2015 UFII Seed Grant Proposal

Title:
MassMine Development and Training Project

PI:
Sid Dobrin, Department of English, College of Liberal Arts & Sciences, sdobrin@ufl.edu

Co-PIs:
Aaron Beveridge, Department of English, College of Liberal Arts & Sciences, aaronbev79@ufl.edu
Norman Lewis, College of Journalism and Communications, nlewis@jou.ufl.edu
Mindy McAdams, College of Journalism and Communications, mmcadams@jou.ufl.edu
Matt Gitzendanner, Research Computing, magitz@ufl.edu
Laurie Taylor, George A. Smathers Libraries, laurien@ufl.edu

Project Start Date:
May 15, 2015

Budget Request:
$49,837
1. Abstract:

The MassMine Development and Training Project represents a new interdisciplinary collaboration among the Department of English, College of Journalism and Communications (CJC), George A. Smathers Libraries, and Research Computing at the University of Florida. The primary collaboration will be between Sidney Dobrin and Aaron Beveridge, Department of English in the College of Liberal Arts & Sciences, and Norman Lewis and Mindy McAdams, in the CJC. Laurie Taylor, Digital Scholarship Librarian in the Smathers Libraries, and Matt Gitzendanner, Training Coordinator in Research Computing, will provide additional support and oversight for the project. The project team respectfully requests funding to support Aaron Beveridge as a graduate fellow for one year, to cover the costs of the graduate stipend, tuition waiver, and research travel funding to present on the project at professional conferences. The project team also requests funding to support a graduate research assistant for Norman Lewis and Mindy McAdams for 10 hours a week during the fall 2015 and spring 2016 semesters. The graduate assistant will support their MassMine research projects on Research Computing’s supercomputer cluster, help develop new data-intensive course content and training materials, and assist in the curation and analysis of new data sets. The total financial request for this project is $49,837. In collaboration with the project team, Beveridge will support the project objectives, provide reporting for the project, and see to the development of the next grant proposals that build from this seed project. The project consists of two key objectives: (1) the development of new teaching and research applications for MassMine and Research Computing within the CJC; and, (2) the pursuit of extramural funding to support: the ongoing development of MassMine, the creation of an interdisciplinary training institute for the collection and analysis of social network data, and future research projects and curriculum development in the CJC. MassMine\textsuperscript{1} is an open source software project that was started at the University of Florida, and it has been extensively tested and supported by UF’s Research Computing. MassMine manages long-term data collection projects for social network API (application programming interface) data.

\textsuperscript{1} http://www.massmine.org
2. Project Description:
   a) Specific Aims/Objectives: Teaching and Research Collaboration

   The first step of the collaboration will be summer 2015 training for MassMine, with the training created and conducted by the graduate fellow. McAdams has Research Computing access from her previous contributions to the MassMine project, and Lewis will create his Research Computing account to learn how to use MassMine over the summer. Using the training materials developed during Beveridge’s “Writing through Big Data” course, Beveridge will assist Lewis and McAdams as they collect and analyze Twitter data over the summer. Lewis currently teaches a graduate course in the CJC titled Big Data Seminar. His course focuses on the theoretical and conceptual aspects of “Big Data” as it relates to journalism and mass communications. The summer training sessions will allow skills to be added to the course so that students may gain practical experience in pulling and analyzing large data sets for mass communication research. Beyond the classroom applications, there are many graduate students in the CJC whose research may be supplemented or expanded by providing access to social network data. For example, a doctoral student in the college is now examining Facebook news feeds to evaluate how people define news in a social media era. The use of MassMine data will supporting improving dissertations by expanding the research skill set of the College’s faculty and students. While Beveridge will focus his summer activities on training Lewis and McAdams, the collaboration will flow in both directions. Lewis and McAdams will provide crucial insights into any unseen limitations of the software, ideas for new modules or features that would be useful for research in journalism, and new teaching applications and extensions for MassMine's training materials and tutorials. In turn, Lewis will be able to use the techniques and data types provided by MassMine to expand his own research into big data. Extending the software to support new research methods, modes of inquiry, and research questions will position the MassMine project team for future grant proposals from additional potential sponsor agencies.

   McAdams teaches a graduate course in CJC titled New Media and a Democratic Society in which uses of social media in protest movements and “citizen journalism” are key topics for study. It may not be impossible to predict when a new Arab Spring or Ferguson will appear, but it is clear that communications on Twitter constituted central and important narratives in both of those cases. MassMine can be deployed quickly after the emergence of such trends to capture and track communications during events of similar social and political importance. Training, practice and expanded tutorials (as well as increased awareness of MassMine’s capabilities) will make it possible for media scholars to move with the necessary speed to analyze conversations and commentary as events unfold. McAdams will focus her work with students on using the flexibility provided by MassMine to Twitter campaigns and then readjust focus as new conversations (trends) arise and change (e.g. new hashtags or keywords emerging from an ongoing trend).

   Following the summer training and development, Beveridge will assist Lewis and McAdams in the teaching of their courses. Beveridge will help students set up the Research Computing accounts, install MassMine, and learn hands-on from trainings on how to manage long-term Twitter research projects. The training includes support for developing research questions that are relevant to the data types available through the Twitter APIs, understanding how to curate and process large Twitter data sets for analysis, and determining which kinds of text mining analyses and data visualizations to use with Twitter data. Since Lewis and McAdams will be fully trained on MassMine by the end of the summer,

---

2. https://www.jou.ufl.edu/assets/syllabi/MMC6936-BgData-Section1E86-Lewis-SP15.pdf
Beveridge's primary role will be in a support and consulting capacity. Additionally, Beveridge will apply feedback from the faculty and classes for the ongoing development of the MassMine software and the MassMine training program in collaboration with the UF Libraries and Research Computing.

