

# **is\_literal()**

**A proposed function for PHP**

**Craig Francis**

# is\_literal()

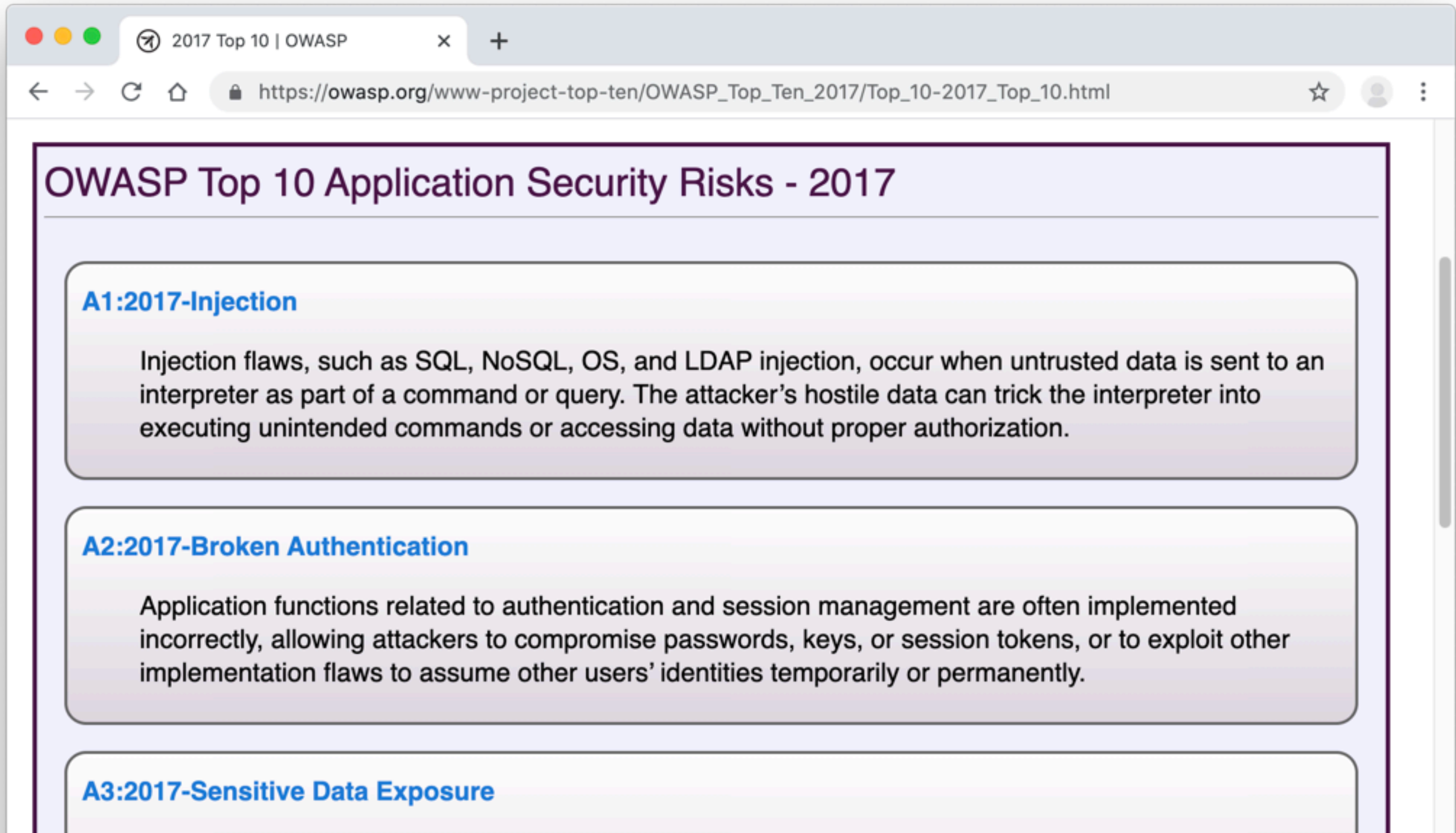
To distinguish between **safe** strings,  
which have been defined in your PHP script.

# `is_literal()`

To distinguish between **safe** strings,  
which have been defined in your PHP script.

**VS unsafe** strings,  
which have been tainted by external sources.

# The Problem





# The Problem


The screenshot shows a web browser window with the URL <https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=XSS>. The page is titled "CVE - Search Results" and displays the MITRE logo and navigation links. A black bar at the top contains links: "Search CVE List", "Download CVE", "Data Feeds", "Request CVE IDs", and "Update a CVE Entry". Below this, it states "TOTAL CVE Entries: 141126". The main content area shows "Search Results" with a message: "There are 16082 CVE entries that match your search." A table follows, listing CVEs and their descriptions.

Name	Description
<a href="#">CVE-2020-9758</a>	An issue was discovered in chat.php in LiveZilla Live Chat 8.0.1.3 (Helpdesk). A blind JavaScript injection lies i full account takeover. The attack fetches multiple credentials because they are stored in the database (stored
<a href="#">CVE-2020-9524</a>	Cross Site scripting vulnerability on Micro Focus Enterprise Server and Enterprise developer, affecting all versio malicious link (reflected XSS).
<a href="#">CVE-2020-9522</a>	Cross Site Scripting (XSS) vulnerability in Micro Focus ArcSight Enterprise Security Manager (ESM) product, Af
<a href="#">CVE-2020-9520</a>	A stored XSS vulnerability was discovered in Micro Focus Vibe, affecting all Vibe version prior to 4.0.7. The vul security context of the target user's browser.
<a href="#">CVE-2020-9485</a>	An issue was found in Apache Airflow versions 1.10.10 and below. A stored XSS vulnerability was discovered i
<a href="#">CVE-2020-9467</a>	Piwigo 2.10.1 has stored XSS via the file parameter in a /ws.php request because of the pwg.images.setInfo f
<a href="#">CVE-2020-9461</a>	Octech Oempro 4.7 through 4.11 allow stored XSS by an authenticated user. The FolderName parameter of th
<a href="#">CVE-2020-9460</a>	Octech Oempro 4.7 through 4.11 allow XSS by an authenticated user. The parameter CampaignName in Camp
<a href="#">CVE-2020-9459</a>	Multiple Stored Cross-site scripting (XSS) vulnerabilities in the Webnus Modern Events Calendar Lite plugin thr
<a href="#">CVE-2020-9447</a>	There is an XSS (cross-site scripting) vulnerability in GwtUpload 1.0.3 in the file upload functionality. Someone other malicious activities like phishing or drive-by hacking.
<a href="#">CVE-2020-9445</a>	Zulip Server before 2.1.3 allows XSS via the modal_link feature in the Markdown functionality.
<a href="#">CVE-2020-9443</a>	Zulip Desktop before 4.0.3 loaded untrusted content in an Electron webview with web security disabled, which
<a href="#">CVE-2020-9440</a>	A cross-site scripting (XSS) vulnerability in the WSC plugin through 5.5.7.5 for CKEditor 4 allows remote attac

# The Problem

CVE - Search Results

https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=SQL%20injection



Common Vulnerabilities and Exposures

CVE List

CNAs

WGs

News & Blog

Board

About

NVD

Go to for:

[CVSS Scores](#)

[CPE Info](#)

