Security-Related Questions about WebAssembly

Bachelor Thesis by Pascal André

Final Presentation – 21 December, 2021

supervised by Mohammad Ghafari
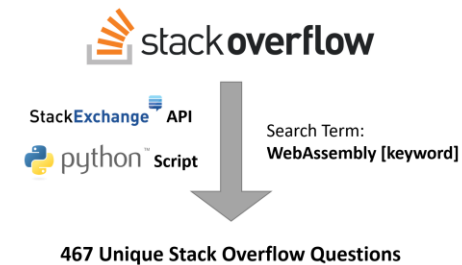
Software Composition Group

University of Bern

Summary 1ˢᵗ Presentation

Motivation and Goals

- What types of security-related questions do developers ask?

- What are important security-related topics that developers discuss?

- Are there groups of developers that ask/answer the most questions?

Collect Dataset – StackExchange API

StackExchange API

python Script

Search Term:
WebAssembly [keyword]
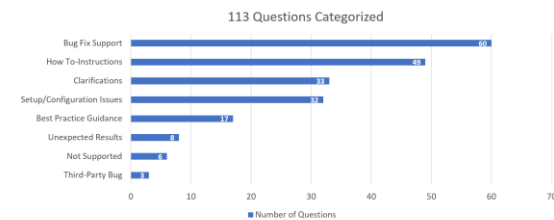
467 Unique Stack Overflow Questions

Methodology

- Evaluating questions based on 5 criteria:
  - Relevancy to our work
  - Identified Categories
  - Reasons for no answers
  - Answered by question owner
  - Correctness of answering status

Results

Which types of questions do developers ask most often?

113 Questions Categorized

	Number of Questions
Bug Fix Support	60
How To-Instructions	49
Clarifications	33
Setup/Configuration Issues	32
Best Practice Guidance	17
Unexpected Results	8
Not Supported	6
Third-Party Bug	3

Cleaning Dataset

Original Dataset:

467 Questions

Relevant Questions:

359 Questions (76.87%)

security-related WebAssembly questions
scored level 2-3 on our scale

Irrelevant Questions:

108 Questions (23.13%)

False-Positives
Scored level 0-1 on our scale
popular keywords: log-in, private, sign

Our Dataset
Answer State

	is_answered == TRUE	is_answered == FALSE	
answers > 0	219 61.00%	34 9.47%	**253** 70.47%
answers = 0	10 2.79%	96 26.74%	**106** 29.53%
	229 63.79%	**130** 36.21%	**359**

24 questions (6.68%) had an incorrect answer state

Evaluating Dataset
Question Tags

Most Popular Tags

Tag	Count	Percentage
blazor	198	55.15%
blazor-webassembly	135	37.60%
webassembly	104	28.97%
asp.net-core	80	22.28%
c#	80	22.28%
blazor-client-side	65	18.11%
authentication	44	12.26%
identityserver4	41	11.42%
javascript	31	8.64%
asp.net	19	5.29%
asp.net-identity	18	5.01%

Most Popular (Programming) Languages

Tag	Count	Percentage
c#	80	22.28%
javascript	31	8.64%
c++	13	3.62%
rust	6	1.67%
typescript	4	1.11%
c	3	0.84%
html	2	0.56%
assembly	1	0.28%
go	1	0.28%

Overall 234 tags. Average of 3.79 tags per question.

11 most popular tags (4.7%) are responsible for 60.5% of all assignments where a tag was added to a question.