In addition to the training and support provided directly to Lewis and McAdams, Beveridge will also support new educational initiatives in the Digital Humanities at UF. Beveridge will complete two sets of three-part trainings in the newly completed Scott Nygren Scholars Studio—a digital studio meant to serve an interdisciplinary training and support role for digital scholarship in the George A. Smathers Libraries. One three-part training series will be completed in the fall, and a second one will be completed during spring semester. These trainings will be open to all students/faculty interested in learning how to conduct research with data from Twitter's API with MassMine. The first of the three parts will focus on setting up Research Computing accounts, installing MassMine, and learning how to manage a long-term data collection project. This includes learning how to use FileZilla to move data from Research Computing servers, how to backup and archive large amounts of data during a long-term project, and how to develop a research question that is based on the data types made available by Twitter. The second part of the three-part training will explain how to use exploratory methods to analyze and adjust research projects while a long-term data collection is underway. This includes learning how to produce correlation cluster-graphs showing the strength of correlation among words, #hashtags, and influential usernames circulating within a particular trend. This will also include basic descriptive statistics such as time series analyses to analyze changes in the data over time and to discover newly emerging trends that may expand the data collection project as data sets are curated. The third part of the training will train participants on how to export their data visualizations, how to redesign them (color choices, size of components, etc.), how to add appropriate labels and descriptions, how to share data and collaborate with other authors, and how to embed visualizations in documents for eventual submission to an academic journal or other type of publication. These trainings, first and foremost, will be marketed to students in the Department of English and the College of Journalism and Communications as a way of extending the collaborative work for this project, but they will also be heavily promoted across campus and available for any students, faculty, and researchers interested in establishing a research project using social network data.

Similar to the support provided to Lewis and McAdams, Beveridge will also support the newly created interdisciplinary program for the “Digital Humanities Graduate Certificate.” The first class in the series will be co-taught by Laurie Taylor and Elizabeth Dale (Department of History), and their students will participate in one of the three-part trainings listed above as part of the course. Where necessary, Beveridge will also consult with individual graduate students who may decide to undertake social network research projects in Dale and Taylor's course, and he will also provide additional support to the course when applicable.

Overall, the MassMine Development and Training Project will extend the use of MassMine and Research Computing to courses and research projects in the College of Journalism and Communications, it will provide valuable feedback on future developments of the MassMine software, and it will extend the available training materials and tutorials for the types of courses and research projects that MassMine will be able to support in the future. All of this work will greatly strengthen the project team's ability to pursue future funding from the multiple funding agencies listed later in this document.

b) Background and Significance

MassMine began as an interdisciplinary collaboration between Aaron Beveridge and Nicholas Van Horn in order to test Sidney Dobrin's theory of hyper-circulation. In the fall of 2013, Beveridge and Van Horn began theorizing a way to use Twitter trends to understand the hyper-circulation of information.

5 http://wiki.hpc.ufl.edu/doc/FileZilla

6 Although Van Horn is not included on this particular grant, he received his PhD in Psychology (Cognitive Neuroscience) from The Ohio State University. Van Horn currently uses MassMine in his social science research methods course at Capital University.

7 Dobrin argues that the human-as-node model of network theory in writing studies cannot currently explain the complex fluctuating-flow of writing that occurs during “hyper-circulation” and “network saturation.”
among locations. Twitter trends are often referenced in news reports and entertainment media, and this reinforces the illusion that social network data is readily available for research and in-depth analysis. Beveridge and Van Horn, however, quickly realized this supposed wide-availability of social network data was a false assumption. Access to social network data is usually limited to three narrow types of access— all of which present interdisciplinary problems for academic scholars who wish to access this data for research. Types of limited access begin with the most common with cloud tools that visualize some aspect of social network API data through web interfaces. Tools like Trendsmap Solutions and SumAll provide initial free access to data and surface-level visualizations, but their more in-depth tools and analyses are only available through paid subscription services. For another type of access, larger providers such as GNIP, Topsy, and Radian6 (owned by Twitter, Apple, and Salesforce respectively) provide access to larger amounts of data than what is available through social network APIs, but these are even more costly than the cloud tools. Even if the costs were not prohibitive, the analytics provided by these tools focus mostly on brand management, marketing, product development, and customer service applications of social network data. In the case of these larger providers, as well as with the API cloud tools, data export functionality is limited and often without transparency regarding the methods employed to provide the analyses given to customers— both of which present large barriers for reproducible academic research. Finally, there is the option of accessing API data directly. While this approach is cost-effective (most APIs can be accessed free of charge), it requires a high level of technical knowledge and programming skills in order to systematically access and archive API data to curate large data sets.⁹

MassMine accesses API data directly, and drastically reduces the amount of technical knowledge and programming skills required in order to collect freely-available API data¹⁰. Twitter is MassMine's only data source currently, and work is underway to add Facebook, Reddit, and Wikipedia as data sources. MassMine is written in R¹¹, one of the top programming languages for data mining and statistical computing, and MassMine is available for free download and open collaborative development (support and user downloads are managed through GitHub). Almost no knowledge of R is required in order for users to collect Twitter data with MassMine. MassMine operates based on a configuration file that users edit in any basic text editor— allowing them to save their API credentials, usernames, and passwords for quickly restarting the software and managing long-term data collection projects. MassMine has full documentation available on pulling data from Twitter's Rest API. Furthermore,MassMine's newly finished Streaming API functionality pulled over 1.3 million Tweets from 6:30pm until 2:00am during the 2015 Academy Awards— successfully tracking the trend “#oscars” as it was occurring on Twitter.

High levels of need and interest in tools for social media research that allow users to collect, access, and query data resulted in a grant proposal submission for ongoing development of MassMine, submitted to the National Endowment for the Humanities (NEH) for the Digital Humanities Start-up Grant category in September 2014.¹² In developing the grant proposal, researchers identified a critical need for training alongside ongoing development. McAdams was one of the first users to test MassMine outside of the Department of English. McAdams has experience as a UF instructor who teaches programming and web development to Journalism students which made her an ideal contributor while MassMine completed its documentation in the fall of 2014. By May of 2015, MassMine will have been successfully tested on Research Computing's supercomputer cluster¹³ for a full year¹⁴. Finally, Beveridge

---

8 Researchers with advanced computer science skills can systematically scrape data from the web directly to create their own large data sets. However, a vast majority of interdisciplinary researchers do not have the skills required to complete such a task.
9 Beveridge has a forthcoming book chapter titled “The Problem of Circulation Data Access: Software Development and eScience” where these problems are explored in further detail.
10 A New York Times article from August 2014 explains how data “janitorial” work remains a crucial component of development for data science.
11 For other humanities data mining examples in R, see Matthew Jockers' Text Analysis with R for Students of Literature.
12 See the full grant proposal at: http://ufdc.ufl.edu/AA00025642/00001
13 http://www.rc.ufl.edu/
and Van Horn are presenting the first version of MassMine's new visualization front-end (see screenshots in appendices) at the Conference on College Composition and Communication (CCCC), the largest international conference in writing studies, on March 20, 2015.

