

ARCHIVING AND VISUAL REGRESSION USING DRUPAL 8

Texas Camp – June 2018

David Stinemetze

Twitter: @davidstinemetze

Drupal.org/GitHub: @WidgetsBurritos



Our Problems

- ▶ What was the content of some particular page on June 19?
- ▶ What did that page look like on October 17?
- ▶ When did some bug get introduced?
- ▶ How will these new code/config changes affect other parts of our website?
- ▶ How will the migration from Drupal 7 to Drupal 8 impact our website visually?

Our Solution

- ▶ **Web Page Archive Module**

- https://www.drupal.org/project/issues/web_page_archive

- ▶ **Project Maintainers**

- David Stinemetze - @WidgetsBurritos
- David Porter - @bighappyface
- Paul Maddern - @pobster

Module Overview

- ▶ The Web Page Archive module allows you to use Drupal to perform and compare periodic snapshots on local and remote websites based on a list of URLs or XML sitemaps.
- ▶ This means snapshots can be captured on any site.
 - Not just your current Drupal instance.
 - Not just Drupal-based websites.
 - Not just sites you host.
 - Not just sites you own (but be nice!)
- ▶ **Requirements**
 - Drupal 8.3+
 - PHP 7.0+
 - Must be installed via composer
 - Some features have additional dependencies

Why Build This In Drupal?

- ▶ Learn Drupal 8
- ▶ Ease-of-use for non technical users
- ▶ Drupal 8's plugin ecosystem makes it easy to support additional use cases
- ▶ Give back to the Drupal community

Capture Utilities

- ▶ Drupal plugins responsible for performing different types of captures
- ▶ Use annotation-based plugin discovery
- ▶ Configurable per instance