Evaluating Dataset
Timeline

Number of questions asked in each quarter since 2016
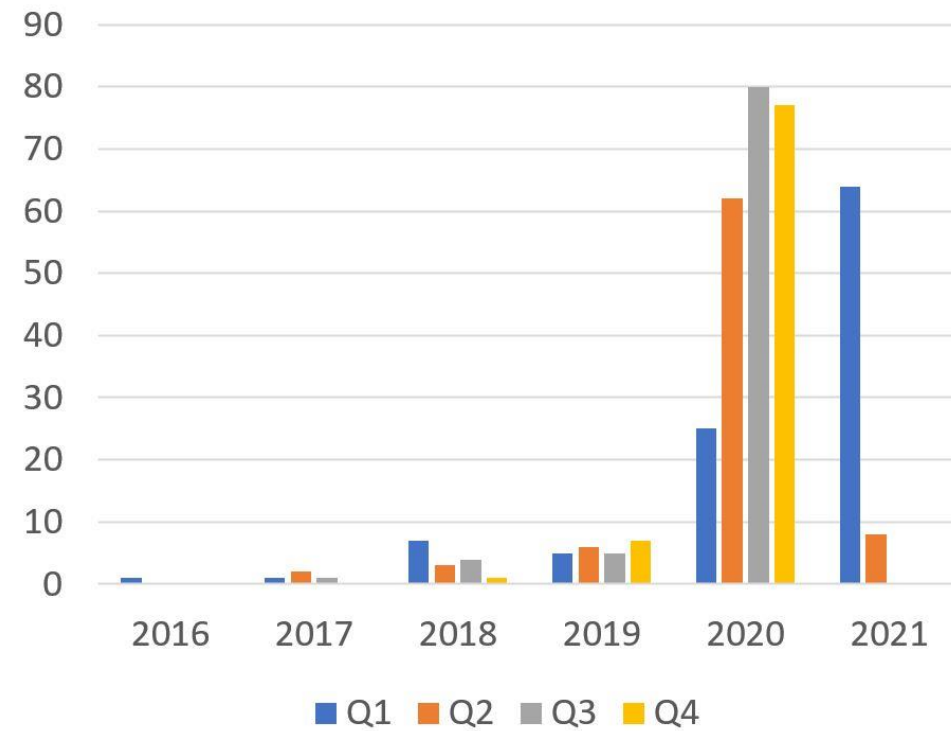
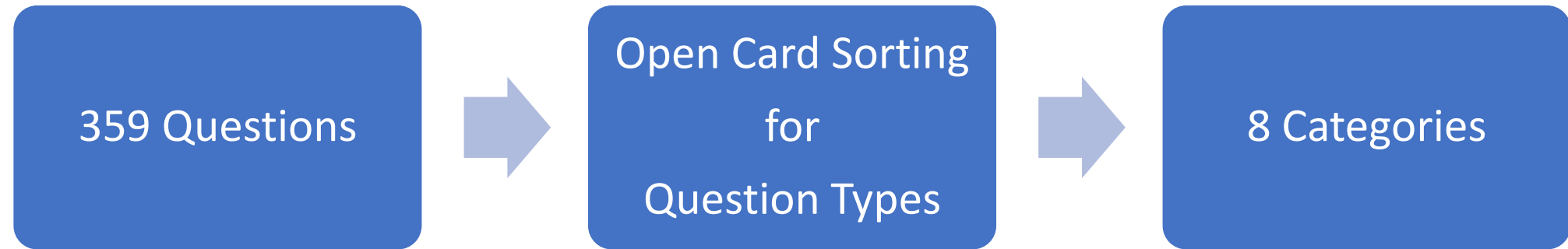|	2016	2017	2018	2019	2020	2021
Q1	1	1	7	5	25	64
Q2	0	2	3	6	62	8
Q3	0	1	4	5	80	
Q4	0	0	1	7	77	

Note: Dataset was collected in April 2021

Research Question #1:

What are the intentions why developers are asking security-related WebAssembly questions?

RQ #1: Developer Intentions
Methodology

359 Questions → Open Card Sorting for Question Types → 8 Categories

RQ #1: Developer Intentions
Number of questions per category

Identified question types (categories) using open card sorting:

Category	Number of Questions	Percentage
Bug Fix Support	181	50.42%
How To-Instructions	168	46.80%
Clarifications	99	27.58%
Setup/Configuration Issues	79	22.01%
Best Practice Guidance	43	11.98%
Unexpected Results	8	2.23%
Not Supported	7	1.95%
Third-Party Bug	4	1.11%

