



H1



Open source data sucks - and
how we learned to love it!



Jonathan Bossenger

Who am I?

- Ex-developer turned educator
- Gamer and coffee snob
- Sponsored open source contributor at Automattic



Sponsored open source contributor at Automattic

- The Cast of HBO's 'Silicon Valley' Cast Explains What Real Startups Do (NSFW)
- <https://automattic.com/work-with-us/job/parse-ly-engineering/>

AUT  MATTIC

Sponsored open source contributor at Automattic

- Commercial
 - WordPress.com, WP VIP, WooCommerce, Newspack, Akismet, Jetpack, Gravatar
- Acquisitions
 - Day One, Longreads, Beeper, Pocket Casts, Tumblr

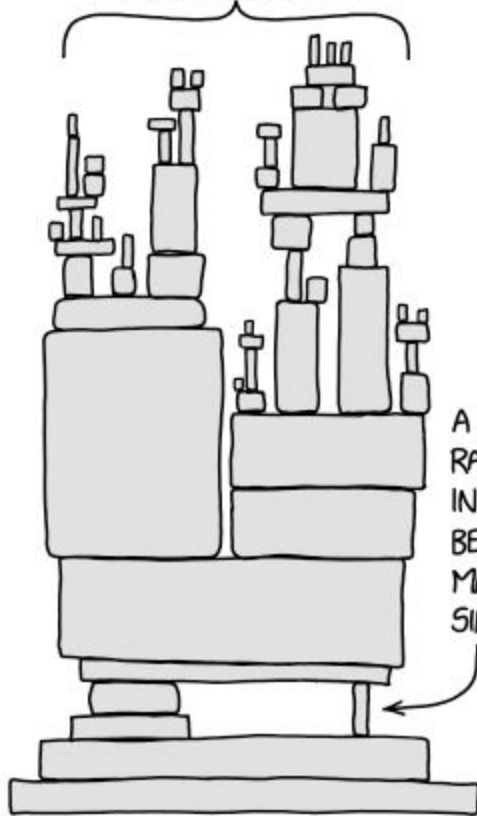
AUT  MATTIC

Sponsored open source contributor at Automattic

- Open source (aka dotOrg)
 - WordPress.org
 - Five for the Future
 - P2
 - WP for iOS
 - BuddyPress
 - bbPress

AUT  MATTIC

ALL MODERN DIGITAL
INFRASTRUCTURE



A PROJECT SOME
RANDOM PERSON
IN NEBRASKA HAS
BEEN THANKLESSLY
MAINTAINING
SINCE 2003

<https://xkcd.com/2347/>

Sponsored open source contributor at Automattic

- WordPress.org (aka the Make teams)
 - Core
 - Design
 - Mobile
 - Accessibility
 - Support
 - Training

AUT  MATTIC

Sponsored open source contributor at Automattic

- The WordPress training team
 - [Learn.WordPress.org](https://learn.wordpress.org)

AUT  MATTIC



Learn WordPress

Grow your WordPress skills with free educational resources for all levels.

Learn WordPress is also available in [English \(UK\)](#).

Search for resources



Get Started

Learning Pathways help you build your skills progressively so you can go from beginner to advanced at your own pace.

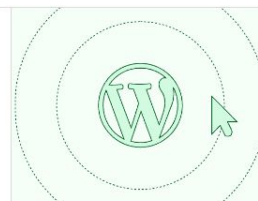
Develop with WordPress

Jump into the code and learn how to extend the software.



Start using WordPress

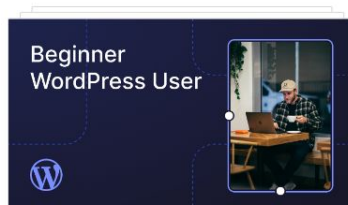
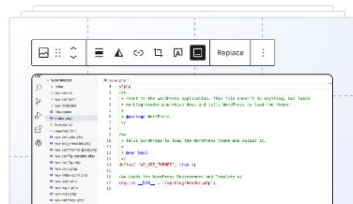
Get the basics to create and manage your own website.



Featured Courses

Focus on building your skills through a series of lessons in various formats.

