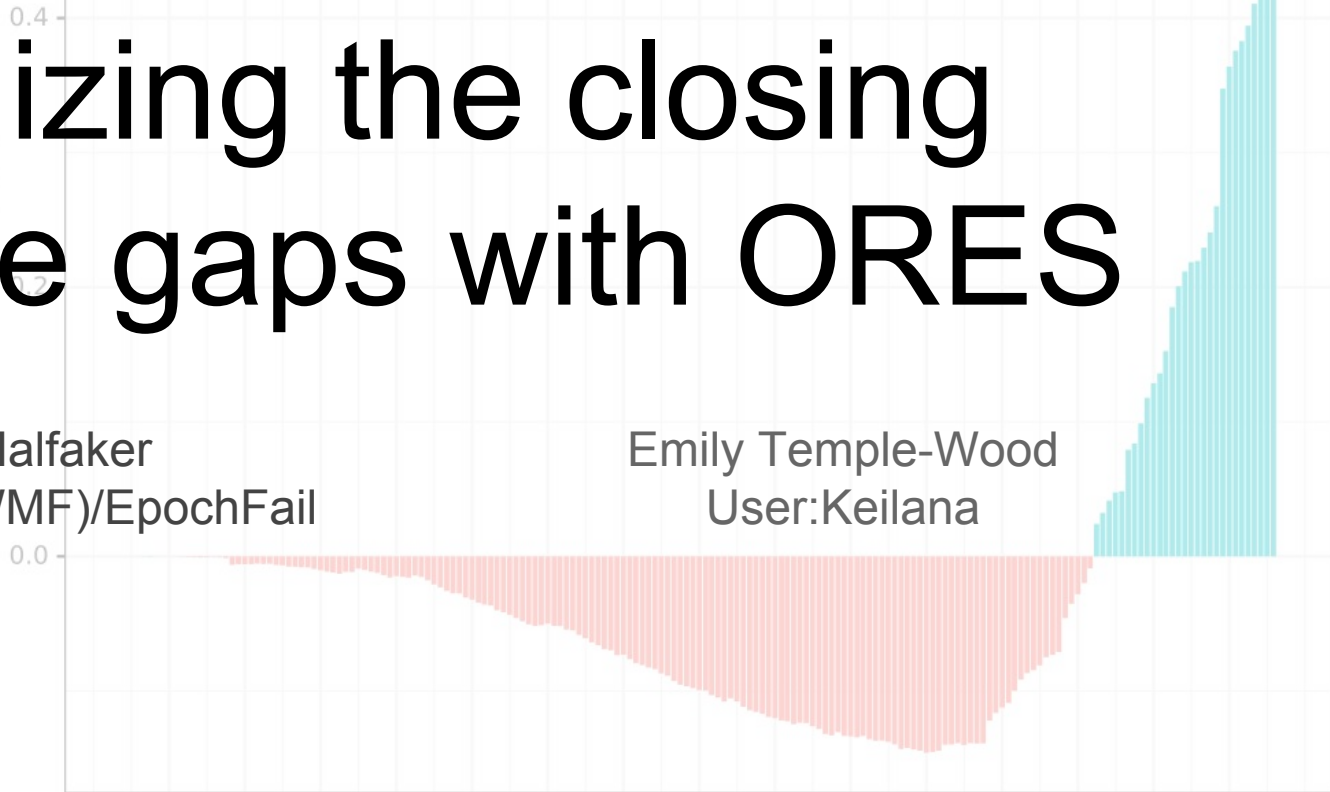# The Keilana Effect: Visualizing the closing coverage gaps with ORES

Aaron Halfaker
User:Halfak_(WMF)/EpochFail

Emily Temple-Wood
User:Keilana

Wikimania 2017, Montreal Canada

Aaron Halfaker

Emily Temple-Wood

Story Time!

Outline

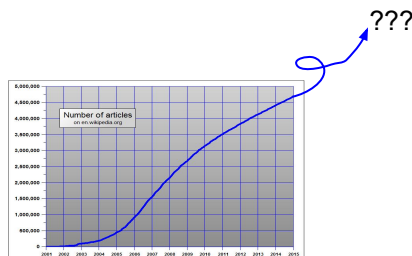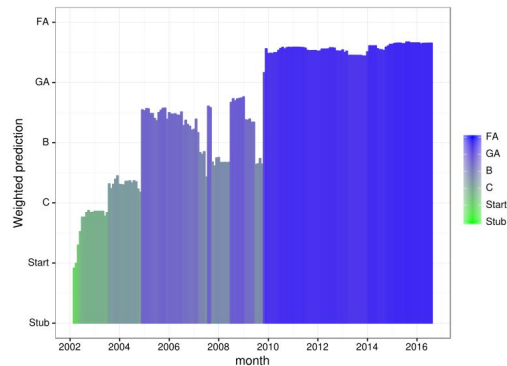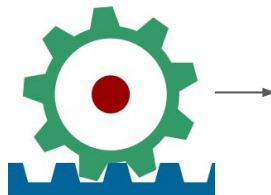1. Content coverage in Wikipedia
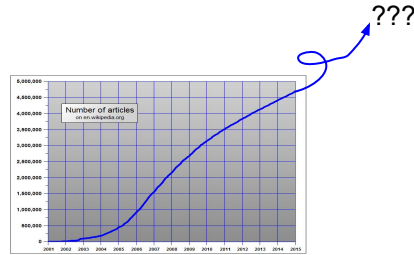
Outline

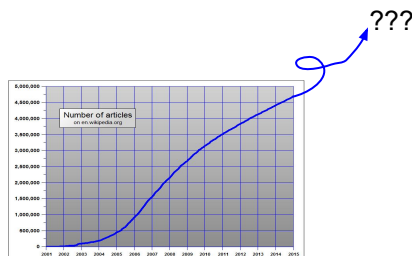1. Content coverage in Wikipedia
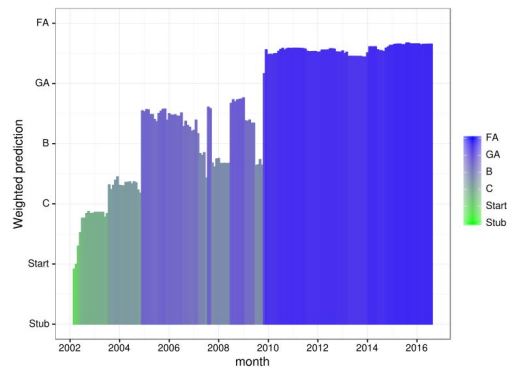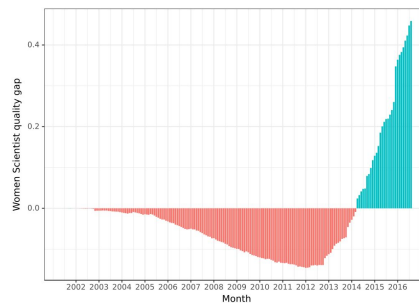
2. Measuring coverage with ORES

Outline

1. Content coverage in Wikipedia

2. Measuring coverage better

3. Measuring a closing coverage gap

Part 1

Content coverage in Wikipedia

Wikipedia:Wikipedia is a work in progress

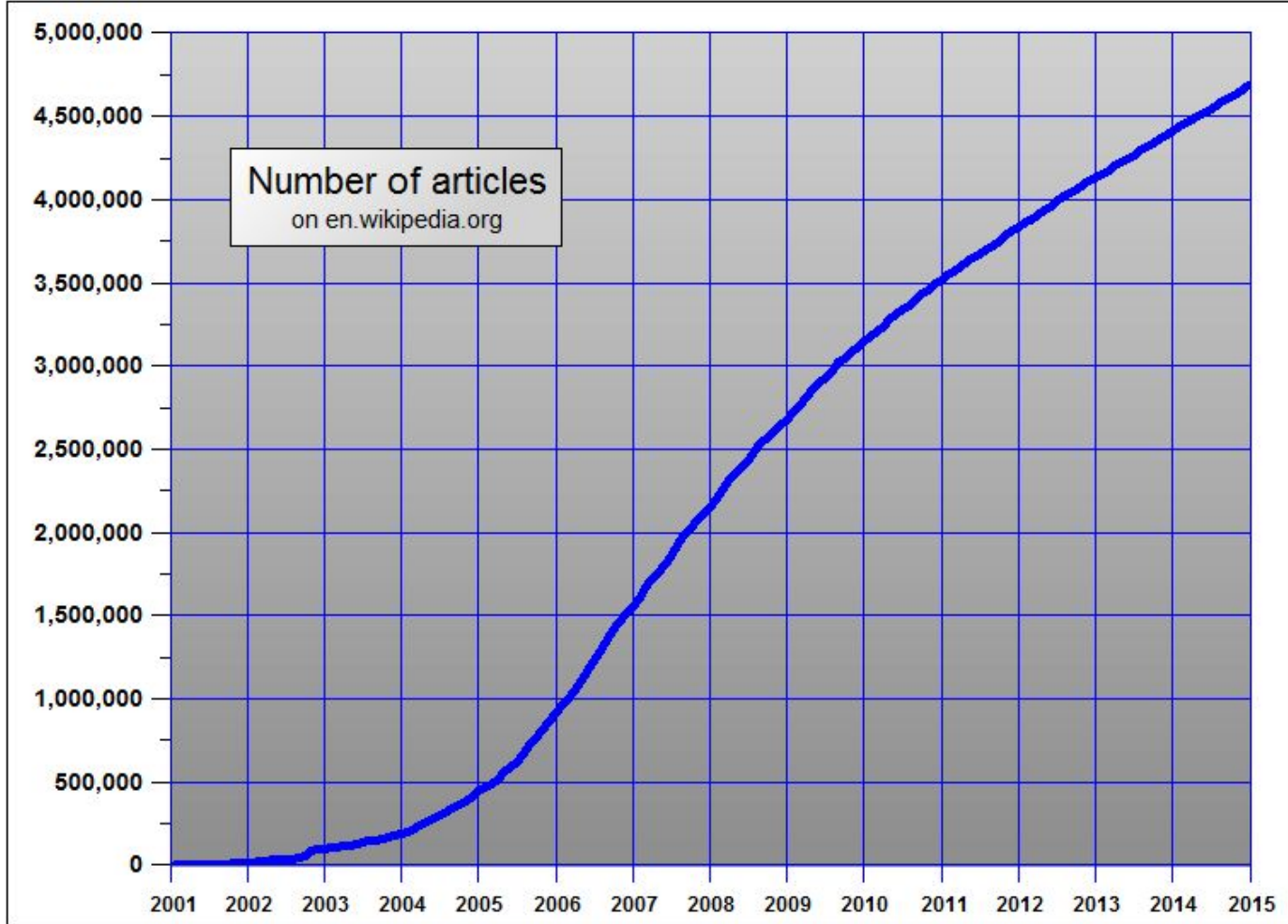*See also Wikipedia:Work in progress (disambiguation) and meta:Eventualism.*

Wikipedia is, by number of articles, the largest encyclopedia ever to exist. It contains a lot of information, and has been edited and viewed by millions of people, many of whom have found it useful. Unfortunately, much of it could be a lot better. Many people have eagerly pointed this out—often failing to give weight to the notion that it has been created entirely by volunteers, from nothing, in just thirteen years—and some have even suggested that the Web would be better off without it. However, in airing their complaints, they frequently miss out one crucial detail: **Wikipedia is not finished**. Not even close. In fact, we're barely getting started.

In its very early days, Wikipedia went through several major software changes. Existing wiki software was not designed for writing encyclopedias, and developing the first version of MediaWiki took time. As a result, much of the earliest page histories have been lost, and while the history of some pages is preserved right back to January 2001, other pages which are equally old have no information from before 2002. It is possible to see Wikipedia as it looked in its entirety in December 2001, as a read-only copy of the pages at that time is hosted at nostalgia.wikipedia.org. To really demonstrate the point, though, it is necessary to go back even further.

The Wikipedia logo. See that big hole in the top? That's there for a reason, and that reason is it is not finished. It's not *just* to make it look like the Second Death Star.

Number of articles
on en.wikipedia.org

???

