



HARVARD
John A. Paulson
School of Engineering
and Applied Sciences

Berkeley | **EECS**
ELECTRICAL ENGINEERING & COMPUTER SCIENCES



Samueli
Computer Science

Enabling Data-Driven API Design with Community Usage Data: A Need-Finding Study



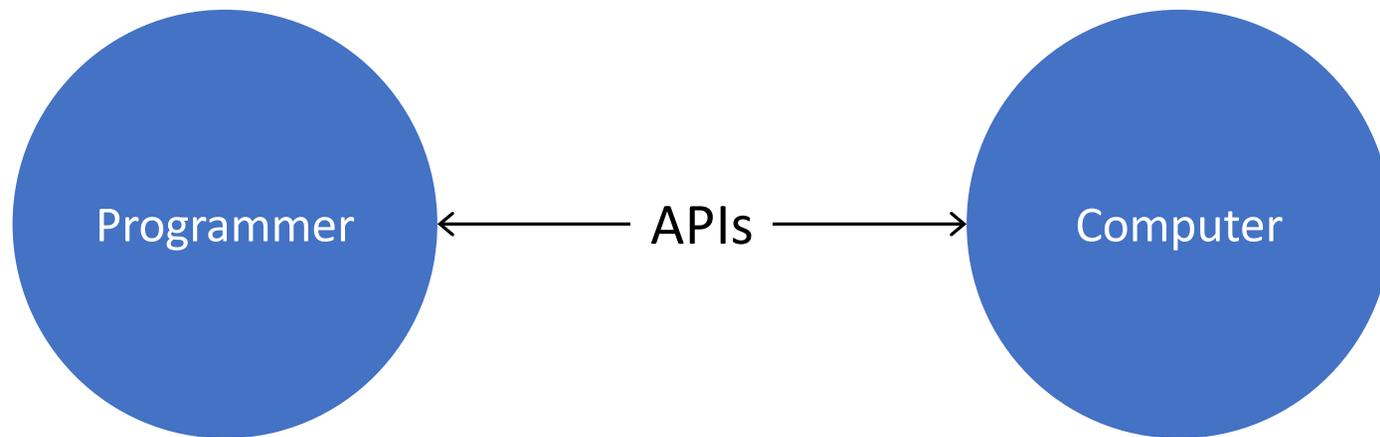
Tianyi Zhang¹, Björn Hartmann², Miryung Kim³, Elena Glassman¹

¹Harvard University ²UC Berkeley ³UCLA

APIs are ubiquitous



APIs are a primary interface between programmers and computers



focus

cooperative and human aspects of SE.....

What Makes APIs Hard to Learn? Answers from Developers

Martin P. Robillard, *McGill University*

A study of obstacles that professional Microsoft developers faced when learning to use APIs uncovered challenges and resulting implications for API users and designers.

Most software projects reuse components exposed through APIs. In fact, current-day software development technologies are becoming inseparable from the large APIs they provide. To name two prominent examples, both the Java Software Development Kit and the .NET framework ship with APIs comprising thousands of classes supporting tasks that range from reading files to managing complex process workflows.

An API is the interface to implemented functionality that developers can access to perform various tasks. APIs support code reuse, provide high-level abstractions that facilitate programming tasks, and help unify the programming ex-

and interviewing developers about the obstacles they faced learning APIs, I discovered many issues that complement those mentioned in API design textbooks and articles. In particular, I found that API learning resources are critically impor-

Human-Centered API Design

- Apply HCI methods to API design, e.g., A/B testing [1,2,3]
- But it is costly.
 - Too many APIs and usage scenarios
 - Participant recruitment

