Crowdsourcing Curriculum Design

Washington I. Helps
Department of Computer Science
University of North Dakota
Grand Forks, ND 58201
washington.helps@gmail.com

Emanuel S. Grant
Department of Computer Science
University of North Dakota
Grand Forks, North Dakota, 58201
grante@cs.und.edu

Abstract

Curriculum design and development is a complex task, and in today's highly connected and digitized world, taking advantage of the tools available to revolutionize this task is a worthwhile endeavor. The purpose of this work is to outline a model for the creation and maintenance of a platform to house a repository of teaching artifacts to include the core content of a given curriculum such as the courses and topics along with supporting content such as lectures, case studies and assessments. The underlying idea behind this platform is that its creation, development and survival will be driven by experts in the given domain along with other relevant contributors including educators who will contribute all the content in an international collaborative setting. Essentially, this is a quality assured crowdsourcing venture, and in the educational space quality must always trump quantity.

This platform will be available to all educators, faculty and universities and because of the inherent attributes of a platform with diversified contributions, curricula will be more flexible, context aware and less biased. Therefore, users will be able to design a curriculum that suits their individual needs and circumstance, while at the same time there is the assurance that the curriculum and associated content will meet certain minimum quality requirements. Developing countries which often lack expertise can have ready access to this pool of resource made available by the best minds in the given field.

This work will present a solution to manage contributions to this platform and explore the feasibility of such a maintenance model which is restrictive enough to ensure only the best submissions get published to the repository, while being relaxed enough to allow for sustained growth of the repository. A solution for the structuring of a curriculum within the platform will also be presented which will allow for efficient resource discovery and the implementation of algorithms that allow for dynamic adaptive selection of course and topic when designing a curriculum.

1 Introduction

Curriculum design is as ubiquitous to education as classrooms are to schools. Yet, despite this fact, the methods and processes of handling this activity remains very antiquated and disjointed. The technologies available today provide many avenues through which this very important activity can be streamlined, made highly available and very open. The motivation behind this paper is not merely to outline another application of the myriad technologies and methodologies available, but also to address very fundamental issues in the higher educational system worldwide.

This work will outline a model for the creation and maintenance of a platform to house a repository of teaching artifacts to include the core content of a given curriculum such as the courses and topics along with supporting content such as lectures, case studies and assessments. The underlying idea behind this platform is that its creation, development and survival will be driven by experts in the given domain along with other relevant contributors including educators who will contribute all the content in an international collaborative setting. Such a model is essentially a quality assured crowdsourcing venture.

1.1 The Curriculum

A curriculum can be described as the embodiment of a program of learning consisting of courses of study in a school and includes but not limited to content, approach and assessment. A curriculum can also be referred to as any document that exists in a school that defines the work of teachers by identifying the content to be taught and the methods to be used [1]. In most educational settings, such a document even in a digital form is quite static, and not flexible to the needs of the beneficiaries. Even the very concept of a curriculum is very limited and not dynamic enough to adapt to the changing needs of the students and the professions they are being prepared for.

In [2], V. Kaprielian suggests that a curriculum is always a work in progress. So it is important to keep in mind the fact that as the environment in which the curriculum is used changes or the market that is being satisfied changes, so too must the curriculum. Failure to keep up with the times and changing needs can result in the institution or stakeholders involved descending into oblivion or obsolescence. Again, this is where the crowd-sourcing model of having continuous contribution, maintenance and development of the content supporting a curriculum will be beneficial.

The dynamic nature of curricula is even more apt in the computing fields and other STEM areas where technology is constantly improving and evolving and in much the same way a curriculum must evolve. This is why the process is really a cycle where following evaluation there should be review which may lead to changes in the curriculum to make it current. Figure 1 clearly highlights the cyclic nature of curriculum design in which it becomes apparent that a curriculum is a dynamic artifact which is constantly

evolving. This is a characteristic which ought to be harnessed especially in the STEM disciplines.

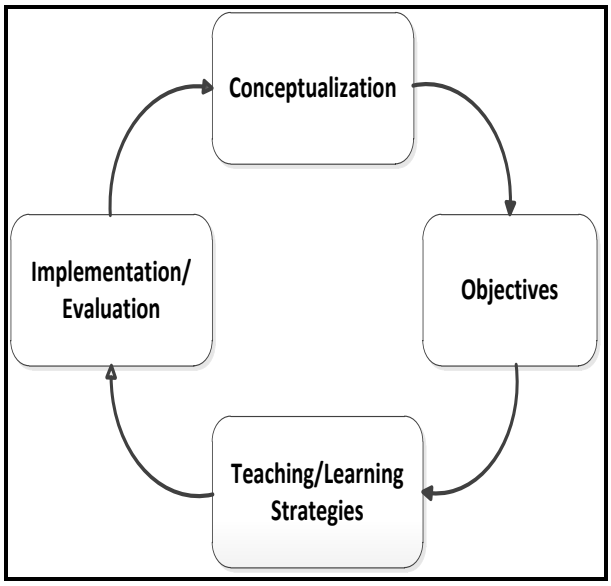

Figure 1: Curriculum Design Process Cycle.