Search CVE List

Download CVE

Data Feeds

Request CVE IDs

Update a CVE Entry

TOTAL CVE Entries: 141126

HOME > CVE > SEARCH RESULTS

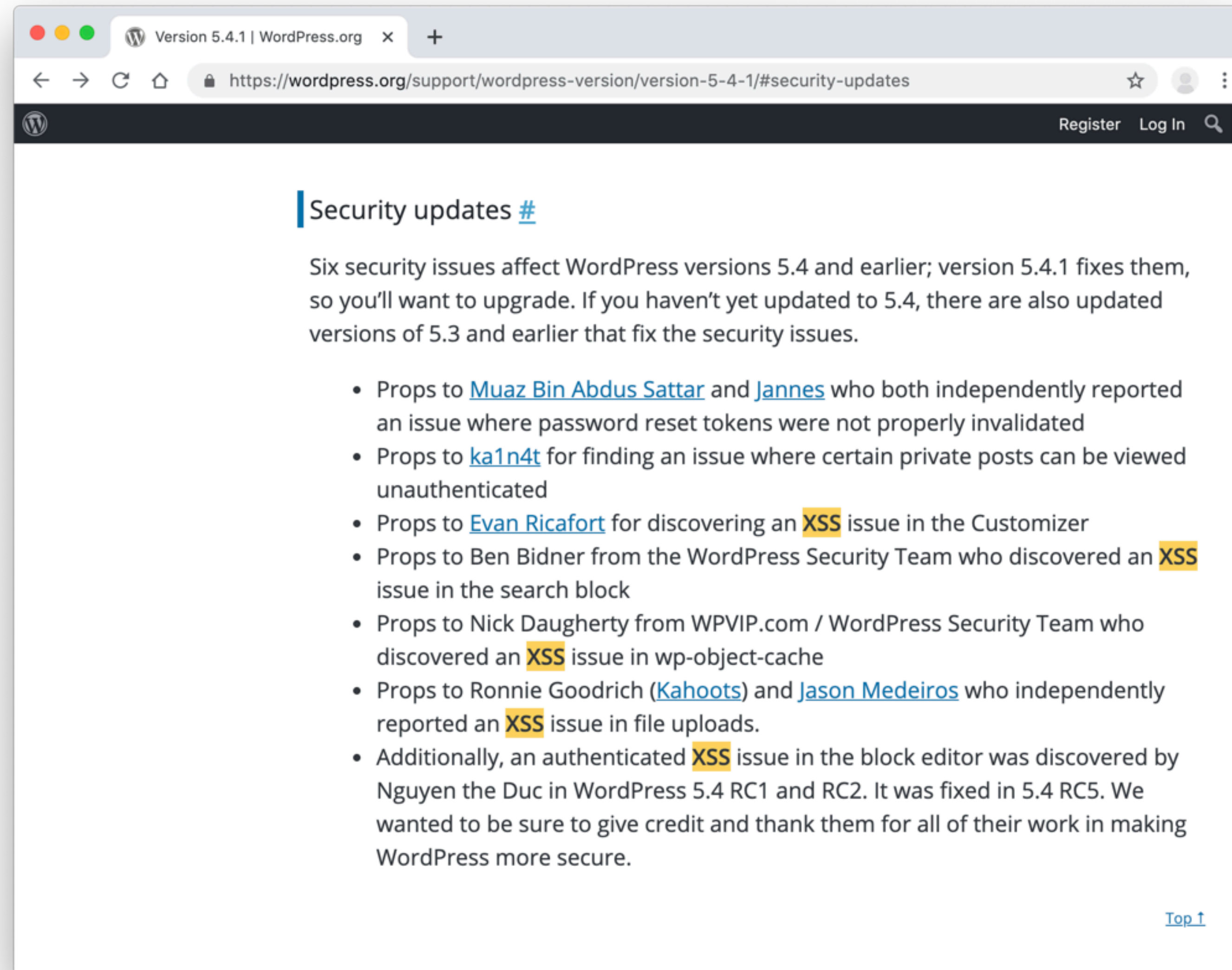
### Search Results

There are 8419 CVE entries that match your search.

Name	Description
<a href="#">CVE-2020-9521</a>	An SQL injection vulnerability was discovered in Micro Focus Service Manager Automation (SMA), affecting versions 2019.08, 2019.05, 2019.02, 2018.08, 2018.05, 2018.02. The vulnerability could allow for the improper neutralization of special elements in SQL commands and may lead to the product being vulnerable to SQL injection.
<a href="#">CVE-2020-9483</a>	<b>**Resolved**</b> When use H2/MySQL/TiDB as Apache SkyWalking storage, the metadata query through GraphQL protocol, there is a SQL injection vulnerability, which allows to access unpexcted data. Apache SkyWalking 6.0.0 to 6.6.0, 7.0.0 H2/MySQL/TiDB storage implementations don't use the appropriate way to set SQL parameters.
<a href="#">CVE-2020-9465</a>	An issue was discovered in EyesOfNetwork eonweb 5.1 through 5.3 before 5.3-3. The eonweb web interface is prone to a SQL injection, allowing an unauthenticated attacker to perform various tasks such as authentication bypass via the user_id field in a cookie.
<a href="#">CVE-2020-9402</a>	Django 1.11 before 1.11.29, 2.2 before 2.2.11, and 3.0 before 3.0.4 allows SQL Injection if untrusted data is used as a tolerance parameter in GIS functions and aggregates on Oracle. By passing a suitably crafted tolerance to GIS functions and aggregates on Oracle, it was possible to break escaping and inject malicious SQL.
<a href="#">CVE-2020-9398</a>	ISPConfig before 3.1.15p3, when the undocumented reverse_proxy_panel_allowed=sites option is manually enabled, allows SQL Injection.
<a href="#">CVE-2020-9340</a>	fauzantrif eLection 2.0 has SQL Injection via the admin/ajax/op_kandidat.php id parameter.
<a href="#">CVE-2020-9318</a>	Red Gate SQL Monitor 9.0.13 through 9.2.14 allows an administrative user to perform a SQL injection attack by configuring the SNMP alert settings in the UI. This is fixed in 9.2.15.
<a href="#">CVE-2020-9269</a>	SOPlanning 1.45 is vulnerable to authenticated SQL Injection that leads to command execution via the



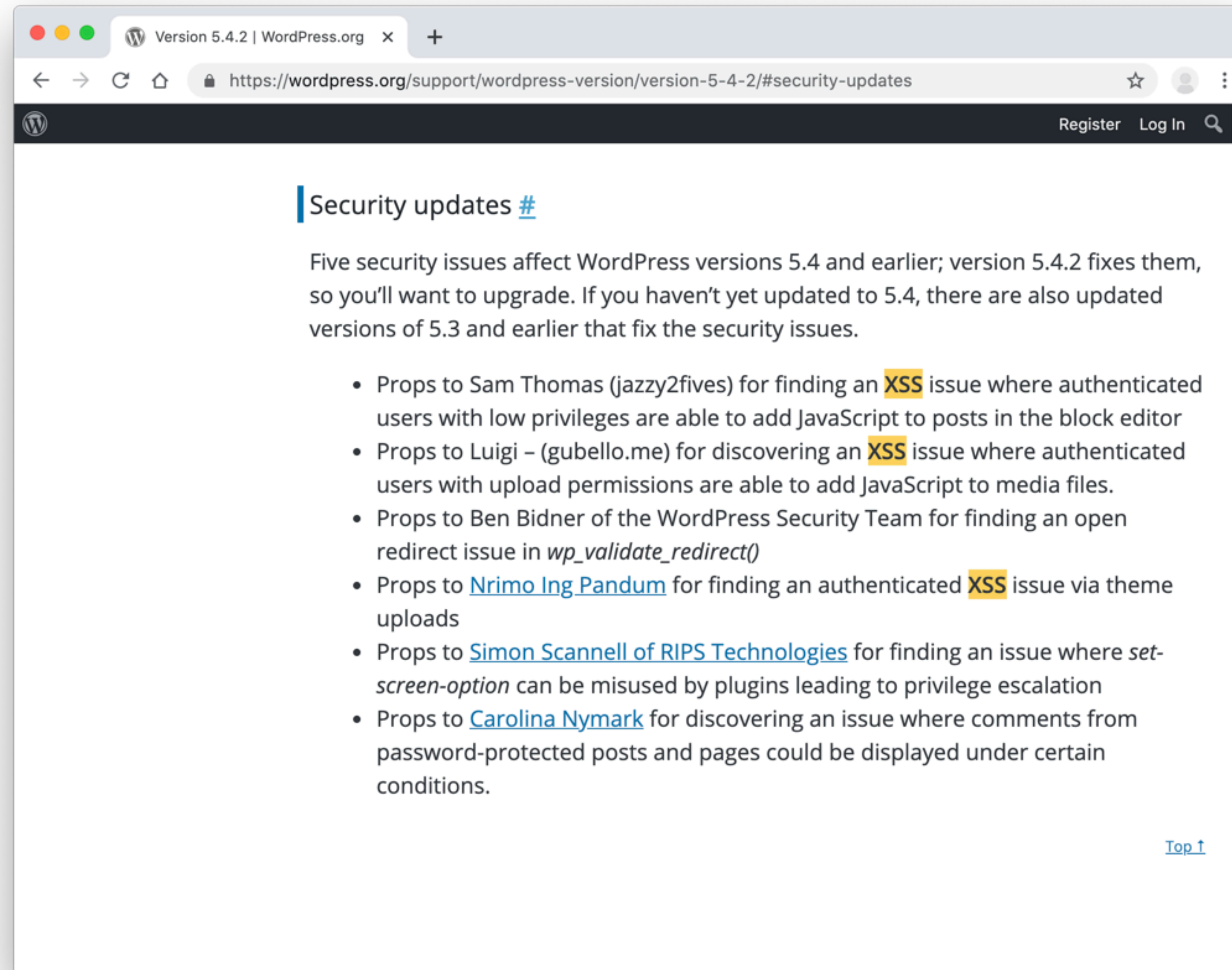
# The Problem



Version 5.4.1

April 29, 2020

# The Problem



Version 5.4.2

June 10, 2020



# Background

Google engineer Christoph Kern, 2015:

