Crowd Documentation
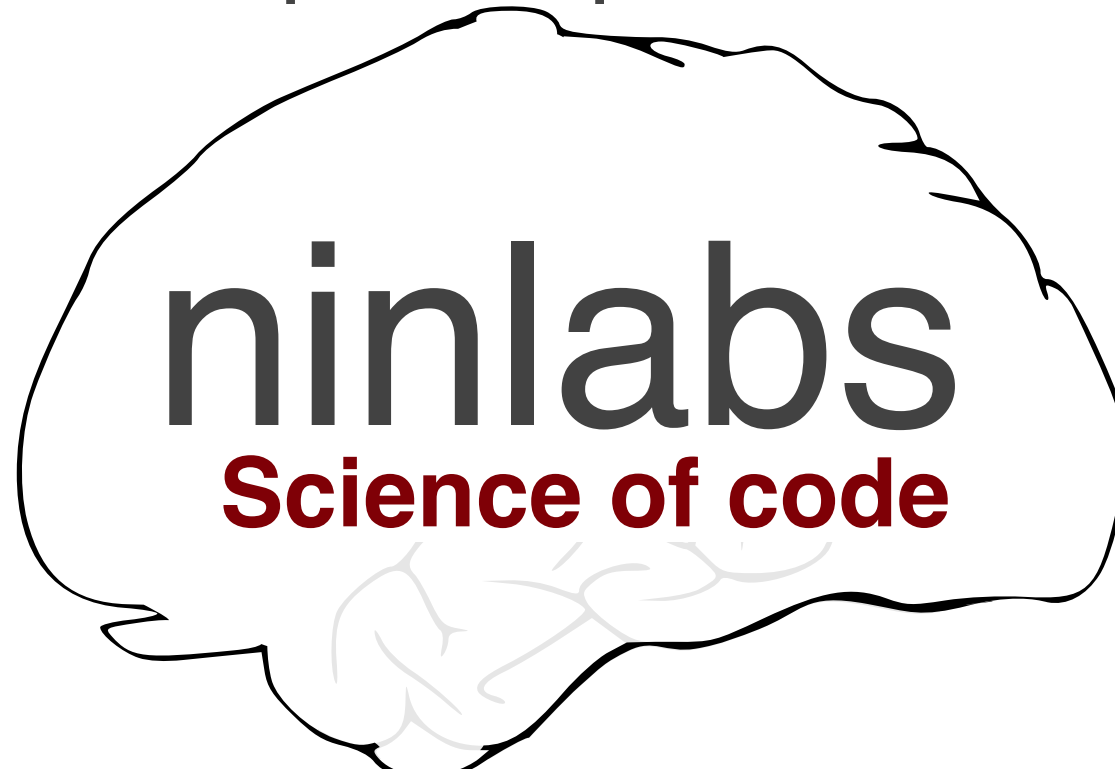
How Programmer Social Communities are Flipping Software Development

Christopher Parnin, @chrisparnin
http://chrisparnin.me

ninlabs
Science of code

Activity

- How do you learn a new technology?

Types of Documentation

Software documentation is created by a few and read by few.

API documentation is created by a few and read by many.

APIs are Big Business

Is costly to create

How big is the MSDN Library – and why should you care?

By JEFF BRAATEN | *Published:* SEPTEMBER 15, 2010

As of September 2010, the MSDN Library provides online access to over 16 million topics of Microsoft technical and product documentation. What exactly is a topic? In layman's terms a topic is simply a page authored for users to view. From a database perspective, it is a unique content item with an XHTML representation. Other statistics:

- **374,000 en-us topics** delivered with Visual Studio 2010 / .NET Framework 4 for the English (United States) locale – our largest single body of product documentation
- **1.4 million en-us topics** across all versions of Visual Studio / .NET Framework (from Visual Studio 6.0 through Visual Studio 2010)
- **3 million en-us topics** in the MSDN Library covering **over 20 Microsoft product families**
- **16 million topics** in the MSDN Library distributed across **14 locales**

Why is the Library so big?

From Office to Windows to Servers to Developer Tools, there are **over 20 product families** represented in the Library. Most product families provide technical

http://thirdblogfromthesun.com/2010/09/how-big-is-the-msdn-library/

Help win platform wars

It sucks, the documentation sucks, the examples are ridiculous. This is part of the reason there are so many junk apps on android as well. iOS stresses interface and user experience from the first time you open the dev portal, android... do they ever even mention it? Not that I have seen. Its all obscure code stuff and external libraries created by who knows who etc.

http://forum.unity3d.com/threads/104567-iOS-vs-Android-from-a-dev-perspective/page2

APIs as the new business model

- Twilio

- Stripe

- Coinbase

```
//require the Twilio module and create a REST client
var client = require('twilio')('ACCOUNT_SID', 'AUTH_TOKEN');
client.makeCall({
    to:'+16515556677',
    from: '+14506667788',
    url: 'http://www.example.com/twiml.php'
}, function(err, responseData) {
    console.log(responseData.from); // outputs "+14506667788"
});
```

Traditional Forms

4000 pages of documentation

What are the sources of dev knowledge now?

Google

Google Search   I'm Feeling Lucky

Unofficial Microsoft Survey

When learning about API documentation % of following resource is used "Often"

73.5% (2,224)

42.5% (1,289)
Code completion (IntelliSense)

40.1% (1,212)
Microsoft's Official Documentation

Searching for Documentation

Examine search results of jQuery

What is jQuery?

jQuery is a fast, small, and feature-rich JavaScript library. It makes things like HTML document traversal and manipulation, event handling, animation, and Ajax much simpler with an easy-to-use API that works across a multitude of browsers. With a combination of versatility and extensibility, jQuery has changed the way that millions of people write JavaScript.

jQuery.post() | jQuery API Documentation
api.**jquery**.com/**jQuery.post**/ ▾
The success callback function is passed the returned data, which will be an XML root
element or a text string depending on the MIME type of the response.

jQuery post() Method - W3Schools
www.w3schools.com/**jquery**/ajax_**post**.asp ▾
Parameter, Description. URL, Required. Specifies the url to send the request to. data,
Optional. Specifies data to send to the server along with the request.

jQuery AJAX get() and post() Methods - W3Schools
www.w3schools.com/**jquery**/**jquery**_ajax_get_**post**.asp ▾
Next Chapter ». The **jQuery** get() and **post**() methods are used to request data from the
server with an HTTP GET or **POST** request. HTTP Request: GET vs. **POST**.

jQuery Ajax POST example - Stack Overflow
stackoverflow.com/.../capture-all-of-the-forms-data-and-submit-it-to-a-p... ▾
Feb 15, 2011 - I am trying to send data from a form to a database. Here is the form I ...
Basic usage of .ajax would look something like this: HTML: <form id="foo"> ...

jQuery - post(url, data, callback, type) Method - Tutorials Point
www.tutorialspoint.com/**jquery**/ajax-**jquery-post**.htm ▾
jQuery - **post**(url, data, callback, type) Method, Learning JavaScript Framework in
simple and easy steps. A beginner's tutorial containing complete knowledge ...

