

Testability Patterns for Web Applications – an OWASP project



Luca Compagna, SAP Security Research 24 November 2023

PUBLIC









About me and others who contribute to this work





Luca Compagna
SAP Security Research



https://www.testable.eu/

joint work with:

Feras Al Kassar (SAP), Giulia Clerici (SAP), Fabian Yamaguchi (SHIFTLEFT), Davide Balzarotti (EURECOM)

[1] Testability Tarpits: the Impact of Code Patterns on the Security Testing of Web Applications. Al Kassar et al. NDSS 2022

An OWASP project: our journey started recently!

Security and privacy **testing** is core for protecting the Web

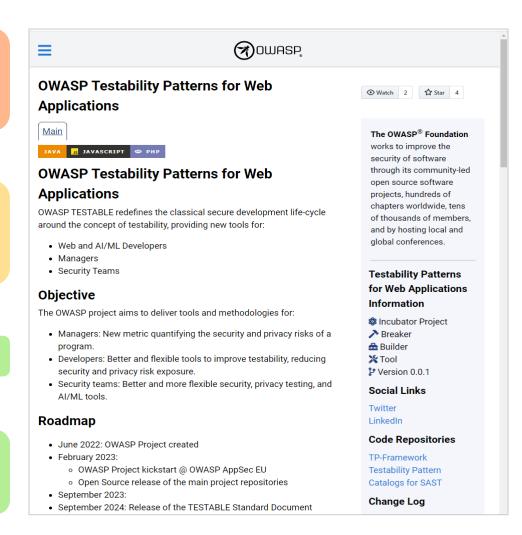
Many challenges impact the testing accuracy!

Motivated by promising **research results**, we started a **new OWASP project** to precisely identify, measure and possibly remediate some of these challenges!

Concrete results for SAST

Join our project and our community to help achieving this vision!

https://owasp.org/www-project-testability-patterns-for-web-applications/



An OWASP project: our journey started recently!

Security and privacy **testing** is core for protecting the Web

Many challenges impact the testing accuracy!

Motivated by promising **research results**, we started a **new OWASP project** to precisely identify, measure and possibly remediate some of these challenges!

Concrete results for SAST

Join our project and our community to help achieving this vision!

https://owasp.org/www-project-testability-patterns-for-web-applications/





Testability patterns for SAST

(first concrete result of our OWASP project)

Testability patterns for SAST



chatGPT 3 (around 01/2023)



what is a testability pattern for static code analysis?



A testability pattern for static code analysis is a set of guidelines or best practices for designing code in a way that makes it easy to test using static analysis tools. This may include things like writing clear, modular code with well-defined interfaces, using dependency injection to decouple components, and following established coding conventions and design patterns. The goal is to make it easier to automatically identify and analyze code for potential bugs or other issues.



7

2021 knowledge base!

Testability

chatGPT 3.5 (today)



You

what is a testability pattern for sta



ChatGPT

Testability patterns in static code testability of software through and "Dependency Injection" pattern.

Dependency Injection (DI) involve class create those dependencies allowing the substitution of deper

For instance, if a class relies on an inject a mock service or a fake datenables better isolation of the codwithout the need for external resources.

By utilizing DI in static code analysis, you can identify areas where dependent or instantiated within classes, making it challenging to isolate Refactoring code to follow DI principles can greatly enhance testability comprehensive and efficient testing processes.



You

are you aware of the wor Testing of Web Applicat



ChatGPT

Yes, that work explores obstacles to security testructures and design c

The study likely examine implications for security meaningful test cases for targeted testing.

The goal of such researd potentially propose stra from a security perspection work towards creati



You

can you refine the definition of testability pattern for static code analysis as presented in that work?



ChatGPT

In the context of the work "Testability Obstacles: The Impact of Code Patterns on the Security Testing of Web Applications," the definition of a testability pattern for static code analysis might encompass specific recurring code structures or design characteristics within web applications that impede effective security testing when analyzed statically.

These patterns could include:

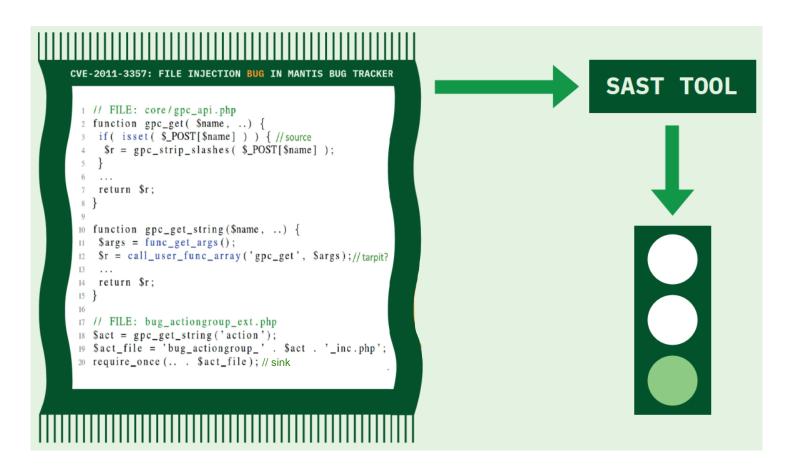
- Tightly Coupled Dependencies: Identifying code sections where dependencies are tightly coupled, making it challenging to isolate components for testing or security analysis.
- Excessive Complexity: Highlighting code structures that are overly complex, making it difficult to understand and create comprehensive tests or conduct thorough security analysis.
- Lack of Modularity: Recognizing areas where the code lacks modularity, hindering the ability to test individual components independently or assess security vulnerabilities in isolation.
- Inadequate Input Validation: Identifying patterns where input validation is insufficient or improperly implemented, leading to potential security vulnerabilities that are hard to detect through static analysis alone.
- Obscure Control Flow: Noting instances where control flow within the code is convoluted or hard
 to trace, making it challenging to identify potential security weaknesses or create effective test
 cases.