<https://www.usenix.org/conference/usenixsecurity15/symposium-program/presentation/kern>

# Background

Google engineer Christoph Kern, 2015:

"Developer education doesn't solve the problem"

# Background

Google engineer Christoph Kern, 2015:

"Bugs are hard to find after the fact"



# Background

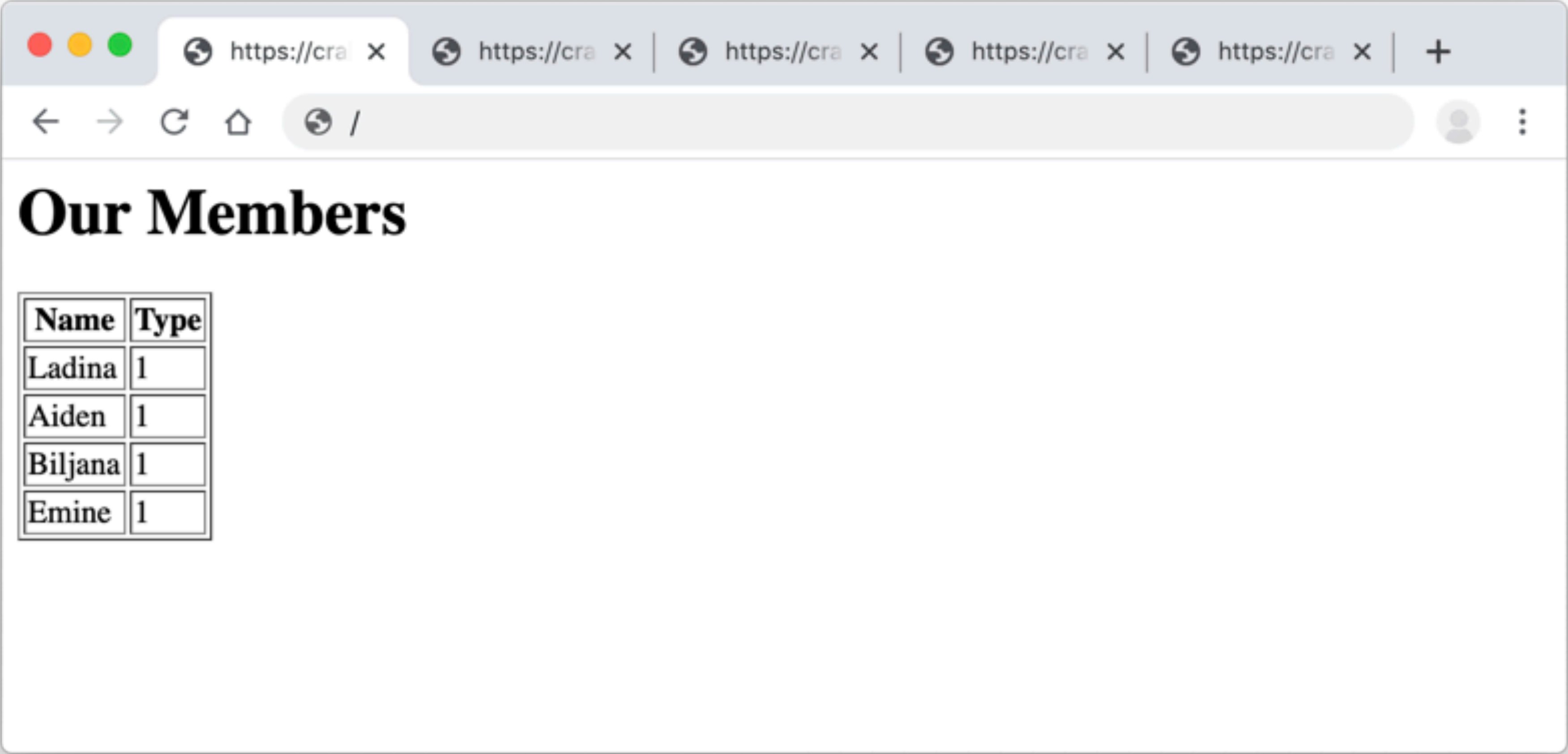
Google engineer Christoph Kern, 2015:

"Bugs are hard to find after the fact"

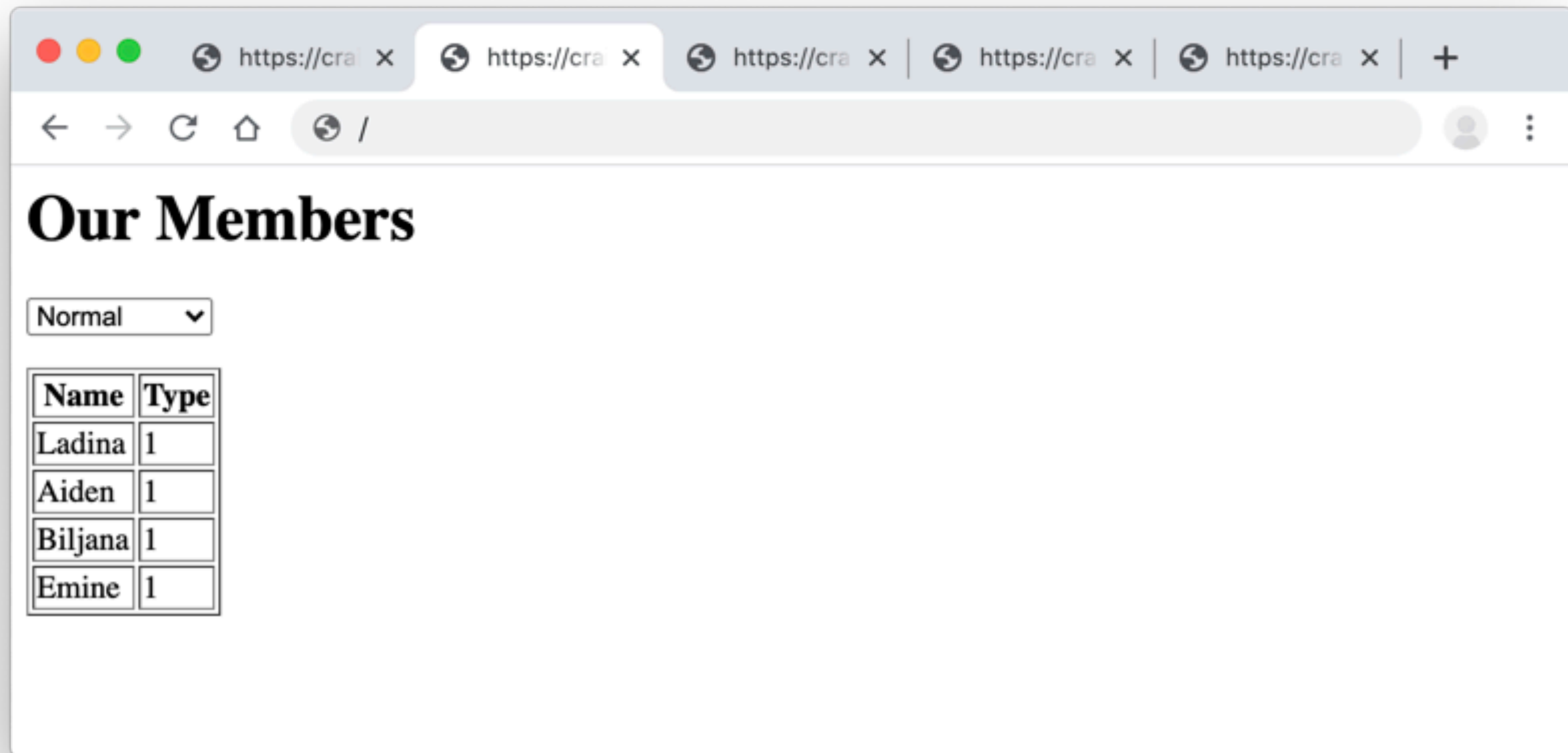


"Manual Testing, Automated Testing, Static Analysis, Human Code Reviews;  
... will find *some* of these bugs..." @04:02