c) Innovation/Potential Impact of Research

In spring 2015, Beveridge is teaching a class at UF titled: “Writing through Big Data.” This is the first humanities course at UF to use Research Computing’s supercomputer cluster in the classroom—allowing students to manage long-term Twitter trend research over the course of a semester with MassMine. In the course each student must choose a specific Twitter trend to research over the course of the semester. The initial research question they are presented with is: what is the exigence of your trend? Students are allowed to pick either a prescriptive #hashtag trend, or they may pick an organic trend that has emerged on its own—Twitter's Rest API allows for both types of trends in the trend data provided. Students are researching everything from the Academy Awards, the NFL draft, and fad diets and health trends—to more serious topics such as human trafficking, women's rights and body politics, #gamergate, trends surrounding the legalization of marijuana, and recent trends involving disease epidemics. In the class students are learning advanced computer skills such as how to use a terminal, command line interface to log into Research Computing and run MassMine, how to archive and carefully save their data over the course of a semester, and finally, how to analyze and visualize their data to present to their peers. Students may choose between completing a 12 minute formal presentation in front of the class at the end of the semester, or they may give a short informal presentation and then attempt to publish their research with some type of outside publication. The course has already completed the project proposals for student research projects—about half the class will be completing formal presentations, and the other half has found blogs or other smaller publications where students will submit their work by the end of the semester.

In addition to development and teaching, the MassMine team is finishing its first academic article for review by Digital Humanities Quarterly. The article, titled “Attention Ecology: Trend Circulation and the Virality Threshold,” discusses the concept of macroscopic “attention” for social networks, and proposes an operational definition for a “virality threshold” within Twitter. The team for this publication includes Beveridge and Van Horn of MassMine, and Sean Morey (Assistant Professor of New Media and Rhetoric, Department of English, Clemson University). The data that forms the basis for this publication was pulled during MassMine's initial test run, and it has motivated further research into the flow of information among locations in Twitter. Along with support from Research Computing, MassMine will be used in a follow-up test of the proposed “virality threshold” among every available trend location in the U.S. provided by Twitter's API. The researchers involved in this project hope that this new data set will form the basis for a second, follow-up publication that further substantiates the “threshold” concept.

Finally, substantial work has gone into writing full documentation through the massmine.org website. Van Horn finished writing the documentation last December, and this provides a comprehensive explanation of how the MassMine software functions. Additionally, Beveridge has a publication in process with The Programming Historian, an open-access, peer reviewed journal that publishes technical work and methodological descriptions of new research approaches in the digital humanities. Beveridge is working with the editors of the publication to provide both video tutorials and technical descriptions of everything from MassMine installation, to managing a long-term project and developing a relevant research questions, and finally to analyzing and visualizing data.

While the use of data science methods in mass communications is in its infancy, some research has been published that suggests it is a fruitful subject both as a tool and as a method. For example, big data techniques have been used to extend agenda-setting theory beyond its traditional conception as a one-way flow toward a reciprocal and dynamic relationship between media and audience. A case study

14 [http://wiki.hpc.ufl.edu/doc/Massmine](http://wiki.hpc.ufl.edu/doc/Massmine)
16 W. Russell Neuman, Lauren Guggenheim, S. Mo Jang, and Soo Young Bae, “The Dynamics of Public
examined the strengths and weaknesses in using big data techniques to augment or supplant human judgment in content analyses. Other scholars have examined the use of big data as a method. A typology was developed to evaluate computational journalism and its epistemological ramifications. Qualitative studies have examined the use of automated reporting techniques and journalistic inspection of algorithms that define and exclude. Several studies have examined Twitter as an information source and as a journalistic tool.

All of this preliminary innovation and research will be directly useful to the MassMine Development and Training Project. Whether it be using MassMine and Research Computing in the classroom, publishing with data collected from Twitter APIs, or the development of technical descriptions and training materials for interdisciplinary work—the preliminary groundwork listed above has made this new collaborative project between the Department of English and the CJC “shovel ready.”

d) Approach/Research Design

At the Informatics Institute Symposium in 2016, Dobrin, Beveridge, Lewis, and McAdams will propose to present a digital poster together titled: “Twitter Data in the UF Classroom”—where examples of student research projects and teaching ideas will be presented. Furthermore, Beveridge will present an additional poster with updated results on the on-going research into the “virality threshold” among all available Twitter locations. Digital copies of all new training materials and resources that are developed as a result of this collaboration will be stored in the Digital Humanities Resources in the George A. Smathers Libraries, as well as new materials being provided to the Research Computing Wiki-page where applicable. The Informatics Institute will receive email notifications when any new digital resources associated with trainings and course materials are created and archived, and these will be created and completed according the project timeline included in the appendices. Finally, as noted below in the “Plans to Obtain External Support” section, two of the four grants listed will be chosen for development and application, and these too will be submitted according to the project timeline. Copies of grant applications will also be stored in the Digital Humanities Resources, and email notifications of their final submission to funding agencies, with links to the digital documents, will be sent to the Informatics Institute as well upon completion.

e) Preliminary Data

Existing training resources for MassMine are available on the MassMine website (www.massmine.org) and new materials are being developed as part of Beveridge’s spring 2015 course. Programming and development on MassMine are ongoing for general needs, with this project specifically to support new collaborative needs with the humanities and journalism, as well as the larger training program to support developing greater critical mass across UF’s campus for big data, data science, digital humanities, and other interdisciplinary research needs related to informatics.

Attention: Agenda-Setting Theory Meets Big Data.” Journal of Communication


3. Project Budget and Justification of Expenses:
Budget costs for Aaron Beveridge is based on his current Graduate School Fellowship (GSF Fellowship\(^\text{23}\)) costs in the Department of English from 2015-2016.

<table>
<thead>
<tr>
<th><strong>Aaron Beveridge</strong></th>
<th>Summer 2015</th>
<th>Fall 2015</th>
<th>Spring 2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary</td>
<td>$20,000.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waivers</td>
<td>$1,795</td>
<td>$4,038</td>
<td>$4,038</td>
<td>$9,871</td>
</tr>
<tr>
<td>Fringe</td>
<td></td>
<td></td>
<td></td>
<td>$1,417</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>$31,288</strong></td>
</tr>
</tbody>
</table>

Budget costs for the research assistant for Norman Lewis and Mindy McAdams is based on the cost of an MA graduate student in the CJC for 10 hours per week.