[See all courses](#)

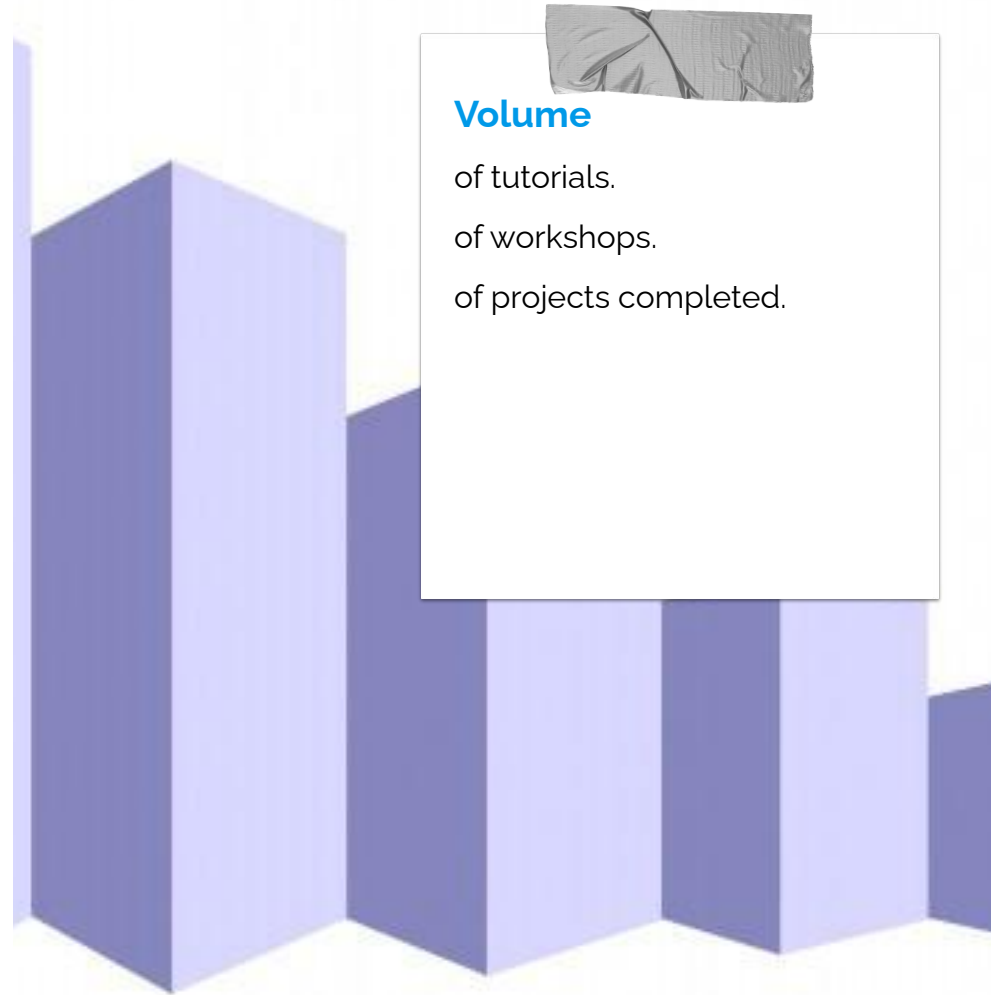




The problem

- Measuring the impact our sponsored time has on a community project, and the benefit that brings back to a commercial company, is difficult.

Our **2023** goals were all
measured by **volume**
which didn't really report
on the **impact** of our
work.





How it started

For the first time, in 2023 teams in dotOrg division were required to set more specific and measurable goals.

→ **S.M.A.R.T goals**

Specific, Measurable, Achievable, Relevant, and Time-Bound

→ **KPIs**

Key performance indicators

→ **Data-driven**

Everything should be driven by data



The solution

- Setting SMART goals, with clear KPIs, and getting leadership buy-in, would give us confidence and security that the work we felt was generating value, was concretely recognized by the company.

—

I'm pleased to say we knocked it out of the park on the first go!



Narrator:

They did not!



2024 planning (v1)

5 goals

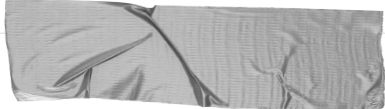
2 - 4 KPIs per goal

—

The process of iteration

2024 goals (v2)

- 2 Goals
- 6 KPIs



Goals:

Launch learning pathways

Improve user engagement on Learn.WordPress.org

KPIs: (monthly)

Completion of learning pathway lessons

25 % Course enrollment rate

15% Course completion rate

80% Learner satisfaction rate

35% Video retention rate

6 pieces of content created by contributors

2024 goals (v2)

- Completion of learning pathway lessons
- 25 % Course enrollment rate
- 15% Course completion rate
- 80% Learner satisfaction rate
- 35% Video retention rate
- 6 pieces of content created by contributors

2024 goals (v2)

- Completion of learning pathway lessons - [published content](#)
- 25 % Course enrollment rate
- 15% Course completion rate
- 80% Learner satisfaction rate - [learner surveys](#)
- 35% Video retention rate
- 6 pieces of content created by contributors - [GitHub issues](#)

2024 goals (v2)

- 25 % Course enrollment rate
- 15% Course completion rate
- 35% Video retention rate



Data dashboards

- At our division meetup in Malaga Spain, we all attended a workshop on building data dashboards using Biterga.
- We all attended with excitement as we envisioned this being the one stop solution to our problems.



