

ORES appropriation and reflection.

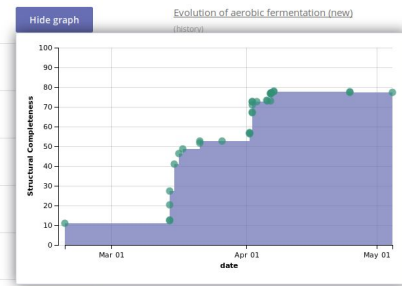
PatruBOT está **detenido de forma indefinida** desde el 27 de marzo de 2018, debido al número de falsos positivos y la imposibilidad de resolverlos a corto plazo. Según el tiempo disponible, en el futuro podría



Improvements [\[edit \]](#)

Diff id	Damaging	Old score	Score1	Score2	Score3	1st improv.	2nd	3rd	Overall
210649690	No	91%	90%	91%	91%	+61%	-25%	+5%	+91%
237996719	No	84%	71%	63%	60%	+13%	-8%	-3%	+24%
242037491	No	95%	40%	74%	71%	+42%	-28%	-3%	+24%
251530750	No	91%	55%	50%	55%	+30%	-1%	+1%	+30%
253045099	No	99%	89%	70%	70%	+10%	+11%	-8%	+29%
257856552	No	91%	30%	4%	1%	+61%	+20%	+3%	+80%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269096263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	40%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

Aaron Halfaker
Scoring Platform @ Wikimedia



Aaron Halfaker

Principal Research Scientist,
Wikimedia Foundation

Think big. Measure what you can. Build better technologies.



About me

Hi. I'm [Aaron Halfaker](#). I'm a scientist. See [projects](#) and [publications](#) below. I've been a Wikipedian since 2008. I mostly build tools and run studies, but I make edits where I can. In 2011, I started working with the Wikimedia Foundation as a [research scientist](#). This is

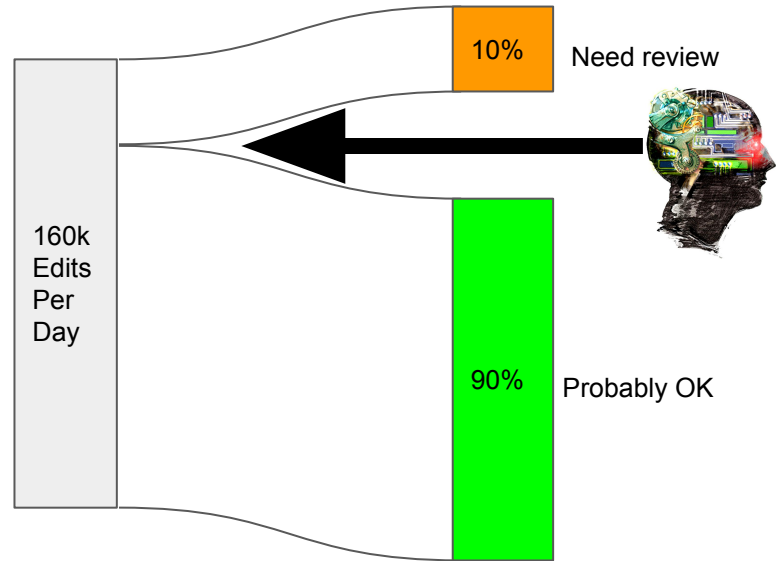
My work

My job is to build understanding about and support for the socio-technical fabric of the Wikimedia movement. I tend to focus on our computer mediated spaces (Wikipedia, Commons, Wikidata, Wikisource, etc.) and quality dynamics (patrolling, curation,

Outline

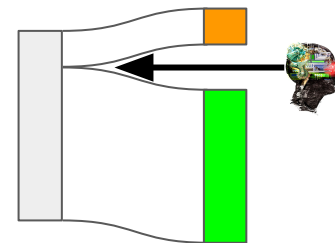
Outline

1. ORES' vision: Efficiency and innovation



Outline

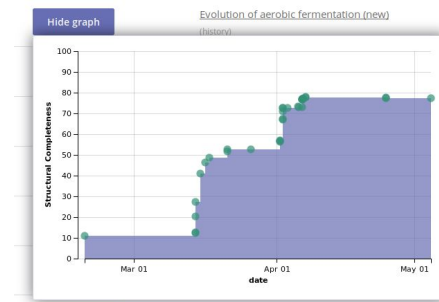
1. ORES' vision: Efficiency and innovation



2. Appropriation and reflection

- Wikidata's "Report mistakes" page
- PatruBOT's bad behavior
- What do you mean, "quality"?

PatruBOT está **detenido de forma indefinida** desde el 27 de marzo de 2018, debido al número de falsos positivos y la imposibilidad de resolverlos a corto plazo. Según el tiempo disponible, en el futuro podría ser reactivado con un nuevo algoritmo o si ORES facilita que sea más efectivo, en función también del interés que pueda tener la comunidad.



Improvements [\[edit\]](#)

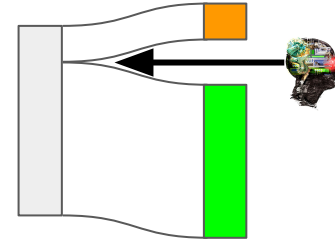
Diff id	Damaging	Old score	Score1	Score2	Score3	1st improv.	2nd	3rd	Overall
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	-1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270083221	No	92%	98%	40%	48%	+6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

Outline

1. ORES' vision: Efficiency and innovation

2. Appropriation and reflection

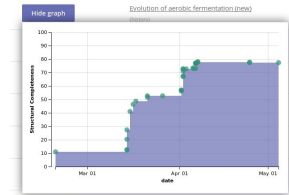
3. Discussion



PatruBOT está **detenido de forma indefinida** desde el 27 de marzo de 2018, debido al número de falsos positivos y la imposibilidad de resolverlos a corto plazo. Según el tiempo disponible, en el futuro podría



Impocventados	1001								
ID#	Damage	Old score	Score	Score	Score	Score	Def #	Def #	Overall
21000490	No	61%	20%	73%	8%	-61%	-12%	10%	-61%
21000500	No	64%	17%	63%	80%	-12%	-4%	-3%	-14%
24001943	No	99%	44%	74%	71%	-44%	20%	13%	-14%
21002792	No	61%	62%	60%	61%	-1%	1%	1%	-14%
21000499	No	99%	69%	79%	70%	-20%	-13%	-6%	-20%
21000462	No	61%	20%	6%	2%	-61%	-10%	0%	-60%
20007020	No	61%	62%	64%	73%	-6%	13%	0%	-14%
20007027	No	61%	64%	79%	80%	-12%	-13%	1%	-14%
20000447	No	99%	100%	10%	71%	2%	-20%	-3%	-19%
20000501	No	99%	100%	61%	70%	0%	-10%	-10%	-19%
20000504	No	99%	69%	6%	43%	0%	-10%	-10%	-14%
20000504	No	99%	69%	6%	43%	0%	-10%	-10%	-14%
20000504	No	99%	100%	64%	71%	0%	-1%	-13%	-19%
20000503	No	99%	69%	6%	84%	-1%	-14%	-10%	-19%
21000522	No	87%	69%	6%	-46%	6%	-16%	6%	-14%
21001703	No	84%	67%	6%	64%	0%	6%	-14%	-19%

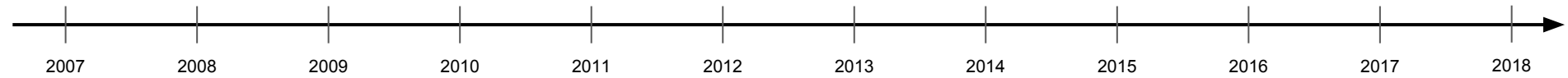


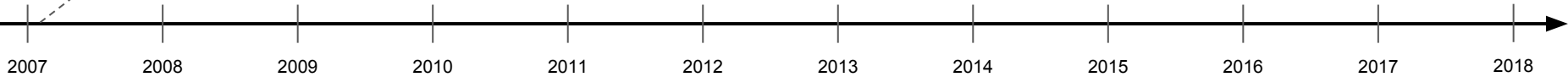
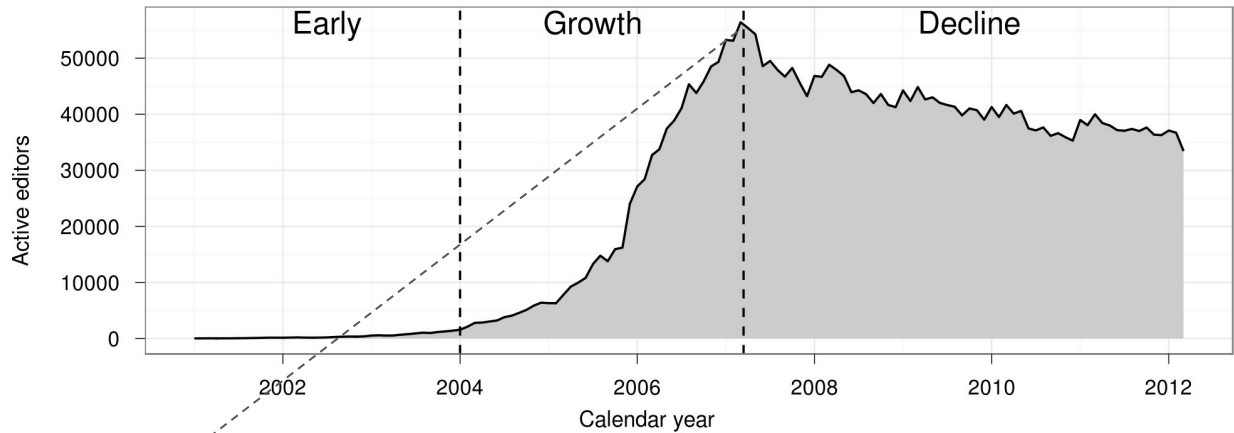
Part 1: ORES' vision

What's ORES?

~~What's ORES?~~

Why is ORES?





A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

Aaron Halfaker
GroupLens Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
halfak@cs.umn.edu

Aniket Kittur Robert Kraut
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
(nkittur, robert.kraut}@cs.cmu.edu

John Riedl
GroupLens Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
riedl@cs.umn.edu

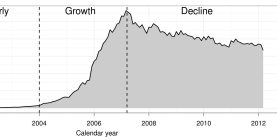
ABSTRACT

Wikipedia is a highly successful example of what mass collaboration in an informal peer review system can accomplish. In this paper, we examine the role that the quality of the contributions, the experience of the contributors and the ownership of the content play in the decisions over which contributions become part of Wikipedia and which ones are rejected by the community. We introduce and justify a versatile metric for automatically measuring the quality of a contribution. We find little evidence that experience helps contributors avoid rejection. In fact, as they gain experience, contributors are even more likely to have their work rejected. We also find strong evidence of ownership behaviors in practice despite the fact that ownership of content is discouraged within Wikipedia.

experience of these teams of volunteers and by their feelings of ownership.

One of the key components of Wikipedia is the review process through which contributions are rejected or accepted. This process is informal and, to an outsider, appears disorganized, with its reliance on watchlists and Internet Relay Chat channels. However, the review process is robust and effective in practice: 42% of vandalistic contributions are repaired within one view and 70% within ten views [1].

Many other systems use peer review, though usually in a more structured manner. For instance, conferences typically have three peers of the authors read each submitted article to decide whether it should be accepted or rejected. Similar peer review systems include NSF grant panels and arts competitions. The goal of these review processes is to



2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018

The Singularity is Not Near: Slowing Growth of Wikipedia

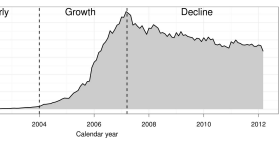
Bongwon Suh, Gregorio Convertino, Ed H. Chi, Peter Pirolli
Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, CA, 94304
+1 (650)912-4800
(suh, convertino, echi, pirolli}@parc.com

ABSTRACT

Prior research on Wikipedia has characterized the growth in content and editors as being fundamentally exponential in nature, extrapolating current trends into the future. We show that recent editing activity suggests that Wikipedia growth has slowed, and perhaps plateaued, indicating that it may have come against its limits to growth. We measure growth, population shifts, and patterns of editor and administrator activities, contrasting these against past results where possible. Both the rate of page growth and editor growth has declined. As growth has declined, there are indicators of increased coordination and overhead costs, exclusion of newcomers, and resistance to new edits. We discuss some possible explanations for these new developments in Wikipedia including decreased opportunities for sharing existing knowledge and increased bureaucratic inertia on the socio-technical system itself.

suggested that Wikipedia shows such exponential growth and that growth is mainly spurred by exponential growth in contributing editors [1].

The existing trends of exponential growth in digital technologies were the basis for Kurzweil's [7] argument that biological evolution and technological evolution follows a law of accelerating returns (i.e. exponential or even super-exponential growth). This led to the notion of the "Singularity" – a point in the near future when technological change becomes "so rapid and profound that it represents a rupture in the fabric of human history." We argue that Wikipedia, one of the world's largest knowledge aggregators, does indeed mirror the growth of natural populations, but, following Darwin [7], we suggest that this growth becomes increasingly constrained and limited, and under those conditions there will be increased evidence of competition and dominance. In this paper, we present data that challenges the notion that



A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

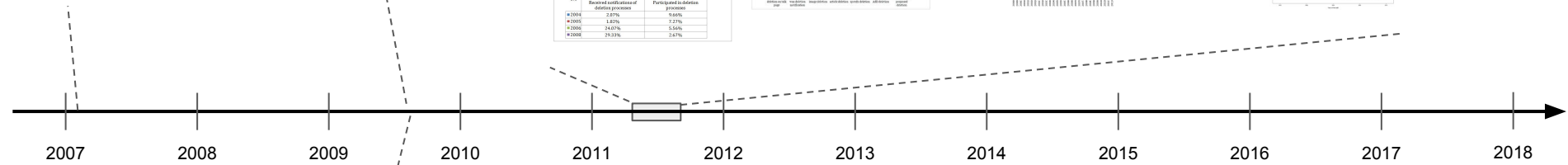
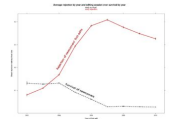
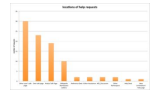
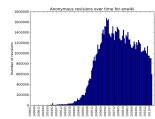
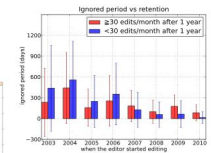
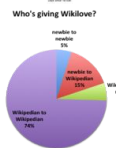
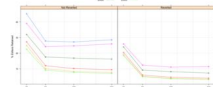
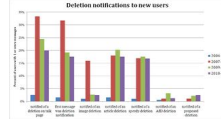
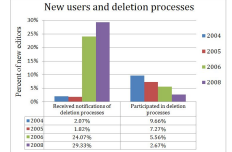
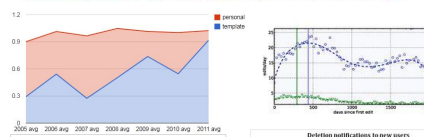
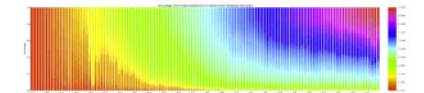
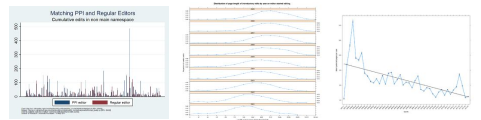
Aaron Halbuak
Cognivity Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
halbuak@cs.umn.edu

Arifket Kitar
Carnegie Mellon University
2000 Forbes Ave.
Pittsburgh, PA 15213
(akitur, robert.kraut)@cs.cmu.edu

Robert Kraut
Carnegie Mellon University
2000 Forbes Ave.
Pittsburgh, PA 15213
robertk@cs.cmu.edu

John Riedl
Carnegie Mellon University
University of Pennsylvania
200 Union St. S.E.
Minneapolis, MN 55455
riedl@cs.umn.edu

ABSTRACT
Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors in the decision-making process play in Wikipedia's edit and delete processes. We investigate and justify a central notion, for successfully reviewing the edit and delete processes, the edit and delete review system before contributions are accepted. In fact, our first experiments demonstrate that users are likely to be more likely to contribute to the system when they are given ownership of the articles they edit. We also find strong evidence of ownership behavior. We present results for edit and delete ownership of content to be accepted within Wikipedia.