The refinement of testability patterns in this context involves pinpointing these recurring code structures or design characteristics that pose obstacles to effective static analysis for security testing. By recognizing and addressing these patterns, the aim is to enhance the overall testability and security posture of web applications.





Static application security testing (SAST) is widely used in industry to detect vulnerabilities [1]

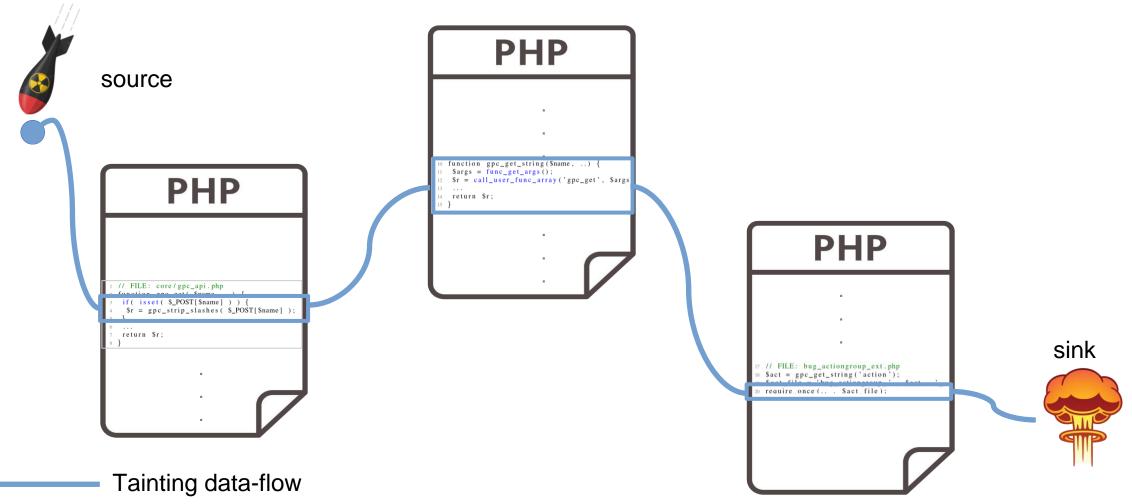


[1] OWASP Code Review Guide v2.0, cf. Figure 1 and Figure 2

Context: Injection vulnerabilities



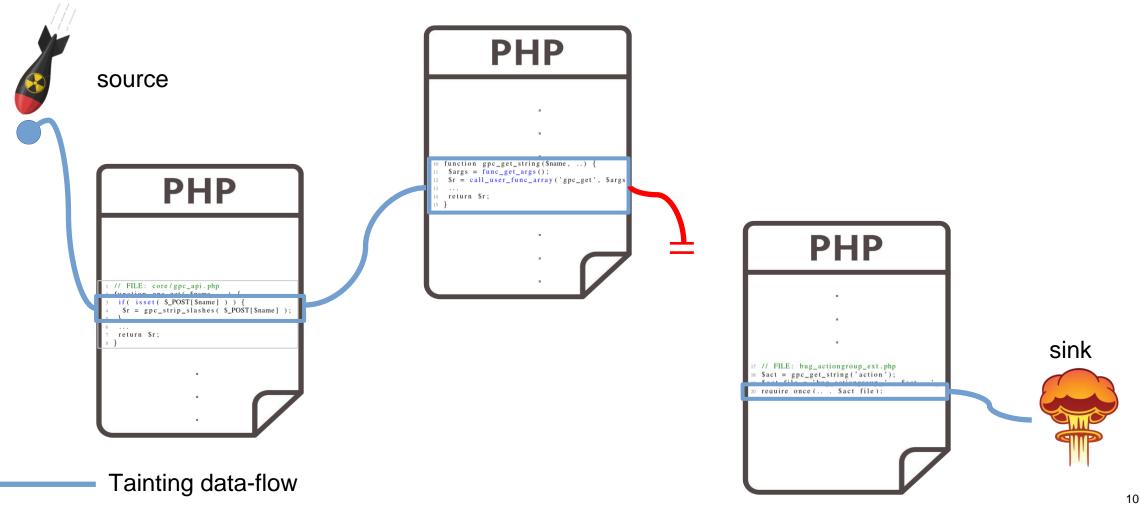
Any attacker-controlled data (source) flowing in a dangerous operation (sink) without sanitization?