Narrator:

It did not!

What did we have?

- Access to student data on our Learn.WordPress.org website?
- Access to analytics for our video content?



Narrator:

They did not!

—

**Why open source
data sucks.**

2024 goals (v2)

Course enrollment rate

- Calculated monthly as (unique enrollments across all courses) / (unique viewers of course landing page) x 100

- Dashboard
- Jetpack
- Posts
- Lesson Plans
- Tutorials
- Media
- Links
- Pages
- Meetings
- Comments
- Feedback
- Sensei LMS
- Home
- Courses
- Modules
- Lessons
- Questions
- Glossary
- Students
- Groups
- Grading 1154
- Reports
- Settings
- Tools
- Appearance
- Plugins
- Users
- Tools
- Settings
- Gutenberg

Reports

Students Courses Lessons

Guide To Using Reports

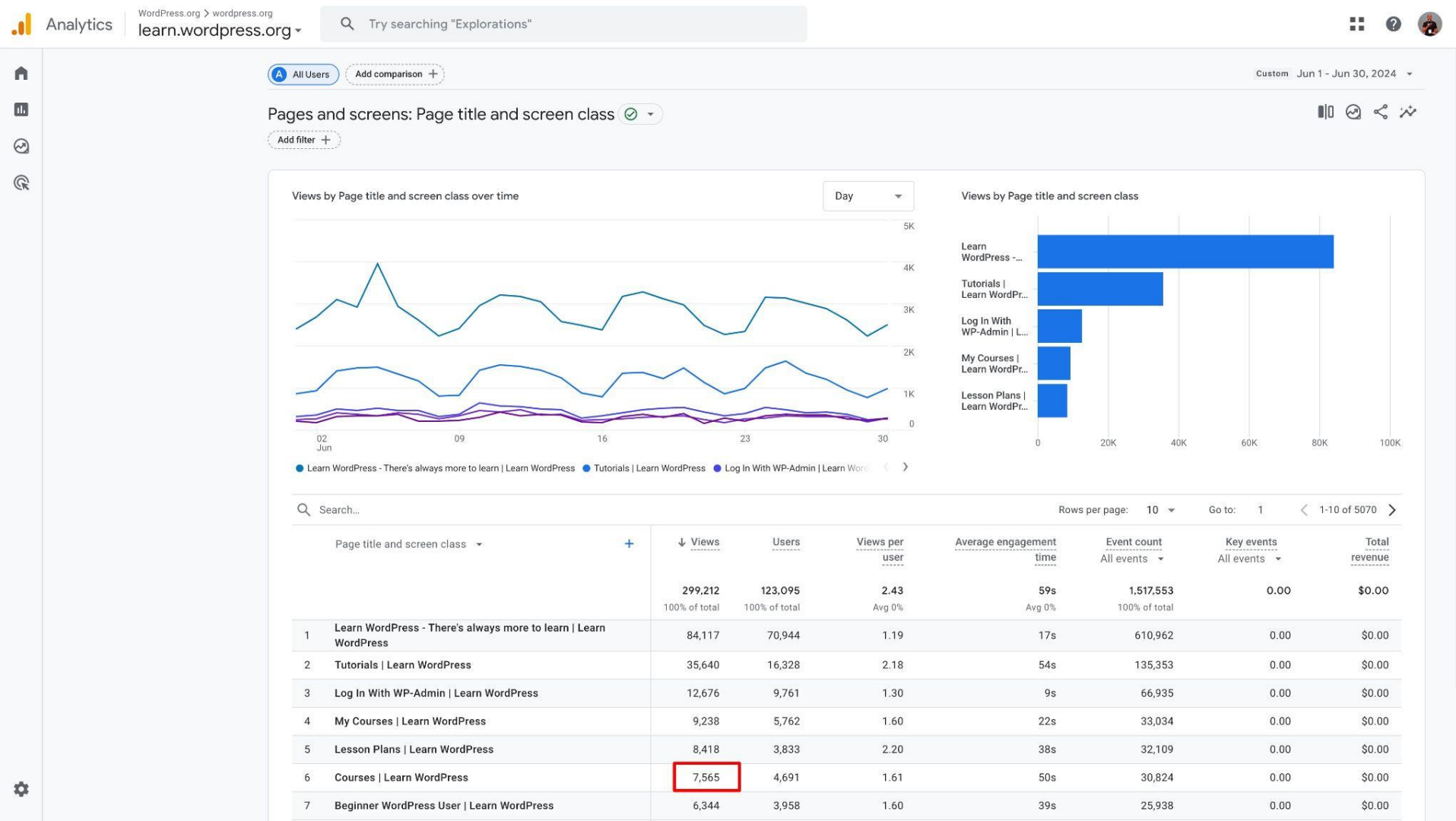
Total number of students 2024-07-01 2024-07-31 1850