■ Number of Questions
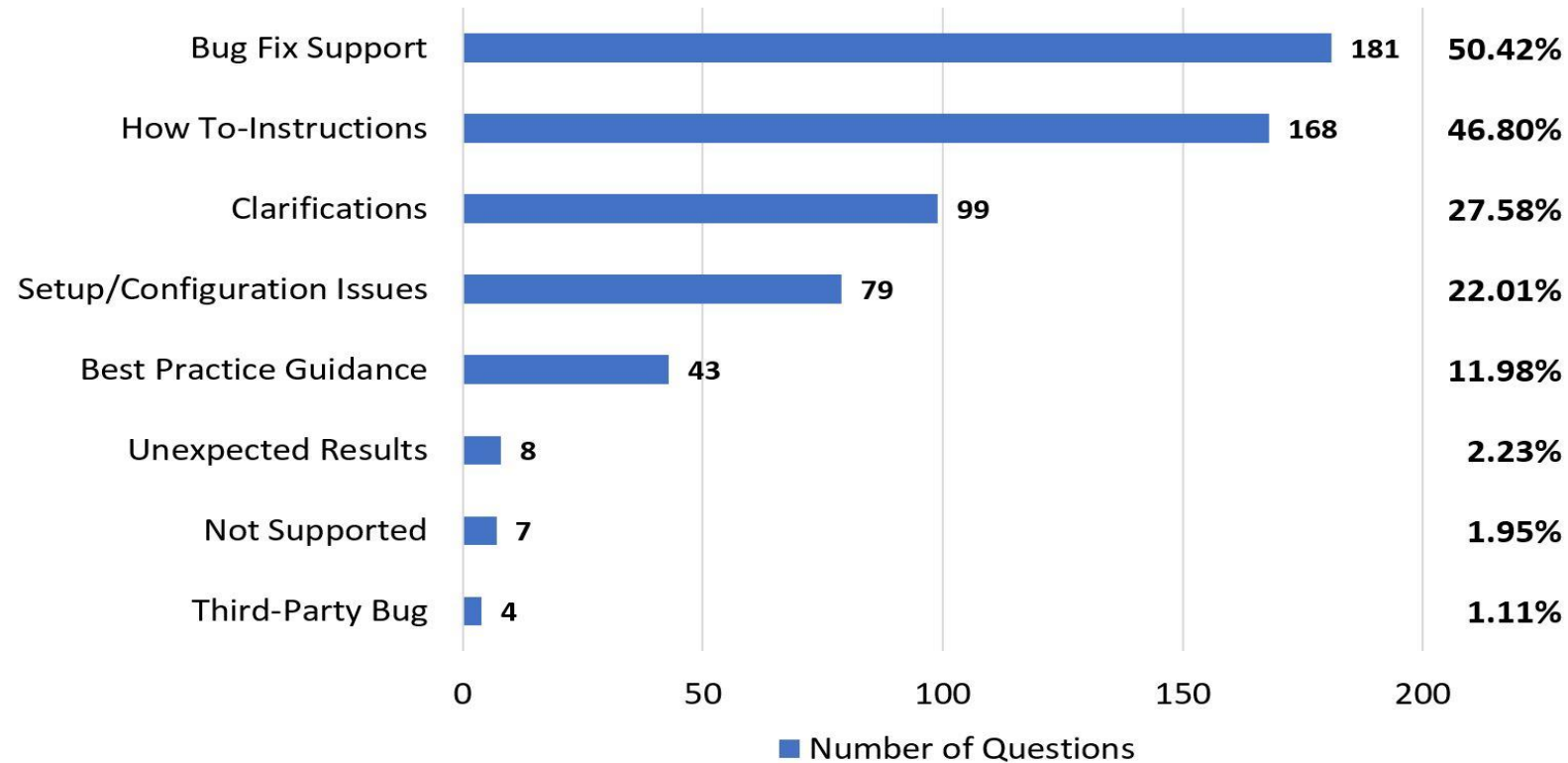
RQ #1: Developer Intentions
Questions answered by its owner

Category	Question Count	Percentage
All Questions	73	20.33%
Setup/Configuration	34	43.04%
Bug Fix Support	55	30.39%
Best Practice Guidance	7	16.28%
How To Instructions	26	15.47%
Clarification	11	11.11%

RQ #1: Developer Intentions
Number of unanswered questions

Category	Question Count	Percentage
All Questions	130	36.21%
Best Practice Guidance	17	39.53%
How To Instructions	63	37.50%
Bug Fix Support	64	35.36%
Clarification	31	31.31%
Setup/Configuration	16	20.25%

Research Question #2:

What are the security-related WebAssembly topics that developers are asking?

RQ #2: Question Topics
Number of questions per topic

Methodology

359 Questions	→

Open Card Sorting for Question Topics	→

19 Topics

281	Blazor
188	Authentication
43	Azure
37	Coding Question
36	User Token
34	HTTP Request
34	Interaction with WASM <-> PL
32	Storage
30	API
29	How WASM works
29	Deployment
21	Navigation/Redirect
20	Emscripten
14	Hosting/Server
14	UI
12	Browser
12	CORS
11	Performance
8	Database