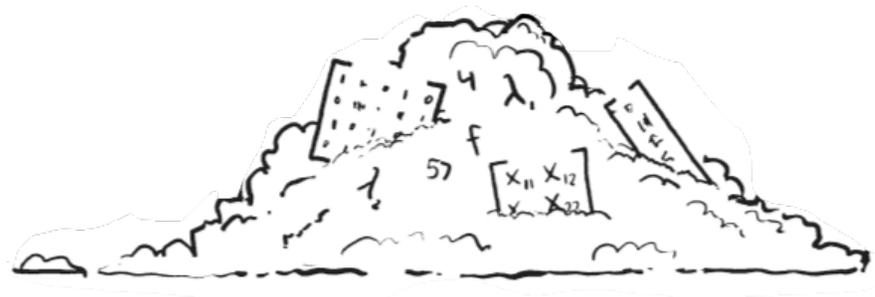
Human-centered design can make application programming interfaces easier for developers to use.

BY BRAD A. MYERS AND JEFFREY STYLOS

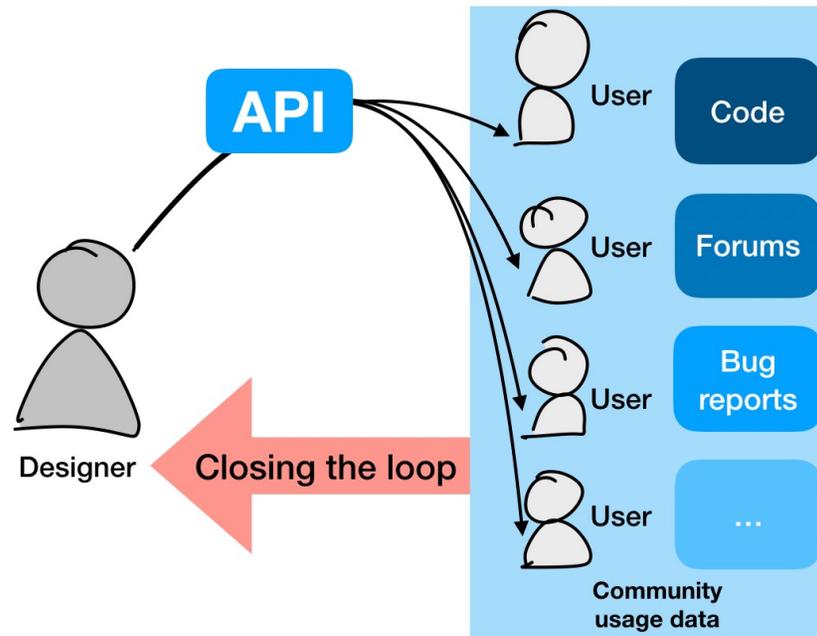
Improving API Usability

1. Ellis et al. The factory pattern in API design: A usability evaluation. ICSE 2007
2. Stylos et al. A case study of API design for improved usability. VL/HCC 2008.
3. Stylos et al. The implications of method placement on API learnability. FSE 2008.

A lot of API usage data has been generated from programmer communities...



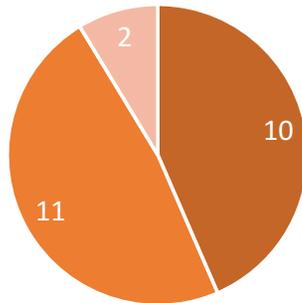
How can we leverage community data to inform better API design?



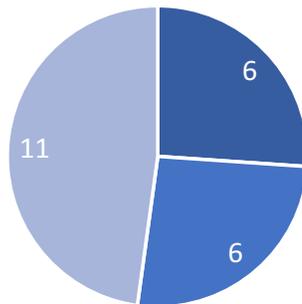
Contributions

- Semi-structured interviews with 23 API designers
- An in-depth analysis of their design styles, usability evaluation methods in practice, and unmet information needs
- Several tool design implications for leveraging community API usage data to inform API design

Interview with 23 API Designers



■ Large Tech Companies ■ Academia ■ Open-source



■ Web APIs ■ DSLs ■ Libraries



Interview Questions

- What **design decisions and tradeoffs** have you made?
- How do you **evaluate** the usability of your APIs?
- What **challenges and frustrations** do you have?
- What **information** would you like to discover?
- What **tool support** do you need?

