

Simplifying the contribution process for both contributors & maintainers

A case study of the  integration in  easybuild

GitHub



FOSDEM 2018

Testing and Automation devroom

Feb 3rd 2018, Brussels (Belgium)

Kenneth Hoste

kenneth.hoste@ugent.be

Twitter: @kehoste



*GitHub is a code hosting service
for git repositories*

<https://github.com>

- launched in April 2008
- currently most popular platform for collaborating on FOSS
 - over 24M users, 67M repositories
 - over 100M *merged* pull requests
- also provides issue tracker, code review facilities, wiki, ...
- used by Docker, Python, TensorFlow, (Linux kernel), ...

Example: one of the EasyBuild repositories on GitHub

The screenshot shows the GitHub interface for the repository 'easybuilders / easybuild-easyconfigs'. The top navigation bar includes 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. On the right, there are icons for notifications, a plus sign, and a user profile. Below the repository name, there are buttons for 'Unwatch' (28), 'Star' (89), and 'Fork' (280). The 'Issues' tab is selected, showing 213 issues. Below the tabs, there are filters for 'is:issue is:open', 'Labels', and 'Milestones', along with a 'New issue' button. The main content area displays a list of issues with columns for checkboxes, status, title, labels, projects, milestones, assignee, and sort. The issues listed are:

- 213 Open** ✓ 248 Closed Author Labels Projects Milestones Assignee Sort
- Fail to build `compiler/GCCcore/6.*.0` under vagrant box/ In: failed to create hard link** **problem report** 2
- OpenMPI-2.1.2-GCC-6.4.0-2.28.eb dependency name** **problem report** 4
- EasyBuild installation aborts after warning on mismatch in automake tools versions** **bug report** 1
- Theano may fail when running in a heterogeneous cluster** **problem report**
- Python 2.7.x incompatible with OpenSSL 1.1.0 (Debian 9)** **problem report**

Example: one of the EasyBuild repositories on GitHub

integrated issue tracker

The screenshot shows the GitHub interface for the repository `easybuilders/easybuild-easyconfigs`. The top navigation bar includes links for "This repository", "Search", "Pull requests", "Issues", "Marketplace", and "Explore". The repository name is displayed as `easybuilders / easybuild-easyconfigs`. On the right, there are buttons for "Unwatch" (28), "Star" (89), and "Fork" (280). Below the repository name, there are tabs for "Code", "Issues 213", "Pull requests 470", "Projects 0", "Insights", and "Settings". The "Issues" tab is selected and highlighted with a white box. Below the tabs, there is a search bar with the filter `is:issue is:open` and buttons for "Labels" and "Milestones". A green "New issue" button is located on the right. The main content area displays a list of issues with the following details:

<input type="checkbox"/>	213 Open ✓ 248 Closed	Author	Labels	Projects	Milestones	Assignee	Sort
<input type="checkbox"/>	Fail to build `compiler/GCCcore/6.*.0` under vagrant box/ In: failed to create hard link problem report						2
	#5685 opened 2 days ago by Falkor 3.x						
<input type="checkbox"/>	OpenMPI-2.1.2-GCC-6.4.0-2.28.eb dependency name problem report						4
	#5649 opened 6 days ago by RvDijk						
<input type="checkbox"/>	EasyBuild installation aborts after warning on mismatch in automake tools versions bug report						1
	#5644 opened 6 days ago by casparvl						
<input type="checkbox"/>	Theano may fail when running in a heterogeneous cluster problem report						
	#5643 opened 6 days ago by smoors						
<input type="checkbox"/>	Python 2.7.x incompatible with OpenSSL 1.1.0 (Debian 9) problem report						
	#5614 opened 12 days ago by vanzod						

Example: one of the EasyBuild repositories on GitHub

GitHub interface for repository `easybuilders / easybuild-easyconfigs`. The page shows navigation options (Code, Issues, Pull requests, Projects, Insights, Settings) and a list of pull requests. The pull requests list includes columns for checkboxes, repository icons, titles, status, update buttons, and comment counts.

Navigation: `<> Code`, `Issues 213`, `Pull requests 470`, `Projects 0`, `Insights`, `Settings`

Filters: `is:pr is:open sort:updated-desc`, `Labels`, `Milestones`, `New pull request`

Clear current search query, filters, and sorts

<input type="checkbox"/>		470 Open	✓ 4,764 Closed	Author ▾	Labels ▾	Projects ▾	Milestones ▾	Reviews ▾	Assignee ▾	Sort ▾
<input type="checkbox"/>		<code>{numlib}[GCCcore/6.4.0]</code>	<code>TetGen v1.5.0</code>	✓	<code>new</code>					9
		#5681 opened 3 days ago by smoors • Changes requested ⌚ updated 25 seconds ago 🚦 next release (3...								
<input type="checkbox"/>		<code>{foss/2017b}</code>	<code>add Tkinter as dependency for ASE & matplotlib using Python 3.6.3</code>	✓	<code>bug fix</code>					20
		#5658 opened 5 days ago by schiotz • Approved ⌚ updated 2 minutes ago 🚦 next release (3...								
<input type="checkbox"/>		<code>{bio}[foss/2018a]</code>	<code>ABySS 2.0.2</code>	✗	<code>update</code>					2
		#5688 opened a day ago by RvDijk • Approved ⌚ updated 6 minutes ago 🚦 next release (3...								
<input type="checkbox"/>		<code>{devel}[foss/2018a]</code>	<code>Boost 1.66.0 (with and without Python)</code>	✗	<code>update</code>					2
		#5687 opened a day ago by RvDijk • Changes requested ⌚ updated 7 minutes ago 🚦 next release (3...								
<input type="checkbox"/>		<code>Add Tkinter sanity check to matplotlib (foss toolchains)</code>	✓	<code>enhancement</code>						3
		#5689 opened a day ago by schiotz • Review required ⌚ updated 16 minutes ago 🚦 next release (3...								

Example: one of the EasyBuild repositories on GitHub

"pull request" interface for incoming contributions

The screenshot displays the GitHub interface for the repository `easybuilders / easybuild-easyconfigs`. The navigation bar includes links for `Code`, `Issues 213`, `Pull requests 470` (highlighted), `Projects 0`, `Insights`, and `Settings`. The main content area shows a list of pull requests with the following details:

<input type="checkbox"/>		470 Open	✓ 4,764 Closed	Author	Labels	Projects	Milestones	Reviews	Assignee	Sort
<input type="checkbox"/>		<code>{numlib}[GCCcore/6.4.0]</code>	<code>TetGen v1.5.0</code>	✓	<code>new</code>					9
#5681 opened 3 days ago by smoors • Changes requested ⌚ updated 25 seconds ago ⚙ next release (3...										
<input type="checkbox"/>		<code>{foss/2017b}</code>	<code>add Tkinter as dependency for ASE & matplotlib using Python 3.6.3</code>	✓	<code>bug fix</code>					20
#5658 opened 5 days ago by schiotz • Approved ⌚ updated 2 minutes ago ⚙ next release (3...										
<input type="checkbox"/>		<code>{bio}[foss/2018a]</code>	<code>ABySS 2.0.2</code>	✗	<code>update</code>					2
#5688 opened a day ago by RvDijk • Approved ⌚ updated 6 minutes ago ⚙ next release (3...										
<input type="checkbox"/>		<code>{devel}[foss/2018a]</code>	<code>Boost 1.66.0 (with and without Python)</code>	✗	<code>update</code>					2
#5687 opened a day ago by RvDijk • Changes requested ⌚ updated 7 minutes ago ⚙ next release (3...										
<input type="checkbox"/>		<code>Add Tkinter sanity check to matplotlib (foss toolchains)</code>	✓	<code>enhancement</code>						3
#5689 opened a day ago by schiotz • Review required ⌚ updated 16 minutes ago ⚙ next release (3...										

