K-12 Web Systems

A Descriptive Web-Crawling Report on Idaho Schools

Royce Kimmons, PhD
March 2014
About the Doceo Center

The University of Idaho Doceo Center for Innovation + Learning leads teaching innovation and technology integration initiatives at the University of Idaho and in Idaho’s K12 schools. The Doceo Center does this by supporting, teaching, modeling, and researching technology integration practices amongst teacher education candidates, school administration candidates, early-career teachers, practicing teachers and administrators, university faculty, and K12 school districts and by providing resources and research to state and national audiences. The Doceo Center is funded by a grant from the J.A. and Kathryn Albertson Foundation.

About the Author

Dr. Royce Kimmons is the Director of the Doceo Center for Innovation + Learning and an Assistant Professor of Learning Technologies at the University of Idaho. He received his PhD from The University of Texas at Austin, and his research agenda focuses upon technology integration and online learning in K12 and higher education settings.

Suggested Citation

K-12 Web Systems

A Descriptive Web-Crawling Report on Idaho Schools

CONTENTS

1 Executive Summary 2
2 Introduction 3
3 Methods 6
4 Popular Systems 8
5 Types and Categories 12
6 Conclusion 17
7 References 19
1 EXECUTIVE SUMMARY

To understand and support large-scale innovations with technology for K-12 education, it is essential to understand the practices, needs, and interests of K-12 institutions. Web-based information systems are of particular interest, because schools and districts use such systems to support a variety of functions ranging from instruction and classroom management to communications and reporting. Since K-12 institutions may utilize a variety of technologies to meet these needs and systems do not exist to track institutional use of web-systems from year-to-year, this report utilizes web-crawling of K-12 school websites to determine which web systems schools are using and categorizes them in terms of function, cost, and openness.

Results from this study reveal that Idaho K-12 schools utilize a variety of web systems to support local needs. The most prominent systems identified in terms of school adoption include: Google Sites, PowerSchool, Blackboard, Wordpress, SchoolLoop, and Blogspot. Each of these systems is adopted in accordance with either a top-down pattern (i.e. school adoption is led by the district or region) or a bottom-up pattern (i.e. district and regional adoption is led by individual schools). Blogspot is a clear example of bottom-up adoption, while Blackboard is a clear example of top-down adoption.

Each of these systems may either serve a generic purpose (e.g., hosting web content) or be designed for education-specific functions (e.g., reporting attendance). Of schools with identified web systems, 44% utilize generic systems and 76% use education-specific systems, with 20% using both. Web systems also have varying costs, ranging from systems that are very expensive to those that have no cost. 44% of schools use a no-cost web system and 75% use a purchased system, with 19% using both. Schools that only use no-cost systems primarily use generic systems, and openness does not seem to be a major factor in system adoption.

Implications of these findings for decision-makers include the following: a single web system solution or set of solutions for all schools seems problematic at best; top-down attempts at scaling innovation could benefit from recognizing which systems have grassroots support; the lack of adoption of open-source systems may reflect lack of local capacity to adapt web systems to support local needs and represent a hurdle to large-scale innovation; and regional differences in adoption exist and may help shed light on region-specific challenges to technology integration.
2 INTRODUCTION

With the advent of the internet and ongoing innovations in networking technologies, schools and school districts have readily adopted a variety of web-based information systems to support their teaching, learning, and administration responsibilities. In addition, most public schools have now adopted publicly accessible websites that allow them to connect with students, parents, the community, and other stakeholders for communication and reporting.

Given the highly localized nature of public education in the United States, wherein schools and districts exercise a large amount of autonomy in determining how to address the needs of the students in their care, schools and districts generally adopt web systems that they view to be valuable for their goals and objectives on a location by location basis. This is problematic for state-level decision-makers and researchers for determining the efficacy of online systems, because data is difficult to collect on usage from local sites and there is extreme variability in which web systems are used, why they are used, and how.

Even in cases when states or regions may provide specific services to schools and districts through group contracts, there is often no guarantee that schools and districts will or should use the provided systems in ways that are educationally or administratively valuable, and just because a system is made available to a school or district, this does not mean that it will be utilized in any meaningful way or with any regularity.