Using jQuery to POST [FromBody] parameters to Web API » Encosia
encosia.com/using-**jquery**-to-**post**-frombody-parameters-to-web... ▾
by Dave Ward - in 329 Google+ circles
Apr 3, 2013 - ASP.NET Web API has been one of my favorite recent additions
to ASP.NET. Whether you prefer a RESTful approach or something more ...

... 1730 search results

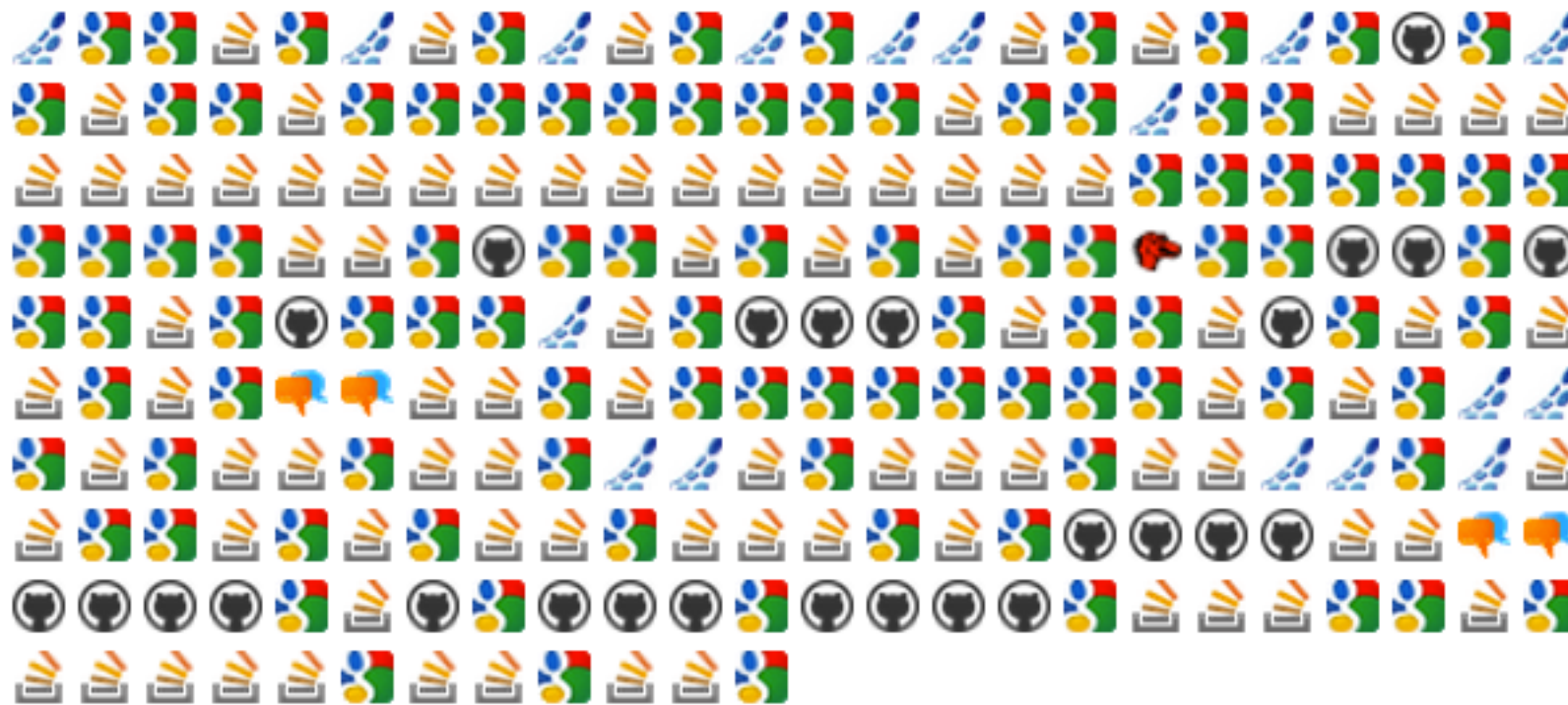| Search Result Type	Coverage	Mean Rank
code snippet	8.7%	9
q&a	9.8%	9
forum	20.2%	8
official bug tracker	21.4%	3
mailing list entry	25.4%	7
official documentation	30.1%	3
official forum	37.0%	3
unofficial documentation	63.6%	6
stackoverflow	**84.4%**	6
blog post	**87.9%**	5
official API	99.4%	1

Crowd Documentation

Knowledge is created and curated by a mostly uncoordinated collective.

An example of **Peer Production**

What sources do devs actually use?