Example: one of the EasyBuild repositories on GitHub

(basic) project management features: labels, milestones

The screenshot shows the GitHub interface for the repository 'easybuilders / easybuild-easyconfigs'. The top navigation bar includes 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. The repository name is displayed as 'easybuilders / easybuild-easyconfigs' with options to 'Unwatch', 'Star' (89), and 'Fork' (280). Below the repository name, there are tabs for 'Code', 'Issues' (213), 'Pull requests' (470), 'Projects' (0), 'Insights', and 'Settings'. A search bar contains the query 'is:pr is:open sort:updated-desc'. To the right of the search bar are buttons for 'Labels' and 'Milestones', and a green 'New pull request' button. Below the search bar, there is a link to 'Clear current search query, filters, and sorts'. The main content area shows a list of pull requests with columns for checkboxes, status (Open/Closed), author, labels, projects, milestones, reviews, assignee, and sort. The pull requests are:

- 470 Open ✓ 4,764 Closed Author Labels Projects Milestones Reviews Assignee Sort
- {numlib}[GCCcore/6.4.0] TetGen v1.5.0 ✓ **new** 9 comments
#5681 opened 3 days ago by smoors • Changes requested updated 25 seconds ago next release (3...)
- {foss/2017b} add Tkinter as dependency for ASE & matplotlib using Python 3.6.3 ✓ **bug fix** 20 comments
#5658 opened 5 days ago by schiotz • Approved updated 2 minutes ago next release (3...)
- {bio}[foss/2018a] ABySS 2.0.2 ✗ **update** 2 comments
#5688 opened a day ago by RvDijk • Approved updated 6 minutes ago next release (3...)
- {devel}[foss/2018a] Boost 1.66.0 (with and without Python) ✗ **update** 2 comments
#5687 opened a day ago by RvDijk • Changes requested updated 7 minutes ago next release (3...)
- Add Tkinter sanity check to matplotlib (foss toolchains) ✓ **enhancement** 3 comments
#5689 opened a day ago by schiotz • Review required updated 16 minutes ago next release (3...)

Example: one of the EasyBuild repositories on GitHub

integration with Continuous Integration services like Travis CI, Circle CI, ...


The screenshot shows the GitHub interface for the repository `easybuilders/easybuild-easyconfigs`. The top navigation bar includes links for "This repository", "Search", "Pull requests", "Issues", "Marketplace", and "Explore". The repository name is displayed as `easybuilders / easybuild-easyconfigs`. On the right, there are buttons for "Unwatch" (28), "Star" (89), and "Fork" (280). Below the repository name, there are tabs for "Code", "Issues" (213), "Pull requests" (470), "Projects" (0), "Insights", and "Settings". A search bar contains the query `is:pr is:open sort:updated-desc`. A "New pull request" button is visible on the right. A "Clear current search query, filters, and sorts" button is located below the search bar. The main content area displays a list of pull requests with the following details:




<input type="checkbox"/>		470 Open	✓ 4,764 Closed	Author	Labels	Projects	Milestones	Reviews	Assignee	Sort
<input type="checkbox"/>		<code>{numlib}[GCCcore/6.4.0] TetGen v1.5.0</code>	✓ new	#5681	opened 3 days ago by smoores	Changes requested	updated 25 seconds ago	next release (3...	9	
<input type="checkbox"/>		<code>{foss/2017b} add Tkinter as dependency for ASE & matplotlib using Python 3.6.3</code>	✓ bug fix	#5658	opened 5 days ago by schiotz	Approved	updated 2 minutes ago	next release (3...	20	
<input type="checkbox"/>		<code>{bio}[foss/2018a] ABySS 2.0.2</code>	✗ update	#5688	opened a day ago by RvDijk	Approved	updated 6 minutes ago	next release (3...	2	
<input type="checkbox"/>		<code>{devel}[foss/2018a] Boost 1.66.0 (with and without Python)</code>	✗ update	#5687	opened a day ago by RvDijk	Changes requested	updated 7 minutes ago	next release (3...	2	
<input type="checkbox"/>		<code>Add Tkinter sanity check to matplotlib (foss toolchains)</code>	✓ enhancement	#5689	opened a day ago by schiotz	Review required	updated 16 minutes ago	next release (3...	3	


Example: one of the EasyBuild repositories on GitHub

code review interface for contributions

build recent PLUMED versions with all (PLUMED) modules enabled #5696

 Open boegel wants to merge 1 commit into `easybuilders:develop` from `boegel:PLUMED_modules_all`


 Conversation **2**  Commits **1**  Files changed **5**

Changes from all commits ▾ Jump to... ▾ **+5 -5** 

Unified

Split

Review changes ▾

2  easybuild/easyconfigs/p/PLUMED/PLUMED-2.3.0-foss-2016b.eb

Show comments

View



@@ -26,7 +26,7 @@ dependencies = [

26 26

]

27 27

28 28

preconfigopts = 'env FC=\$MPIF90 LIBS="\$LIBLAPACK \$LIBS" '

29

-configopts = ' --exec-prefix=%(installdir)s --enable-gsl'

29

+configopts = ' --exec-prefix=%(installdir)s --enable-gsl --enable-modules=all'

Write

Preview

AA ▾ B i



Why this isn't enabled by default is beyond me...

Attach files by dragging & dropping, [selecting them](#), or pasting from the clipboard.

 Styling with Markdown is supported

Cancel

Add single comment

Start a review



*Git is a free and open source (GPLv2)
distributed version control system*

<https://git-scm.com>

- created in 2005 by Linus Torvalds
- currently (by far) most popular version control system (VCS)
- design goals:
 - good performance, even for large projects (Linux kernel)
 - support for non-linear development & distributed workflow
 - to not be like CVS (Concurrent Versions System)
- internals similar to modern *blockchain* technologies (Merkle trees)
- used by Linux kernel, Windows, Android, Python, TensorFlow, ...

Git workflow: branch, stage, commit, push

```
# edit file to contribute
```

```
$ vim file
```

```
# create branch, and check it out
```

```
$ git checkout -b branch_name
```

```
# stage file
```

```
$ git add file
```

```
# commit changes with sensible commit message
```

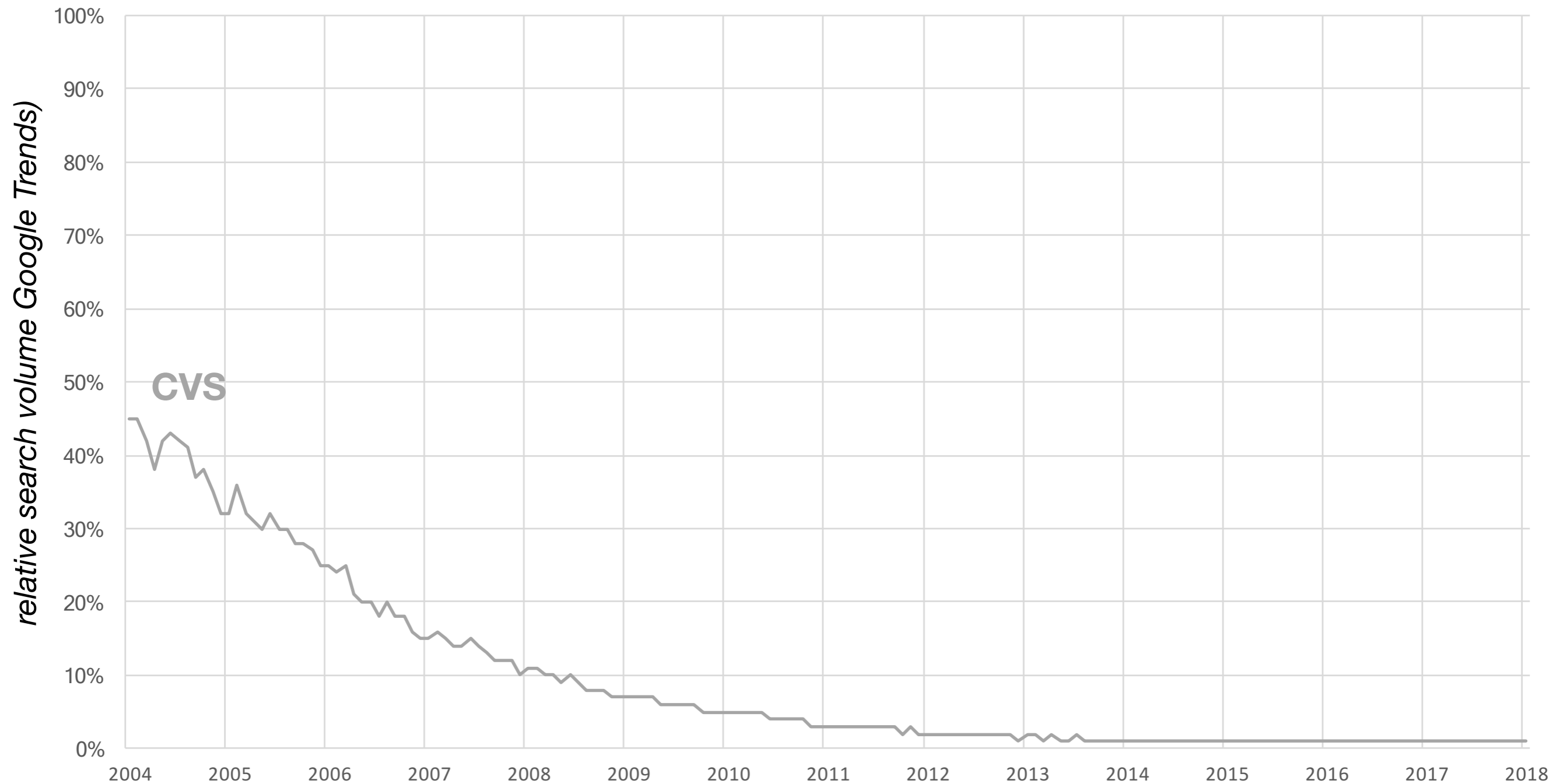
```
$ git commit -m "this is a sensible commit message"
```

```
# push branch to remote repository
```

```
$ git push remote_repo branch_name
```