To effectively support and improve our educational institutions, it is important for us to determine which web systems schools are using, why they are using them, how they are using them, and how this impacts student learning and achievement. However, ascertaining this information is difficult, and decision-makers may often rely upon promises and isolated case studies provided by vendors in making decisions on which web systems to implement without having any real data on large-scale usage and outcomes of such systems.

Using web crawling, this report seeks to provide a descriptive view of web system adoption by K-12 schools in the state of Idaho and to give decision-makers necessary information for identifying common good services and common needs. Though this report only focuses on the state of Idaho, provided descriptive statistics may also be useful to those in other states and contexts for making decisions regarding web systems as processes and findings may be transferrable.
Types of Web Systems

Schools and districts may use a variety of web information systems to support their day-to-day activities. Usage of these systems may include supporting instruction through lecture capture and streaming, supporting collaborative group work through document sharing, reporting of numbers to district or state agencies, tracking of student progress and attendance, sharing of resources and announcements with the community or groups of students, and so forth. As a result when exploring web information systems, we are actually exploring a wide array of different families of systems that each may provide a variety of general or specialized uses. This report divides web information systems into two main categories: generic systems and education-specific systems.

**Generic systems** include any web information system that is designed for flexible use in multiple industries or for a variety of different types of users. Content management systems (CMS) are examples of generic systems, because they provide users with the ability to manage a variety of content (e.g., files, webpages, videos, audio) online in a flexible manner that might be utilized by a business, government agency, nonprofit, or any other organization. In this sense the design of the system is not influenced by education-specific goals and objectives but rather has as its aims merely serving as an organizational or communication tool that could be used in education just as well as another field.

In contrast, **education-specific systems** refer to web systems that are specifically designed and engineered to support the needs of educational institutions. Two major examples of education-specific systems include student information systems (SIS) and learning management systems (LMS), or variants thereof. These types of systems have specific goals implicit in their design that lead them to address specific challenges or needs of educational institutions. For instance, an education-specific system that allows for the creation of user accounts may organize users according to education-specific categories like student, teacher, administrator and so forth, and mechanisms developed in such a system will be geared to replicating or supporting existing education processes, like the recording of attendance, reporting on student progress, and so forth.

In this report we differentiate between general systems and education-specific systems to provide the reader with an understanding of the presence and prevalence of each among K-12 schools adoptions and to open up conversations for determining effective uses.

Additional Categories of Web Systems

This report further utilizes two more web system categorizations based on cost and openness of the source code utilized for the system. In the first category of cost, some systems are provided at no-cost to the public or to educational institutions in particular while others must be purchased (e.g., via a license or hosting). Two examples of no-cost web systems are Google Sites, which is freely available to anyone who would like to create an account, or Moodle, which
may be installed on a school server at no cost. In contradistinction to these no-cost examples, other web systems either require licensing, hosting, or the purchasing of accounts. For example, in order to use Blackboard, institutions incur a use fee that is negotiated based on a variety of factors including users and type of hosting. Distinctions between no-cost and purchased web systems are important, because web systems can be very expensive for schools or districts to purchase, and we need a clear understanding of which schools and districts are utilizing purchased versus no-cost solutions in order to support their effective, sustainable, and scalable use.

Regarding the second category of openness, some web systems open their source code to the community and allow it to be edited, revised, and shared without the consent of the original author, while other web systems maintain copyrighted, proprietary ownership of their source code. Generally speaking, open source solutions may often be appealing to institutions with the expertise necessary to implement them, adjust them, and troubleshoot them on their own, while proprietary solutions may be more useful for institutions that seek to offload such technical support work to a third-party. Examples of open source web systems include open source content management systems like Wordpress, Joomla, or Drupal and open source learning management systems like Moodle. With these systems, schools and districts can install, use, and edit the system’s source code to meet their specific needs. Proprietary systems, on the other hand, do not make their source code available and prevent local adaptation. Examples of proprietary systems include popular learning management systems like Blackboard but also include many no-cost services like Google Sites, Wix, and Weebly. This report differentiates between cost and no-cost systems and between open source and proprietary systems for the purpose of improving understanding of school needs and values.
3 METHODS