Number of articles
on en.wikipedia.org

Number of articles
on en.wikipedia.org

???
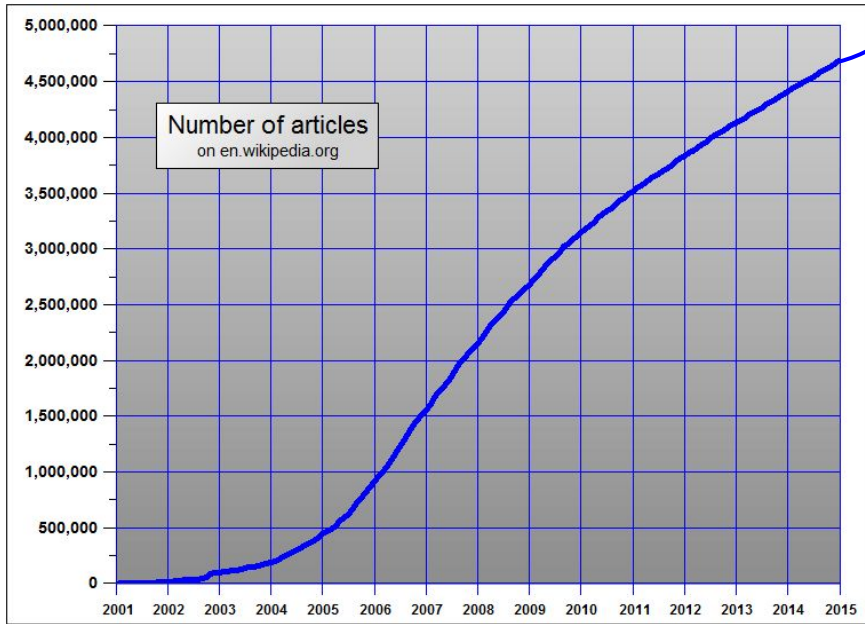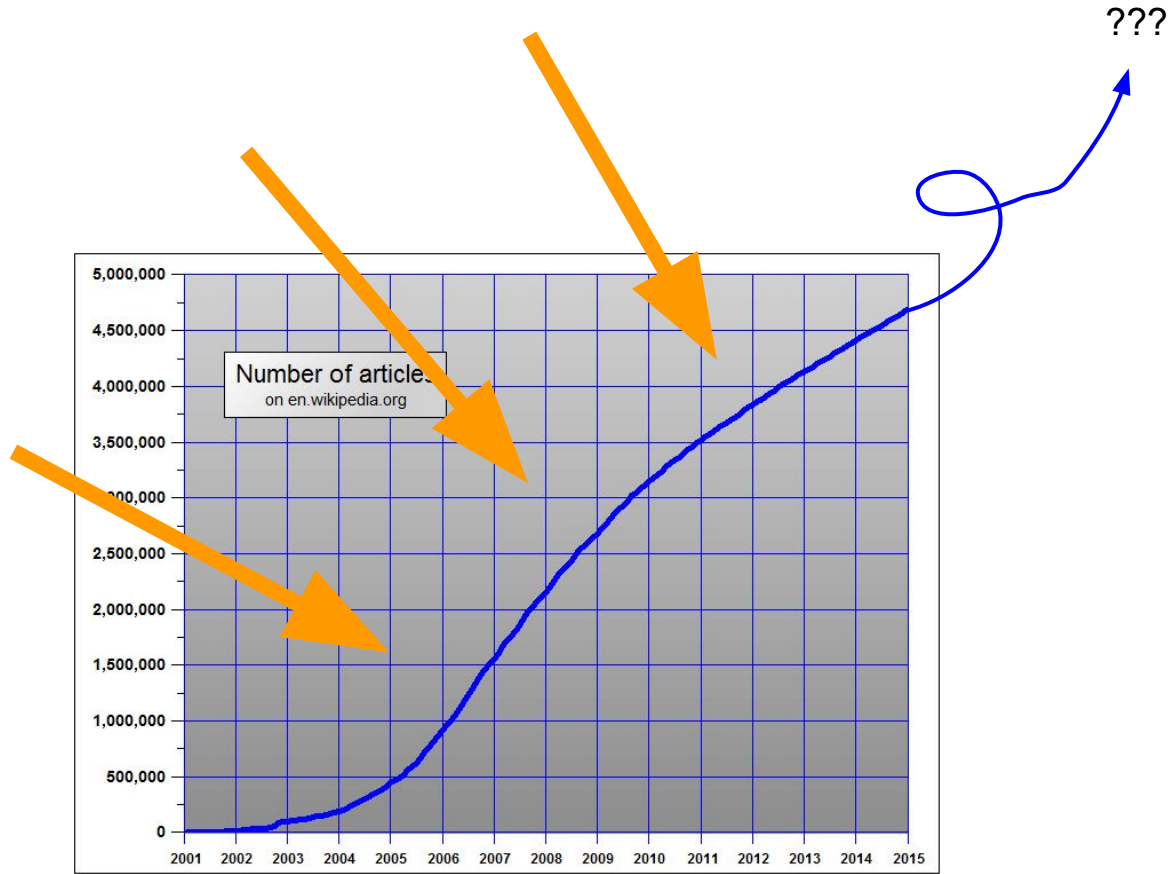
Nelle Morton

A start-class article from Wikipedia, the free encyclopedia

Nelle Morton (?–July 1987) was a feminist activist, civil rights leader, and professor. She taught at Drew University, and after she retired in 1970, wrote essays on feminist theory. In 1985, she published *The Journey is Home*, a collection of her essays discussing feminism, spirituality, religion, and intersectionality.[1]

References

1. **^** Keller, Catherine (Fall 1988). "Goddess, Ear, and Metaphor: On the Journey of Nelle Morton". *Journal of Feminist Studies in Religion*. **4** (2): 51–67. (subscription required (help)).

Nelle Morton

Nelle Morton (?–July 1987) was a feminist activist, civil rights leader, and professor. She taught at Drew University, and after she retired in 1970, wrote essays on feminist theory. In 1985, she published *The Journey is Home*, a collection of her essays discussing feminism, spirituality, religion, and intersectionality.[1]

References

1. ^ Keller, Catherine (Fall 1988). "Goddess, Ear, and Metaphor: On the Journey of Nelle Morton"🔗. *Journal of Feminist Studies in Religion*. **4** (2): 51–67. (subscription required (help)).

Alan Watts

Alan Wilson Watts (6 January 1915 – 16 November 1973) was a British philosopher, writer, and speaker, best known as an interpreter and populariser of Eastern philosophy for a Western audience. Born in Chislehurst, England, he moved to the United States in 1938 and began Zen training in New York. Pursuing a career, he attended Seabury-Western Theological Seminary, where he received a master's degree in theology. Watts became an Episcopal priest in 1945, then left the ministry in 1950 and moved to California, where he joined the faculty of the American Academy of Asian Studies.

Watts gained a large following in the San Francisco Bay Area while working as a volunteer programmer at KPFA, a Pacifica Radio station in Berkeley. Watts wrote more than 25 books and articles on subjects important to Eastern and Western religion, introducing the then-burgeoning youth culture to *The Way of Zen* (1957), one of the first bestselling books on Buddhism. In *Psychotherapy East and West* (1961), Watts proposed that Buddhism could be thought of as a form of psychotherapy and not a religion. He considered *Nature, Man and Woman* (1958) to be, "from a literary point of view—the best book I have ever written."[2] He also explored human consciousness, in the essay "The New Alchemy" (1958), and in the book *The Joyous Cosmology* (1962).

Towards the end of his life, he divided his time between a houseboat in Sausalito and a cabin on Mount Tamalpais. Many of his books are now available in digital format and many of his recorded talks and lectures are available on the Internet. According to the critic Erik Davis, his "writings and recorded talks still shimmer with a profound and galvanizing lucidity."[3]

Alan Watts	
Born	Alan Wilson Watts 6 January 1915 Chislehurst, Kent, England
Died	16 November 1973 (aged 58) Mt. Tamalpais, California, United States
Nationality	British and American[1]
Era	Contemporary philosophy
Region	Eastern Philosophy
School	Zen Buddhism Hinduism Pantheism Christianity Religious naturalism Taoism
Main interests	Personal identity Higher consciousness Aesthetics Cultural criticism Public ethics
Influences	[show]
Influenced	[show]

Contents [show]

Early years [edit]

Watts was born to middle class parents in the village of Chislehurst, Kent (now south-east London), in 1915, living at 3 (now 5) Holbrook Lane, which was subsequently lived in by author John Hemming-Clark in the early 2000s. Watts' father, Laurence Wilson Watts, was a representative for the London office of the Michelin Tyre Company; his mother, Emily Mary Watts (née Buchan), was a housewife whose father had been a missionary. With modest financial means, they chose to live in pastoral surroundings and Alan, an only child, grew up playing at brookside, learning the names of wildflowers and butterflies.[4] Probably because of the influence of his mother's religious family[5] the Buchans, an interest in "ultimate things" seeped in. But it mixed with Alan's own interests in storybook fables and romantic tales of the mysterious Far East.[6]

It's good to have any coverage at all.

It's better to have *complete* articles.

What's a complete article?

Class	Criteria	Reader's experience
⭐ FA	The article has attained featured article status by passing an official review. **More detailed criteria** [show]	Professional, outstanding, and thorough; a definitive source for encyclopedic information.
◍ A	The article is well organized and essentially complete, having been reviewed by impartial reviewers from this WikiProject or elsewhere. Good article status is not a requirement for A-Class. **More detailed criteria** [show]	Very useful to readers. A fairly complete treatment of the subject. A non-expert in the subject would typically find nothing wanting.
⊕ GA	The article has attained good article status by passing an official review. **More detailed criteria** [show]	Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.
B	The article is mostly complete and without major problems, but requires some further work to reach good article standards. **More detailed criteria** [show]	Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.
C	The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup. **More detailed criteria** [show]	Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.
Start	An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources. **More detailed criteria** [show]	Provides some meaningful content, but most readers will need more.
Stub	A very basic description of the topic. However, all very-bad-quality articles will fall into this category. **More detailed criteria** [show]	Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.

Wikipedia:Version 1.0 Editorial Team/Statistics

All rated articles by quality and importance

Quality	Importance					
	Top	High	Mid	Low	???	Total
⭐ FA	1,168	1,793	1,697	1,044	191	**5,893**
⭐ FL	141	561	650	600	116	**2,068**
◍ A	216	424	577	365	77	**1,659**
⊕ GA	2,059	4,695	9,173	9,872	1,687	**27,486**
B	11,951	22,639	34,618	27,432	13,619	**110,259**
C	10,127	29,141	64,890	89,320	42,720	**236,198**
Start	17,105	75,083	302,452	770,461	288,260	**1,453,361**
Stub	4,214	30,501	224,407	1,827,987	836,433	**2,923,542**
List	2,980	11,075	33,668	91,122	60,902	**199,747**
Assessed	49,961	175,912	672,132	2,818,203	1,244,005	**4,960,213**
Unassessed	138	409	1,790	16,329	518,042	**536,708**
Total	50,099	176,321	673,922	2,834,532	1,762,047	**5,496,921**

About this table

Article Rating
- ⭐ FA: 2,081
- A-class: 2,017
- GA-class: 2,093
- B-class: 2,107
- C-class: 2,095
- Start: 2,119
- Stub: 2,087
- ⭐ FL: 2,028
- List: 2,076

About this pie chart

Biology

A good article from Wikipedia, the free encyclopedia

> *"Biological science" redirects here. It is not to be confused with life science.*
>
> *For other uses, see Biology (disambiguation).*

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[*clarification needed*] However, despite the broad scope of biology, there are certain general and unifying concepts within it that govern all study and research, consolidating it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the synthesis and creation of new species. It is also understood today that all the organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living organisms.

- top: *E. coli* bacteria and gazelle
- bottom: Goliath beetle and tree fern

Sub-disciplines of biology are defined by the scale at which organisms are studied, the kinds of organisms studied, and the methods used to study them: biochemistry examines the rudimentary chemistry of life; molecular biology studies the complex interactions among biological molecules; botany studies the biology of plants; cellular biology

Biology

A good article from Wikipedia, the free encyclopedia

> *"Biological science" redirects here. It is not to be confused with life science.*
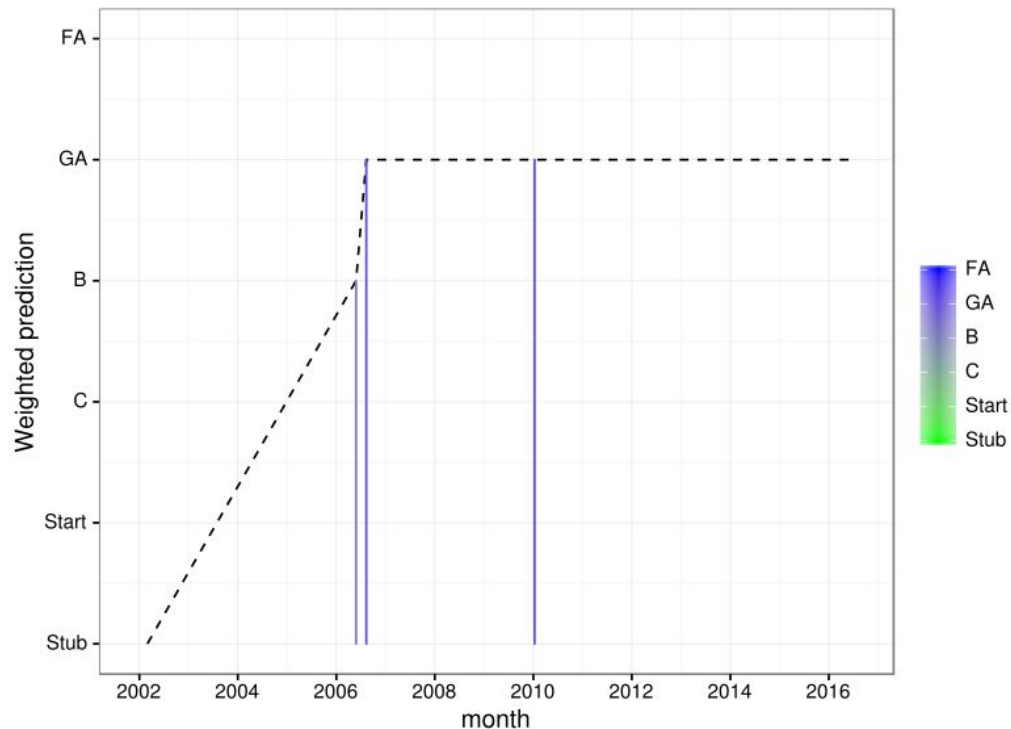>
> *For other uses, see Biology (disambiguation).*

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[*clarification needed*] However, despite the broad scope of biology, there are certain general and unifying concepts within it that govern all study and research, consolidating it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the synthesis and creation of new species. It is also understood today that all the organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living organisms.

- top: *E. coli* bacteria and gazelle
- bottom: Goliath beetle and tree fern

Sub-disciplines of biology are defined by the scale at which organisms are studied, the kinds of organisms studied, and the methods used to study them: biochemistry examines the rudimentary chemistry of life; molecular biology studies the complex interactions among biological molecules; botany studies the biology of plants; cellular biology

Biology

A good article from Wikipedia, the free encyclopedia

> *"Biological science" redirects here. It is not to be confused with life science.*
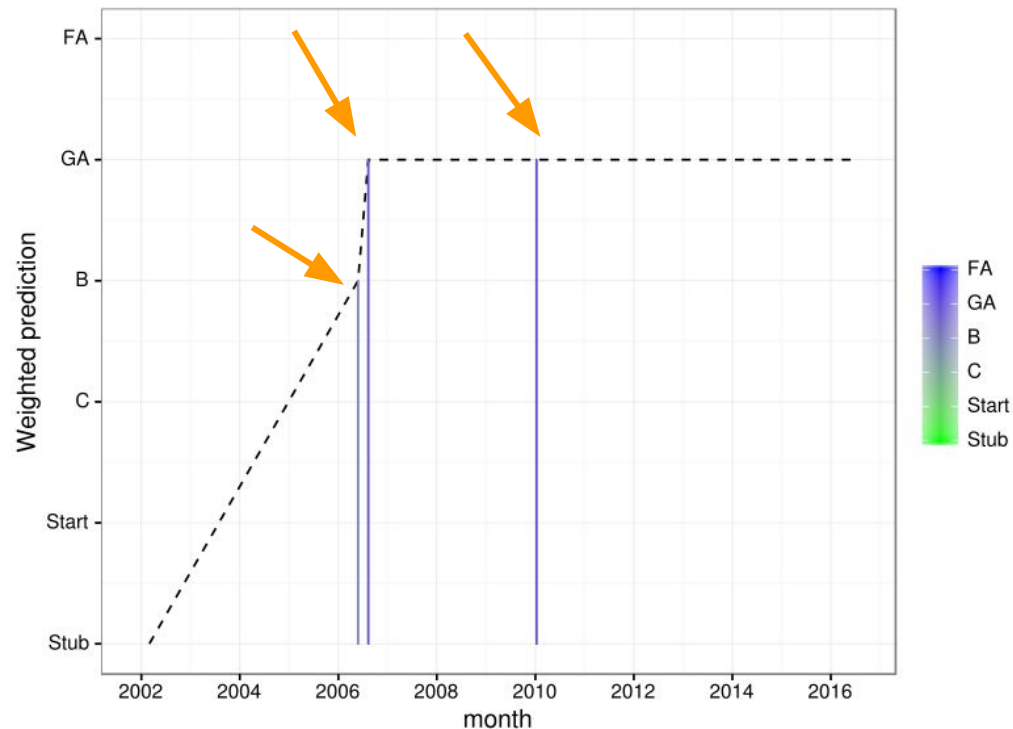>
> For other uses, see *Biology (disambiguation).*

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[*clarification needed*] However, despite the broad scope of biology, there are certain general and unifying concepts within it that govern all study and research, consolidating it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the synthesis and creation of new species. It is also understood today that all the organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living organisms.
- top: *E. coli* bacteria and gazelle
- bottom: Goliath beetle and tree fern

Sub-disciplines of biology are defined by the scale at which organisms are studied, the kinds of organisms studied, and the methods used to study them: biochemistry examines the rudimentary chemistry of life; molecular biology studies the complex interactions among biological molecules; botany studies the biology of plants; cellular biology

Biology

A good article from Wikipedia, the free encyclopedia

"Biological science" redirects here. It is not to be confused with life science.