1.2 Motivation

In many developing countries there is a scarcity in classroom human resources as well as physical resources for teaching which can be viewed as a barrier to standardized world class curricula. The scarcity in human resources leads to high student to teacher ratios, a scenario which is very undesirable in higher education institutions.

As far as having a standardized curriculum in a given domain goes, there is a need for there to be collaboration from domain experts and teachers worldwide such that a common pool of resources can be established and made accessible. A framework that is set up in the proposed way will allow for a common pool of resources to be available internationally, thus strengthening the collaborative efforts between higher learning institutions and the sharing of knowledge.

Another motivation for this work surrounds the importance of having curricula in a given domain are of a high standard regardless of where it is delivered. It can be argued that a degree program in the same domain offered in different regions of the world can have differing outcomes and quality. Therefore it is important for there to be a minimum quality threshold that is quantitative as well as qualitative.

The implementation of the proposed framework facilitated by content contribution via a crowd-sourced model will provide the benefits of having curricula that is unbiased and which is not reliant on any one person, region or group. The framework will also provide added value in the way of automated assessment generation from the repository content.

2 Related Works

In this section, several papers will be briefly summarized to highlight efforts deemed relevant to this work.

2.1 Predictive Analytics Reporting

In this section, several papers will be briefly summarized to highlight efforts deemed relevant to this work. To start off, mention must be made of a very interesting undertaking involving the collaboration of universities and their willingness to contribute student records for a good cause. The Predictive Analytics Reporting (PAR) Framework was developed to identify factors that influence student retention and progression, and to guide decision-making designed to improve postsecondary student completion in the United States [3]. PAR now includes more than 1.7 million anonymized and institutionally de-identified student records and 8.1 million course-level records. This initiative relies on institutions to supply the students' records, without which the framework would be useless. Essentially, PAR is a collaborative multi-institutional data mining effort geared at improving student success. This is a wonderful example of the power of many coming together for a common cause. This is the beauty of crowd-sourcing, and in the education sphere, PAR has demonstrated this.

2.2 Joint Information Systems Committee

The Joint Information Systems Committee (JISC) [4] in the United Kingdom is in charge of several projects at different universities that are looking into transforming curriculum design and delivery through technology. The Higher Education Funding Council for England [5], also based in the UK is doing a similar sort of work. A committee based in the United Kingdom (UK) has embarked on an ambitious program to tackle the problem of managing curriculum change. This committee which is called JISC is seen as the UK's expert on information and digital technologies for education and research. JISC initiated a four-year program which started in 2009 to investigate how processes involved in the design of programs of study can be made more agile and responsive through the use of technology. This program comprises 12 projects led by teams in UK universities.

Key visions for curriculum design as outlined by JISC include the ability to respond creatively and flexibly to changing cultural and economic climates as well as learning resources that are searchable, accessible and sharable. This is the essence of the proposed crowd-sourced curriculum framework in this paper, even though the implementation differs in that JISC's focus is on an end-to-end integration of technology in the curriculum processes by utilizing various tools and pre-existing systems.

2.3 Internet Based Education Model for Caribbean Countries

In 2000, Grant et al. published a paper entitled "Towards an Internet-based Education Model for Caribbean Countries" which looked into leveraging the usage of low-cost technology to offer Virtual Classroom environments to teaching institutions in remote areas via the Internet. The work in [6] is mentioned here because it echoes one of the motivating factors for this work. In developing countries there is a high student-teacher ratio and further still within those countries, the majority of high-quality teachers and teaching facilities are located in the urban areas.

One of the main motivating factors in [6] is to ensure that there is the capability to provide consistent, high-quality educational materials to students across a wide geographical area, in a timely manner and with limited human and physical resources. Some of the benefits highlighted in the paper include:

- The availability of high quality teaching materials and teaching methodologies to geographically remote schools in developing countries.
- Rural schools can now obtain the same material and pedagogical skills which are available to their counterparts in the best urban schools.
- The negative effects of a large student/teacher ratio can be suppressed and not extend to students utilizing Virtual Classrooms in a great way, because of the required concentration and attention demand on such students.

Figure 2 depicts this contrast beautifully where the Virtual Classroom Over the Internet (VCOIN) on the right side of the diagram shows at a low level, the physical setup of the infrastructure.