<table>
<thead>
<tr>
<th><strong>CJC Student</strong></th>
<th>Summer 2015</th>
<th>Fall 2015</th>
<th>Spring 2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary</td>
<td>$0</td>
<td>$3,301</td>
<td>$3,301</td>
<td>$6,602</td>
</tr>
<tr>
<td>Waivers</td>
<td>$0</td>
<td>$5,048</td>
<td>$5,048</td>
<td>$10,096</td>
</tr>
<tr>
<td>Fringe</td>
<td>$0</td>
<td>$432</td>
<td>$419</td>
<td>$851</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>$17,549</strong></td>
</tr>
</tbody>
</table>

Aaron Beveridge has presented at CCCC (Conference on College Composition and Communication, the largest international conference in writing studies) for the past two years. Travel funds will support his presentations at similar international conferences in 2016.

<table>
<thead>
<tr>
<th><strong>Conference Presentation &amp; Travel</strong></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaron Beveridge</td>
<td><strong>$1,000</strong></td>
</tr>
</tbody>
</table>

**Total Grant Budget**

**$49,837**

\(^{23}\) [http://graduateschool.ufl.edu/finances-and-funding/graduate-school-fellowship-program-guidelines](http://graduateschool.ufl.edu/finances-and-funding/graduate-school-fellowship-program-guidelines)
4. Literature Cited

5. Key Personnel
The list provided below gives bulleted descriptions of each project participant’s role in the grant. This list corresponds directly to the project timeline provided in the appendices of this document.

Sidney Dobrin, Department of English
- Principal Investigator
- Project advisor
  - Oversee Aaron Beveridge's work with Norman Lewis, Mindy McAdams, Laurie Taylor, and training program delivery including in the Scott Nygren Scholars Studio
  - Assist in grant selection, grant writing, and act as advisor and principal investigator on future funding applications

Aaron Beveridge, Department of English
- Co-Principal Investigator
- Training and classroom support for Norman Lewis, Mindy McAdams, Laurie Taylor, and training program delivery including in the Scott Nygren Scholars Studio
- Grant writing
- Continue collaborative research into “virality threshold” for Twitter trends

Norman Lewis, College of Journalism and Communications
- Principal Investigator
- Project advisor
  - Oversee Aaron Beveridge's work with the CJC
- Develop new course content using MassMine and Research Computing
- Grant selection, grant writing, and act as advisor and principal investigator on future funding applications
- Provide feedback on limitations and/or future developments of MassMine for academic research
- Conduct research into social media using big data methods

Mindy McAdams, College of Journalism and Communications
- Co-Principal Investigator
- Development new course content using MassMine and Research Computing
  - Write tutorials and training materials specific to students and faculty in the CJC
- Provide feedback on limitations and/or future developments of MassMine for academic research
- Grant writing

Laurie Taylor, George A. Smathers Libraries
- Co-Principal Investigator
- Oversee Aaron Beveridge's work with the Scott Nygren Scholars Studio
- Integrate MassMine trainings into the interdisciplinary Graduate Certificate in Digital Humanities
- Provide feedback on limitations and/or future developments of MassMine for academic research
• Grant selection, grant writing, and act as advisor and principal investigator on future funding applications

Matt Gitzendanner, Research Computing
• Co-Principal Investigator
• Provide Research Computing accounts for all courses and training sessions associated with the project
• Provide additional support for Aaron Beveridge if necessary
• Act as principal investigator on future funding applications
6. Plans to Obtain Continuing External Support:
The collaboration between the Department of English and the CJC, with support from Research Computing and the UF Libraries, for the UF Informatics Institute Seed Grant will accomplish two primary goals regarding the pursuit of extramural funding. First, the Seed Grant will provide support for ongoing software development and training as applied to the specific needs of the CJC, which has many shared interests and methods but which extends beyond the core existing community of users in English and Writing Studies, to build from shared interests and methods which ensure that the collaboration can efficiently and productively accomplish the stated goals, and which will result in improvements to the software and training program to enable MassMine to seek additional grants from the NEH, including the Digital Humanities Implementation and Advanced Training Institute awards. Second, because Journalism is a separate college and discipline, there are additional data-intensive and informatics related grants that will be pursued as a result of the collaboration. The Alfred P. Sloan Foundation has the “Digital Information Technology” grant that supports interdisciplinary research in the areas of “Data and Computational Research” and “Universal Access to Knowledge.” With support from MassMine's data mining technology and Research Computing, Norman Lewis will be competitive to pursue funding from this agency. The “Digital Information Technology” grant funds project budgets over $125,000, and has numerous options for supporting the interdisciplinary development of educational and research programs that utilize data-intensive methodologies.\(^\text{24}\) Additionally, Google Research awards regular grants up to $125,000 to university faculty research projects that seek to “Organize the world's information and make it universally accessible and useful.”\(^\text{25}\) As big data technologies continue to affect the way citizens access information, developing critical resources and new methodologies for quickly analyzing social network data is a crucial component of Journalism and Mass Communication’s role in evaluating and responding to the growing use of predictive analytics and personalized media in ways that are profoundly changing every aspect of the field.

MassMine submitted its first grant in fall 2014 to the NEH “Digital Humanities Start-Up Grant” program\(^\text{26}\) to fund $60,000 toward post-doctoral funding for Nicholas Van Horn and additional project costs. NEH normally provides notification of awards in April 2015. There are two additional NEH grants that are directly relevant to this project. One is the NEH “Digital Humanities Implementation Grant”\(^\text{27}\) program that is “designed to fund the implementation of innovative digital-humanities projects that have successfully completed a start-up phase and demonstrated their value to the field.” This value is demonstrated, according to the NEH, through both the creation of new methods and tools, and through well-established scholarship and teaching. The DH Implementation Grant funds projects up to $325,000. The MassMine project team would propose a grant for this source to support the hiring of additional post-doctoral fellows to continue MassMine's development, conduct new social network research and publish results, and continue development of interdisciplinary relationships with other departments/colleges at the University of Florida and partner institutions. The second NEH grant is the “Institutes for Advanced Topics in the Digital Humanities”\(^\text{28}\) which supports the

\(^{24}\) http://www.sloan.org/major-program-areas/digital-information-technology/data-and-computational-research/  
\(^{25}\) http://research.google.com/university/relations/research_awards.html  
\(^{26}\) http://ufdc.ufl.edu/AA00025642/00001  
\(^{27}\) http://www.neh.gov/grants/odh/digital-humanities-implementation-grants  
development of regional, multi-institutional training institutes for advanced topics in the digital humanities, with maximum awards of $250,000. Two of NEH’s primary categories for this funding are “textual analysis and text mining” and “high-performance computing or supercomputing.” MassMine's successful partnership with Research Computing at UF—using the supercomputer cluster to manage long-term Twitter mining projects—uniquely positions MassMine to develop an interdisciplinary training institute in coordination with the UF Informatics Institute and Research Computing.

