

Fingerprinting the Fingerprinters: Learning to Detect Browser Fingerprinting Behaviors

Umar Iqbal
University of Iowa

Steven Englehardt
Mozilla Corporation

Zubair Shafiq
University of California-Davis

Overview

- ❑ Mainstream browsers are implementing countermeasures against third-party cookie-based cross-site tracking
- ❑ Trackers are expected to migrate to browser fingerprinting, which does not require cookies, to track users
- ❑ Existing countermeasures against fingerprinting limit website functionality, cause website breakage, and are not scalable
- ❑ We propose FP-Inspector, a syntactic-semantic machine learning approach that detects browser fingerprinting
- ❑ FP-Inspector detects 26% more scripts than the state-of-the-practice with an accuracy of 99%

Fingerprinting Behavior

- ❑ Script's internal and external context
- ❑ Perspective about the inner workings of the script
 - ❑ Does the script mostly contain APIs that have fingerprinting potential?
- ❑ Perspective about the script's collaboration with external entities
 - ❑ Does the script interact with external entities?

```
<script>
// Canvas font fingerprinting script.
Fonts = ["monospace", ... , "sans-serif"];
CanvasElem = document.createElement("canvas");
CanvasElem.width = "100";
CanvasElem.height = "100";
context = CanvasElem.getContext('2d');
FPDict = {};

for (i = 0; i < Fonts.length; i++) {
  CanvasElem.font = Fonts[i];
  FPDict[Fonts[i]] = context.measureText("example").width;
}

var img = document.createElement("img");
img.src = "tracker.com/track_user/?userId={FPDict}";
</script>
```