# Examples

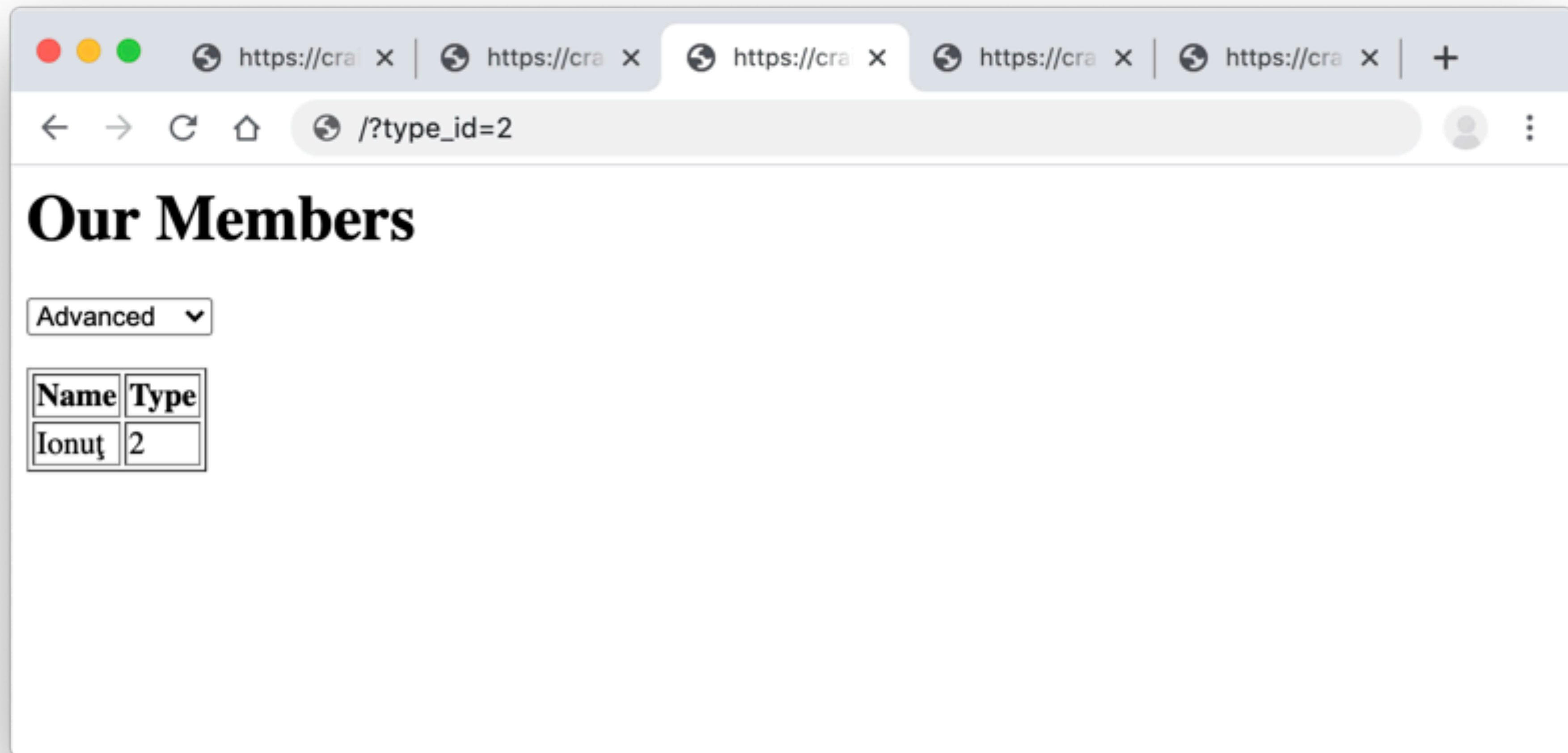


# Examples





# Examples



# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = 1')
```

```
->getQuery()
```

```
->getResult();
```

Doctrine QueryBuilder



# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = ?1')
```

```
->setParameter(1, $_GET['type_id'])
```

```
->getQuery()
```

```
->getResult();
```



# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = 1')
```

```
->getQuery()
```

```
->getResult();
```

# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

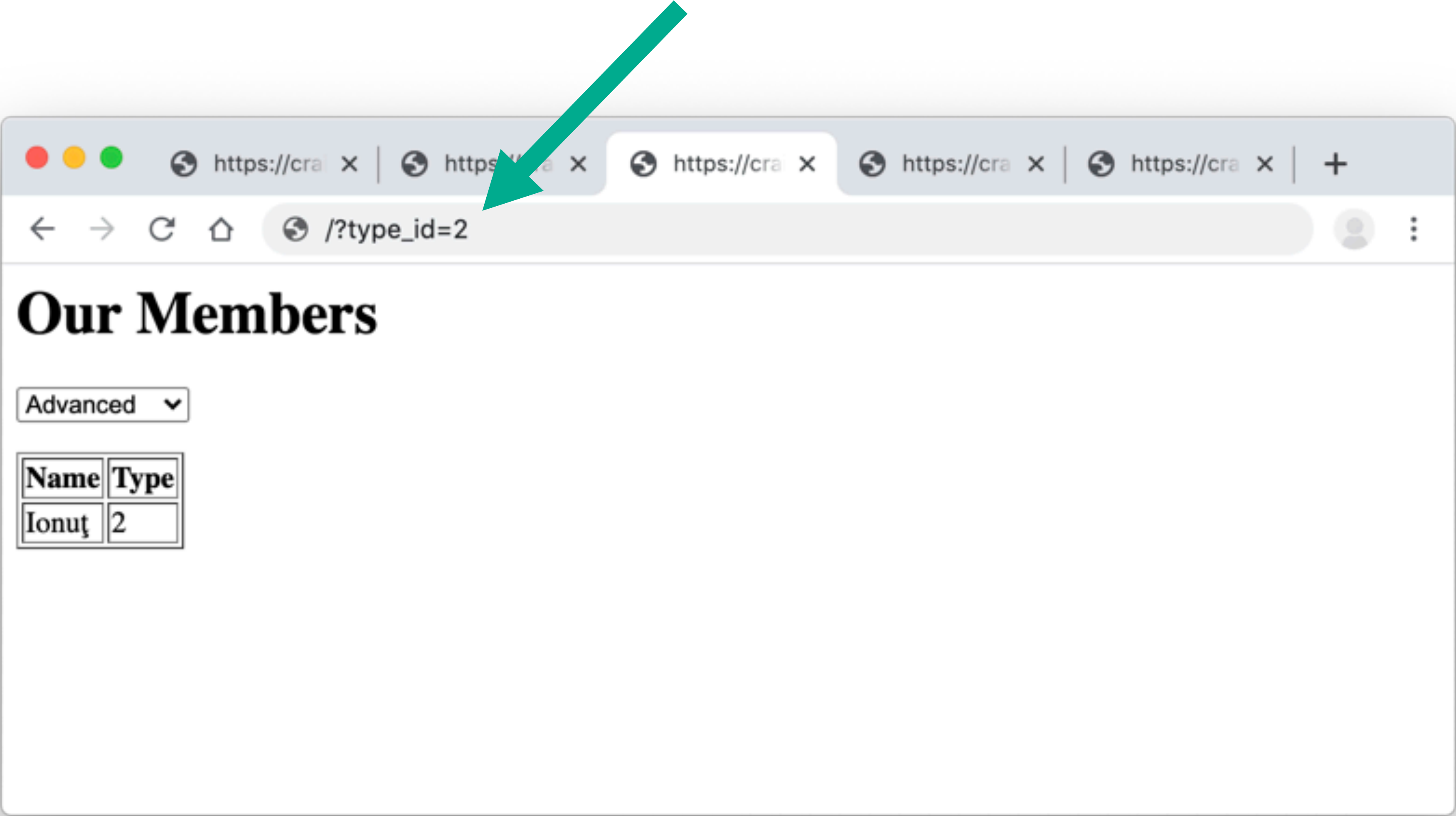
```
->from('User', 'u')
```