The Singularity is Not Near: Slowing Growth of Wikipedia

Borghwan Suh, Ginoago Conventuro, Ed H. Chi, Peter Fritzl
9400 Wilshire Blvd, Suite 1000, Beverly Hills, CA 90212
suh, conventuro, edh, pfritz@parc.com

ABSTRACT
Prior research on Wikipedia has characterized the growth in content and editors as being fundamentally exponential in nature. This paper suggests that Wikipedia growth has slowed and is approaching a singularity. We present evidence for this claim in terms of growth rate, content quality, and ownership. We present evidence that the growth rate is slowing and that the quality of content is declining. We also present evidence that ownership is becoming more concentrated among a smaller number of users. We argue that these findings suggest that Wikipedia is approaching a singularity and that the growth rate will continue to decline. We discuss the implications of these findings for the future of Wikipedia and for other crowdsourcing systems.

Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors in the decision-making process play in Wikipedia's edit and delete processes. We investigate and justify a central notion, for successfully reviewing the edit and delete processes, the edit and delete review system before contributions are accepted. In fact, our first experiments demonstrate that users are likely to be more likely to contribute to the system when they are given ownership of the articles they edit. We also find strong evidence of ownership behavior. We present results for edit and delete ownership of content to be accepted within Wikipedia.

Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors in the decision-making process play in Wikipedia's edit and delete processes. We investigate and justify a central notion, for successfully reviewing the edit and delete processes, the edit and delete review system before contributions are accepted. In fact, our first experiments demonstrate that users are likely to be more likely to contribute to the system when they are given ownership of the articles they edit. We also find strong evidence of ownership behavior. We present results for edit and delete ownership of content to be accepted within Wikipedia.

Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors in the decision-making process play in Wikipedia's edit and delete processes. We investigate and justify a central notion, for successfully reviewing the edit and delete processes, the edit and delete review system before contributions are accepted. In fact, our first experiments demonstrate that users are likely to be more likely to contribute to the system when they are given ownership of the articles they edit. We also find strong evidence of ownership behavior. We present results for edit and delete ownership of content to be accepted within Wikipedia.

Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors in the decision-making process play in Wikipedia's edit and delete processes. We investigate and justify a central notion, for successfully reviewing the edit and delete processes, the edit and delete review system before contributions are accepted. In fact, our first experiments demonstrate that users are likely to be more likely to contribute to the system when they are given ownership of the articles they edit. We also find strong evidence of ownership behavior. We present results for edit and delete ownership of content to be accepted within Wikipedia.

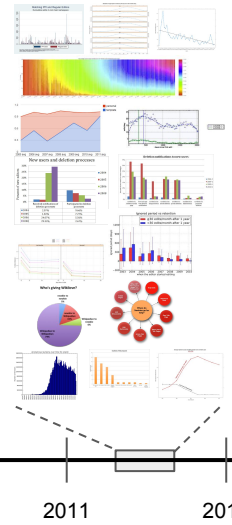
The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline

Aaron Halfaker¹, R.Stuart Geiger², Jonathan T. Morgan³, and John Riedl¹

Abstract

Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer contributors to remain relevant. Wikipedia was created through a tremendous number of contributions by millions of contributors. However, recent research has shown that

WSoR 2011



A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

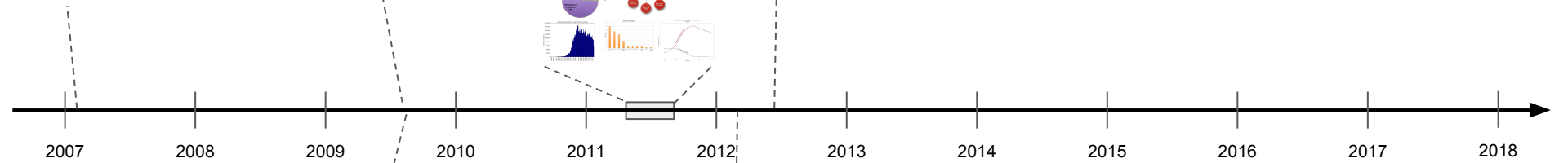
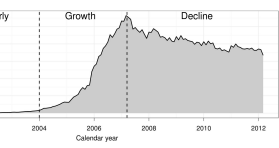
Aaron Halfaker
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 halfak@cs.umn.edu

Arvind Kittur
 Carnegie Mellon University
 5000 Forbes Ave.
 Pittsburgh, PA 15213
 kittur@cmu.edu

Robert Kraut
 Carnegie Mellon University
 5000 Forbes Ave.
 Pittsburgh, PA 15213
 rkraut@cs.cmu.edu

John Riedl
 Carnegie Mellon University
 5000 Forbes Ave.
 Pittsburgh, PA 15213
 riedl@cs.cmu.edu

ABSTRACT
 Wikipedia is a highly successful example of what some call "collaboration" or an internet peer review system run voluntarily. In this paper, we examine the role that the quality of the contributions, the experience of the contributors and the ownership of the contributions play in the discussion article contributions process. We find that the quality of the contributions is a strong predictor of the number of contributions. We also find that the experience of the contributors is a strong predictor of the number of contributions. We find that the ownership of the contributions is a strong predictor of the number of contributions. We also find that the ownership of the contributions is a strong predictor of the number of contributions.



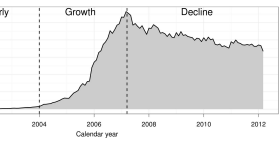
The Singularity is Not Near: Slowing Growth of Wikipedia

Borghvorn Sukh, Gungoro Conventore, Ed H. Chi, Peter Pritzl
 3303 Canyon Rd, Redwood Falls, MN, CA, 56454
 {sukh, conventore, edh, ppritzl}@epcc.com

ABSTRACT
 Wikipedia has experienced the growth in content and activity in a way that is exponentially increasing in nature. This paper examines the growth of Wikipedia and the impact of the singularity hypothesis. We find that the growth of Wikipedia is slowing down and that the singularity hypothesis is not supported by the data. We find that the growth of Wikipedia is slowing down and that the singularity hypothesis is not supported by the data. We find that the growth of Wikipedia is slowing down and that the singularity hypothesis is not supported by the data.



Welcome to the
teahouse
 A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.



A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

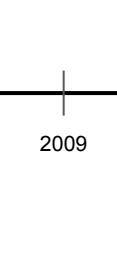
Aaron Halfaker
 Complex Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 halfak@cs.umn.edu

Arvind Kittur
 Complex Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 kittur@cs.umn.edu

Robert Kraut
 Complex Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 robert.kraut@cs.umn.edu

John Riedl
 Complex Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 riedl@cs.umn.edu

ABSTRACT
 Wikipedia is a highly successful example of other users' collaboration in an informal peer review system not available, to the paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors play in determining article quality. We investigate and justify a number of ways for experimentally measuring the quality of a contributor. We find little evidence that experience helps contributors avoid repetition. In fact, on first-time experience, contributors are more likely to have their work removed. We also find strong evidence of ownership behavior. We present results that show that ownership is associated with decreased within-Wikipedia...

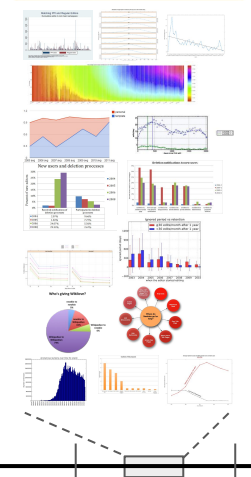


The Singularity is Not Near: Slowing Growth of Wikipedia

Bongwon Suh, Gregorio Conzatti, Ed H. Chi, Peter Fritsch
 3500 Canyon Rd, Redwood City, CA, 94064
 {suh, conzatti, echi, pfritsch}@parc.com

ABSTRACT
 Peer research on Wikipedia has characterized the growth in content and editors as being fundamentally dependent on some underlying growth rate. We show that the rate of growth is slowing and suggest that Wikipedia growth has slowed not because of saturation, but that there are some agents in the system that are slowing growth. We consider agents: individual users, groups of editors, and administrative sub-systems, considering their impact and effects over periods. We show that administrative sub-systems are slowing growth. As groups have declined, there are a number of user experiences that may be slowing down growth. We discuss the implications of these user experiences for Wikipedia and suggest some possible explanations for these user experiences. We discuss the potential implications for slowing growth, including the role of the informal institutions that are the microstructural system that...

WSor 2011





Welcome to the **teahouse**

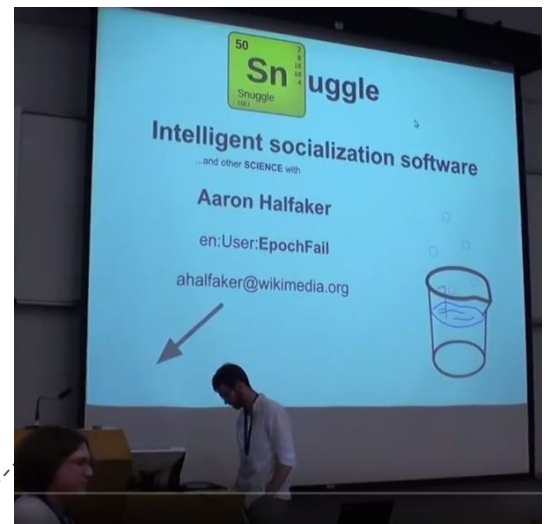
A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.

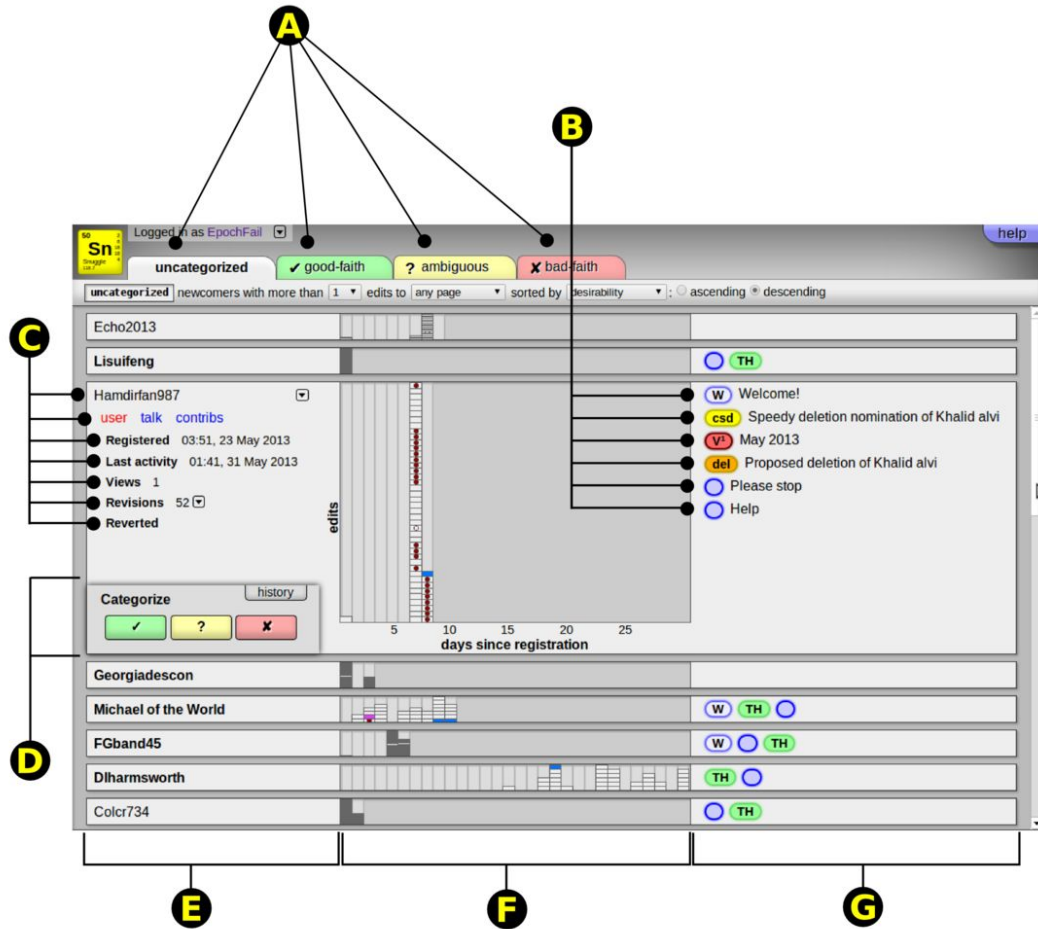
The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity is Causing Its Decline

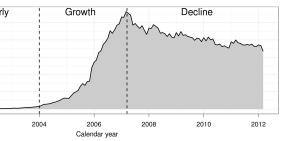
Aaron Halfaker¹, R. Stuart Geiger², Jonathan T. Morgan¹, and John Riedl¹

Accepted for publication: 10/11/12
 © 2012 ACM Publications
 Permission to publish: http://dx.doi.org/10.1145/2201111
 http://dx.doi.org/10.1145/2201111

Abstract
 Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer contributors to remain relevant. Wikipedia was created through a tremendous number of...







A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

Aaron Halfaker
 Complex Systems Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 halfak@cs.umn.edu

Arvind Kittur
 Complex Systems Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 kittur@cs.umn.edu

Robert Kraut
 Complex Systems Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 robert.kraut@cs.umn.edu

John Riedl
 Complex Systems Research
 University of Minnesota
 200 Union St. S.E.
 Minneapolis, MN 55455
 riedl@cs.umn.edu

ABSTRACT
 Wikipedia is a highly successful example of what some call 'crowdsourcing' in an informal peer review system not accessible, in this paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors play in determining which contributions become part of Wikipedia and which ones are rejected by the community. We introduce and justify a complex system for automatically measuring the quality of a contributor. We find little evidence that experience helps contributors avoid rejection. In fact, on first page experience, contributions are more likely to be rejected than ever. We also find strong evidence of ownership behavior. It is a general finding that the main motivation of users to discontinue within Wikipedia.

