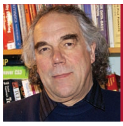
IMPACT

Editor: **Michiel van Genuchten**
VitalHealth Software
genuchten@ieee.org



Editor: **Les Hatton**
Oakwood Computing Associates
lesh@oakcomp.co.uk

WordPress

A Content Management System to Democratize Publishing

Jordi Cabot

From the Editors

Another Impact article about an open source software product. WordPress is supporting many of the websites around us and has created a multibillion-dollar business. —*Michiel van Genuchten and Les Hatton*

WORDPRESS IS THE most popular content management system (CMS), powering up to 29 percent of all websites.[1] A CMS is any system that facilitates the creation and publication of digital content. This category includes platforms for creating static sites, blogs, forums, online stores, and everything in between. Other well-known CMSs include Joomla, Drupal, Shopify, and Squarespace, all of them with a much lower market share. Sixty percent of all sites using a CMS go with WordPress, the only one with a double-digit market share. (Joomla takes second place with a 6.4 percent market share.)

WordPress started in 2003 when Matt Mullenweg and Mike Little forked the b2 blogging platform and created the first version of WordPress. WordPress is an open source project released under GPL (GNU General Public License) v2.0 (or later). Mullenweg started the WordPress Foundation in 2010 (inspired by the Free Software Foundation and Mozilla Foundation) to further support WordPress's sustainability and promote the project. The foundation owns the WordPress trademark.

From the beginning, the WordPress mission has been to democratize publishing, ensuring that any nontechnical person can create a website, while at the same time building a product that can scale all the way up to enterprise clients with complex needs (for example, e-commerce, multilingual, or mobile). The recent addition of the WordPress REST (Representational State Transfer) API is a step forward in this direction. Thanks to the API, you can now use WordPress as a headless CMS to build your own web apps on top of, while benefitting from all its core back-end functionalities (for example, for collaboration and content and user management).

The WordPress Code Base

WordPress comprises more than 500 KLOC consisting mainly of PHP code serving HTTP requests by querying a MySQL database. However, the use of JavaScript is quickly increasing, especially in all the components implied in the front-end aspects of WordPress. React serves as the core framework for WordPress's new JavaScript-based developments.

Nowadays, PHP still represents over half of the total LOC, with an additional 30 percent of JavaScript. CSS, HTML, and XML make up the rest. The code base's compound annual growth rate (CAGR) over 14 years (from 16,314 LOC in Dec. 2003 to

0740-7459/18/$33.00 © 2018 IEEE

565.917 LOC in Dec. 2017, according to Open Hub; www.openhub.net/p/wordpress/analyses/latest/languages_summary) was 28.8 percent. This puts WordPress in the upper end (but still within the typical range) of the calculated CAGRs for software projects.[2]

The database schema is rather simple, with only 12 core tables. (The entity–relationship schema at codex.wordpress.org/Database_Description shows the relationships between the tables, although the corresponding foreign keys aren't part of the WordPress database SQL Data Definition Language script.)

The WordPress code is organized into a few dozen core components on the basis of their functionality (rather than size or language). Each component is coded using a mix of procedural and object-oriented programming techniques. Typically, new functionalities are developed with a more object-oriented style, and a few class wrappers have been added to better encapsulate related sets of functions.

WordPress Development

WordPress uses Subversion (SVN) as the version control system, with the current development version available at develop.svn.wordpress.org/trunk. Trac (core.trac.wordpress.org) is the bug-tracking system; all important discussions take place in it, with Slack as a complement for real-time communication. There's also a GitHub mirror, but only as a read-only version of the SVN repository.

More than 30 versions of WordPress have been released; the current one (at the time of writing) is 4.9.4. WordPress strives never to break backward compatibility. This is unlike other CMSs such as Drupal, which is prepared to break it at every major release if that simplifies radical code improvements.

For historical reasons, WordPress doesn't use semantic versioning. So, the first two digits of the version number identify a major release, with the third digit identifying a minor release (mainly for security patches and bug fixes). Until recently, a new major release occurred regularly—approximately every four months. In 2016, WordPress announced that it would move toward a more feature-driven approach, with each new version focusing on the release of a major functionality. Editing, customization, and the REST API were the first three releases planned. In the end, we've witnessed a mixed approach, with different release models happening in parallel.

Each release has a lead developer and can now also have a lead designer. (This recognition of designers' key role in development is occurring in a number of tech companies that are drastically improving their developer-to-designer ratio.[3]) Leaders rotate with each release, which helps involve more people in key positions of the project. This in turn helps improve the project's *bus factor*, a measure of the risk of concentrating too much information in a small number of developers.

The release leads decide on all technical aspects of the release, but they depend on the WordPress community to move the code forward. Code evolution also includes adding unit tests (WordPress uses PHPUnit and QUnit for automatic testing of PHP and JavaScript, respectively). For instance, 443 contributors participated in the latest WordPress major release, of whom 185 were new contributors.

Beyond the lead developers, a number of core committers have write access to SVN and can therefore commit the patches submitted by those contributors. Sometimes WordPress grants temporary commit access to people working on specific components. A number of people have permanent commit access; they constitute the WordPress core team.[4] Promotion from being an external contributor to being a member of the core team is based mainly on merit. This meritocracy takes into account not only technical skills but also attitude, professionalism, and respect for the project's core philosophies. At the top of the chain, Mullenweg supervises everything under his (unofficial) role of Benevolent Dictator for Life.

An open bug triage meeting is held weekly, but security vulnerabilities are addressed immediately. (The list of declared security vulnerabilities is at www.openhub.net/p/wordpress/security.) If necessary, a new minor security release is prepared, and all WordPress sites are automatically updated. The WordPress security team comprises approximately 50 experts.