Screenshot Capture Utility

- ▶ **What it does:**

- Captures screenshot of specified URL

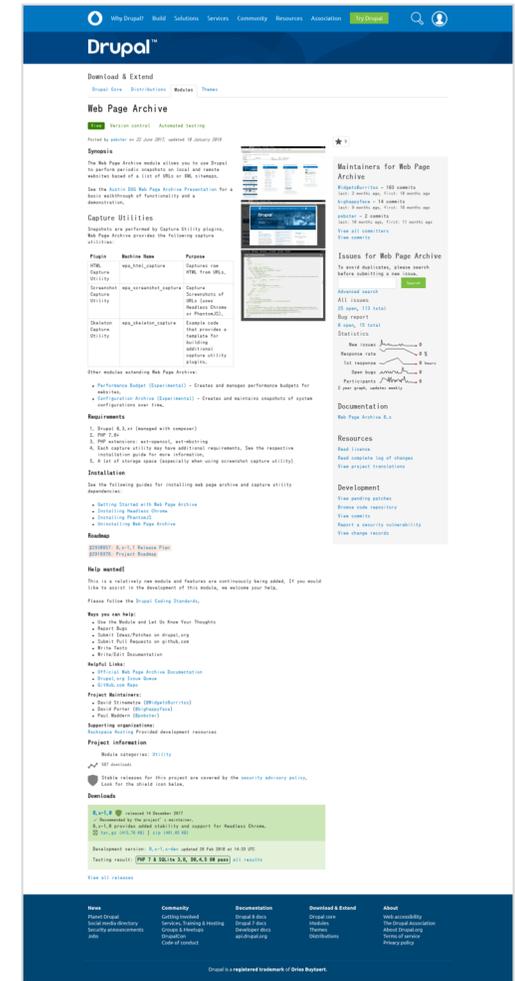
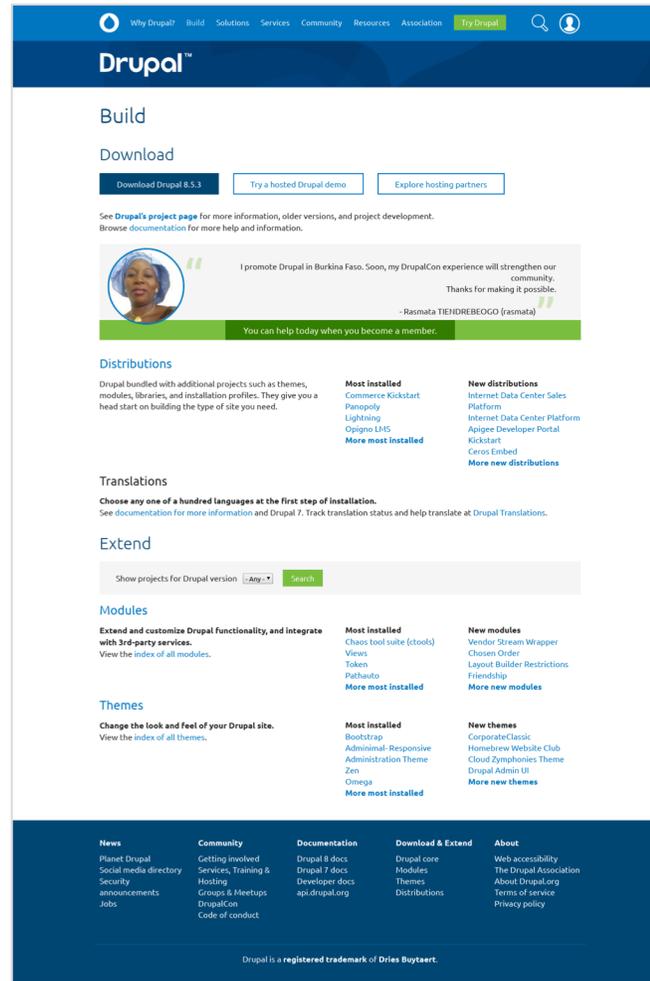
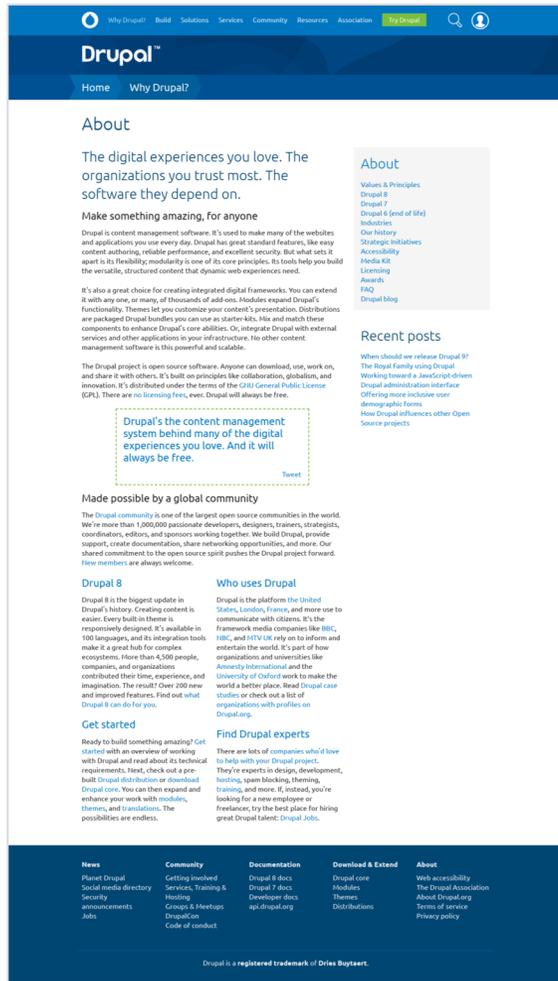
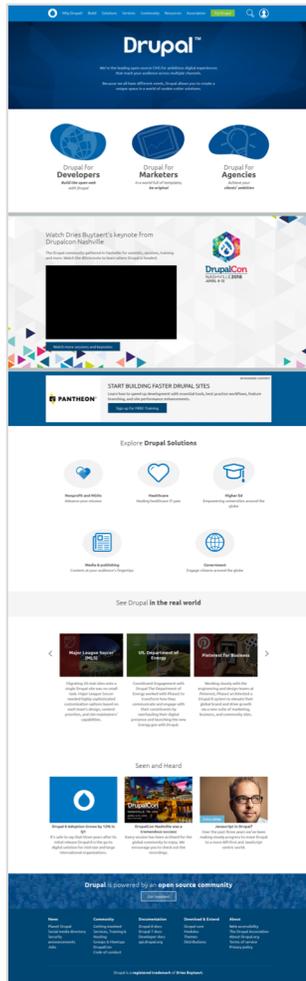
- ▶ **How it works:**

- Uses either Headless Chrome or PhantomJS to capture screenshots
- Can capture specific widths (e.g. Mobile vs Desktop)

- ▶ **Where it lives:**

- *wpa_screenshot_capture* submodule

Screenshot Examples



HTML Capture Utility

- ▶ **What it does:**

- Captures raw HTML of specified URL.

- ▶ **How it works:**

- Uses Guzzle, which is part of Drupal Core, to capture raw HTML.

