COMPUTER SUPPORT FOR
A CROSS-DISCIPLINE RESEARCH CONSORTIUM

William B. Frakes, Gabriella Belli, Julián Urbano* and Reghu Anguswamy

Virginia Tech, Haycock Rd., Falls Church, VA, U.S.A.

*University Carlos III of Madrid, Avda. Universidad, Leganés, Madrid, Spain

Keywords: Graduate education, Research support, Website support, Statistical packages.

Abstract: This paper describes the use of web based and freeware tools for a cross-discipline research consortium to support graduate education. The use of WordPress and Simple Machine Forums as the main environment to provide web support is discussed, as are the use of statistics tools such as R and SAS-JMP. Usage data for these tools are provided.

1   INTRODUCTION

This paper describes a multi-year effort on tools and services to support a cross discipline Research Methods Consortium for the Virginia Tech National Capital region (NCR-RMC). The purpose of the NCR-RMC is to provide better support for research at VT-NCR by coordinating research resources across colleges. This involves improved information importation and sharing for consulting support and coordinated teaching. The VT-NCR offers degrees at the masters and doctorate levels. There are now 6 colleges represented in the region: Architecture and Urban Studies, Business, Engineering, Science, Liberal Arts and Human Sciences, and Natural Resources. These colleges currently offer more than 45 graduate degrees and certificates. There are 18 recognized centers represented in the region.

The colleges employ about 105 full-time teaching, research and administrative faculty, and 50 support staff. There are currently around 1,000 students of whom about one third are full-time. The part time students are primarily working professionals who spend limited time on campus.

A preliminary survey produced a list of available courses and some information about requirements and needs. Many VT-NCR programs have a need for statistical and/or research methods instruction, which may or may not be offered within their own department or college. Also, research or analysis topics are sometimes included as part of content courses. We found that several programs have to rely on adjuncts or were in the process of trying to hire faculty to help meet the research needs of students. Occasionally, a methods course that students planned on taking was cancelled, forcing students to go to other universities. In at least one case, a viable course for students in one program could have been available from another program, but that information was not known. In another case, professors in two different colleges were teaching essentially the same multivariate statistics course with very small class sizes.

While certain specialized courses for particular needs in one program may not be suited to another program, some courses could serve multiple programs. Likewise, useful resources or materials could be shared if there was some way to do so efficiently. One problem was a lack of effective communication in a timely manner and of a good vehicle to share and/or obtain such information. We did an in-depth exploration of available resources and specific needs for statistics and research methods across VT-NCR programs. We determined that one way to support these needs was with online information, tools, and training. Section 2 describes these resources and reports measures of their usage. Section 3 presents the supporting activities of the consortium, section 4 discusses the collaboration activities and section 5 provides the conclusions of the paper.

2 COMPUTER SUPPORT

In this section we describe several types of computing resources that were used to support the research consortium effort.

2.1 Statistical Analysis Tools

We decided to primarily support three statistical analysis tools, R (The R Project for Statistical Computing, 2010), Microsoft Excel®, and SAS JMP (JMP Software, 2010). We chose R because it is a powerful, widely used, freeware environment and SAS-JMP because of its ease of use and the fact that it is supplied to all Virginia Tech students. We chose Excel because many of the students already knew how to perform basic tasks with it. These tools are used for thesis and dissertation research by students and research by faculty. These tools are also used to support classes.

To support training in R, we developed several screencast videos with CamStudio, containing tutorials. These videos were to be uploaded to Youtube, but due to its lack of video quality, we finally chose to use Google Video. The R training videos on Google Video are:

1. Introduction. This first video tutorial covers how to download R, variables and simple data types, operations, functions, and how to get help within R.
2. Complex Data Types. The second tutorial covers the basics of creation and modification of vectors, lists, data frames, and tables in R.
3. Indexes. The third tutorial covers indexing vectors, lists, data frames, tables, and indexing by conditions.
4. Data Import & Export. The forth tutorial covers how to import different types of data structures both from files and the clipboard. Exporting data is also covered.

As of October 13, 2009, in a year and a half of availability, these videos has been watched 350 times and downloaded 9 times.

