

Analyzing Moodle Plug-ins Across Several Thousand Sites

Mike Churchward

CIO, Remote-Learner.net Inc., Canada/UK/USA, mike@remote-learner.net

Abstract

Remote-Learner.net Inc., manages, maintains and supports thousands of Moodle sites with hundreds of different Moodle plug-ins. The data produced across all of these sites, combined with CRM data, provides a wealth of information used to plan for the needs of that userbase.

This paper discusses how we mine the large amount of data we have, to learn about the plug-ins our users want, need and use. And how we use that data to plan for the add-ons we will provide, support and manage for our users.

Keywords

Moodle, plug-ins, modules, blocks, trends.

Introduction and Overview

As a large Moodle service provider, we manage and maintain thousands of Moodle sites with hundreds of different Moodle plug-ins. Understanding the needs of this large userbase is tantamount to our management success. This paper discusses how we mine the large amount of data we have to learn about the plug-ins our users want, need and use. And how we use that data to plan for the add-ons we will provide, support and manage for our users.

We will discuss the type of data we have available to us, the collection techniques used, the analysis we do and the results and actions we determine. Specific examples of our data and analyses will be discussed.

Background

At Remote Learner, we manage our client and site information using a variety of technologies. Our base client information is kept in the Salesforce CRM system. This provides us with data specific to each client account, including information such as market segment. We have automation tools that work across our server network that provides very specific data about each installation back to our central databases. These tools interact with Salesforce to connect the client data to the site data.

As part of our business, we provide installations of Moodle with all standard Moodle plug-ins, as well as a number of non-core add-ons we refer to as “Golden Add-ons” or GAO’s. Additionally, any client can request the inclusion of any add-on not supplied as part of the previous collection. These are reviewed for performance and security and if approved, installed into the client’s codebase.

We maintain multiple versions of Moodle for the client base. Currently, we have around 380 different releases of Moodle across all of our inventory, with almost half still on a version of 1.9. We are working to standardize these releases to no less than three major releases with one version per release.

Figure 1 shows the breakdown of our Moodle releases:

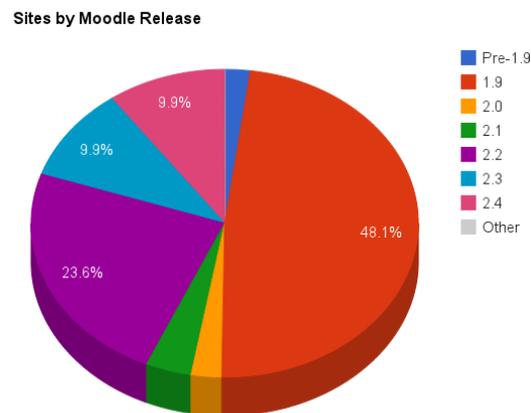


Figure 1: Breakdown of Moodle Releases

To maintain all of these releases, we have to create a new codebase for each new release, consisting of the Moodle code and all of our extra add-ons. As a result, we “lag” behind the major release date by up to 90 days.

Our current systems have automated much of this process, allowing us to collect much of the data used in this report, and get a much clearer picture into what our clients are using and need.

Data Available

Our unique position as a Moodle service provider gives us a great sample of data. The summary of data used for this report is as follows:

Table 1: Data Summary

Moodle Sites	3,574
Moodle Courses	2,540,009
Users	18,356,395
Data Sizes	23 TB of data / 86 TB of data plus content
Total Unique Blocks	556
Total Unique Modules	225
Total Unique Auth Plug-ins	49
Total Unique Enrol Plug-ins	34
Number of Pre-1.9 Sites	76
Number of 1.9 Sites	1721
Number of 2.0 Sites	91
Number of 2.1 Sites	137
Number of 2.2 Sites	843
Number of 2.3 Sites	354
Number of 2.4 Sites	354

Data Collection Techniques

All of the sites installed and managed, are maintained with strict processes and scripts. These scripts also retrieve information about each site and return that information to our master database. These scripts run on regular schedules, designed to minimize the impact on the actual sites.

For example, basic configuration information about a Moodle site is returned hourly. Information on installed plug-ins is returned weekly. This gives us very accurate data on versions, configurations, sizes and more.

Data is also retrieved from our Salesforce application, and the records associated with the sites, correlated with the records associated with the CRM. This allows us to (among other things) report on specific client demographics, such as market sector.