Context: Injection vulnerabilities



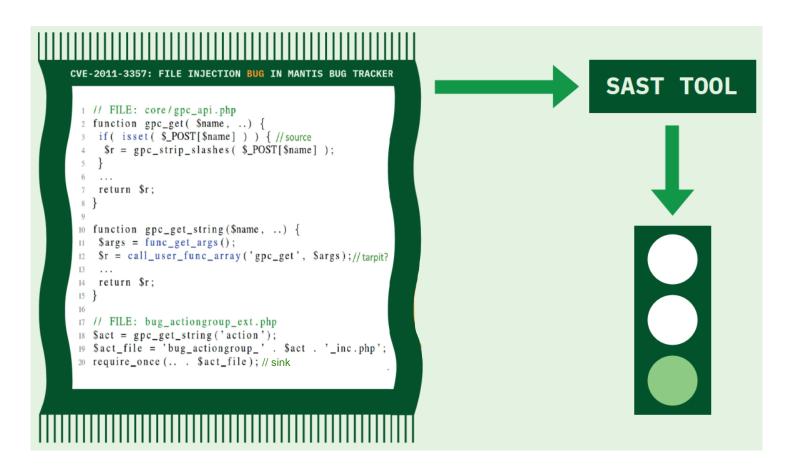
Any attacker-controlled data (source) flowing in a dangerous operation (sink) without sanitization?







Static application security testing (SAST) is widely used in industry to detect vulnerabilities [1]

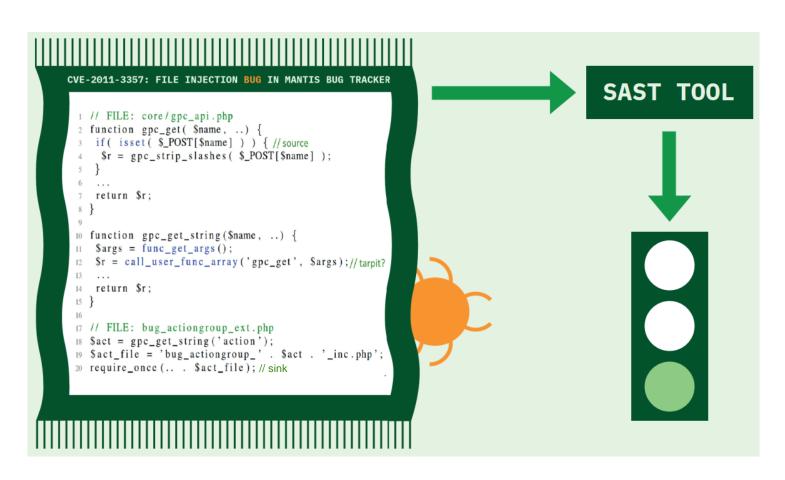


[1] OWASP Code Review Guide v2.0, cf. Figure 1 and Figure 2

Context: SAST and testability



Static application security testing (SAST) is widely used in industry to detect vulnerabilities [1]



Was all the code analyzed? No bugs under the carpet?

Technical questions

Code obstacles impacting SAST?

Can we measure these obstacles?

Can we discover / remediate them?

[1] OWASP Code Review Guide v2.0, cf. Figure 1 and Figure 2





20 require_once (... act_file); \angle **EXECUTE** THE CONTENT OF A FILE < ABUSABLE FUNCTION > **SINK LIST**





```
20 require_once(.. . $act_file); // sink
```

CVE-2011-3357: File inclusion in mantis bug tracker



```
FUNCTION CALL < PROPAGATES ... > TAINTING CAPABILITIES

STRING CONCAT < PROPAGATES THE CONTROL OF A STRING TO ANOTHER > TAINTING CAPABILITIES

17  // FILE: bug_actiongroup_ext.php

18  $act = gpc_get_string('action');

19  $act_file = 'bug_actiongroup_' . $act . '_inc.php';

20  require_once(... $act_file); // sink
```





```
function gpc_get_string($name, ...) {
    $args = func_get_args();
    $r = call_user_func_array('gpc_get', $args);
    ...
    return $r;

}

// FILE: bug_actiongroup_ext.php

8    $act = gpc_get_string('action');
    $act_file = 'bug_actiongroup_' . $act . '_inc.php';
    require_once(... $act_file); // sink
```



CVE-2011-3357: File inclusion in mantis bug tracker

```
1 // FILE: core/gpc_api.php
2 function gpc_get( $name, ..) {
                                          $ POST <TAKEN FROM UNTRUSTED SOURCE> SOURCE LIST
   if ( isset ( \$_POST[\$name] \_)
   $r = gpc_strip_slashes( $_POST[$name] );
   return $r;
  function gpc_get_string(\$name, \dots) \{
   $args = func_get_args();
   $r = call_user_func_array('gpc_get', $args);
  return $r;
16
  // FILE: bug_actiongroup_ext.php
  $act = gpc_get_string('action');
  $act_file = 'bug_actiongroup_' . $act . '_inc.php';
 require_once(.. . $act_file); // sink
```