WSoR 2011

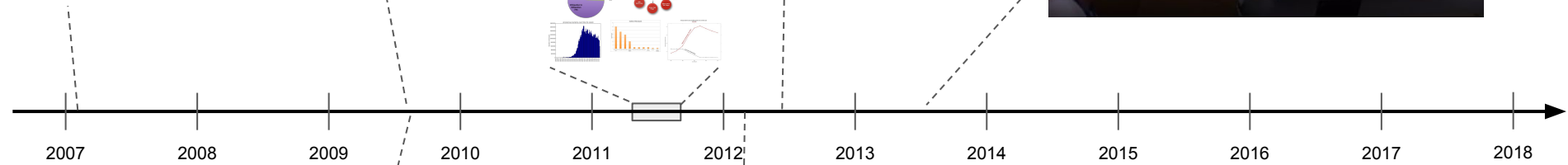
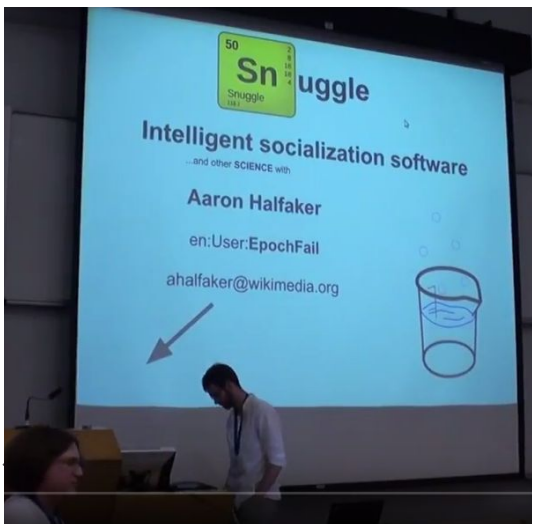


The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity is Causing Its Decline

Aaron Halfaker¹, R. Stuart Geiger², Jonathan T. Morgan¹, and John Riedl¹

Abstract
 Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer contributors to remain relevant. Wikipedia was created through a randomized number of contributors to remain relevant.

Accepted for publication: 10/11/11
 © 2012 ACM Publications
 Permission to publish this article is granted by the ACM Publications Department. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.
 ACM, 475 Washington St., Boston, MA 02116-1099
 http://www.acm.org
 958-0446



The Singularity is Not Near: Slowing Growth of Wikipedia

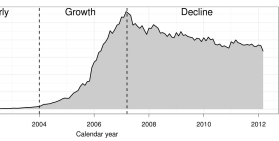
Bongwon Suh, Gregorio Conzatti, Ed H. Chi, Peter Fritsch

Ed H. Chi Research Center
 3500 Canyon Rd. #1000, Palo Alto, CA, 94304
 echi@stanford.edu
 (suh, conzatti, echi, pfritsch)@stanford.edu

ABSTRACT
 Peer research on Wikipedia has characterized the growth in content and editors as being fundamentally dependent on some underlying growth rate. We show that the rate of growth is slowing and suggest that Wikipedia growth has slowed not because of saturation, but because of a change in the underlying growth rate. We measure growth, saturation, and the rate of growth in terms of editor and article counts, and find that the rate of growth is slowing. We find that the rate of growth is slowing because of a change in the underlying growth rate. We find that the rate of growth is slowing because of a change in the underlying growth rate. We find that the rate of growth is slowing because of a change in the underlying growth rate.

Welcome to the **teahouse**

A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.



A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

Aaron Halfaker
Cognitive Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
halfak@cs.umn.edu

Arvind Kittur
Cognitive Research
200 Union St.
Pittsburgh, PA 15213
akittur@cs.cmu.edu

Robert Kraut
Cognitive Research
200 Union St.
Pittsburgh, PA 15213
rkraut@cs.cmu.edu

John Riedl
Cognitive Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
riedl@cs.umn.edu

ABSTRACT
Wikipedia is a highly successful example of other users' collaboration in an internet peer review system not available, to this paper, we examine the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors play in the success of Wikipedia and search new contributors. We include and justify a wide range of metrics for measuring the quality of a contributor. We find little evidence that experience helps contributors avoid repetition, but that, on the other hand, more contributions can only help to find good contributors. We also find strong evidence of ownership behavior. We in general suggest that the main reason of success to be investigated within Wikipedia.

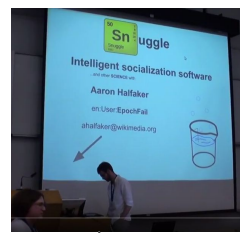
WSoR 2011



The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline

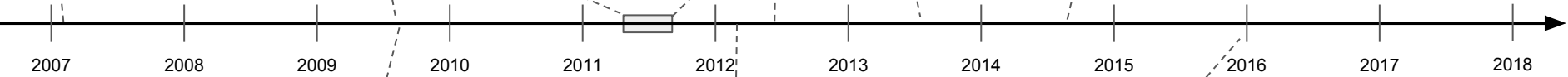
Aaron Halfaker¹, R. Stuart Geiger², Jonathan T. Morgan³, and John Riedl⁴

Abstract
Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer contributors to remain relevant. Wikipedia was created through a series of social...



Wikipedia as a socio-technical system

by Aaron Halfaker



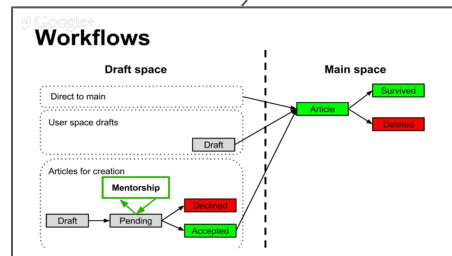
The Singularity is Not Near: Slowing Growth of Wikipedia

Borghese Sukh, Gungoro Conventore, Ed H. Chi, Peter Fritzl

ABSTRACT
Peer research on Wikipedia has characterized the growth in content and editors in large, traditionally unregulated, in-house, encyclopedia systems like the... The new data and existing evidence suggest that Wikipedia growth has slowed and further plateauing, especially in the last few years, appears as likely as growth. We consider growth, Wikipedia article, and patterns of edit and administrator activities, including their impact and trends over periods. Both the historical growth and the current growth are linked. As growth has declined, there is a downward trend in user engagement and, only the decline in user participation for these two developments in Wikipedia article growth appears to be strong evidence of a plateau. We discuss the reasons behind the current plateauing and the historical behavior, with an eye on the socio-technical system that...

Welcome to the teahouse

A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.



ORES

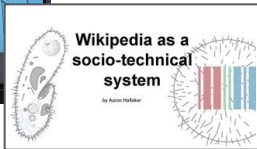
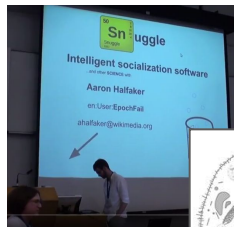
The people's classifier
service
Towards an open model for algorithmic
infrastructure



The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline

Aaron Halfaker¹, R. Stuart Geiger²,
Jonathan T. Morgan³, and John Riedl⁴

Abstract
Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer
contributors to remain relevant. Wikipedia was created through a snowball number



WSoR 2011



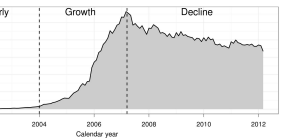
A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

Aaron Halfaker¹, Robert Kraut², John Riedl⁴
Geography Research Center, University of Pittsburgh
200 Union St., S.E., Pittsburgh, PA 15213
halfak@cs.umms.edu, robert.kraut@pitt.edu, riedl@cs.umms.edu

ABSTRACT

Wikipedia is a highly successful example of what some scholars call an informal peer review system on a voluntary basis. In this paper, we examine the role that the quality of the contributors, the experience of the contributors and the amount of time spent in the discussion area contribute to the quality of Wikipedia and search for an underlying model. We include and justify a wide range of variables for our regression analysis. Our results indicate that experience helps contribute to article quality. In fact, we find that each year, contributors who come back to help edit each other are more likely to have higher quality articles. We also find strong evidence of expertise behavior in the peer review process and that many editors of content to be accepted within Wikipedia.

experience of these users is related and by their length of experience. One of the key components of Wikipedia is the review process through which contributions are accepted or rejected. This process is informal and is an essential aspect of Wikipedia, with its reliance on volunteer and informal Editor that thereby, however, the review process is neither well defined nor precisely. We find little evidence that experience helps contribute to article quality. In fact, we find that each year, contributors who come back to help edit each other are more likely to have higher quality articles. We also find strong evidence of expertise behavior in the peer review process and that many editors of content to be accepted within Wikipedia.



Abstract
Wikipedia is a highly successful example of what some scholars call an informal peer review system on a voluntary basis. In this paper, we examine the role that the quality of the contributors, the experience of the contributors and the amount of time spent in the discussion area contribute to the quality of Wikipedia and search for an underlying model. We include and justify a wide range of variables for our regression analysis. Our results indicate that experience helps contribute to article quality. In fact, we find that each year, contributors who come back to help edit each other are more likely to have higher quality articles. We also find strong evidence of expertise behavior in the peer review process and that many editors of content to be accepted within Wikipedia.

The Singularity is Not Near: Slowing Growth of Wikipedia

Borghvorn Subh, Gungoro Conventore, Ed H. Chi, Peter Pirolli
Rice University
3300 Cypress Rd, P.O. Box 4800, CA, 48040
+1 281 357 8400
{subh, conventore, echi, pirolli}@rice.com

ABSTRACT

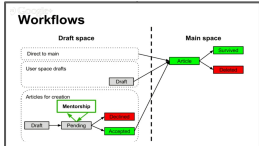
Peer research on Wikipedia has characterized the growth in content and editors in being fundamentally exponential in some respect. However, our new analysis suggests that Wikipedia growth has slowed and perhaps plateaued. We find a clear trend in our data on growth rates. We measure growth rates across a wide range of editors and administrator activities, including those that are not visible to the public. We find that growth rates are slowing across all dimensions. As growth has slowed, there is an increase in the number of editors who are active on a regular basis. We argue that the slowing of growth is a result of the increasing number of editors who are active on a regular basis. We argue that the slowing of growth is a result of the increasing number of editors who are active on a regular basis. We argue that the slowing of growth is a result of the increasing number of editors who are active on a regular basis.

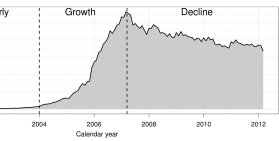
Abstract
Peer research on Wikipedia has characterized the growth in content and editors in being fundamentally exponential in some respect. However, our new analysis suggests that Wikipedia growth has slowed and perhaps plateaued. We find a clear trend in our data on growth rates. We measure growth rates across a wide range of editors and administrator activities, including those that are not visible to the public. We find that growth rates are slowing across all dimensions. As growth has slowed, there is an increase in the number of editors who are active on a regular basis. We argue that the slowing of growth is a result of the increasing number of editors who are active on a regular basis. We argue that the slowing of growth is a result of the increasing number of editors who are active on a regular basis.



Welcome to the teahouse

A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.





A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia

Aaron Halfaker
Cognition Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
halfak@cs.umn.edu

Arvind Khetur
Cognition Research
200 Union St.
Pittsburgh, PA 15213
(khetur.arvind@cs.cmu.edu)

Robert Kraut
Cognition Research
200 Union St.
Pittsburgh, PA 15213
rkraut@cs.cmu.edu

John Riedl
Cognition Research
University of Minnesota
200 Union St. S.E.
Minneapolis, MN 55455
riedl@cs.umn.edu

ABSTRACT
Wikipedia is a highly successful example of open social collaboration in an internet peer review system not available, to the paper, we measure the role that the quality of the contributors, the experience of the contributors and the ownership of the contributors play in the success of Wikipedia and search for new contributors. We include and justify a wide range of variables for measuring the quality of a contributor. We find little evidence that experience helps contribute to success. In fact, on the peer review system, contributions are more likely to be rejected than accepted. We also find strong evidence of ownership behavior. It is a general finding that the team ownership of content is discouraged within Wikipedia.

The Singularity is Not Near: Slowing Growth of Wikipedia

Borghese Suk, Gungor Acemoglu, Ed H. Chi, Peter Friesl
1500 University Ave
3300 Cypress Hill Road, Palo Alto, CA, 94304
+1650251-9480
{suk, acemoglu, edh, friesl}@parc.com

ABSTRACT
Peer review on Wikipedia has characterized the growth in content and quality in many traditionally open-ended systems including scientific journals. We find that the rate of adding new content suggests that Wikipedia growth has slowed with further experience. We identify a set of factors that are unique to open-ended systems. We consider content, experience, quality, and ownership as factors that affect growth. We find that content quality and ownership are particularly important in slowing growth. As growth slows, the rate of ownership of content increases and the rate of ownership of content increases. We find that ownership is particularly important in slowing growth. We find that ownership is particularly important in slowing growth. We find that ownership is particularly important in slowing growth.

WSor 2011

Welcome to the teahouse

A friendly place to help new editors become accustomed to Wikipedia culture, ask questions, and develop community relationships.

The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline

Aaron Halfaker¹, R. Stuart Geiger², Jonathan T. Morgan¹, and John Riedl¹

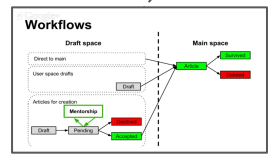
Abstract
Open collaboration systems, such as Wikipedia, need to maintain a pool of volunteer contributors to remain relevant. Wikipedia was created through a virtuous cycle that

Intelligent socialization software

Aaron Halfaker
en@User:EnFal
a.halfaker@hawaii.edu

Wikipedia as a socio-technical system

by Aaron Halfaker



ORES

The people's classifier service

Towards an open model for algorithmic infrastructure

Backlogs!

The Internet → NEW CONTENT CMG → Wikipedia

Backlogs everywhere...

Aaron Halfaker, Sunit Adhwa → Wikidata Research Showcase - Feb, 2018

Building Automated Vandalism Detection Tools for Wikidata

Andrius Dabravka, Aaron Halfaker, Dan Stransky
Ludoviko@protonmail.com, a.halfaker@wikimedia.org, danstransky@gmail.com

ABSTRACT
The Wikidata project has been successful in providing a structured data store for Wikipedia. However, it is still facing significant challenges in terms of vandalism detection. We present a new tool for detecting vandalism in Wikidata, which uses machine learning to identify suspicious edits. This tool is designed to be used by Wikidata administrators to identify and remove vandalism. We present a new tool for detecting vandalism in Wikidata, which uses machine learning to identify suspicious edits. This tool is designed to be used by Wikidata administrators to identify and remove vandalism.