Many places in our website are dedicated to these statistical analysis tools, with references to books and other tutorials, as well as other pointers to resource material.

2.2 Website

A website was created containing information about research courses, funding sources, cross-discipline research, and tools and methods for research. It can be found at http://rmc.ncr.vt.edu. The blogging platform WordPress (WordPress > Blog Tool and Publishing platform, 2010) was used for this website. It allowed us to add information and resources so that it is easy for readers to keep track of new content. Blogs are thought to be useful for this kind of situation as users can easily post feedback comments and also follow other comments. We chose WordPress because it is one of the easiest and most used free platforms. Moreover, the website supports RSS (Really Simple Syndication) syndication, allowing readers to be instantly notified of new content either in their e-mail or feed reader. At a given time, on average, 10 users are subscribed to this feed.

Despite being focused on the management, categorization and display of new content, WordPress also allowed us to have static information (i.e. lists of courses, funding sources, etc.) always visible.

Another point in favor of a platform like this, is the possibility of having multiple users adding new content at the same time which we considered a major advantage.

The website is running on a regular desktop computer, with a standard freeware LAMP platform, using Ubuntu Linux, Apache, MySQL and PHP. Being the most common platform on the Internet, there is plenty of documentation and its maintenance does not require much effort.

As of October 13, 2009, the website has been visited 6,897 times by 4,223 unique visitors, who displayed a total of 17,782 pages.

2.2.1 Mindmaps

In order to organize the information and make navigation easy for the user, we developed several mindmaps using FreeMind (FreeMind – free mind mapping software, 2010). Figure 1 shows a mindmap of the research resources and information available. Each of the topics in the mindmap is hotlinked to the relevant section in the website.

2.2.2 Forum

It was clear from the beginning that the purpose of the consortium was not just for its members to share resources, but also to debate and ask for help on any research-related topic. Of course, this is crucial from the student point of view, so we later decided to incorporate a web forum. It allowed us to discuss topics about statistics, writing, joint research groups and some others, as shown in Figure 2.
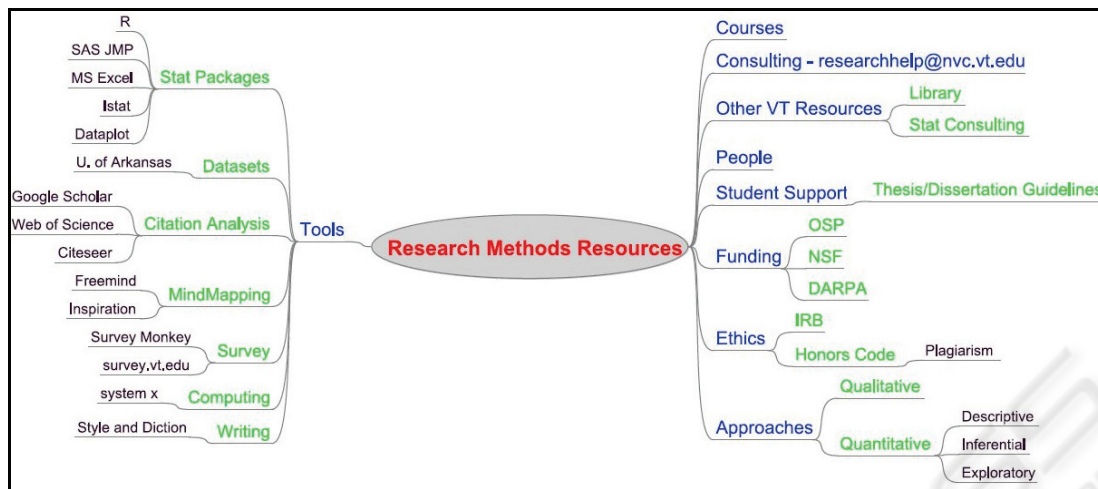
Figure 1: Mindmap of RMC resources.

The platform chosen was Simple Machines Forum (Home of SMF : Free PHP and MySQL forum software, 2010) because it is, again, one of the most widely used and simple systems. It is freeware, offers good anti-spam protection and is also capable of RSS feed syndication, which was a plus. Its system requirements are similar to WordPress', so the same machine is used to run both services

As of October 13, 2009, 59 registered members participate in the forum, having a total of 254 posts organized in 209 topics. Each of these topics has been seen 530 times on average, summing up to 110,775 views in total.