Determining actual school adoption of web systems can be a daunting task given that so much decision-making regarding web system use occurs on the local level, which means that each school may use a different system for different purposes and may not report this use (even if asked). For this report, we utilize a unique data collection approach to exploring this issue by utilizing web crawling technologies to discover what web systems schools and districts utilize for their primary web presence and which web systems they link to therefrom. By collecting and categorizing such information, we hope to illustrate web system usage in the state of Idaho and to provide understanding on the different types and frequencies of web system use for the purpose of informing large-scale decision-making on the state level and beyond that is based on an understanding of the needs and capacity of local schools and districts.

Beginning with a list of all school websites in the state of Idaho (Idaho State Department of Education, 2013b), we utilized PHP and MySQL to design a web crawling script that would visit each website, record some identifying information about the site, record all links provided on the main page of the site and categorize these links in a variety of ways. The script would then visit all linked pages that remained within the school website and perform indexing of links on those pages as well. By developing our script in this way, we are assuming that schools that utilize web systems to support their day-to-day activities will most likely provide links to these systems on their websites and that the types of systems linked to will reflect the schools’ needs and uses. Intranet systems, like some student information systems that are not intended to be accessible off-campus, are excluded from this collection process, because they serve a different purpose than web information systems (internal vs. external, intranet vs. internet). This report focuses on internet-based systems only.

A Linux server was utilized to run scripts developed for the study and to store collected data. In total, seven hundred thirty-two (n = 732) school websites in the state of Idaho were indexed, accounting for 1.7 million links and representing 12,357 websites. A set of heuristics were then developed to identify web systems used by schools based on url patterns, keywords, and html identifiers (e.g., school sites that had a url including the string “wp-login.php” were identified as using Wordpress, those that had a url including the string “bblearn” were identified as using Blackboard). In this process, it was assumed that schools might utilize more than one web system; therefore, web system coding was inclusive (e.g., a school might be identified as using both Wordpress and Blackboard). Data were then coded within the MySQL relational database, and cross-referenced tables of web system types and categorizations were generated. Other data were also introduced into the database from publically available sources, including data
sets on school multi-year academic ratings, results of standardized testing, enrollment, and budget expenditures (Idaho State Department of Education, 2013a, 2013b, 2013c, 2013d). State-assigned school and district identification numbers were used as relational identifiers to connect data sets together. Relational data were then aggregated and exported to a .csv file for analysis in SPSS, which was used to provide descriptive statistics of the data and to conduct analyses on research questions.

This report deals primarily with reporting on descriptive statistics and answers the following research questions:

1. **What are the popular web systems in use among Idaho schools, both statewide and across regions?**
2. **What types and categories of web systems are in use both statewide and across regions, and what is the relationship between them?**
In exploring our collected data, a number of major web systems were identified based upon URL patterns, domains, and page level identifiers. Twenty two major web systems were identified through this process and include: PowerSchool, Google Sites, Wordpress, Weebly, Blackboard (including Blackboard and EdLine), Blogspot, SchoolLoop, Moodle, SchoolNet, Joomla, MyBigCampus, Drupal, Brainhoney, InfiniteCampus, Webs.com, SharpSchool, Wix, SchoolWires, EPivot, DotNetNuke, Apex, and SchoolStream. Each of these systems is included in Figure 1 and Table 1, along with information on adoption rates.

School adoption rates were determined by the number of schools that utilized the system as their primary web presence or that linked to the system from their primary web presence, while district representation rates were determined by the adoption of a single school within the district. For instance, if a web system was adopted by two schools in the same district, it was counted as being adopted by two schools and one district, even if additional schools in the district did not also adopt the system.

**Figure 1: Popular web system adoption and representation**
As Figure 1 illustrates, adoption rates of systems varied both between systems and in the relationship between school adoption and district representation. The most utilized systems by schools in the state included Google Sites, PowerSchool, Blackboard, and Wordpress, while the most utilized systems by districts in the state included PowerSchool, Google Sites, Wordpress, and Weebly. Differences between school and district representation rates reveal that some systems may be used by only one or two schools within their adopting districts, like Drupal and Blackboard, while other systems may be used by many schools within comparatively few districts (like Blackboard or Wix).