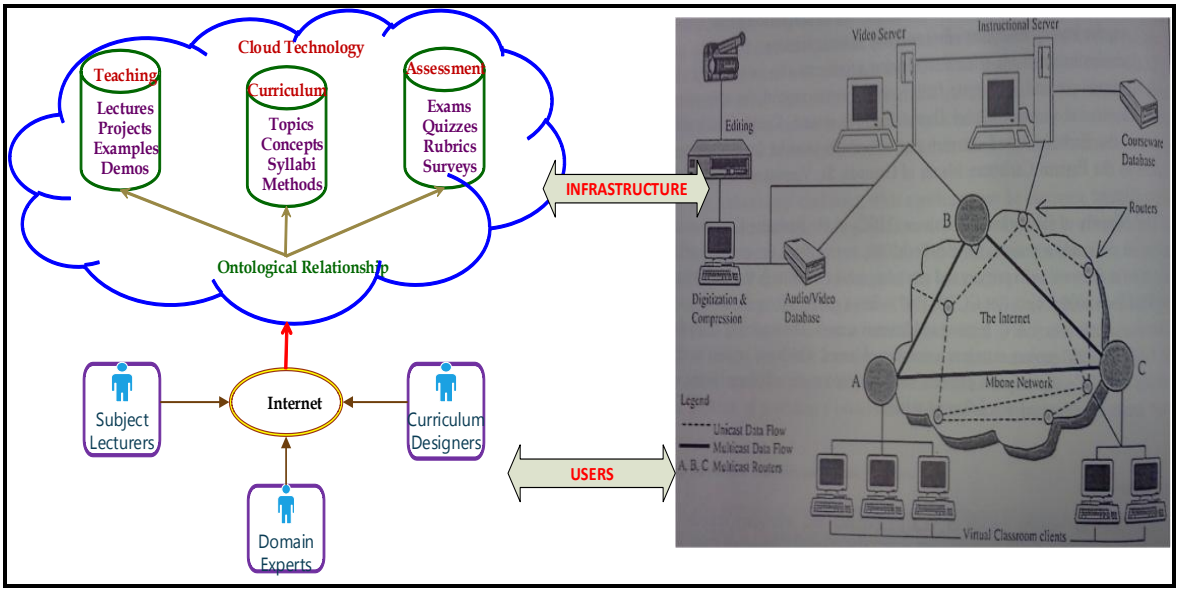

Figure 2 – Diagram showing User and Infrastructure View of the proposed framework being contrasted with the architecture of VCOIN as seen in [6]

3 Research Methodology

The scope of this work places focus on the teaching and learning aspects of a curriculum. At the heart of this framework is the Repository which describes the common pool of teaching artifacts and resources. This repository will house all content which can either be core content or supporting content. Core content is content consisting of essential material required for the delivery of the given degree or field of study and can include topics, modules, courses and knowledge areas [9]. Supporting content on the other hand is content which complements the core content and include material such as lectures (audio, video or written), case studies, projects and assessment questions.

The methodology will be broken down into three distinct phases:

1. A model that captures the building of the repository
2. The crowd-sourcing model
3. The modeling of a curriculum

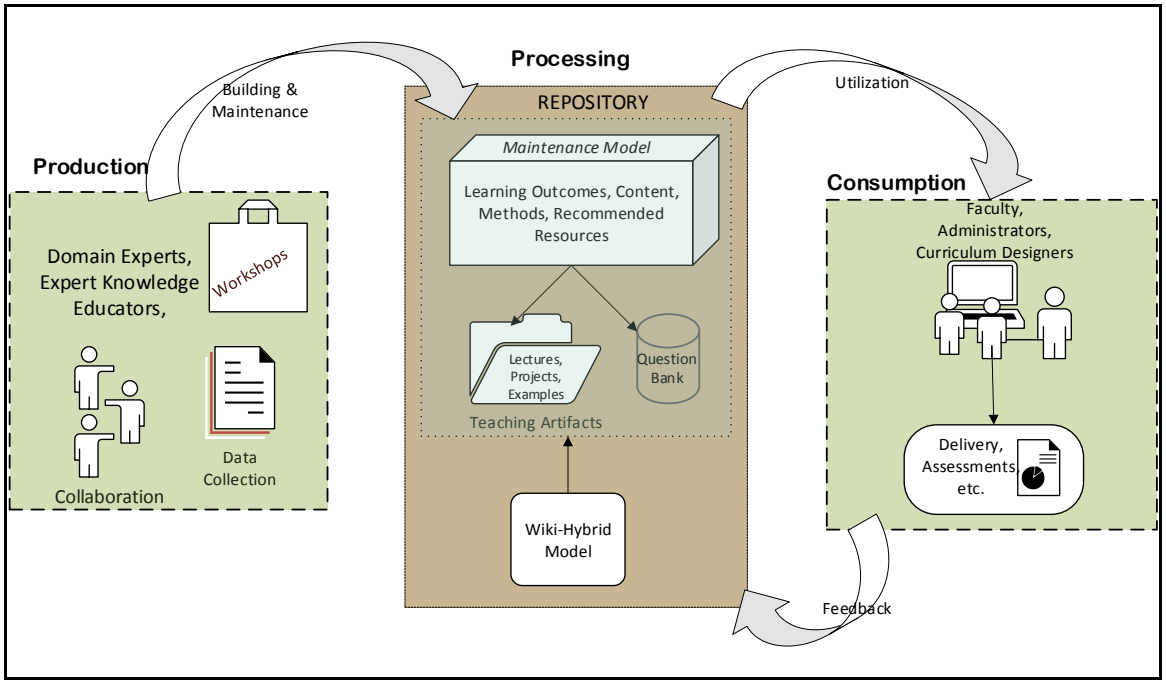

Figure 3 – Overview of the Framework's contribution and utilization

Figure 3 gives a broad overview of what the methodology will entail. The Production component details the framework for the initial building of the repository and its future maintenance. This will build on the work already done by Grant et al. in [7][8] in order to clearly outline a path to the realization of the repository. Processing is the most involved of the three components and this is where an attempt is made to model a curriculum given the input from the first component. Finally, the Consumption component demonstrates the utilization of the repository and the modeled curriculum, such as a faculty designing a program of study and accessing various teaching artifacts, the delivery of the content, student assessments as well as automated feedback to the repository. Such feedback can

be stealthy in nature where various statistical and anonymous data are uploaded to the repository. Of course, for countries like the US, this approach would steer clear of any potential legislative issues like FERPA (Family Educational Rights and Privacy Act)[1]. Only aspects of Production and Processing components will be covered in this paper.

3.1 The Repository Development

The foundation of the proposed framework must be put in place before and crowd-sourced contribution or maintenance efforts can be considered. This is because there must be a clear understanding of what it is that contributors will be contributing to and furthermore, a rigid structure must be put in place. The major aspect of the production component is the collaboration of various domain experts and educators in the given field. Such collaboration is crucial to the successful implementation of such a framework and similar approaches such as those undertaken in the IEEE/ACM 2004 Development Process in [9] have proven to be very useful.
This collaboration has already started as outlined in [8] with the workshops at HAU in the Philippines and the ongoing widening of that pool of contributors since then. It is important to mention the workshop because it shows the approach that should be taken to develop curriculum. It shows the beginnings (foundation) of a platform that other domains can emulate. It lays the groundwork required for the building and future maintenance of the repository which lies at the heart of this new paradigm.

3.2 The Crowd-Sourced Model

As mentioned earlier, a curriculum is always a work in progress and inevitably changes over time. Therefore, any framework that supports curriculum design and development must be inherently dynamic, flexible and fluid. In order for the repository to be maintained and continuously developed, a structure must be put in place to ensure this. One approach being explored is one that is similar to the Wikipedia[2] model. Wikipedia is a web based encyclopedia made up of a large number of interconnected web pages. The idea behind this is that it is a community of users who develop and maintain these pages; adding, updating and editing the content. Wikipedia's content is written by volunteers. According to the site, "anyone with Internet access can write and make changes to Wikipedia articles (except in certain cases where editing is restricted to prevent disruption or vandalism)" [10]. The proposed framework will build on the positives of this Wikipedia model and adapt it to this new paradigm where experts across the world can contribute to the repository and participate in its continued maintenance and development.

The open nature of Wikipedia means that not much can be done to prevent the publishing of inaccurate, incomplete or biased information in the first place. Therefore, much of

1 http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html
2 http://en.wikipedia.org/wiki/Wikipedia:About

Wikipedia's quality control can be viewed as reactive stemming from the constant editing that all of Wikipedia's content is subject to. In the open web, this is an accepted limitation, provided that consumers of this content know that what they are getting has no guarantees.

It is important to point out that Wikipedia was predated by Nupedia , a project to produce a free, open source, collaborative online encyclopedia which lasted from March 2000 until September 2003. In fact, Wikipedia was founded as an offshoot of Nupedia. Nupedia had an elaborate peer review system, encompassing a rigorous seven-step approval process to control content of articles and which required highly qualified contributors [11]. So arduous was this review system that it ultimately led to the demise of Nupedia. Nevertheless, its motives were pure, in that the aim of Nupedia was to publish articles of a very high quality, comparable to professional hard-copy encyclopedias at the time. Figures 4 and 5 below demonstrate the contrasting review processes between Wikipedia and Nupedia. Figure 4 shows the seven-step editorial process, while figure 5 shows the simplicity in the Wikipedia publishing flow with reactive editorial review.