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



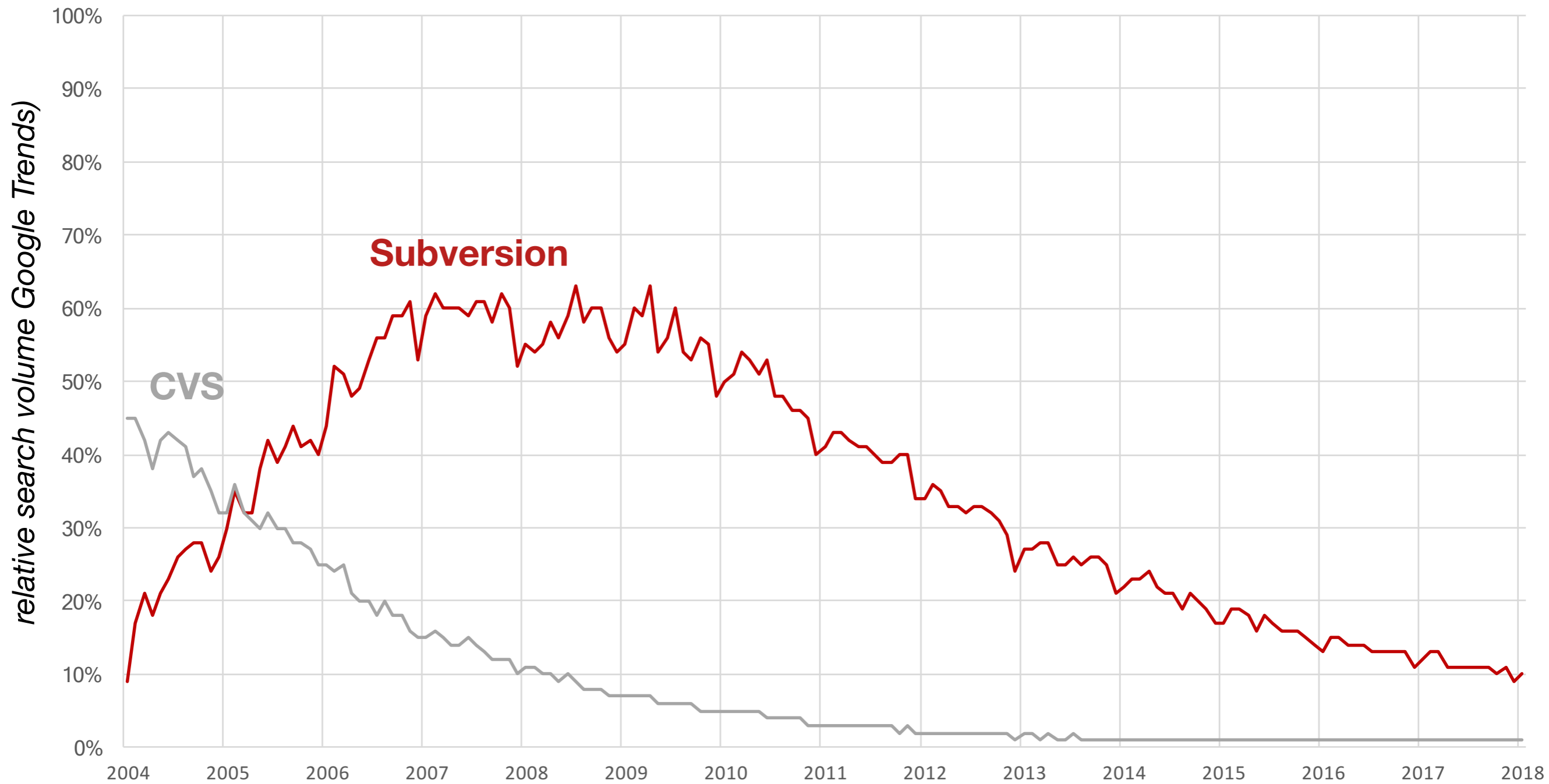
Concurrent Versions System (CVS)

<http://www.nongnu.org/cvs>

initial release: Nov 1990, last release: May 2008

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



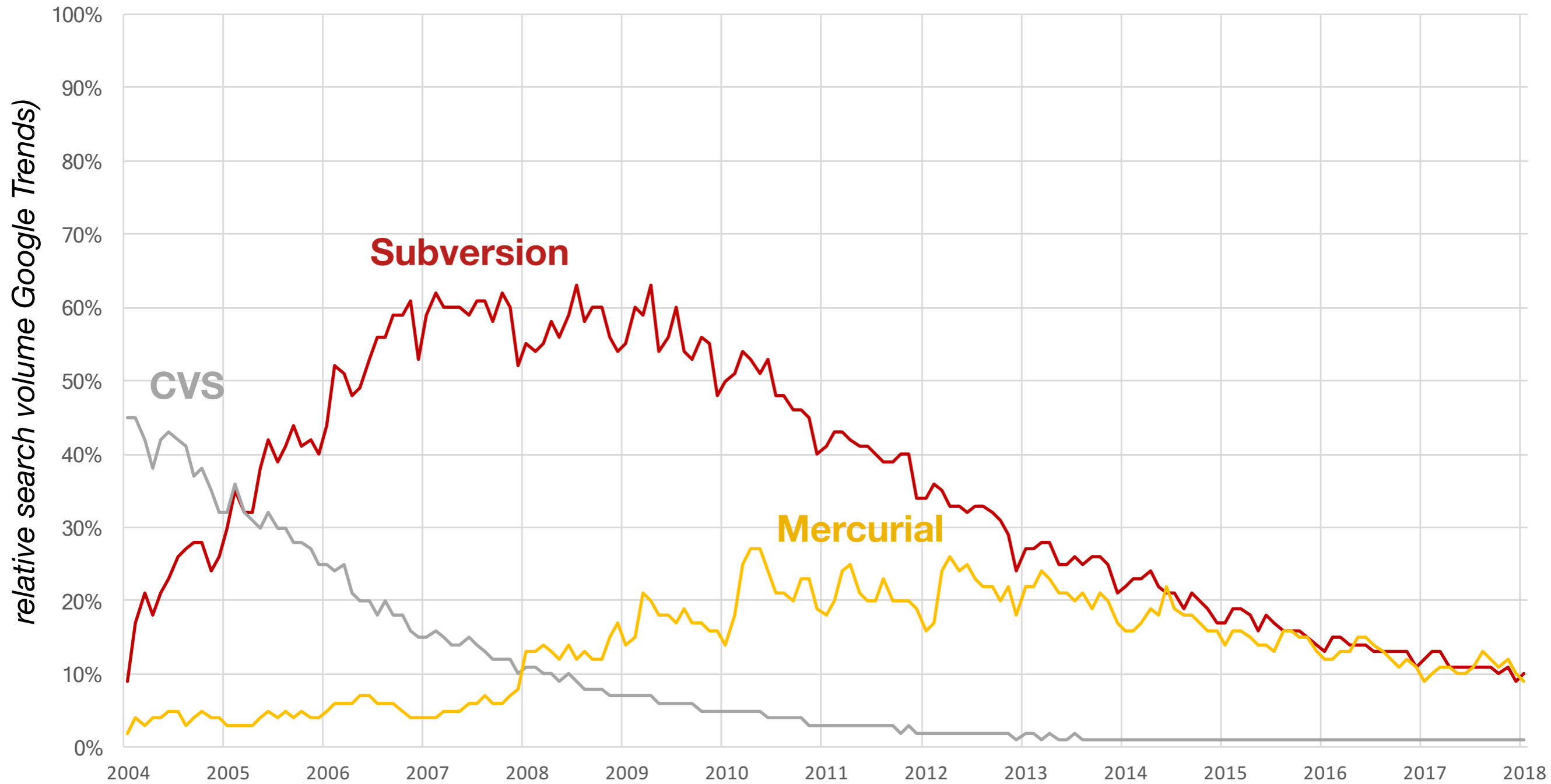
Apache Subversion (SVN)