Search Students

981 items 1 of 50

Select a group Last Activity: Start Date End Date Filter

Student (981)	Email	Date Registered	Last Activity	Active Courses (1305)	Completed Courses (471)	Average Grade (95%)
		13 June 2024	N/A	1	0	0%
		13 June 2024	N/A	1	0	0%
		9 June 2024	9 June 2024	1	0	100%
		5 June 2024	8 June 2024	1	1	79.17%
		16 June 2024	N/A	0	0	0%
		16 June 2024	16 June 2024	1	0	100%
		6 June 2024	6 June 2024	1	0	0%
		10 June 2024	N/A	1	0	0%
		7 June 2024	7 June 2024	1	0	0%
		12 June 2024	N/A	1	0	0%
		9 June 2024	N/A	0	0	0%
		9 June 2024	9 June 2024	1	0	0%
		3 June 2024	3 June 2024	0	1	97.57%
		14 June 2024	14 June 2024	1	0	100%
		10 June 2024	10 June 2024	1	0	100%
		7 December 2012	3 January 2023	1	2	95.97%
		9 June 2024	9 June 2024	1	0	0%
		31 May 2024	31 May 2024	1	0	83%
		5 June 2024	9 June 2024	1	2	87.82%



2024 goals (v2)

- ~~25 % Course enrollment rate~~
- 15% Course completion rate
- 35% Video retention rate

2024 goals (v2)

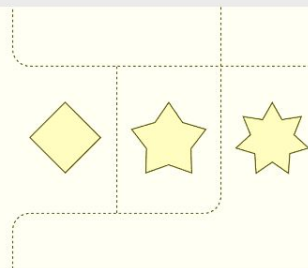
Course completion rate

- Calculated for each learning pathway as the average progress of all students who enrolled in that course that month.

Learn WordPress is also available in [English \(UK\)](#).

Develop with WordPress

Jump into the code and learn how to extend the software.

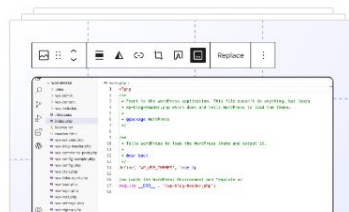


Search



Topic

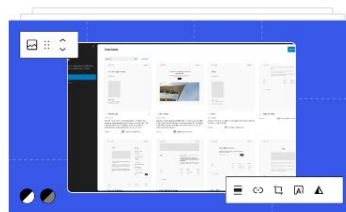
Level: All



Beginner WordPress Developer

Start your WordPress development journey here. Learn how WordPress works, and how to extend it.

9 hours 59 lessons In progress



Intermediate Theme Developer

Learn how to develop a custom block theme, from global styles to release on WordPress.org.

8 hours 32 lessons

Why open source data sucks.

(Part deux)

- Dashboard
- Jetpack
- Posts
- Lesson Plans
- Tutorials
- Media
- Links
- Pages
- Meetings
- Comments
- Feedback
- Sensei LMS
 - Home
 - Courses
 - Modules
 - Lessons
 - Questions
 - Glossary
 - Students
 - Groups
 - Grading 1154
 - Reports**
 - Settings
 - Tools
- Appearance
- Plugins
- Users
- Tools
- Settings
- Gutenberg

[Reports](#) > [Beginner WordPress Developer](#)

Lessons in this Course | **Students taking this Course**

Search Students

Date Started:

2024-07-01

2024-07-31

Filter

Student	Email	Date Started	Date Completed	Status	Percent Complete
		2024-07-01 07:44:51		In Progress	2%
		2024-07-01 09:52:11		In Progress	0%
		2024-07-01 10:09:23		In Progress	0%
		2024-07-01 11:31:55		In Progress	2%
		2024-07-01 12:11:40		In Progress	0%
		2024-07-01 12:48:15		In Progress	0%
		2024-07-01 13:16:16		In Progress	0%
		2024-07-01 13:37:57		In Progress	0%
		2024-07-01 13:54:30		In Progress	0%
		2024-07-01 11:12:57		In Progress	2%
		2024-07-01 15:02:50		In Progress	2%
		2024-07-01 15:36:51		In Progress	0%
		2024-07-01 20:00:58		In Progress	0%
		2024-07-01 20:56:19		In Progress	0%
		2024-07-01 21:13:16		In Progress	2%
		2024-07-02 02:47:24		In Progress	0%
		2024-07-01 11:44:02		In Progress	29%
		2024-07-02 03:59:05		In Progress	0%
		2024-07-02 03:41:36		In Progress	31%
		2024-07-02 07:33:58		In Progress	0%
Student	Email	Date Started	Date Completed	Status	Percent Complete

Export all rows (CSV)

H2	=AVERAGE(F:F)														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Student	Email	Date Started	Date Completed	Status	Percent Complete									
2			2024-06-05 14:18:52		In Progress	0		9.705882353							
3			2024-06-06 17:11:09		In Progress	0									
4			2024-06-07 14:13:19		In Progress	0									
5			2024-06-09 0:52:17		In Progress	0									
6			2024-06-09 1:28:13		In Progress	19									
7			2024-06-10 16:19:02		In Progress	56									
8			2024-06-11 17:21:18		In Progress	0									
9			2024-06-12 12:04:38		In Progress	0									
10			2024-06-14 9:20:22		In Progress	0									
11			2024-06-14 10:59:50		In Progress	0									
12			2024-06-15 0:21:07		In Progress	0									
13			2024-06-15 8:53:27		In Progress	0									
14			2024-06-13 23:40:16		In Progress	26									
15			2024-06-16 13:42:42		In Progress	26									
16			2024-06-18 7:11:56		In Progress	2									
17			2024-06-22 6:12:27		In Progress	0									
18			2024-06-22 13:52:08		In Progress	0									
19			2024-06-22 14:43:40		In Progress	0									
20			2024-06-22 14:49:51		In Progress	0									
21			2024-06-22 15:59:56		In Progress	2									
22			2024-06-22 16:29:19		In Progress	3									
23			2024-06-22 20:19:26		In Progress	3									
24			2024-06-23 10:07:27		In Progress	27									
25			2024-06-23 14:54:16		In Progress	0									
26			2024-06-23 20:48:23		In Progress	0									
27			2024-06-23 21:00:53		In Progress	0									
28			2024-06-23 21:13:21		In Progress	0									
29			2024-06-23 22:36:32		In Progress	7									
30			2024-06-24 1:30:43		In Progress	0									
31			2024-06-24 2:37:05		In Progress	0									
32			2024-06-24 3:45:52		In Progress	0									
33			2024-06-24 6:27:22		In Progress	63									
34			2024-06-24 16:11:02		In Progress	0									
35			2024-06-22 17:56:05		In Progress	12									
36			2024-06-24 22:26:17		In Progress	0									
37			2024-06-22 4:17:31		In Progress	10									
38			2024-06-25 5:27:00		In Progress	0									
39			2024-06-25 7:12:37		In Progress	0									
40			2024-06-24 6:38:31		In Progress	10									
41			2024-06-24 9:25:12		In Progress	27									

Course completion rate script

- Download all csv files to a folder
- Run `php course-completion.php`
- ...
- Profit?

2024 goals (v2)

- ~~25 % Course enrollment rate~~
- ~~15% Course completion rate~~
- 35% Video retention rate

Why open source data sucks.

(The threequel)

WordPress.tv

- Surely the video platform for one of the largest open source software projects on the web has built in analytics?



Narrator:

It did not!

And don't call me Shirley!

Enter YouTube

2024 goals (v2)

Video retention

- Calculated monthly as (the average percentage viewed) across all Tutorials and Lessons combined.