For other uses, see Biology (disambiguation).

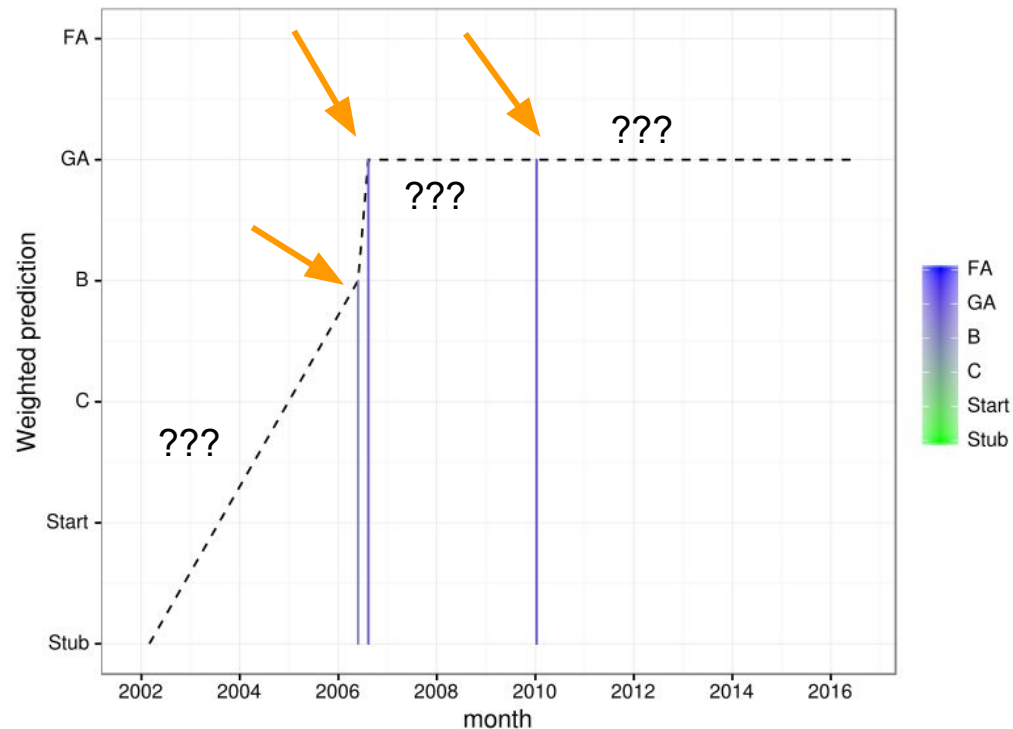**Biology** is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[clarification needed] However, despite the broad scope of biology, there are certain general and unifying concepts within it that govern all study and research, consolidating it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the synthesis and creation of new species. It is also understood today that all the organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living organisms.

- top: *E. coli* bacteria and gazelle
- bottom: Goliath beetle and tree fern

Sub-disciplines of biology are defined by the scale at which organisms are studied, the kinds of organisms studied, and the methods used to study them: biochemistry examines the rudimentary chemistry of life; molecular biology studies the complex interactions among biological molecules; botany studies the biology of plants; cellular biology
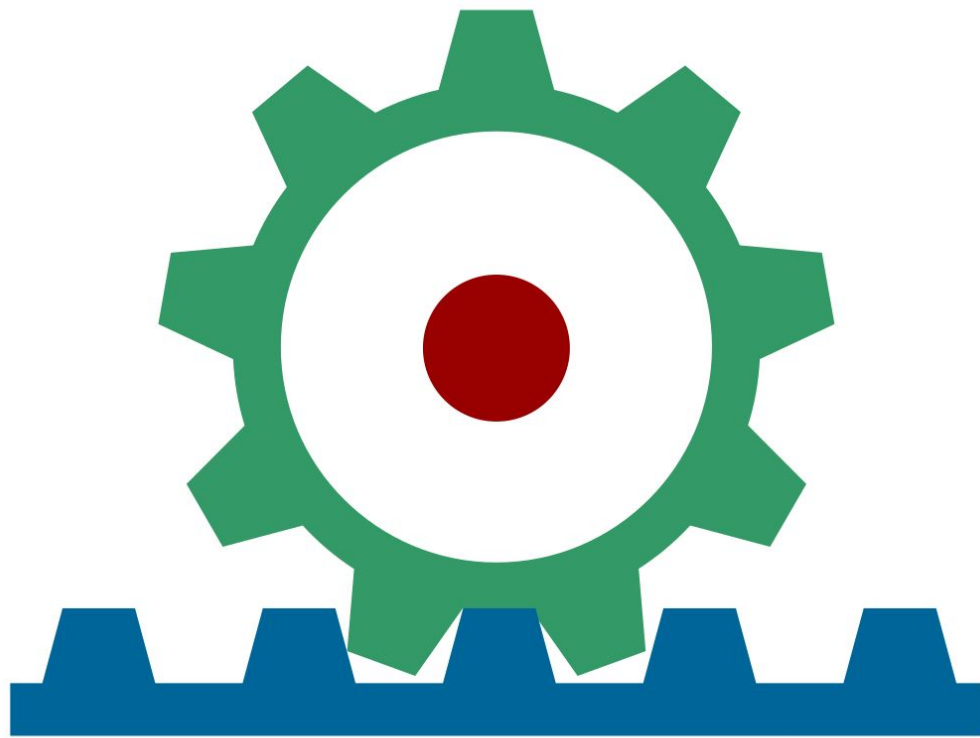
Where are our coverage gaps?

When did the gaps appear?

Have our initiatives helped?
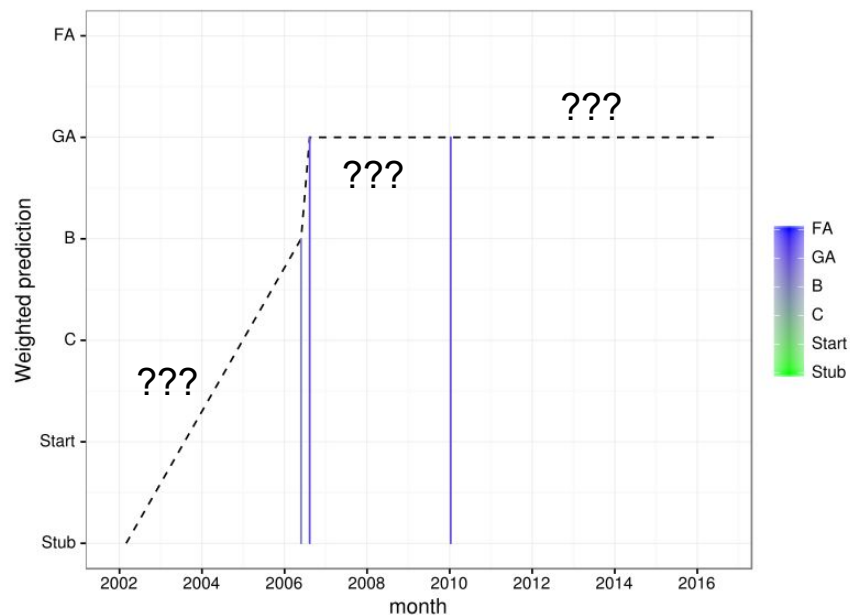
Part 2

Measuring coverage with ORES

ORES (/ɔɹz/)

The Objective Revision Evaluation Service (ORES) is a web service that provides machine learning as a service for Wikimedia Projects. The system is designed to help automate critical wiki-work -- for example, vandalism detection and removal. This service is developed as part of the Revision scoring as a service research project.
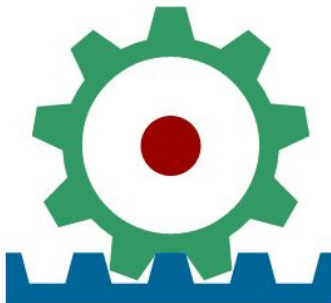
Scores API

ORES is intended to be used as a source of information by tool developers. To access ORES, scores, a simple RESTful API is provided. There are two versions of the scoring API that differ slightly in their behavior. Version 2 provides access to model info in a scoring request. Version 1 is preserved for backwards compatibility.

Version 2 (docs)
   The current end recommended version of the API.
   Accessible via /v2/scores/.
Version 1 (docs)
   Preserved for backwards compatibility -- this version of the API implements ORES original behavior.
   Accessible via /scores/ and /v1/scores/.

Reference scoring user interface

ORES provides a simple user interface for obtaining scores. See /ui/.

On-wiki documentation

To read more about the service and the models available, see:

https://meta.wikimedia.org/wiki/Objective_Revision_Evaluation_Service

Privacy policy

See the Wikimedia policy here: https://wikimediafoundation.org/wiki/Privacy_policy

Wiki Thing → Prediction

Objective Revision Evaluation Service logo by Offnfopt (PD 0)

WIKIMEDIA
FOUNDATION

"Intelligence" is hard to define, whether in dogs, other animals, or humans. The ability to learn quickly might be taken as a sign of [[intelligence (trait)|intelligence]], but such evidence must be interpreted with care, because learning speed may be affected by such things as the effectiveness of the rewards used in training or the motivation or activity level of the dog. For example, some breeds, such as [[Siberian Husky|Siberian Huskies]], are said to be not particularly rewarded by pleasing their owners, but quickly learn to escape from yards or catch small animals, often using ingenious ways of doing both. **LLAMAS GROW ON TREES**

https://en.wikipedia.org/wiki/?diff=642215410

Good!

Bad!

WIKIMEDIA
FOUNDATION

Biology

A good article from Wikipedia, the free encyclopedia

"Biological science" redirects here. It is not to be confused with *life science*.
For other uses, see *Biology (disambiguation)*.

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[*clarification needed*] However, despite the broad scope of biology, there are certain unifying concepts within it that consolidate it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the creation of new species. It is also understood today that all organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living

https://en.wikipedia.org/wiki/Biology

Prediction

Biology

A good article from Wikipedia, the free encyclopedia

"Biological science" redirects here. It is not to be confused with life science.
For other uses, see Biology (disambiguation).

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[clarification needed] However, despite the broad scope of biology, there are certain unifying concepts within it that consolidate it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the creation of new species. It is also understood today that all organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living

[[:en:Biology]]

Class	Criteria	Reader's experience
★ FA	The article has attained featured article status by passing an official review. **More detailed criteria** [show]	Professional, outstanding, and thorough; a definitive source for encyclopedic information.
⊕ GA	The article has attained good article status by passing an official review. **More detailed criteria** [show]	Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.
B	The article is mostly complete and without major problems, but requires some further work to reach good article standards. **More detailed criteria** [show]	Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.
C	The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup. **More detailed criteria** [show]	Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.
Start	An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources. **More detailed criteria** [show]	Provides some meaningful content, but most readers will need more.
Stub	A very basic description of the topic. However, all very-bad-quality articles will fall into this category. **More detailed criteria** [show]	Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.

[[Wikipedia:WikiProject_assessment]]

WIKIMEDIA
F O U N D A T I O N

Biology

A good article from Wikipedia, the free encyclopedia

"Biological science" redirects here. It is not to be confused with *life science*.
For other uses, see *Biology (disambiguation)*.

Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, identification and taxonomy.[1] Modern biology is a vast and eclectic field, composed of many branches and subdisciplines.[clarification needed] However, despite the broad scope of biology, there are certain unifying concepts within it that consolidate it into single, coherent field. In general, biology recognizes the cell as the basic unit of life, genes as the basic unit of heredity, and evolution as the engine that propels the creation of new species. It is also understood today that all organisms survive by consuming and transforming energy and by regulating their internal environment to maintain a stable and vital condition known as homeostasis.

Biology deals with the study of the many living

[[:en:Biology]]

WIKIMEDIA
FOUNDATION

Class	Criteria	Reader's experience
★ FA	The article has attained featured article status by passing an official review. **More detailed criteria** [show]	Professional, outstanding, and thorough; a definitive source for encyclopedic information.
⊕ GA	The article has attained good article status by passing an official review. **More detailed criteria** [show]	Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.
B	The article is mostly complete and without major problems, but requires some further work to reach good article standards. **More detailed criteria** [show]	Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.
C	The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup. **More detailed criteria** [show]	Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.
Start	An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources. **More detailed criteria** [show]	Provides some meaningful content, but most readers will need more.
Stub	A very basic description of the topic. However, all very-bad-quality articles will fall into this category. **More detailed criteria** [show]	Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.

[[Wikipedia:WikiProject_assessment]]

User:Nettrom

From Wikipedia, the free encyclopedia

I have a PhD in computer science from GroupLens Research at the University of Minnesota. My research is focused on the production and consumption of quality content in peer production communities, communities like Wikipedia and OpenStreetMap. As part of that research I helped develop the article quality model that's used in ORES, which allows you to get reasonably accurate predictions of article quality. I also run SuggestBot, a bot that will send you suggestions of articles you might like to edit, and they're all articles in need of improvement.

In my spare time I like to play guitar and squash.