- ▶ **Where it lives:**

- *wpa_html_capture* submodule

HTML Example

https://www.drupal.org/about



```
<a href="/about" class="active">About</a> </li>

<div class="pane-content">
  <ul class="menu"><li class="first leaf"><a href="/about/values-and-principles" title="">Values & Principles</a></li>
<li class="leaf"><a href="/8" title="">Drupal 8</a></li>
<li class="leaf"><a href="/about/drupal-7" title="">Drupal 7</a></li>
<li class="leaf"><a href="/about/drupal6-eol">Drupal 6 (end of life)</a></li>
<li class="leaf"><a href="/industries" title="">Industries</a></li>
<li class="leaf"><a href="/about/history">Our history</a></li>
<li class="leaf"><a href="/about/strategic-initiatives" title="">Strategic Initiatives</a></li>
<li class="leaf"><a href="/about/features/accessibility">Accessibility</a></li>
<li class="leaf"><a href="/about/media-kit" title="">Media Kit</a></li>
<li class="leaf"><a href="/about/licensing">Licensing</a></li>
<li class="leaf"><a href="/about/awards">Awards</a></li>
<li class="leaf"><a href="/about/faq">FAQ</a></li>
<li class="last leaf"><a href="/blog">Drupal blog</a></li>
</ul> </div>

</div>
<div class="panel-separator"></div><div class="panel-pane pane-views-panes pane-blog-panel-pane-2" >

  <h2 class="pane-title">
Recent posts </h2>
```

Skeleton Capture Utility

- ▶ **What it does:**

- Absolutely nothing.

- ▶ **Why does it exist:**

- Prototype for how to build other capture utilities.
- Can be used to test surrounding behavior without having to perform actual captures.

- ▶ **Where it lives:**

- `wpa_skeleton_capture` submodule

Web Page Test Capture Utility

▶ **What it does:**

- Captures web performance test results.

▶ **How it works:**

- Starts and captures performance test results from webpagetest.org.
- While this module can presently capture data, we don't have a way of actually showing this data visually yet.

▶ **Where it lives:**

- (Experimental) Performance Budget module:
https://www.drupal.org/project/performance_budget

Other PROPOSED CAPTURE UTILITIES

- ▶ Configuration Archive Capture Utility
- ▶ Rendered DOM Capture Utility
- ▶ Google Analytics Threshold Monitor
- ▶ Security Scan Utility

Capture Entities

- ▶ **web_page_archive**
 - Configuration entities containing settings for distinct jobs.
 - Identify which capture utilities should be used.
- ▶ **web_page_archive_run**
 - Content entities containing the actual capture results.
 - Each run is stored as a separate revision.
- ▶ Each *web_page_archive* entity is mapped to a *web_page_archive_run* entity.

Configuring A Capture Job

Back to site Manage Shortcuts wpa_admin Devel

Label *
Drupal.org Machine name: drupal_org
Label for the Web page archive entity.

Run capture job automatically.
Yes

Crontab schedule (relative to PHP's default timezone)
0 5 * * *
Crontab format (see https://crontab.guru/)
Next run: 2018-05-29 @ 5:00am CDT

Timeout (ms)
500
Amount of time to wait between captures, in milliseconds.

Capture Type
URL

Honor robots.txt restrictions.
Yes
If checked, capture utility will respect robots.txt crawling rules.

Browser user agent
WPA
Specify the browser user agent. e.g. "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_5) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/59.0.3071.115 Safari/537.36"

URLs to Capture *
https://www.drupal.org/
https://www.drupal.org/about
https://www.drupal.org/download
https://www.drupal.org/community
https://www.drupal.org/project/web_page_archive
A list of urls to capture.

CAPTURE UTILITY	OPERATIONS
✚ Screenshot capture utility	Edit
✚ Select a new capture utility	Add

Update archive Delete

Back to site Manage Shortcuts wpa_admin Devel

Edit Screenshot capture utility capture utility

Home » Administration » Configuration » System » Web Page Archive » Drupal.org

Browser *
Headless Chrome
Specify the browser you would like to use to perform captures with.

Capture width (in pixels) *
1280
Specify the width you would like to capture.

Image type *
png

Delay (ms) *
500
How long to delay before capturing a screenshot. This is helpful if you need to wait for javascript or resources to load.

Update capture utility Cancel

Viewing Capture Results

The screenshot shows the Drupal.org website's Web Page Archive interface. The breadcrumb trail is "Home » Administration » Configuration » System » Web Page Archive". It displays "Displaying 1 - 5 of 5" items. The interface includes a search filter set to "Contains", a "Sort by" dropdown set to "Capture URL", and an "Items per page" dropdown set to "20". Below the controls, four thumbnails of captured pages are visible, each with its URL: <https://www.drupal.org/>, <https://www.drupal.org/about>, <https://www.drupal.org/community>, and <https://www.drupal.org/download>.