None of the data retrieved includes any personal information identifiable to any particular person. The only data on users retrieved at all is the number of users on a particular site.

The plug-in data collected includes the following data for each site:

- plug-in name and type,
- version and release information,
- whether the plug-in is enabled (visible) on the site,
- whether the code is present on the site,
- the source of the plug-in (core, GAO, non-standard),
- the number of instances in use.

There is other information that can be derived by linking the site records to the other collected records.

For purposes of our analyses, we focused on the plug-ins that were used on the sites. For this analysis, that meant modules, blocks, enrolment plug-ins and authentication plug-ins that were visible (enabled) on the site. We cross-referenced this data with site data to compare market segments and Moodle releases.

The data used in this report was gathered in the May/June 2013 time period.

Uses

At a general level, being able to report on usage of Moodle add-ons, allows us a peek into what our users are doing with Moodle from a much wider scope. At an operational level, the data allows us to plan for client support and training, allowing us to make sure our staff is focused in the right areas. At the sales level the data allows us to target specific solutions to specific sectors of clients. At the service level, seeing the non-standard plug-in usage allows us to identify other potential high-value add-ons that our clientbase may have use for. And at the maintenance level, seeing potential upgrade problems well ahead of time provides us the ability to manage risk better.

Results and Analysis

Our analysis allowed us to track all module usage by number of clients, number of sites and number of instances. For general plug-in statistics, the number of sites using the modules proved to be less useful data. All of our sites are installed with all of our standard modules (core and GAO) installed and visible. Very few clients ever disable a module, so the data for modules per site for all standard installed modules was pretty much the same as the number of sites.

We also had the number of installed instances for each module which was more useful, but in some cases a very few sites had a very large number of installed instances of a specific module, skewing the results. To come up with a more useful analysis, we created a weighted value by multiplying the number of sites by the number of instances. We called this value the “Rank Value”. For our needs, this helped to bring the most used plug-ins to the forefront, highlighting the *truly* most used modules. Figure 2 shows the results:

All Modules by Rank Value

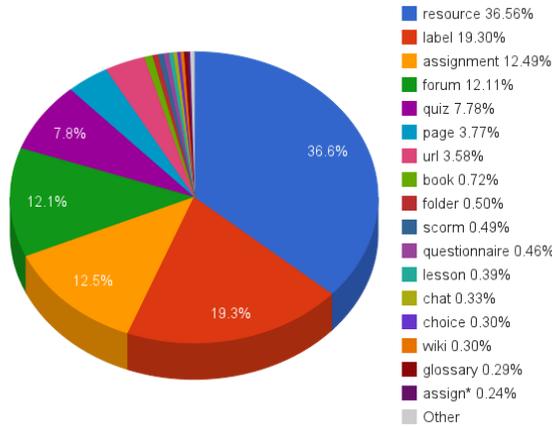


Figure 2: All Modules Ranked

The interesting part of this data is that 44% of the modules used are information-only resources (resource, page, url). If you include “label” in that total, it increases to 64%. Assignment and quiz total another 20%. These are “assessment” activities. And then the forum module, the collaborative basis for Moodle, accounts for 12%. The least used current core modules were “feedback” and “lti”. Although at the bottom, “lti” is very new in comparison to the other core modules, and was used over 6000 times. This still makes it an interesting module to watch and see how the trends over time go.

Our data allowed us to separate the core Moodle modules from our GAO and non-standard modules installed. The core module statistics, even using the weighted, ranked calculation, pretty much showed the same results as the general data above.

For our purposes, getting to the next level was more interesting. Analysing the number of non-core modules installed in our userbase. Figure 3 displays this data:

Non-core Modules by Rank Value

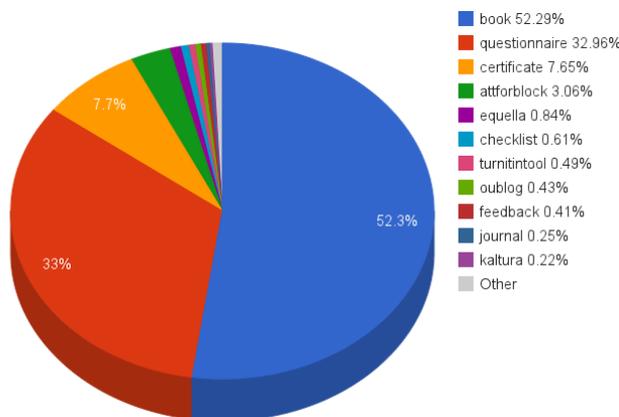


Figure 3: Non-core Modules Ranked

This data shows that, once-again, information-only modules lead the way with the “book” module (now core) taking over 50% of the ranked instances. The next module “questionnaire” accounts for 32%, well above its functional counterpart “feedback” at 0.41%. This could be due to legacy - “questionnaire” has been available a lot longer than “feedback”. When we extend our analysis to include date trends, we may be able to determine which one is more likely to be used now.