CVE-2011-3357: File inclusion in mantis bug tracker

```
1 // FILE: core/gpc_api.php
2 function gpc_get( $name, ..) {
   if ( is set ( $_POST[$name] _) _) { // source
   $r = gpc_strip_slashes( $_POST[$name]_);
                                                         POST <a href="https://mantisb.com/service">https://mantisb.com/service</a>
                                                         i action=<ATTACKER_PAYLOAD>
   return $r;
 function gpc_get_string(\$name, \dots) \{
   $args = func_get_args();
   $r = call_user_func_array('gpc_get', $args);
  return $r;
  // FILE: bug_actiongroup_ext.php
  $act = gpc_get_string('action');
  $act_file = 'bug_actiongroup_' . $act . '_inc.php';
 require_once(.. . $act_file); // sink
```





Many SAST tools (including commercial ones) do not find that File inclusion

pattern

creation (1)

vulnerable app

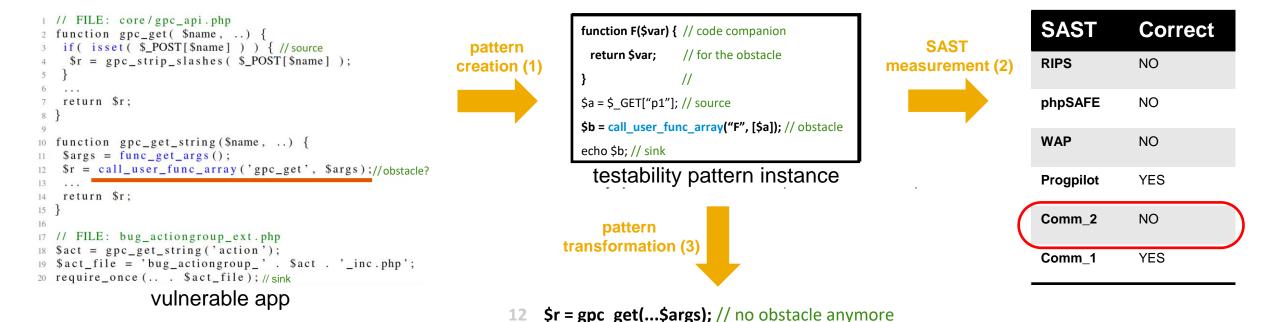
```
// replace with
// code companion for the obstacle
//
$a = $_GET["p1"]; // source
$b = $a // replace with obstacle!
echo $b; // sink
```

testability pattern skeleton (baseline XSS)

Toward testability patterns for SAST



Many SAST tools (including commercial ones) do not find that File inclusion



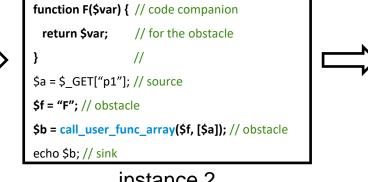
After that transformation, commercial tool Comm_2 finds the File inclusion!

Many variants of that pattern can be created...

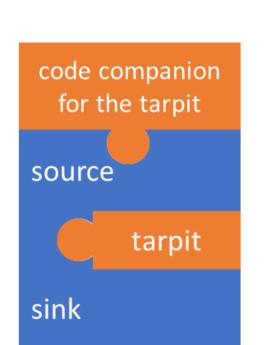


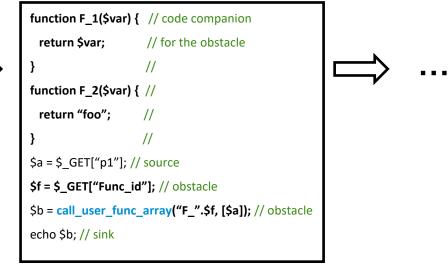
```
function F($var) { // code companion
 return $var;
                 // for the obstacle
                  //
$a = $_GET["p1"]; // source
$b = call_user_func_array("F", [$a]); // obstacle
echo $b; // sink
```

instance 1



instance 2





instance 3

Testability Patterns for SAST

Targeted 3 popular languages

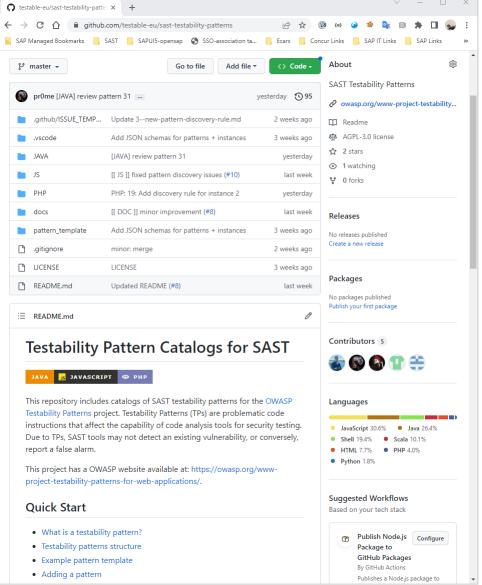
- PHP: ~120 pattern instances (mature: 8/10)
- JS: ~150 pattern instances (mature: 6/10)
- Java: ~200 pattern instances (mature: 3/10)

We inspected the entire language manual

What do we want to do with them?