2.2.3 Writing Support

The forum section on writing support currently includes links to the style and diction tools for checking writing style, the Queequeg Grammar (Queequeg, A Tiny English Grammar Checker, 2010) checker, and to discussions of Latex (LaTeX project: LaTeX - A document preparation system, 2010) and Microsoft PowerPoint. Students get training on these tools in some courses. These courses also require students to use Google Scholar (Google Scholar, 2010), Citeseer (Computer and Information Science Papers Citeseer Publications Research Index, 2010), and Web of Science (ISI Web of Knowledge, 2010) to measure the citation rates for references they choose for course research projects.

2.2.4 Research Resources

The research resources section of the forum includes

links to a variety of information sources, from decision trees for selecting inferential statistics to how to measure research quality, as well as links to publicly available datasets. Also included is information specific to our own students such as how to get library help and a seminar series in Computer Science.

2.3 Research Groups

There are two research groups listed on the forum, Service Oriented Architectures, and Software Reuse and Domain Analysis. These group listings provide a way for researchers to identify each other and collaborate, and provides a convenient way to link web resources relevant to the groups.

2.3.1 Dissertation and Thesis Support

Since January 1, 1997, Virginia Tech. (VT) mandated that all graduate students submit their theses and dissertations online in order that research and scholarship conducted by them be freely available and openly accessible. It also allows for the use of audio and video as part of their final scholarly products. This section of the website contains a summary of available online guidelines, which appear in various VT web pages. Additionally, students are provided with an overview of the traditional five chapter format and guidelines for an alternative format, which refers to the use of manuscripts and/or book chapters to replace standard chapters. Other useful information, such as advice on how to successfully complete a thesis, is included in a forum topic area.
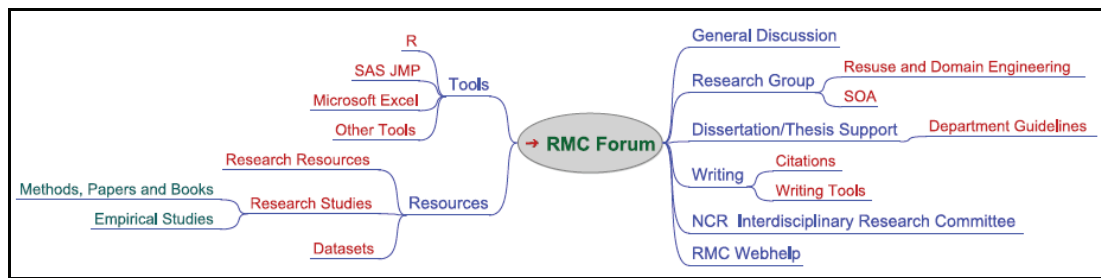
Figure 2: Mindmap of RMC forum.

2.3.2 Institutional Review Board

Any research conducted by students or faculty that involves human subjects must be reviewed and approved by the Institutional Review Board (IRB). This page alerts researchers to that requirement, providing links to VT's IRB page. Additional links lead researchers to the pages for the Office of Research Compliance, which provides support to assure that appropriate laws and regulations regarding the conducting of research are met. This also addresses Post Approval Monitoring for both animal and human researchers, which is a way to ensure compliance with both University and Federal regulations after a research protocol has been approved.

2.4 Course Support

The NCR-RMC effort has been used to support and enhance several courses, and to make more courses available across college boundaries. R, SAS-JMP, and Excel, for example, have been introduced into the courses: CS5014 - *Research Methods in Computer Science*, CS5744 - *Software Design and Quality,* STAT 5615 - *Statistics in Research I* and EDRE 6605 & 6606 - *Quantitative Research Methods in Education I & II*. The Google Video tutorials on R, discussed above, are used to introduce students to that tool. In both courses the RMC website is used as an information source and the forum is used for discussion of course topics, which serves as a knowledge body for future years. For example, the discussion of the use of simulation versus experimentation, which began in CS5014, has been accessed over 2000 times. Other contributions to the forum discussions by students include the topics of Google technical talks on information retrieval and the use of formal methods in software engineering.