Research publications [edit]

Peer-reviewed papers about Wikipedia:

- **Warncke-Wang, M.**, Ranjan, V., Terveen, L., and Hecht, B. "Misalignment Between Supply and Demand of Quality Content in Peer Production Communities", ICWSM 2015. pdf 📄 See also: Signpost/Research Newsletter coverage
- **Warncke-Wang, M.**, Ayukaev, V. R., Hecht, B., and Terveen, L. "The Success and Failure of Quality Improvement Projects in Peer Production Communities", CSCW 2015. pdf 📄
- **Warncke-Wang, M.**, Cosley, D., and Riedl, J. "Tell Me More: An Actionable Quality Model for Wikipedia", WikiSym, 2013. pdf 📄
- **Warncke-Wang, M.**, Uduwage, A., Dong, Z., and Riedl, J. "In Search of the Ur-Wikipedia: Universality, Similarity, and Translation in the Wikipedia Inter-language Link Network", WikiSym, 2012. pdf 📄

User:Nettrom at Wikimania 2013, photo by Daniel Schwen, CC-BY-SA

User:Nettrom

From Wikipedia, the free encyclopedia

I have a PhD in computer science from GroupLens Research at the University of Minnesota. My research is focused on the production and consumption of quality content in peer production communities, communities like Wikipedia and OpenStreetMap. As part of that research I helped develop the article quality model that's used in ORES, which allows you to get reasonably accurate predictions of article quality. I also run SuggestBot, a bot that will send you suggestions of articles you might like to edit, and they're all articles in need of improvement.

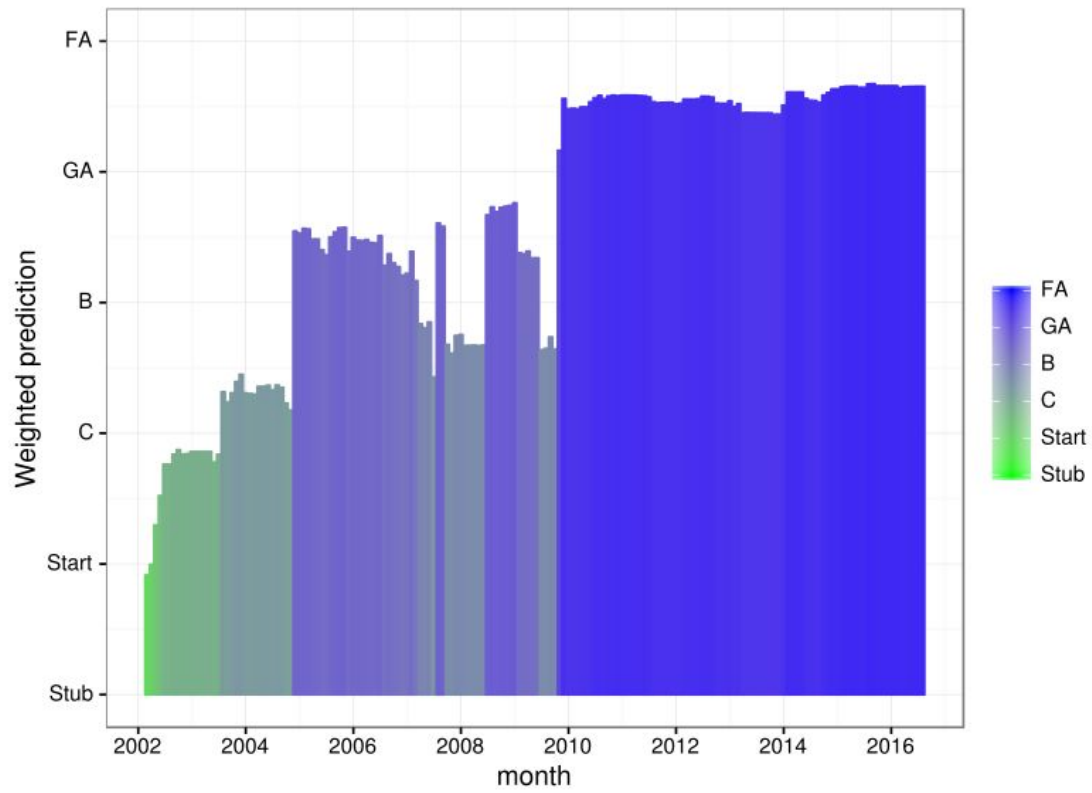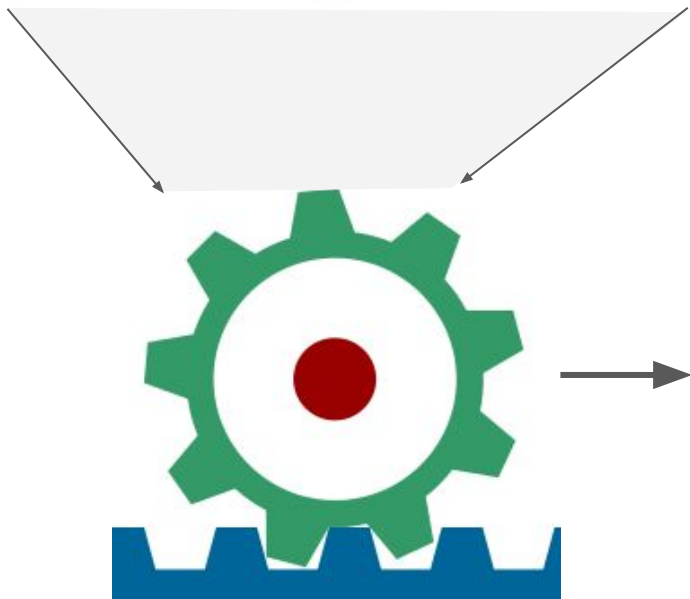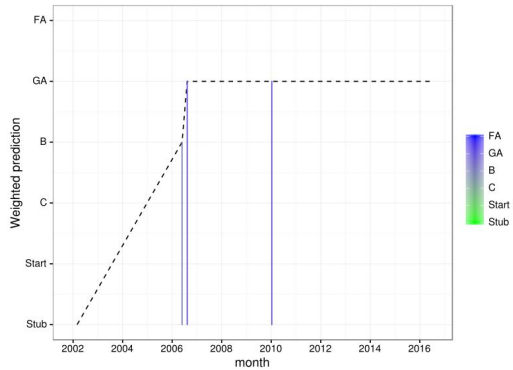In my spare time I like to play guitar and squash.
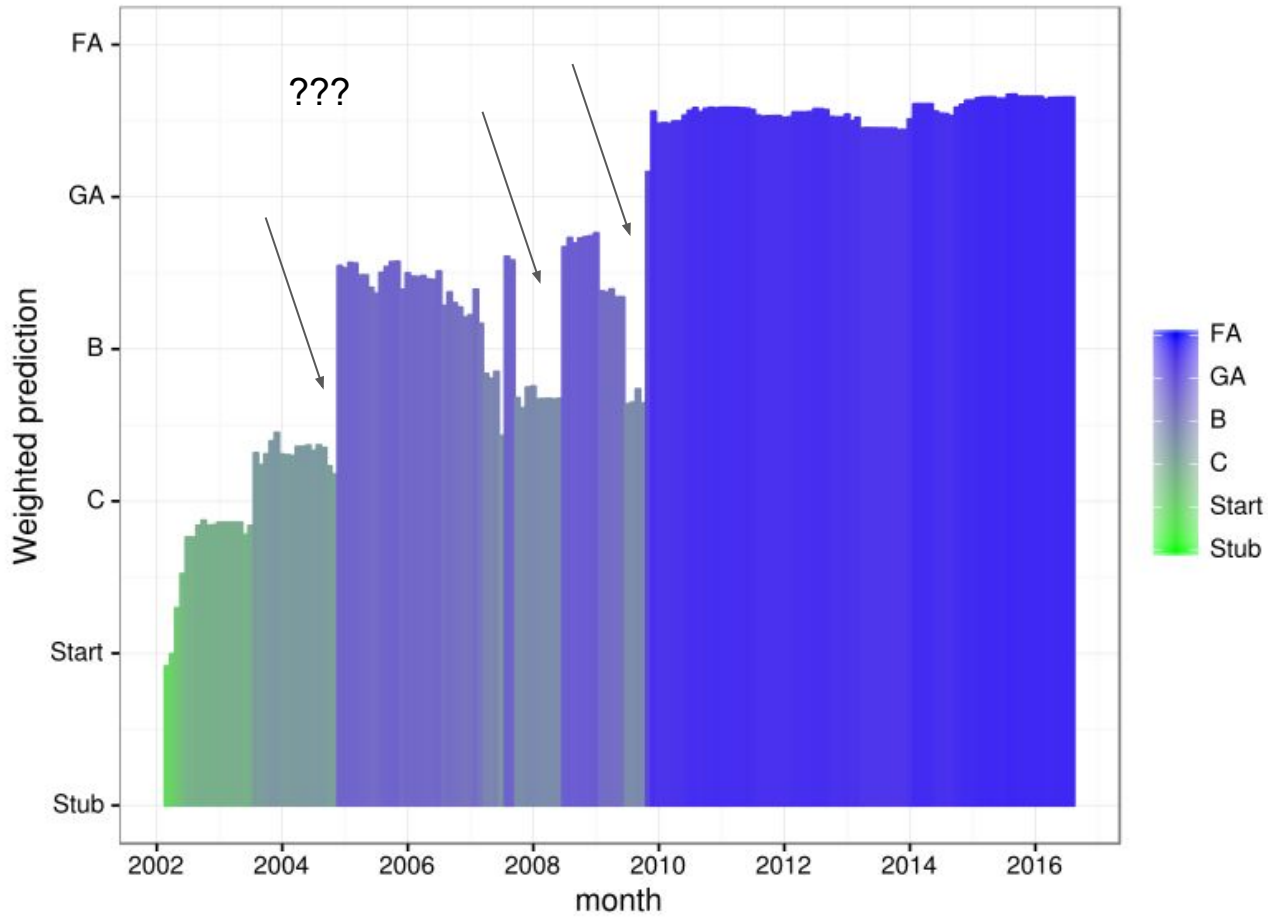
Research publications [edit]
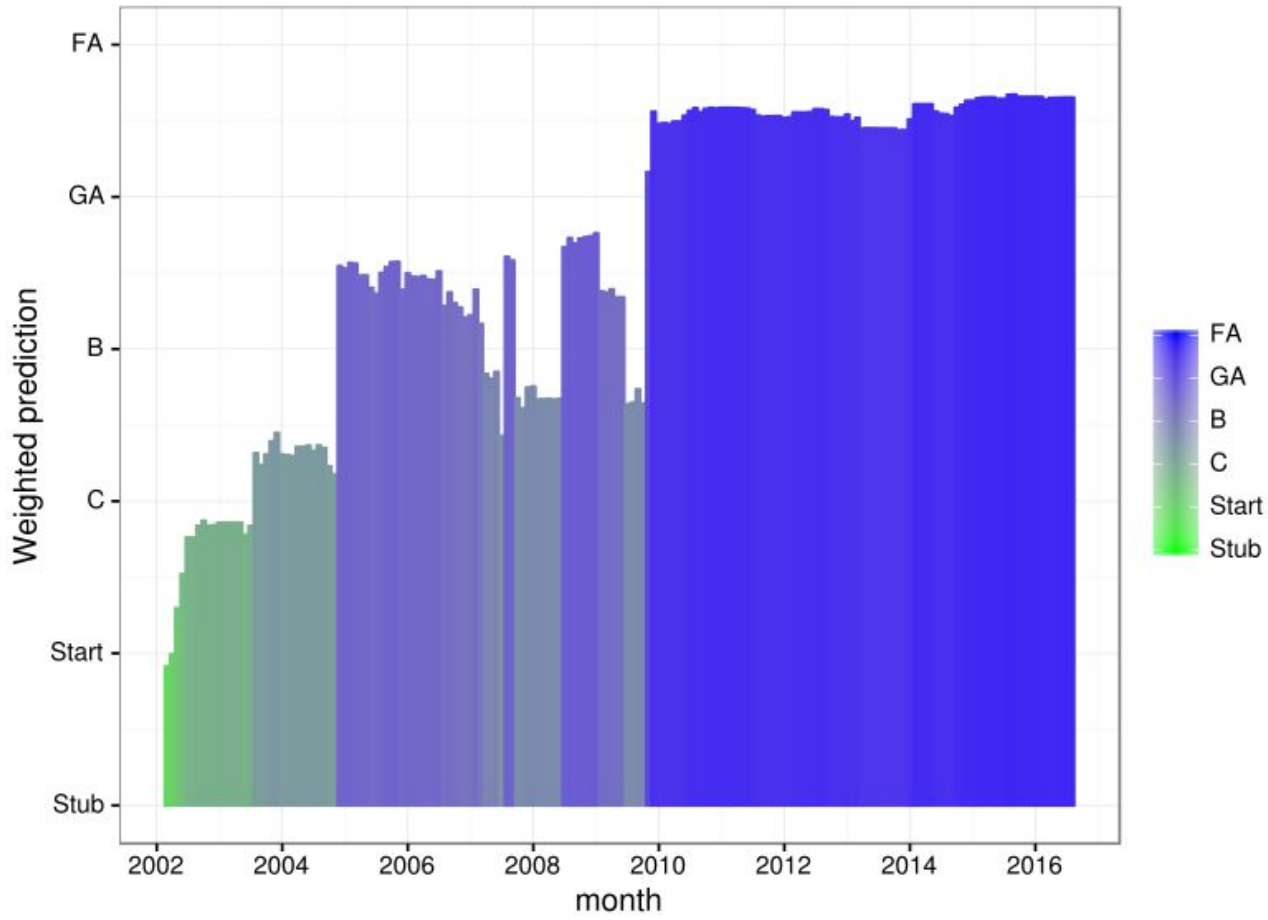
Peer-reviewed papers about Wikipedia:

- **Warncke-Wang, M.**, Ranjan, V., Terveen, L., and Hecht, B. "Misalignment Between Supply and Demand of Quality Content in Peer Production Communities", ICWSM 2015. pdf See also: Signpost/Research Newsletter coverage
- **Warncke-Wang, M.**, Ayukaev, V. R., Hecht, B., and Terveen, L. "The Success and Failure of Quality Improvement Projects in Peer Production Communities", CSCW 2015. pdf
- **Warncke-Wang, M.**, Cosley, D., and Riedl, J. "Tell Me More: An Actionable Quality Model for Wikipedia", WikiSym, 2013. pdf
- **Warncke-Wang, M.**, Uduwage, A., Dong, Z., and Riedl, J. "In Search of the Ur-Wikipedia: Universality, Similarity, and Translation in the Wikipedia Inter-language Link Network", WikiSym, 2012. pdf

User:Nettrom at Wikimania 2013, photo by Daniel Schwen, CC-BY-SA

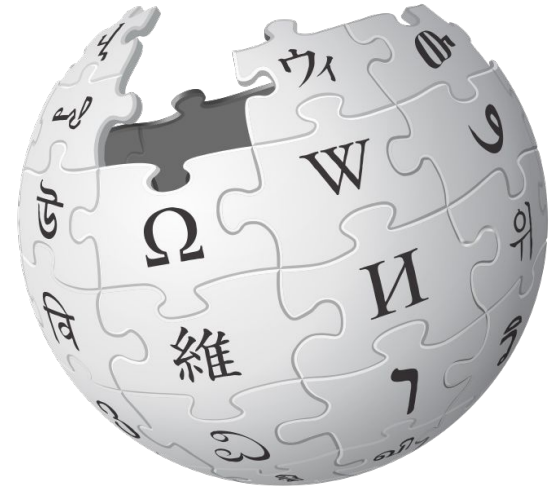# Cool, right?