The project team will choose between one of the two NEH Digital Humanities grants by the end of May 2015. If Norman Lewis and Mindy McAdams research projects produce promising datasets with MassMine, the project team will additionally apply for a second grant—choosing between the Alfred P. Sloan Foundation and Google Research Grants by the end of the summer. All four primary collaborators will be included as PIs or co-PIs in both grants (unless, of course, it appears more advantageous to apply for particular grants separately), and commitments of support for the pursuit of future funding have already been approved by Research Computing at UF. Laurie Taylor, a Digital Scholarship Librarian for the University of Florida, will collaborate with the drafting and conceptualizing of the NEH grant. Grant specialist Yulia Streklova will support the project team if a second grant is applied for during the 12 month funding period.
7. Investigator Biosketches

Biographical Sketch

Sidney I. Dobrin

Department of English, University of Florida, FL
Tel: (352)-332-5526, Email:

Professional Preparation
Virginia Wesleyan College   English   B.A., 1989
Old Dominion University    English   M.A., 1991
University of South Florida English   Ph.D., 1995

Appointments
1997-present  Professor, Department of English, University of Florida
2012-2015    University of Florida Research Foundation Professor
1995-1997    Assistant Professor, University of Kansas
1991-1995    Graduate Teaching Assistant, University of South Florida
1990-1991    English Instructor, ECPI (an A.A. accredited proprietary college),
             Virginia Beach, VA
1990-1991    English Instructor, Volunteer English Tutor for Soviet Immigrants
1987-1989    Peer Instructor, Virginia Wesleyan College

Five Products Most Closely Related to Proposed Project

   University of New York Press.

Five Other Significant Products


**Synergistic Activities**

1. Founding Director of TRACE innovations initiative, the Department of English. TRACE is a research endeavor developed and maintained by the University of Florida’s Department of English. TRACE works at the intersection of ecology, posthumanism, and writing studies. Providing an interdisciplinary forum for scholars, we focus on the ethical and material impact of media. TRACE acts as a hub for several distinct projects including an online journal, Vectors, ARCs, and MassMine for which we are always seeking submissions. We welcome submissions that contribute to the theorization and invention of nature, technology, and culture.


5. Editorial Board, *Composition Studies* (2001-present)

**Collaborators & Co-Editors** J.A. Rice (Western Kentucky University) Sean Morey (Clemson University) Christopher J. Keller (University of Texas Pan-American) Kenneth Kidd (University of Florida) Stephen G. Brown (University of Nevada) Christian Weisser (Penn State Berks) Lynn Worsham (Idaho State University) Gary Olsen (Daemen College) Anis S. Bawarsh (University of Washington)

**Graduate Advisors** Gary A. Olsen (University of South Florida) Patricia Bizzel (College of Holy Cross)

**Graduate Advisees** Jeff A. Rice (PhD) Micahel Vastola (PhD) Sean Morey (PhD) Rania Williams (MA) Clay Arnold (PhD) Sean Morey (PhD) Brenda Maxey-Billings (PhD) Clay Arnold (MA) Sean Morey (MA) Mike Mulhauser (MA) Linda Howell (PhD) Scott Reed (MA) Lloyd Willis (MA) Christopher J. Keller (Ph.D.) Dion Cautrell (Ph.D)
Biographical Sketch

Norman P. Lewis

College of Journalism and Communications, University of Florida, FL 32611
Telephone: 352-392-5137. Email: nplewis@ufl.edu. Website: www.bikeprof.com

Professional Preparation

<table>
<thead>
<tr>
<th>Institution</th>
<th>Program</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Illinois University</td>
<td>Journalism</td>
<td>B.A., 1979</td>
</tr>
<tr>
<td>University of Maryland</td>
<td>Journalism &amp; Public Communication</td>
<td>Ph.D., 2007</td>
</tr>
</tbody>
</table>

Appointments

- 2013 to present: Associate professor, University of Florida
- 2007 to 2013: Assistant professor, University of Florida
- 2004 to 2007: Graduate teaching assistant, University of Maryland (while doc student)
- 1986 to 1989: City editor, *La Crosse Tribune*, La Crosse, WI

Five Publications Most Closely Related to Proposed Project


Five Other Significant Publications


Synergistic Activities

- Since summer 2014, I have collaborated with four other UF faculty from four other departments to create and teach one of two courses President Machen selected for pilot
for a “grand challenge” reimaging of undergraduate core curriculum. Our class is called People and Data. It involves the social sciences and touches on big data.

- As lead author on a genetics research project published in *New Genetics and Society* in 2011, I collaborated with faculty and staff at UF Shands and the medical school to blend their technical expertise with a social science approach to analyzing direct-to-consumer websites promising genetic information.

- I am on the editorial board of Media Law and Ethics Journal and regularly review research articles submitted to three leading journals in our field.

- For the College of Journalism and Communications, I created a multi-disciplinary graduate seminar in big data being taught for the first time this spring and am creating a “flipped” course in media numeracy that could be required of all 2,000 undergraduates in all disciplines.

- I actively participate in a data journalism organization, the National Institute for Computer-Assisted Reporting (NICAR), to extend university-driven knowledge and research to data-minded journalism professionals.

**Collaborators & Co-Editors** W.L. Allen (University of Florida), F. Gao (Zhejiang University), S.I. Hsu (Prometheon Pharma), K. Hull (UF), K. McAdams (University of Maryland); J. Neely (University of Tampa), J. Starr (UF), Y. Takata (IQ Solutions), D. Treise (UF), K. Walsh-Childers (UF), Q. Xie (American Association for Public Opinion Research), B. Zhong (Penn State University).

**Graduate and Postdoctoral Advisors** K. Bartol (U. of Maryland), T. Kunkel (St. Norbert College).

**Graduate Advisees (total of 10)** Ph.D.: K. Anderson (UF), P. Braun (self-employed), J. Cox (Salisbury University), L. Darm (UF), A. Hall (UF), M. Mallicoat (UF), E. Newport (UF), E. Rice (UF). Master’s: M. Gillespie (WOLO-TV ABC, Columbia, S.C.); W. Lavaux (UF)

**Postdoctoral Associates (total of 0)** None
Biographical Sketch

Melinda J. (“Mindy”) McAdams

Department of Journalism, University of Florida, FL 32611
Tel: (352) 392-8456 Email: mmcadams@ufl.edu

Professional Preparation
Pennsylvania State University Journalism B.A., 1981
The New School for Social Research Media Studies M.A., 1993

Appointments
1999 – present Professor and Knight Chair for Journalism Technologies and the Democratic Process, University of Florida
1997 – 1999 Web strategist, American Press Institute, Reston, VA
1995 – 1997 Consultant, McAdams New Media, Toronto and Washington, DC
1984 – 1988 Copy chief, copy editor, MIS Week, New York

Publications Most Closely Related to Proposed Project

Other Significant Publications

Synergistic Activities
- Twitter: 9,000+ followers (none paid or bought). https://twitter.com/macloo
- Social Media Management: undergraduate course developed and taught by me, 2014–15, exploring uses of various social media platforms and tools in journalism organizations. https://socmeduf.wordpress.com/sked/
- Mellon Scholar-in-Residence, Rhodes University, Grahamstown, South Africa, May–July 2014. Invited by the director of the university’s research office and the head of the
journalism school to develop and deliver workshops for faculty, graduate and undergraduate students in digital journalism, and to consulted with faculty on curriculum reform and a new master’s program in digital journalism.