Finding 1. Designer Spectrum



User-driven



Self-driven



Visionary



Closed-world

User-Driven API Designer

- User-centered design process
- Survey user needs
- API stability



Self-Driven API Designer

- Make their own decisions
- Clear goals and priorities
- API expressiveness & extensibility



Visionary API Designer

- Lack communication channels to users
- Design with imagined use cases
- Eager to know API usage in the wild



Closed-World API Designer

- APIs used internally in a company
- Work closely with all stakeholders
- Least tension between API designers and users



Finding 1. Designer Spectrum



User-driven



Self-driven



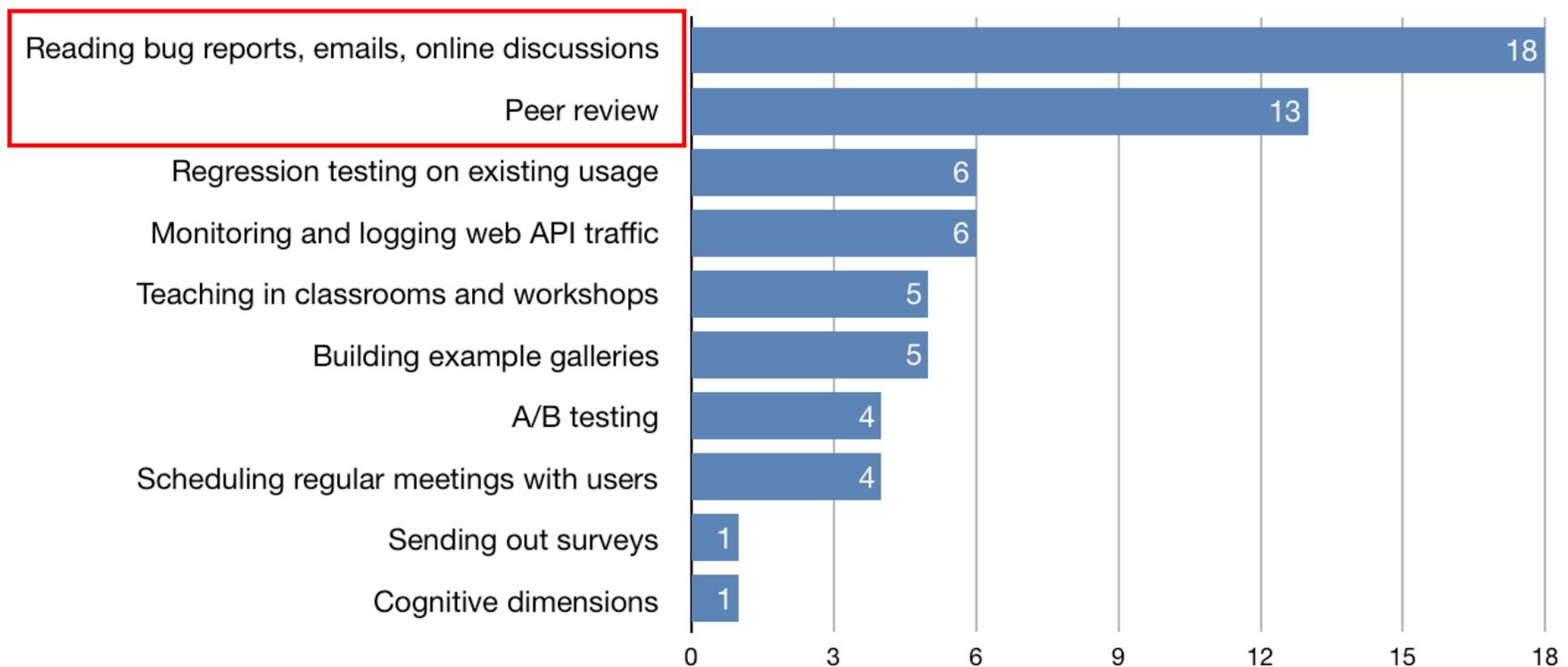
Visionary



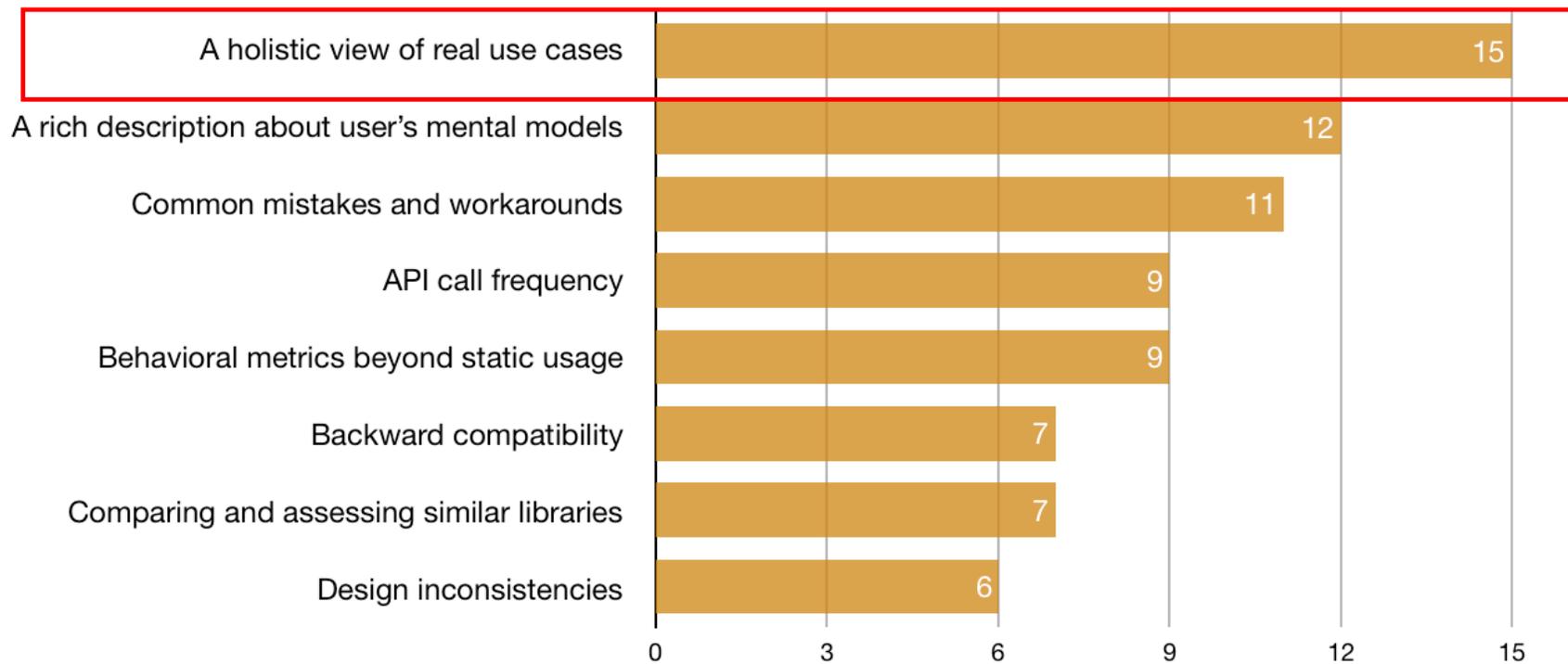
Closed-world

They all acknowledged the importance of keeping users in mind.