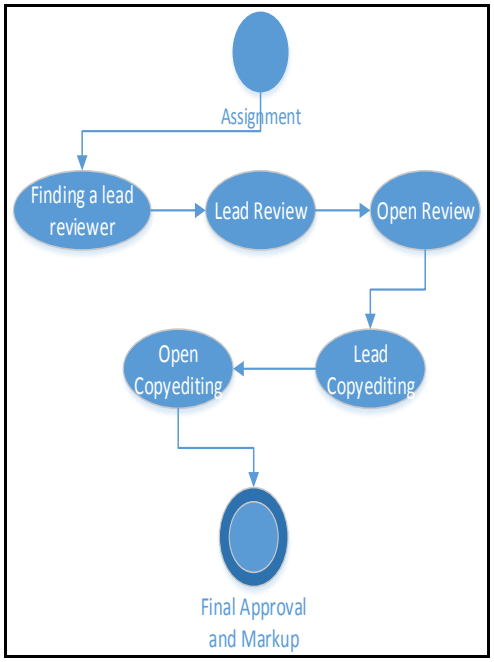

Figure 4 – Nupedia Editorial Process

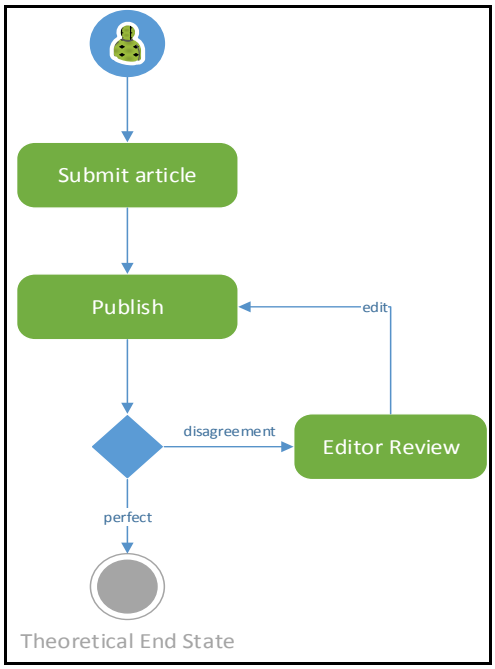

Figure 5 – Wikipedia Publishing Flow

3.2.1 Two Contrasting App-Store Models

It this section, two contrasting platforms are examined to find meaningful patterns that the proposed framework can borrow. The Google Play Store and Apple's App Store (iTunes) provide such a contrast. Google's Play Store is the marketplace for Android device applications, and with Android being an open source operating system is it not surprising that the model Google has employed for publishing apps is open distribution

which is unrestricted. To get an app published in the Play Store, one only needs to have a Google account and as long as all the basic requirements are met such as having the correct image sizes and a valid APK[3], the app once submitted is available in the store in a matter of minutes. In the case of Apple on the other hand, iOS app submissions are subject to approval by Apple and can take weeks for the approval process to be completed.

There are advantages and disadvantages to the approaches used by either company. In Google's case, the main disadvantage is app quality and the possibility of malicious apps that may contain malware or other nefarious code. The openness has the benefits that developers do not have to feel restricted and confined to the ideals and interests of Google. For Apple, the reliability testing and analysis that is done in there strict approval process ensure the approval of only high quality apps which enhance user experience and does not degrade the platform. The downside for Apple's approach is that developers' creativity may be stifled somewhat and they are left at the 'mercy' of Apple. There have been many controversial cases where Apple took steps to deliberately delay approval for apps that either competed against Apple or that were not in the financial best interest. Apple is often seen to be hypocritical and seen to be an authority that censors app submissions only when it benefits their own interest.

3.2.2 The Wiki-Hybrid Model

The strengths of the publishing quality of Nupedia can be coupled with the openness of the Wikipedia model to create a more streamlined model suited to the proposed framework. With the app stores, Google's Play Store openness leads to significantly more security flaws and quality issues which can only be remedied after the fact (reactive), similar to how Wikipedia operates. On the other hand, Apple takes a more proactive approach ensuring that bad apps never make it into the marketplace, an approach akin to the now defunct Nupedia. The proposed framework would take advantage of Google's and Wikipedia's open approach by welcoming collaboration and contribution from a wide community and consortium of users, and take advantage of Apple's closed ecosystem ensuring that all contribution are properly and efficiently vetted before they have any impact on the repository.

There will be a need to have stricter control core content such as courses and topics. On the other hand, supporting content such as assessment artifacts will not need to be as tightly guarded and so lower levels of restriction would suffice. Figure 6 is a proposed crowd-sourcing model for maintenance and development of the Repository through active collaboration and contribution.