- SAST tools measurement
- Discover patterns into applications
- Remediate patterns to increase testability





Testability Patterns Framework for SAST

Framework provides operations for

- SAST tools measurement (e.g., codeql integrated)
- Discover patterns into apps (via Joern and Scala queries)

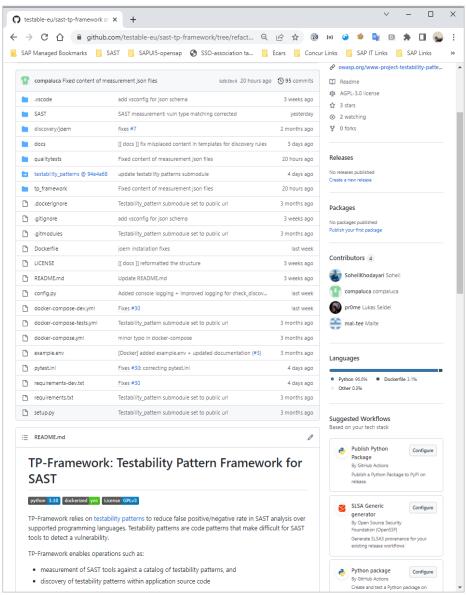
```
tpframework measure -1 JS -p 1 2 --tools codeq1:2.13.1
```

tpframework discovery -t MYAPP/ -l PHP -a --tools T1:V1 T2:V2

Results spoiler

- Measurement: many SAST tools struggle on our patterns
- Discovery: many apps in Github use these patterns
- Testability for SAST can be improved!



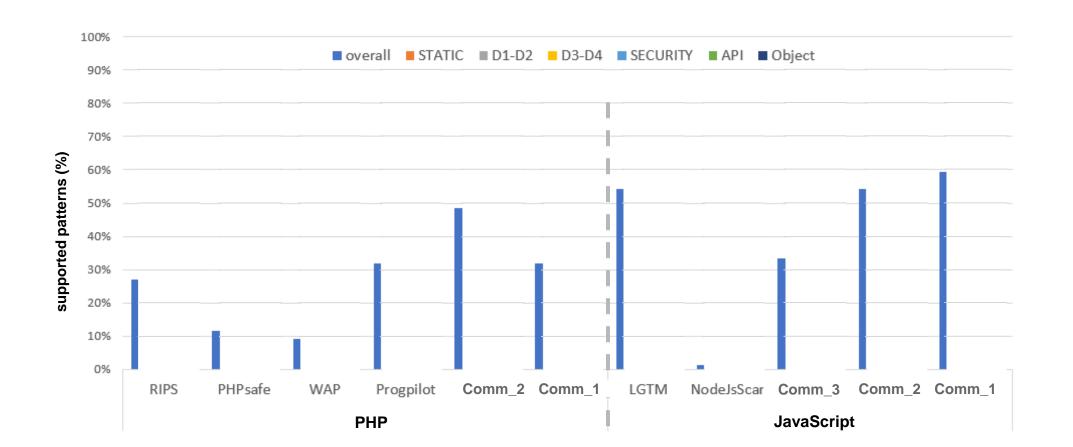






Measured our pattern instances against SAST

- Overall: <50% support for PHP and <60% for JS
- Tools Combination: 66% PHP, 82% JS

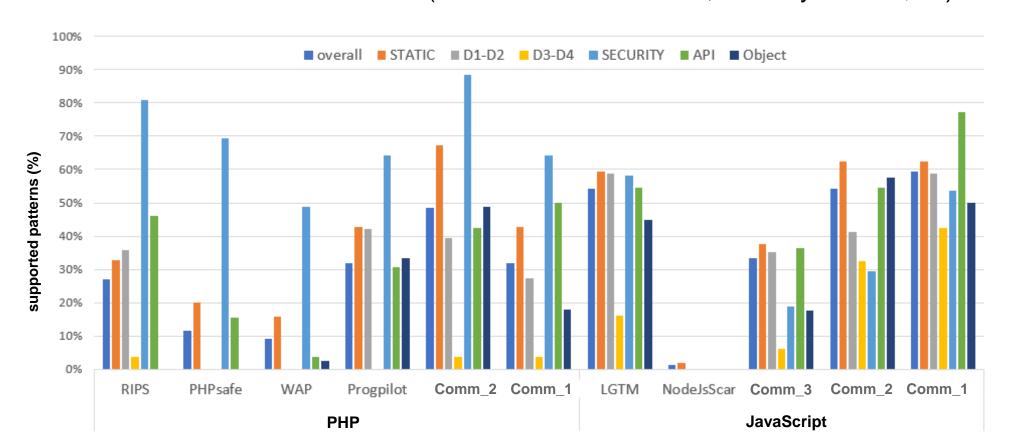






Measured our pattern instances against SAST

- Overall: <50% support for PHP and <60% for JS
- Tools Combination: 66% PHP, 82% JS
- Studied different dimensions (internal API vs Feature, security-related, ...)