```
->where('u.type_id = ' . $_GET['type_id'])
```

```
->getQuery()
```

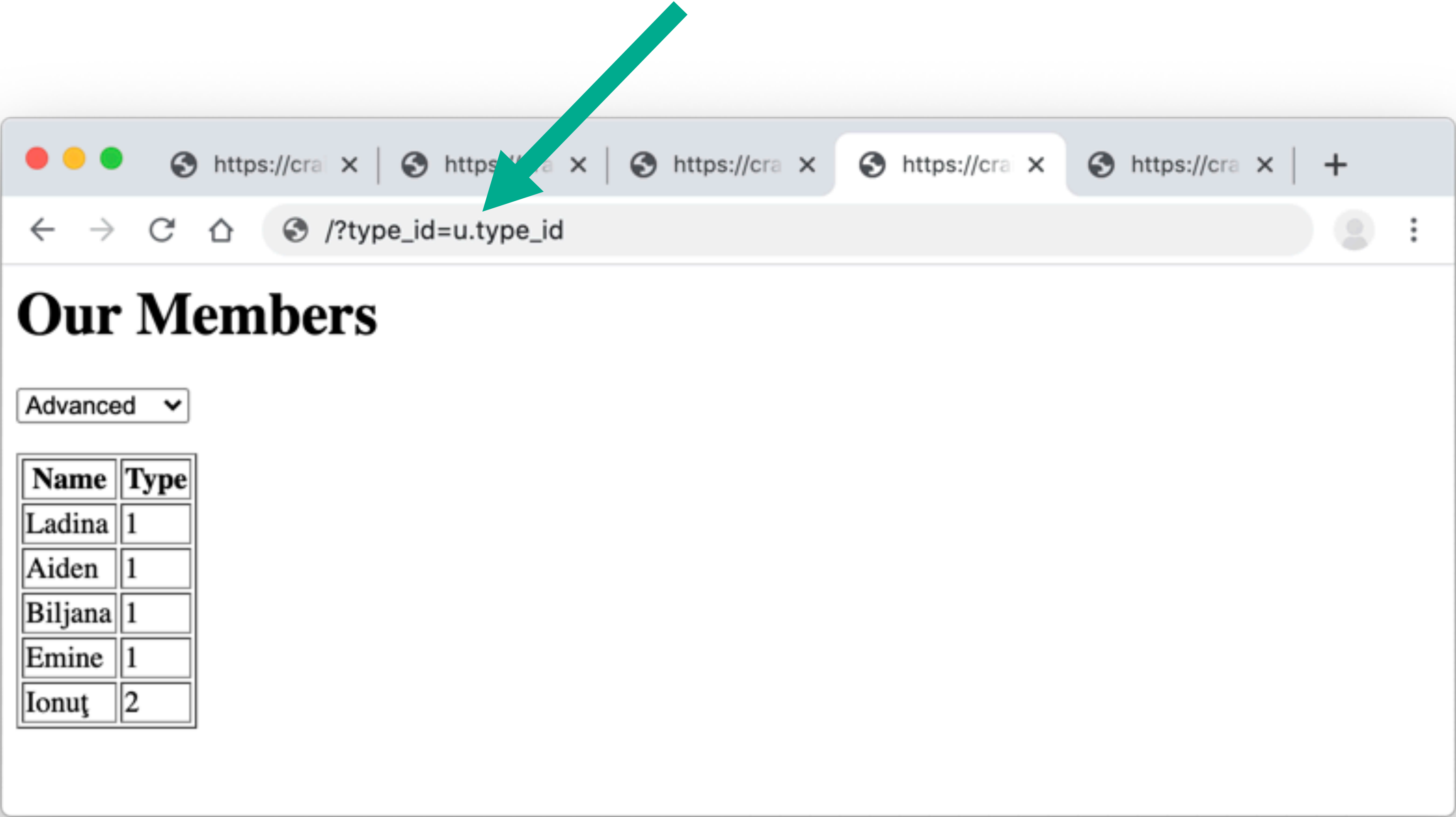
```
->getResult();
```

# Examples





# Examples



# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = 1')
```

```
->getQuery()
```

```
->getResult();
```

Safe String, a Literal



# Examples

```
$users = $queryBuilder
```

```
->select('u')
```

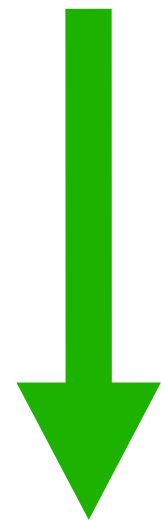
```
->from('User', 'u')
```

```
->where('u.type_id = ' . $_GET['type_id'])
```

```
->getQuery()
```

```
->getResult();
```

Safe String



Unsafe String



# Examples

```
$users = $QueryBuilder
```

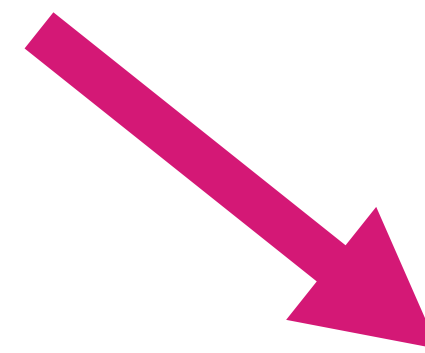
```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = ' . $_GET['type_id'])
```

```
->getQuery()
```

```
->getResult();
```

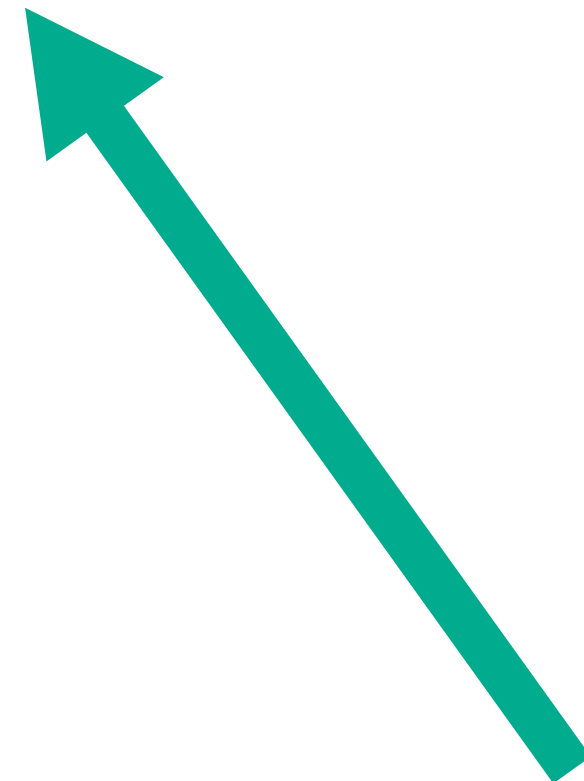


'WHERE u.type\_id = u.type\_id'

# Examples

```
$sql = 'SELECT u FROM User u WHERE u.type_id = ' . $_GET['type_id'];
```

```
$query = $entityManager->createQuery($sql);
```

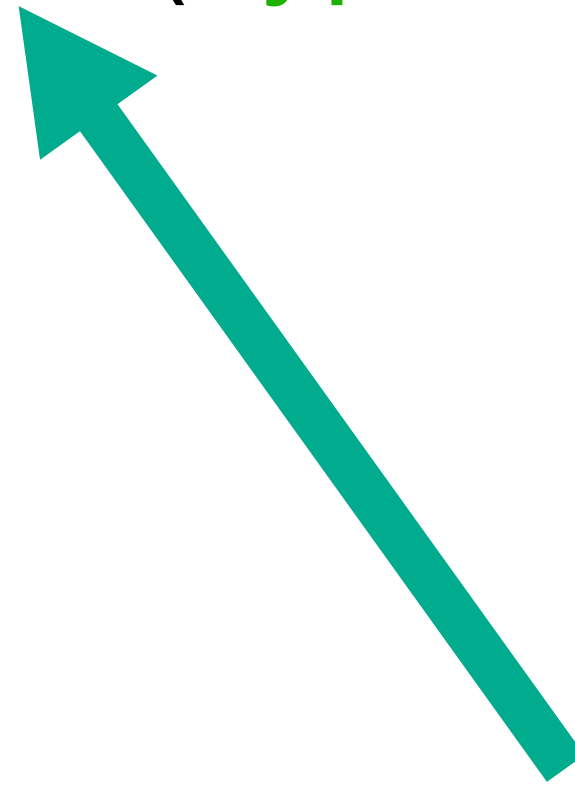


Doctrine - CreateQuery, DQL



# Examples

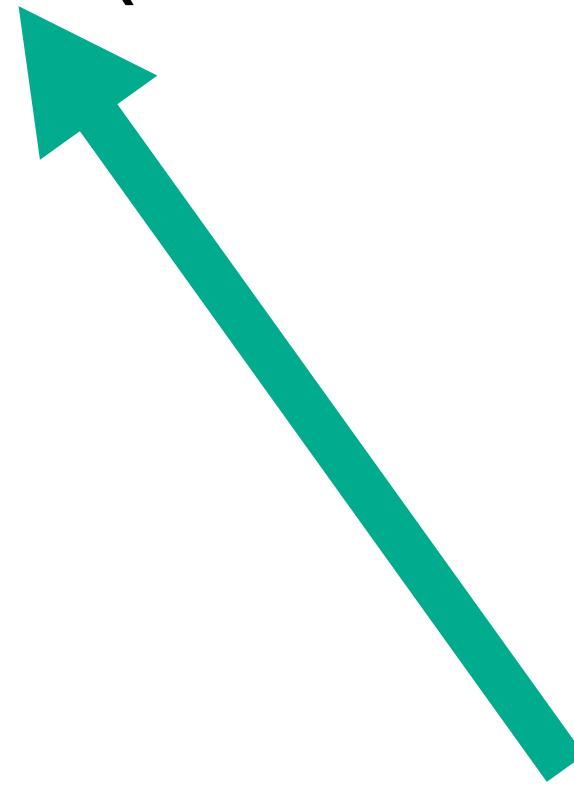
```
$users = UserQuery::create()->where('type_id = ' . $_GET['type_id'])->find();
```