As stated previously, we distinguished between two types of web systems (i.e. generic systems and education-specific systems), and we utilized two additional categorizations to understand issues of cost and openness. We used these types and categories to label each identified web system (as determinable by information obtainable about the web systems online), and these labels may be found in Table 1 along with more detailed adoption rates.

Table 1: Popular web system identification and adoption

<table>
<thead>
<tr>
<th>Web System</th>
<th>Type</th>
<th>Cost</th>
<th>Openness</th>
<th>Districts Represented</th>
<th>Schools Adopted</th>
</tr>
</thead>
<tbody>
<tr>
<td>PowerSchool</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>48 (36%)</td>
<td>210 (29%)</td>
</tr>
<tr>
<td>Google Sites</td>
<td>General</td>
<td>No Cost</td>
<td>Proprietary</td>
<td>35 (26%)</td>
<td>110 (15%)</td>
</tr>
<tr>
<td>Wordpress</td>
<td>General</td>
<td>No Cost</td>
<td>Open-source</td>
<td>32 (24%)</td>
<td>59 (8%)</td>
</tr>
<tr>
<td>Weebly</td>
<td>General</td>
<td>No Cost</td>
<td>Proprietary</td>
<td>21 (16%)</td>
<td>34 (5%)</td>
</tr>
<tr>
<td>Blackboard</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>20 (15%)</td>
<td>133 (18%)</td>
</tr>
<tr>
<td>Blogspot</td>
<td>General</td>
<td>No Cost</td>
<td>Proprietary</td>
<td>17 (13%)</td>
<td>37 (5%)</td>
</tr>
<tr>
<td>SchoolLoop</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>14 (10%)</td>
<td>61 (8%)</td>
</tr>
<tr>
<td>Moodle</td>
<td>Education-Specific</td>
<td>No Cost</td>
<td>Open-source</td>
<td>13 (10%)</td>
<td>23 (3%)</td>
</tr>
<tr>
<td>SchoolNet</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>13 (10%)</td>
<td>31 (4%)</td>
</tr>
<tr>
<td>Joomla</td>
<td>General</td>
<td>No Cost</td>
<td>Open-source</td>
<td>12 (9%)</td>
<td>59 (8%)</td>
</tr>
<tr>
<td>MyBigCampus</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>10 (7%)</td>
<td>20 (3%)</td>
</tr>
<tr>
<td>Drupal</td>
<td>General</td>
<td>No Cost</td>
<td>Open-source</td>
<td>8 (6%)</td>
<td>9 (1%)</td>
</tr>
<tr>
<td>Brainhoney</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>5 (4%)</td>
<td>8 (1%)</td>
</tr>
<tr>
<td>InfiniteCampus</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>5 (4%)</td>
<td>37 (5%)</td>
</tr>
<tr>
<td>Webs.com</td>
<td>General</td>
<td>No Cost</td>
<td>Proprietary</td>
<td>5 (4%)</td>
<td>6 (1%)</td>
</tr>
<tr>
<td>SharpSchool</td>
<td>Education-Specific</td>
<td>Cost</td>
<td>Proprietary</td>
<td>3 (2%)</td>
<td>11 (2%)</td>
</tr>
</tbody>
</table>
Adoption rates not only varied between schools, districts, and web systems, but also across Idaho’s six educational regions. This variation may be the result of a number of factors including vendor marketing patterns, the size of school districts in terms of enrollment and number of schools, the diversity of school districts within a given region, and the types of schools served in the region (for example, rural schools may have different needs, capacities, expectations, and opportunities than urban schools). Illustrations of adoption rates within all six regions of six selected web systems are provided in Figure 2 and Figure 3.