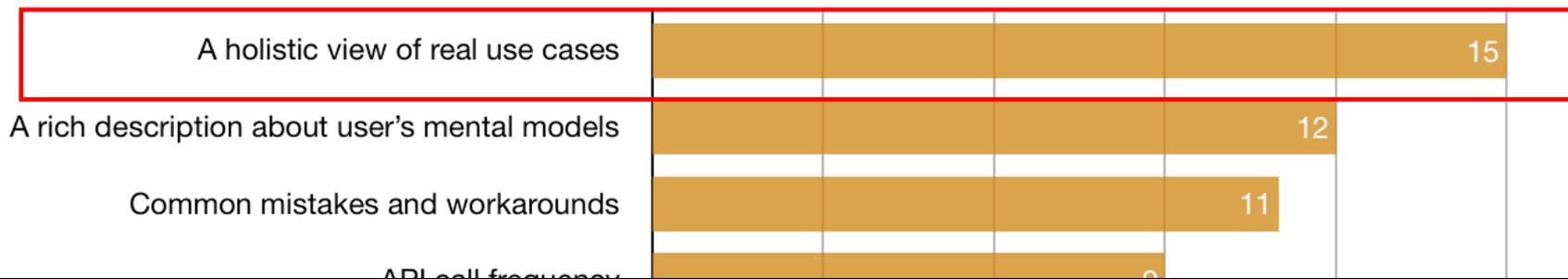
Finding 2. Usability Evaluation Methods



Finding 3. Unmet Information Needs

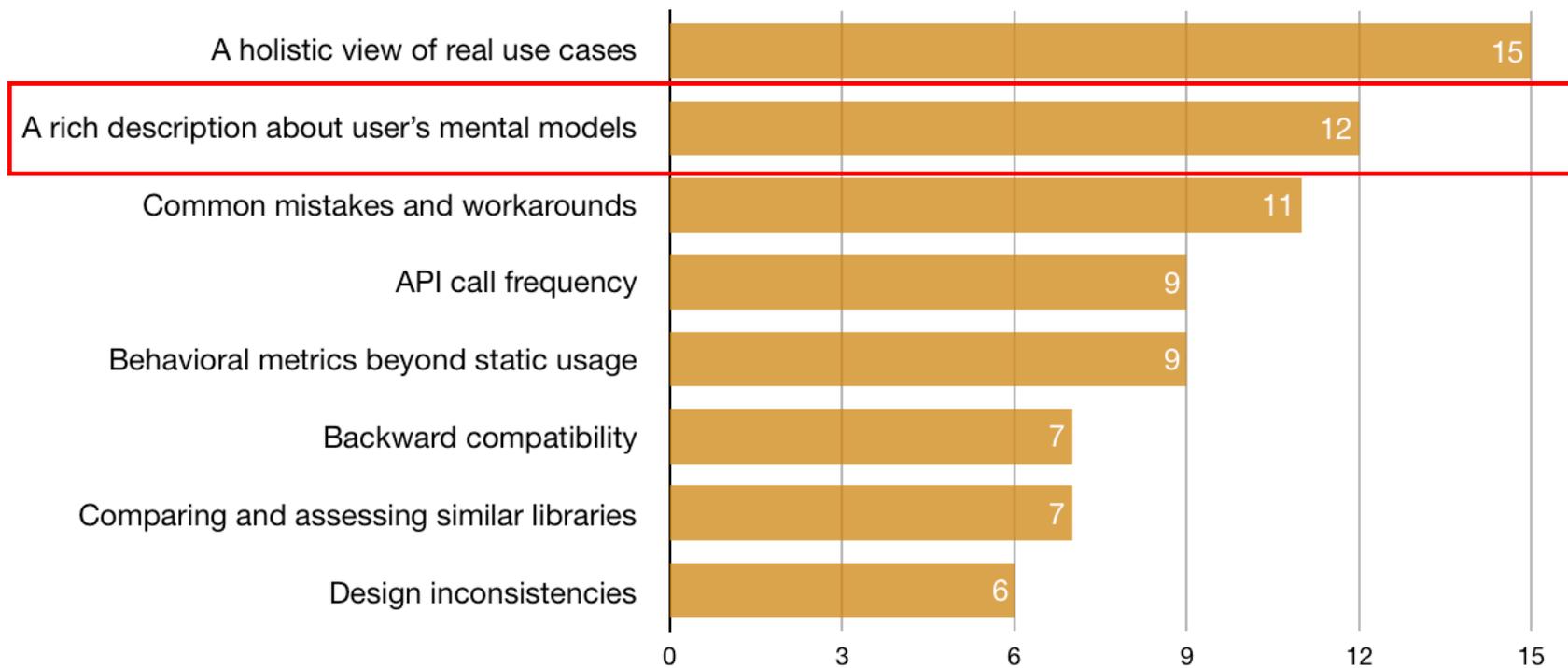


Finding 3. Unmet Information Needs



P13: “I’d like to **look at their code** and see if they write code **in the ideal way we want them to write**. Because every API has its own purpose when they’re being designed, and we have some use cases in mind. If not, we’ll probably think if this API is designed in a proper way or we should **create some more obvious APIs to deal with this case**.”