Part 3

Measuring a closing coverage gap

WikiProject Women Scientists

Portrait of Marie Curie
(Public domain)

vs.

All articles in Wikipedia

?
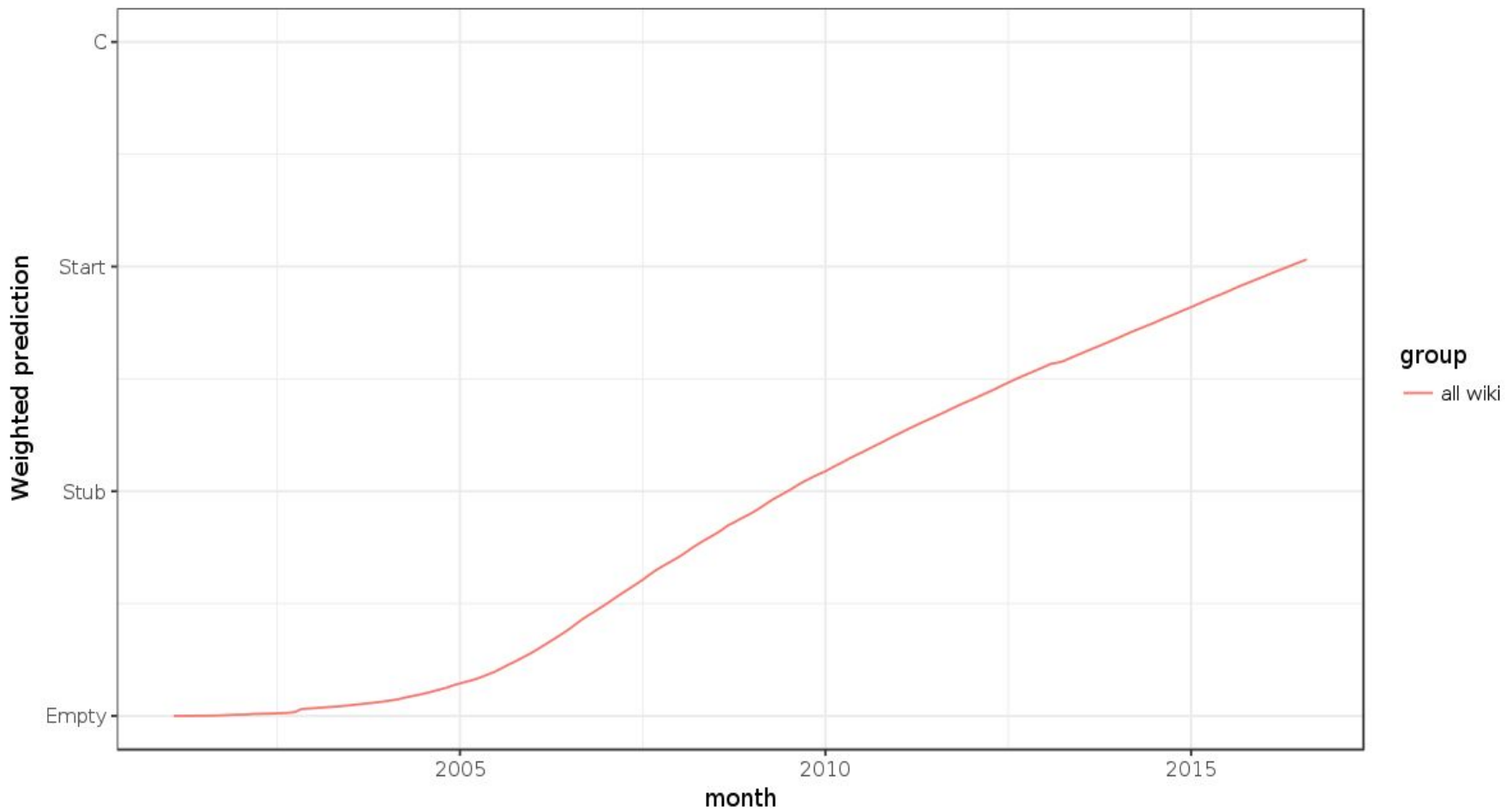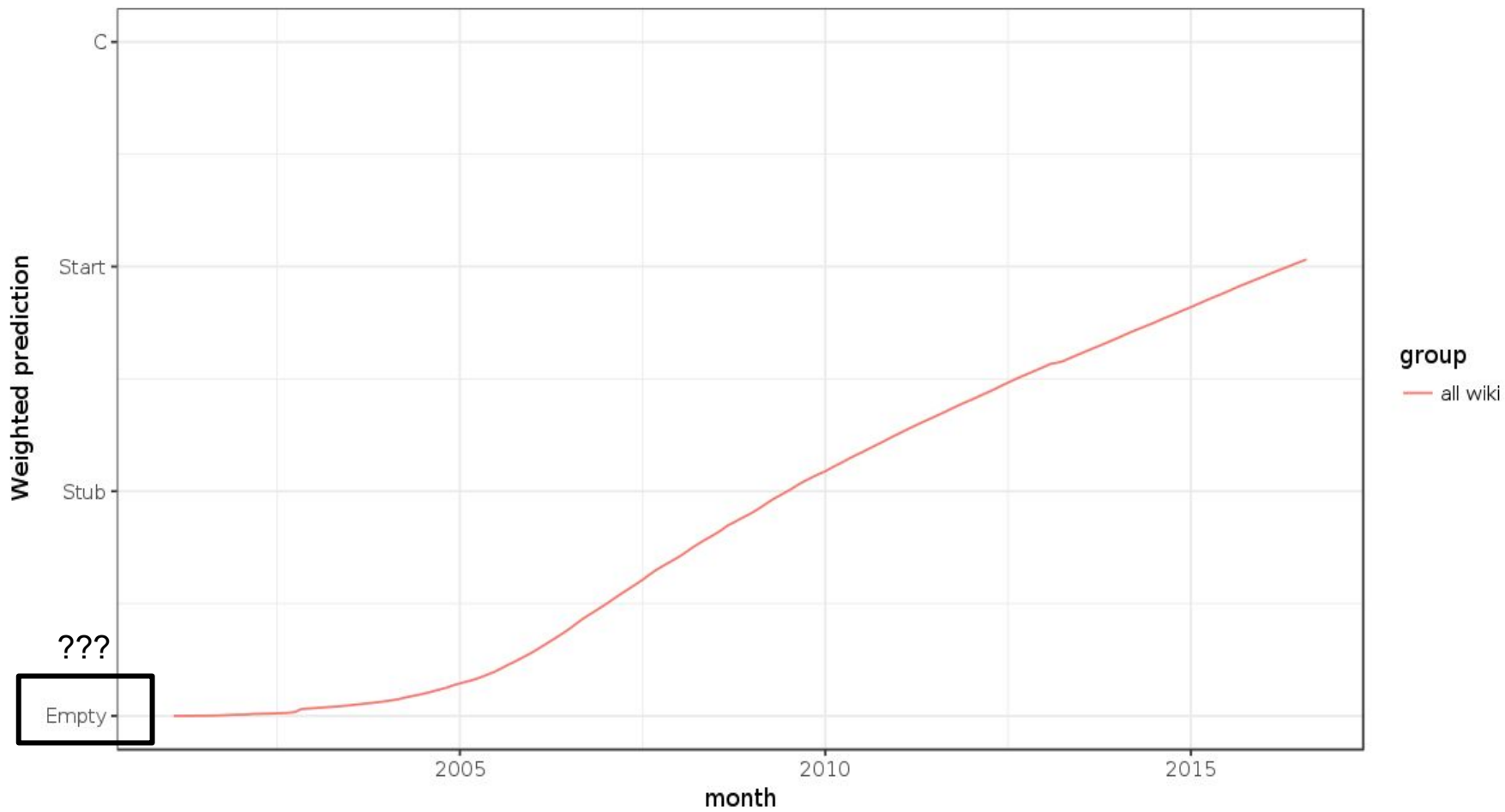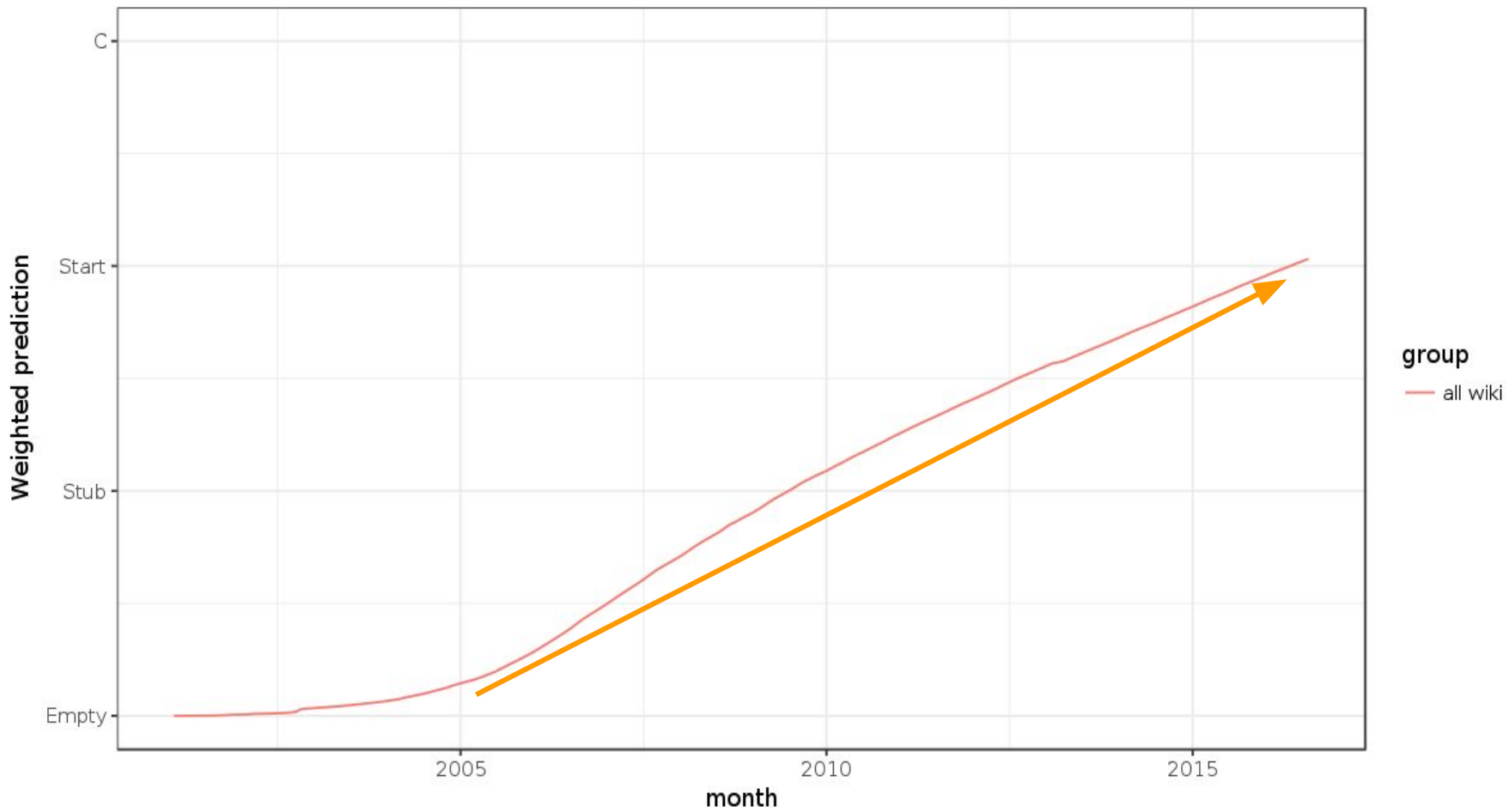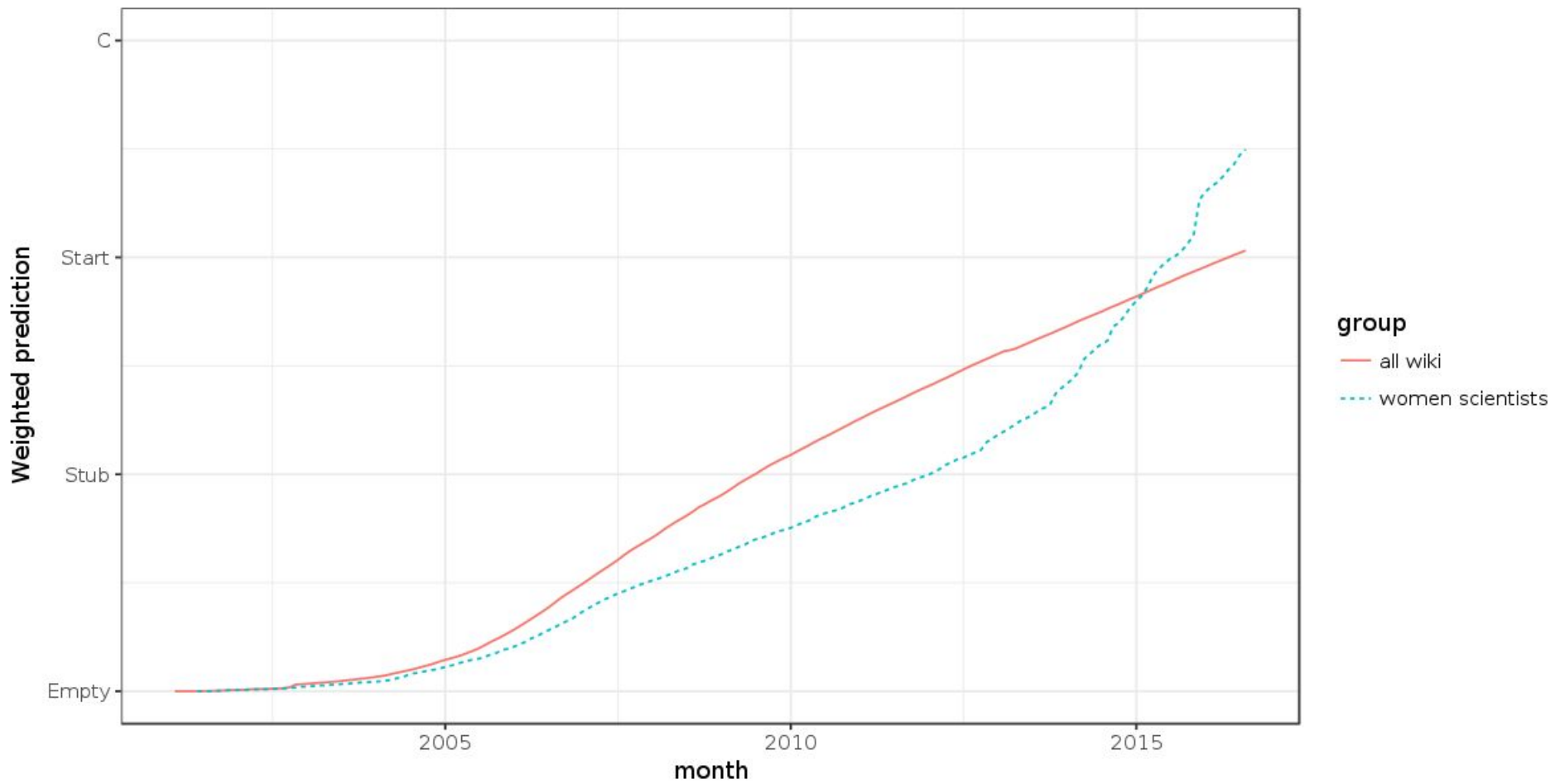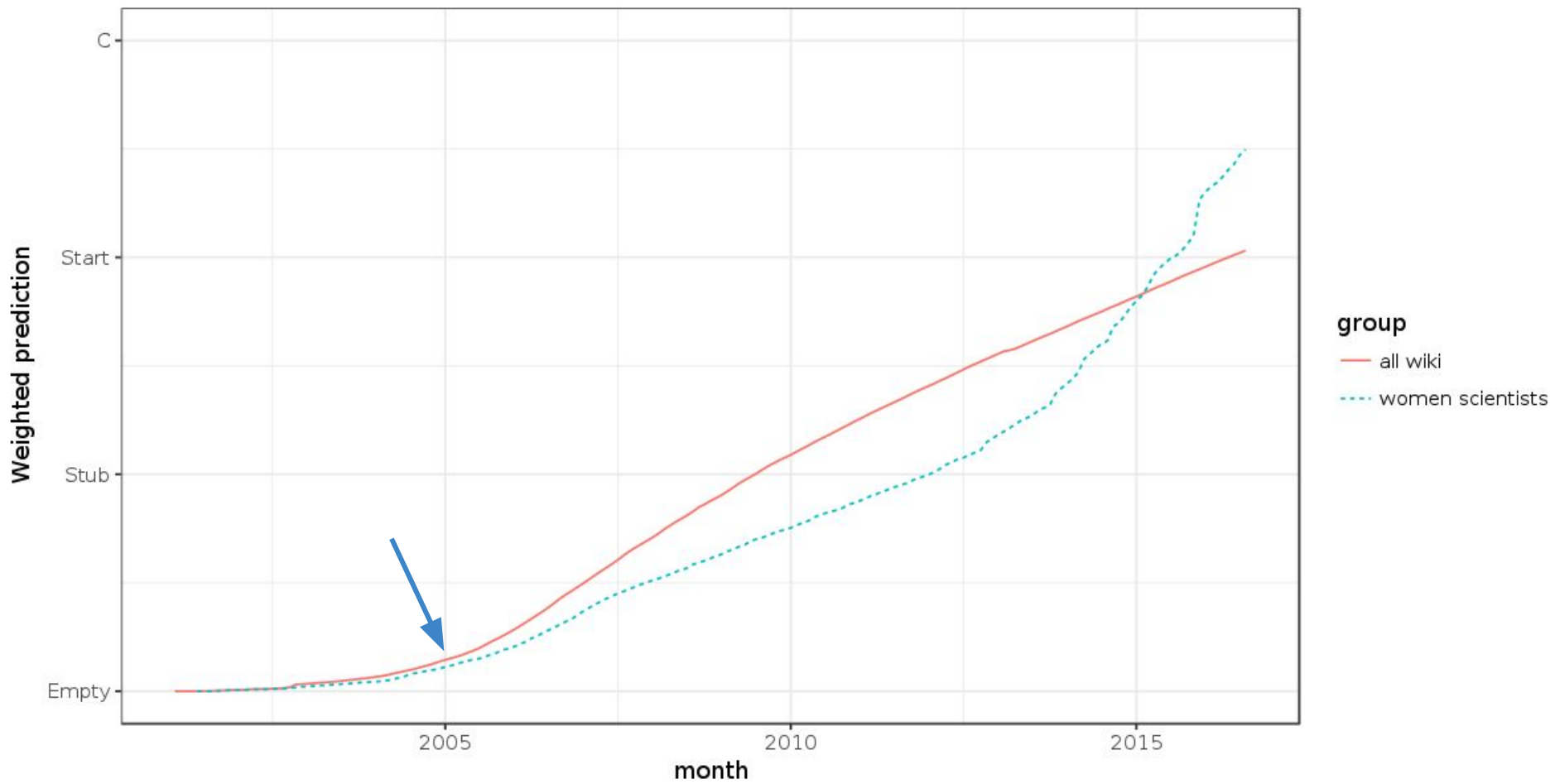