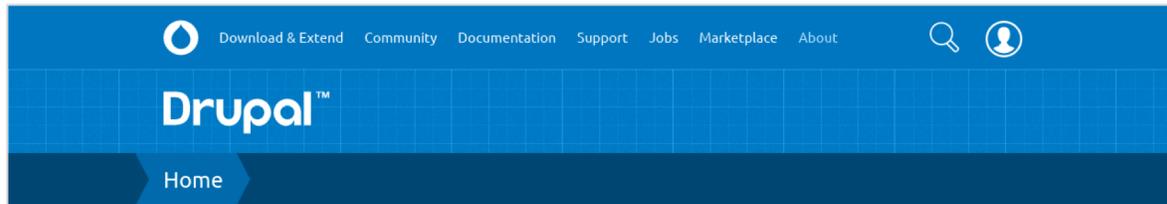
The screenshot shows the HTML source code of the page <https://www.drupal.org/about>. The code is displayed in a dark-themed editor. The visible HTML includes a closing `</div>` tag, followed by a `<div class="pane-content">` block. Inside this block, there is a `<div class="field-items">` container with a `<div class="field-item even">` element. This element contains a paragraph of text starting with "The digital experiences you love. The organizations you trust most. The software they depend on." followed by a heading `<h3>Make something amazing, for anyone</h3>` and several paragraphs of text describing Drupal's features and licensing. The code ends with a `</div>` tag.

Current Project Status: 1.0

- ▶ Reliably captures screenshots and HTML
- ▶ Integration with Headless Chrome and PhantomJS
- ▶ Captures manually, and automatically via crontab-based scheduling
- ▶ Parses XML sitemaps
- ▶ Honors robots.txt restrictions
- ▶ Basic garbage collection
- ▶ Views-driven user interface
- ▶ Test coverage for most functionality
- ▶ Well documented

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

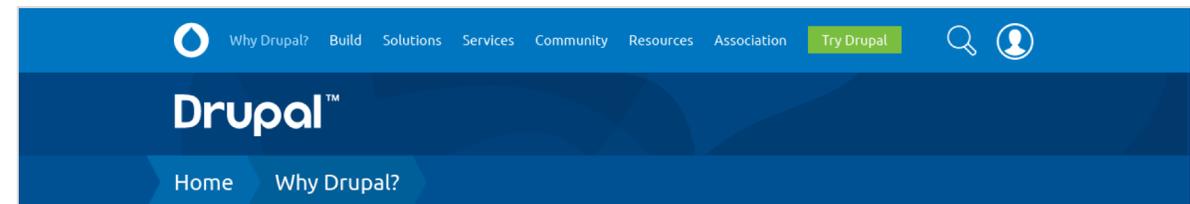
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

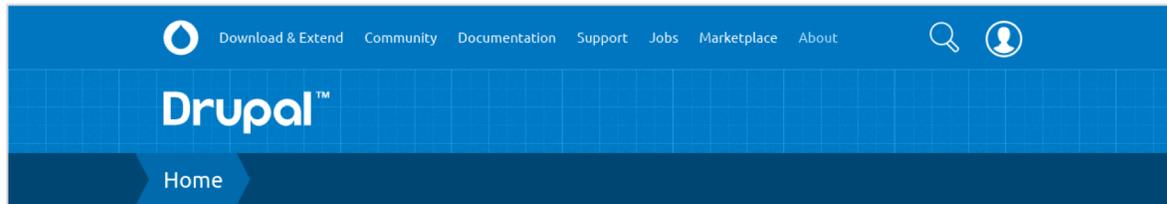
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

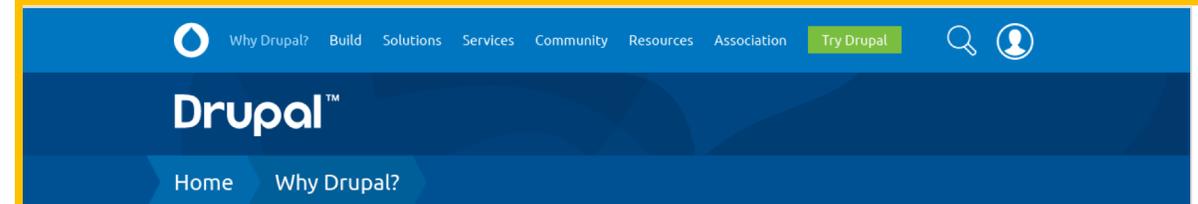
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