Looking at module usage by sector - Higher Education (universities and colleges), K-12 and Corporate provided some interesting differences.

For Higher Ed, information-only modules still led the way, but the forum module ranked above quiz and assignment. Assignment at 11.8% ranked above quiz at 8.6%. The forum and assignment rankings show a high instructor involvement in courses for Higher Ed.

For K-12, again resources and labels were ranked high. Interestingly, the assignment module ranked first in interactive modules, followed by the forum. The quiz module trailed after that. This shows an even higher instructor involvement, expected at this level. An interesting difference between Higher Ed and K-12 is the higher ranking of the SCORM module in K-12 over Higher Ed. This may be due to the often higher multimedia content in SCORM’s. It will be interesting to see if LTI shows a trend increase in K-12.

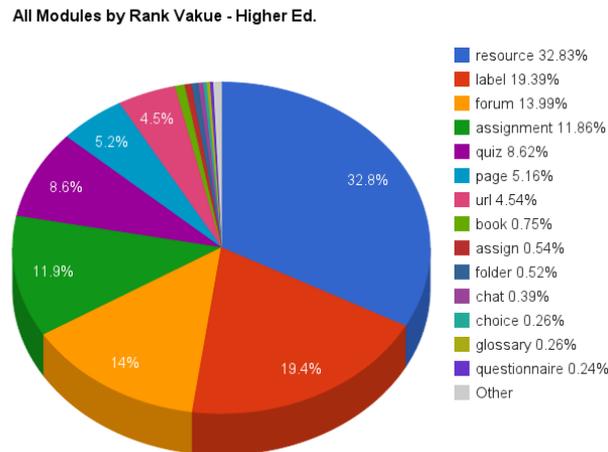


Figure 4: All Modules Ranked – Higher Ed

All Modules by Rank Value - K-12

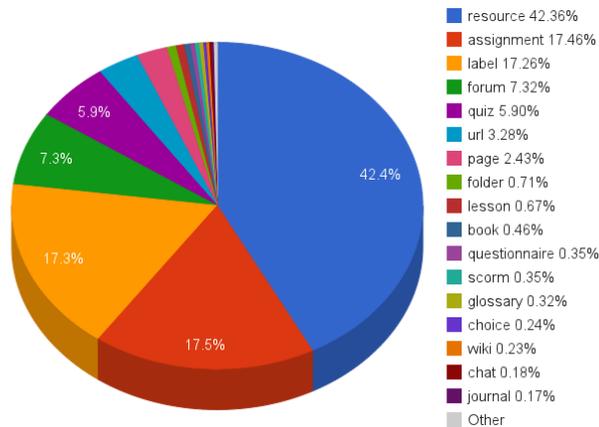


Figure 5: All Modules Ranked – K-12

The corporate sector shows similar results to the education sector, surprising in some cases. We expected that corporates would have their courses set up primarily on “auto-pilot”, yet assignment and forum ranked right near the top at 9.5% and 9.3% respectively. Both of these expect some form of high instructor interaction. The quiz module was next at 8.7%. The SCORM module captured the highest ranking yet at 1.6%. This was expected as we see a lot of corporate clients relying heavily on SCORM content.

All Modules by Rank Value - Corporate

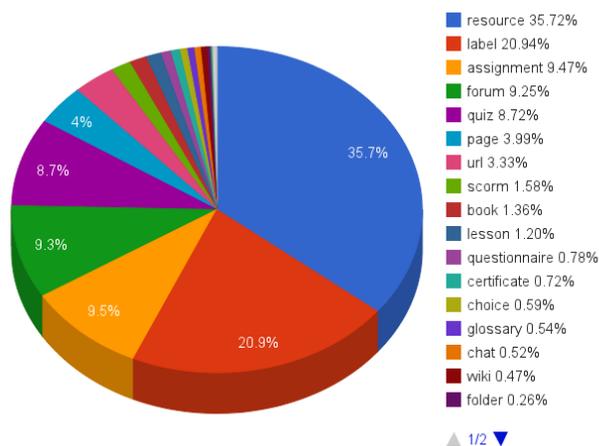


Figure 6: All Modules Ranked - Corporate

The final data set we looked at was the overall ranking of non-standard modules. These are the modules that are not enabled by default on our installations. Some of these are installed but hidden while others are specifically requested by clients. This data provided some surprises and some concerns. The surprises included the “game”

and “lightboxgallery” modules. Neither of these had been considered for our GAO list, but clearly they are in demand.