A day of a dev

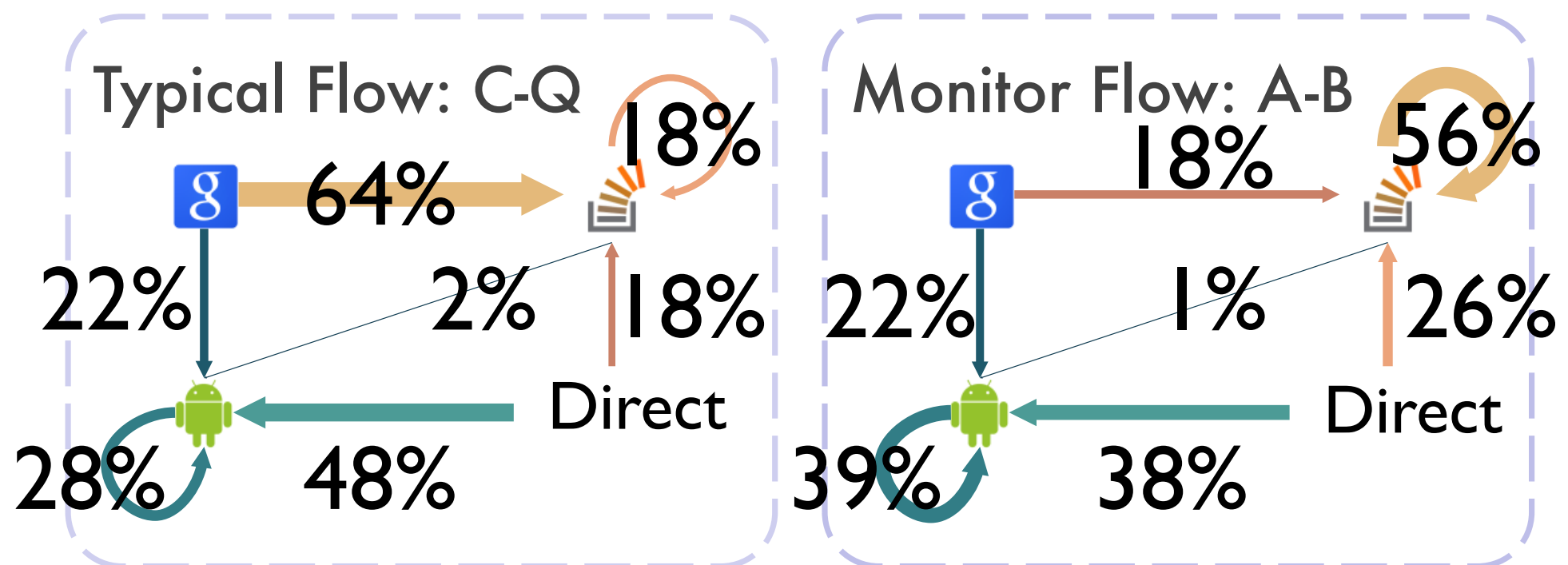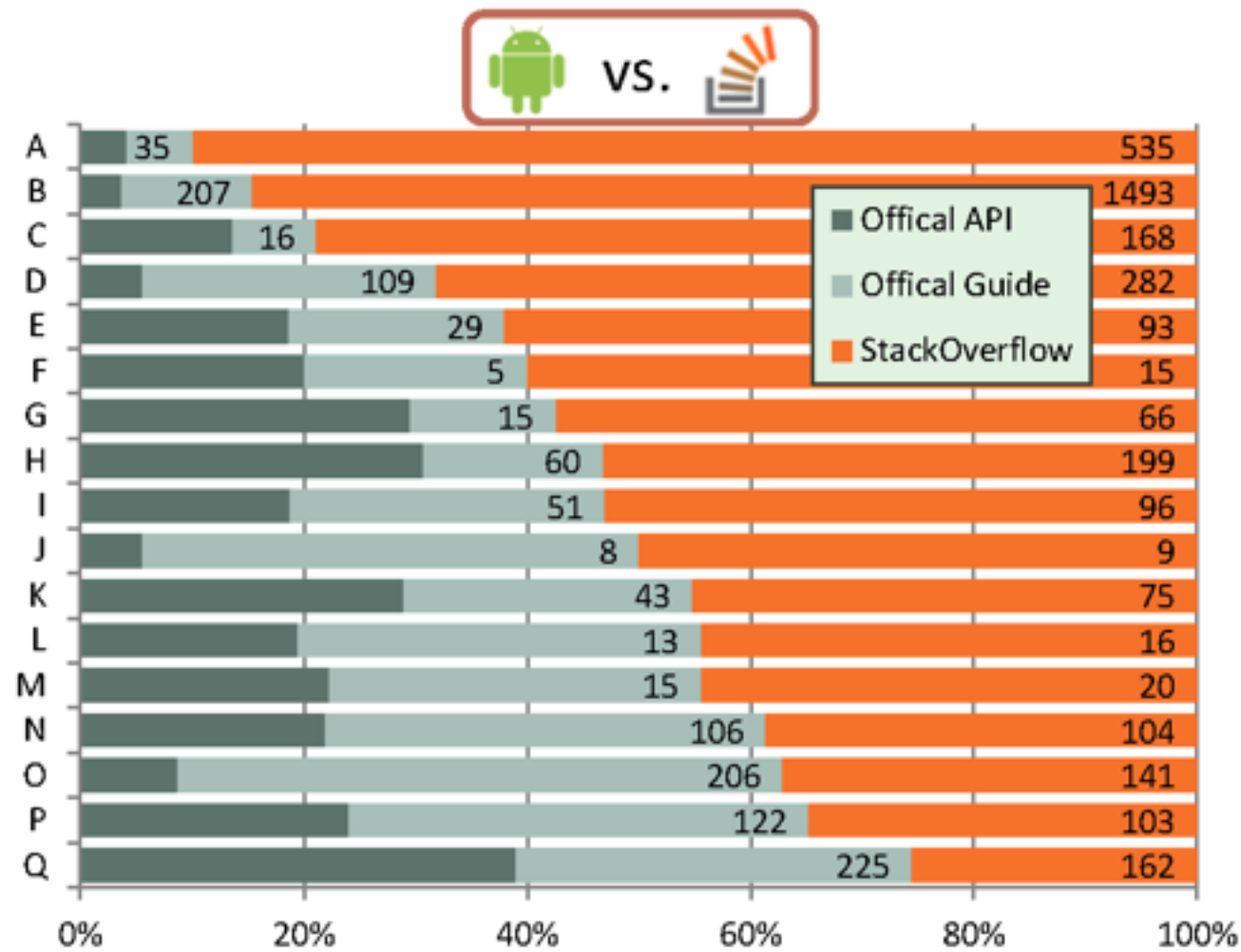Wednesday, May 15 2013

...1,316 days of developer browser history

Typical dev >

Typical Flow: C-Q

18%
64%
22%   2%   18%
Direct
28%   48%

Monitor Flow: A-B

56%
18%
22%   1%   26%
Direct
39%   38%

Consistent with self-reported surveys.

This is the data stacked up by percentage. Seems about 50% of people use Stack Overflow for about 40%+ of their documentation needs. Next, we'd need to break down troubleshooting category a little more.

Web Documentation Usage

■ Stack Overflow   ■ Trouble Shooting
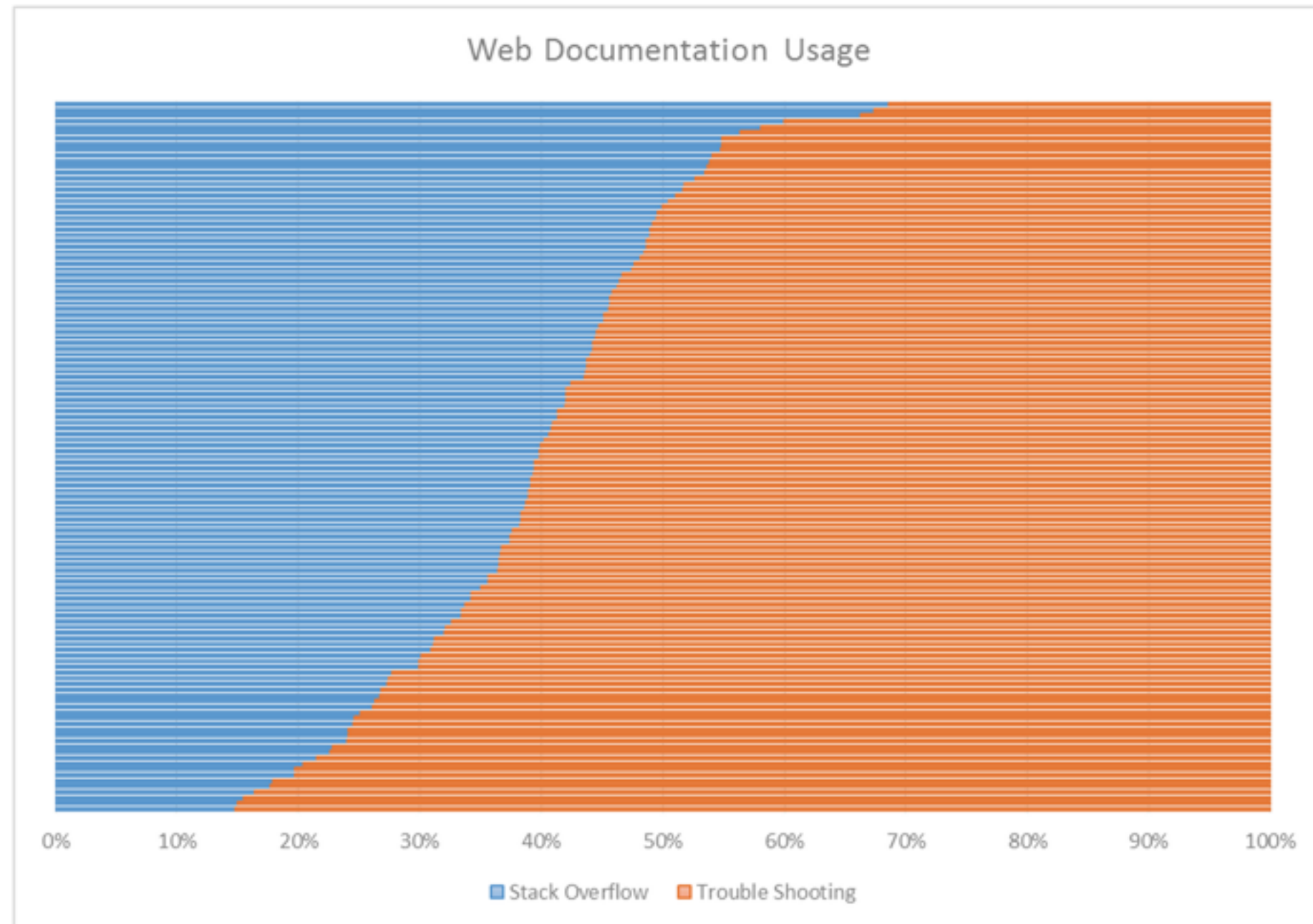
The troubleshooting category may be broken according to