Figure 2 illustrates raw school adoption rate within a region, while Figure 3 considers districts as containers for schools and illustrates adoption rates of districts. Each illustration is useful for understanding certain aspects of state web system adoption, and comparisons of these two illustrations may be useful for understanding actual adoption patterns.
One issue that arises upon considering adoption of web systems is top-down versus bottom-up (or grassroots) adoption. In a top-down model, districts, regions, or the state may provide schools with access to a web system, while a bottom-up adoption pattern would begin at individual schools and scale up to the district, region, or state. Using the provided illustrations, it seems that some systems follow a top-down model while others are more grassroots.

Consider the differences between adoption of Blogspot and adoption of Blackboard. Blogspot is utilized by 5% of all schools in the state, these schools represent 13% of all districts in the state, and they are diffused fairly evenly across all six regions. This suggests that Blogspot is being adopted at a school level (i.e. bottom-up) by innovators at schools spread across the state. Blackboard, on the other hand, is adopted by 18% of all schools in the state, but these schools only represent 15%.

When contrasted in this way, it appears that of these major web systems, Weebly, Blogspot, Google Sites, and Wordpress all follow a bottom-up pattern of adoption, while Blackboard, PowerSchool, and SchoolLoop follow a top-down pattern. In the case of bottom-up systems, diffusion seems to be fairly uniform across regions and district representation exceeds school adoption (meaning that the schools adopting these systems come from districts in which other schools are not adopting them). In the case of top-down systems, adoption is isolated or exaggerated in select regions, and school adoption exceeds district representation (meaning that adopting schools are clustered in a few districts).

Identifying popular web systems and recognizing adoption patterns in this way helps us to recognize that there is a high degree of experimentation and innovation occurring in the state with regards to web system adoption and that some systems might be more appealing to schools at a grassroots level than others. However, we have not yet considered aspects of the systems themselves that may influence adoption patterns, and in the next section we will take into consideration such characteristics.
5 TYPES AND CATEGORIES

Based on the identification of major web systems, we will now explore breakdowns by the type of system and categorizations of cost and openness. Because schools have a variety of tasks they need to accomplish via the internet, it might be expected that schools would utilize a variety of tools and that some of these tools would be education-focused and that some would be more generic. Around 44% of schools used a generic system like a content management system (e.g., Wordpress) or a website hosting platform (e.g., Google Sites) and around 76% used an education-specific system like a learning management system (e.g., Blackboard) or a student information system (e.g., PowerSchool). There is also a large overlap between generic systems and education specific systems, as 20% of schools used both (cf. Figure 4).

![Figure 4: Types of systems used by schools](image)

Given this breakdown, it seems that schools clearly have a need for both generic systems and education-specific systems and that there are also needs that each of these types of systems alone cannot address. Just as a school might use something like Blackboard to manage student learning, it would likely also need a content management system of some sort to manage its communications with students and parents. Though some systems may attempt to address all
of these needs in one package, the fact that schools are evenly split between generic and education-specific systems suggests that neither type provides a clear and simple all-in-one solution for schools.

Further, cost seems to be a major factor in exploration, utilization, and adoption of web systems. A breakdown of web system adoption based on cost reveals that around 44% of schools utilize a no-cost system and around 75% of schools utilize a purchased system, with 19% utilizing both (cf. Figure 5). Similar to findings on types of systems, there seems to be a clear need and interest for both types of systems, but there might be more to this, because there is a clear advantage to no cost systems over purchased systems for budgetary reasons. Nonetheless, purchased systems maintain a strong presence, suggesting that either no-cost systems serve a different purpose for schools than purchased systems afford or they are not fully up to the task of replacing purchased systems.

![Figure 5: Cost of systems used by schools](image)

By further breaking down no-cost systems to types as illustrated in Figure 6, we discover that 97% of all schools that only use no-cost systems do not use any education specific system. This suggests that interest in no-cost systems is highly connected with interest in generic systems like content management systems and that either education-specific no-cost systems may be lacking in comparative functionality to their purchased counterparts or that schools that use generic systems do not recognize the need for education-specific systems as well.
Figure 6: Types of no-cost systems used by schools that do not purchase a system

This suggests that there is great interest in no-cost solutions but that schools are only finding widespread value in no-cost solutions to meet generic needs.