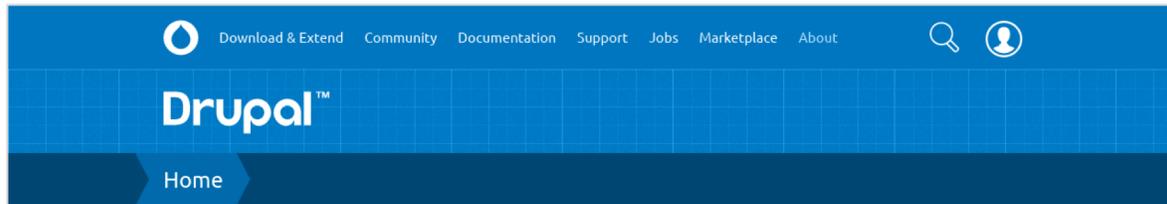
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

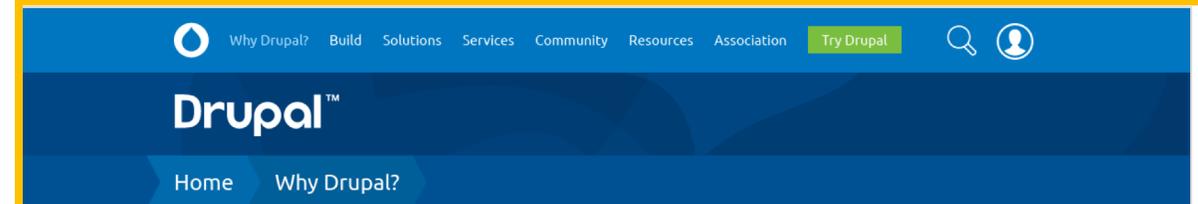
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

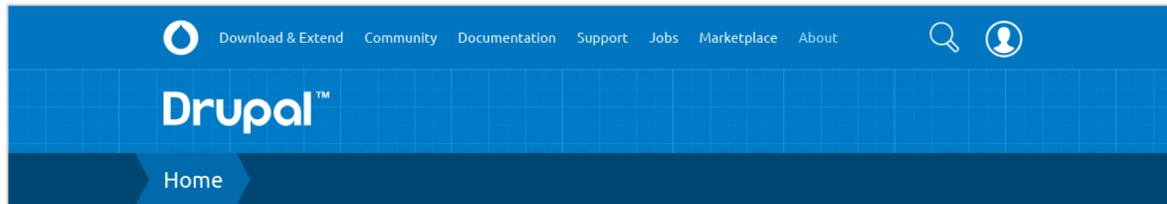
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

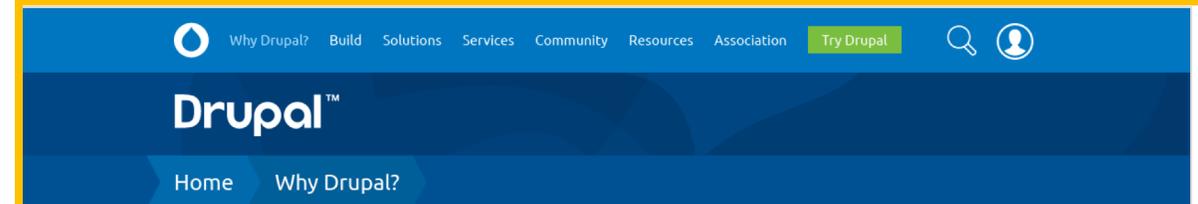
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

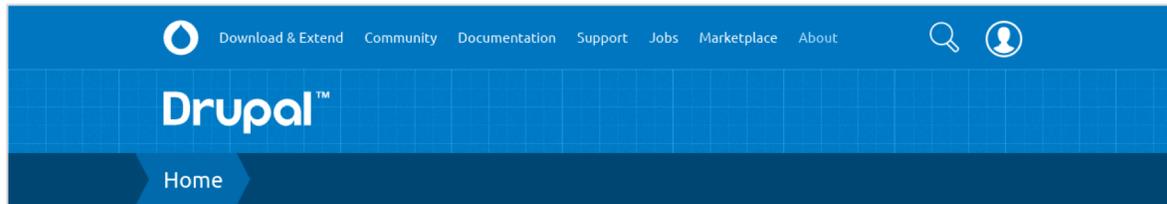
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