Activity

- What documentation style fits you?

What makes it different?

add blobs or trees to a created repository using pygit2 (libgit2)?

▲

5

▼

I am trying to use pygit2 library.

seems I got stuck on the first step. its documentation doesn't explain how to create a blob and add it to a tree. It is mostly around how to work with an existing git repository but I want to create one and add

1 Answer

▲

5

▼

The python bindings don't let you create a blob from a file directly, so you'll have to read in the file to memory and use `Repository.write(pygit2.GIT_OBJ_BLOB, filecontents)` to create the blob.

You can then create trees with the `TreeBuilder` , for example, like

```
import pygit2 as g

repo = g.Repository('.')
# grab the file from wherever and store in 'contents'
```

5   Thanks a lot, how can I learn more about pygit2? none of what you mentioned is in pygit2 document, How did you learn how to use it? –   Peqi Hash May 5 '12 at 11:29

1   @PeqiHash Carlos is one of the developers of libgit2 – T0xicCode Dec 11 '12 at 5:19

add comment

Twice as many examples can be found on Stack Overflow than the official documentation guide.

- Developers may be getting as much as 50% of their documentation from Stack Overflow.

- More examples can be found on Stack Overflow than the official documentation guide.

- In web searches, Stack Overflow questions are visited 2x-10x more often than official documentation.

Coverage of API Classes

Speed of coverage

100
80
60
40
20%
0%

n=5
n=20
n=50
n=100

7/31/2008 7/31/2009 7/31/2010 7/31/2011

100%
80%
60%
40%
20%
0%

7/31/2008 7/31/2009 7/31/2010 7/31/2011

100%
80%
40%
20%

/31/2011
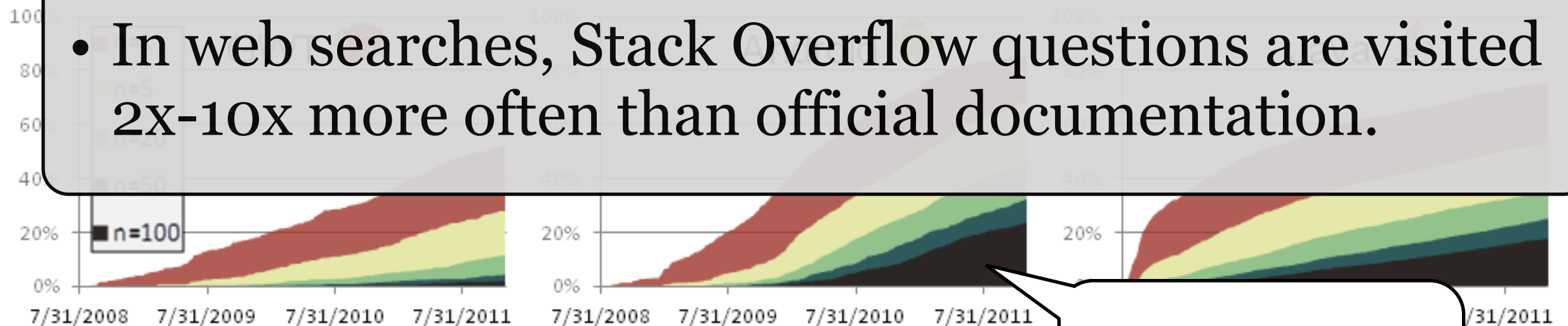
User participation
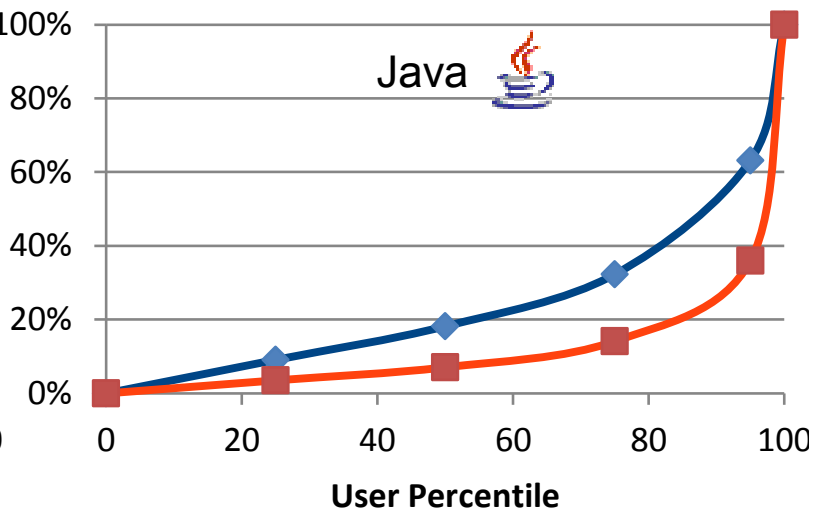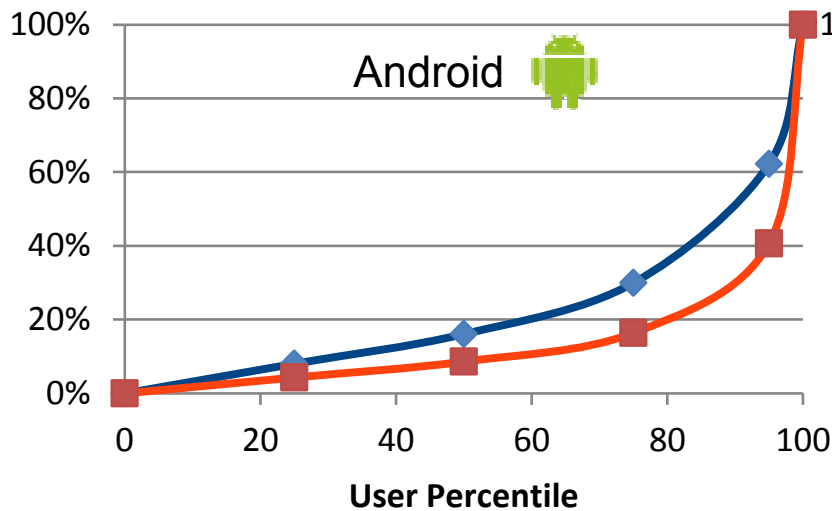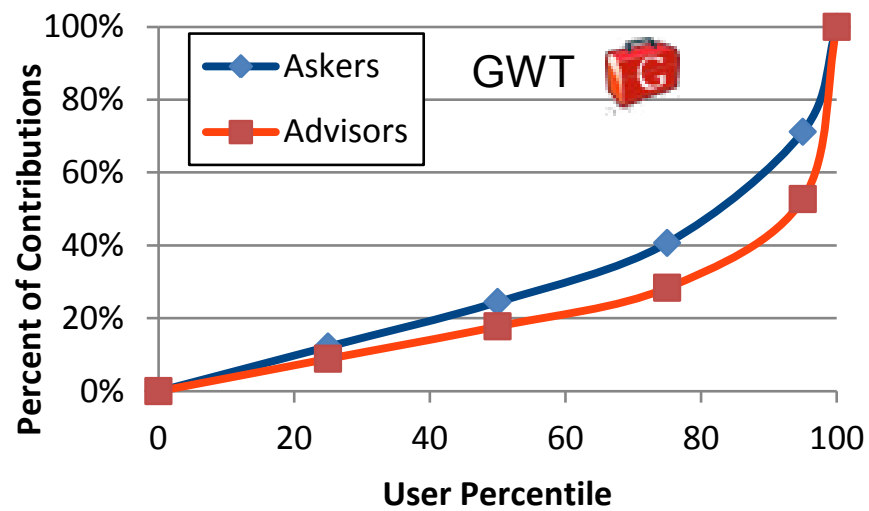
Automatic Documentation