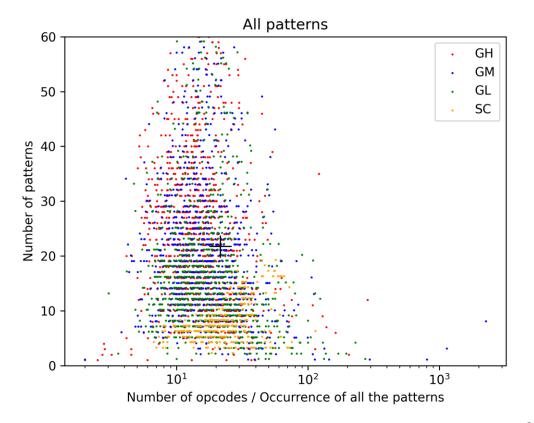




PHP: created discovery rules for our patterns and run them over >3000 open-source PHP apps (from Github and Sourcecodester)

Our patterns are very prevalent in the real world

	Unique obstacles per Project	obstacles per LoC			
AVG	21	20			



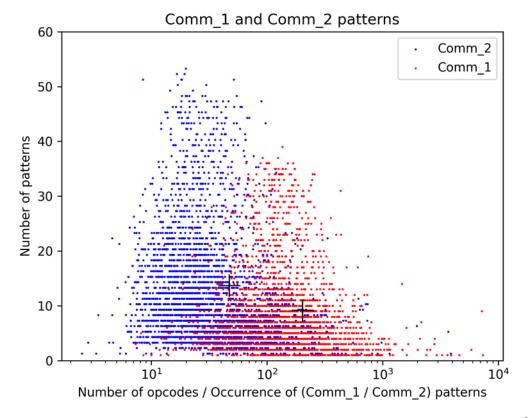




PHP: created discovery rules for our patterns and run them over >3000 open-source PHP apps (from Github and Sourcecodester)

Our patterns are very prevalent in the real world

	Unique obstacles per Project	obstacles per LoC			
AVG	21	20			
Comm_1	8	203			
Comm_2	13	47			



2. Discovery: many apps in Github use these patterns



2 glo 3 glo 4 con 5 con 6 coa 7 stri	atic_variables obal_variables obal_array nditional_assignment mbined_operator alesce ing_arithmetic_operations	1 1 1 1 1 1	✓	S S S S S			 -SXY WP-Y R-WPXY	prj 50 89 33	med 13 10 6	prj 443 203 138	med 4 7 4.0	prj 635 213 162	M med 7 10 5.0	712 210 179	med 14.0 12.0
2 glo 3 glo 4 con 5 con 6 coa 7 stri	obal_variables obal_array nditional_assignment mbined_operator alesce ing_arithmetic_operations mple_reference	1 1 1 1 1 1	>	S S S			-SXY WP-Y R-WPXY	89 33	10	203	7	213	10	210	12.0
3 glo 4 con 5 con 6 coa 7 stri	obal_array nditional_assignment mbined_operator alesce ing_arithmetic_operations mple_reference	1 1 1 1 1	✓	S S S			WP-Y R-WPXY	33							
4 con 5 con 6 coa 7 stri	nditional_assignment mbined_operator alesce ing_arithmetic_operations mple_reference	1 1 1 1	✓ .	S S			R-WPXY		6	138	4.0	162	5.0	179	` ~
5 con 6 coa 7 stri	mbined_operator alesce ing_arithmetic_operations mple_reference	1 1 1		S								I			7
6 coa 7 stri	alesce ing_arithmetic_operations mple_reference	1 1 1	,					221	74	795	18	890	31.5	908	59.0
7 stri	ing_arithmetic_operations mple_reference	1 1	✓	S			R-WPXY	335	170	919	33	942	64.0	934	97.5
	mple_reference	1	 	ı			RS-XY	0	0	0	0	280	6.0	433	11
8 sim				S			RSW-XY	277	10	523	6	636	9.0	707	11
		1		S			X-	42	39	163	9	231	5	292	6.0
	ference_argument	1		S			XY	208	14	387	7	399	10	486	9.0
	turn_by_reference	1		S				19	11	83	4	95	4	132	4.0
11 fore	reach_with_reference	1		S				41	7	182	3.0	238	4.0	321	4
12 mai	ake_ref	2		S		 	$P-\overline{Y}$	25	6	116	6.0	134	4.0	180	4.0
13 assi	sign_static_prop_ref	1		S			PX-	9	1.0	19	1	10	1.0	17	1
14 obj	ject_assigned_by_reference	1		S			X-	22	21	100	6.5	107	7	155	5
15 nes	sted_function	1		S			-SWPX-	66	3.5	166	4.0	222	4.0	283	5
	riadic_functions	1		S			XY	1	12	59	3	143	3	239	3
1	t_arguments	1		S			RY	23	5.0	106	4.0	137	4	191	4
	nd_unpack	1		S			RSX-	1	31	73	3	146	3.0	264	4.0
19 clos	osures	2		D2			XY	36	1	543	7	733	11	782	25.5
20 use	e_with_closures	2		D2		✓	$X\overline{Y}$	25	1	321	4	524	5.0	614	12.0
21 sim	nple_object	1		S	✓		PXY	336	199	968	350.5	977	863	974	1536.5
22 ass	sign_object	1		S	✓		W_X_	30	3	138	4.0	212	4.0	325	4
23 obj	ject_argument	1		S	✓		X_	119	30	591	23	718	53.5	804	79.5
24 nev	w_self	1		S	✓		X-	41	6	162	2.0	249	3	351	3
25 clos	one	1		S	✓		PX-	41	6	147	3	238	5.0	338	5.0
26 late	e_static_binding	2		D2	✓	🗸	$\overline{X}-$	13	1.0	165	4	279	5	386	8.0
27 get	t_called_class	1		D2	✓			0	0	16	1.0	28	1.0	34	2.0
28 stat	tic_methods	1		S	✓		PX-	119	17	792	29.0	865	61	898	126.5