The Identification of cross-discipline research methods courses led to the adoption of new cognate courses for Computer Science majors. Cognate courses are courses taught outside the computer science department that are used to give students breadth of learning. These courses, such as EDRE5644, *Questionnaire Design and Survey Research* help to broaden the research approaches for theses and dissertations in Computer Science.

2.5 Usage Statistics

Although NCR-RMC is intended for Virginia Tech faculty and students, we soon noticed that our website was being visited from many other states in the United States, and eventually from many other countries. Eighty one percent of the total visits to our website, excluding the forum, came from within the US. The other 19% came from 118 different countries. The surprising figure is that only 55% of the total visits came from the State of Virginia. This leaves us with about 45% of the visits coming from people that presumably had nothing to do with Virginia Tech.

A more detailed analysis shows that 38% of the total visits were the result of queries to search engines like Google or Yahoo. Again, only 16% of these visits came from the State of Virginia.

We examined the 2000+/- queries people used when they reached us, and found that over 55% of the queries were about:

- "research methods" with 466 visits.
- "quantitative research" with 344 visits.
- "education" with 328 visits.
- "statistics" with 228 visits.

We think these numbers help to support the point in favour of efforts like the NCR-RMC, because many visits to the site seem to come from people external to the University.

3 SUPPORTING ACTIVITIES

Although the primary motivation for the RMC was related to better communication across the faculty, it soon became apparent that student needs beyond

coursework and resource information needed to be considered. As stated above, most of our student body consists of employed adults working on their advanced degrees on a part time basis. This translates into a need for information and support in ways that differ from the needs of full-time students.

3.1 Support Person

A full-time graduate student at VT-NCR is employed as a support person for 20 hours a week to support the RMC. The support person is given an office with a desktop computer. The support person chosen was required to have basic background experience in computer science, statistics and research methods. An email id (researchhelp@nvc.vt.edu) is setup as a contact email id for the support person specific to RMC activities. Responsibilities of the support person include maintaining and updating the website and forum as its administrator.

The support person also acts as a consultant on statistics and statistical tools such as R, SAS-JMP, and Excel. For example, the support person assisted a professor from the College of Business to implement a Principal Component Analysis and Factor Rotation in SAS-JMP. At least 3 students from the College of Engineering were trained in R to build boxplots and interpret them. Students also approach the support person for help in thesis and dissertation formatting, as well as for the analysis of their data.

The support person serves as a point of contact to the RMC resources. The person also guides students to related resources such as the different support groups, online resources, courses offered etc.

3.2 Dissertation Support Group

One venue for such support was the Dissertation Support Group, which was jointly sponsored by the NCR-RMC and the NCR Graduate Student Services Office. Meetings were held on a monthly basis for the purpose of providing general information and discussions regarding research and the dissertation process. After the NCR-RMC was initiated, the first Dissertation Support Group meeting consisted of a presentation of the resources available on the website and a discussion of how to make use of them. Other sessions included presentations by invited faculty and discussions on topics such as

- using spreadsheets to organize the literature search,

- the differences between a typical five-chapter dissertation format and a manuscript format with several chapters intended to be stand-alone manuscripts suitable for submission,
- issues concerning selecting members for the dissertation committee
- guidelines for getting approval from the Institutional Review Board for their research.

Participating students were from different colleges and departments, as well as from different physical locations in the Northern Virginia area. Although attendance was small (typically, 10 to 25 students per session), those who attended were very positive about the presentations.

3.3 Panels on Qualitative Research

Although much of the research at VT-NCR is quantitative, more and more students are interested in pursuing qualitative research (Creswell, 2007), (Schwandt, 2001). We have limited course opportunities for these students. In response to this need, we organized two panel discussions for the purpose of demonstrating some of the breadth of qualitative research. Five VT faculties and one doctoral student from different disciplines provided their perspectives via overviews of the use of qualitative techniques in their research. The six presentations covered auto-ethnography, ethnography and case studies, oral histories, historical research, phenomenology, and grounded theory and the application of qualitative methods to inform building simulation models in systems engineering. MP3 files of the presentations and subsequent discussions were made available through the NCR-RMC website.