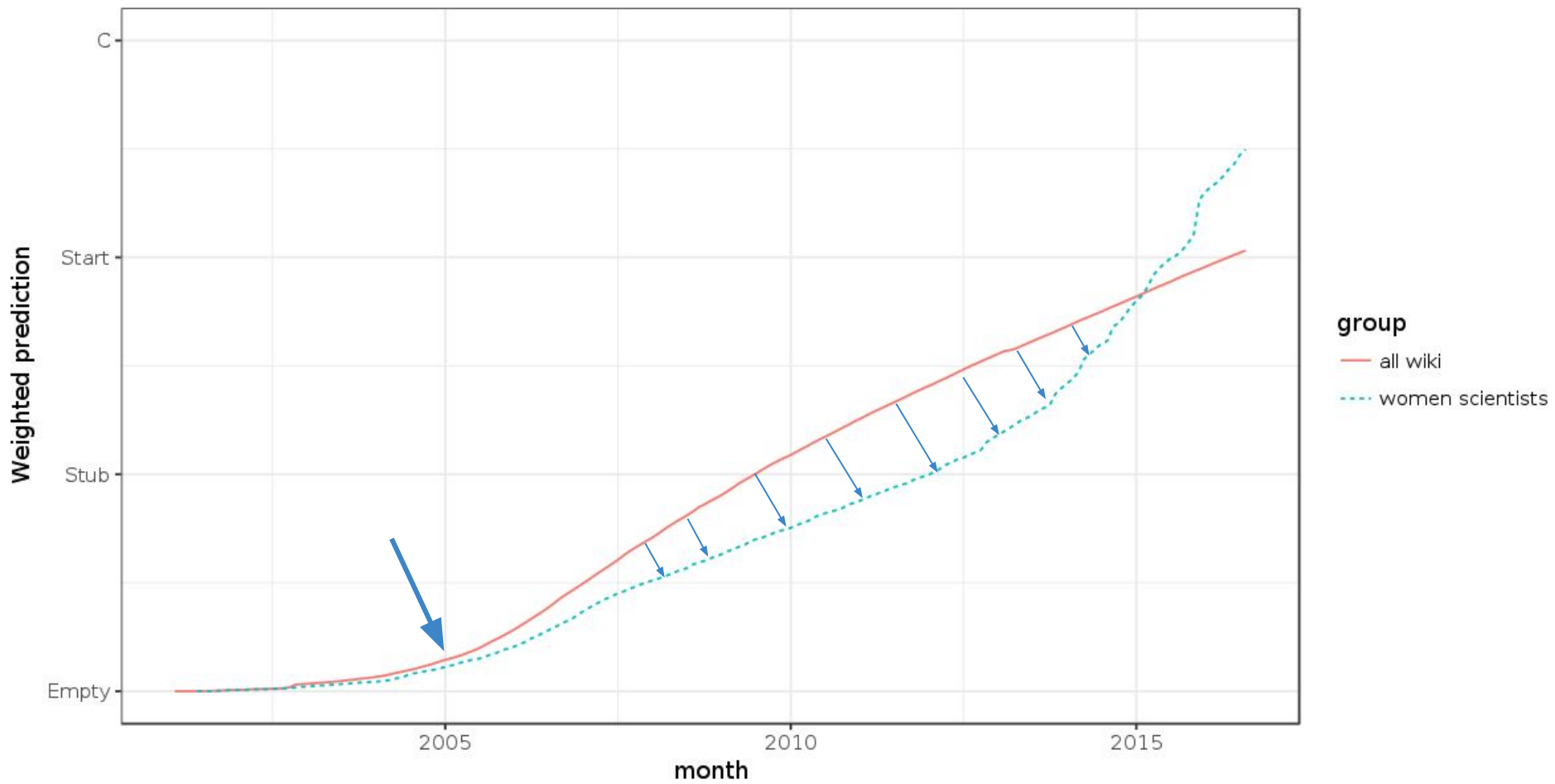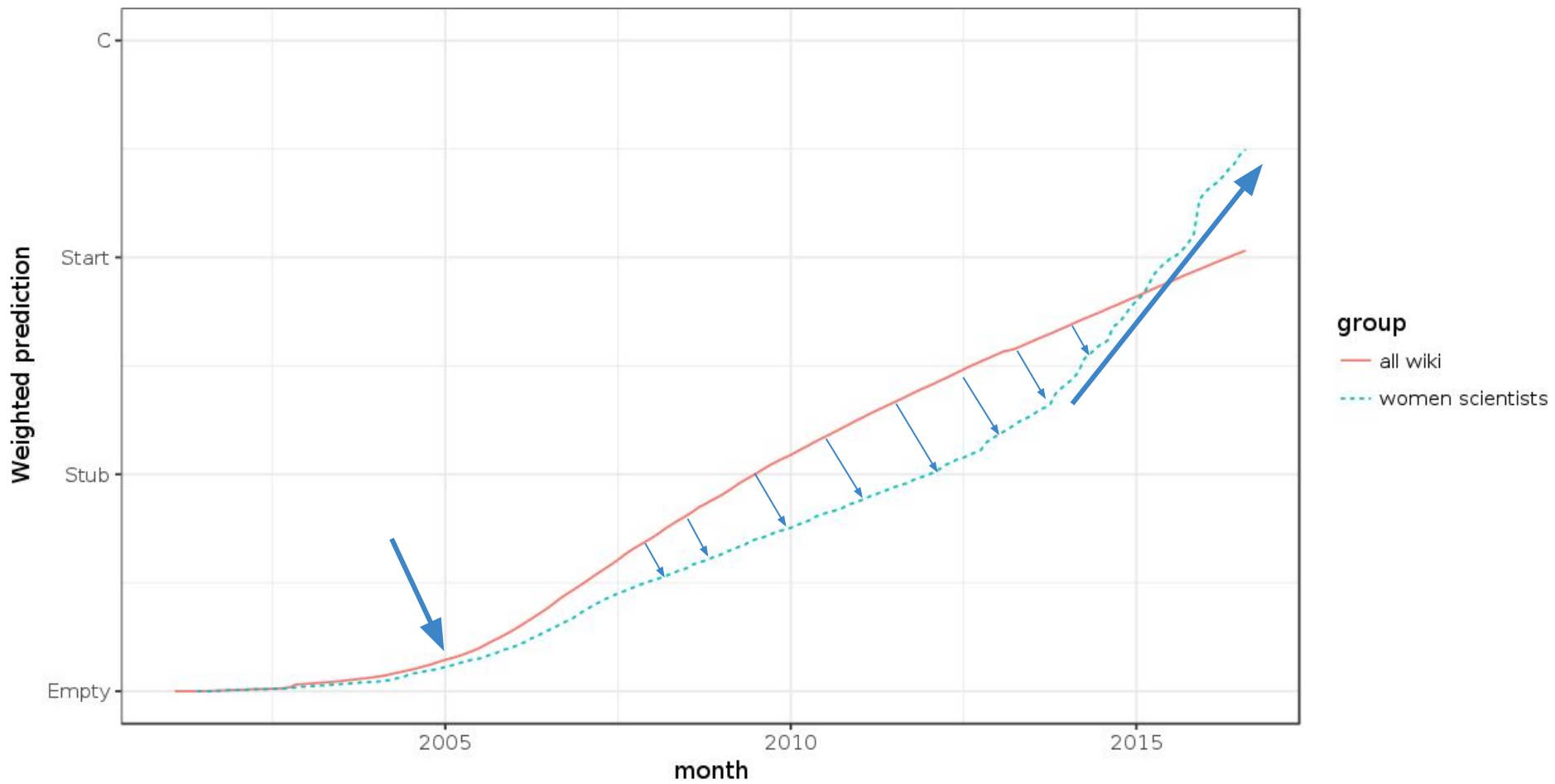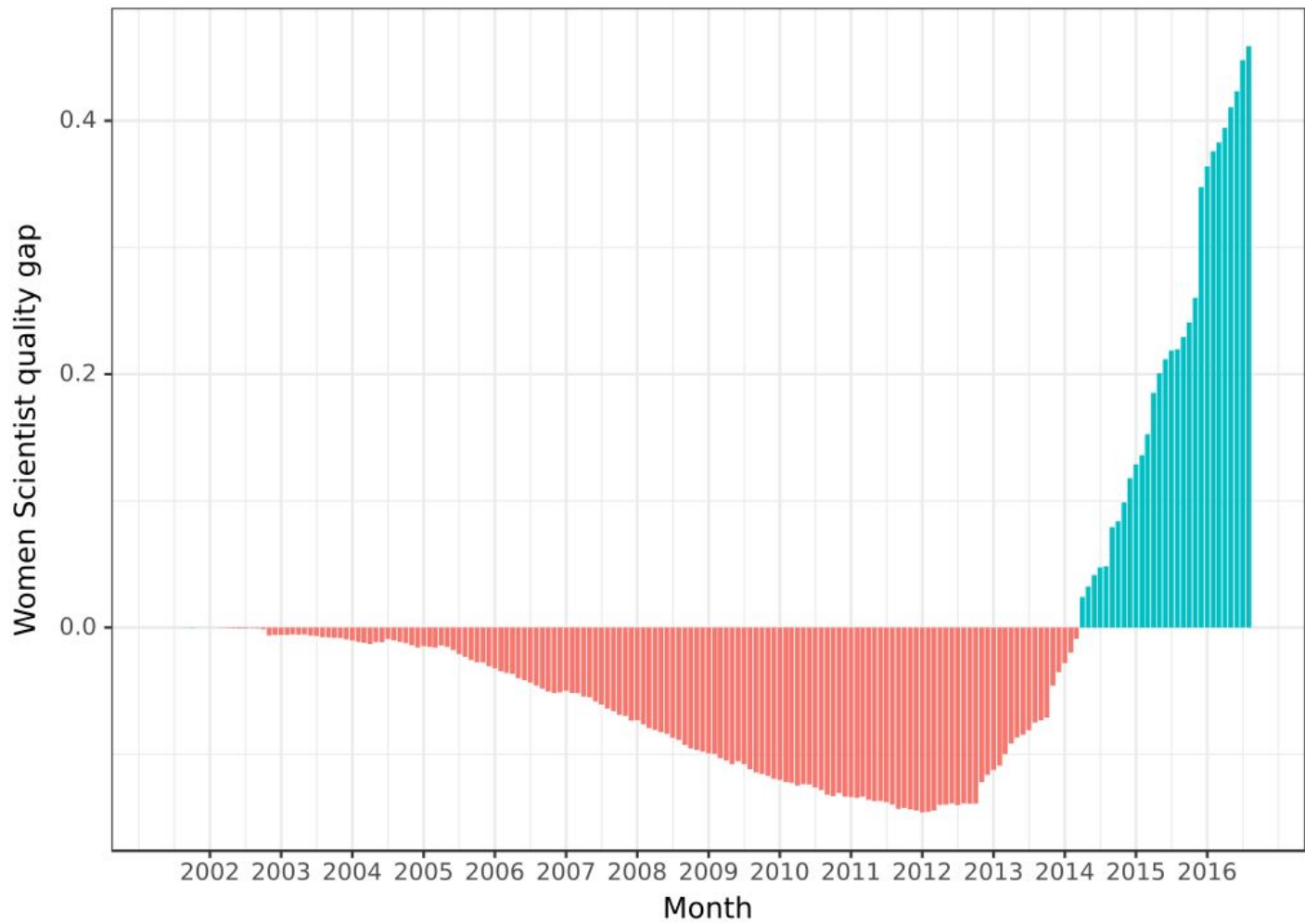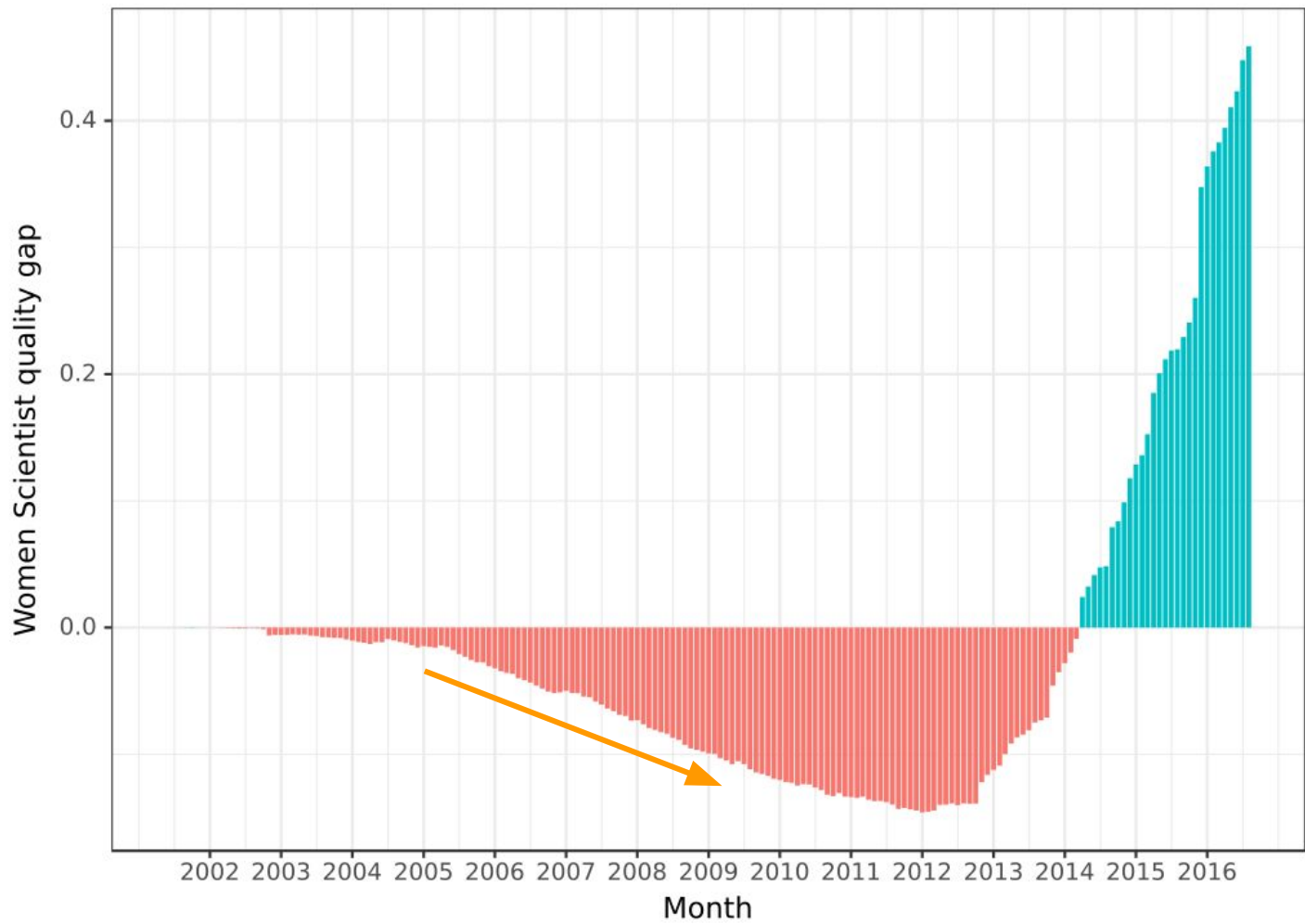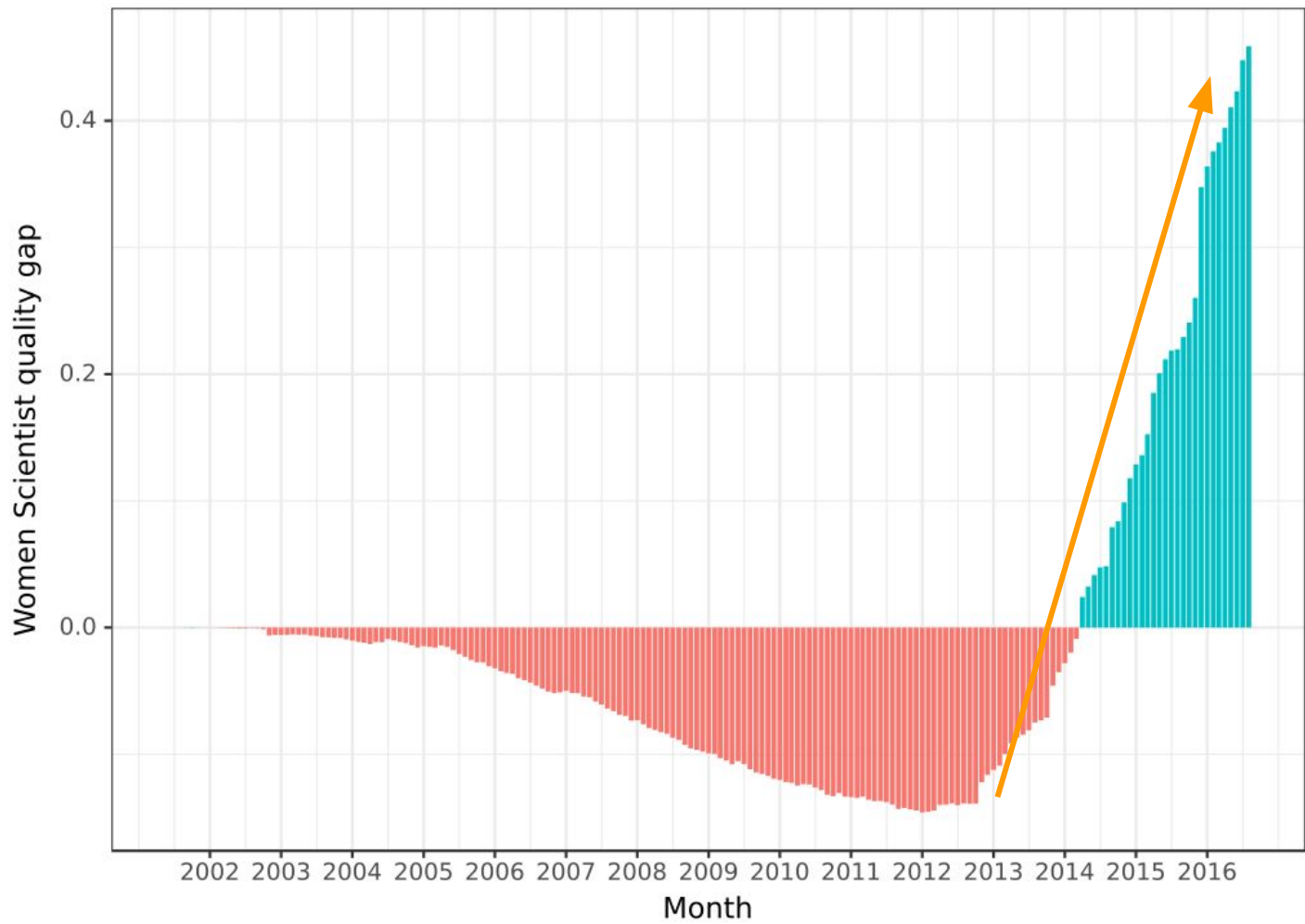
[[:mw:Wikimedia Research/Showcase]] (Dec. 2016)
"English Wikipedia Quality Dynamics and the Case of WikiProject Women Scientists"