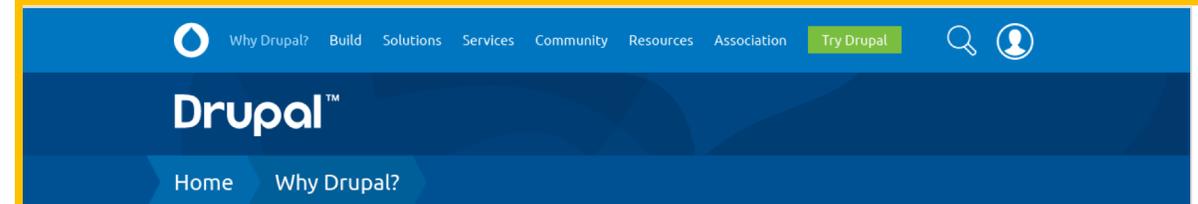
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

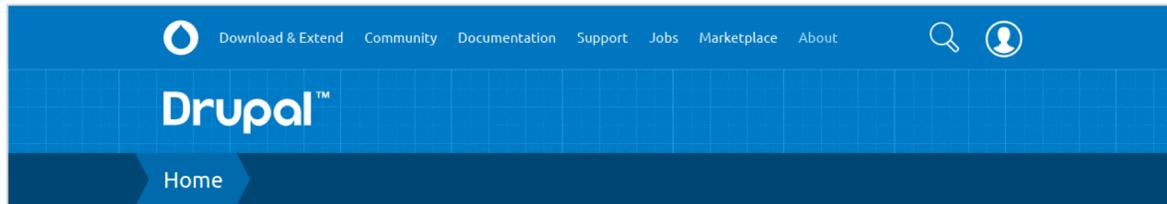
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

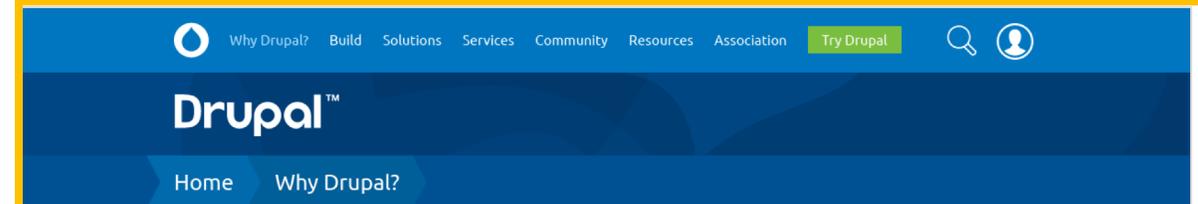
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

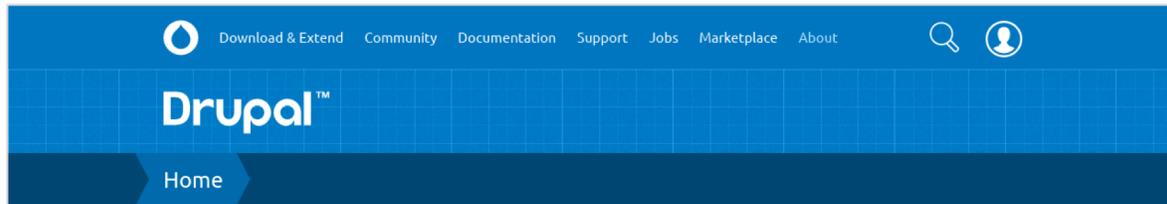
- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Spot the Differences

January 5, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

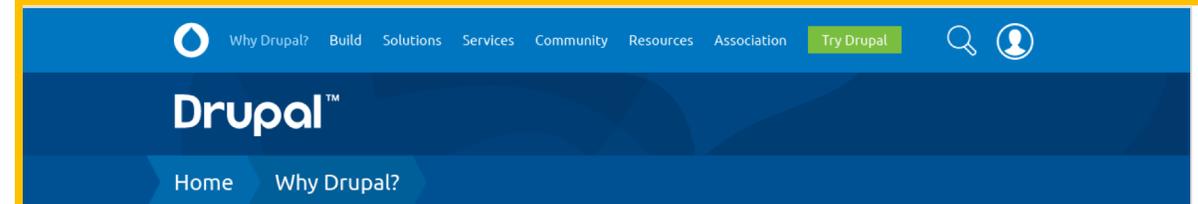
About

- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Mission and principles](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

Recent posts

- [Accelerate Drupal 8 by funding a Core Committer](#)
- [Massachusetts launches Mass.gov on Drupal 8](#)
- [We have 10 days to save net neutrality](#)
- [An update on the Workflow Initiative for Drupal 8.4/8.5](#)
- [An update on the Layout Initiative for Drupal 8.4/8.5](#)

May 26, 2018



About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

Drupal is content management software. It's used to make many of the websites and applications you use every day. Drupal has great standard features, like easy content authoring, reliable performance, and excellent security. But what sets it apart is its flexibility; modularity is one of its core principles. Its tools help you build the versatile, structured content that dynamic web experiences need.