English Wikipedia Quality Dynamics

By Aaron Halfaker, User:Halfak, User:EnFal

...and the Case of WikiProject Women Scientists

Deploying and maintaining AI in a socio-technical system

Aaron Halfaker
@aaronhalfaker

References

- Suh, B., Convertino, G., Chi, E. H., & Pirolli, P. (2009, October). The singularity is not near: slowing growth of Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (p. 8). ACM.
- Halfaker, A., Kittur, A., Kraut, R., & Riedl, J. (2009, October). A jury of your peers: quality, experience and ownership in Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (p. 15). ACM.
- https://meta.wikimedia.org/wiki/Research:Wikimedia_Summer_of_Research_2011
- <https://meta.wikimedia.org/wiki/Research:Teahouse>
- Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2012). The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist*, 0002764212469365.
- <https://www.youtube.com/watch?v=bwMBLrAHJLM> (Snuggle @ Wikimania 2013)
- https://www.mediawiki.org/wiki/File:Wikimedia_Research_%26_Data_Showcase_-_February_2014.webm (Wikipedia article creation analysis)
- <https://www.youtube.com/watch?v=AUupsnvV1oA#t=35m26s> (Articles for Creation workflow analysis)
- <https://www.youtube.com/watch?v=-We4GZbH3lw#t=34m10s> (Wikipedia as a socio-technical system "The Paramecium Talk")
- <https://www.youtube.com/watch?v=Hj7o5d-OEis#t=3m25s> (ORES -- The people's classifier service)
- <https://www.youtube.com/watch?v=nmrLu5qTgyA#t=1m15s> (The Keilana Effect -- English Wikipedia quality dynamics)
- <https://www.youtube.com/watch?v=rsFmqYxt9w#t=29m40s> (Deploying and maintaining AI in a socio-technical system. Lessons learned)
- https://www.youtube.com/watch?v=fpmRWCE7F_l#t=30m05s (Backlogs -- Backlogs everywhere)

Why is ORES?

- Wikipedia has socio-technical problems with newbies.



Why is ORES?

- Wikipedia has socio-technical problems with newbies.
- Many of Wikipedia's problems are due to its scale.



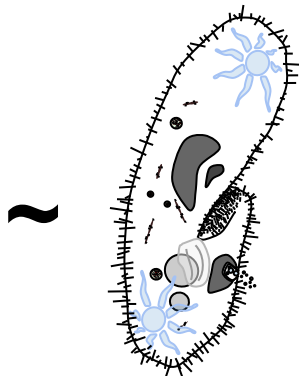
https://commons.wikimedia.org/wiki/File:Flickr_-_Official_U.S._Navy_Imagery_-_Sailor%27s_daughter_operates_a_fire_hose_with_crew_member_assistance.jpg

Why is ORES?

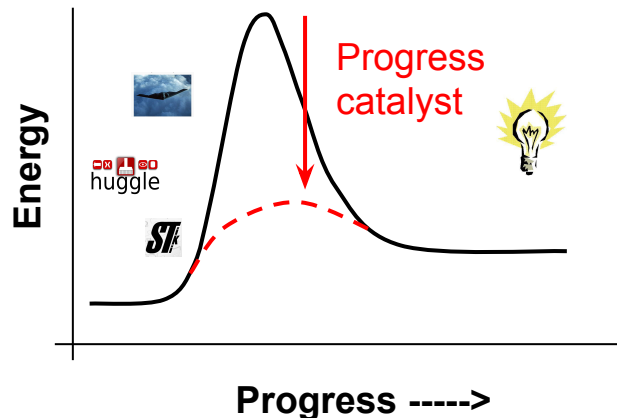
- Wikipedia has socio-technical problems with newbies.
- Many of Wikipedia's problems are due to its scale.
- ORES is an attempt to address both at the same time.
 - ... in a way that accounts for the complex dynamics of Wikipedia.



System with specialized
sub-systems



System with specialized
sub-systems



Why is ORES?

- Wikipedia has socio-technical problems with newbies.
- Many of Wikipedia's problems are due to its scale.
- ORES is an attempt to address both at the same time.
 - ... in a way that accounts for the complex dynamics of Wikipedia.

ORES -- The people's classifier service

<https://www.youtube.com/watch?v=Hj7o5d-OEis#t=3m25s>

Why is ORES?

- Wikipedia has socio-technical problems with newbies.
- Many of Wikipedia's problems are due to its scale.
- ORES is an attempt to address both at the same time.
 - ... in a way that accounts for the complex dynamics of Wikipedia.

What is ORES?

The machine classifier



https://commons.wikimedia.org/wiki/File:Artificial_intelligence.jpg
Public Domain

The machine classifier

is_anon
chrs_added
chrs_removed
cust_comment
repeated_chrs
longest_token
badwords_added

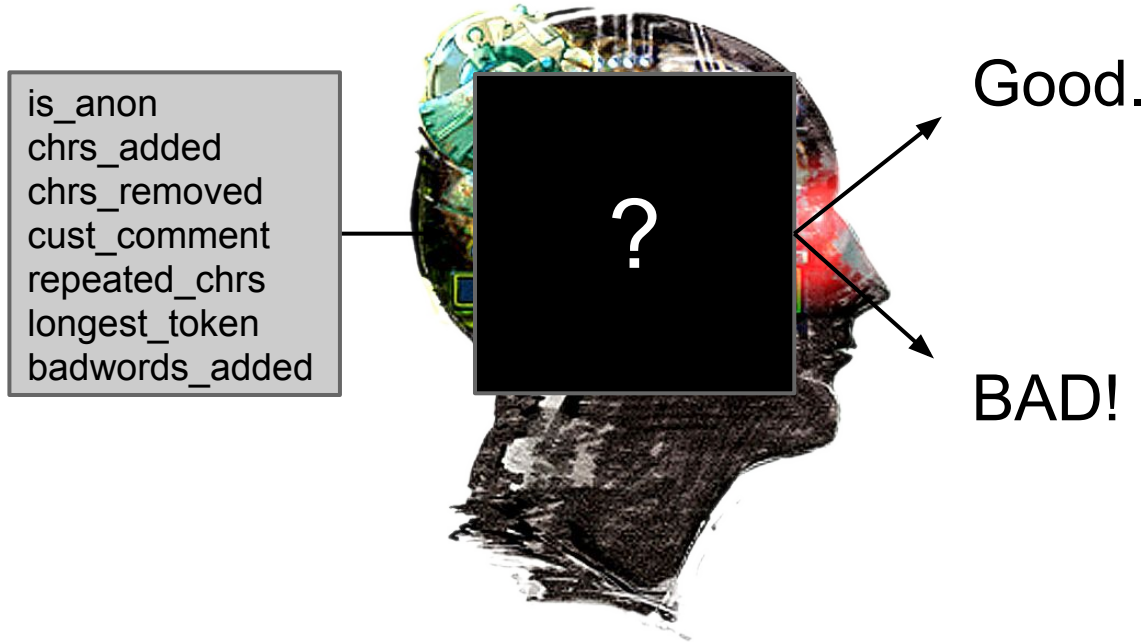


The machine classifier

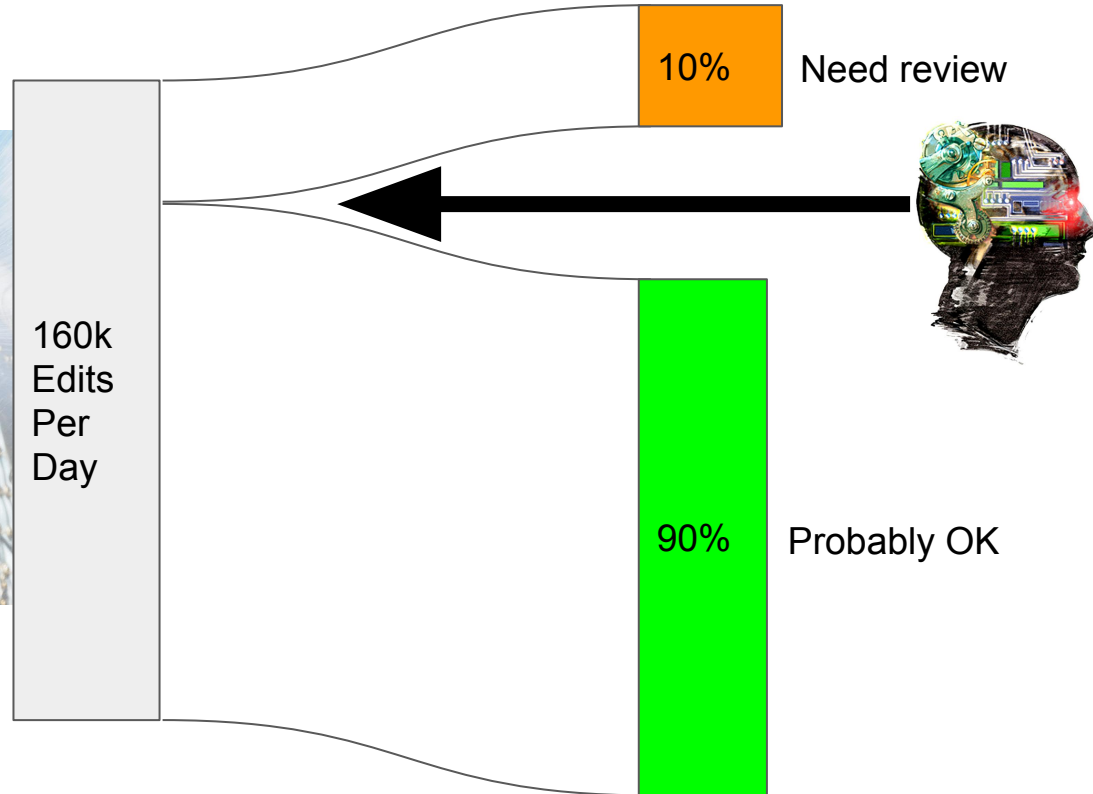
is_anon
chrs_added
chrs_removed
cust_comment
repeated_chrs
longest_token
badwords_added



The machine classifier



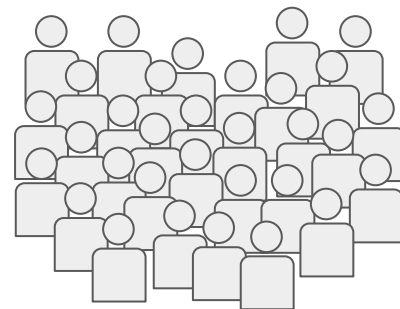
Counter vandalism



Without ORES: Reviewing 160k edits per day...

267 Hours

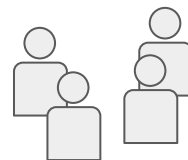
(33 people * 8 hours)



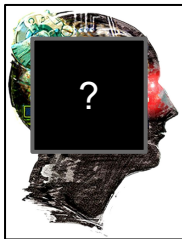
With ORES: Reviewing 16k edits per day...

27 Hours

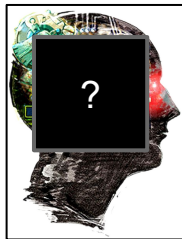
(4 people * 8 hours)



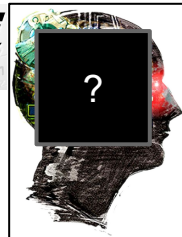

huggle



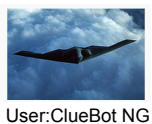
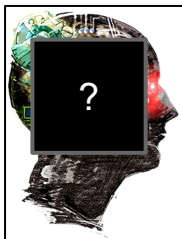
User:ClueBot NG



STK



   
huggle

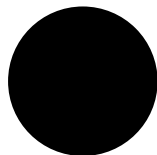


Wikipedia is a firehose ✓

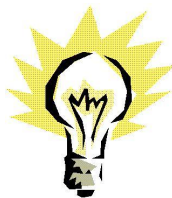
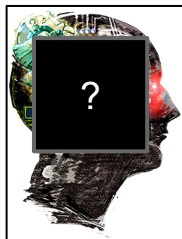
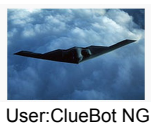
Bad edits must be reverted ✓

Minimize effort wasted on
quality control work ✓

Socialize & train the newcomers! ✓




huggle

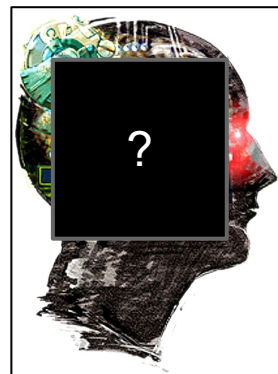
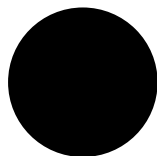


Wikipedia is a firehose ✓

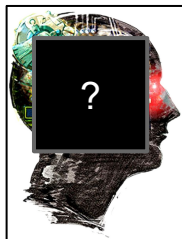
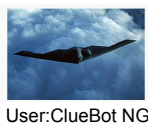
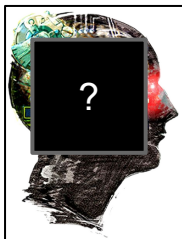
Bad edits must be reverted ✓

Minimize effort wasted on
quality control work ✓

Socialize & train the newcomers! ✓




huggle

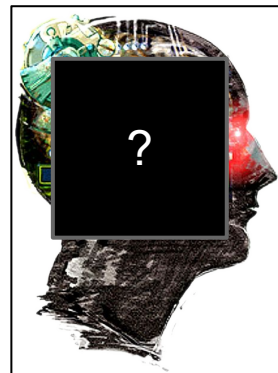
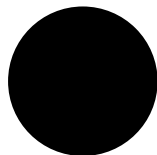


Wikipedia is a firehose ✓

Bad edits must be reverted ✓

Minimize effort wasted on
quality control work ✓

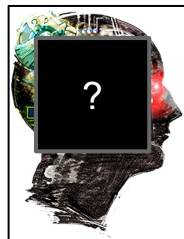
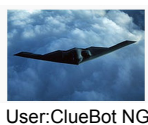
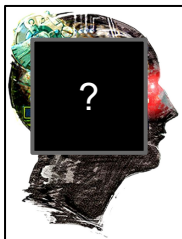
Socialize & train the newcomers! ✓



- 20+ research papers on Wikipedia damage detection
- Machine classification not part of standard CS degree
- Many volunteer tool devs don't have a CS degree anyway.

Labor intensive, performance considerations, etc.


huggle

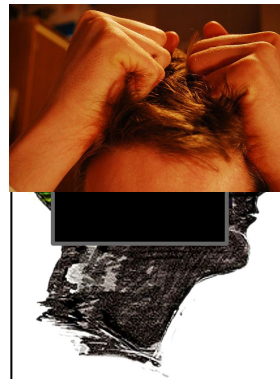
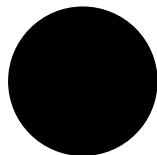


Wikipedia is a firehose ✓

Bad edits must be reverted ✓

Minimize effort wasted on
quality control work ✓

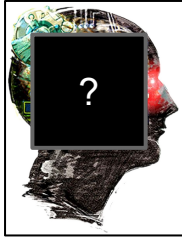
Socialize & train the newcomers! ✓



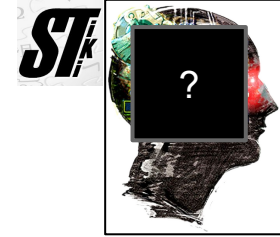
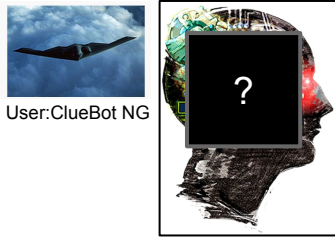
- 20+ research papers on Wikipedia damage detection
- Machine classification not part of standard CS degree
- Many wiki-tool devs don't have a CS degree anyway.