<https://subversion.apache.org>

initial release: Oct 2000, last release: Aug 2017

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



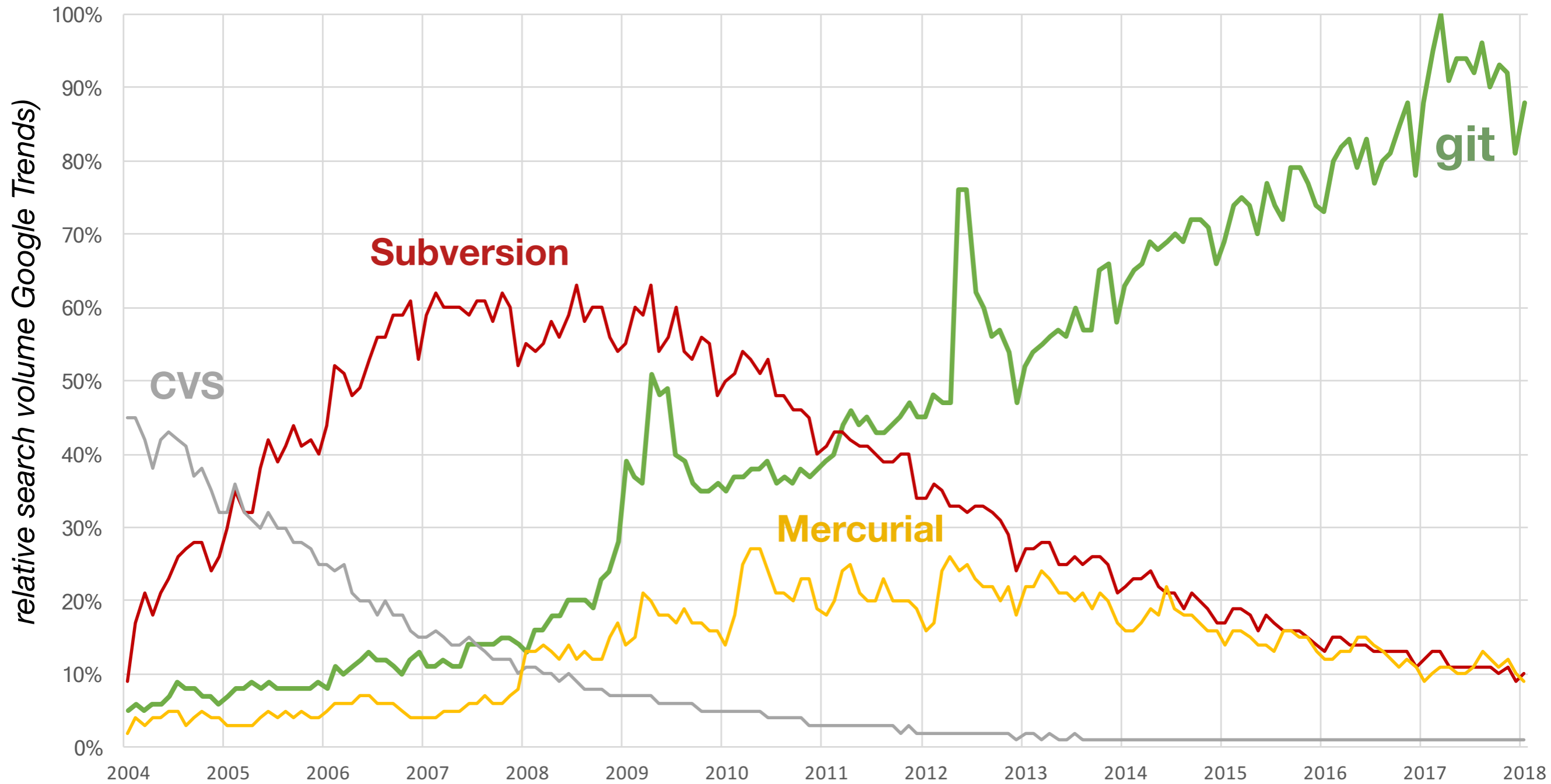
Mercurial

<https://www.mercurial-scm.org>

initial release: April 2005, last release: Dec 2017

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



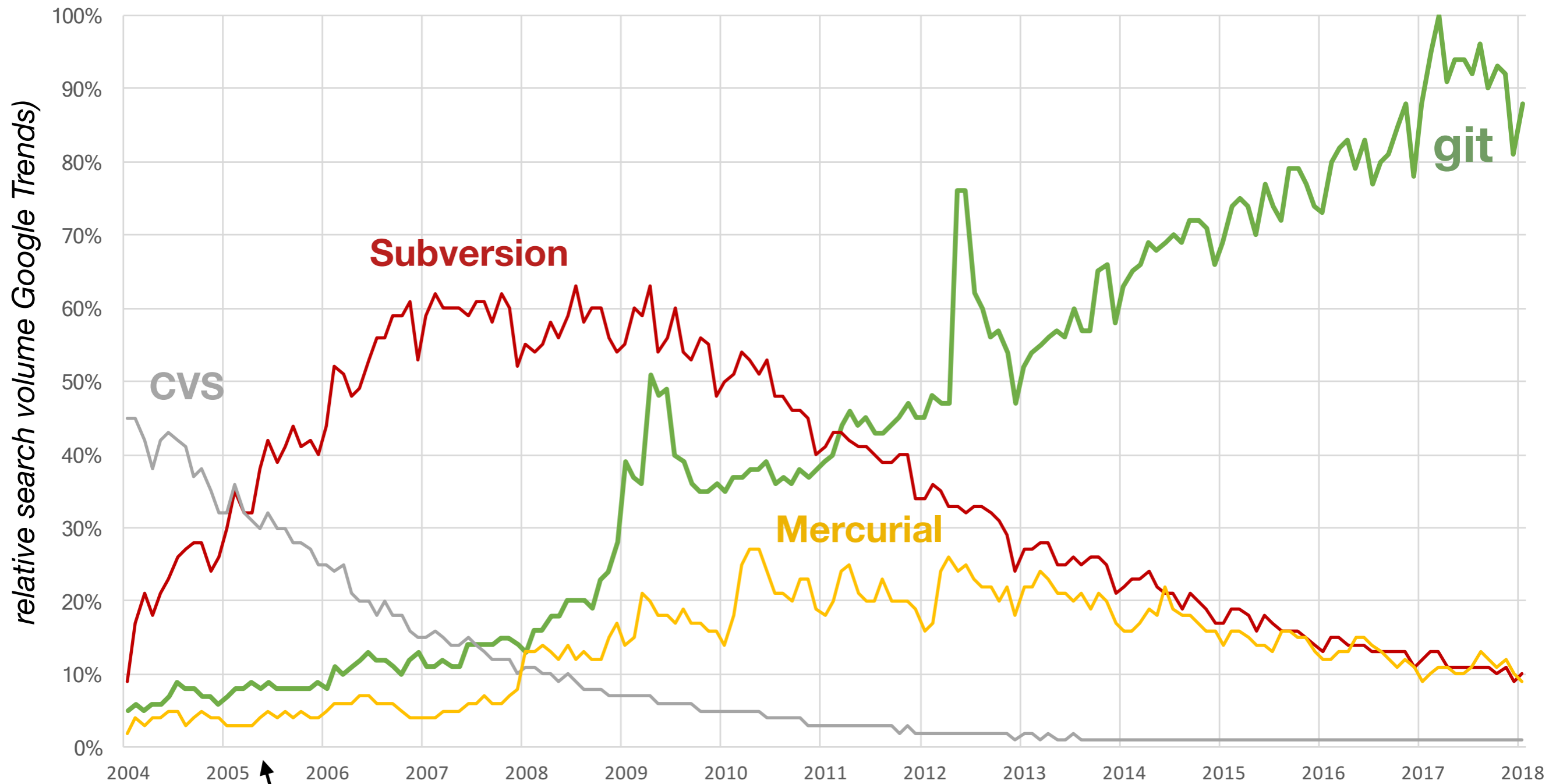
git

<https://git-scm.com>

initial release: April 2005, last release: Jan 2018

Why is git so popular?