MessageDigest - Autogenerated Crowd Documention

"MessageDigest is your friend. Call getInstance("MD5") to get an MD5 message digest you can use."

Top Threads

30 In Java, how do I convert a byte array to a string of hex digits while keeping leading zeros?

23 Performance of ThreadLocal variable

18 How do I use 3des encryption/decryption in Java?

16 Using SHA1 and RSA with java.security.Signature vs. MessageDigest and Cipher

12 How do you determine the ideal buffer size when using FileInputStream?

Popular Snippet

```java
public static String MD5hash(String text) throws NoSuchAlgorithmException {
    byte[] hash = MessageDigest.getInstance("MD5").digest(text.getBytes());
    return String.format("%032x",new BigInteger(1, hash));
}
```

Other Resources

http://commons.apache.org/codec/apidocs/org/apache/commons/codec/digest/DigestUtils.html

http://download.oracle.com/javase/6/docs/api/java/security/MessageDigest.html

http://stackoverflow.com/questions/415953/generate-md5-hash-in-java

http://en.wikipedia.org/wiki/MD5

https://www.owasp.org/index.php/Hashing_Java

Activity

- Discuss the last time your were frustrated during a programming task?

Python Tutor

So far, over 1.2 million people in over 165 countries have used Online Python Tutor to visualize over 11 million pieces of code, often as a supplement to textbooks, lecture notes, and online programming tutorials.

```
1  var array = [1,2,3,4,5];
2  var removed = array.splice(3,5);
3
4  removed;
5  array;
```

Frames

Global frame

array

removed

Objects

array

0	1	2
1	2	3

array

0	1
4	5

How frustrated were you in last coding session?

What was the cause of the problem?

60% of questions answered by 5% of users.

Diminishing game mechanisms

2 years to reach 80% coverage

Failure of online communities

http://michael.richter.name/blogs/why-i-no-longer-contribute-to-stackoverflow/

Encouraging User Behaviour with Achievements: An Empirical Study
Scott Grant and Buddy Betts

What did developer's think?

Developers May Be Getting 50% of Their Documentation From Stack Overflow

"Software companies, such as Microsoft, create documentation for millions of topics concerning its APIs, services, and software platforms. Creating this documentation comes at a considerable cost and effort. And after all this effort, much documentation is rarely consulted (citation) and lacking enough examples (citation). A new study suggests that developers are increasingly consulting Stack Overflow and crowd-sourced sites over official documentation, using it as much as 50% of time. How should official documentation be better redesigned? What are the implications of software created from unruly mashups?"

Mary N
"**Devel**
docume
Expand

John Patterson @jasper_(
MT @Kinvey developers
from **Stack Overflow:** go
Expand

Abby Fichtner @HackerCh
Yes! RT @Kinvey: Is this t
be getting **50%** of their d
Expand

Kinvey @Kinvey
Is this true for you? Study
50% of their documentat
Expand

Simon Dooms @sidooms   7 Mar
I sure do! | **Developers** May Be Getting **50%** of Their Documentation
From **Stack Overflow** developers.slashdot.org/story/13/03/05...
Expand

Jeff Atwood ✔
@codinghorror

👤▾   🐦 **Follow**

Developers get 50% of their documentation from Stack Overflow?
blog.ninlabs.com/2013/03/api-do...

← Reply   ⟲ Retweet   ⭐ Favorited   ◥ Pocket   ••• More

143
RETWEETS

67
FAVORITES

11:34 PM - 4 Mar 13

8 Mar
り **Stack Overflow** と
発者コミュニティから
ジャパン デベロッパ

developers.slashdot.jp/story/13/03/08...

8 Mar
であったのに対し、Android 用の公式ド
47 回であった": プログラマの多くは開
以上の情報を集める   goo.gl/VJW1d

7 Mar
トでしょう。"@slashdotjp:
3/03/08... #プログラミング
らくは開発者コミュニティから **50 %** 以

7 Mar
Isso é tão babadamente óbvio o porquê, que não preciso nem
comentar. fb.me/2x70Aji9y
Expand

But then...everything you thought you knew was wrong

"DLLs are tossed over the fence to English majors"

"Documentation isn't meant to be read by everyone, it's meant to be authoritative"

"We sometimes take a break from documentation and try building something for a month or two. Then we realize how bad it is"

"We contact the top 100 Stack Overflow contributors"

"You don't understand how much this has impacted us. No, really you don't."

Stakeholders Concerns

- Will we have a job in 5 years?

- What should our voice be?

- How can we "regain" authority over the crowd (e.g. people suggesting unsupported/non-public API methods in android development). [*]

- Getting feedback from the crowd: sentiment analytics over the frustrating parts, frequent questions, confusing topics, etc.

* https://code.google.com/p/android/issues/detail?id=62220

Why not participating?

Activity: Why Don't You Blog or Post on Stack Overflow?

Fear of Being "Wrong" on the Internet

"I don't blog because I fear being wrong on the internet... I share gists, or fix documentation and then tweet about it.

Blogs are disconnected from official sources, like project documentation. I can't imagine someone reading what I write. I would rather associate the content with the documentation."

Saturation:
Nothing Left to Answer

"About 2 years ago, I used to answer a lot more. But now, it feels like most questions have already been answered -- it is saturated. That might change when something new comes out, some of the newer stuff we started working on, I might have a chance to answer questions on it."