Propel ORM - Where

# Examples

```
$result = R::find('user', 'type_id = ' . $_GET['type_id']);
```



RedBeanPHP - Find

# Examples

Safe String, a Literal



```
$template = $twig->createTemplate('<p>Hello {{ name }}</p>');
```

```
echo $template->render(['name' => $_GET['name']));
```

# Examples

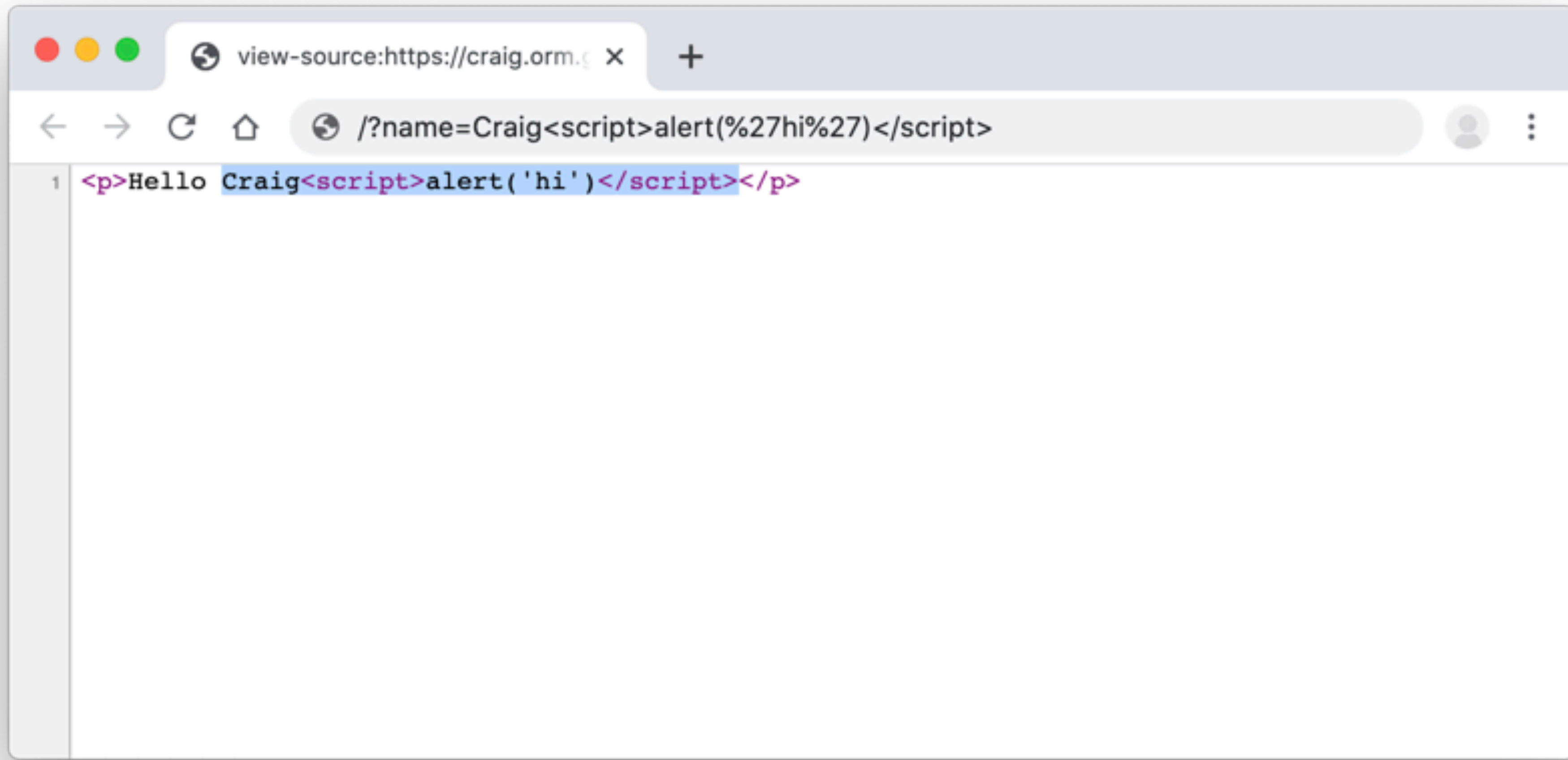
Unsafe String



```
$template = $twig->createTemplate('<p>Hello ' . $_GET['name'] . '</p>');
```

```
echo $template->render();
```

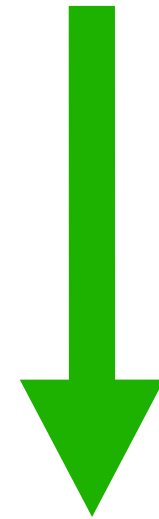
# Examples





# Examples

Safe String, a Literal



```
$output = shell_exec('ls /');
```

# Examples

Unsafe String



```
$output = shell_exec('ls ' . $_GET['path']);
```

# Taint Checking

<https://pecl.php.net/package/taint>

# Taint Checking

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . 'Name';
```

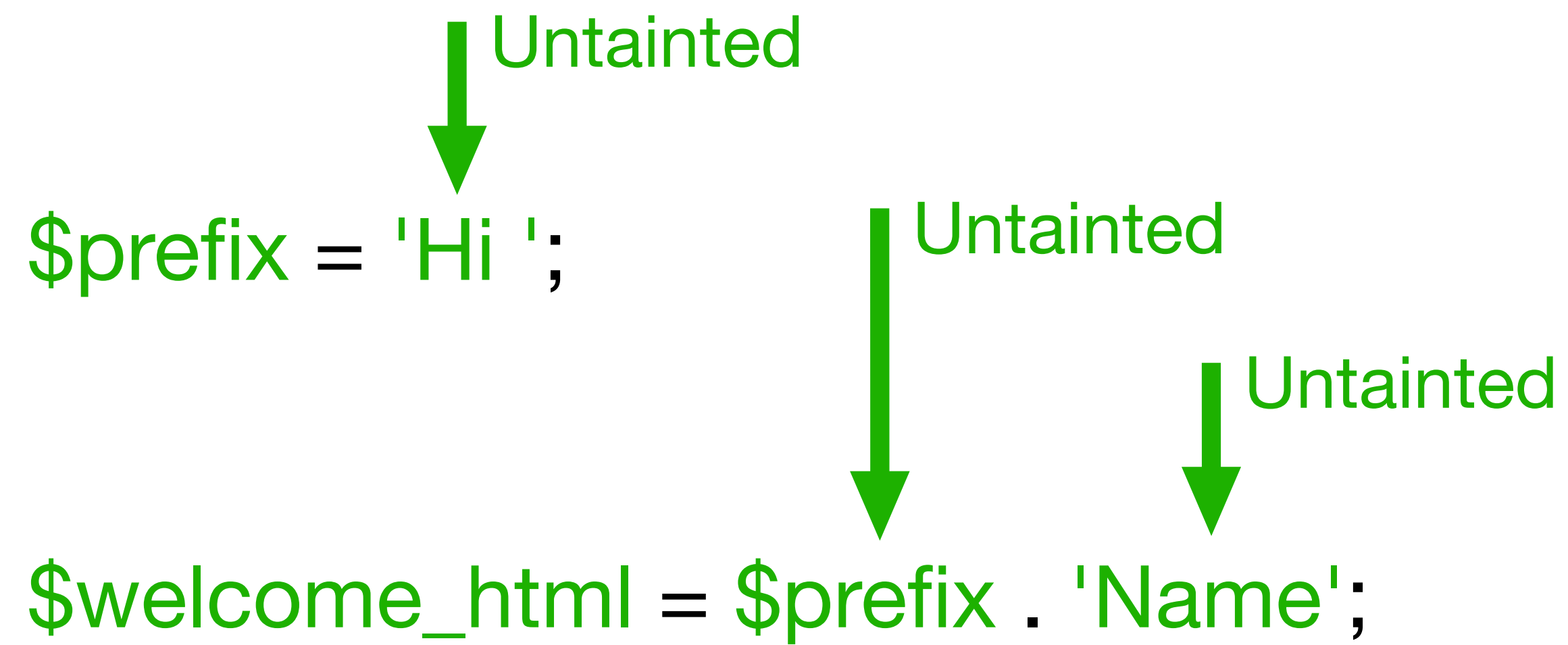
# Taint Checking

↓ Untainted  
\$prefix = 'Hi ';

\$welcome\_html = \$prefix . 'Name';

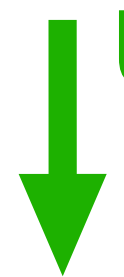


# Taint Checking



# Taint Checking