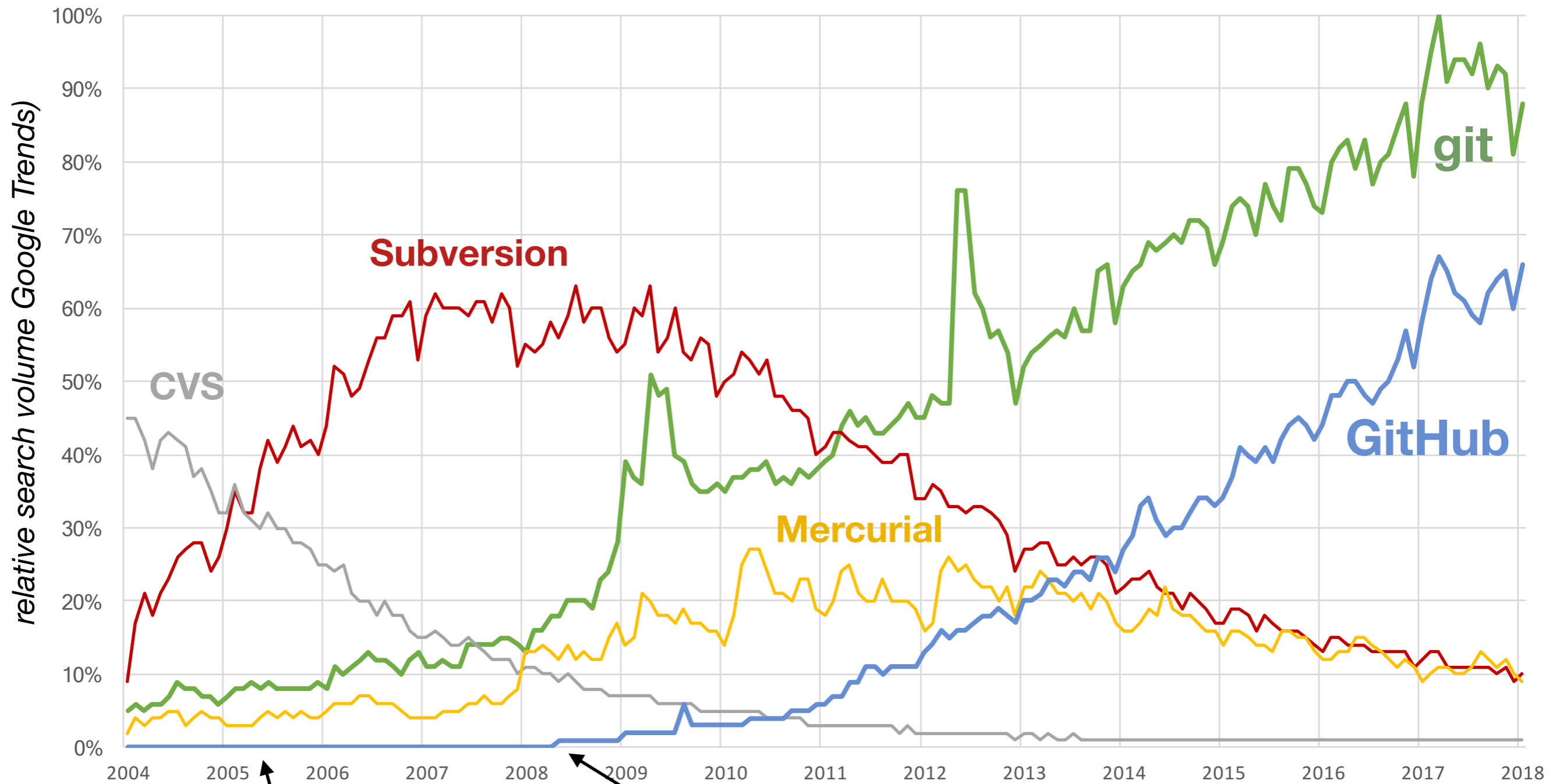
(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



 **git** initial release
(April 2005)

Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



git

initial release
(April 2005)

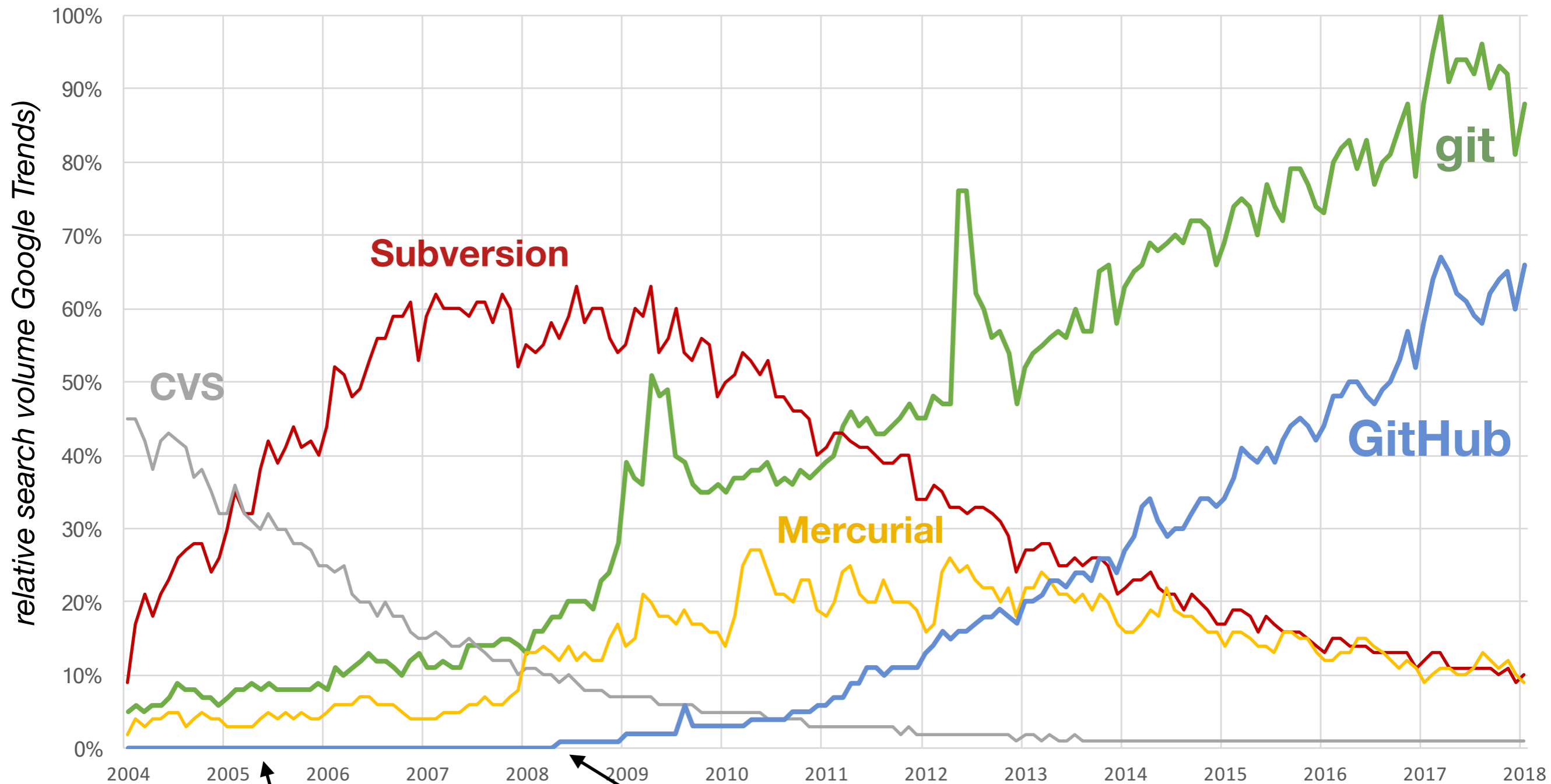


GitHub

public launch
(April 2008)

Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



git

initial release
(April 2005)