Abstraction Process	Nothing Left to Answer
Posting is Hard, Friends are Easy	Onboarding Hoops
Onboarding Hoops	**Fear of Negative Feedback**
Qualifications	Intimidating Community Size
Intimidating Community Size	No "Good Answer" Guarantee
No "Good Answer" Guarantee	Qualifications
Stranger Discomfort	Stranger Discomfort

Via interviews with developers, we've identified 14 barriers that **female participants** frequently face:

of asking, "Have you done your research?" These pressures are reinforced here where it is common for the user to enforce a culture of doing homework before asking a question. Questioning the abilities of users has discouraged them to post questions on the site.

"I think there's only one case where I was close to posting a question, but then it said do your research, this question might already be there, things like that." (P4)

Why are there fewer female users on Stack Overflow? [duplicate]

-12

This question already has an answer here:

What can Stack Overflow do to persuade female programmers to participate more? [closed]

16 answers

I've been here for a little while, but I've never seen a single female user. I know maybe there are some, but I've honestly never seen any.

What can Stack Overflow do to persuade female programmers to participate more? [closed]

48

15

I know a lot of female programmers, and I know there are a good number of them out there.

But I don't recall ever having one of my questions answered by, nor have I ever answered a question by a female programmer here at Stack Overflow.

Now I mostly deal in the Delphi tag, and that may make a difference. But I've got the feeling there is something about the programming social network that Stack Overflow is, along with its form of mini-competition (via reputation) that turns most ladies off.

When I go down the first few pages of top users by rep, there is nary a lady to be found.

Go over to Quantcast and look up Stack Overflow. Go down to the Demographics, and you'll see that a full 20% of the visits are by females. So they are here, lurking and using the wonderful info that Stack Overflow provides.

Then why aren't there more of them participating, both with questions and answers?

What can Stack Overflow do to persuade female programmers to participate more?

So after a dozen downvotes offset by a dozen upvotes, and accusations by many that I am a bigot, the bottom line seems to be that Stack Overflow appeals to males more than females.

This is my observation and I believe a proper analysis would support the hypotheses that a much higher percentage of male programmers use Stack Overflow than the percentage of female programmers.

Assuming that is the case, then Stack Overflow is geared more towards the male psyche than the female. If so, I again ask, what can they do about that?

This question was asked in StackOverflow Meta on November 20, 2009. It was made Community Wiki. It was asked in a constructive manner and should be considered an appropriate question for Meta.

Yet it was closed over two years after it was asked as "not constructive" after more than 40 upvotes and 10 favorites and 16 answers.

Now it resides here on Meta StackExchange, still closed. I just shake my head.

discussion   stackoverflow   statistics   diversity   participation

share improve this question   edited Dec 22 '14 at 5:44   community wiki
                                                                     7 revs, 3 users 77%
                                                                     lkessler

closed as not constructive by Rosinante, Pops ♦, Chris Frederick, Adam Lear ♦ Dec 31 '11 at 23:12

As it currently stands, this question is not a good fit for our Q&A format. We expect answers to be supported by facts, references, or expertise, but this question will likely solicit debate, arguments, polling, or extended discussion. If you feel that this question can be improved and possibly reopened, visit the help center for guidance.

If this question can be reworded to fit the rules in the help center, please edit the question.

Welcome!

Welcome! Meta Stack Exchange is intended for bugs, features, and discussions that affect the whole Stack Exchange family of Q&A sites.

about »   help »

asked   6 years ago

viewed   6535 times

active   1 year ago

BLOG

Why Stack Overflow Doesn't Care About Ad Blockers

72 People Chatting

Tavern on the Meta
1 hour ago - bjb568

Shadow's Den
1 hour ago - Mooseman

Love this site?

Get the **weekly newsletter!** In it, you'll get:

- The week's top questions and answers
- Important community announcements
- Questions that need answers