Siko Bouterse @sikob · 21 Dec 2016

Wow! @halfak's research shows big increase in quality of women scientist articles starting in mid-2013. I think the reason is @keilanawiki

> **Dario Taraborelli** @ReaderMeter
>
> How much high-quality content are gender-focused programs adding to Wikipedia? Watch @halfak's upcoming talk today twitter.com/WikiResearch/s…

1   5   6

Dario Taraborelli @ReaderMeter · 21 Dec 2016

hadn't thought of @keilanawiki's work <facepalm>, totally makes sense. @halfak

2   2

WIKIMEDIA
FOUNDATION

Adam Hyland @therealprotonk · 21 Dec 2016
we also had to adjust our estimates of student efficacy in the sciences due to the @keilanawiki effect. :)

↩ 2   ↻   ♥ 2

Emily Temple-Wood
@keilanawiki

Following

Replying to @therealprotonk @halfak and 2 others

I've always wanted an effect named after me. The Keilana Effect has a pretty good ring, no? ;P

LIKES
4

Aaron Halfaker
@halfak

I'm down. Sounds like a good term to me.

2:34 PM - 21 Dec 2016

1 Like

1   1

Tweet your reply

Yuvi Panda @yuvipanda · 21 Dec 2016
Replying to @halfak @keilanawiki and 3 others
I second this.

1   1

Dario Taraborelli @ReaderMeter · 21 Dec 2016
quick, let's get the term covered in some reliable source.

2   2

Keilana's work

Ann Bishop (biologist)

Ann Bishop (19 December 1899 – 7 May 1990) was a British biologist from Girton College at the University of Cambridge and a Fellow of the Royal Society, one of the few female Fellows of the Royal Society. She was born in Manchester but stayed at Cambridge for the vast majority of her professional life. Her specialties were protozoology and parasitology; early work with ciliate parasites, including the one responsible for blackhead disease in the domesticated turkey, lay the groundwork for her later research. While working towards her doctorate, Bishop studied parasitic amoebae and examined potential chemotherapies for the treatment of amoebic diseases including amoebic dysentery.

Her best known work was a comprehensive study of *Plasmodium*, the malaria parasite, and investigation of various chemotherapies for the disease. Later she studied drug resistance in this parasite, research that proved valuable to the British military in World War II. She discovered the potential for cross-resistance in these parasites during that same period. Bishop also discovered the protozoan *Pseudotrichomonas keilini* and

Ann Bishop	
Born	19 December 1899 Manchester, England

Women in science

Women have made significant contributions to science from the earliest times. Historians with an interest in gender and science have illuminated the scientific endeavors and accomplishments of women, the barriers they have faced, and the strategies implemented to have their work peer-reviewed and accepted in major scientific journals and other publications. The historical, critical and sociological study of these issues has become an academic discipline in its own right.

The involvement of women in the field of medicine occurred in several early civilizations and the study of natural philosophy in ancient Greece was open to women. Women contributed to the proto-science of alchemy in the first or second centuries AD. During the Middle Ages, convents were an important place of education for women, and some of these communities provided opportunities for women to contribute to scholarly research. While the eleventh century saw the emergence of the first universities, women were, for the most part, excluded from university education.[1] The attitude to educating women in medical fields in Italy

"Woman teaching geometry"
Illustration at the beginning of a medieval translation of Euclid's *Elements* (c. 1310 AD)

WikiProject Women Scientists invites your participation

By Sarah Stierch

December 18th, 2012

Starting a WikiProject is a pretty big undertaking, what with sorting out the templates, tagging thousands of articles, and recruiting new members. But I have to say, starting WikiProject Women Scientists with Sarah Stierch has been one of the most rewarding experiences I've had in my 5 1/2 years contributing to Wikipedia. I had never consciously noticed the gender gap until it was pointed out to me that, as a female Wikipedian, I was a pretty rare commodity. That was a bit of a wake-up call, and I started to think about the systemic bias inherent in Wikipedia, a reference work largely

Starting a WikiProject is a pretty big undertaking, what with sorting out the templates, tagging thousands of articles, and recruiting new members. But I have to say, starting WikiProject Women Scientists with Sarah Stierch has been one of the most rewarding experiences I've had in my 5 1/2 years contributing to Wikipedia.