i

callback_functions | 1 || | D1 | | | ---P-Y || 41 | 2 | 128 | 3.0 | 159 | call_user_func_array('gpc_get', \$args);

208 2.5

3. Testability for SAST can be improved!

Remediation 1: Two transformation experiments targeting PHP and JS

- transformations intended as code rewriting for obstacles
- >9000 new alerts: 370 true positives in 48 apps (over ~2700 alerts inspected)
- 182 true positives already confirmed from 31 projects: 38 impacting popular Github projects

Remediation 2: improve SAST tools

- Force SAST tools to collaborate to reduce false negatives, see our paper at USENIX 2023, [2]
- SAST tool vendors improving their tool on top of the testability patterns, e.g., our little story with codeQL

Remediation 3: provide custom rules to make SAST tools overcoming obstacles (on-going work)

[2] WHIP: Improving Static Vulnerability Detection in Web Application by Forcing tools to Collaborate. Feras Al-Kassar et al. USENIX Sec. 2023

Story 1: SAST tool improvement

SAP Security Research

- measuring SAST tools
- internal dissemination

SAP Security Testing Team

- processing the results
- selection of some results
- feedback to SAST tools
- e.g., codeql tickets

codeQL

- processing the results
- improving the tool
- releasing new version

SAP Security Research

- measuring new SAST tools' versions
- evaluating the improvement

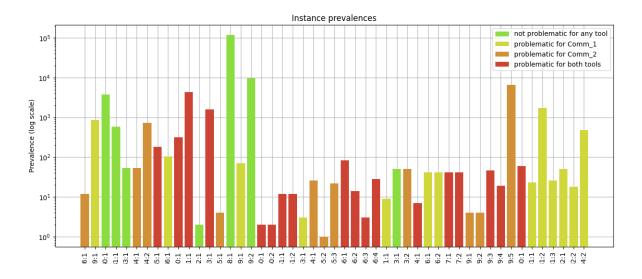
Story 2: Usage at SAP

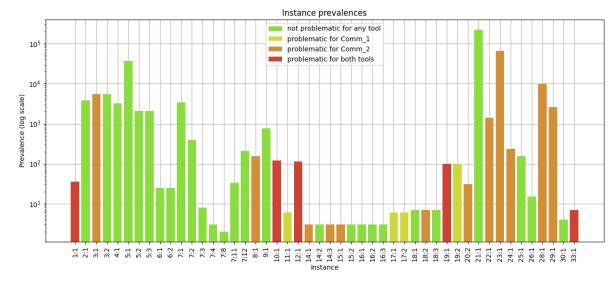
Testability of a SAP application wrt SAST

Which patterns are there?

Which patterns are more problematic?

Anything we can do to remediate them?





Call for action

Join our project and our community!

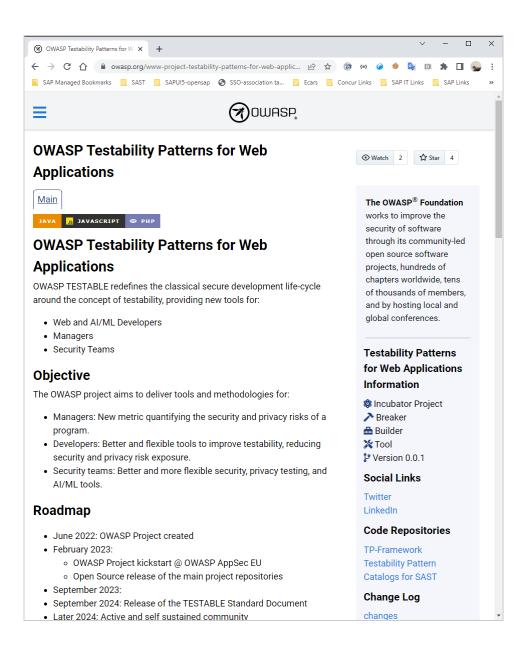
Contribute

- https://github.com/testable-eu/sast-testability-patterns
- https://github.com/testable-eu/sast-tp-framework

There is a lot to do

- testability patterns: create new ones, mature existing, target other programming languages...
- framework (python): improve features, integrate a new SAST tool, implement new features, ...