```
$prefix = 'Hi ';
```



Untainted, Safe

```
$welcome_html = $prefix . 'Name';
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . $name;
```

# Taint Checking



Tainted

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . $name;
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . $name;
```

# Taint Checking

```
$name = $_GET['name'];
```

↓ Untainted

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . $name;
```

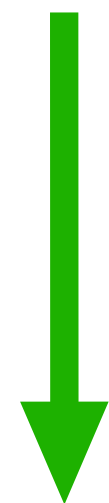


# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

Untainted



Tainted



```
$welcome_html = $prefix . $name;
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

↓ Tainted, not safe... complain when printed.

```
$welcome_html = $prefix . $name;
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

```
$welcome_html = $prefix . htmlentities($name);
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```

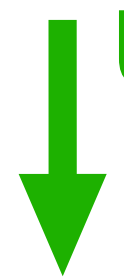


```
$welcome_html = $prefix . htmlentities($name);
```

# Taint Checking

```
$name = $_GET['name'];
```

```
$prefix = 'Hi ';
```



Untainted, should be safe.

```
$welcome_html = $prefix . htmlentities($name);
```

# Taint Checking

```
$url = $_GET['url'];
```

```
$name = $_GET['name'];
```

```
$link_html = '<a href="' . htmlentities($url) . '"' . '>' . htmlentities($name) . '</a>';
```

# Taint Checking



```
$url = $_GET['url'];
```



```
$name = $_GET['name'];
```

```
$link_html = '<a href="' . htmlentities($url) . '">' . htmlentities($name) . '</a>';
```



# Taint Checking

```
$url = $_GET['url'];
```

```
$name = $_GET['name'];
```

```
$link_html = '<a href="" . htmlentities($url) . ">' . htmlentities($name) . '</a>';
```

The diagram illustrates taint checking. Two magenta arrows point from the variables `$url` and `$name` in the code above to their respective arguments in the `htmlentities` function call in the code below. Each arrow is labeled "Tainted" in magenta text, indicating that the data from these inputs is tainted and flows into the output string.

# Taint Checking

```
$url = $_GET['url'];
```

```
$name = $_GET['name'];
```



```
$link_html = '<a href="' . htmlentities($url) . '">' . htmlentities($name) . '</a>';
```

# Taint Checking

```
$url = $_GET['url'];
```

```
$name = $_GET['name'];
```

↓ Untainted, surely this is safe?

```
$link_html = '<a href="' . htmlentities($url) . '"> . htmlentities($name) . '</a>';
```

# Taint Checking

`$url = $_GET['url'];` ← `'javascript:alert("hi")'`

`$name = $_GET['name'];`

`$link_html = '<a href="' . htmlentities($url) . '>' . htmlentities($name) . '</a>';`

# Taint Checking

`$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);`

↓ Tainted

# Taint Checking

↓ Untainted      ↓ Un-Taint

```
$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);
```

# Taint Checking

↓ Untainted, surely this is safe?

```
$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);
```

# Taint Checking

```
$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);
```



Missing quotes



# Taint Checking

```
$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);
```



A diagram illustrating a taint flow. A magenta arrow points from the string `"id"` to the `$_GET['id']` expression in the code snippet above. This indicates that the value of the `id` parameter from the HTTP GET request is being used in a SQL query, which is a potential security vulnerability if the input is not properly sanitized.

# Taint Checking

`$sql .= 'WHERE id = ' . mysqli_real_escape_string($link, $_GET['id']);`



`$sql .= 'WHERE id = id';`



# Taint Checking

↓ Untainted

↓ Tainted

↓ Untainted

```
$img_html = '<img src=' . htmlentities($_GET['url']) . '>';
```

# Taint Checking



```
$img_html = '<img src=' . htmlentities($_GET['url']) . '>';
```

# Taint Checking

↓ Untainted, surely this is safe?

```
$img_html = '<img src=' . htmlentities($_GET['url']) . '>';
```

# Taint Checking

```
$img_html = '<img src=' . htmlentities($_GET['url']) . '>';
```




Missing quotes

# Taint Checking

`$img_html = '<img src=' . htmlentities($_GET['url']) . '>';`

`"/ onerror=evil-js"`

A pink arrow points from the string `"/ onerror=evil-js"` to the `$_GET['url']` variable in the code above.A pink arrow points from the `<img src=` part of the code above to the `<img src=` part of the code below.

`$img_html = '<img src=/ onerror=evil-js>';`

# Taint Checking

How about character encoding issues?



# Taint Checking

The PDO `Quote` function clearly says it's *only* "theoretically safe".

<https://www.php.net/pdo.quote>

# Static Analysis

# Static Analysis

Injection problems are rarely solved by Static Analysis.

# Static Analysis

Injection problems are rarely solved by Static Analysis.



Still do use this.

# Static Analysis

Injection problems are rarely solved by Static Analysis.



But it's hard to follow *every* variable from Source to Sink.

# Static Analysis

Injection problems are rarely solved by Static Analysis.



And don't expect every programmer to use.

# Static Analysis

```
$result = $mysqli->query('SELECT * FROM user WHERE id = ' . $_GET['id']);
```

```
if ($result instanceof mysqli_result) {
```

```
print_r($result->fetch_all());
```

}

# PHPStan

A terminal window titled "Terminal" with standard macOS window controls (red, yellow, green buttons). The prompt is "craig\$". The command executed is "./vendor/bin/phpstan analyse --level max public/". The progress bar shows "1/1 [████████████████████] 100%". A large green box contains the message "[OK] No errors". The final prompt is "craig\$ " followed by a cursor.

```
craig$ ./vendor/bin/phpstan analyse --level max public/  
1/1 [████████████████████] 100%  
  
[OK] No errors  
  
craig$
```

# Static Analysis

```
$mysqli = new mysqli('localhost', 'test', '???', 'test');
```

```
$id = (string) $_GET['id']; ← Avoid MixedAssignment
```

```
$result = $mysqli->query('SELECT * FROM user WHERE id = ' . $id);
```

```
if ($result instanceof mysqli_result) {
```

```
    print_r($result->fetch_all());
```

```
}
```

Psalm



```
Terminal
craig$ ./vendor/bin/psalm
Scanning files...
Analyzing files...

-----
No errors found!
-----

Checks took 0.18 seconds and used 61.097MB of memory
Psalm was able to infer types for 100% of the codebase
craig$
```



# Static Analysis

```
$mysqli = new mysqli('localhost', 'test', '???' , 'test');
```

```
$id = (string) $_GET['id']; ← Avoid MixedAssignment
```

```
$result = $mysqli->query('SELECT * FROM user WHERE id = ' . $id);
```

```
if ($result instanceof mysqli_result) {
```

```
    print_r($result->fetch_all());
```

```
}
```

I'm not using "mysqli\_query()"

<https://github.com/vimeo/psalm/issues/4155>

Psalm,  
Taint Analysis



```
Terminal
craig$ ./vendor/bin/psalm --taint-analysis
Scanning files...
Analyzing files...

-----
No errors found!
-----

Checks took 0.18 seconds and used 61.202MB of memory
Psalm was able to infer types for 100% of the codebase
craig$
```

# Summary

Do not mix safe strings (literals), with anything that may be attacker controlled.

# Summary

We need a way to ensure this does not happen.

```
is_literal('This is a Literal');
```

```
is_literal('This is a Literal');
```



true

```
$a = 'This is a Literal';
```

```
is_literal($a);
```



true

```
$a = 'This is a Literal';
```

```
is_literal($a . 'And this');
```



true

```
$a = 'This is a Literal';
```

```
is_literal($a . $_GET['name']);
```



false



```
$a = 'This is a Literal';
```

```
is_literal($a . htmlentities($_GET['name']));
```



Still false

```
$a = 'This is a Literal';
```

```
is_literal($a . htmlentities($_GET['name']));
```



Still false

is\_literal() is not the same as Taint Checking.

```
$a = 'This is a Literal';
```

```
is_literal($a . strtoupper('abc'));
```



Sorry, this is false - 'ABC' is no longer the string defined in the PHP script.

```
$sql .= ' AND id = ' . mysqli_real_escape_string($db, $_GET['id']) . '';
```

```
is_literal($sql);
```