Making Money with WordPress

WordPress is a multibillion-dollar business, and everything indicates that it will continue to grow. A key factor in this growth is the huge WordPress user community. For example, WordPress organizes WordCamps—locally organized events featuring local speakers. Last year, there were 128 WordCamps and more than 4,000 Meetups (informal face-to-face gatherings), totaling more than 100,000 participants.

Beyond offering all kinds of consulting services (installing, configuring, tuning, migrating, and so on) to serve the WordPress community, many people build and sell plug-ins and themes. Plug-ins extend the core WordPress functionality; themes customize WordPress sites' look and feel. WordPress offers many predefined *hooks*,[5] which plug-ins and themes can "hook into" to provide their functionality without modifying core WordPress files. Around 2,000 hooks are available; each hook corresponds to a common WordPress event (such as saving a post, approving a comment, or creating a user). In response to those events, plug-ins and themes can perform a specific action or filter the current content to change how it's displayed to the user.

The official repository contains approximately 47,000 plug-ins, which have been downloaded more than 600 million times. Plug-ins (and themes) can be free, be paid (sometimes as part of a subscription service), or follow a freemium model. The plug-ins' quality varies wildly; usually, they're at fault when a WordPress site gets hacked.

A plug-in review team[6] or a theme review team checks that each submitted plug-in or theme adheres to the guidelines (verifying security, quality, and the lack of "spammy" aspects). Like all sets of rules, the guidelines try to be precise but are open to interpretation. So, there has been controversy regarding decisions on the inclusion or exclusion of certain plug-ins and themes. The upcoming Tide initiative (github.com/wptide) will make the process more transparent by helping developers run automated tests to check for PHP compatibility errors and warnings before submitting their plug-ins or themes.

Although the creators of plug-ins and themes are mainly independent developers or small agencies, larger companies offer WordPress hosting services with a more predictable and recurring revenue model. Although you could install WordPress on any Internet hosting provider, some provide more dedicated support for WordPress sites, offering, for instance, staging sites or integrated cache systems. One such company is Mullenweg's Automattic, which offers hosting services under the domain wordpress.com (not to be confused with wordpress.org, the home of the open source project itself[7]).

The Road Ahead

WordPress has gone a long way from a humble blogging platform to the flexible CMS it is today. But it will need to continue evolving if it wants to stay on top. The CMS market is appealing, with new competitors popping up every year in all areas of the CMS spectrum. Many of them are trying to become the best CMS for specific customer profiles or sectors, in contrast to WordPress's aim to be the one-size-fits-all solution.

In reaction to this threat, the WordPress Growth Council has been recently created. The Five for the Future initiative complements this by requesting all companies living on WordPress to dedicate 5 percent of their people to contribute back to the WordPress core—be it development, documentation, security, support forums, theme reviews, training, testing, translation, or whatever helps move the WordPress mission forward.

On the technical side, the next WordPress version will ship with Gutenberg (github.com/WordPress/Gutenberg), a major architectural shift for WordPress and the longest feature development effort in WordPress's history. Gutenberg aims to simplify all previous concepts of WordPress (for example, the menu, the widgets, and the shortcodes) into one elegant concept: the block. According to Mullenweg, Gutenberg will be the future of WordPress writing, editing, and customization for years to come.[8] As with any major change, Gutenberg has stirred much controversy in the community because it will force many developers to rewrite their plug-ins and themes and rethink how they work. It will also make many of the popular page builders obsolete unless their developers invest considerable effort to adapt to the Gutenberg methodology.

These changes are in line with WordPress's goal of dominating the CMS market, from large enterprises to individual bloggers with little technical knowledge. However, some communities of WordPress users could feel that the project is evolving in a direction that no longer represents their views and decide to fork the project and create a specialized version to better fit their needs. This is a risk that all open source projects face and must prepare for.

One way or the other, the research community has much to contribute to WordPress's future. It's somewhat surprising that so few research articles focus on WordPress, compared to, for instance, papers analyzing Linux from every possible perspective. The richness and importance of the WordPress code base and ecosystem pose many interesting challenges for the research community, especially for researchers working on mining software repositories. As a longtime WordPress user and researcher, I encourage you all to contribute to the growth of WordPress and its community.

ABOUT THE AUTHOR

JORDI CABOT is an ICREA Research Professor at the Internet Interdisciplinary Institute, Universitat Oberta de Catalunya. Contact him at jordi.cabot@icrea.cat.

References

1. “Content Management Systems,” W3Techs, 2018; w3techs.com.
2. M. van Genuchten and L. Hatton, “Compound Annual Growth Rate for Software,” *IEEE Software*, vol. 29, no. 4, 2012, pp. 19–21.
3. D. Field, “6 Major Tech Companies Have Doubled Their Design Hiring Goals in the Last Half Decade,” TechCrunch, 31 May 2017; techcrunch.com/2017/05/31/here-are-some-reasons-behind-techs-design-shortage.
4. “Make WordPress Core,” WordPress, 2018; make.wordpress.org/core/handbook/about/organization.
5. A. Brown, “WordPress Hooks Database,” 2017; adambrown.info/p/wp_hooks.
6. “Make WordPress Plugins,” WordPress, 2018; make.wordpress.org/plugins/handbook.
7. “WordPress.com and WordPress.org,” WordPress, 2018; en.support.wordpress.com/com-vs-org.
8. B. Krogsgard, “Interview with Matt Mullenweg on the WordPress Ecosystem—Draft Podcast,” *Post Status*, 18 Nov. 2017; poststatus.com/interview-matt-mullenweg-wordpress-ecosystem-draft-podcast.