Finding 3. Unmet Information Needs

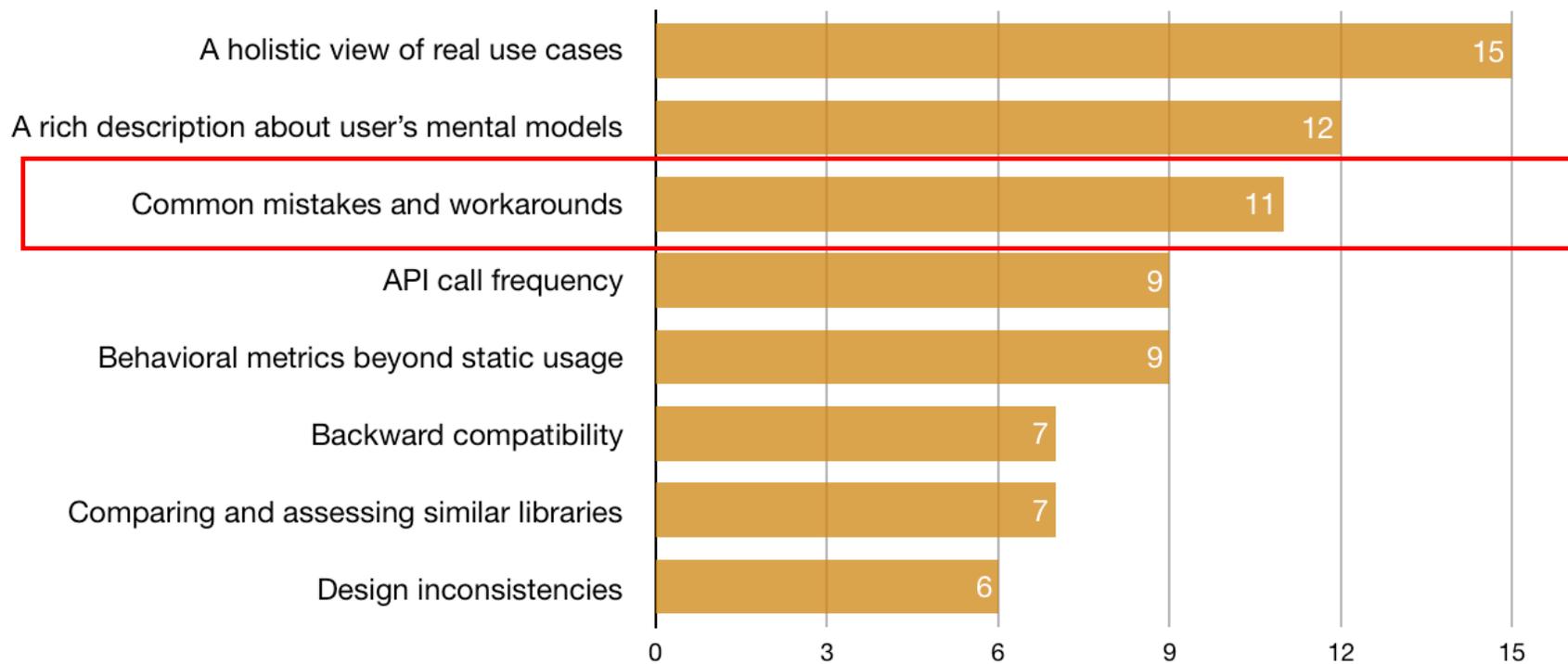


Finding 3. Unmet Information Needs

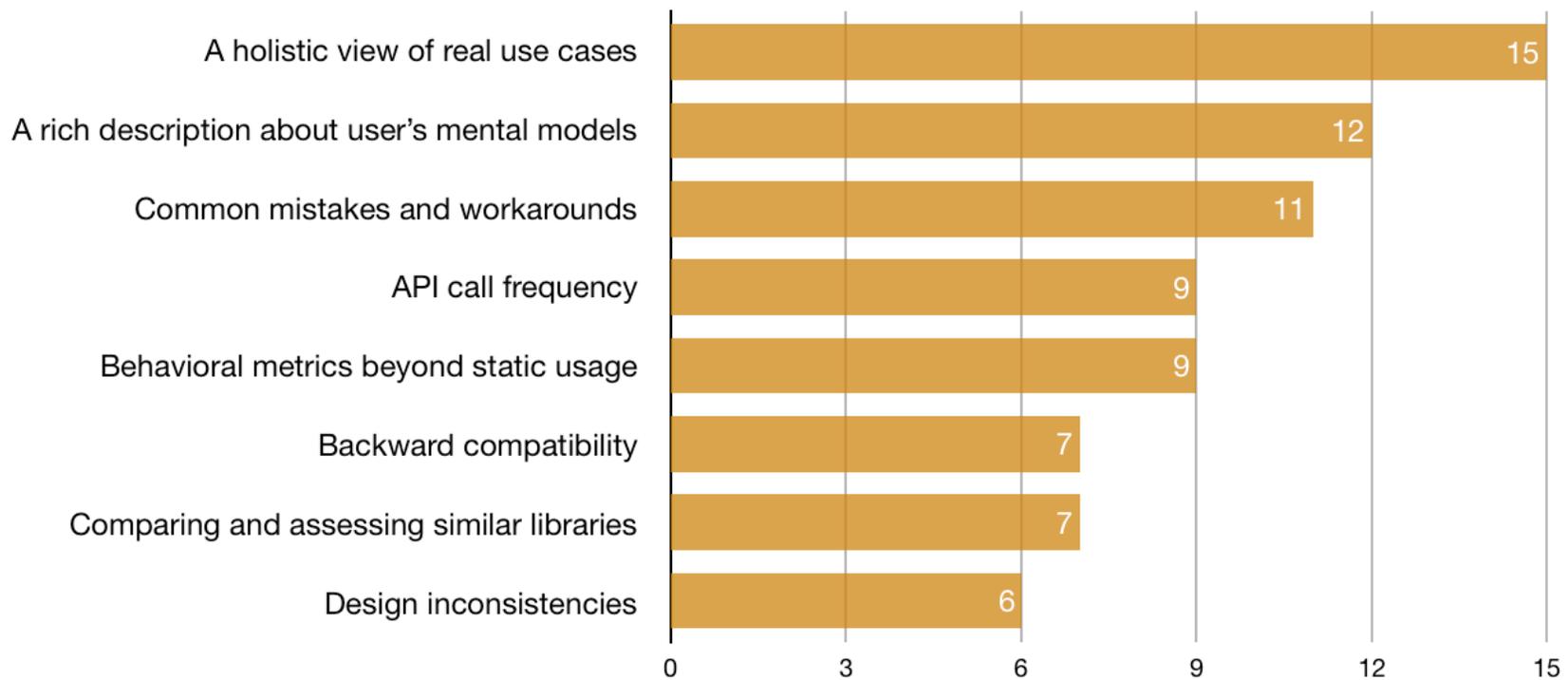


P6: “I want to get what people are thinking, **not just whether they are able to use this API**, like that’s what you would get from like mining GitHub repos. But what you are not getting is like this person still **spent three hours googling** how to use this API, and he **spent an hour on Stack Overflow** trying to figure out what was going on, in the end, he got that, but it was a **frustrating attempt**.”