4 COLLABORATION

Faculty collaboration across departments is encouraged at Virginia Tech, and particularly so in our off-campus location. An initial underlying goal for the NCR-RMC was to foster the development of a community of practice that crossed departmental lines and with a common interest in teaching research methods. One example of this was a joint research methods course taught by faculty from computer science and college of education.

Communities of practice are "groups of people informally bound together by shared expertise and passion for a joint enterprise" (Wenger and Snyder, 2000), (Wenger, McDermott and Snyder, 2002). There was and continues to be a lot of interest and support for such an endeavour in the NCR, but we

are still working on increasing the levels of participation.

(Wenger et al., 2002) described three levels of participation in communities of practice:

1. Core members, small group or regular participants who contribute regularly to the forum.
2. Active members, a somewhat larger group who participate periodically with lower levels of engagement.
3. Peripheral members, who typically constitute the largest group but who do not contribute and maintain passive participation.

This certainly has been our experience. While gratified by the apparent peripheral usage of the website and supporting activities, we are working toward increasing both active and core participation.

Two activities planned for this academic year revolve around different formats for dissertations and theses. The first will bring interested faculty together to discuss various options so that a coherent set of guidelines may be developed for students from different programs. The second is to facilitate a series of presentations by current and former students to discuss their use of different formats and approaches. It is hoped that these activities will foster continued involvement through the forum, thereby providing a vibrant discussion of options and procedures.

5 CONCLUSIONS

In this paper we have described the use of web based, freeware, and other tools to support a cross-discipline research consortium in an off-campus setting, with several locations and many part-time professional students. Based on data we have collected, and on our experience using the tools in courses and for research, we conclude that this technology is an important aid to our research consortium effort.

The website serves as an information repository and discussion venue for the consortium. Its contents are secondary and tertiary literature. That is, some information and resources about research methods, tools, and coursework contained in the NVC-RMC web page is available elsewhere, but we thought it necessary to have a single location where both faculty and students could easily find such information and, more importantly, be able to share their knowledge.

The usage statistics only tell part of the story.

There are also intangible benefits from bringing together a community of researchers, both faculty and students.

It is our hope that the model we used in creating and maintaining the Research Methods Consortium for Virginia Tech's National Capital Region will be useful to others who might want or need a similar stimulus to aid in interdisciplinary research and in building a research community.

ACKNOWLEDGEMENTS

REFERENCES

Computer and Information Science Papers Citeseer Publications ResearchIndex 2010, viewed: 09 Jan 2010, http://citeseer.ist.psu.edu

Creswell, J. W., 2007. *Qualitative inquiry and research design: Choosing among five approaches. Thousand Oaks*, CA: Sage Publications.

FreeMind – free mind mapping software 2010, viewed: 09 Jan 2010, http://freemind.sourceforge.net/wiki

Google Scholar 2010, viewed: 09 Jan 2010, http://scholar.google.com

Home of SMF : Free PHP and MySQL forum software 2010, viewed: 09 Jan 2010, http://www.simplemachines.org

ISI Web of Knowledge 2010, viewed: 09 Jan 2010, www.isiknowledge.com

JMP Software 2010, viewed: 09 Jan 2010, http://www.jmp.com

LaTeX project: LaTeX - A document preparation system 2010, viewed: 09 Jan 2010, http://www.latex-project.org

Queequeg, A Tiny English Grammar Checker 2010, viewed: 09 Jan 2010, http://queequeg.sourceforge.net/index-e.html

Schwandt, T. A., 2001. *Dictionary of qualitative inquiry*, 2nd ed. Thousand Oaks, CA: Sage Publications.

The R Project for Statistical Computing 2010, viewed: 09 Jan 2010, http://www.r-project.org

Wenger, E.C. & Snyder, W.M, 2000. *Communities of practice: The organizational frontier*. Harvard Business Review, January-February, 139-145.

Wenger, E.C., McDermott, R., & Snyder, W.M., 2002. *Cultivating communities of practice: A guide to managing knowledge*. Cambridge: MA: Harvard Business School Press.

WordPress > Blog Tool and Publishing platform 2010, viewed: 09 Jan 2010, http://wordpress.org