3 http://developer.android.com/distribute/open.html

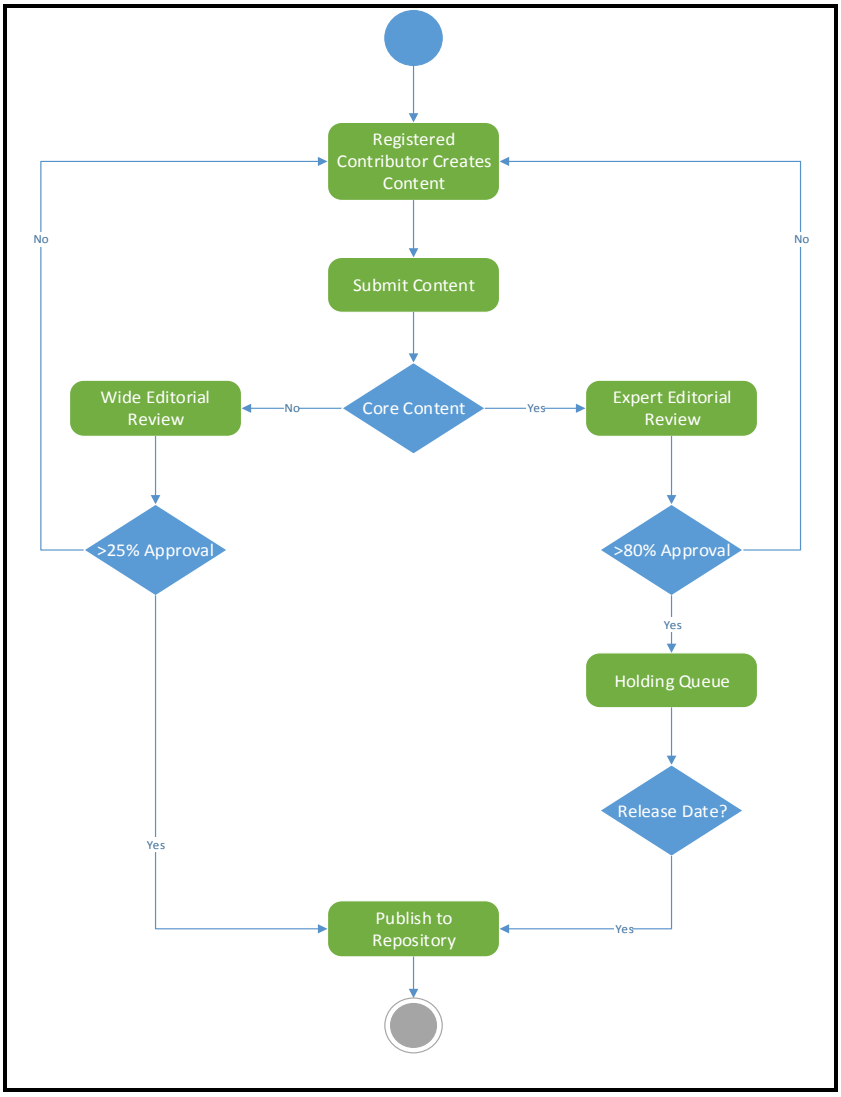

Figure 6 – Proposed Model for Repository maintenance

3.2.3 Description of Maintenance Model

1. Core content must be evaluated by experts (for example, contributors with PhDs only). This step is similar to the review process of Nupedia or Apple's app store, but far less cumbersome. This review step goes off the simple premise that there is strength in numbers; where the greater the number of experts that agree on something, the greater its quality and value. A set threshold of expert approval must be surpassed in order for a submission to be accepted.

a. This is much like a Peer Review where all the editors have equal weighting and if for instance, there is high consensus among the reviewers, then the submission is accepted.
b. If this threshold is not met, the submission fails and is sent back to the submitter. Just to point out how streamlined this review step is, a submission that needs to be edited for whatever reason can be denied by an editor and if enough editors deny this submission, it will not meet the threshold and will eventually end up back with the submitter to make the corrections. So the theme of collaboration and general consensus is what makes this approach practical and quality-controlled. Compared with the Wikipedia model, this approach ensures that only quality content gets published since the review is done before the fact and not after.

c. One possible limitation of this approach is that it may not scale well. This is because too many core content submissions may result in backlog and longer approval times. A possible fix for this limitation would be to require a fixed number of expert editors for threshold approval, rather than requiring a fixed percentage or a hypothetical "high approval". But this can only be done if there are enough experts available, such as setting an acceptance threshold of 20 expert approvals if there are 50 total experts instead of a fixed percentage like 80% which would require approval of 40 out of the 50 experts.

2. Supporting Content is content which has lower precedence than Core content and as such can be reviewed by any registered contributor. The acceptance threshold for acceptance can therefore be more relaxed and flexible. A possible acceptance threshold could be to require moderate approval rates (such as hypothetically having at least 20% of all registered contributors approving it. Fast approvals would be suited as it would lead to rapid growth of the repository from a supporting content perspective. Accepted submissions are immediately published to the Repository.

