

Presentation

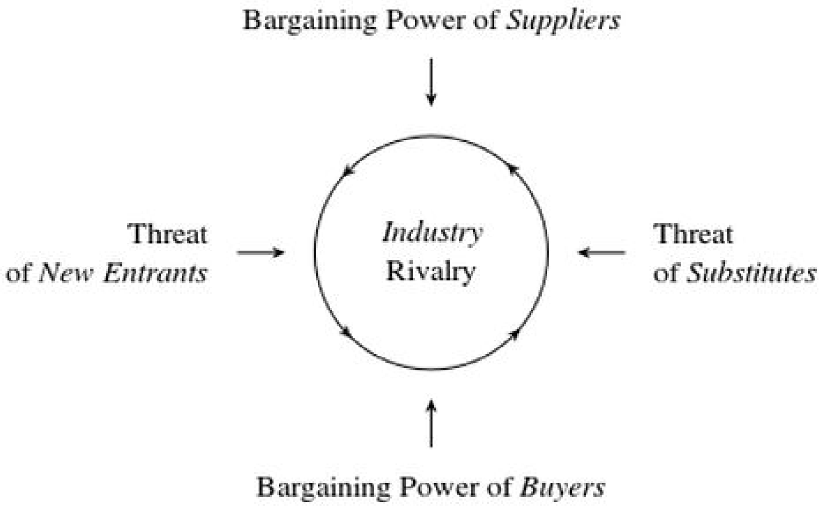
Syndication

Segment Attractiveness

1. Segments are used to identify areas where we can differentiate ourselves
2. After identifying the cascade’s segment(s), we need to analyze its “attractiveness”, IOW how easy it will be to succeed in that segment.
3. We’ll use Michael Porter’s Five Forces model to frame this analysis.

Segment Attractiveness (5 Forces)

Everybody knows how Michael Porter models an industry![1] Let’s model ours!



[1]https://en.wikipedia.org/wiki/Porter_five_forces_analysis
[2] https://en.wikipedia.org/wiki/File:Elements_of_Industry_Structure.svg#file

Five Forces Analysis - New Entrants

Force	Comments
Economies of scale	No one has the capital, time and expertise to recreate Wikipedia’s corpus. But they don’t have to since the license is permissive.
Cost Advantages	Cost advantages for digital companies are generally economies of learnings connected with already being in an industry[1] such as process. Our fundraising capabilities are probably the best example.
Capital requirements	We don’t have that much data and it’s easy to scale with incremental costs. Startup capital is also pretty easy in 2015
Differentiation	Differentiation is connected to brand awareness and customer loyalty. We’re strong in awareness (I think) but perhaps weaker in (reader) loyalty
Distribution	• Our position in Google’s SERPs is critical to our distribution. It’s unlikely that anyone could get consistently ranked higher in average, but we’re definitely vulnerable here. • We do not have good social or app distribution strategies (we’re working on it)
Government and Legal Barriers	I can’t think of any here
Retaliation	Our culture, license and open-access philosophy make our retaliation against new entrants unlikely

[1] Grant, Robert, Contemporary Strategy Analysis, Blackwell Publishing, 2002

Five Forces Analysis - Rivalries

Force	Comments
Concentration	Concentration is measure of how many companies offer similar products. In the digital world, the concentration of social networks is very high (1), while the concentration of publishers is very low.
Cost Advantages	Our running costs (~US$ 60mm for 2105-16) are much lower than other industry players, largely because we don't maintain large advertising and targeting systems and hardware footprints.
Differentiation	Differentiation is connected to brand awareness and customer loyalty. We're strong in awareness (I think) but perhaps weaker in (reader) loyalty in GN; however none of these things apply in GS.

Segments	Competitors
Lookup	Google, Bing, IMDB, Quora, Stack Overflow, Reddit, Yahoo Answers, Baidu
Learning	Google, Stack Overflow, Quora, Khan Academy, Coursera, MIT, Online degrees, Code Academy, Wikihow
Browse	New York Times, New Yorker, many other publishers both online and offline

[1] Grant, Robert, Contemporary Strategy Analysis, Blackwell Publishing, 2002

Five Forces Analysis - Buyers

Force	Comments
Concentration	Refers to number/size of buyers but mainly in enterprise where # of buyers is small. This is generally not an issue in consumer spaces. (is it?)
Switching Costs	Essentially no cost in switching to another provider.
Differentiation	Differentiation is connected to brand awareness and customer loyalty. We’re strong in awareness (I think) but perhaps weaker in (reader) loyalty.
Number/Growth	An industry is attractive when the numbers of buyers is increasing

[1] Grant, Robert, Contemporary Strategy Analysis, Blackwell Publishing, 2002

Five Forces Analysis - Sellers

Force	Comments
Concentration	In this case, how many options do the editors have? (Stack overflow, quora, etc) How many options do we have? (bots, mechanical ingestion, etc)
Switching Costs	More suitable to enterprise situations; although Editors can fork.
Differentiation	Due to size and scope, Wikipedia is well differentiated to editors (I would think?)

[1] Grant, Robert, Contemporary Strategy Analysis, Blackwell Publishing, 2002

Five Forces Analysis - Substitutes

Force	Comments
Buyer’s desire to substitute	Will Wikipedia users actually use other sources of info? How loyal are they
Relative performance	• Considering our license, I doubt there’s much difference here. • Our competitors (Google) can deliver the content faster due to data center footprint

[1] Grant, Robert, Contemporary Strategy Analysis, Blackwell Publishing, 2002