Labor intensive, performance considerations, etc.

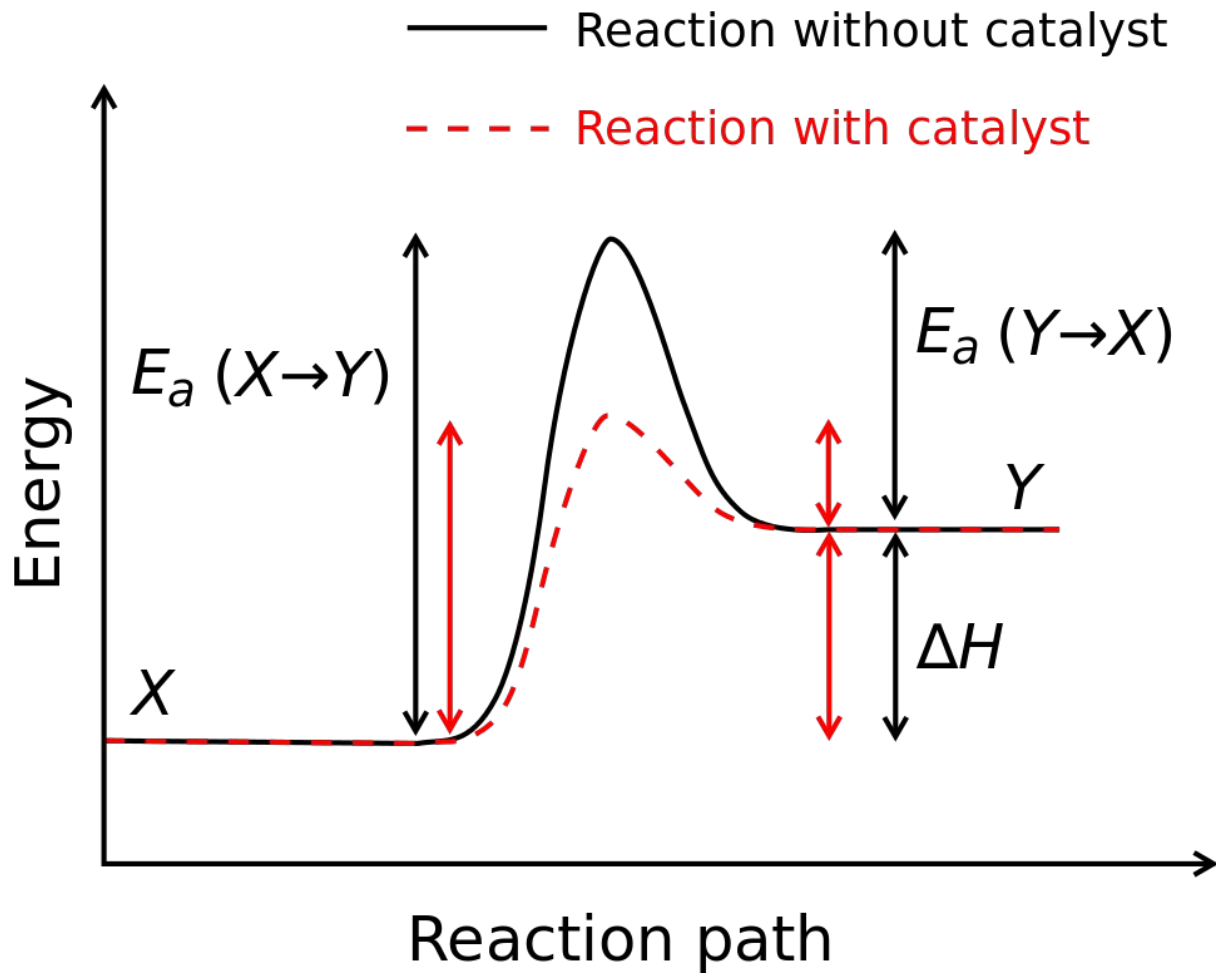
huggle

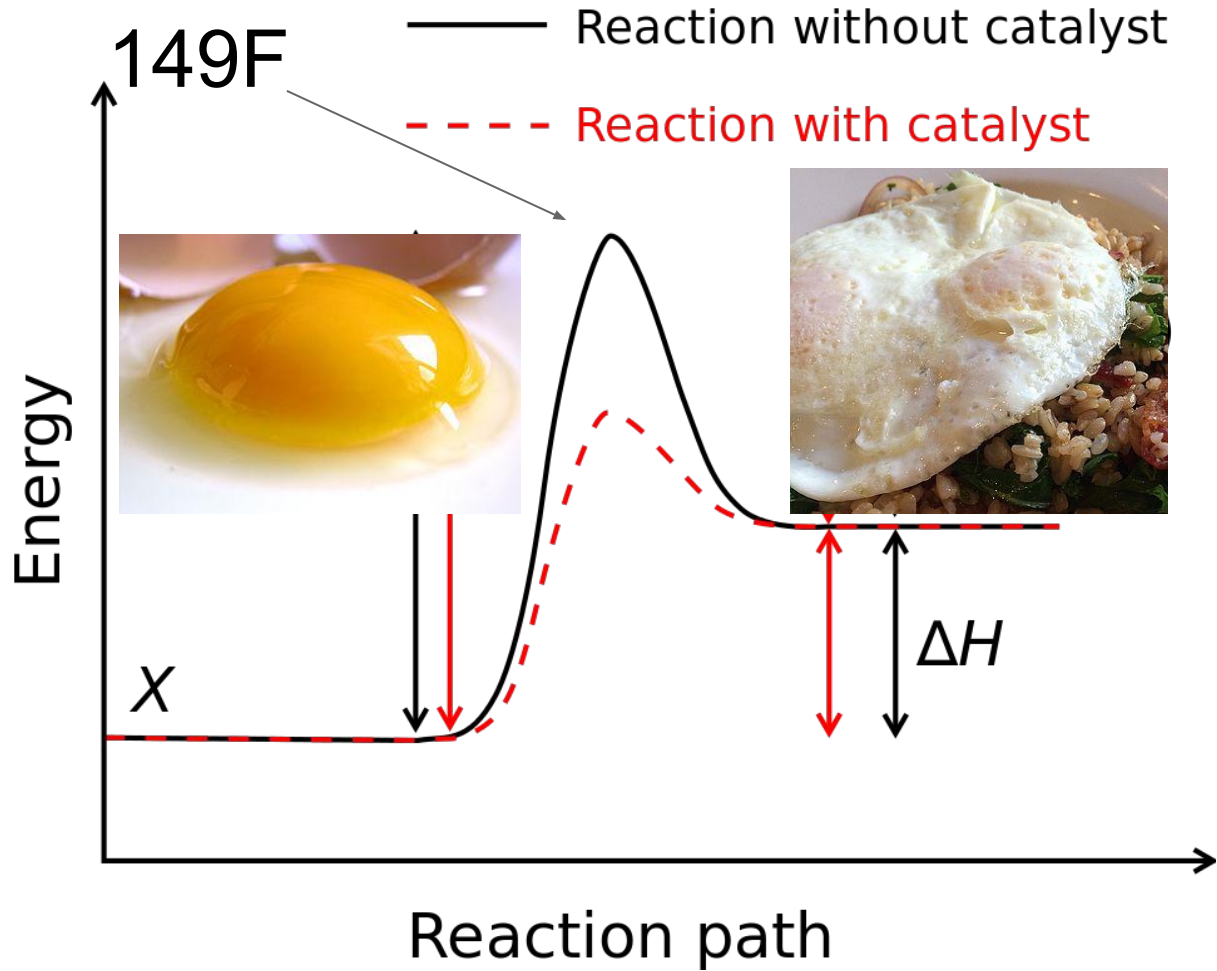


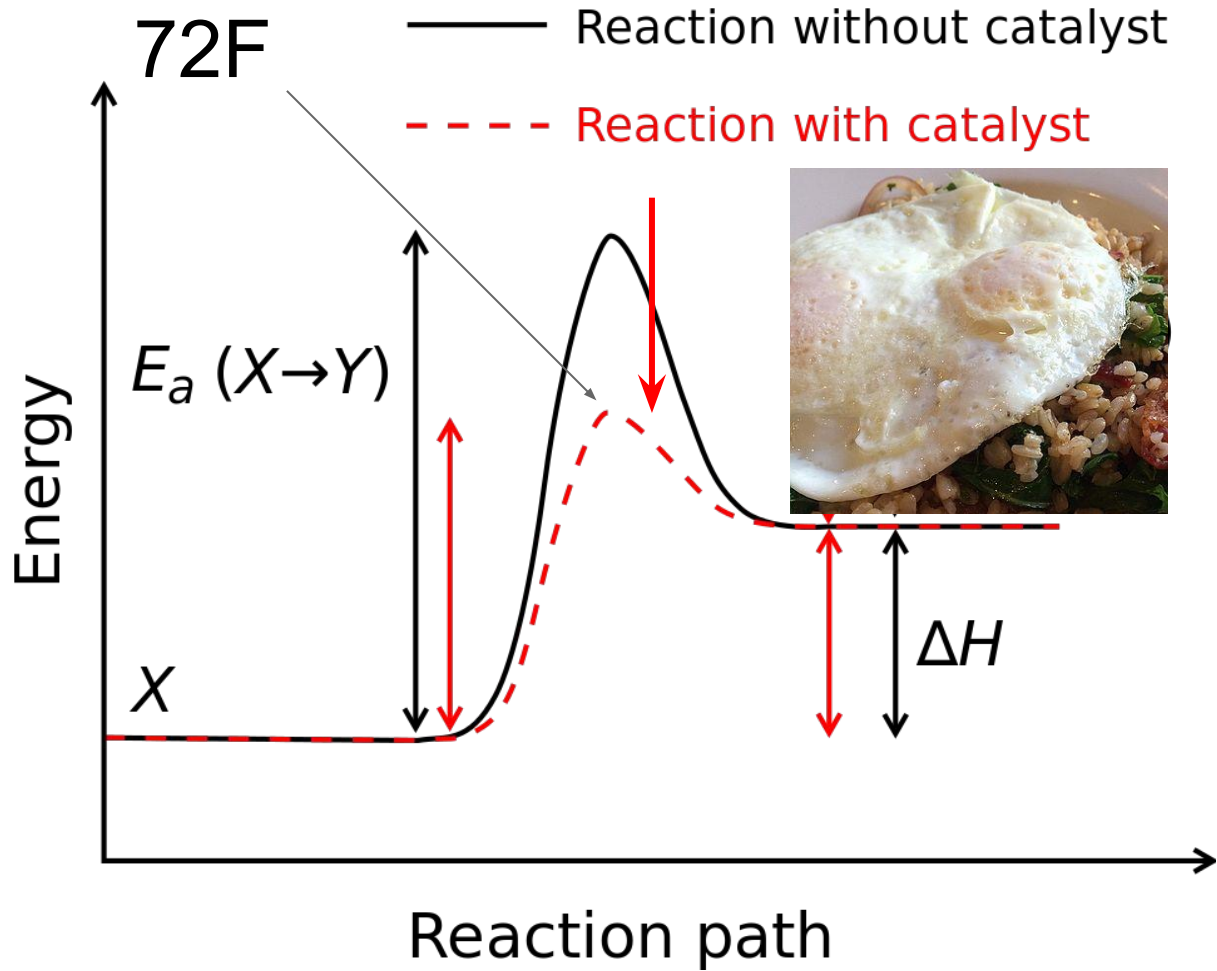
User:ClueBot NG



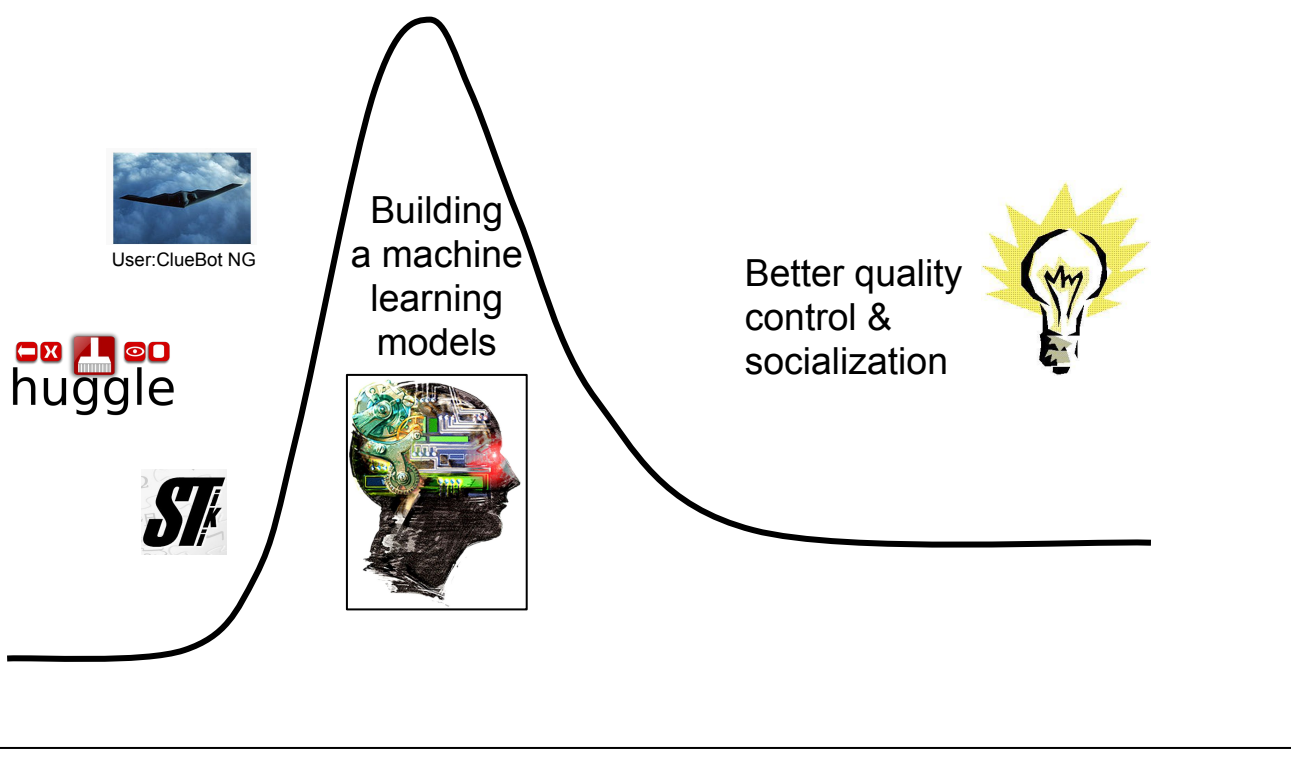
Authored by computer scientists with extensive skills in machine learning and distributed systems.