3. Accepted Core content submissions are held in a container called the Holding Queue. This is to ensure that changes to the Core are not made too frequently, thus ensuring stability. Content in this queue are published at fixed times throughout the year such as every 6 months or at the end of every semester. This is akin to already released software that has a fixed release cycle to provide updates or patches.

3.3 The Modeling of a Curriculum in the Repository

The structure representing the logical groupings of individual pieces of data in the repository is very important. This is because this structure or modeling gives both syntax

and meaning to the content of the repository and by extension the curriculum it embodies. This section will give a brief overview of how curricula can be modeled in the framework by utilizing various data structures.
Software Engineering will be used here as a prototype degree program to demonstrate the modeling in the Repository. Consider the following terms:

- Software Engineering Education Knowledge (SEEK): the body of knowledge that is appropriate for an undergraduate program in software engineering.
- Knowledge Area: a particular sub-discipline of software engineering such as Software Design or Software Management.
- Course: a subset of a knowledge area. This term is used interchangeably with the word unit.
- Unit: a subset or module of a knowledge area. This term is used interchangeably with the word course throughout this paper.
- Topic: the lowest level of the SEEK hierarchy. A topic is a subset of a unit.

Given that a curriculum consists of the above listed entities that can be represented hierarchically, it is therefore feasible to model such entities and the relationships among them as an ordered tree. A general tree is a nonlinear data structure in computing in which each node may have zero or more children. It turns out that it is more suitable to model the curriculum as a linear tree because there is implicit ordering in the relationships both horizontally (siblings) and vertically (parents-children). This ordering exits because for instance, it makes sense to teach a certain topic before another, or teaching all year 1 courses before year 2 and so on. Trees provide a natural organization for data and as such have become ubiquitous structures in file systems, databases and other computer systems. The relationships in a tree are hierarchical, with the generic parent-child (family tree) layout, a theme that fits perfectly in the model outlined in figures 5 and 20. There are properties that every tree subscribes to; however, there are other properties that are specific but not unique to the curriculum model. These properties are outlined in figure 7.

Nodes that share the same parent are siblings, meaning that they are at the same level in the tree. Useful information can be gleaned from deciphering this property such as ranking the importance of a piece of content in comparison with other content. Representing a curriculum using the tree structure offers many benefits. For instance, for a given leaf node (topic) in the tree, a number of valuable information can be learned by tracing its ancestry.

Such an approach allows for the derivation of both implicit and explicit information from a simple traversal of the tree. When this model is tied in with the supporting content it is clear to see how intuitive the framework becomes. For example, if questions are linked with core content entities when they are submitted to the repository, then generating assessments becomes very easy and can be well targeted to specific content in the curriculum.