Filter

Jan 1 – Aug 27, 2024

Content

Traffic source

Geography

Cities

Viewer age

Viewer gender

Date

Subscription status

Subscription source

Content type

Playlist

Podcast

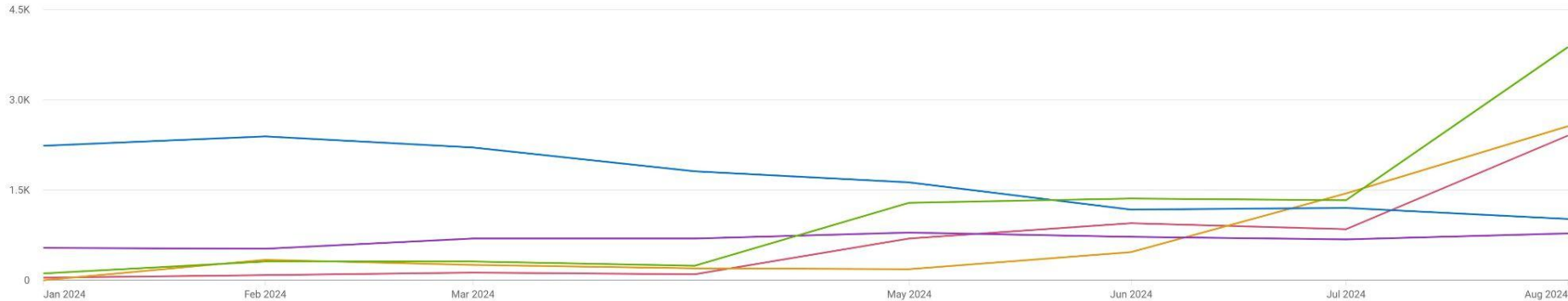
More

Views by: Content

Select secondary metric

Line chart

Monthly



Content	Views	Watch time (hours)	Subscribers	Average view duration	Impressions	Impressions click-through rate
Total	10,000	10,000	1,000	1:00	10,000	1.0%
Using the WordPress REST API	1,000	1,000	100	1:00	1,000	1.0%
Introduction to WordPress	1,000	1,000	100	1:00	1,000	1.0%
WordPress and web servers	1,000	1,000	100	1:00	1,000	1.0%
How to add a blog or news page to any WordPress we...	1,000	1,000	100	1:00	1,000	1.0%
Choosing and installing a theme	1,000	1,000	100	1:00	1,000	1.0%
WordPress REST API – custom routes and endpoints	1,000	1,000	100	1:00	1,000	1.0%

Content

Views by:

4.5K

3.0K

1.5K

0

Jan 2024

Feb 2024

Mar 2024

May 2024

Jun 2024

Jul 2024

Aug 2024

gender

Date

Subscription status

Subscription source

Content type

Playlist

Podcast

More

Line chart

Monthly

Recent

97

Learn Tutorials

Video group

Aug 21, 2024

Created

153

Learn Lessons

Video group

Apr 5, 2024

Created

273

Tutorials + Lessons

Video group

Apr 5, 2024

Created

Create new group

Content	Views	Watch time (hours)	Subscribers	Average view duration	Impressions	Impressions click-through rate
Total						
13:24 Using the WordPress REST API						
4:25 Introduction to WordPress						
6:36 WordPress and web servers						
5:19 How to add a blog or news page to any WordPress we...						
10:22 Choosing and installing a theme						
17:20 WordPress REST API – custom routes and endpoints						

YouTube Analytics

- Surely YouTube has an easy way to fetch all video IDs for a playlist?



Narrator:

It does not!

Get all video IDs

```
var els = document.getElementsByClassName('yt-simple-endpoint style-scope  
ytd-playlist-video-renderer');  
var show="";  
for(i = 0;i<els.length;i++){  
    var el = els[i];  
    show += (el.href.split('?v=')[1].split('&list')[0] + "\n");  
}  
console.log(show);
```

Get all video IDs

```
yt-dlp --flat-playlist --get-id "playlist-url"
```

2024 goals (v2)

- ~~25% Course enrollment rate~~
- ~~15% Course completion rate~~
- ~~35% Video retention rate~~

Why open source data sucks.

(and how we learned to love it)

What did we learn?

- Data is messy, open source data even more so.
- What you can control, vs what you only have influence on.
- Never underestimate the power of an spreadsheet for v1.
- Validate your data.
- What is your acceptable margin of error?
- Linus's law - given enough eyeballs, all (data reporting) bugs are shallow.
- Report, automate, improve.

Thank you.

- jonathanbossenger.com
- fosstodon.org/@jonathanbossenger
- linkedin.com/in/jonathanbossenger
- twitter.com/jon_bossenger



Questions/feedback?