RQ #2: Question Topics
Popularity of pairs of topics

Topics	Blazor	Authentication	Azure	Coding Question	User Token	HTTP Request	Interaction with WASM <-> PL	Storage	API	How WASM Works	Deployment	Navigation/Redirect	Emscripten	Hosting/Server	UI	Browser	CORS	Performance	Database
Blazor	-	188	43	20	36	34	8	15	27	4	27	21		14	14	6	8	7	6
Authentication	188	-	26	7	32	16	1	5	16	1	16	17		5	11	1		5	1
Azure	43	26	-	3	6	4		2	7		7	3					3		1
Coding Question	20	7	3	-			5	4	1		1		3			1	1	1	
User Token	36	32	6		-	3		3	6		1	1					2		
HTTP Request	34	16	4		3	-	1		6		4	1		1		2	2	1	
Interaction with WASM <-> PL	8	1		5		1	-	7	1	11			7		1	1	1	1	1
Storage	15	5	2	4	3		7	-		6			4		1	3	1		2
API	27	16	7	1	6	6	1		-		3		1	1			1		
How WASM works	4	1					11	6		-			6			1			
Deployment	27	16	7	1	1	4			3		-	3		5	1	1		1	
Navigation/Redirect	21	17	3		1	1					3	-			1			2	
Emscripten				3			7	4	1	6			-			2	1		
Hosting/Server	14	5				1			1		5			-			1		
UI	14	11					1	1			1	1			-				1
Browser	6	1		1		2	1	3		1	1		2			-	1		
CORS	8		3	1	2	2	1	1	1					1	1	1	-		1
Performance	7	5		1		1	1				1	2						-	
Database	6	1	1				1	2							1		1		-

RQ #2: Question Topics
Most popular specific questions asked

Looking at the most popular pairs of topics (Authentication & Blazor etc.),

in descending order, the following specific problems were identified most often:

P1: Implement 3rd Party Authentication

P2: Authentication works locally in development mode but not after deployment

P3: User navigation not working after login

P4: How to customize the pre-existing UI elements of the login screen

P5: How to assign user roles

P6: How to secure API endpoints to only give access to group of selected/authenticated users

Research Question #3:

Can we identify groups of developers that are asking or answering particularly many questions?

Experience based on Stack Exchange Reputation Leagues

All developers within dataset (wrote at least one question/answer/comment)

631 developers wrote:

- 359 questions
- 357 answers
- 1097 comments

Min. Total Reputation	Level
0	Level 0
1	Level 1
200	Level 2
500	Level 3
1'000	Level 4
2'000	Level 5
3'000	Level 6
5'000	Level 7
10'000	Level 8
25'000	Level 9
50'000	Level 10
100'000	Level 11
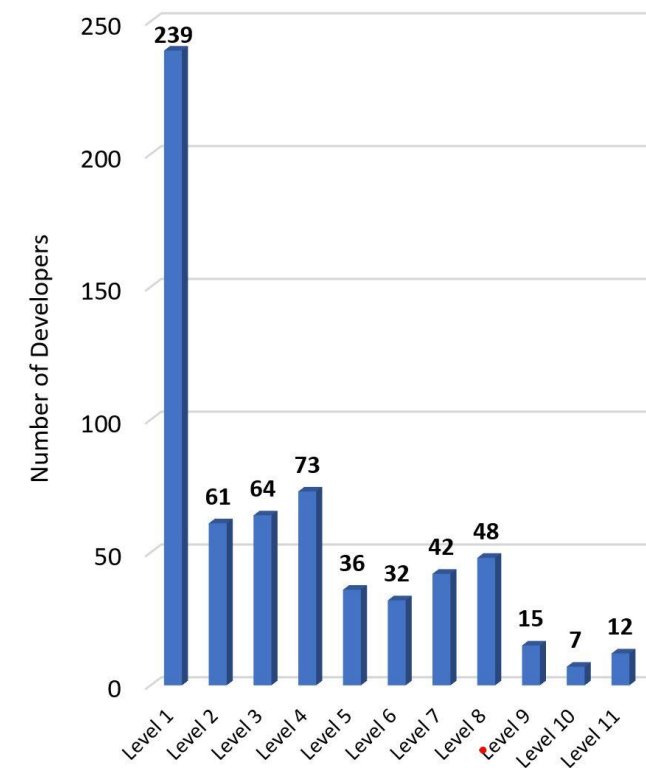
RQ #3: Developers
Developers asking questions