Finding 3. Unmet Information Needs



Finding 3. Unmet Information Needs



Finding 4. Opportunities for Tool Support

- Mining and visualizing API usage mistakes
- Interactively analyzing population-level API usage
- Adapting elicitation mechanisms to understand API users' mental models
- Exploring the design space of similar APIs
- Live API documentation

Mining and Visualizing Documented Mistakes

<> Code **Issues 3,419** Pull requests 204 Actions Projects 1 Security 2 Insights

tf.keras.metrics.MeanIoU API is practically unusable without a threshold #39173

New issue

Open dd1923 opened this issue 9 days ago · 3 comments



dd192

[TF 2.0] tf.assert_equal([], [1.0]) doesn't raise error #32

Closed

David-Mao opened this issue on Aug 29, 2019 · 14 comments

Please
GitHub
we only
build/ir

System

- Ha
- Tei
- OS
- Mc
- Tei
- Tei
- Py
- Ba
- GC
- CL
- GF

Descri



David-Mao commented on Aug 29, 2019



System ir

Does keras.tokenizer.text_to_sequence simply translate into number vectors, or something more?

- Have
- Tensi

Yes Asked today Active today Viewed 14 times

- OS P

Darw

- Mobi

N/A

- Tensi

binar

Microsoft Azure logo and a 3D cube graphic. Text: "Experiment, learn, and build with our top services, free for 12 months". Button: "Try Azure free >"

im currently trying to learn the ins and outs of keras. in working with a dataset containing sentences, I m doing the following

1

The Overflow Blog

Podcast 234: We're do

Ensuring backwards cc distributed systems

Capturing **Undocumented** Learning Barriers

- Participants suspect a lot of issues were not reported at all.

P9: *“I think the people that use the API with no prior experience perhaps hit the wall quickly. That’s the real issue. But for the most part, I suspect that I **missed most of the easy problems** going on there because **they don’t convey that first barrier** with their remote communication.”*

Capturing Undocumented Learning Barriers

- Compilation/runtime errors are good indicators of undocumented learning barriers.

```
11 raw_name: str = input("Please enter your name: ")
12 first_name: str = get_first_name(raw_name)
13
14 # If the user didn't type anything in, use the fallback name
15 i not first_name:
16     first_name = get_first_name(fallback_name)
17
18 print(f"Hi, {f Expected type 'str', got 'Dict[str, str]' instead more... (%F1)
```

```
9 public static void foo(String s) {
10     System.out.println(s.toLowerCase());
11 }
12
13
```

```
Map.java
1 package com.industriallogic.collections;
2
3 public class Map {
4     private static int INITIAL_CAPACITY = 10;
5     protected List keys = new Object[INITIAL_CAPACITY];
6     protected List values = new Object[INITIAL_CAPACITY];
7     private int size = 0;
8     private int indexWhereKeyFound;
9     private boolean readOnly;
10
11     public boolean isEmpty() {
12         return size == 0;
```

```
Problems @ Javadoc Declaration Console
<terminated> Temp [Java Application] /Library/Java/JavaVirtualMachines/jdk1.8.0
Exception in thread "main" java.lang.NullPointerException
    at Temp.foo(Temp.java:10)
    at Temp.main(Temp.java:5)
```

Conclusion

- API designers desire to have a holistic view of real API use cases.
- Such real use cases are gathered in an informal way due to a lack of tool support.
- New interactive systems are needed to distill a variety of API usage insights from community data.