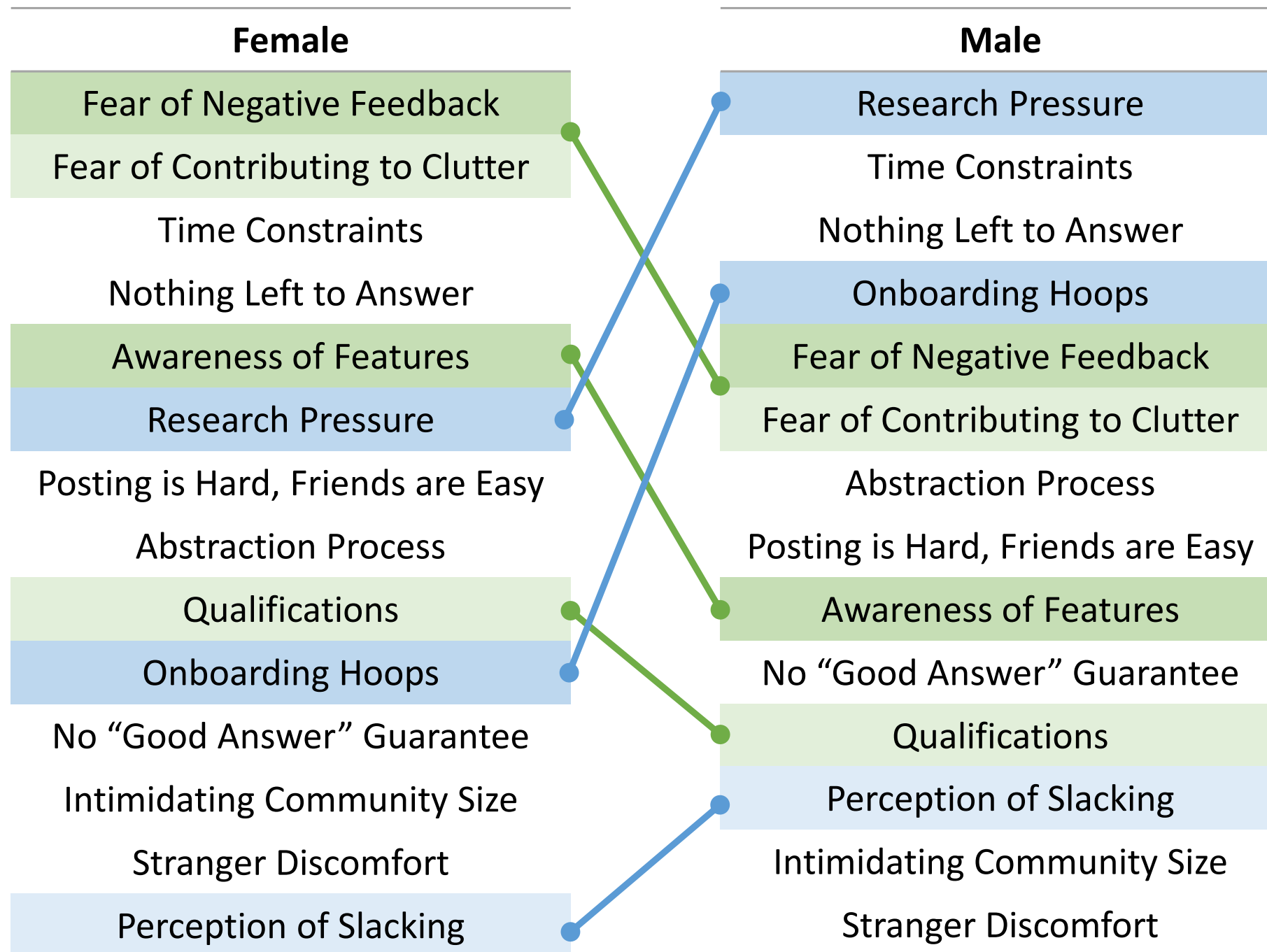Sign up for the newsletter

see an example newsletter
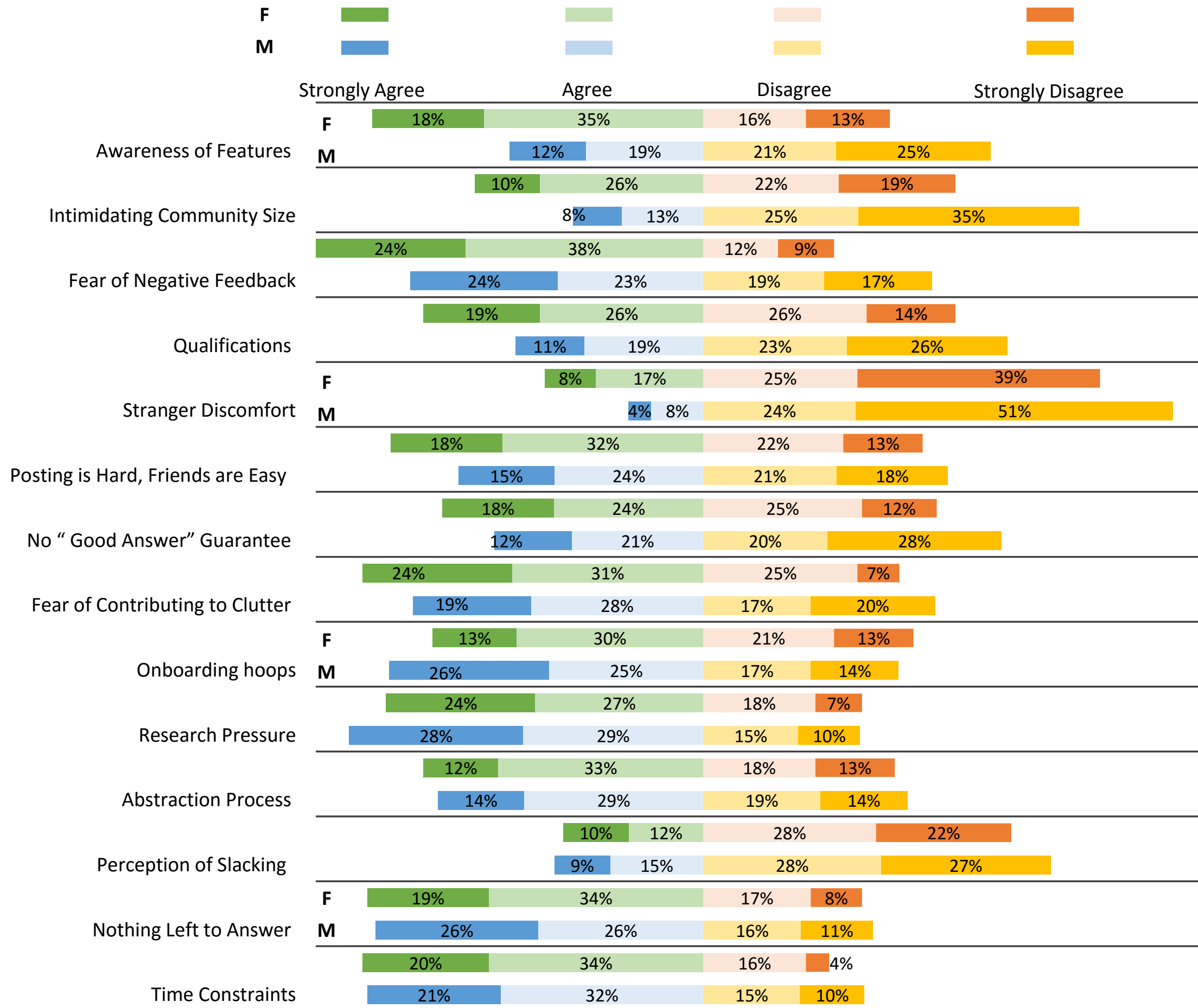
Linked

1   Stackoverflow is male dominated

-12 Why are there fewer female users on Stack Overflow?
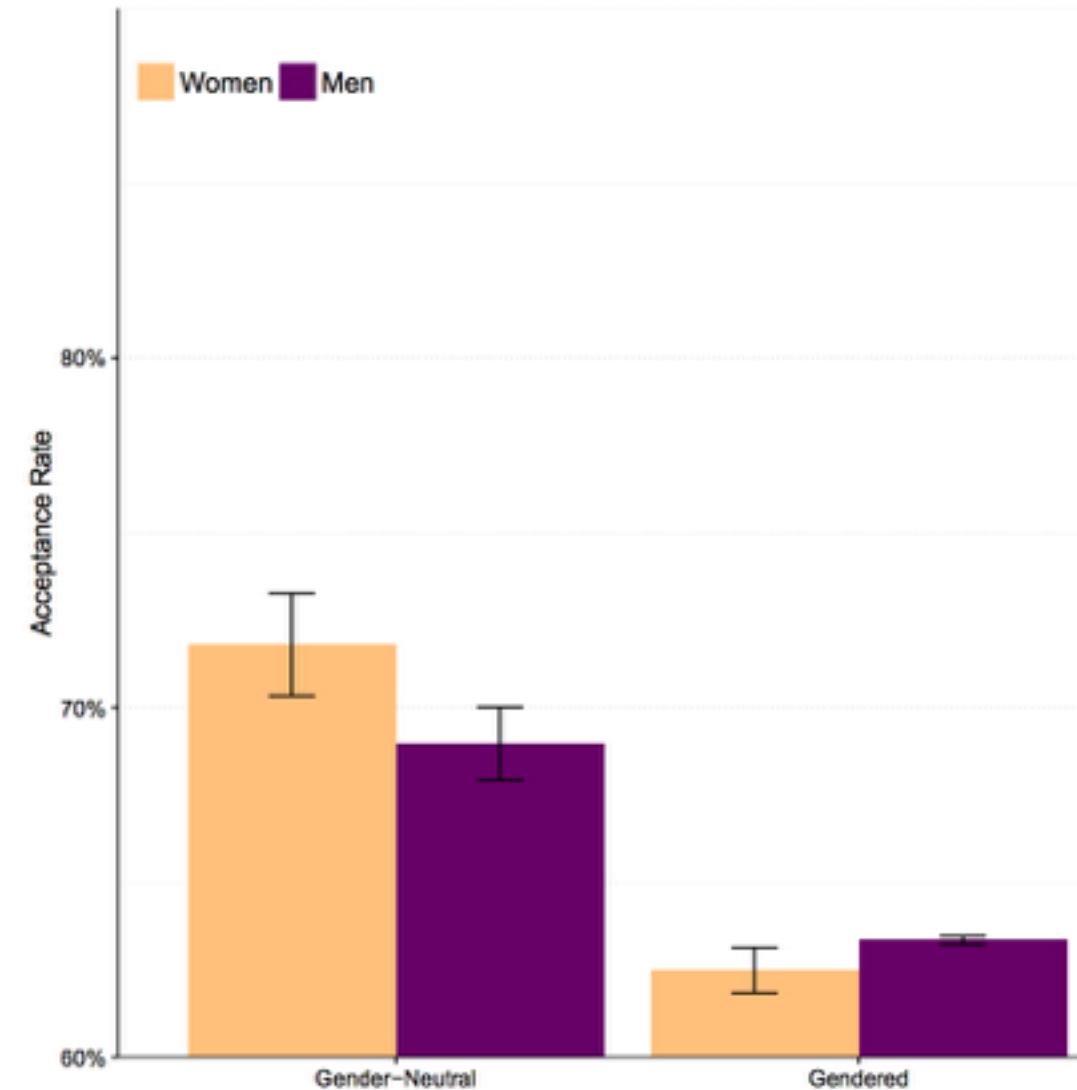
508 The Many Memes of Meta