- New Media and a Democratic Society: graduate course developed and taught by me, 1999–2014, exploring new technologies for communication and how they impact free speech, political participation, and agency. [https://mmc6612.wordpress.com/category/assignments/](https://mmc6612.wordpress.com/category/assignments/)

- Advanced Online Media Production: undergraduate course developed and taught by me, 2000–present, focused on digital storytelling techniques and best practices (Web code and apps). [https://mmc4341.wordpress.com/sked/](https://mmc4341.wordpress.com/sked/)

Collaborators & Co-Editors: Eisa Al Nashmi (Kuwait University), C. Armstrong (University of North Texas), Johanna Cleary (University of Florida), J.-C. Molleda (University of Florida)

Graduate Advisor (total of 1): Wolfgang Schirmacher, Ph.D., professor and Arthur Schopenhauer Chair, European Graduate School (EGS).

Thesis Advisor (total of 16) Eisa Al Nashmi (Kuwait University), Matt Blake (California State, Chico), Katherine Blasewitz (unknown), Denise Bortree (Penn State), Hyeri Choi (unknown), Janna Crabb (Council for Advancement and Support of Education, Washington, D.C.), Joseph Crinkley (NPR Digital Services, Boston), Myriam Delouvrier (Farner Consulting AG, Zurich), Alan Flaten (RE/MAX), Fangfang Gao (Zhejiang University), Brian Hagen (illustrator, Cincinnati), Ben Holtzclaw (FishTrack.com, Huntington Beach, CA), Kecia Johnson (unknown), Hedda Prochaska (University of Texas, Austin), Adriana Pulido (Ministry of Information Technologies and Communications, Bogotá, Colombia), Amy Zerba (New York Times, New York)
Biographical Sketch

Laurie N. Taylor

George A. Smathers Libraries, University of Florida, FL 32611
Tel: (352)-273-2902, Email: laurien@ufl.edu

Professional Preparation

Jacksonville University          English          B.A., 1999
University of Florida           English          M.A., 2002
University of Florida           English          Ph.D., 2006

Appointments

2011–present              Digital Scholarship Librarian, Scholarly Resources & Services, Smathers Libraries, University of Florida
2008–2011                  Interim Director, Digital Library Center, Smathers Libraries, University of Florida

Five Products Most Closely Related to Proposed Project


Five Other Significant Products


Synergistic Activities

- Technical Director and Scholarly Advisory Board Ex-Officio Member, the Digital Library of the Caribbean (dLOC, www.dloc.com), an international collaborative open access digital library with 37 partner institutions; supporting needs for academia, scholarly research, teaching, and service, as well as for GLAMs (galleries, libraries, archives, and museums); supporting data curation, digital preservation, creation of new digital scholarship, and ongoing community development. Liaison and collaborator on projects to leverage dLOC as scholarly cyberinfrastructure for scholarly research needs, as with “The Archive of Haitian Religion and Culture”, “The Florida and Puerto Rico Newspaper Project”, and “The Panama Canal” (2008-present).
- Co-Chair, UF’s campus-wide Data Management/Curation Task Force (http://library.ufl.edu/datamgmt) of Research Computing, the Division of Sponsored Programs, and Libraries, undertaking various activities (assessments; trainings; developing tools, policies, and procedures; focused projects to define needs and develop integrated data support; e.g., “little data” and small databases) to develop comprehensive data support and create a data literate culture (2012-present)
- Co-Convener, UF Digital Humanities Working Group, supporting activities, events, etc., (http://www.humanities.ufl.edu/digitalhum.html; 2011-present)
- Co-Chair, Association of Caribbean University, Research and Institutional Libraries (ACURIL) Information Technology Special Interest Group (IT-SIG), organizing panels and presentations for the annual conference and sharing practices throughout the year (2010-present)


Graduate Advisees (total of 1) Lukasz Ziemba (U. Florida, 2011)

Postdoctoral Associates (total of 1) Dunstan Newman (U. West Indies Mona, 2010)
MATTHEW A. GITZENDANNER

Contact information: Department of Biology, University of Florida, Gainesville, FL 32611; e-mail: magitz@ufl.edu; Phone: 352-273-1960; Fax: 352-846-2154

a. Professional preparation:
Reed College, Portland, OR Biology B.A., May 1994
Washington State University, Pullman WA Botany Ph.D., May 2000

b. Appointments:
Associate Scientist, Department of Biology, University of Florida, Aug 2008-present
Bioinformatics Specialist, UF Research Computing, Univ. Fl., July 2011- present
Assistant Scientist, Department of Botany, University of Florida, 2001-Aug 2008
Lecturer, School of Biological Sciences, Washington State University, 2000-2001
Graduate Teaching Assistant, Department of Botany and School of Biological Sciences,
Washington State University, 1995-2000

c. Products:
Five publications most closely related to the proposed project:
Five additional relevant publications:


d. Synergistic activities:

User support and training for UF Research Computing: I assist researcher across campus make use of the High-Performance Computing Center and other resources. I conduct regular training sessions and provide one-on-one demonstrations of data analysis. In addition to general support, I am one of the bioinformatics specialists on staff at Research Computing and assist users in analyses.

Florida Museum of Natural History’s Informatics Committee. This committee is responsible for coordinating the Museum’s informatics efforts. We are commencing an overhaul of the Museum’s informatics resources, bringing us to the forefront of museum informatics.