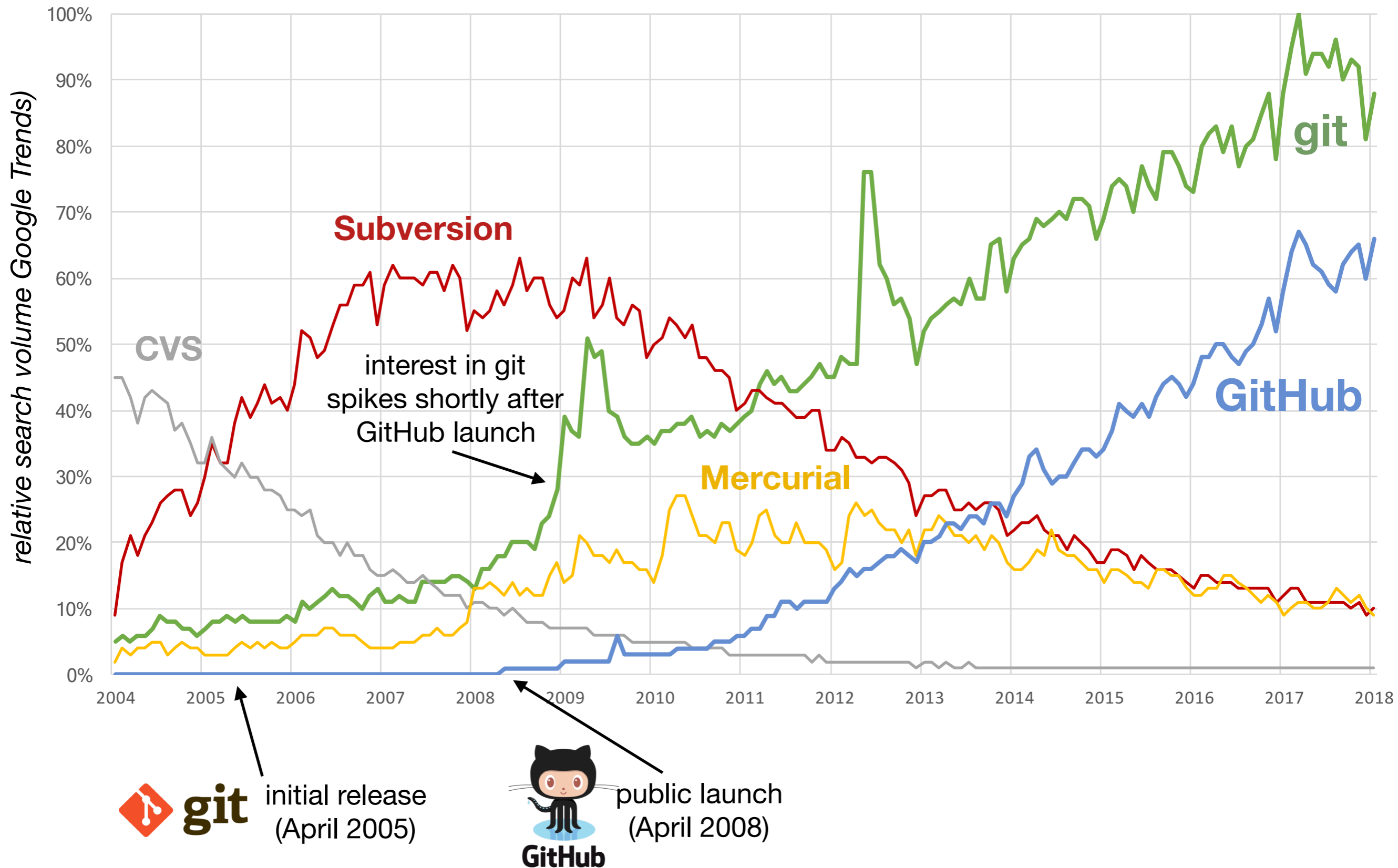


GitHub

public launch
(April 2008)

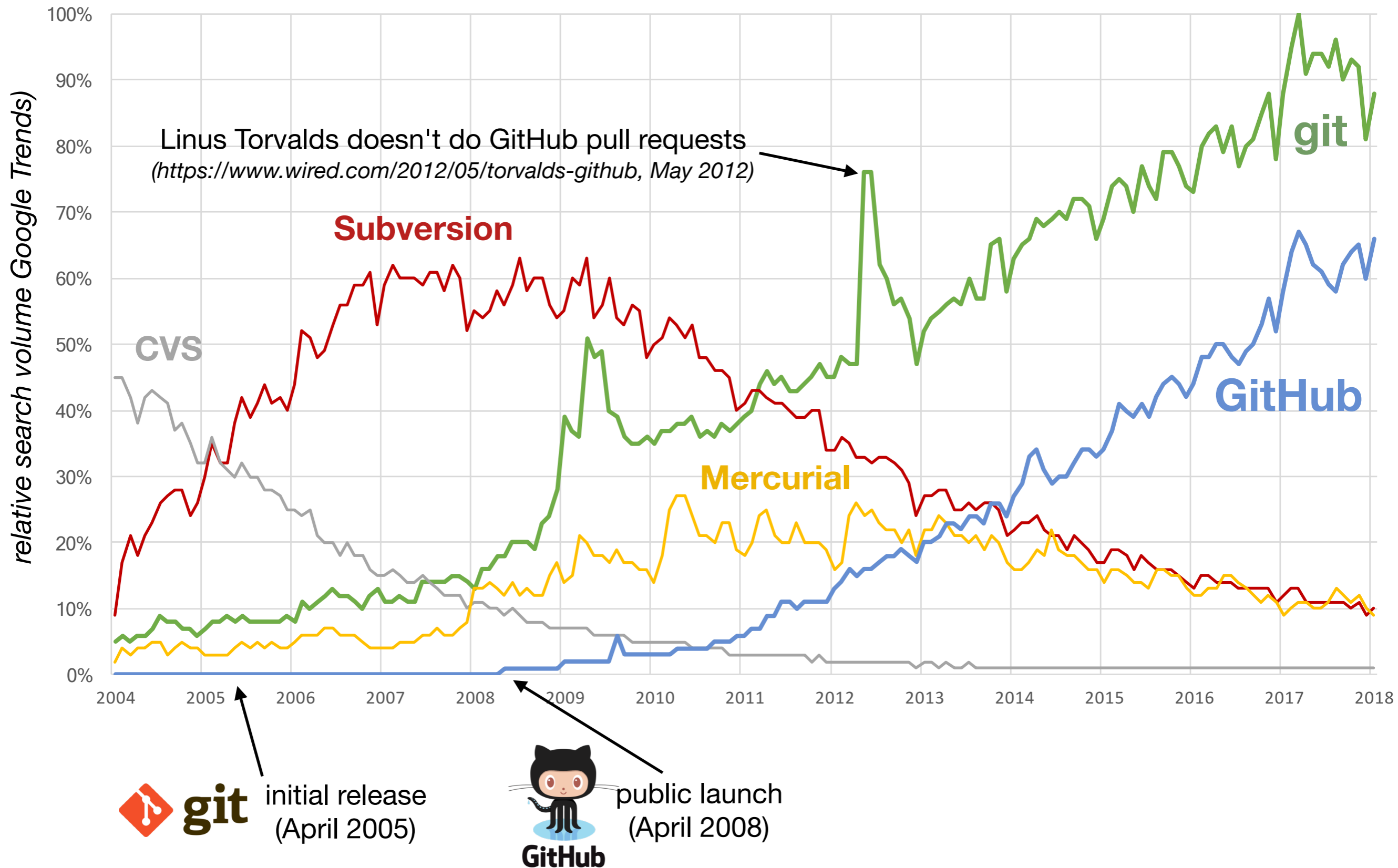
Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



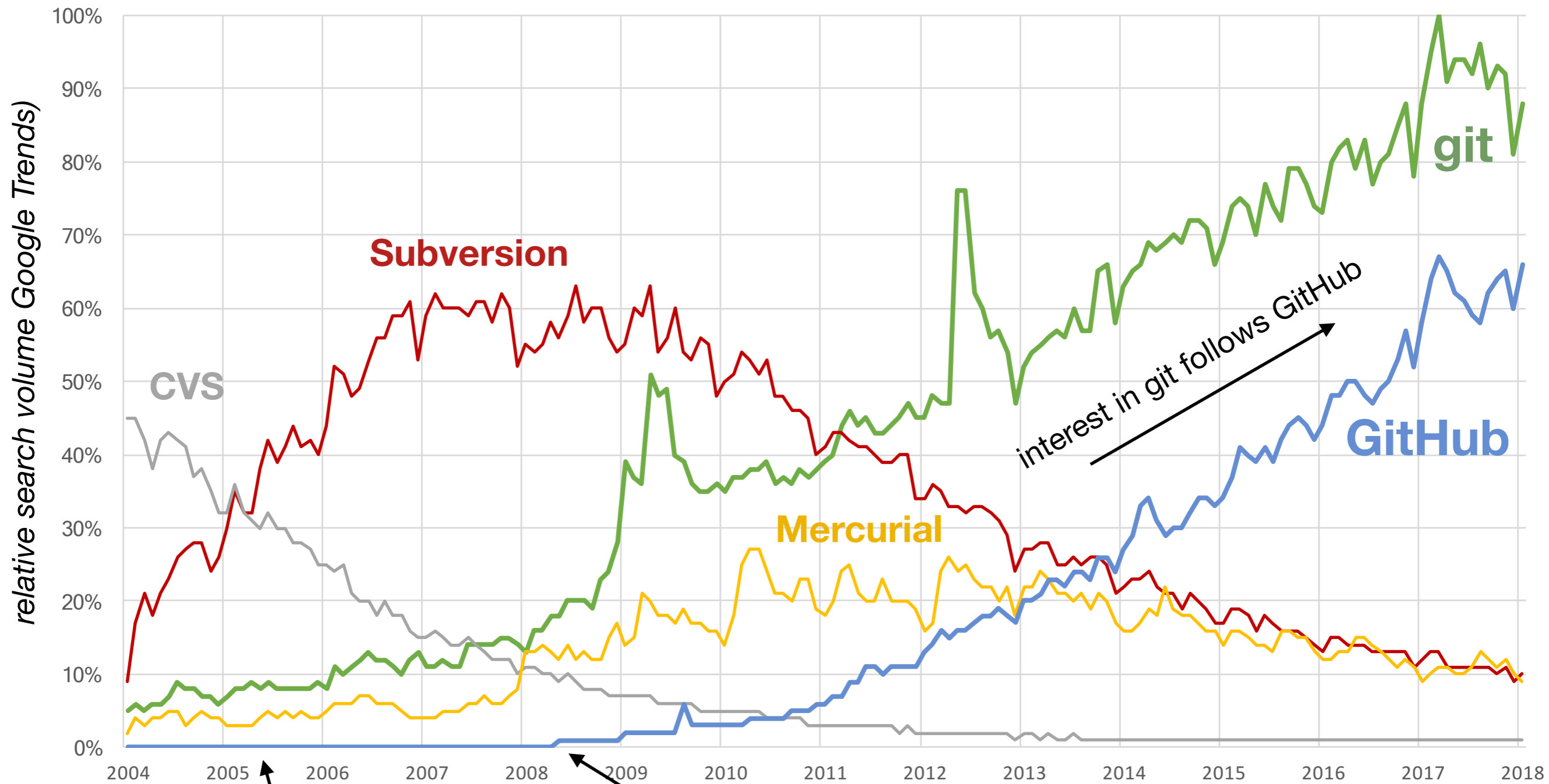
Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)




Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)

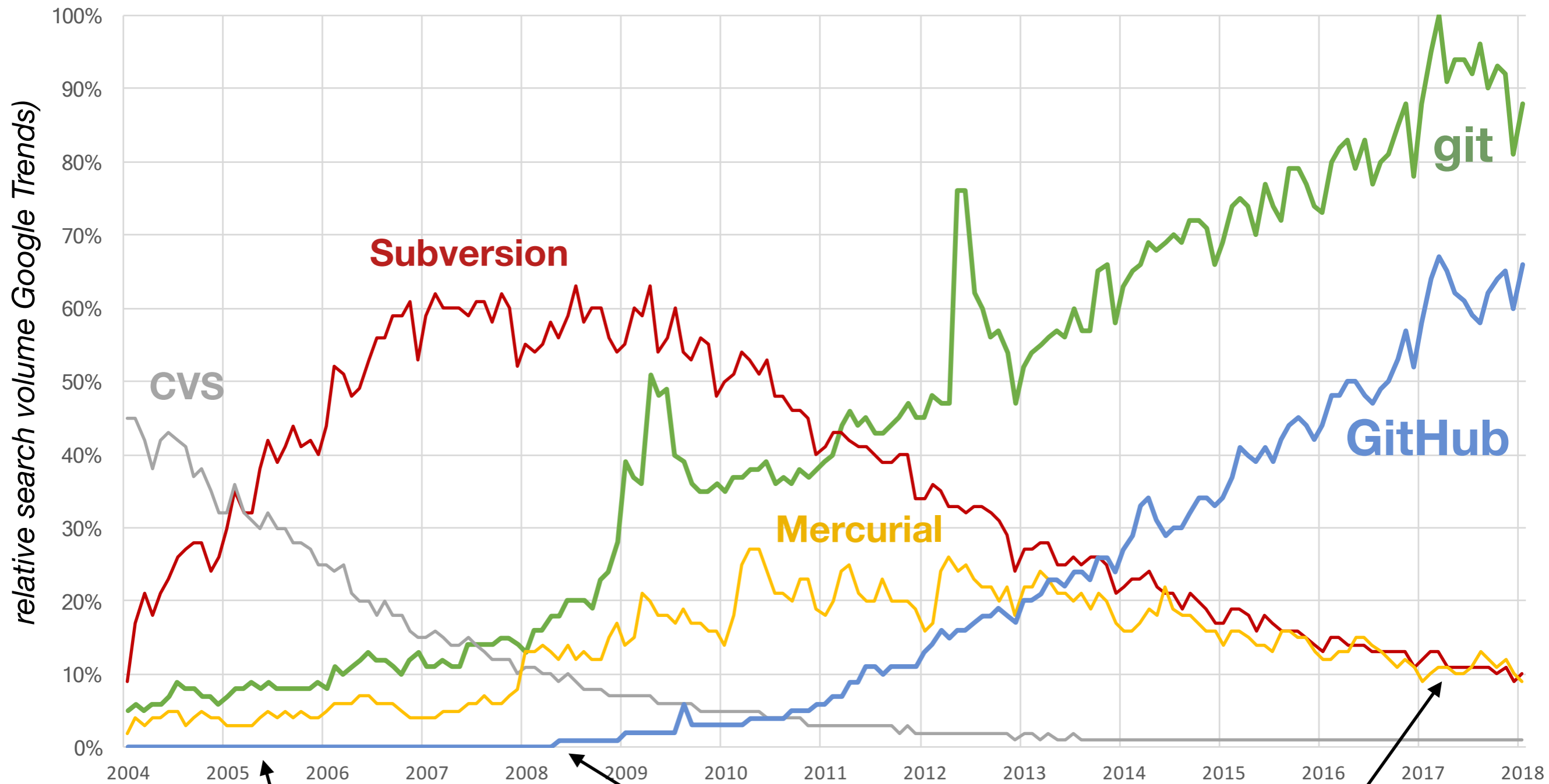


 **git** initial release
(April 2005)

 public launch
(April 2008)

Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



git

initial release
(April 2005)



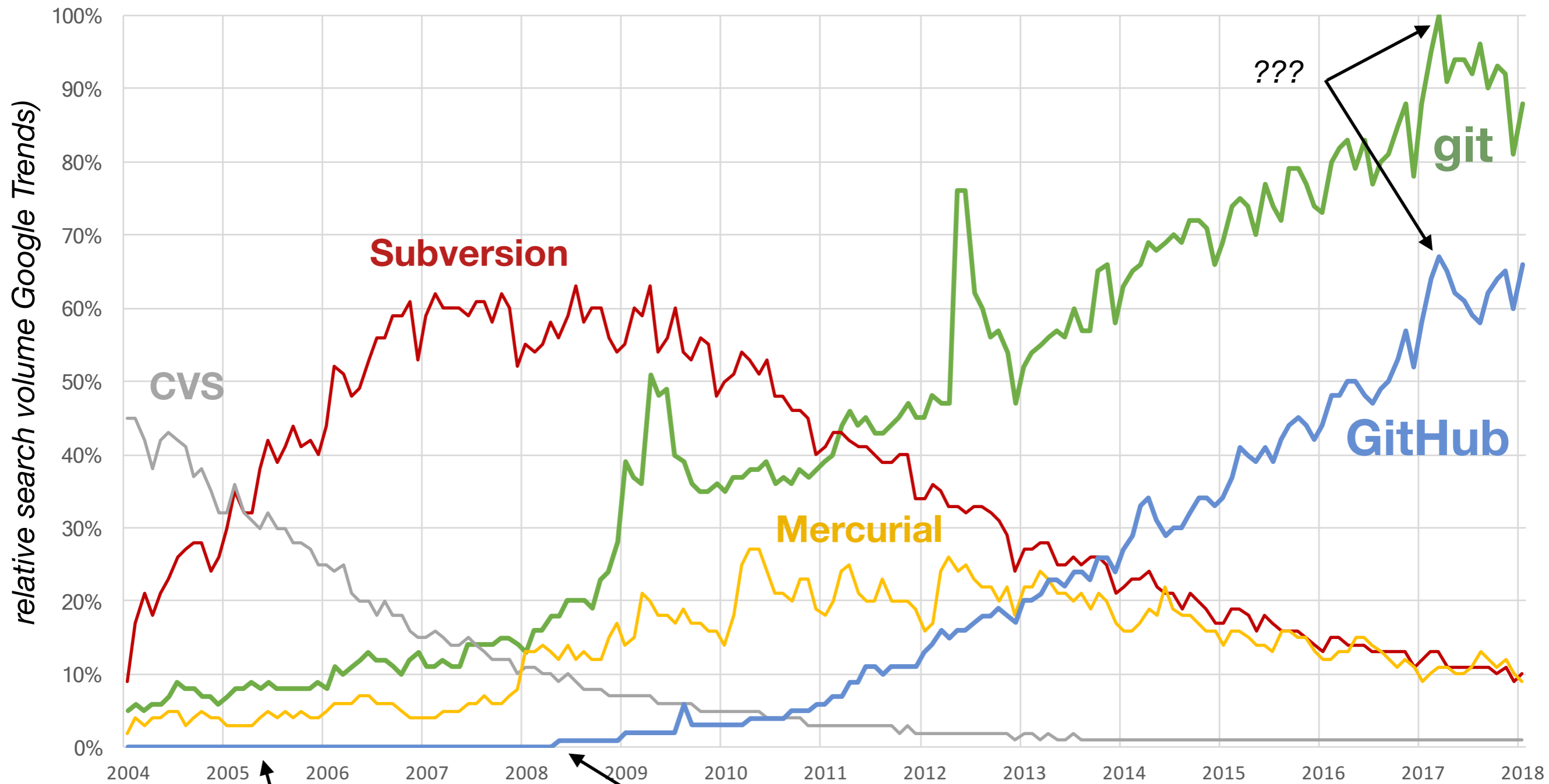
GitHub

public launch
(April 2008)