•



E.g. Developer of a SAST Tool



Lukas Seidel

Qwiet Al
formerly known as
ShiftLeft

Benefits you get

Measure SAST tools against Testability Patterns: identify strengths / weakness (B1)

Discover patterns in open source apps (B2)

(B1)+(B2): Prioritize which patterns to support better

Incentives to contribute

Define/improve patterns on which your tool outperform

Question/improve patterns on which your tool underperform

E.g. Web Developer



Malte Wessels



Benefits you get

Identify testability obstacles in your app (pattern discovery)

Remediate these obstacles to make your app more testable

Select which SAST tool to use on your app

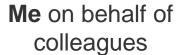
Incentives to contribute

Capture as testability patterns the challenges you are facing with SAST

Question/improve patterns that cannot be removed from your app

E.g. Security central team







Benefits you get

Identify testability patterns in the entire organization (pattern discovery)

Devise remediation strategies for top 10 patterns

- code refactoring practices can be introduced and disseminated to dev teams in the organization
- channel request for improvements to SAST tools

Select which SAST tool(s) to buy (also depending on the SAST measurement on the top 10 patterns)

Incentives to contribute

Capture as testability patterns the challenges the organization is facing with SAST

Question/improve patterns that cannot be removed from your organization

E.g. Auditors







Benefits you get

Identify testability patterns in their customer apps (pattern discovery)

Consult their customers to make their apps more testable

Select which SAST tool(s) to use on which customer app

Incentives to contribute

Capture as testability patterns the challenges experienced while using SAST on customer projects

Define/improve/remediate patterns to demonstrate thought leadership to their customers

OWASP Top 10 testability patterns for SAST?

Preliminary results for PHP

SAST measurement score

- weighted score of SAST tools failing on the pattern
- top performer tools (often commercial) get higher weight
- e.g., 3.9 == all the 6 SAST tools failed

Pattern discovery score

- how many findings of that pattern in a reasonable large and representative dataset
- e.g., for PHP
 - Github low popularity: 1,000 projects [20, 70] stars
 - Github medium popularity: 1,000 projects [200, 700] stars
 - Github high popularity: 1,000 projects > 1,000 stars
 - Sourcecodester: 341 Projects (customizable web apps)

	nottorn nomo	SAST measurement	
1	throw exception	score 3,9	score 2189
2	throw_exception static_variables	3,9 3,9	1802
3		3,9	1464
	variable_variables	•	
4	function_variable	3,9	1400
5	object_to_array	3,9	1182
6	methods_variable	3,9	834
7	foreach_with_reference	3,9	758
8	static_instance	3,9	699
9	new_from_variable	3,9	691
10	array_map	3,9	678
11	buffer	3,9	548
12	object_callable	3,9	487
13	dirname	3,9	347
14	dynamic_include	3,9	347
15	return_by_reference	3,9	317
16	extract	3,9	289
17	get_overloading	3,9	207
18	array_walk	3,9	181
19	call_overloading	3,9	178
20	callback_functions	3,9	178 ³⁷

An OWASP project: journey started, join us!

Targeting the **Testability** dimension

First concrete result: Testability Patterns for SAST

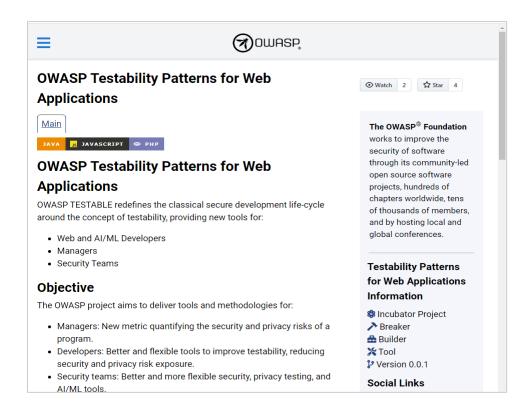
- https://github.com/testable-eu/sast-testability-patterns
- https://github.com/testable-eu/sast-tp-framework

Help us with your valuable expertise

https://owasp.org/www-project-testability-patterns-for-web-applications/

Let us devise OWASP top 10 testability patterns for SAST!

Can we do the same for DAST, Privacy, ML?





SAP Security Research

Thank you!

People

- 20 Security Researchers
- + PhDs (positions available)
- + interns (positions available)



Portfolio