Energy



User:ClueBot NG



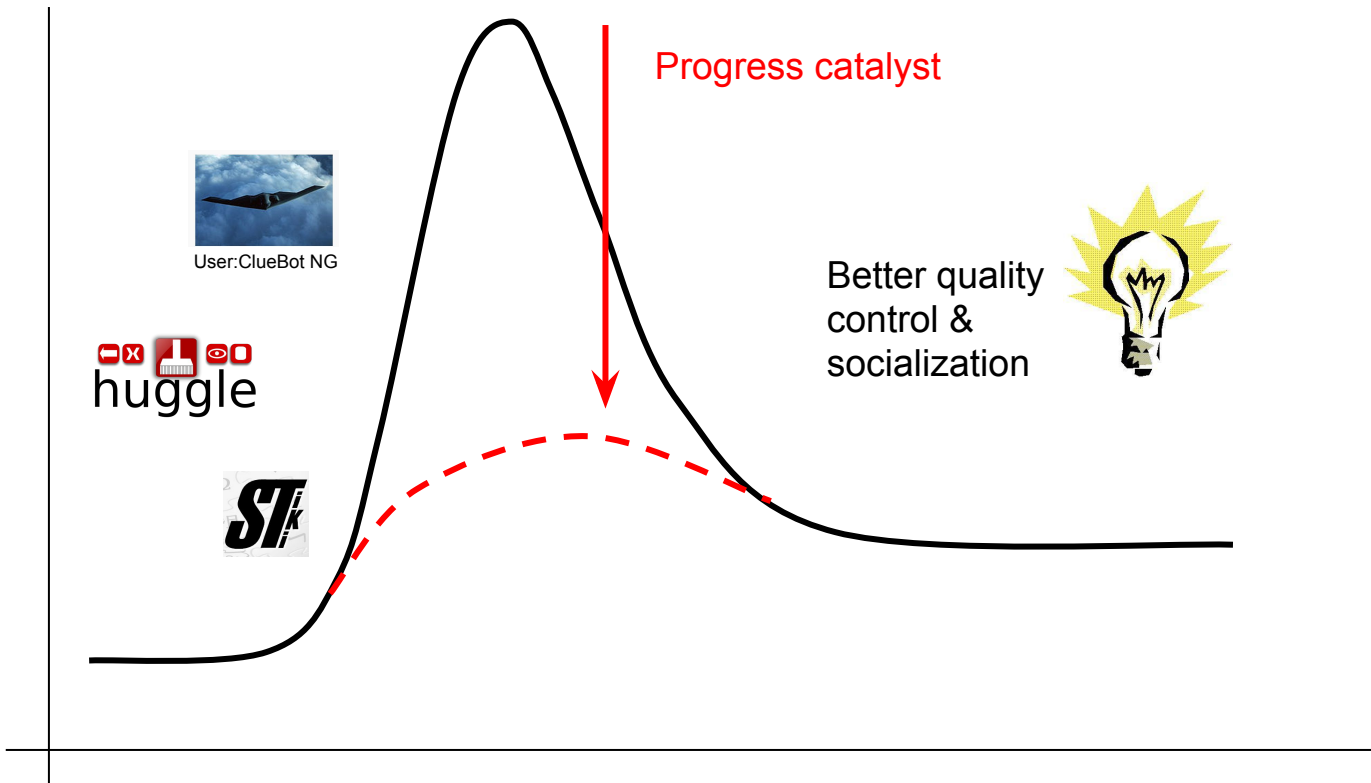
Building a machine learning models

Better quality control & socialization



Progress ----->

Energy



Progress catalyst



User:ClueBot NG

huggle

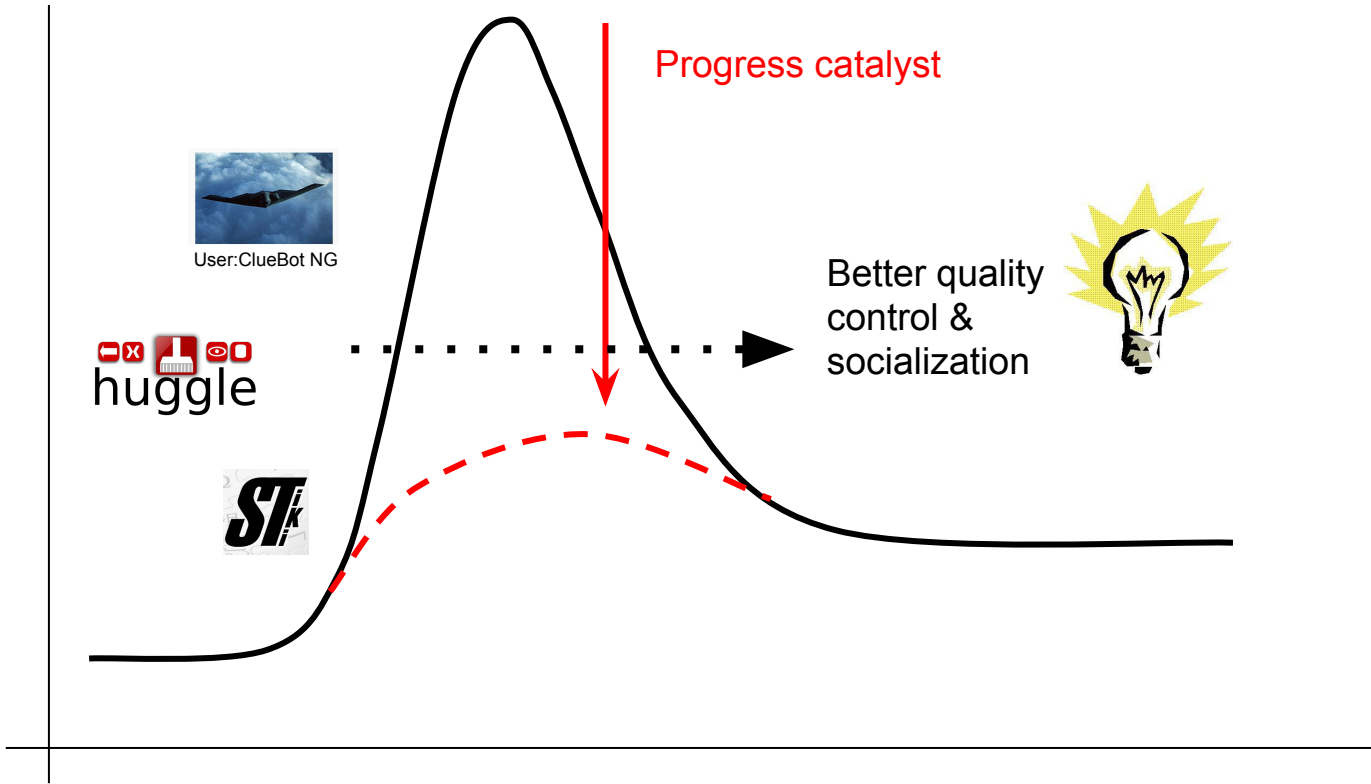


Better quality control & socialization



Progress ----->

Energy



Progress catalyst

User:ClueBot NG

huggle

STK

Better quality control & socialization

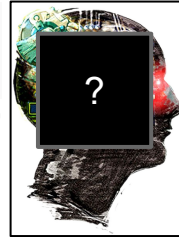
Progress ----->

Scoring Platform


huggle



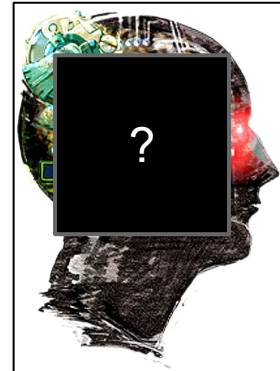
User: ClueBot NG



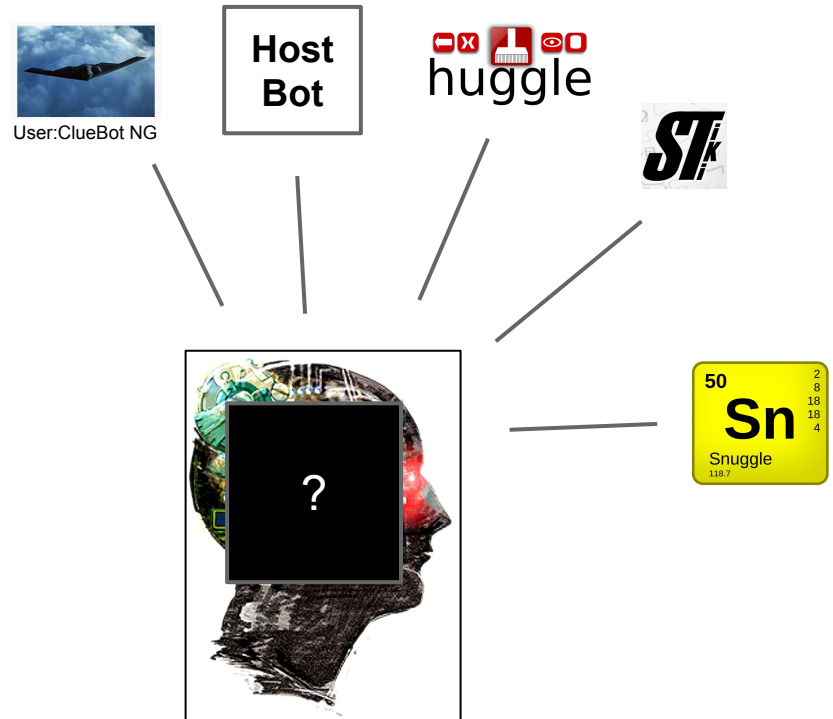
Scoring Platform



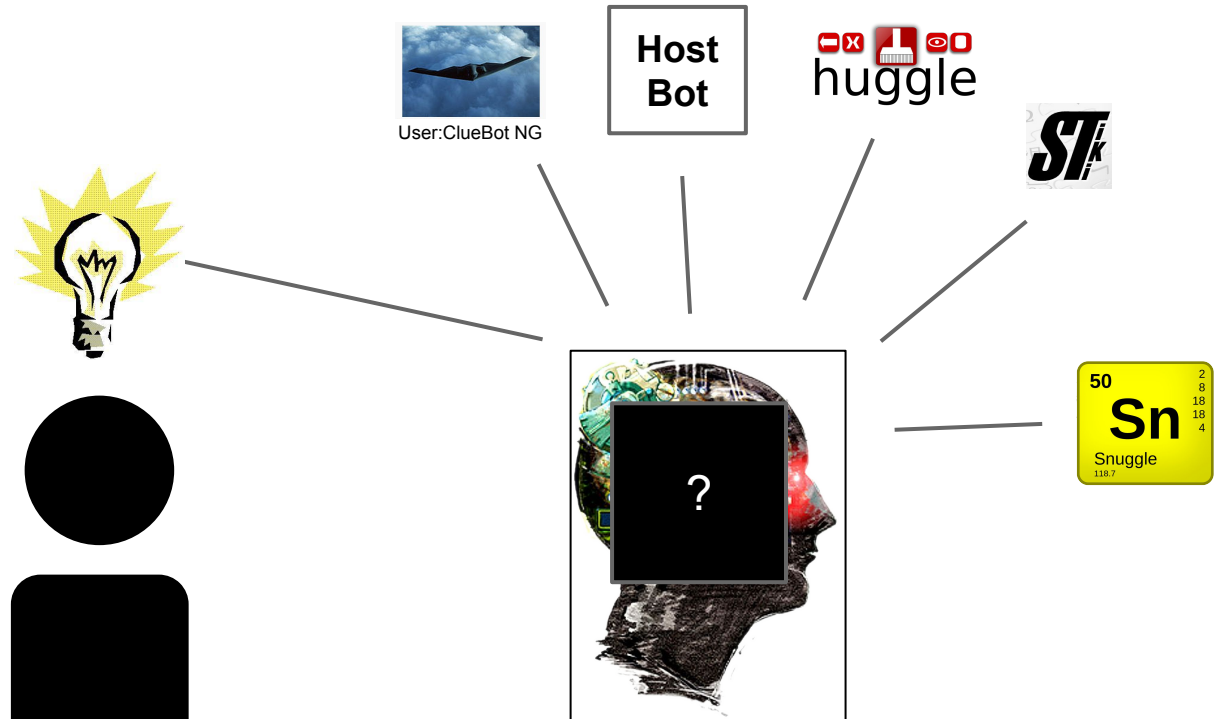
User:ClueBot NG



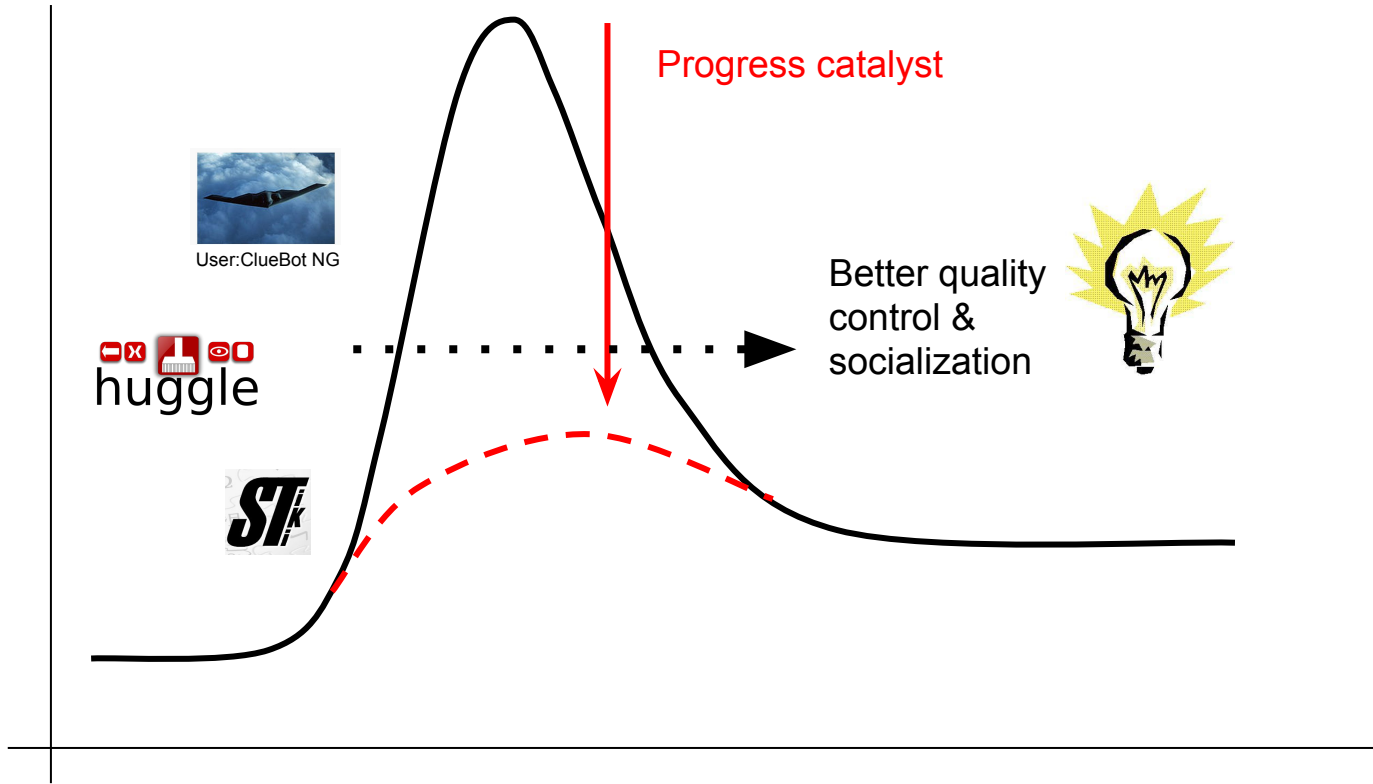
Scoring Platform



Scoring Platform



Energy



Progress catalyst



User:ClueBot NG

huggle

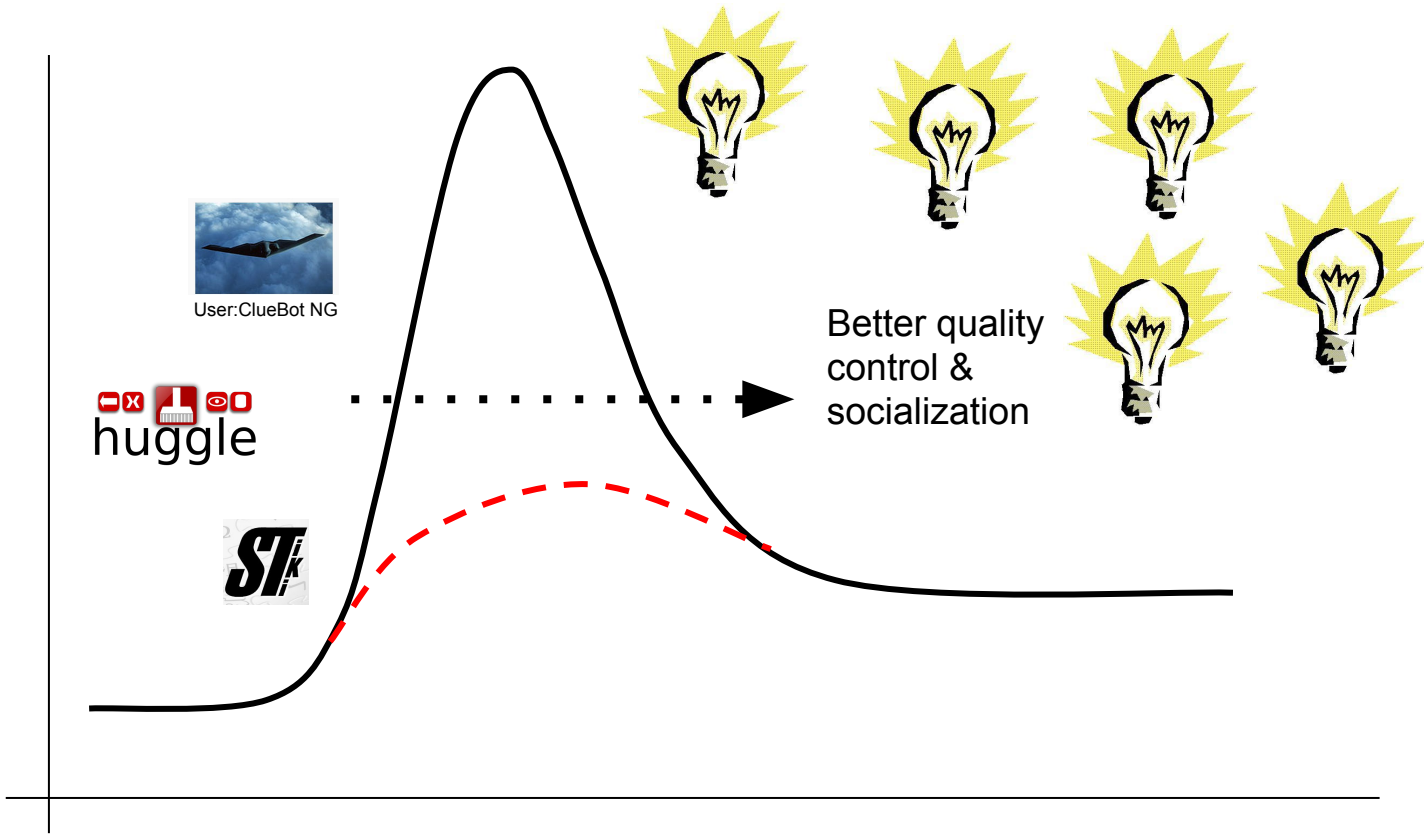


Better quality control & socialization



Progress ----->

Energy



User:ClueBot NG

huggle



Progress ----->

Better quality control & socialization

Part 2: Case studies

Part 2: Case studies

- Wikidata false-positives
- PatruBOT's banning
- Structural completeness

Wikidata:ORES/Report mistakes

< Wikidata:ORES

List [edit]

List here. Please include revision id (or a link to the diff) and why you think it's misclassified. Thanks :) [Amir \(talk\)](#) 21:49, 5 November 2015 (UTC)

[Special:Diff/269077027](#) - This is the deletion of a page on a local wiki. That can't be vandalism as it's something only a local admin can do. When a page is deleted, the sitelink needs to be removed to. In this case it's an automated deletion of the sitelink because of the page being deleted and not a user manually removing a link. [Mbch331 \(talk\)](#) 22:10, 5 November 2015 (UTC)

✓ **Done** Added in PR#13, it may take a while to work in wikidata [Amir \(talk\)](#) 16:58, 6 November 2015 (UTC)

[Special:Diff/269077025](#) - Same as above. [Mbch331 \(talk\)](#) 22:10, 5 November 2015 (UTC)

✓ **Done** Same as above

[Special:Diff/269086457](#) - We do get a lot of vandalism of Commons category statements, but in this case, the edit is fixing the way the category name was entered (changing "Category:Name" to "Name" is also fine). - [Nikki \(talk\)](#) 23:46, 5 November 2015 (UTC)

I'm trying to find out best way to include them. [Amir \(talk\)](#) 16:58, 6 November 2015 (UTC)

[Special:Diff/251530750](#) - This is just adding a reference, it doesn't look suspicious to me. - [Nikki \(talk\)](#) 23:46, 5 November 2015 (UTC)

Same as above [Amir \(talk\)](#) 16:58, 6 November 2015 (UTC)

[Special:Diff/243937491](#) - This is just adding a sitelink, it doesn't look suspicious to me. There are quite a few more similar sitelink additions on [Triton \(Q3359\)](#) which are marked. - [Nikki \(talk\)](#) 23:46, 5 November 2015 (UTC)

Again [Amir \(talk\)](#) 16:58, 6 November 2015 (UTC)

[Special:Diff/269090263](#) - Here I'm undoing bad edits. :P - [Nikki \(talk\)](#) 23:59, 5 November 2015 (UTC)

✓ **Done** Added in PR#13, it may take a while to work in wikidata [Amir \(talk\)](#) 16:58, 6 November 2015 (UTC)

Insights

- Client edits can't be “damaging”
- Merge edits are not reverts
- Vandalism to commons categories

Improvements [[edit](#)]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Improvements [edit]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Improvements [edit]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Improvements [edit]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Improvements [edit]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Improvements [[edit](#)]

Diff id ↕	Damaging ↕	Old score ↕	Score1 ↕	Score2 ↕	Score3 ↕	1st improv. ↕	2nd ↕	3rd ↕	Overall ↕
210649590	No	91%	30%	5%	0%	+61%	+25%	+5%	+91%
237999679	No	84%	71%	63%	60%	+13%	+8%	+3%	+24%
243937491	No	95%	46%	74%	71%	+49%	-28%	+3%	+24%
251530750	No	91%	55%	56%	55%	+36%	-1%	+1%	+36%
253584599	No	99%	89%	78%	70%	+10%	+11%	+8%	+29%
257856652	No	91%	30%	4%	1%	+61%	+26%	+3%	+90%
269077025	No	91%	82%	64%	73%	+9%	+18%	-9%	+18%
269077027	No	91%	89%	79%	86%	+2%	+10%	-7%	+5%
269086457	No	98%	100%	74%	71%	-2%	+26%	+3%	+28%
269090263	No	95%	100%	81%	76%	-5%	+19%	+5%	+19%
269093456	No	89%	99%	49%	44%	-10%	+50%	+5%	+45%
269186604	No	95%	100%	84%	71%	-5%	+16	+13%	+24%
269609233	No	89%	88%	46%	58%	+1%	+42%	-12%	+31%
270093221	No	92%	98%	40%	+48%	-6%	+58%	-8%	+44%
274775730	No	84%	87%	93%	64%	-3%	-6%	+29%	+20%

From [\[5\]](#) and [\[6\]](#). I'll check to improve it [Amir \(talk\)](#) 23:04, 4 December 2015 (UTC)

Insights

- False positive reporting is very valuable to model engineering
- Reporting improvements builds trust

PatruBOT ~ ClueBot NG

Spanish Wikipedia

English Wikipedia

PatruBOT ~ ClueBot NG

Spanish Wikipedia

English Wikipedia



Too Many Mistakes!



PatruBOT está **detenido de forma indefinida** desde el 27 de marzo de 2018, debido al número de falsos positivos y la imposibilidad de resolverlos a corto plazo. Según el tiempo disponible, en el futuro podría ser reactivado con un nuevo algoritmo o si ORES facilita que sea más efectivo, en función también del interés que pueda tener la comunidad.

Parada de PatruBOT [[editar código](#)]