Regular participation in Florida Museum of Natural History outreach events such as Sensational Science (an event for 7th grade students) and Dr. Discovery program (presenting research activities to visitors to the Discovery Room in the Museum).

e. Collaborators and other affiliates:

i. Collaborators:

David Althoff, Syracuse University; André Amorin, Universidade Estadual de Santa Cruz, Brazil; Riccardo Baldini, Università degli Studi, Italy; Charles Bell, University of New Orleans; Martin Cheek, Royal Botanic Gardens, Kew; Josh Clayton, University of Florida; Charles Davis, Harvard University; Amit Dhingra, Washington State University; Tory Hendry, University of Michigan; Shuguang Jian, South China Botanical Garden; Eric Menges, Archbold Biological Station; Ruiqi Li, University of Michigan; Michael Moore, Oberlin College; Yin-Long Qui, Université of Michigan; Vincent Savolainen, Royal Botanic Gardens, Kew; Kari Seagraves, Syracuse University; Douglas Soltis, University of Florida; Pamela Soltis, University of Florida; Cristina Vieira, University of Porto, Portugal;
Carl Weekly, Archbold Biological Station.

ii. Graduate advisor (Ph.D): Pamela Soltis, University of Florida.


f. Percent Effort:
   Instructional: 26%
   Research: Total: 71%
      Individual Research: 15% (5% would be dedicated to this project)
      Research Support and Coordination: 56%
   Service: 3%
Biographical Sketch

Aaron Beveridge

Department of English, University of Florida, FL
Tel: (419)-612-5786, Email: aaronbev79@ufl.edu

Professional Preparation
The Ohio State University        English        B.A., 2005
Akron University                  English        M.A., 2008
University of Florida             English        Ph.D., Expected 2017

Appointments
2013-present  Graduate Research Fellow, Department of English, University of Florida
2006-2008    Graduate Teaching Assistant, Department of English, Akron University
2004-2006    Writing Center Consultant, The Ohio State University

Five Products Most Closely Related to Proposed Project
4. Writing Studies and Data Science in the 4th Paradigm” College Conference on Composition and Communication. Indianapolis, IN (March 2014).

Synergistic Activities
1. Founding member of TRACE innovations initiative, the Department of English. TRACE is a research endeavor developed and maintained by the University of Florida’s Department of English. TRACE works at the intersection of ecology, posthumanism, and writing studies. Providing an interdisciplinary forum for scholars, we focus on the ethical and material impact of media. TRACE acts as a hub for several distinct projects including an online journal, Vectors, ARCs, and MassMine for which we are always seeking submissions. We welcome submissions that contribute to the theorization and invention of nature, technology, and culture. <http://trace.english.ufl.edu/>
2. Co-creator with Nicholas Van Horn of MassMine research software. Open source software in use at University of Florida, Clemson University, and Capital University (2014 – present) <https://github.com/n3mo/massmine>
requested: $60,000.00 (September 2014). <http://ufdc.ufl.edu/AA00025642/00001/pdf>

4. Developer of interdisciplinary Arduino prototyping and programming trainings for Marston Science Libraries at the University of Florida (April 2015)

5. Editor, TRACE journal. <http://trace.english.ufl.edu/journal/>

Collaborators Nicholas Van Horn (The Ohio State University) Sean Morey (Clemson University)
8. Appendices

Project Timeline

2015
May
• Sidney Dobrin, Aaron Beveridge, Norman Lewis, and Laurie Taylor will meet to discuss which NEH grant they will pursue
  ◦ Complete grant “feasibility meeting” with George A. Smathers Libraries
  ◦ Aaron Beveridge begins preliminary grant research/work
    • Find other similarly funded grants with NEH
    • Create to-do list and timeline for the project
    • Meet with potential grant participants to gain interdisciplinary support for the project
  ◦ Norman Lewis and Mindy McAdams hire graduate assistant
  ◦ Begin preliminary groundwork for grant
    • Find other similarly funded projects at funding agencies of choice
    • Create to-do list and timeline for the projects
    • Meet with potential grant participants to gain interdisciplinary support for the project
• Aaron Beveridge revises syllabus on “Writing through Big Data” course, rewrites notes collected during the semester for sharing with Norman Lewis, Mindy McAdams, and Laurie Taylor, collects any other relevant tutorials or training materials in preparation for summer trainings.

June
• Final decisions on grants made
• Continue grant research/development
• Meeting with Lewis, McAdams, and Beveridge to determine the dates for the weekly summer training schedule for July and August
• Beveridge continues collaborative research into the “virality threshold” for Twitter trend data

July
• Lewis signs up for Research Computing account and installs MassMine
• Begin weekly training and development sessions for Lewis and McAdams
  ◦ Train on how to use MassMine through Research Computing
    • Including setting up and managing their own long-term data collection projects
  ◦ Begin discussing how to use MassMine in the classroom
    • Discuss notes/syllabus from Beveridge’s “Writing through Big Data” course
    • Examine student projects from Beveridge's class for ideas that may be relevant to CJC students
    • Look at current training materials and documentation
• Continue grant research/development
• Beveridge continues collaborative research into the “virality threshold” for Twitter trend data

August
• Continue weekly training and development sessions for Lewis and McAdams
• McAdams and CJC research assistance will begin drafting new training materials specific to students and faculty in the CJC
• Lewis begins new individual research project based on data collected during trainings and development sessions
• Continue grant research/development
• Beveridge meet with Taylor to set exact dates for fall 2015 three-part training sessions in Scott Nygren Scholars Studio
• Beveridge continues collaborative research into the “virality threshold” for Twitter trend data
• Beveridge will work with Van Horn to make any necessary adjustments to MassMine software as a result of training, development, and collaboration with Lewis, McAdams, and Taylor
  ◦ Includes:
    • Improvements or additions to MassMine's visual front-end
    • If possible, the addition of new data sources or types from other APIs
    • Updates to MassMine's processing functionality for working with large data sets

September
• Begin drafting/grant writing
• Complete training and development with Lewis and McAdams
  ◦ Begin planning for Spring semester courses
    • Scheduling Beveridge's course assistance and support time for Lewis, McAdams spring courses
• Beveridge will meet with Laurie Taylor to discuss course assistance and support for spring certificate course in the digital humanities
• Beveridge will continue work with Van Horn to make any necessary adjustments to MassMine software
• Beveridge will complete the “virality threshold” data collection for Twitter trends, and begin collaboratively processing and analyzing the 6 month data set (collection will begin in April 2015)

October
• Continue grant drafting/writing
• Beginning of October: Beveridge will complete the first of the three-part MassMine and Research Computing training series in the Scott Nygren Scholars Studio
• End of October: Beveridge will complete the second of the three-part MassMine and Research Computing training series in the Scott Nygren Scholars Studio
• Beveridge will continue collaboratively analyzing “virality threshold” data, and begin additional research for the project