interest in Subversion/Mercurial
is 1/10th that of git...

Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)

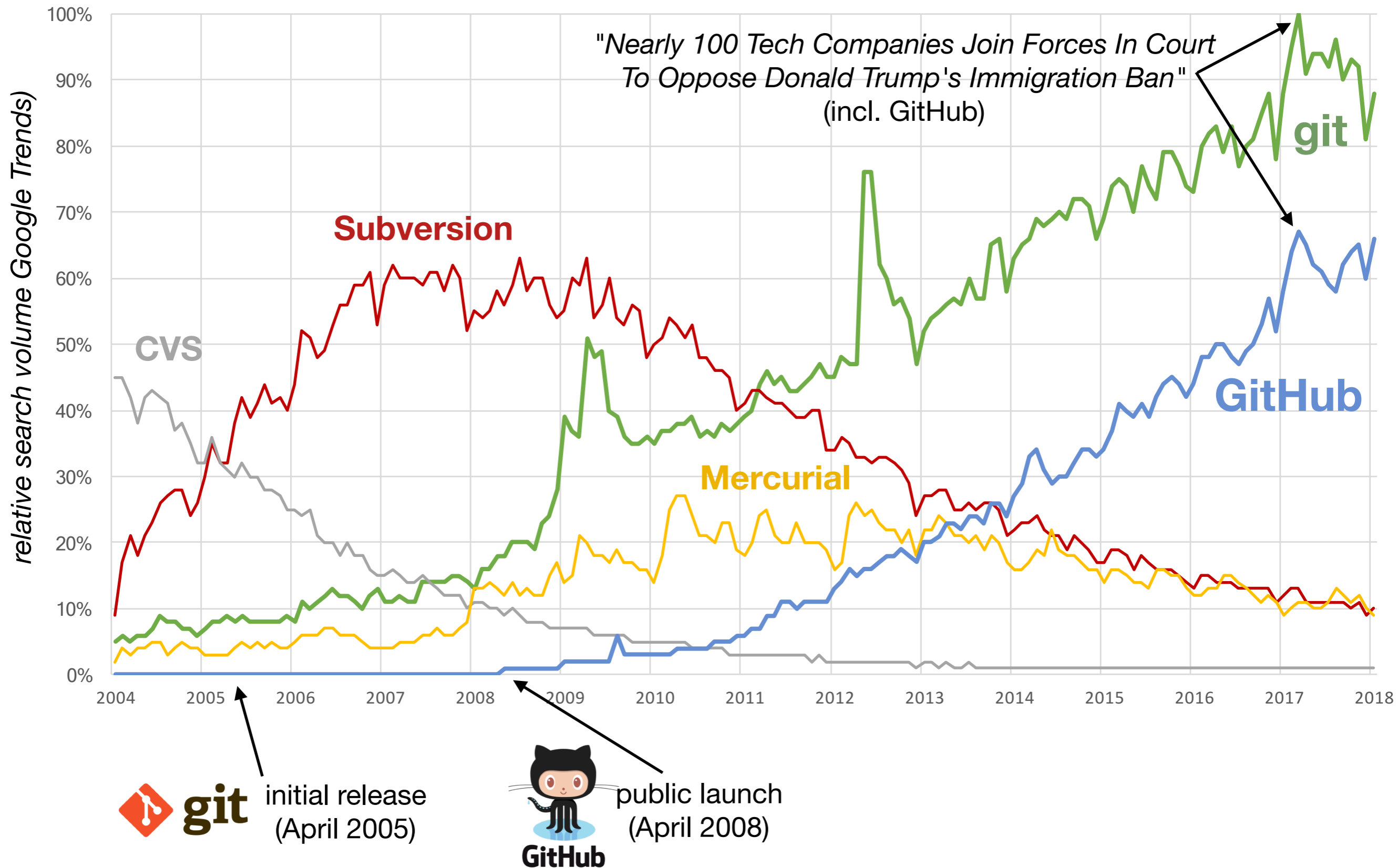


 **git** initial release
(April 2005)

 public launch
(April 2008)
GitHub




Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



Is **git** so popular thanks to **GitHub**?



- Hard to be sure, but it's likely  has had a major impact.
GitHub
- Maybe there even was some sort of feedback loop...
(**GitHub** increases interest in  **git**, which in turns makes **GitHub** more popular?)
- ... or maybe it's just because **GitHub** has 'git' in its name, and I can't get Google Trends to behave...
- Whatever the reason,  **git** is the most popular (D)VCS today.
- You may need it in order to contribute to a FOSS project.
- Which means you'll need to learn how to use it...

git has a steep learning curve



Mike Coutermarsh

@mscccc

Just googled "undo last commit but keep changes"

And I work for GitHub.

This stuff is hard!

6:10 PM - 18 Jan 2018

244 Retweets 855 Likes



 34

 244

 855



git has a steep learning curve

(from <https://www.quora.com/Why-is-Git-so-hard-to-learn>)



Levente Huszko, a pragmatic guy who's keen on theory too

Answered Dec 7, 2016

Git, in my opinion is hard to learn because it takes on a hard subject, distributed version control. Version control is difficult on its own.



Jesse McGrew, 15+ years as a professional and open source developer

Answered Jan 8, 2017

Git is hard to learn because its developers gave up on making it easy.



Ori Shalev, a software engineer

Answered Jan 2, 2017

Because probably like most people, you skipped the basics.

git has a steep learning curve

(from <https://www.quora.com/Why-is-Git-so-hard-to-learn>)



Terry Lambert, former Sr. Software Engineer at Google (2011-2012)

Answered Nov 10, 2016

The primary problem with git is that there are too many ways to do the same few, conceptually simple things.

***It's the perl of source code control systems.
This is not a compliment.***



William Estoque, Ruby, Ruby on Rails, Javascript, Engineer, Startups

Answered Nov 4, 2016

Git is hard to learn because people try to learn everything at once.

git is not for everybody

- main target audience is *software developers*
- ... especially *teams* of people working together on software
- basic experience with command line tools is desirable
- a lot of people (think they) don't need it for their *daily work*
- some feel they don't have enough *time* to learn it
- others *try* to learn, but can't seem to get their head around it
- not everybody knows an experienced git user to *help* them
- some people don't **want** to learn it (and that's fine)

Easyconfig files

Easyconfig files are software build recipes for EasyBuild.

```
easyblock = 'PythonPackage'

name = 'scikit-learn'
version = '0.19.1'
versionsuffix = '-Python-%(pyver)s'

homepage = 'http://scikit-learn.org/stable/index.html'
description = "Scikit-learn is a Python library for machine learning algorithms."

toolchain = {'name': 'intel', 'version': '2017b'}

source_urls = [PYPI_SOURCE]
sources = [SOURCE_TAR_GZ]
checksums = ['5ca0ad32ee04abe0d4ba02c8d89d501b4e5e0304bdf4d45c2e9875a735b323a0']

dependencies = [
    ('Python', '3.6.3'),
    ('matplotlib', '2.1.0', versionsuffix),
]

options = {'modulename': 'sklearn'}

sanity_check_paths = {
    'files': [],
    'dirs': ['lib/python%(pyshortver)s/site-packages/sklearn'],
}

moduleclass = 'data'
```

- Python syntax
- key-value parameter definitions
- easy to tweak:
 - software version
 - toolchain
 - dependencies
 - configure options
 - ...