Comunico que acabo de detener indefinidamente a PatruBOT. He visto que se han mantenido las quejas, y dado que anuncié esa posibilidad y que por ahora apenas si tengo tiempo de revisar (bastante por detrás) mi lista de seguimiento y de estar atento a IRC, creo que es lo más prudente. Notifico a [Mar del Sur](#) y [Ganimedes](#). Según mi tiempo disponible es posible que reconsidere en un futuro una reactivación con un nuevo algoritmo o si ORES facilita que sea más efectivo, en función también del interés que pueda tener la comunidad. En este sentido, leeré también aquello que se quiera comentar en este hilo. Saludos, - [José Emilio –jem– Tú dirás...](#)
09:34 27 mar 2018 (UTC)

Yo me desentendí hace bastante del asunto, pero gracias por el aviso. --Saludos. [Ganimedes](#) 12:14 27 mar 2018 (UTC)

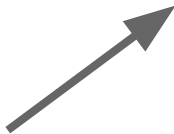
Una vez corregidos, al menos en parte, los problemas debería reactivarse, creo que es positivo. --[PePeEfe \(discusión\)](#) 13:11 27 mar 2018 (UTC)

Aunque cometiera (bastantes) errores, ayudaba muchísimo. De todos modos, si es la decisión de su operador, lo respetaré y aceptaré, por supuesto. Un bot antivandalismo me parece necesario en una Wikipedia tan visitada por vándalos que se empeñan en dejar su *huella* aquí. Espero que algo podamos hacer, porque con los reversiones o wikipedistas en general interesados en revisar los cambios recientes de los que disponemos actualmente no creo que tengamos suficiente. Un saludo cordial. --Fdo.: [Gonzalo P.M.G.](#) • 13:40 27 mar 2018 (UTC)

Considero que, a pesar de sus errores, el bot ayudaba mucho, ahora que ya no está activo los reversiones tendremos una tarea ardua. Espero que lo puedas activar de nuevo pronto. Saludos. --[Chico⁵¹²](#) 13:44 27 mar 2018 (UTC)

No me extraña que lo desactive por las constantes quejas que recibe. Algunas de estas quejas proceden de usuarios que llegan para editar un tiempo limitado, pero que se quejan de que se les han borrado datos, cuando en muchos casos lo que aportan son ediciones incorrectas. ¿Comete errores? Está claro, yo también, pero no me hago a la idea del trabajo que va a dejar de hacer. Los vándalos estarán encantados, y no me extrañaría que alguno fuese precisamente el que genera quejas para conseguir esto. Un saludo. --[vanbasten_23 \(discusión\)](#) 13:55 27 mar 2018 (UTC)

No extrañándome la decisión, que comprendo perfectamente, solo nos queda distribuirnos el trabajo lo mejor que podamos. Yo tengo marcas de seguimiento en todos los días y meses del año (efemérides), que son las páginas que patrullo continuamente. Supongo que hay otras formas de patrullar por intereses (no me veo patrullando deportes, por ejemplo). ¿Cómo se pueden implementar listas de patrullaje por categorías?. Saludos,--[Jmrebes \(déjame un mensaje aquí\)](#) 14:02 27 mar 2018 (UTC)

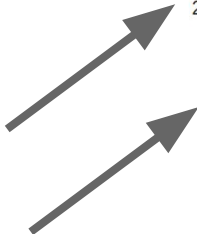


Parada de PatruBOT [[editar código](#)]

Comunico que acabo de detener indefinidamente a PatruBOT. He visto que se han mantenido las quejas, y dado que anuncié esa posibilidad y que por ahora apenas si tengo tiempo de revisar (bastante por detrás) mi lista de seguimiento y de estar atento a IRC, creo que es lo más prudente. Notifico a [Mar del Sur](#) y [Ganimedes](#). Según mi tiempo disponible es posible que reconsidere en un futuro una reactivación con un nuevo algoritmo o si ORES facilita que sea más efectivo, en función también del interés que pueda tener la comunidad. En este sentido, leeré también aquello que se quiera comentar en este hilo. Saludos, - [José Emilio -jem-](#) *Tú dirás...* 09:34 27 mar 2018 (UTC)

Yo me desentendí hace bastante del asunto, pero gracias por el aviso. --Saludos. [Ganimedes](#) 12:14 27 mar 2018 (UTC)

Una vez corregidos, al menos en parte, los problemas debería reactivarse, creo que es positivo. --[PePeEfe](#) ([discusión](#)) 13:11 27 mar 2018 (UTC)



Aunque cometiera (bastantes) errores, ayudaba muchísimo. De todos modos, si es la decisión de su operador, lo respetaré y aceptaré, por supuesto. Un bot antivandalismo me parece necesario en una Wikipedia tan visitada por vándalos que se empeñan en dejar su *huella* aquí. Espero que algo podamos hacer, porque con los reversiones o wikipedistas en general interesados en revisar los cambios recientes de los que disponemos actualmente no creo que tengamos suficiente. Un saludo cordial. --Fdo.: [Gonzalo P.M.G.](#) • 13:40 27 mar 2018 (UTC)

Considero que, a pesar de sus errores, el bot ayudaba mucho, ahora que ya no está activo los reversiones tendremos una tarea ardua. Espero que lo puedas activar de nuevo pronto. Saludos. --[Chico](#)⁵¹² 13:44 27 mar 2018 (UTC)

No me extraña que lo desactive por las constantes quejas que recibe. Algunas de estas quejas proceden de usuarios que llegan para editar un tiempo limitado, pero que se quejan de que se les han borrado datos, cuando en muchos casos lo que aportan son ediciones incorrectas. ¿Comete errores? Está claro, yo también, pero no me hago a la idea del trabajo que va a dejar de hacer. Los vándalos estarán encantados, y no me extrañaría que alguno fuese precisamente el que genera quejas para conseguir esto. Un saludo. --[vanbasten_23](#) ([discusión](#)) 13:55 27 mar 2018 (UTC)