Female and Male Rankings

	Strongly Agree	Agree	Disagree	Strongly Disagree
Awareness of Features F	18%	35%	16%	13%
M	12%	19%	21%	25%
Intimidating Community Size F	10%	26%	22%	19%
M	8%	13%	25%	35%
Fear of Negative Feedback F	24%	38%	12%	9%
M	24%	23%	19%	17%
Qualifications F	19%	26%	26%	14%
M	11%	19%	23%	26%
Stranger Discomfort F	8%	17%	25%	39%
M	4%	8%	24%	51%
Posting is Hard, Friends are Easy F	18%	32%	22%	13%
M	15%	24%	21%	18%
No " Good Answer" Guarantee F	18%	24%	25%	12%
M	12%	21%	20%	28%
Fear of Contributing to Clutter F	24%	31%	25%	7%
M	19%	28%	17%	20%
Onboarding hoops F	13%	30%	21%	13%
M	26%	25%	17%	14%
Research Pressure F	24%	27%	18%	7%
M	28%	29%	15%	10%
Abstraction Process F	12%	33%	18%	13%
M	14%	29%	19%	14%
Perception of Slacking F	10%	12%	28%	22%
M	9%	15%	28%	27%
Nothing Left to Answer F	19%	34%	17%	8%
M	26%	26%	16%	11%
Time Constraints F	20%	34%	16%	4%
M	21%	32%	15%	10%

Github Gender Bias

Actual Crowd

Warlords of Documentation: A Proposed Expansion of Stack Overflow

It's been 7 years and 10,000,000+ Questions since Stack Overflow was launched. The amount of good that has been done for the field - all the developers helped, all the person-hours saved, all the beginners who grew into professionals - is hard to overstate. I cannot express how proud I am of what we've built together (and not just because it let me see our arch-enemy vanish from my Google results); this community has realized the goal of "making the internet a better place" beyond what anyone could have predicted.

Lately we've been asking ourselves "what else could we do to improve developers' lives on the internet?". Jeff's original announcement of Stack Overflow said this:

> There's far too much great programming information trapped in forums, buried in online help, or hidden away in books that nobody buys any more. We'd like to unlock all that. Let's create something that makes it easy to participate, and put it online in a form that is trivially easy to find.

Stack Overflow has made all of that a *lot* better, but there's one area that is still hanging around: **Documentation**. Just like Q&A in 2008, Documentation in 2015 is something every developer needs regularly, and something that by most appearances stopped improving in 1996. We think, together, we can make it a lot better.

How would adding Documentation to Stack Overflow improve Documentation?

1. Documentation is often an afterthought, obviously done just to say that it exists, with little concern for its quality. We've all seen plenty of this in our time.

Jamie Dixon
@jamiedixon   ⚙   +🧑 Follow

Got to love some commercial API documentation: "GetRowEnumerator(); - Gets the rows enumerator" - No further information.

RETWEETS   FAVORITES

Big Code

- https://sourcegraph.com/

- statistical linting

- API Miner

Crowd Programming

- We have seen "crowd documentation" how might the other pieces look like?

Rather than having a rich standard API, Javascript essentially has a "**crowd API**" assembled from Stack Overflow snippets and Github repositories.

Project DNA: Track and understand creation and use of javascript apis.
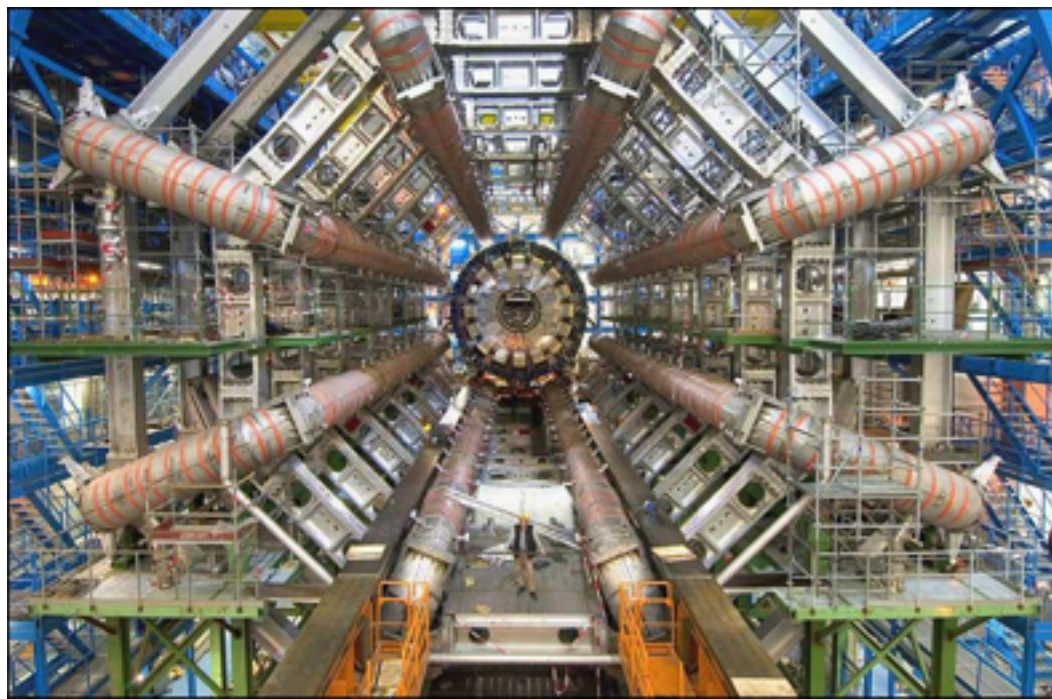
two decades

Massively distributed software engineering

In the next 50 years, as governments increasingly turn legal policy and services into source code and public APIs, often created in the timespan of a president's term, we must be prepared to build massively-sized software systems on a regular basis. This will often require cooperation of many diverse stakeholders.

health care insurance marketplace

At our current place, imagine how these would fare if needed in a few years:

- A government api to calculate taxes on all online purchases for any location?
- A distributed traffic regulation system for a network of self-driving cars and drones.
- ...