It's also a great choice for creating integrated digital frameworks. You can extend it with any one, or many, of thousands of add-ons. Modules expand Drupal's functionality. Themes let you customize your content's presentation. Distributions are packaged Drupal bundles you can use as starter-kits. Mix and match these components to enhance Drupal's core abilities. Or, integrate Drupal with external services and other applications in your infrastructure. No other content management software is this powerful and scalable.

The Drupal project is open source software. Anyone can download, use, work on, and share it with others. It's built on [principles](#) like collaboration, globalism, and innovation. It's distributed under the terms of the [GNU General Public License](#) (GPL). There are [no licensing fees](#), ever. Drupal will always be free.

Drupal's the content management system behind many of the digital experiences you love. And it will always be free.

[Tweet](#)

Made possible by a global community

The [Drupal community](#) is one of the largest open source communities in the world. We're more than 1,000,000 passionate developers, designers, trainers, strategists,

About

- [Values & Principles](#)
- [Drupal 8](#)
- [Drupal 7](#)
- [Drupal 6 \(end of life\)](#)
- [Industries](#)
- [Our history](#)
- [Strategic Initiatives](#)
- [Accessibility](#)
- [Media Kit](#)
- [Licensing](#)
- [Awards](#)
- [FAQ](#)
- [Drupal blog](#)

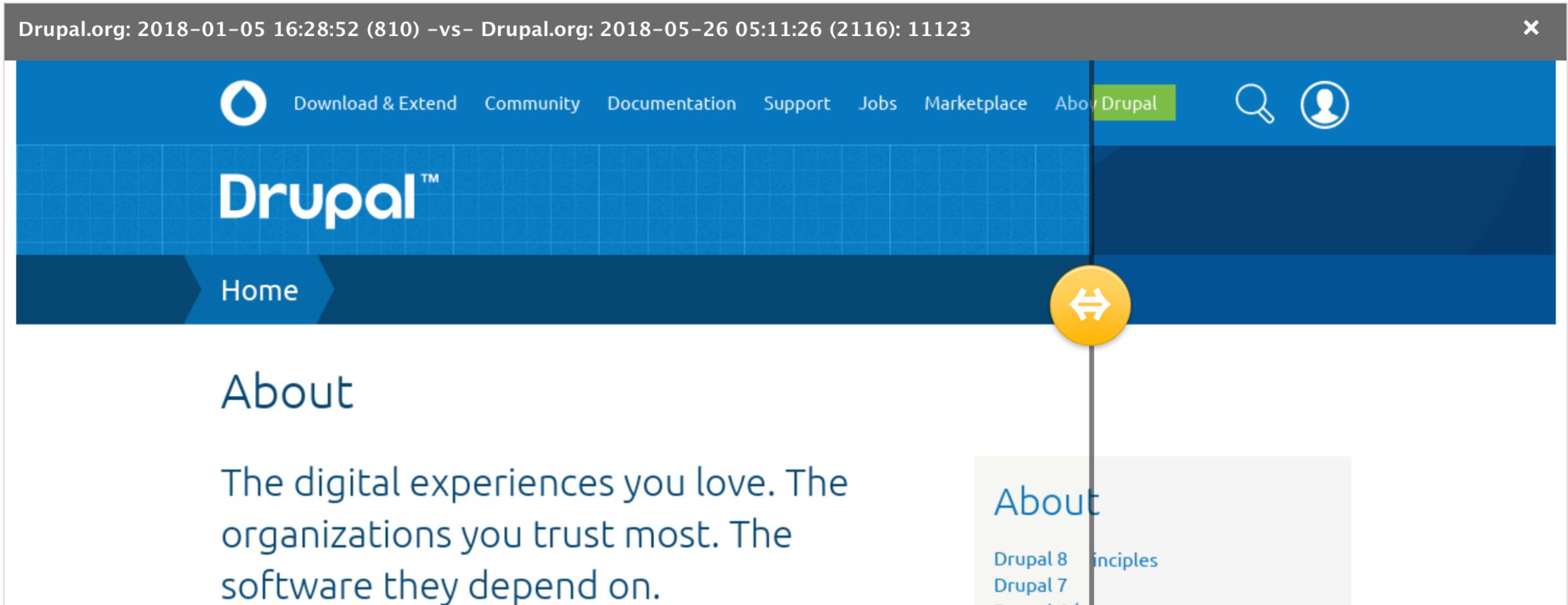
Recent posts

- [When should we release Drupal 9?](#)
- [The Royal Family using Drupal](#)
- [Working toward a JavaScript-driven Drupal administration interface](#)
- [Offering more inclusive user demographic forms](#)
- [How Drupal influences other Open Source projects](#)