Grants:IEG/Women Scientists Workshop Development

< Grants:IEG

🔑 This project is funded by an **Individual Engagement Grant**

proposal	people	timeline & progress	finances	midpoint report	final report

🐱 This Individual Engagement Grant is **renewed**

renewal scope   timeline & progress   finances   midpoint report   final report

Project idea [edit]

This project is aiming to create a new model for bringing women into the Wikimedia movement and creating more content to fill the coverage gap with topics related to women, especially biographies of women. Building on the model of the workshops I am

status: **selected**

project:

Women Scientists Workshop Development

"This has been a long hard road but I've greatly enjoyed engaging with the grants team and the grants."

Keilana starts article
about Ann Bishop

WikiProject
Women Scientists
Announcement

Siko Bouterse @sikob · 21 Dec 2016

Wow! @halfak's research shows big increase in quality of women scientist articles starting in mid-2013. I think the reason is @keilanawiki

> **Dario Taraborelli** @ReaderMeter
>
> How much high-quality content are gender-focused programs adding to Wikipedia? Watch @halfak's upcoming talk today twitter.com/WikiResearch/s…

↩ 1   ⟲ 5   ♥ 6

Dario Taraborelli @ReaderMeter · 21 Dec 2016

hadn't thought of @keilanawiki's work <facepalm>, totally makes sense. @halfak

↩ 2   ⟲   ♥ 2

WIKIMEDIA
FOUNDATION

Adam Hyland @therealprotonk · 21 Dec 2016

we also had to adjust our estimates of student efficacy in the sciences due to the @keilanawiki effect. :)

⤷ 2   ⟳   ♥ 2

Emily Temple-Wood
@keilanawiki

Following

Replying to @therealprotonk @halfak and 2 others

I've always wanted an effect named after me. The Keilana Effect has a pretty good ring, no? ;P

LIKES
4

Aaron Halfaker
@halfak

Replying to @keilanawiki @therealprotonk and 2 others

I'm down. Sounds like a good term to me.

2:34 PM - 21 Dec 2016

1 Like

💬 1   🔁   ♡ 1   ᵢₗᵢ

Tweet your reply

Yuvi Panda @yuvipanda · 21 Dec 2016
Replying to @halfak @keilanawiki and 3 others
I second this.

💬 1   🔁   ♡ 1   ✉

Dario Taraborelli @ReaderMeter · 21 Dec 2016
quick, let's get the term covered in some reliable source.

💬 2   🔁   ♡ 2   ✉

The International Symposium on Open Collaboration

Galway, Ireland
Aug, 22nd

Interpolating Quality Dynamics in Wikipedia and Demonstrating the Keilana Effect

Aaron Halfaker
Wikimedia Research
San Francisco, California, USA
ahalfaker@wikimedia.org

ABSTRACT
For open, volunteer generated content like Wikipedia, quality is a prominent concern. To measure Wikipedia's quality, researchers have historically relied on expert evaluation or assessments of article quality by Wikipedians themselves. While both of these methods have proven effective for answering many questions about Wikipedia's quality and processes, they are both problematic: expert evaluation is expensive and Wikipedian quality assessments are sporadic and unpredictable. Studies that explore Wikipedia's quality level or the processes that result in quality improvements have only examined small snapshots of Wikipedia and often rely on complex propensity models to deal with the unpredictable nature of Wikipedians' own assessments. In this paper, I describe a method for measuring article quality in Wikipedia historically and at a finer granularity than was previously possible. I use this method to demonstrate an important coverage dynamic in Wikipedia (specifically, articles about women scientists) and offer this method, dataset, and open API to the research community studying Wikipedia quality dynamics.

Wikipedia has advanced our understanding of the open encyclopedia's quality and the processes by which crowds of volunteers can manage such an information artifact.

Our first major leaps in understanding of Wikipedia's quality dynamics happened around the time that Jim Giles published a report in *Nature* (2005)[7] that surprised the world. This seminal report showed that Wikipedia's coverage of scientific content compared favorably (and in some ways, better) than dominant, traditional, print-based encyclopedias. Since that surprising result was published, researchers have been pushing toward greater understanding of how open, volunteer processes could have generated such a high quality information resource.

While we do know a lot about quality dynamics in Wikipedia, there are still many questions that remain. Where are Wikipedia's coverage gaps? What types of editing patterns lead to efficient quality improvements? These questions are important for the science and the practices of Wikipedians–the volunteers who write and curate the encyclopedia's content. In this paper, I detail the development of a measurement strategy and the release of a public dataset that I believe will make answering these questions far easier than ever before. In the following sections, I'll summarize the state of the art with regards to quality dynamics and measurement in Wikipedia,

ACM Classification Keywords
C.4 PERFORMANCE OF SYSTEMS: Measurement techniques. Modeling techniques

$$\text{weighed sum} = \sum_{c \in C} I(c)P(c)$$

The International
Symposium on Open
Collaboration

Galway, Ireland

Aug, 22nd

Demonstrating the Keilana Effect

Aaron Halfaker
Wikimedia Research
San Francisco, California, USA
ahalfaker@wikimedia.org

ABSTRACT
For open, volunteer generated content like Wikipedia, quality is a prominent concern. To measure Wikipedia's quality, researchers have historically relied on expert evaluation or assessments of article quality by Wikipedians themselves. While both of these methods have proven effective for answering many questions about Wikipedia's quality and processes, they are both problematic: expert evaluation is expensive and Wikipedian quality assessments are sporadic and unpredictable. Studies that explore Wikipedia's quality level or the processes that result in quality improvements have only examined small snapshots of Wikipedia and often rely on complex propensity models to deal with the unpredictable nature of Wikipedians' own assessments. In this paper, I describe a method for measuring article quality in Wikipedia historically and at a finer granularity than was previously possible. I use this method to demonstrate an important coverage dynamic in Wikipedia (specifically, articles about women scientists) and offer this method, dataset, and open API to the research community studying Wikipedia quality dynamics.

ACM Classification Keywords
C.4 PERFORMANCE OF SYSTEMS: Measurement techniques, Modeling techniques

Wikipedia has advanced our understanding of the open encyclopedia's quality and the processes by which crowds of volunteers can manage such an information artifact.

Our first major leaps in understanding of Wikipedia's quality dynamics happened around the time that Jim Giles published a report in *Nature* (2005)[7] that surprised the world. This seminal report showed that Wikipedia's coverage of scientific content compared favorably (and in some ways, better) than dominant, traditional, print-based encyclopedias. Since that surprising result was published, researchers have been pushing toward greater understanding of how open, volunteer processes could have generated such a high quality information resource.

While we do know a lot about quality dynamics in Wikipedia, there are still many questions that remain. Where are Wikipedia's coverage gaps? What types of editing patterns lead to efficient quality improvements? These questions are important for the science and the practices of Wikipedians–the volunteers who write and curate the encyclopedia's content. In this paper, I detail the development of a measurement strategy and the release of a public dataset that I believe will make answering these questions far easier than ever before. In the following sections, I'll summarize the state of the art with regards to quality dynamics and measurement in Wikipedia,

$$\text{weighed sum} = \sum_{c \in C} I(c)P(c)$$

Int...

	n	ROC-AUC	acc (within-1)	Stub	Start	C	B	GA	FA
Stub	4968	98.5%	85.5% (99.2%)	4247	685	27	9	0	0
Start	4982	91.2%	64.3% (91.5%)	600	3205	754	358	58	7
C	4994	86.6%	48.9% (86.1%)	44	870	2443	986	558	93
B	4990	84.1%	40.3% (79.8%)	51	617	1258	2012	710	342
GA	5000	92.1%	62.7% (93.3%)	1	19	313	304	3135	1228
FA	4454	96.1%	77.4% (94.4%)	5	2	23	220	757	3447

Wikipedians' own assessments. In this paper, I describe a method for measuring article quality in Wikipedia historically and at a finer granularity than was previously possible. I use this method to demonstrate an important coverage dynamic in Wikipedia (specifically, articles about women scientists) and offer this method, dataset, and open API to the research community studying Wikipedia quality dynamics.

While we do know a lot about quality dynamics in Wikipedia, there are still many questions that remain. Where are Wikipedia's coverage gaps? What types of editing patterns lead to efficient quality improvements? These questions are important for the science and the practices of Wikipedians–the volunteers who write and curate the encyclopedia's content. In this paper, I detail the development of a measurement strategy and the release of a public dataset that I believe will make answering these questions far easier than ever before. In the following sections, I'll summarize the state of the art with regards to quality dynamics and measurement in Wikipedia,

ACM Classification Keywords
C.4 PERFORMANCE OF SYSTEMS: Measurement techniques. Modeling techniques

"The Keilana Effect"

Women Scientist quality gap

Keilana starts article about Ann Bishop

WikiProject Women Scientists Announcement

group

— all wiki

···· visual arts

– – women scientists

Quality of articles & proportion of FAs

Original research limitations

Take aways

Grants:IEG/Women Scientists Workshop Development

< Grants:IEG

This project is funded by an **Individual Engagement Grant**

proposal	people	timeline & progress	finances	midpoint report	final report

This Individual Engagement Grant is **renewed**

renewal scope	timeline & progress	finances	midpoint report	final report

Project idea [edit]

This project is aiming to create a new model for bringing women into the Wikimedia movement and creating more content to fill the coverage gap with topics related to women, especially biographies of women. Building on the model of the workshops I am

status: **selected**

project:

Women Scientists Workshop Development

Grants:IEG/Women Scientists Workshop Development

< Grants:IEG

This project is funded by an **Individual Engagement Grant**

proposal	people	timeline & progress	finances	midpoint report	final report

This Individual Engagement Grant is **renewed**

renewal scope   timeline & progress   finances   midpoint report   final report

Project idea [edit]

This project is aiming to create a new model for bringing women into the Wikimedia movement and creating more content to fill the coverage gap with topics related to women, especially biographies of women. Building on the model of the workshops I am

status: **selected**

project:

Women Scientists Workshop Development

Are there other projects like Keilana's?

With ORES, we can find other inflection points.

WIKIMEDIA
F O U N D A T I O N

Siko Bouterse
@sikob

Replying to @ReaderMeter @halfak @keilanawiki

never doubt that one person with good support and community organizing can make a big difference!

LIKES
6

2:15 PM - 21 Dec 2016

1

6

Emily Temple-Wood
@keilanawiki

Replying to @sikob @halfak @ReaderMeter

Thanks for giving an untested weirdo kid you'd met once a chance, Siko. <3

LIKES
4

2:17 PM - 21 Dec 2016

1

4

What's next?

WIKIMEDIA
FOUNDATION

What's next?

- Analyze more content gaps

Historical article quality of "WikiProject India" in English Wikipedia

Click to add description...

SQL

```sql
SELECT
    timestamp,
    SUM(weighted_sum) AS weighed_sum,
    SUM(prediction = "Stub") AS stub_n,
    SUM(prediction = "Start") AS start_n,
    SUM(prediction = "C") AS c_n,
    SUM(prediction = "B") AS b_n,
    SUM(prediction = "GA") AS ga_n,
    SUM(prediction = "FA") AS fa_n
FROM templatelinks
INNER JOIN page AS talk ON
    talk.page_id = tl_from AND
    talk.page_namespace = 1
INNER JOIN page AS article ON
    article.page_namespace = 0 AND
    article.page_title = talk.page_title
INNER JOIN u2041__ores_p.monthly_wp10_enwiki AS wp10 ON
    wp10.page_id = article.page_id
WHERE
    tl_namespace = 10 AND
    tl_title = "WikiProject_India"
GROUP BY wp10 timestamp;
```

Query status: **complete**

Resultset (185 rows)

Show [100 ▾] entries

Search: [_____]

| timestamp	weighed_sum	stub_n	start_n	c_n	b_n	ga_n	fa_n
20010401000000	2.085	6	0	0	0	0	0
20010501000000	2.595	7	0	0	0	0	0
20010601000000	8.246	16	1	0	0	0	0
20010701000000	8.246	16	1	0	0	0	0
20010801000000	8.67	17	1	0	0	0	0
20010901000000	22.547	31	2	0	2	0	0
20011001000000	33.963	38	5	0	4	0	0
20011101000000	58.314	66	12	0	4	0	0
20011201000000	87.536	92	21	0	7	0	0
20020101000000	100.129	103	26	0	7	0	0
20020201000000	109.516	115	25	0	9	0	0
20020301000000	144.794	166	33	0	7	0	0
20020401000000	171.205	186	42	0	8	0	0
20020501000000	183.61	193	47	0	9	0	0
20020601000000	189.552	203	47	0	9	0	0
20020701000000	207.828	212	56	0	11	0	0
20020801000000	218.527	221	58	0	13	0	0

What's next?

- Analyze more content gaps
- Support more wikis

What's next?

- Analyze more content gaps
- Support more wikis
  - English Wikipedia
  - French Wikipedia
  - Russian Wikipedia
  - Turkish Wikipedia
  - Wikidata
  - … your wiki? Talk to me!

Thanks!

SCIENCE

SCIENCE

SCIENCE

SCIENCE

SCIENCE

SCIENCE

SCIENCE
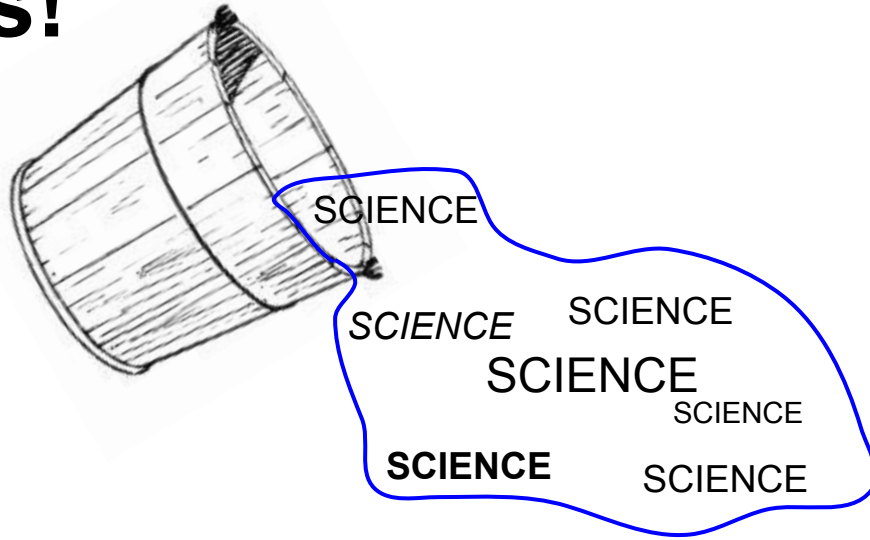
[[WP:WikiProject Women scientists]]

Aaron Halfaker

ahalfaker@wikimedia.org

User:EpochFail

https://twitter.com/halfak

Emily Temple-Wood

User:Keilana

https://twitter.com/keilanawiki