312 developers asked 359 questions
→ Only 8 developers asked 3+ many questions

RQ #3: Developers
Developers writing accepted answers
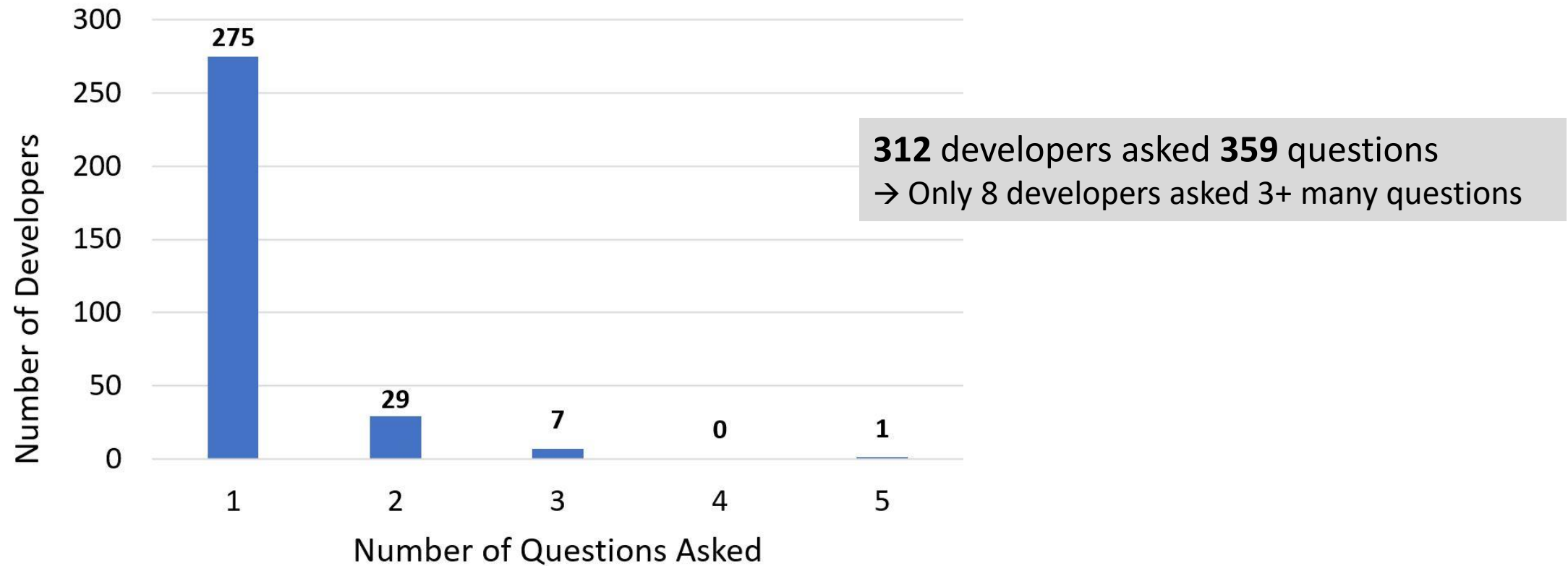
109 developers wrote **157** accepted answers
→ 6.5% of all developers provided 28% of all accepted answers

Summary

Cleaning Dataset

Original Dataset:
467 Questions

Relevant Questions:
359 Questions (76.87%)

Irrelevant Questions:
108 Questions (23.13%)

security-related WebAssembly questions
scored level 2-3 on our scale

False-Positives
Scored level 0-1 on our scale
popular keywords: log-in, private, sign

3

RQ #1: Developer Intentions
Number of questions per category

Identified question types (categories) using open card sorting:

Category	Number of Questions	Percentage
Bug Fix Support	181	50.42%
How To-Instructions	168	46.80%
Clarifications	99	27.58%
Setup/Configuration Issues	79	22.01%
Best Practice Guidance	43	11.98%
Unexpected Results	8	2.23%
Not Supported	7	1.95%
Third-Party Bug	4	1.11%

■ Number of Questions

9

RQ #2: Question Topics
Number of questions per topic

Methodology

359 Questions → Open Card Sorting for Question Topics → 19 Topics

281	Blazor
188	Authentication
43	Azure
37	Coding Question
36	User Token
34	HTTP Request
34	Interaction with WASM <-> PL
32	Storage
30	API
29	How WASM works
29	Deployment
21	Navigation/Redirect
20	Emscripten
14	Hosting/Server
14	UI
12	Browser
12	CORS
11	Performance
8	Database

13

RQ #3: Developers
Developers writing accepted answers

109 developers wrote **157** accepted answers
→ 6.5% of all developers provided 28% of all accepted answers

Number of Developers (y-axis): 0–100
Number of Accepted Answers (x-axis): 1, 2, 3, 4, 6, 8, 9, 14

Values: 94, 8, 2, 1, 1, 1, 1, 1

19