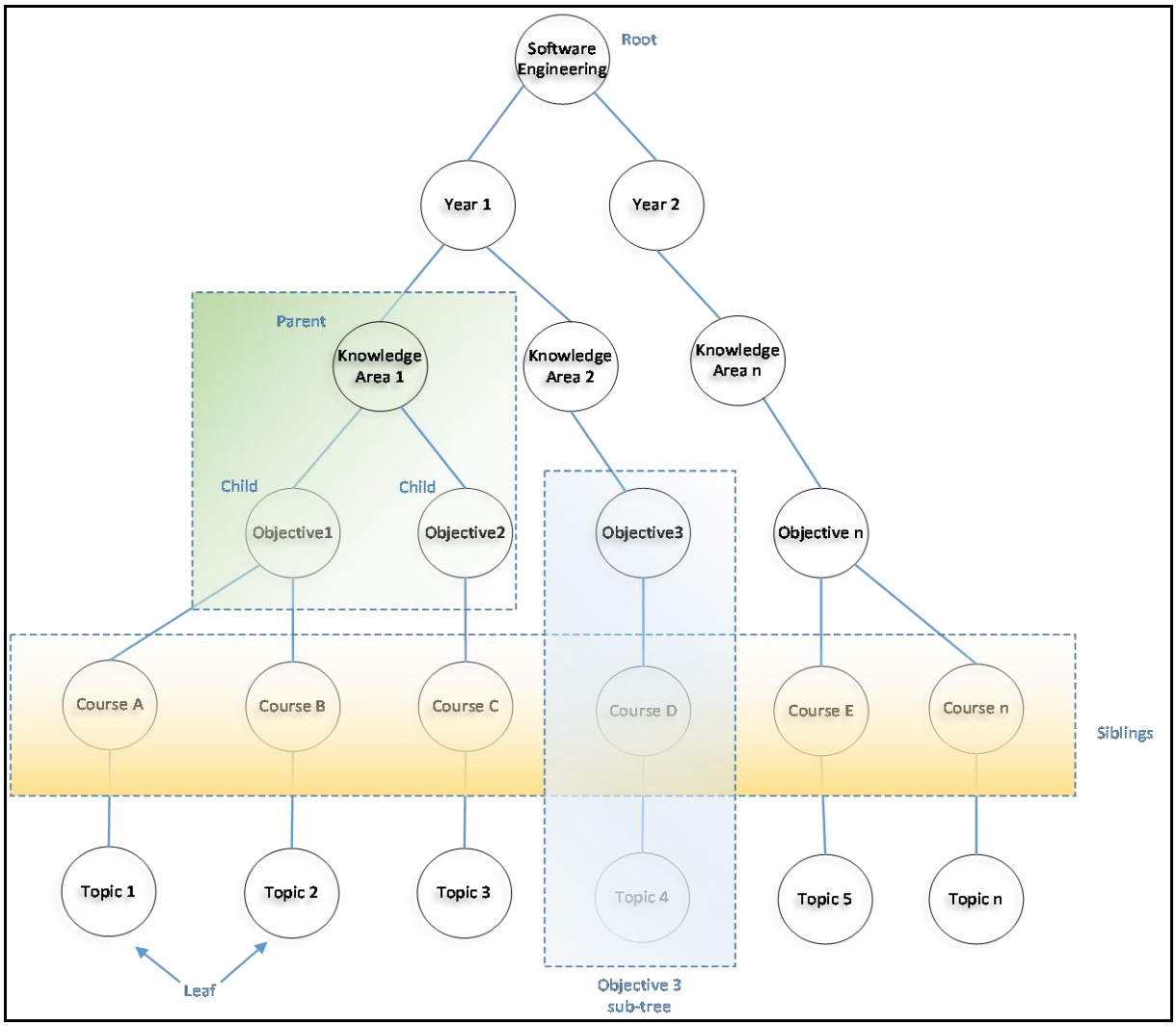

Figure 7 – Curriculum hierarchy annotated with properties of the Tree Data Structure

4 Conclusion

A new framework for the design and development of curricula was presented in this paper. This framework relies on a consortium of users to contribute content and expertise to a repository and a crowd-sourcing model that captures the essence of such collaboration was outlined. This paper presents a roadmap to a new and exciting forefront in the fusion of technology and knowledge to reap tremendous benefits in education worldwide. There is still much work and investigation needed to determine the interest and willingness of institutions, individuals, accreditation bodies and other authorities to get on board such an initiative in both the near and long term.

References

[1] English, F.W. (2000). *Deciding what to teach and test: Developing, aligning and auditing the curriculum*. Thousand Oaks, CA: Corwin Press.

[2] Kaprelian, V. S. *Faculty Development Module Instruction*. Duke University Medical Center 2002. Retrieved from http://cfmmodules.mc.duke.edu/curriculum/intro/process.html on March 20, 2012.

[3] Morrison, C. *PAR (Predictive Analytics Reporting) Framework*. Retrieved from http://wcet.wiche.edu/wcet/docs/par/PARFrameworkPOCSummary.pdf on March 21, 2013.

[4] Knight, S. *Institutional approaches to curriculum design*. Joint Information Systems Committee 2009. Retrieved from http://www.jisc.ac.uk/curriculumdesign on August 25, 2012.

[5] Browne, T., Hewitt, R., Jenkins, M., Walker, R. *2008 Survey of Technology Enhanced Learning for higher education in the UK*. Universities and Colleges Information Systems Association. University of Oxford, November 2008.

[6] Grant, E., France, R. Towards an Internet-based Education Model for Caribbean Countries. Journal of Education Media, Vol. 25, No. 1, 2000.

[7] Grant, E. *Online Collaborative Teaching of Software Engineering in an International Multi-Campus Environment*. First HAU International Workshop on Enhancing Teaching & Learning of Software Engineering, Holy Angel University Angeles City, Philippines August 12, 2011

[8] Grant, E., Stokke, T., Helps, W., Esquivel, J., Concepcion, W. *Enhancing Teaching & Learning of Software Engineering in an International Environment*. In Proceedings of the 2nd Annual International Conference on Computer Science Education: Innovation and Technology (CSEIT 2011), pages 36-41, 2011.

[9] Software Engineering 2004 Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. A Volume of the Computing Curricula Series, August 23, 2004

[10] Wikipedia: About. Retrieved from http://en.wikipedia.org/wiki/Wikipedia:About on October 14, 2012.

[11] Lanxon, Nate (June 5, 2008). "The greatest defunct Web sites and dotcom disasters". CNET UK. p. 5. Retrieved from http://crave.cnet.co.uk/gadgets/the-greatest-defunct-web-sites-and-dotcom-disasters-49296926/5/ on January 21, 2013.