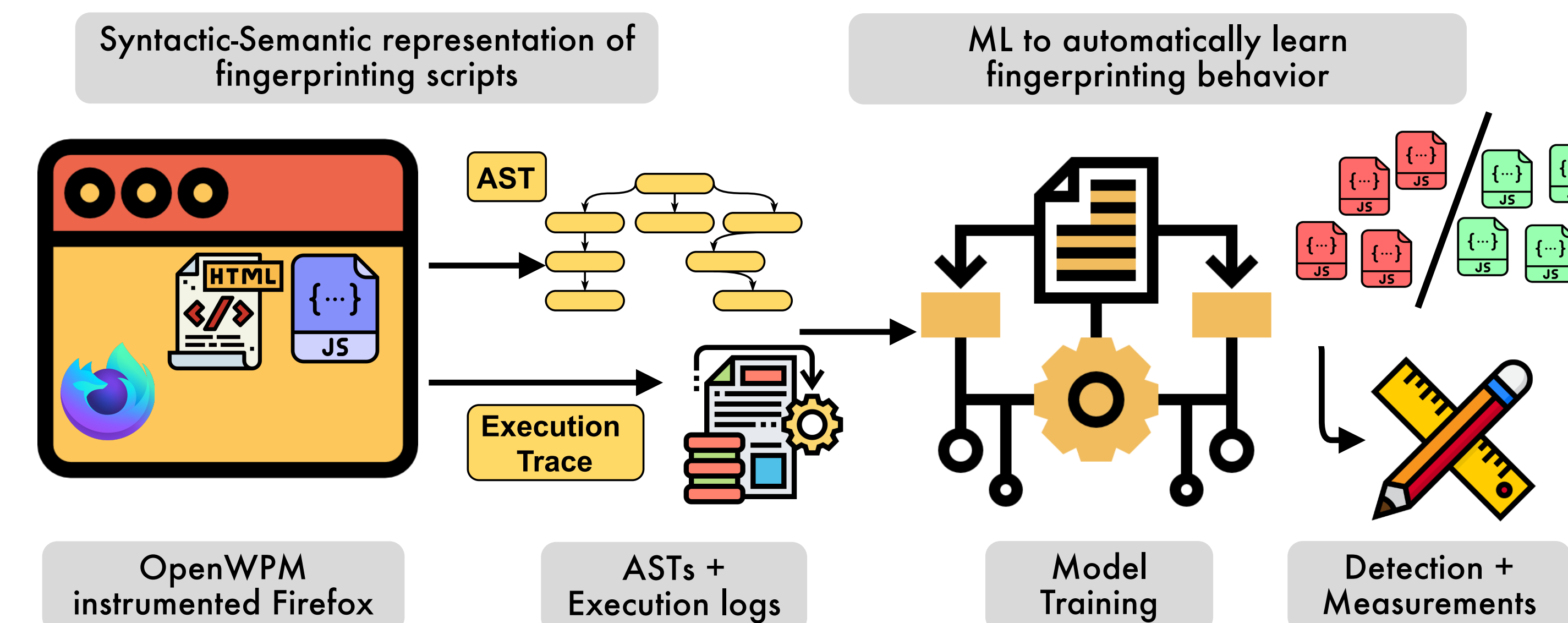
HTML canvas element

APIs with fingerprinting potential

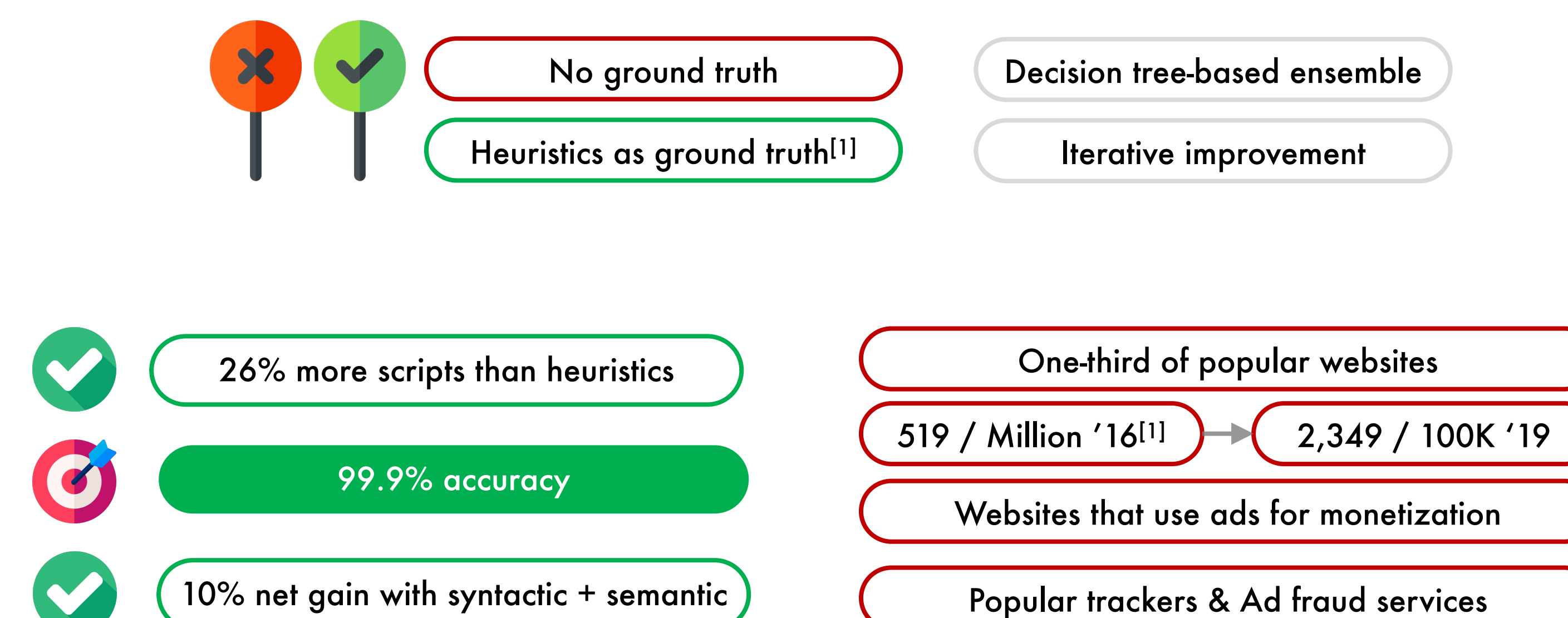
Network request with fingerprint

FP-Inspector

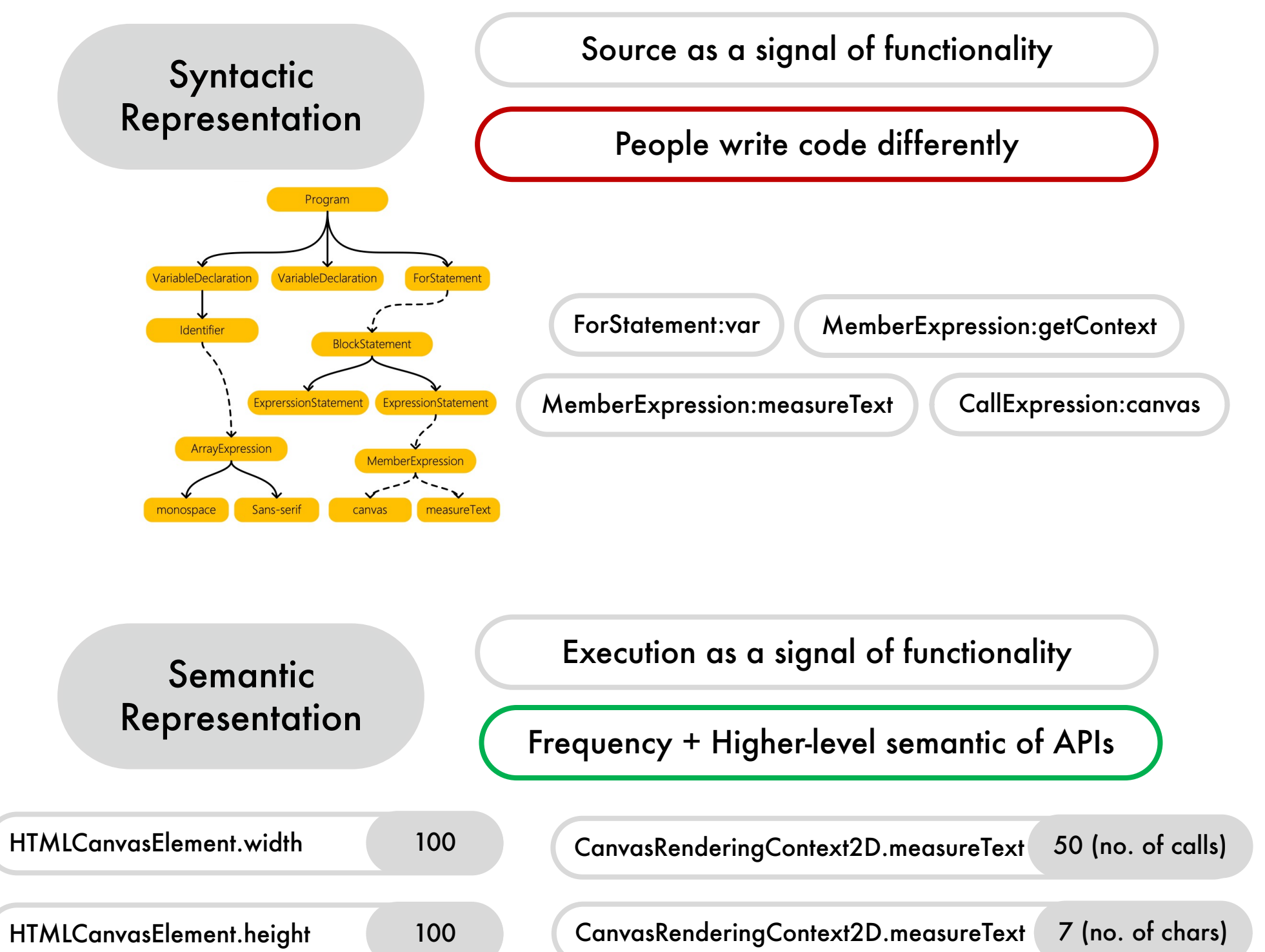
- ❑ A syntactic-semantic machine learning approach that detects browser fingerprinting
- ❑ Syntactic-semantic representation to model script's behavior
 - ❑ Syntactic representation is created through Abstract Syntax Trees (ASTs)
 - ❑ Semantic representation is created through script's execution
- ❑ Machine learning to learn fingerprinting patterns
 - ❑ Combination of APIs commonly used for fingerprinting
 - ❑ Limited interaction with the webpage
 - ❑ Communication with external entities



Evaluation



Syntactic & Semantic



Key Takeaways

- ❑ FP-Inspector improves the state-of-the-art in browser fingerprinting detection by incorporating syntactic-semantic representation
- ❑ Fingerprinting adoption is on the rise with more than one-third of top 1K popular websites using fingerprinting
- ❑ We open source our implementation and detections so that the community can benefit from them
- ❑ FP-Inspector's detections are incorporated by popular filter lists, such as EasyPrivacy, Disconnect, and DuckDuckGo
- ❑ More analysis like fingerprinting APIs discovery in the paper!

Fingerprinting the Fingerprinters: Learning to Detect Browser Fingerprinting Behaviors
Umar Iqbal, Steven Englehardt, and Zubair Shafiq
IEEE S&P 2021



Paper



Source & Data

Umar Iqbal
@umaarr6
umariqbal.com

Contact details