Comparison Utilities

- ▶ Drupal plugins responsible for performing comparisons of different types of captured data.
- ▶ Can be used across environments/domains
- ▶ Store comparison results as ComparisonResponse entities

Compare Screenshots: Before & After



Compare Screenshots: Heatmap

The screenshot shows the Drupal.org website with a heatmap overlay. The heatmap is a grid of red squares of varying intensity, indicating user engagement. The most intense red areas are in the top navigation bar, the 'Home' button, and the 'About' section. The 'About' section is highlighted with a red background and contains the following text:

About

The digital experiences you love. The organizations you trust most. The software they depend on.

Make something amazing, for anyone

About

- Drupal 8
- Drupal 7
- Drupal 6 (end of life)
- Industries
- Our history

Compare HTML: Text Diffing

Rackspace.com (en-us, HTML): 2017-12-31 04:11:39 (767) -vs- Rackspace.com (en-us, HTML): 2018-01-12 04:08:27 (863): 795



<pre>class="dropdown-menu navbar-dropDownMenu"><li class="navbar-dropDownItem">Managed Hosting</pre>		<pre>class="dropdown-menu navbar-dropDownMenu"><li class="navbar-dropDownItem">Managed Hosting</pre>
<pre><li class="navbar-tertiary-dropDownItem">Private Cloud</pre>		<pre><li class="navbar-tertiary-dropDownItem">Private Cloud</pre>
<pre><li class="navbar-tertiary-dropDownItem">Public Cloud</pre>		<pre><li class="navbar-tertiary-dropDownItem">Public Cloud</pre>
<pre>- <li class="navbar-tertiary-dropDownItem">Hybrid Cloud</pre>	+	<pre><li class="navbar-tertiary-dropDownItem">Hybrid Cloud</pre>
<pre><li class="navbar-tertiary-dropDownItem">Multi-Cloud</pre>		<pre><li class="navbar-tertiary-dropDownItem">Multi-Cloud</pre>
<pre></pre>		<pre></pre>
<pre><li class="navbar-tertiary-dropDownItem">Ecommerce</pre>		<pre><li class="navbar-tertiary-dropDownItem">Ecommerce</pre>
<pre><li class="navbar-tertiary-dropDownItem">Enterprise Resource Planning (ERP)</pre>		<pre><li class="navbar-tertiary-dropDownItem">Enterprise Resource Planning (ERP)</pre>
<pre>- <li class="navbar-tertiary-dropDownItem">Web Content Management</pre>	+	<pre><li class="navbar-tertiary-dropDownItem">Web Content Management</pre>
<pre></pre>		<pre></pre>
<pre><li class="navbar-dropDownItem">Professional Services</pre>		<pre><li class="navbar-dropDownItem">Professional Services</pre>
<pre><li class="navbar-topItem">Platforms<ul class="dropdown-menu navbar-dropDownMenu"><li class="navbar-dropDownItem">Dedicated Servers</pre>		<pre><li class="navbar-topItem">Platforms<ul class="dropdown-menu navbar-dropDownMenu"><li class="navbar-dropDownItem">Dedicated Servers</pre>

1.1 Release Plan

- ▶ 1.1 Release Plan:
 - <https://www.drupal.org/node/2930857>
- ▶ Some Remaining Issues:
 - #2936374: Comparison Criteria for ScreenshotCaptureResponse - Image Diff
 - #2932535: Run Comparison Documentation

Future Roadmap

- ▶ Roadmap:

- <https://www.drupal.org/node/2916976>

- ▶ Some Upcoming Issues:

- [#2937227: Variance filter](#)
- [#2908636: Inject javascript into session](#)
- [#2922937: Don't save duplicate captures](#)
- [#2896502: Run entity performance on large sitemaps](#)
- [#2975917: Drush 9 support](#)
- [#2977032: Pre-configured Docker Container](#)

How You Can Help

- ▶ Project issue queue
- ▶ Submit patches or pull requests
- ▶ Review code
- ▶ Report bugs
- ▶ Submit feature requests
- ▶ Write tests
- ▶ Help improve documentation
- ▶ **Help improve UX**



DEMONSTRATION



ARCHIVING AND VISUAL REGRESSION USING DRUPAL 8

https://www.drupal.org/project/issues/web_page_archive

Contributions are welcome!

David Stinemetze

Twitter: @davidstinemetze

Drupal.org/GitHub: @WidgetsBurritos