No extrañándome la decisión, que comprendo perfectamente, solo nos queda distribuirnos el trabajo lo mejor que podamos. Yo tengo marcas de seguimiento en todos los días y meses del año (efemérides), que son las páginas que patrullo continuamente. Supongo que hay otras formas de patrullar por intereses (no me veo patrullando deportes, por ejemplo). ¿Cómo se pueden implementar listas de patrullaje por categorías?. Saludos,--[Jmrebes](#) ([déjame un mensaje aquí](#)) 14:02 27 mar 2018 (UTC)

Me atrevo a decir que el BOT tiene una tasa de error cercana al **40%**, eso implica por lógica que acertaría un 60%. Eso como en la teoría del vaso medio lleno o medio vacío se puede ver desde dos ópticas. Desde la óptica de que acierta mucho vandalismo y por ende lo evita o desde la óptica de que falla bastante y por lo tanto evita ediciones que podrían ser correctas. Es evidente que la tasa de error del bot es excesiva pero también que desactivarlo hará que el 100% de los vandalismos que sí era capaz de evitar ahora no se eviten. En todo caso yo sigo confiando en la buena y generalmente poco valorada labor de los que patrullan cambios recientes, además de la efectividad de los filtros de edición. Por lo demás, quizás algún otro compañero sea capaz de correr un bot o de colaborar con Jem en el código de Patrubot para tener algo que no necesariamente haga cientos de reversiones al día pero sí sea capaz de que las que haga sean correctas. Porque quizás el problema de PatruBOT sea ese, que quería abarcar demasiado y al final ha terminado abarcando nada. Saludos. [Bernard - Et voilà!](#) 14:27 27 mar 2018 (UTC)

Desde la página [Wikipedia:Mantenimiento/Revisión de errores de PatruBOT](#) saqué la estadística de cuántas reversiones del bot son correctas e incorrectas. De las 236 reversiones reportadas, 153 eran reversiones incorrectas y 83 eran reversiones correctas (un 64% de error). Aun así no lo podemos tomar como definitivo. Si alguien vandaliza una página y el bot lo revierte, lo más probable es que no reporte el error. Esa reversión es correcta, pero como nadie la reporta no quedó plasmada en este porcentaje. No estoy seguro, pero creo que una gran cantidad de las reversiones correctas reportadas eran por temas de referencias, plagios, formato, estilo, cosas que un editor novato no conoce y no se puede decir que sean vandalismo. También se pueden revisar [las contribuciones del bot](#) y con ello sacar unos resultados más precisos. Habrá que ver cómo va este tiempo sin el bot, luego puede ser reactivado si vemos que efectivamente el vandalismo es demasiado para ameritar su uso.--

[Santiago142857](#) 17:17 27 mar 2018 (UTC)



Wikipedia:Mantenimiento/Revisión de errores de PatruBOT/Análisis

< [Wikipedia:Mantenimiento](#) - [Revisión de errores de PatruBOT](#)



Esta página está destinada a analizar una muestra aleatoria de las contribuciones de PatruBOT.

Muestra aleatoria [[editar código](#)]

100 [[editar código](#)]

- 00:39 10 feb 2018 (+283) Berenguer de Cruïlles** (Revertidos los cambios de [95.16.136.45](#) a la última edición de TheRichic)-Revirtió borrado de texto y una parte de una plantilla - Gani
- 02:50 10 feb 2018 (-142) Unicanal (Paraguay)** (Revertidos los cambios de [181.120.150.204](#))- revirtió agregado de texto sin fuentes - Gani
- 02:59 10 feb 2018 (-40) Televisión digital terrestre en Paraguay** (Revertidos los cambios de [181.120.150.204](#) a la última edición de 2800:810:55D:17C9:957F:7965:2D6C:17B4)- revierte vandalismo - Gani
- 06:01 10 feb 2018 (+437) Sharknado 2: The Second One** (Revertidos los cambios de [179.41.146.136](#) a la última edición de FrescoBot)- revierte vandalismo - Gani
- 07:05 10 feb 2018 (-407) Colorina (telenovela peruana)** (Revertidos los cambios de [181.66.165.246](#) a la última edición de 179.7.208.250)- revierte edición que cambia apellidos, sin fuentes (posible vandalismo) - Gani
- 10:53 10 feb 2018 (-26) Configuración electrónica** (Revertidos los cambios de [83.56.97.82](#))- retira adición de cuenta de Twitter (posible spam) - Gani
- 11:08 10 feb 2018 (-613) PAW Patrol** (Revertidos los cambios de [190.46.1.9](#))- revierte correcciones de ortografía y adición de texto sin fuentes - Gani
- 13:36 10 feb 2018 (-156) Jueves Lardero** (Revertidos los cambios de [79.159.47.242](#))- revierte corrección menor y elimina adición de texto sin fuentes - Gani

Índice [[ocultar](#)]

- Muestra aleatoria
 - 100
 - 200
 - 300
 - 400
 - 500
 - 600
 - 700
 - 800
 - 900
 - 1000
- Reservas
- Origen de los datos
 - Estadísticas globales
- Análisis de los datos
- Conclusiones

https://ores.wikimedia.org/v3/scores/eswiki?models=damaging&model_info=statistics.thresholds.true.'

maximum recall @ precision \geq 0.9,

```
{"!f1": 0.95, "!precision": 0.905, "!recall": 0.998,  
  "accuracy": 0.906, "f1": 0.271,  
  "filter_rate": 0.981, "fpr": 0.002, "match_rate": 0.019,  
  "precision": 0.916, "recall": 0.159,  
  "threshold": 0.962}
```

https://ores.wikimedia.org/v3/scores/eswiki?models=damaging&model_info=statistics.thresholds.true.'

maximum recall @ precision \geq 0.9,

```
{"!f1": 0.95, "!precision": 0.905, "!recall": 0.998,  
  "accuracy": 0.906, "f1": 0.271,  
  "filter_rate": 0.981, "fpr": 0.002, "match_rate": 0.019,  
  "precision": 0.916, "recall": 0.159,  
  "threshold": 0.962}
```

https://ores.wikimedia.org/v3/scores/eswiki?models=damaging&model_info=statistics.thresholds.true.'

maximum recall @ precision \geq 0.9,

```
{"!f1": 0.95, "!precision": 0.905, "!recall": 0.998,  
  "accuracy": 0.906, "f1": 0.271,  
  "filter_rate": 0.981, "fpr": 0.002, "match_rate": 0.019,  
  "precision": 0.916, "recall": 0.159,  
  "threshold": 0.962}
```

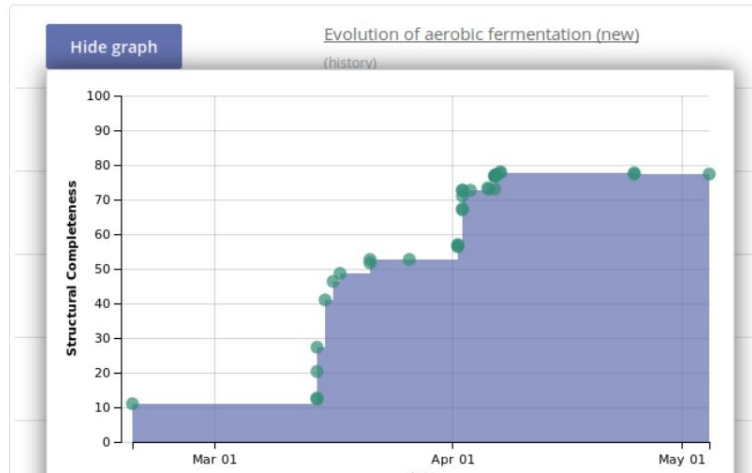
Insights

- Precision/Recall is critical to auto-revert bots
- Bot developers don't have time to manage the nuances of precision
- Developers don't know how to apply our threshold optimizations

Visualizing article history with Structural Completeness

By Sage Ross on September 16, 2016

You may have noticed a recent addition to the Articles tab of dashboard.wikiedu.org course pages: “structural completeness”. This feature is an experiment in visualizing the history of articles as they develop.



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

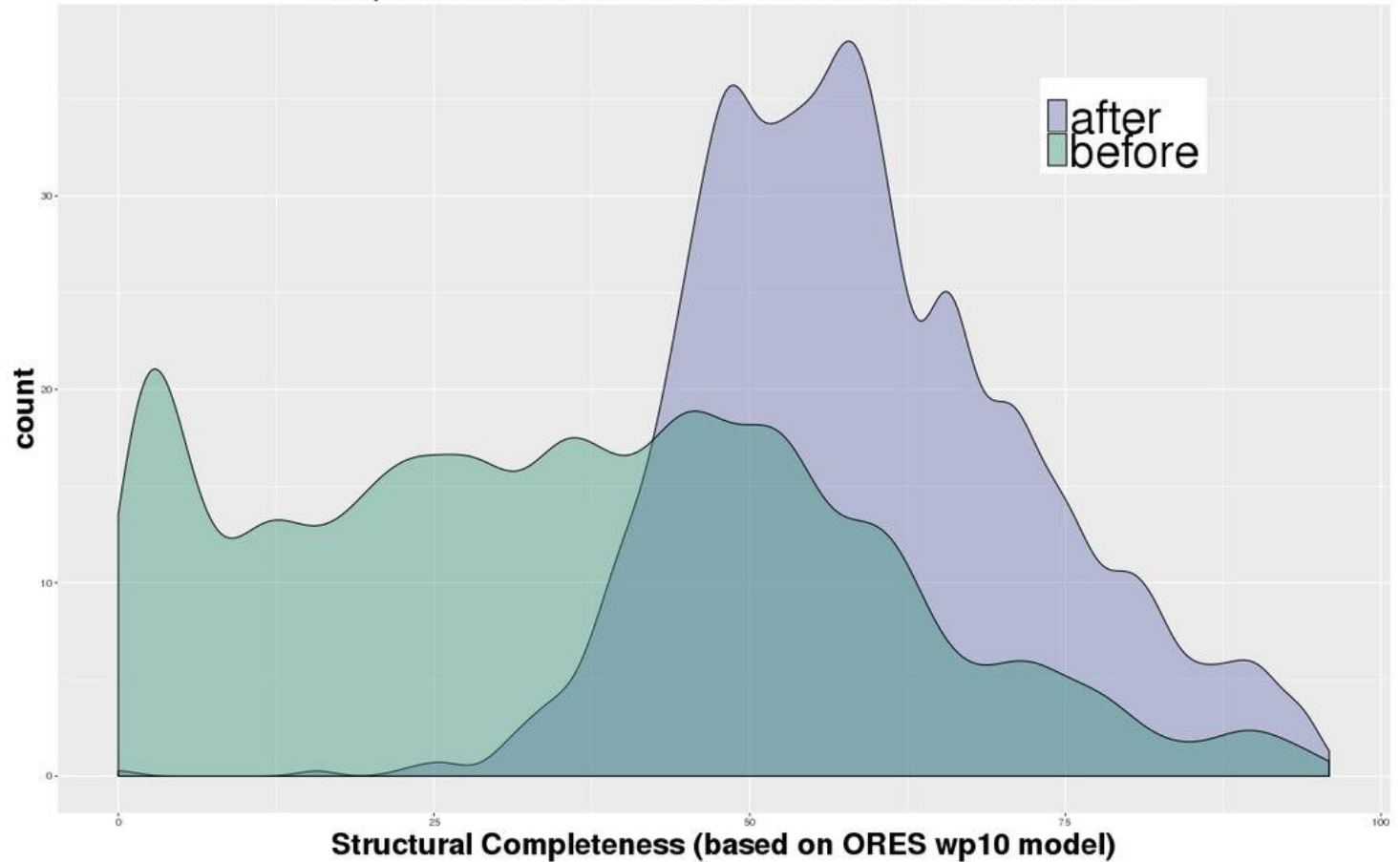
`[[[:en: Biology]]`



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]]`

Classroom Program Spring 2017 existing article improvement
min bytes added: 6000 - articles: 1156 - ave. before: 35.7 - ave. after: 59.0



Part 3: Discussion

Part 3: Discussion

[ORES: Facilitating re-mediation of Wikipedia's socio-technical problems.pdf](#)

(Halfaker et al. *under review*)

ORES: Facilitating re-mediation of Wikipedia's socio-technical problems

AARON HALFAKER, Wikimedia Foundation, USA

JONATHAN T. MORGAN, Wikimedia Foundation, USA

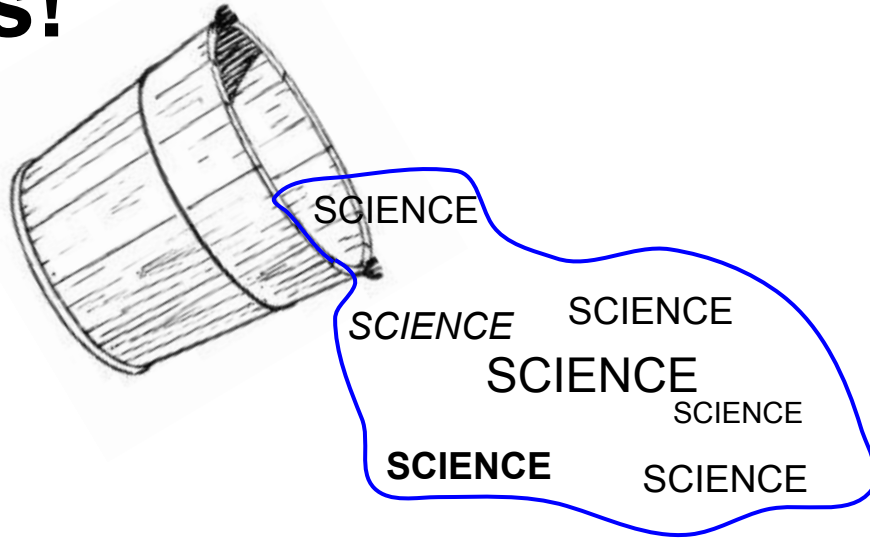
AMIR SARABADANI, Wikimedia Deutschland, Germany

ADAM WIGHT, Wikimedia Foundation, USA

Intelligent algorithms have a long history of making curation work in peer production tractable. From counter-vandalism to task routing, basic machine prediction allows open knowledge projects like Wikipedia to scale to the largest encyclopedia in the world. However, the ideologies and values of the community were captured in the development of these algorithms and the processes they support. Wikipedia's challenges and the community's values have changed in the last decade, but its algorithmic support systems have remained largely stagnant. The conversation about what quality control should be and what place algorithms have remains restricted to a few expert engineers. In this paper, we describe ORES: an algorithmic service designed to open up socio-technical conversations in Wikipedia to a broader set of participants. In this paper, we argue the theoretical mechanisms of social change ORES enables and we describe the phenomena around ORES from the 3 years since ORES' deployment.

CCS Concepts: • **Networks** → **Online social networks**; • **Computing methodologies** → **Supervised learning by classification**; • **Applied computing** → **Sociology**; • **Software and its engineering** → **Software design techniques**; • **Computer systems organization** → **Cloud computing**;

Thanks!



Props to my collaborators

- User:Ladsgroup
- User:Platonides
- User:Ragesoss

Aaron Halfaker

ahalfaker@wikimedia.org

[User:EpochFail](#)

<https://twitter.com/halfak>