Contrarily, if we consider the schools that only use purchased systems, we find that these schools do not use generic systems at all but rather utilize their education-specific systems to address generic needs as well, like web hosting, or do not use systems for these purposes at all.

Openness is the third category by which K-12 web systems were considered. As depicted in Figure 7, the vast majority of web system utilization by schools is proprietary in nature (i.e. the source code is not open, and licensing prevents editing, revising, modding, etc.). Only 26% of schools use some type of open source system, and only 6% use only open source systems. This may suggest that openness and the ability to share, remix, and adapt web systems to the specific needs of schools may not be a top priority for schools when considering a web system for adoption but rather may merely be a byproduct of the adoption of no-cost solutions. In other words, systems like Moodle, WordPress, and Joomla may be adopted or rejected by schools solely on the basis of their cost rather than openness. If so, then this likely means that very few schools approach web systems with the intent of modifying the system on the level of code to meet their specific needs or to share local adaptations with the larger community.
A look at school adoption percent by region reveals that some regions find far greater value in no-cost web systems versus purchased web systems than others and that some regions find little to no value in open source web systems at all (cf. Figure 8). Region one is of particular interest, because it relies much more heavily upon purchased, proprietary web systems in comparison to no cost and open-source systems than the other districts.
Reasons for these regional differences are not clear, but the rural/urban nature of each region, inter-region collaborations, accessibility of supporting resources (e.g., university partnerships), and comparative wealth of districts within each region may have an impact. What is clear, however, is that both no-cost and purchased systems are widely used and that proprietary systems are much more heavily adopted than their open-source counterparts.
CONCLUSION

This report is intended to provide a high-level overview of web system adoption in K-12 schools throughout the state of Idaho as determinable by heuristic analysis of schools’ public websites. Some overarching conclusions that may be of interest to decision-makers at local, district, and state levels may include the following:

1. Schools in the state have adopted a variety of web systems, which range from being very generic in nature to being very education-specific, and there seems to be clear interest in utilizing both types of systems in conjunction with one another.
2. There seems to be great interest in no-cost web systems, but such systems have not supplanted purchased systems, which primarily seem to serve education-specific goals, while no-cost systems are primarily generic in nature (e.g., web hosting, blogging, content management).
3. The openness of a system does not seem to be a major factor in web system adoption, and currently adopted open-source web systems may more generally reflect schools’ interests in no-cost solutions rather than commitment to openness and sharing.
4. Adoption of systems and types and categories of systems varies by region and may reflect a variety of factors including wealth, capacity, connections to other institutions (e.g., universities or other schools), top-down adoption from district leaders, and so forth.

These findings have a variety of implications for decision-makers. First, it seems clear that since schools utilize such a variety of systems that may reflect local needs and will have very different experiences and local capacities for utilizing web systems, a single solution or set of solutions for all schools seems problematic at best. Reading this positively, this diversity of local use may reflect a high level of local innovation for adopting and utilizing technologies that are beneficial in local contexts, and such variability perhaps should be applauded as indicative of schools’ willingness to innovate without top-down support.

Second, it seems that some system adoption patterns reflect bottom-up diffusion while others reflect top-down diffusion from districts. It is beyond the scope of this report to make suggestions on suitable policies for supporting innovation along these two methods, but it seems that some systems are readily being adopted across regional and district lines by schools that have no connection with one another and that these types of systems would likely gain a large amount of buy-in from local districts if they were part of top-down projects intended to support large-scale innovation.
Third, the lack of adoption of open-source systems may reflect both lack of local capacity to adapt web systems to support local needs and also may reflect a hurdle to large-scale innovation in the sense that local school adaptations of web systems are not made with the intent or in a manner conducive to sharing with the larger community. This means that if we want local innovation to scale up between schools, across districts, and through the state, it seems that the development of local capacity in the area of web system adaptation should be a priority.

And fourth, regional differences in adoption should be considered both to determine sources of innovative practice in the state and hurdles for scaling innovation. While the adoption of a particular web system or innovative practice in one region might have clear benefit for statewide application, regional challenges and barriers are real and seem to influence web system adoption within and between regions.
REFERENCES