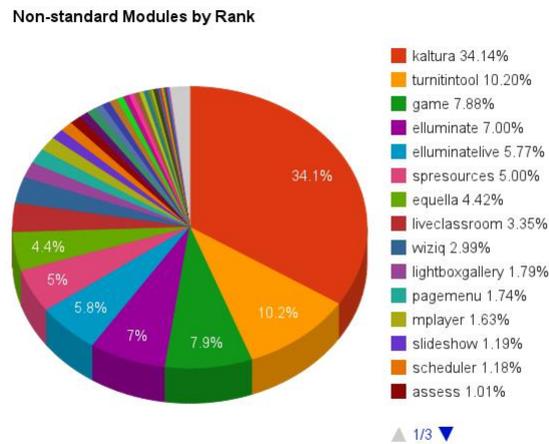


Figure 7: Non-standard Modules Ranked

Of concern are the “elluminate” modules. Elluminate is a discontinued product, replaces by different applications. These may need to be addressed for the clients using them.

Blocks

Our data for blocks showed a fairly common pattern across all of our user sectors.

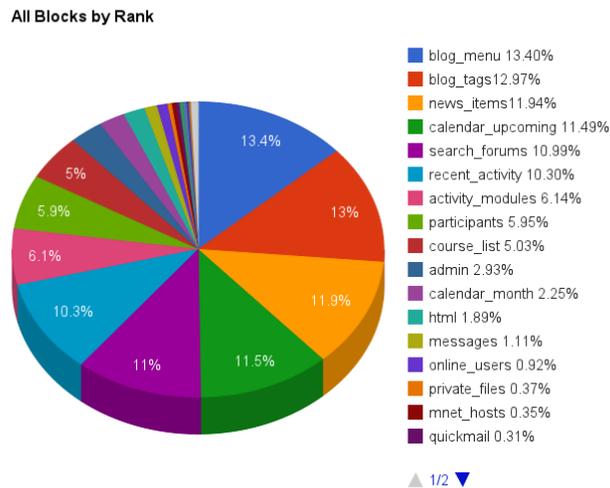


Figure 7: Non-standard Modules Ranked

Of note were the “quickmail” and “grade me” blocks. These two non-core blocks were used most frequently. The “quickmail” block shows a need for easily integrated external messaging while the “grade me” blocks

shows a need for easier to identify grading tasks by instructors. More analysis will be done on blocks in the coming months.

Conclusions

Our first pass through the data showed the potential for a huge wealth of analysis that can and will help with our planning needs in almost all sectors of our business. More correlations is needed to more client factors, and this will continue to grow.

Our first analysis has already helped us determine where we can expect training and support needs. And our business sector analysis has helped us design demo sites highlighting the most used plug-ins.

We will continue to grow the data set and improve the analysis.

Next Steps

More Market-specific Analysis

We currently have more meta data available to us on specific market demographics per client. More emphasis on gathering that data into the master database will provide the ability to give even more granular usage analysis.

More Categorization of “type” of Add-on

At present, we are only tracking modules, blocks, authentication and enrolment plug-ins. Moodle offers many more plug-ins. We will increase our gathering efforts to expand the numbers of plug-ins we can track and analyse.

Trend Analysis

Our data needs to store historical data, through archiving existing retrieved data, and combining the date and history fields of our sites. With this data, we will be able to see if certain usage patterns of add-ons resolve.

Correlate Data to Config Settings

Data that we currently gather includes all configuration settings. At present, we have not used this data for analysis. We need to create analyses that look for interesting trends in plug-in usage with certain patterns of configurations.

References

Remote-Learner.net (2013), Client dashboard database. not publically available.
Moodle.org (2013), Moodle Plugins Directory. <https://moodle.org/plugins/>.