November
• Continue grant drafting/writing
  ◦ Meet with all project participants to make sure that substantial progress is made on grant drafts
  ◦ Create any necessary timelines or to-do lists to see that grants are polished and finished for timely submissions to funding agencies
• Beginning of November: Beveridge will complete the third of the three-part MassMine and Research Computing training series in the Scott Nygren Scholars Studio
• Beveridge will begin collaboratively drafting the second, follow-up article on the “virality threshold” based on the new data set and analyses
December
- Continue grant drafting/writing
- Beveridge will continue collaboratively drafting the second, follow-up article on the “virality threshold”
- Beveridge will work with Van Horn to make any new adjustments to MassMine software as a result of fall three-part trainings

2016 January
- Final meetings to ensure that grant participants have everything completed and ready for final submission of grants to funding agencies
- Beveridge will begin supporting Lewis and McAdams in CJC courses
- Beveridge will begin supporting Taylor in certificate course for digital humanities
- Beveridge will meet with Taylor to schedule exact dates for three-part MassMine and Research Computing training series
- Beveridge will continue collaboratively drafting the second, follow-up article on the “virality threshold”

February
- Submission of grants
  - Copies of grant documents uploaded to Digital Humanities Resources in George A. Smathers Libraries
  - Email sent to UFII with links to documents in order to report on final completion of grants submission
- Beveridge will continue supporting Lewis, McAdams, and Taylor's courses
- End of February: Beveridge will complete the first of the three-part MassMine and Research Computing training series in the Scott Nygren Scholars Studio

March
- Beveridge will continue supporting Lewis, McAdams, and Taylor's courses
- Beveridge will meet with Dobrin, Lewis, McAdams, and Taylor to create “Twitter Data in the UF Classroom” for presentation at the UFII Symposium
- Beveridge will create poster on the new “virality threshold” data and analyses for presentation at UFII Symposium

April
- Beveridge will finish supporting Lewis, McAdams, and Taylor's courses
- Entire project team will meet to discuss the previous year's work and future collaborations
- Beveridge will finish drafting second, follow-up article on “virality threshold” and submit for publication
- Beveridge will submit final email to UFII including notes, from project team final meeting, any updates on grant submissions, and to communicate successful completion of grant activities
March 12, 2015

Dear Dr. Dobrin;

It is my pleasure to join you and your colleagues on The MassMine Development and Training Project and to offer my support and training expertise. I have worked with Aaron Beveridge in getting MassMine up and running on Research Computing resources, produced documentation for this on the Research Computing wiki page, and lectured for his course on how to use MassMine at Research Computing. I am happy to support additional users and provide training for them to be able to effectively conduct their social media research using Research Computing facilities.

I will plan on dedicating 5% of my time to this project.

Sincerely,

Matt Gitzendanner
Associate Scientist, Department of Biology
Training Coordinator, UFIT Research Computing
March 15, 2015

Dr. Sidney I. Dobrin and
Aaron Beveridge
Department of English
CAMPUS

Dear Sid and Aaron,

The College and I are excited about your potential participation in the MassMine Development and Training Project, through your application for seed funding from the UF Informatics Institute. As I understand it, such funding is for new projects and collaborative partnerships such as the one you're developing with other stakeholders. I'm excited about the partnership of English with Smathers Libraries, the College of Journalism and Communications, and Research Computing. The MassMine project is a significant and worthy one for the humanities and social sciences especially, especially given the long-term prospects for social network data collection and analysis. I am happy to support your application, and appreciate your initiative and leadership.

Sincerely,

Kenneth Kidd
Professor and Chair
kbkidd@ufl.edu

The Foundation for The Gator Nation
An Equal Opportunity Institution
March 13, 2015

Dr. George Michailidis
Director, UF Informatics Institute

Dear Dr. Michailidis,

I am writing in support of the seed grant application from doctoral student Aaron Beveridge and Dr. Sidney Dobrin of the Department of English and Professor Mindy McAdams and Dr. Norm Lewis from the College of Journalism and Communications.

Their application is to develop and apply a big data research tool that can simplify the acquisition and analysis of large amounts of information from social media. Mr. Beveridge and a collaborator have already begun development of this tool, MassMine, which in a trial shows considerable promise.

Our interest is in using this tool in the classroom and to foster research. The tool can be used in classes that Professor McAdams teaches in programming and that Dr. Lewis teaches through a graduate seminar on big data. Through this seed grant, we’ll gain both the tools and the time to develop the expertise to learn and teach the tool. In addition, it will open opportunities for Dr. Lewis and others to conduct research projects suitable for publication in peer-reviewed journals.

Further, we anticipate that this seed grant will build sufficient expertise within our college that we will be able to apply for larger grants from outside agencies.

Professor McAdams and Dr. Lewis are two of our very best professors in the College of Journalism and Communications. They are forward-thinking (and forward-acting) when it comes to developing projects that have major implications for improving the skills and knowledge of our students, as well as when it comes to improving practices in both the media industry and journalism education. Their talents and interest in this area of big data, social media and technology are a perfect match for this project. We appreciate the opportunity to collaborate with the College of Liberal Arts and Sciences in this project, and thank you for its consideration.

Sincerely,

Ted Spiker
Chair, Associate Professor
Date: March 13, 2015
To: Dr. George Michailidis
   Director, UF Informatics Institute

Dear Dr. Michailidis,

I am writing in support of the seed grant application from doctoral student Aaron Beveridge and Dr. Sidney Dobrin of the Department of English and Professor Mindy McAdams and Dr. Norm Lewis from the College of Journalism and Communications.

Their application is to develop and apply a big data research tool that can simplify the acquisition and analysis of large amounts of information from social media. Mr. Beveridge and a collaborator have already begun development of this tool, MassMine, which in a trial shows considerable promise.

Our interest is in using this tool in the classroom and to foster research. The tool can be used in classes that Professor McAdams teaches in programming and that Dr. Lewis teaches through a graduate seminar on big data. Through this seed grant, we’ll gain both the tools and the time to develop the expertise to learn and teach the tool. In addition, it will open opportunities for Dr. Lewis and others to conduct research projects suitable for publication in peer-reviewed journals.

Further, we anticipate that this seed grant will build sufficient expertise within our college that we will be able to apply for larger grants from outside agencies.

We appreciate the opportunity to collaborate with the College of Liberal Arts and Sciences in this project, and thank you for its consideration.

Sincerely,

Yulia A. Strekalova
Grants Development Director
College of Journalism and Communications
(352) 846-2399 | yulias@ufl.edu