Still false

```
$sql .= ' AND id = ?';
```

```
$parameters[] = $_GET['id'];
```

```
is_literal($sql);
```



true

```
$users = $queryBuilder
```

```
->select('u')
```

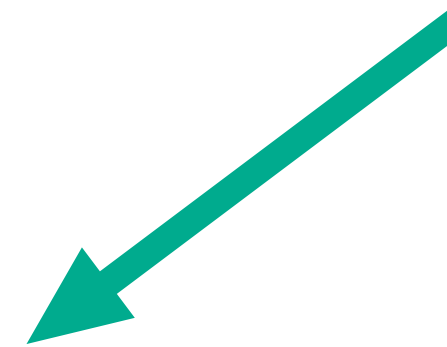
```
->from('User', 'u')
```

```
->where('u.type_id = 1')
```

```
->getQuery()
```

```
->getResult();
```

Doctrine



\$users = \$QueryBuilder

->select('u')

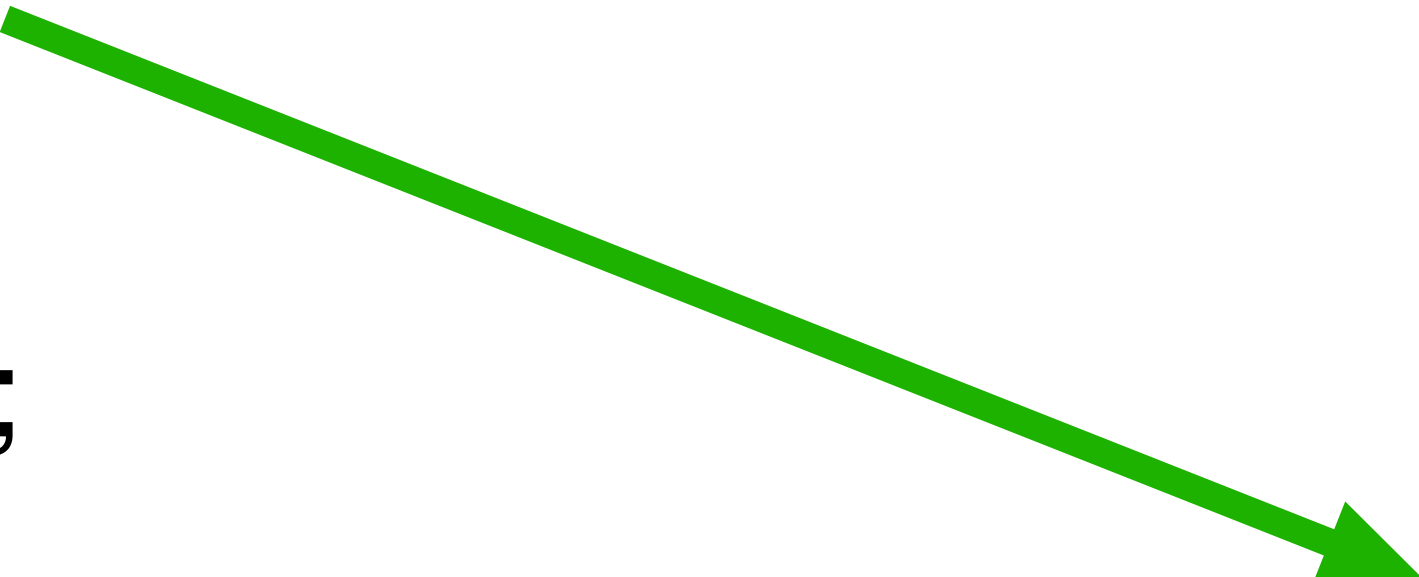
->from('User', 'u')

->where('u.type\_id = 1')

->getQuery()

->getResult();

Is a Safe Literal? →



```
public function where($predicates)
{
    if (!is_literal($predicates)) {
        throw new Exception('Can only accept a literal');
    }
    ...
}
```

```
$users = $queryBuilder
```

```
->select('u')
```

```
->from('User', 'u')
```

```
->where('u.type_id = ' . $_GET['type_id'])
```

```
->getQuery()
```

```
->getResult();
```



Not a Safe Literal →

```
public function where($predicates)
{
    if (!is_literal($predicates)) {
        throw new Exception('Can only accept a literal');
    }
    ...
}
```



```
$template = $twig->createTemplate('<p>Hello {{ name }}</p>');
```

```
echo $template->render(['name' => $_GET['name']]);
```



```
$template = $twig->createTemplate('<p>Hello {{ name }}</p>');
```

```
echo $template->render(['name' => $_GET['name']])
```



Is a Safe Literal? →

```
public function createTemplate(string $template, string $name = null): TemplateWrapper
{
    if (!is_literal($template)) {
        throw new Exception('Can only accept a literal');
    }
    ...
}
```

```
$template = $twig->createTemplate('<p>Hello ' . $_GET['name'] . '</p>');
```

```
echo $template->render()
```



Not a Safe Literal



```
public function createTemplate(string $template, string $name = null): TemplateWrapper
{
    if (!is_literal($template)) {
        throw new Exception('Can only accept a literal');
    }
    ...
}
```

Backwards compatibility



```
if (function_exists('is_literal') && !is_literal($a)) {  
    trigger_error('Can only accept a literal', E_USER_NOTICE);  
}
```



Notices might be safer for legacy projects.

Backwards compatibility



```
if (!function_exists('is_literal')) {  
    function is_literal($variable) {  
        return true;  
    }  
}
```

What about Table and Field names?



```
$sql .= 'ORDER BY ' . $field_name;
```

```
$fields = [  
    'name',  
    'created',  
    'admin',  
];
```

```
$field_id = array_search($_GET['sort'] ?? 'created', $fields);
```

```
$sql = ' ORDER BY ' . $fields[$field_id];
```

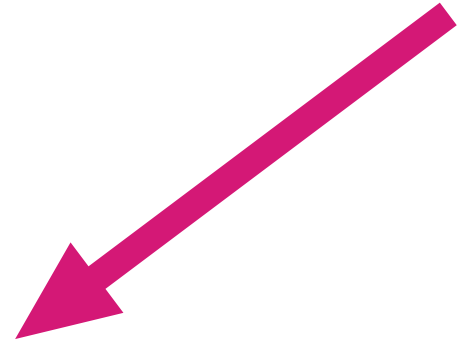
What about WHERE IN?



```
$sql .= 'WHERE id IN (1, 2, 3)';
```



From an unknown source



```
$ids = [1, 2, 3];
```

```
$sql .= 'WHERE id IN (' . implode(', ', $ids) . ')';
```

```
$ids = [1, 2, 3];
```

```
$sql .= 'WHERE id IN (' . implode(',', array_fill(0, count($ids), '?')) . ')';
```

```
$ids = [1, 2, 3];
```

```
$sql .= 'WHERE id IN (' . implode(',', array_fill(0, count($ids), '?')) . ')';
```



```
$sql .= 'WHERE id IN (?, ?, ?)';
```

# Other Implementations

# Other Implementations

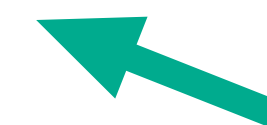
Google have a similar solution in Go:

HTML Injection / XSS



<https://github.com/google/go-safeweb/tree/master/safehttp>

<https://blogtitle.github.io/go-safe-html/>



By Roberto Clapis

# Other Implementations

Google have a similar solution in Go:

SQL Injection



<https://github.com/google/go-safeweb/tree/master/safesql>

A Query Builder, with an Append() method that only accepts **compile-time constant strings** (literal or const).

# Other Implementations

JavaScript may get a similar solution:

<https://github.com/tc39/proposal-array-is-template-object>

"Distinguishing strings from a trusted developer,  
from strings that may be attacker controlled"

# Concerns



# Concerns

- Performance

# Concerns

- Performance

Might not be able to justify string concatenation.

# Concerns

- Performance

Might not be able to justify string concatenation.

But could introduce a new function to combine an array of literals:

# Concerns

- Performance

Might not be able to justify string concatenation.

But could introduce a new function to combine an array of literals:

```
$sql = ['SELECT u.name FROM user AS u'];
```

```
if ($id) {  
    $sql[] = 'WHERE id = ?';  
    $parameters[] = $id;  
}
```

```
$sql = literal_implode(' ', $sql);
```

# Concerns

- Performance
- Can `array_fill()` + `implode()` pass through the "is\_literal" flag? (WHERE IN)
- Name it something else?
- Variables, like a table prefix, being stored in an INI / JSON / YAML file?

# The Future

- Internal PHP functions, like `mysqli_query()`, `preg_match()`, `exec()`, etc; could all be setup to only accept safe literals.
- When printing strings (i.e. sending HTML to the browser), this could be blocked, with an exception for your trusted HTML templating engine. Maybe PHP could only accept a specific value object that has a `__toString()` method?

# Thank You

[https://wiki.php.net/rfc/is\\_literal](https://wiki.php.net/rfc/is_literal)

<https://github.com/craigfrancis/php-is-literal-rfc>